



HAL
open science

Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem

Didier Henrion, Adrien Le Franc, Victor Magron

► **To cite this version:**

Didier Henrion, Adrien Le Franc, Victor Magron. Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem. 2024. hal-04500960v3

HAL Id: hal-04500960

<https://hal.science/hal-04500960v3>

Preprint submitted on 26 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem

Didier Henrion^{1,2}, Adrien Le Franc¹, Victor Magron^{1,3}

Draft of November 26, 2024

Abstract

We describe a parametric univariate quadratic optimization problem for which the moment-SOS hierarchy has finite but increasingly slow convergence when the parameter tends to its limit value. We estimate the order of finite convergence as a function of the parameter.

1 Introduction

The moment-SOS (sum of squares) hierarchy proposed in [12] is a powerful approach for solving globally non-convex polynomial optimization problems (POPs) at the price of solving a family of convex semidefinite optimization problems (called moment-SOS relaxations) of increasing size, controlled by an integer, the relaxation order. Under standard assumptions, the resulting sequence of bounds on the global optimum converges asymptotically, i.e., when the relaxation order goes to infinity [9, 8, 21]. We say that *convergence is finite* if the bound matches the global optimum at a finite relaxation order, i.e., solving the relaxation actually solves the POP globally. In this case, we also say that the *relaxation is exact*. It is known that convergence is finite generically [20], which means that POPs for which convergence is not finite are exceptional. In the specific case of minimizing a univariate polynomial of degree d over either the real line, a given closed interval or $[0, \infty)$, finite convergence systematically occurs at the minimal relaxation order $\lceil d/2 \rceil$ as a consequence of [5, 24] (see also [13, § 2.3] for a modern exposition) if the interval is described with a proper set of linear/quadratic inequality constraints. In addition to finite convergence occurring systematically in such univariate POPs, one can even bound the running time required to compute the resulting SOS decompositions when the input data involve rational coefficients [16, 17].

¹CNRS; LAAS; Université de Toulouse, 7 avenue du colonel Roche, F-31400 Toulouse, France.

²Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czechia.

³Institute of Mathematics of Toulouse, 118 route de Narbonne, F-31062 Toulouse, France.

Restricting our attention to POPs on compact sets, several elementary examples without finite convergence are known. A bivariate example is $\min_{x \in \mathbb{R}^2} (1 - x_1^2)(1 - x_2^2)$ s.t. $1 - x_1^2 \geq 0, 1 - x_2^2 \geq 0$, see [3, Prop. 29]. Many results on the speed of convergence of the moment-SOS hierarchy are now available, see, e.g., [2, 26]. However, they are practically not useful since the rates are asymptotic, for very large values of the relaxation order. In practice, when implementing the moment-SOS hierarchy on low-dimensional benchmark POPs, we observe a finite and fast convergence of the hierarchy, see, e.g., the original experiments carried out in [10]. To the best of the authors' knowledge, almost nothing is known about the speed of convergence for small relaxation orders, when the convergence is finite.

The contribution of this note is to describe and study the elementary bounded univariate quadratic POP

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & x \\ \text{s.t.} \quad & 1 - x^2 \geq 0 \\ & x + (1 - \varepsilon)x^2 \geq 0, \end{aligned}$$

parametrized by a scalar $\varepsilon \in [0, 1]$, such that the convergence of the moment-SOS hierarchy is finite, but arbitrarily slow, when ε tends to zero. The feasible set is the segment $[0, 1]$ for every $\varepsilon > 0$. When $\varepsilon = 0$, the feasible set is the union of the segment $[0, 1]$ and the point -1 . We estimate the order of finite convergence as a function of ε .

Our contributions can be summarized as follows:

- Finite convergence holds for all $\varepsilon \in [0, 1]$ (Proposition 1) and there exist steps / threshold values $\varepsilon_d \in [0, 1]$ for the exactness of the relaxation of order d (Proposition 2);
- We provide lower and upper bounds on ε_d to estimate the order at which finite convergence occurs (Theorem 1).

Interestingly, this implies that we are able to generate simple univariate POPs with arbitrarily large finite convergence orders.

2 POP design

In this section we explain how our parametric univariate quadratic POP is designed.

