

Response to “Counterexample to global convergence of DSOS and SDSOS hierarchies”

Amir Ali Ahmadi*and Anirudha Majumdar †

Abstract

In a recent note [8], the author provides a counterexample to the global convergence of what his work refers to as “the DSOS and SDSOS hierarchies” for polynomial optimization problems (POPs) and purports that this refutes claims in our extended abstract [4] and slides in [3]. The goal of this paper is to clarify that neither [4], nor [3], and certainly not our full paper [5], ever defined DSOS or SDSOS hierarchies as it is done in [8]. It goes without saying that no claims about convergence properties of the hierarchies in [8] were ever made as a consequence. What was stated in [4, 3] was completely different: we stated that *there exist* hierarchies based on DSOS and SDSOS optimization that converge. This is indeed true as we discuss in this response. We also emphasize that we were well aware that some (S)DSOS hierarchies do not converge even if their natural SOS counterparts do. This is readily implied by an example in our prior work [5], which makes the counterexample in [8] superfluous. Finally, we provide concrete counterarguments to claims made in [8] that aim to challenge the scalability improvements obtained by DSOS and SDSOS optimization as compared to sum of squares (SOS) optimization.

1 Introduction

In [8], the author considers our recently proposed (S)DSOS optimization framework [5]. Based on it, he constructs two particular hierarchies, which [8] calls “the DSOS and SDSOS hierarchies”, and shows that they do not always converge to the global optimal value of polynomial optimization problems (POPs) via an example. The author then makes two primary claims: (i) that this example refutes statements made in our extended abstract [4] and slides [3], and (ii) that DSOS and SDSOS optimization do not provide tractable alternatives to SOS optimization. This response shows that (i) is not true and the main technical premise on which (ii) is based is also not true.¹

We refer the reader to [5] for a formal introduction to DSOS and SDSOS programming.² Briefly, dsos (resp. sdsos) polynomials are sos polynomials whose Gram matrices are diagonally dominant (resp. scaled diagonally dominant). Optimizing a linear function over these cones subject to affine constraints can be performed using linear and second-order cone programming respectively. The resulting DSOS and SDSOS programs have shown improvements in scalability over SOS programming for a range of different problems [5, Section 4].

*Department of Operations Research and Financial Engineering, Princeton University. (a.a.a@princeton.edu, <http://aaa.princeton.edu/>)

†Department of Mechanical and Aerospace Engineering, Princeton University. (ani.majumdar@princeton.edu, <https://irom-lab.princeton.edu/majumdar>).

¹We take this opportunity to point out that the reference for our work on DSOS and SDSOS optimization should be our full paper on the topic [5]. In both our slides [3], and our extended abstract [4], details have intentionally been omitted (see e.g. the disclaimer on the first page of [4]).

²Our notation here is consistent with [5]; in particular, we use upper case to refer to (S)DSOS cones and hierarchies, while lower case is reserved for (s)dsos polynomials.

2 Claim 1: Non-convergence of DSOS and SDSOS hierarchies on POPs

In reference to our work, the abstract of [8] states that the result presented in [8] “refutes the claim in the literature according to which the DSOS and SDSOS hierarchies can solve any polynomial optimization problem to arbitrary accuracy”.

2.1 Rebuttal to Claim 1

The “claim in the literature” that [8] purports to refute does not appear anywhere in our main manuscript [5], which was made publicly available in 2017. It does not even feature (contrary to what [8] claims) in our extended abstract [4] from 2014 or presentation slides [3] from 2013. In particular, our prior work did not present “the DSOS and SDSOS hierarchies” that [8] considers to be convergent hierarchies for the POP. The results in [8] therefore do not provide counterexamples to any results proven in [5, 4] (or, to our knowledge, any other claims made in the literature).

