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 be written

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

   With Ilias Diakonikolas, Sushrut Karmalkar, and Aaron Potechin

   FOCS 2024

   Description. Non-Gaussian Component Analysis (NGCA) is the task of finding a non-Gaussian direction in the distribution of the samples, i.e., about detecting a signal among noise. Specifically, we're given i.i.d. samples from a distribution P_v on $\mathbb{R}^n$ which we know behaves like some distribution A in a hidden direction v and like a standard Gaussian in $v^\perp$; the task is to find v. Of particular interest is the situation where the distribution A matches all low-order moments with the standard Gaussian. In this case, it is known (under mild conditions on A) that current learning algorithms such as *statistical-query* and *low-degree polynomial tests* require a large number of samples to complete the task.

   Here we study the complexity of NGCA in the Sum-of-Squares (SoS) framework, an algorithm class that broadly strengthens the above ones. We prove that SoS in super-constant degree still requires a large number of samples, which translates to sample lower bounds for the corresponding

superpolynomial-time algorithms. Specifically, we show that if A matches the first $(k-1)$ moments of $\mathcal{N}(0,1)$ and satisfies mild conditions, then the degree $\widetilde{o}(\sqrt{\log n})$ SoS algorithm will almost always confirm there's some direction v along which the distribution is A, despite the truth being that all $n^{\frac{(1-\varepsilon)k}{2}}$ many samples are i.i.d. Gaussian. This lower bound is nearly tight, as there are spectral algorithms that excludes v using $n^{\frac{k}{2}}$ samples [Dudeja-Hsu 2022]. As a corollary, SoS lower bounds for several other problems in robust statistics and learning of mixture models are obtained.

The lower bound proof introduces a new technique and a few improvements of 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 Q is PSD while errors are small. What's new is the following. First, instead of the minimum weight vertex separator, we use the minimum square separator. Secondly, compared to previous works, there is a shift in approach for analysing pseudo-random matrices motivated by an intrinsic reason: in all prior works, the target matrix Q modulo negligible errors is a "real matrix" whose entries are numbers. Here, however, it is a nontrivial linear combination of equally dominating pseudo-random matrices (i.e., *graph matrices*, whose entries are polynomials). To show this linear combination is positive-definite something more than norm-based analysis is necessary, so we introduce a representation-theoretic method, which may be of more general interest. Namely, we model the multiplications of "important" pseudo-random matrices by an $\mathbb{R}$-algebra, construct all irreducible representations of this algebra, and use this information to show the special element Q is positive-definite (w.h.p.). Interestingly, the PSDness of Q boils down to the multiplicative identity of Hermite polynomials.

4. *Graph Colouring Is Hard on Average for Polynomial Calculus and Nullstellensatz*
With Jonas Conneryd, Susanna DeRezende, Jakob Norström, and 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 then by the known size-degree relation.

The proof proceeds by constructing an Alekhnovich-Razborov pseudo-reduction operator, where the essential combinatorial task is to define and construct a useful notion of *closure* of small vertex sets. To see what it means, imagine a prover reasons about the purported k-colouring χ, starting from a vertex set $S \subseteq V(G)$ and trying to deduce all possible information about $\chi|_S$. She can draw all conclusions from the assumption that $\chi|_S$ is a valid colouring of $G[S]$, the induced subgraph, sure. But she can do more—from the assumption that χ is a valid colouring on a larger part of the graph, she might deduce strictly more information about $\chi|_S$. The closure $cl(S) \supseteq S$ we seek, then, is a kind of surrogate of globality: we want to guarantee that for all $T \supseteq cl(S)$ which is not too large (say $|T| \leq \varepsilon|V(G)|$ for some global constant ϵ), $G[T]$ compared to $G[cl(S)]$ provides no more information about $\chi|_S$.

(Once we find such closures and show certain natural properties hold, the degree lower bound will not be far: the insight here is that for algebraic proofs, the word "information" above can be made concrete in terms of *monomial reducibility*.)

To construct the closures, we use a technique introduced by [Romero-Tunçel 2021] for graphs with large girth. We extend this technique to get rid of the girth assumption, where the key is an upper bound on the size of the closure assuming only that the graph is *everywhere-sparse*, a property random graphs enjoy. In particular, on a d-regular graph this property amounts to expansion. For random graphs, the set of large-degree vertices can cause trouble in the size estimate, but this can be handled by using a variant of closure.

3. *SoS Lower Bound for Exact Planted Clique*
   CCC 2021

   Description. This work proves Sum-of-Squares (SoS) degree lower bounds for Exact Planted Clique on random graphs $G(n, 1/2)$. In this problem, a parameter ω is chosen large enough that, with probability $> 99.99\%$, a random graphs does not contain a clique of size ω. Yet, determining whether a *given sample* G contains such a clique is a different task. We show that SoS algorithms cannot feasibly complete this task—unless their degree is so high that the algorithm essentially performs brute force. We focus on $\omega \ll \sqrt{n}$, since with larger ω-values the task becomes feasibly solvable by a spectral (i.e., degree-2) algorithm [Alon-Krivelevich-Sudakov

1998].

The word *exact* in the title means that the algorithm has access to the full set of axioms of the problem. This includes 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, say). We show a degree lower bound $d = \Omega(\frac{\varepsilon^2 \log n}{\log \log n})$ for ω up to $O(n^{\frac{1}{2}-\varepsilon})$, which is almost optimal in both d and ω.

A key motivation for our work is to further develop average-case lower bound techniques for SoS. Our approach expresses the *pseudo-expectation operator* $\widetilde{E}(\cdot)$ based on Fourier characters as in prior works, but we heavily explore the freedom in designing their coefficients, with no useful reference distribution in hand for calibrating the design. This is different from prior works. Cost is, we need to face the large number of possible choices with limited guiding principles, and with no clear feedback mechanism from the failure of analysis to the refinement of the choice.

After committing to a design, our analysis proceeds as follows. On a high-level, we decompose the moment matrix into a (positive) sum of Hadamard-products $\sum_{i=0}^{d} A_i \circ B_i$, and we show the positive definiteness of the two kinds of factors separately.

We define each A_i to be a *Johnson scheme*, inspired by the work [Feige-Krauthgamer 2003], whose positive definiteness follows relatively easily. The corresponding factors B_i are more complicated; we call the idea of analysis *relative factorization*, meaning the following. We decompose B_i as a sum $\sum_R B_{i,R}$ where R ranges over vertex subsets below a size threshold; the (vague) intuition is that the summand $B_{i,R}$ reflects *how the B_i factor should be if we condition on the event that the purported clique contains R.* Expression of $B_{i,R}$ is obtained by binomial transform, and it turns out to have a much more "factorizable form" than B_i, leading to a combinatorial $LQL^\top$-shaped factorization. Finally, we show the middle Q matrix is positive definite in the major case, $i = 0$. For that purpose, we study the analytical properties of a special matrix family which we call *factorial Hankels*.

In retrospect, the whole analysis is possible because $\widetilde{E}(\cdot)$ is chosen "correctly" in the first place—but is there an a priori explanation for this choice, which appears admittedly 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-DeRezende-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.