2.1 Circle and two lines

We introduce the following toy problem, which turns out to be a bivariate quadratic POP:

$$\begin{aligned} v^*(\varepsilon) = \min_{x \in \mathbb{R}^2} \quad & x_1 \\ \text{s.t.} \quad & x_1^2 + x_2^2 = 1 \\ & 1 - \varepsilon + x_1 - (1 - \varepsilon)x_2 \geq 0 \\ & 1 - \varepsilon + x_1 + (1 - \varepsilon)x_2 \geq 0, \end{aligned} \tag{1}$$

$\log \varepsilon^{-1}$	$v_1(\varepsilon)$	$v_2(\varepsilon)$	$v_3(\varepsilon)$	$v_4(\varepsilon)$	$v_5(\varepsilon)$	$v_6(\varepsilon)$	$v_7(\varepsilon)$	$v_8(\varepsilon)$	$v^*(\varepsilon)$
1	-0.900000	-0.014972	0	0	0	0	0	0	0
2	-0.990000	-0.639195	0	0	0	0	0	0	0
3	-0.999000	-0.952678	-0.270070	0	0	0	0	0	0
4	-0.999900	-0.995117	-0.850752	-0.014093	0	0	0	0	0
5	-0.999990	-0.999510	-0.983399	-0.601522	0	0	0	0	0
6	-0.999999	-0.999951	-0.998321	-0.945228	-0.228194	0	0	0	0
7	-1.000000	-0.999995	-0.999832	-0.994312	-0.830694	-0.004678	0	0	0
8	-1.000000	-1.000000	-0.999983	-0.999429	-0.980874	-0.560890	0	0	0
9	-1.000000	-1.000000	-0.999998	-0.999943	-0.998062	-0.937191	-0.188023	0	0
∞	-1	-1	-1	-1	-1	-1	-1	-1	-1

Table 1: Lower bounds $v_d(\varepsilon)$ on the value $v^*(\varepsilon)$ of POP (2) obtained with the moment-SOS hierarchy, for increasing relaxation orders d and different values of ε .

where $\varepsilon \in [0, 1]$ is a given parameter. This toy program is motivated by sparse optimization problems involving the l_0 pseudonorm: it was recently observed that such problems are well-structured when projected onto the unit sphere, as $l_0(x) = \mathcal{L}_0\left(\frac{x}{\|x\|_2}\right)$, where \mathcal{L}_0 is a proper lower semicontinuous convex function, introduced in [4]. Consequently, such problems can be reformulated as convex programs over the unit sphere, where \mathcal{L}_0 — hence l_0 — can be approximated as closely as desired by a polyhedral function [14].

For every $\varepsilon > 0$, the non-convex feasible set of (1) is the half circle $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, x_1 \geq 0\}$. The first and second affine inequality constraints are tight for $x \in \{(-1 + \varepsilon, 0), (0, 1)\}$ and $x \in \{(-1 + \varepsilon, 0), (0, -1)\}$, respectively. Geometrically it follows that $v^*(\varepsilon) = 0$ for all $\varepsilon \in (0, 1]$ and $v^*(0) = -1$, so that the value function is lower semi-continuous, with a discontinuity at 0.

2.2 Circle and parabola

Let us replace the two affine constraints in POP (1) with a parabolic inequality constraint that is tight for $x \in \{(-1 + \varepsilon, 0), (0, 1), (0, -1)\}$:

$$\begin{aligned}
v^*(\varepsilon) = \min_{x \in \mathbb{R}^2} \quad & x_1 \\
\text{s.t.} \quad & x_1^2 + x_2^2 = 1 \\
& 1 - \varepsilon + x_1 - (1 - \varepsilon)x_2^2 \geq 0,
\end{aligned} \tag{2}$$

where $\varepsilon \in [0, 1]$ is a given parameter. The value function $v^*(\varepsilon)$ is unchanged.

On Table 1 we report the values of the lower bounds v_d on the value v^* of POP (2) obtained with the moment-SOS hierarchy and the high-precision semidefinite solvers ClusteredLowRankSolver [15] and Loraine [7] (both output similar results), for increasing relaxation orders $d = 1, \dots, 8$ and different values of $\log 1/\varepsilon$. We report 0 when the bound is less than 10^{-60} in absolute value.

We observe that the hierarchy converges to the optimal value at finite relaxation order, but convergence is slower when ε tends to zero. A similar table can be obtained with standard semidefinite solvers, e.g., Mosek [1] or SeDuMi [28], implemented in double floating-point precision, but the output values happen to be very inaccurate. For instance the first two entries of the first row obtained with either Mosek