What we do claim in [4, 3] is that there *exist* converging hierarchies for POPs based on DSOS and SDSOS optimization. This statement is certainly true; in fact there are many weaker converging hierarchies that as an immediate corollary imply existence of a (S)DSOS-based converging hierarchy (see, e.g., those by Lasserre [10], [11, Chap. 9], Peña, Vera, and Zuluaga [14], [9, Theorem 2], and Ahmadi and Hall [1]). In the first two references, the LP hierarchy presented only requires a search over nonnegative scalars. If one replaces scalar optimization variables by dsos polynomials, one trivially obtains new LP-based globally convergent hierarchies (since the modified hierarchy automatically inherits the convergence properties of the original LP hierarchy). In the latter reference, the sos polynomials that feature in the hierarchy all have a *diagonal* Gram matrix. These polynomials are also a very special case of dsos polynomials. The assumptions that these hierarchies make on the POP (particularly in the case of [1] and [11]) are very similar to those made for SOS-based converging hierarchies; see, e.g., the discussion in [11, Section 2.4.3]. In general, we expect that there are many more possible ways of constructing hierarchies based on (S)DSOS optimization. We further note that [5, Section 3.2] was already providing a convergent hierarchy for the important special case of copositive programs [6].

Finally, we note that while globally convergent hierarchies are valuable from a theoretical perspective, they are generally of limited utility to practitioners without concrete and practical bounds on convergence rates. This is a challenge for the usual hierarchies based on SOS optimization, and it is also a challenge faced by convergent hierarchies based on (S)DSOS optimization. Due to this reason, we did not focus on explicitly presenting convergent hierarchies for POP in [5], which is meant to be a broadly-accessible application-oriented paper. Rather, in [5], we chose to focus on demonstrating the scalability of (S)DSOS programming using large-scale numerical examples drawn from a diverse range of application areas ([5, Section 4]), and algorithms for iteratively improving the quality of solutions obtained from (S)DSOS programs ([5, Section 5]). In a different paper [1], one can find a systematic approach for constructing converging hierarchies for POPs based on minimal requirements. In fact, the methodology presented there produces a converging hierarchy that does not even use optimization.

2.2 An example from our prior work makes the main claim in [8] superfluous

We point out in this subsection that our paper [5], which is referenced in [8], already gave an example of a very simple polynomial optimization problem (see problem (2) below) for which the first level of a well-known SOS hierarchy is exact but for which no level of the analogous DSOS or

the SDSOS hierarchies are exact. This is an immediate corollary of [5, Proposition 14] in our paper. It shows that we were cognizant of the fact that one cannot take any SOS hierarchy, replace the sos constraints with dsos or sdsos constraints (or even r -dsos and r -sdsos constraints), and expect convergence. As a consequence, we would not have made such a claim, and we certainly never stated that the Lasserre hierarchy with sos constraints replaced by dsos/sdsos constraints converges. Our claim was simply that one can construct hierarchies with dsos/sdsos constraints that converge.

Recall that a polynomial f is said to be r -sos (resp. r -dsos, r -sdsos) if $f(x) \cdot (\sum x_i^2)^r$ is sos (resp. dsos, sdsos). Consider the problem of minimizing a form p of degree $2d$ on the sphere:

$$\begin{aligned} p^* &:= \min_{x \in \mathbb{R}^n} p(x) \\ \text{s.t.} \quad & \sum_{i=1}^n x_i^2 = 1. \end{aligned} \tag{1}$$

The most standard SOS hierarchy for this problem (indexed by $r \in \mathbb{N}$) reads [13]:

$$\begin{aligned} l_{\text{sos},r} &:= \max_{\gamma \in \mathbb{R}} \gamma \\ \text{s.t.} \quad & p(x) - \gamma \left(\sum_{i=1}^n x_i^2 \right)^d \text{ is } r\text{-sos}. \end{aligned}$$

These are semidefinite programs that produce a sequence of lower bounds $\{l_{\text{sos},r}\}$ on p^* with $l_{\text{sos},r} \rightarrow p^*$ as $r \rightarrow \infty$. This hierarchy is shown to be equivalent to the Lasserre hierarchy applied to problem (1) in [7, Proposition 2].

