Sum-of-Squares Reading

July 26, 2020

1 Introduction

Roughly speaking, the Sum-of-Squares (SOS) hierarchy is a semi-definite programming hierarchy to reason about polynomial optimization subject to a system of polynomial equalities and inequalities constraints [BPT12, Las15]. Given a concrete instance of a polynomial optimization problem and a degree parameter [1] d, the SOS hierarchy [2] at level d specifies a semi-definite relaxation program, whose size grows with d. The lager the degree d the tighter this relaxation becomes as additional consistency constraints are imposed by the hierarchy and a harsher positive semi-definite condition must be satisfied. The versatility of polynomials as a language to model combinatorial optimization, statistics and machine learning problems makes the SOS hierarchy quite handy. More importantly, SOS gives the best known polynomial time guarantees for several of these problems [FKP19]. This striking success as a powerful algorithmic tool makes SOS lower bounds quite appealing as evidence of hardness. This is specially true for some average case problems for which NP-hardness may not be an option.

Being a strengthening of Linear Programming (actually the corresponding Sherali–Adams hierarchy), it is expected that the SOS hierarchy may be instrumental in upcoming state-of-the-art algorithms in diverse fields. It has been successfully applied to give subexponential time algorithms for Unique Games. SOS provides the best known approximation guarantees for all Max k-CSPs. More recently, it has been used in a variety of (robust) Statistics and Machine Learning problems, where several improved polynomial time guarantees have been obtained. Also recently, SOS has been employed in Coding Theory. On one hand SOS relieves the algorithm designer from defining the semi-definite relaxation, on the other hand rounding (when needed) can be quite challenging. A systematic theory of rounding that takes full advantage of the SOS hierarchy is an ambitious open problem.

As Computer Science goes beyond worst-case guarantees, several average case problems seem to defy our best algorithmic techniques leading one to conjecture their hardness. However, here the sophisticated machinery of Probabilistically Checkable Proofs tailored to worst-case analysis may not be available. Yet, one stills needs an answer regarding the hardness of these average case problems (hardness can also lead to useful applications). Confronting an average case problem against the SOS hierarchy and obtaining a confirmation that SOS fails (to provide the required guarantees within a given degree range) can serve as some evidence of hardness. Of course, this is not an evidence against all possible algorithms, but at least it says that a powerful meta-algorithm (the SOS hierarchy) cannot do it. Showing an SOS lower bound entails showing a limitation on

[1] Sufficiently large to for the SDP to be well-defined.

[2] This is dual view of SOS hierarchy.

this powerful meta-algorithm, which can be quite challenging. This study has produced quite interesting objects and tools involving random matrix theory and Fourier analysis (the so-called graph matrices). Similarly to the algorithmic front, a systematic theory of SOS lower bounds is an ambitious open problem.

Before we reach a systematic understanding of the SOS hierarchy both in terms of upper and lower bounds, it seems natural that we might need to develop it in a myriad of concrete problems hopefully helping in pushing the state-of-the-art in several other fields along the way.

2 Algorithms

The following is a list of some SOS (or SOS-related) algorithms.

- In [CMY20], Cherapanamjeri–Mohanty–Yau show how to do list decode mean estimation in nearly linear time. This improves the running time of the SOS based result for mean estimation of [RY20].
- In [BK20], Bakshi and Kothari give a list decoding algorithm for subspace recovery.
- In [BBK+20], Bafna et al. show how to approximate unique games on constraint graphs whose small-set expansion can be certified.
- In [RY20], Raghavendra and Yau give list decoding algorithms for learning (regression and mean estimation).
- In [KKK19], Karmalkar–Klivans–Kothari also give list decoding algorithms for learning (regression and mean estimation).
- In [AJQ+20], Alev et al. give list decoding algorithm for distance amplified codes using HDX and expander graphs.
- In [AJT19], Alev et al. give approximation algorithms for k-CSPs on HDXs extending the result of Barak–Raghavendra–Steurer [BRS11].
- In [RSS18], Raghavendra–Schramm–Steurer surveys the growing literature of SOS for high-dimensional estimation (deducing parameters of distributions given their samples).
- In [BKS17], Barak–Kothari–Steurer give a $\exp(\widetilde{O}(\sqrt{n}))$-time algorithm to find a rank one operator acting on an input subspace of ambient dimension n. This gives an $\exp(\widetilde{O}(\sqrt{n}))$-time algorithm for QMA(2) with each witness dimension of n.
- In [BGG+17], Bhattiprolu et al. give approximation algorithms to compute the maximum absolute value of a multivariate homogeneous polynomial over the sphere. Their result is based on a weak decoupling result (de Finetti like result).
- In [HKP+17], Hopkins et al. show that that spectral algorithms using matrices whose entries are low degree polynomials on the input variables provide similar guarantees of SOS. They also show nearly tight SOS lower bounds for tensor and sparse PCA.
- In [RRS17], Raghavendra–Rao–Schramm give SOS refutation algorithms for random CSP (see [KMOW17] for nearly matching lower bounds).