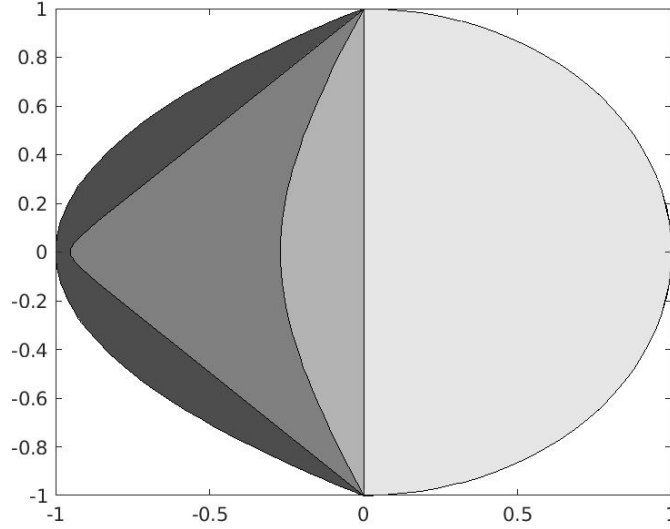


Figure 1: Nested projections on the first degree moments of the moment relaxations for increasing orders $d = 1, \dots, 4$, from dark to clear gray.

or SeDuMi are approximately equal to -0.63 and 0 , instead of -0.90 and -0.149 , respectively.

On Figure 1 we represent the nested projections on the first degree moments of the moment relaxations for orders $d = 1, \dots, 4$, for the case $\varepsilon = 3 \cdot 10^{-3}$. They are outer approximations of the convex hull of the feasible set, namely the half disk $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, x_1 \geq 0\}$. While the relaxation of order 4 is the half disk, we observe that the lower order relaxations are not tight.

2.3 Univariate parabolic reduction

Letting $x_2^2 = 1 - x_1^2$, we can reformulate bivariate POP (2) as the univariate POP

$$\begin{aligned}
 v^*(\varepsilon) = \min_{x \in \mathbb{R}} \quad & x \\
 \text{s.t.} \quad & 1 - x^2 \geq 0 \\
 & x + (1 - \varepsilon)x^2 \geq 0,
 \end{aligned} \tag{3}$$

where the constraint $1 - x^2 \geq 0$ keeps track of positivity of x_2^2 . The feasible set of POP (3) is the segment $[0, 1]$ if $\varepsilon \in (0, 1]$. If $\varepsilon = 0$, the feasible set is the non-convex union of $[0, 1]$ and $\{-1\}$. Next we recall background on real algebraic geometry and the moment-SOS hierarchy of semidefinite relaxations to approximate as closely as desired the solution of POP (3).

Let $\mathbb{R}[x]_d$ denote the vector space of polynomials of x of degree up to d , and let $\Sigma[x]_{2d} \subset \mathbb{R}[x]_{2d}$ denote the convex cone of polynomials that can be expressed as sums of squares (SOS) of polynomials of degree up to d . For more details on the definitions and results mentioned hereafter, we refer the interested reader to [18].

Given a set of polynomials $S = \{g_1, \dots, g_m\} \subset \mathbb{R}[x]$ and an integer d , let $\mathcal{M}(S)_{2d}$ be the *truncated quadratic module* of degree $2d$ generated by S , i.e., $\mathcal{M}(S)_{2d} := \{\sum_{j=0}^m \sigma_j g_j : \sigma_j \in \Sigma[x]_{2d}, \sigma_j g_j \in \mathbb{R}[x]_{2d}\}$ with $g_0 := 1$. The *quadratic module* $\mathcal{M}(S)$ generated by S is $\mathcal{M}(S) := \bigcup_{d \in \mathbb{N}} \mathcal{M}(S)_{2d}$. It is clear that every polynomial from $\mathcal{M}(S)$ is nonnegative on $K(S) := \{x \in \mathbb{R} : g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$. We say that $\mathcal{M}(S)$ is *saturated* if it coincides with the full cone of polynomials nonnegative on $K(S)$. We say that $\mathcal{M}(S)$ is *stable* if, for every degree d , there exists $k = k(d)$ such that every polynomial of degree d in $\mathcal{M}(S)$ can be represented in $\mathcal{M}(S)_{2k}$. Given $p(x) = \sum_{a=0}^{2d} p_a x^a \in \mathbb{R}[x]_{2d}$ and $y \in \mathbb{R}^{2d+1}$, define the linear functional ℓ_y such that $\ell_y(p(x)) = \sum_{a=0}^{2d} p_a y_a$. The relaxation of order d of the moment-SOS hierarchy for POP (3) consists of solving the primal moment problem

$$\begin{aligned} \text{mom}_d(\varepsilon) &= \inf_{y \in \mathbb{R}^{2d+1}} \ell_y(x) \\ \text{s.t.} & \quad \ell_y(p(x)) \geq 0, \forall p \in \Sigma[x]_{2d} \\ & \quad \ell_y((1-x^2)p(x)) \geq 0, \forall p \in \Sigma[x]_{2(d-1)} \\ & \quad \ell_y((x-(1-\varepsilon)x^2)p(x)) \geq 0, \forall p \in \Sigma[x]_{2(d-1)} \\ & \quad \ell_y(1) = 1, \end{aligned} \quad (4)$$

and the dual SOS problem

$$\begin{aligned} \text{sos}_d(\varepsilon) &= \sup_{q,r,s,v} v \\ \text{s.t.} & \quad x - v = q(x) + r(x)(1-x^2) + s(x)(x + (1-\varepsilon)x^2) \\ & \quad q \in \Sigma[x]_{2d}, r \in \Sigma[x]_{2(d-1)}, s \in \Sigma[x]_{2(d-1)}, v \in \mathbb{R}. \end{aligned} \quad (5)$$