One can now define an analogous DSOS/SDSOS hierarchy:

$$\begin{aligned} l_{\text{dsos},r} \text{ (resp. } l_{\text{sdsos},r}) &:= \max_{\gamma \in \mathbb{R}} \gamma \\ \text{s.t.} \quad & p(x) - \gamma \left(\sum_{i=1}^n x_i^2 \right)^d \text{ is } r\text{-dsos (resp. } r\text{-sdsos)}. \end{aligned}$$

We clearly have $l_{\text{dsos},r} \leq l_{\text{sdsos},r} \leq l_{\text{sos},r}$ for all r . Consider now the following optimization problem:

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & (x_1 + x_2 + x_3)^2 \\ \text{s.t.} \quad & x_1^2 + x_2^2 + x_3^2 = 1. \end{aligned} \tag{2}$$

The optimal value of this problem is zero, which is achieved by the first-order sos relaxation; i.e., $l_{\text{sos},0} = 0$. (This is because all nonnegative quadratic forms are sums of squares.) However, one has

$$l_{\text{dsos},r} \leq l_{\text{sdsos},r} \leq -1$$

for all r . This is an immediate consequence of the following proposition, which we have already proven in [5].

Proposition 2.1 (Proposition 14 in [5]). *For any $0 < a < 1$, the quadratic form*

$$f(x_1, x_2, x_3) = (x_1 + x_2 + x_3)^2 + a(x_1^2 + x_2^2 + x_3^2) \tag{3}$$

is positive definite but not r -sdsos for any r .

We remark that in contrast to our example in (2), the example given in [8] is in two variables and constitutes a convex problem, but these considerations are secondary.

3 Claim 2: (S)DSOS is not necessarily more tractable than SOS

The abstract of [8] states: “We further observe that the dual to the SDSOS hierarchy is the moment hierarchy where every positive semidefinite constraint is relaxed to all necessary second-order conic constraints. As a result, the number of second-order conic constraints grows exponentially as the order of the SDSOS hierarchy increases. Together with the counterexample, this suggests that DSOS and SDSOS are not necessarily more tractable alternatives to sum-of-squares.”

3.1 Rebuttal to Claim 2

As observed already in earlier work [15, Section 3.2], [2, Section 3.3], the dual to SDSOS programs can indeed be obtained by relaxing semidefinite constraints to necessary second-order conic constraints. However, the number of SOCP constraints grows *polynomially* as the order of the r -sdsos hierarchy increases. Hence the claim above about the exponential increase is false. Indeed, consider a polynomial in n variables and degree $2d$. When this polynomial is required to be r -sdsos, the scaled diagonal dominance constraint needs to be imposed on a symmetric matrix of size $N \times N$, where $N := \binom{n+d+r}{n}$. This leads to $O(N^2) = O(n+d+r)^{2n}$ SOCP constraints. Hence, the dependence on r is indeed polynomial. (Here, n is constant as the number of variables to any polynomial optimization problem that goes through the hierarchy is clearly fixed.)

We shall emphasize that the size of the matrices that SDSOS optimization deals with is exactly the same as that of the SOS approach. The difference is that an expensive semidefinite constraint of size $N \times N$ is replaced with $\binom{N}{2}$ cheap SOCP constraints. We have shown with numerous examples (see [5, Section 4]) that for a range of problem sizes of practical value, this can lead to significant improvements in scalability. We also refer the reader to Theorem 10 and Theorem 12 of [5], where we show that from a theoretical standpoint, polynomial-time solvability of r -(S)DSOS programs is identical to that of r -SOS programs.