- In [HSSS16], Hopkins et al. show faster spectral algorithms based on SOS proofs. Although (standard) SOS is not the fastest algorithmic tool (in terms of finer polynomial running time), its insights can lead to more efficient implementation.
- In [BKS14], Barak–Kelner–Steurer design some reasonably general rounding scheme for SOS. In particular, they apply their techniques to polynomial optimization, with non-negative coefficients, over the sphere.
- In [BRS11], Barak–Raghavendra–Steurer use the SOS hierarchy to give an approximation algorithm for 2-CSPs with guarantees based on the threshold rank of a graph. They also show that degree-$n^{\mathrm{poly}(\epsilon)}$ SOS can decide unique games. Qualitatively speaking their results have some similarities to [GS11].
- In [GS11], Guruswami and Sinop use the SOS hiearchy to give approximation algorithms with guarantees based on the threshold rank of a graph. They also show subexponential time algorithm to decide unique games. Qualitatively speaking their results have some similarities to [BRS11].

Although the SOS hierarchy specifies how to write an SDP program relxation given a polynomial optimization problem and a degree d, it does not refrain the user from rounding if an integral solution is need. Rounding can be quite challenging and it is not always clear what is the best rounding scheme. Therefore, an ambitious open-ended question is the following.

Open Problem 2.1 (Open-ended Problem)**.** *Is there a general theory of SOS rounding?*

The SOS hierarchy had some early successes in approximating CSPs, where under the Unique Games Conjecture it provides the best approximation guarantees. It can decide unique games in subexponential time and give efficient good approximations for CSPs on reasonably structured graphs (low threashold rank). More recently, it has been successfully applied to machine learning, some (robust) statistics and coding theory problems.

Open Problem 2.2 (Open-ended Problem)**.** *Which fields and algorithmic questions can (algorithmically) benefit from the SOS hierarchy?*

3 Lower Bounds

The following is a list of some SOS lower bounds.

- In [MRX20], Mohanty–Raghavendra–Xu quite surprisingly show how to lift degree-2 lower bounds to degree-4 lower bounds under suitable conditions.
- In [Pot18], Potechin shows nearly tight upper and lower bounds on the ordering principle.
- In [KOS19], Kothari, O'Donnell and Schramm show how to incorporate global cardinality constraints in SOS lower bounds. They obtain strong degree lower bounds for Min-Bisection and Max-Bisection.
- In [Pot19], Potechin explores symmetry to give a framework for SOS lower bounds for symmetric problems.

- In [KM18], Kothari and Mehta show SOS hardness of finding any equilibrium rather than for those optimizing some objective function.

- In [KMOW17], Kothari et al. obtain nearly tight SOS lower bounds for refuting CSPs whose predicates support a t-wise uniform distribution. Their PSD analysis is based on a Gram–Schmidt orthoganalization process.

- In a breakthrough result [BHK$^+$16], Barak et al. show nearly tight SOS lower bounds for the planted clique problem. Their result introduce the general framework of pseudocalibration to find good candidate integrality gap solutions.

- In [LRS15], Lee–Raghavendra–Steurer introduce reasonably general SDP lower bounds techniques. In particular, they obtain that for MAX CSPs any polynomial-size SDP is equivalent in power to degree-$O(1)$ SOS.

- In [Tul09], Tulsiani obtain SOS integrality gaps for MAX k-CSPs. Then using reductions Tulsiani obtains SOS gaps for Max Independent Set, Approximate Coloring, Chromatic Number and Min Vertex Cover.

SOS lower bounds are oftentimes obtained in a case-by-case manner, though some general elegant tools and techniques have emerged such as graph matrices and pseudo-calibration. Proving that a given candidate solution is PSD can be quite challenging. One might ambitiously (and possibly very naively) ask the following.

Open Problem 3.1 (Open-ended Problem)**.** *Is there a general theory of SOS lower bounds?*

The preceding open problem may be too ambitious to ever see a positive resolution. However, more modestly and of great interest is the following.

Open Problem 3.2 (Open-ended Problem)**.** *More tools and techniques for SOS lower bounds.*

4 SOS Machinery

A noteworthy tool that has emerged in the study of SOS lower bounds are the so-called graph matrices. They provide an elegant mix of Fourier analysis, combinatorics and random matrix theory with (possibly) non independent entries.

- In [AMP16], Kwangjun et al. give a detailed account of the graph matrices machinery, which is a powerful tool used in some SoS lower bound results. Graph matrices can be seen as a matrix analogue of Fourier characters. Spectral norm bounds follow from simple combinatorial properties of underlying graphs.

- In [CP20], Cai and Potechin pin-point the singular values spectra of some graph matrices called Z-shaped.

5 Misc

- In [RSST18], Raymond et al. show an equivalence of (the semi-define method of) flag algebras and (symmetric) Sum-of-Squares. Their techniques are based on the representation theory of the symmetric group.
- In [GP04], Gatermann and Parrilo show how to explore symmetries of a polynomial optimization problem to greatly reduce the size of the corresponding SOS SDP program. They use representation theory to find this succinct description.
- In [BBH$^+$12], Barak et al. give an SOS proof of the hypercontractivity inequality obtaining several applications. In particular, they rule out the short-code graph as a hard (super polynomial) instance for SOS.

References

[AJQ$^+$20] Vedat Levi Alev, Fernando Granha Jeronimo, Dylan Quintana, Shashank Srivastava, and Madhur Tulsiani. List decoding of direct sum codes. In *Proceedings of the 31st ACM-SIAM Symposium on Discrete Algorithms*, 2020.

[AJT19] Vedat Levi Alev, Fernando Granha Jeronimo, and Madhur Tulsiani. Approximating constraint satisfaction problems on high-dimensional expanders. 2019.

[AMP16] Kwangjun Ahn, Dhruv Medarametla, and Aaron Potechin. Graph matrices: Norm bounds and applications, 2016.

[BBH$^+$12] Boaz Barak, Fernando G.S.L. Brandao, Aram W. Harrow, Jonathan Kelner, David Steurer, and Yuan Zhou. Hypercontractivity, sum-of-squares proofs, and their applications. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*, STOC 12, page 307326, 2012.

[BBK$^+$20] Mitali Bafna, Boaz Barak, Pravesh Kothari, Tselil Schramm, and David Steurer. Playing unique games on certified small-set expanders, 2020.

[BGG$^+$17] V. Bhattiprolu, M. Ghosh, V. Guruswami, E. Lee, and M. Tulsiani. Weak decoupling, polynomial folds and approximate optimization over the sphere. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1008–1019, 2017.

[BHK$^+$16] B. Barak, S. B. Hopkins, J. Kelner, P. Kothari, A. Moitra, and A. Potechin. A nearly tight sum-of-squares lower bound for the planted clique problem. In *2016 IEEE 57th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 428–437, 2016.

[BK20] Ainesh Bakshi and Pravesh Kothari. List-decodable subspace recovery via sum-of-squares, 2020.

[BKS14] Boaz Barak, Jonathan A. Kelner, and David Steurer. Rounding sum-of-squares relaxations. In *Proceedings of the Forty-Sixth Annual ACM Symposium on Theory of Computing*, STOC 14, page 3140, 2014.

[BKS17] Boaz Barak, Pravesh K. Kothari, and David Steurer. Quantum entanglement, sum of squares, and the log rank conjecture. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, page 975988, 2017.