4 Discussion

From an application viewpoint, we believe that the (S)DSOS optimization approach can provide practitioners in diverse application domains with powerful and tractable alternatives to SOS optimization. As we demonstrate in [5, 12] with numerical examples from a wide range of domains (including copositive and combinatorial optimization, machine learning and statistics, control theory, and robotics), the (S)DSOS approach can provide significant gains in computation times (as much as orders of magnitude in certain cases) when compared to the SOS approach. For example, in [12], we used SDSOS programming to design stabilizing feedback controllers for a 30 dimensional state space and 14 dimensional control input space model of a humanoid robot³. The scale of this problem is well beyond what SOS programming can currently handle. The computational gains in the problems considered in [5, 12] are obtained without exploiting any particular structure (e.g., sparsity or symmetry) in the problems. As noted in [5, Section 2], the (S)DSOS approach could potentially be combined with approaches that exploit the structure of the problem at hand in order to obtain even more significant computational gains. We further note that while there is conservatism inherent in the (S)DSOS approach as compared to SOS optimization, the examples in [5] demonstrate that this conservatism can be small and mitigated by the techniques presented in [5, Section 5].

³A video is available online: <http://youtu.be/lmAT556Ar5c>

We conclude by noting that the (S)DSOS framework is not meant as a general replacement for hierarchies based on SOS programming (e.g., the Lasserre/Parrilo hierarchies). Rather, the (S)DSOS approach provides a way to tackle problems that are beyond the reach of SOS programming due to computational limitations. Indeed, we believe that the true power of the (S)DSOS framework comes from the fact that it can provide solutions in situations where even the first level of the SOS relaxation hierarchy is simply too expensive to solve (see [5, 12] for numerous examples of this kind).

References

- [1] A. A. Ahmadi and G. Hall. On the construction of converging hierarchies for polynomial optimization based on certificates of global positivity. Available at <https://arxiv.org/abs/1709.09307>, 2017.
- [2] A. A. Ahmadi and G. Hall. Sum of squares basis pursuit with linear and second order cone programming. *Contemporary Mathematics*, 2017.
- [3] A. A. Ahmadi and A. Majumdar. DSOS and SDSOS: More tractable alternatives to SOS. *Presentation slides at the meeting on Geometry and Algebra of Linear Matrix Inequalities, CIRM, Marseille*, 2013.
- [4] A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: LP and SOCP-based alternatives to sum of squares optimization. In *proceedings of the 48th annual IEEE Conference on Information Sciences and Systems (CISS)*, 2014. Extended abstract.
- [5] A. A. Ahmadi and A. Majumdar. DSOS and SDSOS optimization: more tractable alternatives to sum of squares and semidefinite optimization. Available at <https://arxiv.org/abs/1706.02586>, 2017.
- [6] S. Burer. Copositive Programming. In *Handbook on Semidefinite, Conic and Polynomial Optimization*, pages 201–218. Springer, 2012.
- [7] E. de Klerk, M. Laurent, and P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. *Positive Polynomials in Control*, pages 580–580, 2005.
- [8] C. Jozs. Counterexample to global convergence of DSOS and SDSOS hierarchies. *arXiv preprint arXiv:1707.02964*, 2017.
- [9] X. Kuang, B. Ghaddar, J. Naoum-Sawaya, and L. Zuluaga. Alternative SDP and SOCP approximations for polynomial optimization. *arXiv preprint arXiv:1510.06797*, 2017.
- [10] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1. World Scientific, 2009.
- [11] J. B. Lasserre. *An Introduction to Polynomial and Semialgebraic Optimization*. Cambridge University Press, 2015.
- [12] A. Majumdar, A. Ahmadi, and R. Tedrake. Control and verification of high-dimensional systems with DSOS and SDSOS programming. In *53rd IEEE Conference on Decision and Control*, pages 394–401. IEEE, 2014.

- [13] P. A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. PhD thesis, California Institute of Technology, May 2000.
- [14] J. Peña, J. Vera, and L. Zuluaga. Positive polynomials on unbounded domains. Available at <https://arxiv.org/abs/1709.03435>, 2017.
- [15] F. Permenter and P. Parrilo. Partial facial reduction: simplified, equivalent SDPs via approximations of the PSD cone. *arXiv preprint arXiv:1408.4685*, 2014.