[BPT12] Grigoriy Blekherman, Pablo A. Parrilo, and Rekha R. Thomas. *Semidefinite Optimization and Convex Algebraic Geometry*. Society for Industrial and Applied Mathematics, USA, 2012.

[BRS11] B. Barak, P. Raghavendra, and D. Steurer. Rounding semidefinite programming hierarchies via global correlation, 2011.

[CMY20] Yeshwanth Cherapanamjeri, Sidhanth Mohanty, and Morris Yau. List decodable mean estimation in nearly linear time, 2020.

[CP20] Wenjun Cai and Aaron Potechin. The spectrum of the singular values of z-shaped graph matrices, 2020.

[FKP19] N. Fleming, P. Kothari, and T. Pitassi. *Semialgebraic Proofs and Efficient Algorithm Design*. 2019.

[GP04] Karin Gatermann and Pablo A. Parrilo. Symmetry groups, semidefinite programs, and sums of squares. *Journal of Pure and Applied Algebra*, 192(1-3):95128, Sep 2004.

[GS11] Venkatesan Guruswami and Ali Kemal Sinop. Lasserre hierarchy, higher eigenvalues, and approximation schemes for graph partitioning and quadratic integer programming with psd objectives. In *Proceedings of the 2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, FOCS 11, page 482491, 2011.

[HKP+17] S. B. Hopkins, P. K. Kothari, A. Potechin, P. Raghavendra, T. Schramm, and D. Steurer. The power of sum-of-squares for detecting hidden structures. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 720–731, 2017.

[HSSS16] Samuel B. Hopkins, Tselil Schramm, Jonathan Shi, and David Steurer. Fast spectral algorithms from sum-of-squares proofs: Tensor decomposition and planted sparse vectors. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC 16, page 178191, 2016.

[KKK19] Sushrut Karmalkar, Adam R. Klivans, and Pravesh Kothari. List-decodable linear regression. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8-14 December 2019, Vancouver, BC, Canada*, pages 7423–7432, 2019.

[KM18] Pravesh K. Kothari and Ruta Mehta. Sum-of-squares meets nash: Lower bounds for finding any equilibrium. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, page 12411248, 2018.

[KMOW17] Pravesh K. Kothari, Ryuhei Mori, Ryan ODonnell, and David Witmer. Sum of squares lower bounds for refuting any csp. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, page 132145, 2017.

[KOS19] Pravesh K. Kothari, Ryan O'Donnell, and Tselil Schramm. SOS lower bounds with hard constraints: Think global, act local. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124, pages 49:1–49:21, 2019.

[Las15] Jean Bernard Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization.* Cambridge Texts in Applied Mathematics. Cambridge University Press, 2015.

[LRS15] James R. Lee, Prasad Raghavendra, and David Steurer. Lower bounds on the size of semidefinite programming relaxations. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC 15, page 567576, 2015.

[MRX20] Sidhanth Mohanty, Prasad Raghavendra, and Jeff Xu. Lifting sum-of-squares lower bounds: Degree-2 to degree-4. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, page 840853, 2020.

[Pot18] Aaron Potechin. Sum of squares bounds for the ordering principle, 2018.

[Pot19] Aaron Potechin. Sum of squares lower bounds from symmetry and a good story. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124, pages 61:1–61:20, 2019.

[RRS17] Prasad Raghavendra, Satish Rao, and Tselil Schramm. Strongly refuting random csps below the spectral threshold. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, page 121131, 2017.

[RSS18] Prasad Raghavendra, Tselil Schramm, and David Steurer. High-dimensional estimation via sum-of-squares proofs, 2018.

[RSST18] Annie Raymond, James Saunderson, Mohit Singh, and Rekha R. Thomas. Symmetric sums of squares over k-subset hypercubes. *Math. Program.*, 167(2):315354, February 2018.

[RY20] Prasad Raghavendra and Morris Yau. List decodable learning via sum of squares. In *Proceedings of the Thirty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA 20, page 161180, USA, 2020. Society for Industrial and Applied Mathematics.

[Tul09] Madhur Tulsiani. Csp gaps and reductions in the lasserre hierarchy. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC 09, page 303312, 2009.