Using the above notations, the equality constraint from (5) can be rewritten as $x - v \in \mathcal{M}(S_\varepsilon)_{2d}$ with $S_\varepsilon := \{1 - x^2, x + (1 - \varepsilon)x^2\}$.

Thanks to the constraint $1 - x^2 \geq 0$, these are semidefinite optimization problems without duality gap [11]:

$$v_d(\varepsilon) := \text{mom}_d(\varepsilon) = \text{sos}_d(\varepsilon).$$

We say that the relaxation of order d is exact when $v_d(\varepsilon) = v^*(\varepsilon)$. A sufficient condition for exactness is when this latter value is attained in the SOS dual, namely

$$x - v^*(\varepsilon) = q(x) + r(x)(1-x^2) + s(x)(x + (1-\varepsilon)x^2), \quad (6)$$

for some SOS polynomials q, r, s of appropriate degrees.

The values of the lower bounds $v_d(\varepsilon)$ on the value $v^*(\varepsilon)$ of POP (3) obtained with the moment-SOS hierarchy are the same as the one reported in Table 1.

3 Analysis

The aim of this section is to explain analytically how the numerical values of Table 1 follow the staircase pattern of Figure 2. The gray region on Figure 2 corresponds to values of d and ε for which the relaxation is exact, i.e., $v_d(\varepsilon) = v^*(\varepsilon)$. The steps of the staircase correspond to limit values ε_d such that for all $\varepsilon \geq \varepsilon_d$, the relaxation of order d is exact.

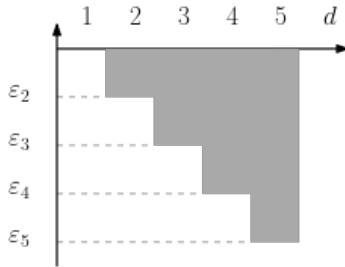


Figure 2: Staircase pattern of the moment-SOS hierarchy: the gray region corresponds to values of relaxation order d and parameter ε for which the relaxation is exact. The steps correspond to threshold values $\varepsilon_2 \geq \varepsilon_3 \geq \dots \geq \varepsilon_5$ for each degree.

Proposition 1 shows that we indeed have a staircase geometry. Proposition 2 shows that the staircase has genuine steps, i.e., the limit values ε_d can be attained. Finally, Theorem 1 gives explicit lower and upper bounds on ε_d as functions of d , to quantify the slope of the staircase.

3.1 Finite convergence

Proposition 1. *For any $\varepsilon \in [0, 1]$, the moment-SOS hierarchy converges in a finite number of steps. In particular $v_d(0) = -1$ for all orders d , and for all $\varepsilon \in (0, 1]$ one has $v_d(\varepsilon) = 0$ for a finite relaxation order d depending on ε .*

Proof. If $\varepsilon = 0$, Equation (6) admits the elementary solution $v_1(0) = v^*(0) = -1$ for $q = 0$, $r = 1$ and $s = 1$, so the first relaxation is exact.

We now consider $\varepsilon \in (0, 1]$, $S_\varepsilon = \{g, g_\varepsilon\}$ with $g = 1 - x^2$ and $g_\varepsilon = x + (1 - \varepsilon)x^2$. POP (3) has a unique local and global minimizer $x^* = 0$ over the feasibility set $K(S_\varepsilon) = [0, 1]$. As suggested by an anonymous referee, we now prove that the quadratic module $\mathcal{M}(S_\varepsilon)$ is saturated. The end points of the closed interval $K(S_\varepsilon) = [0, 1]$ satisfy $g_\varepsilon(0) = 0$, $g'_\varepsilon(0) = 1 > 0$, $g(1) = 0$ and $g'(1) = -1 < 0$. Then it follows from Theorem 9.3.1 and Theorem 9.3.3 in [18] that $\mathcal{M}(S_\varepsilon)$ is saturated, and in particular x belongs to $\mathcal{M}(S_\varepsilon)$. In other words, the moment-SOS hierarchy converges in finitely many steps for POP (3). \square

Remark 1. *For the case $\varepsilon \in (0, 1]$ we present an alternative proof of Proposition 1 based on the results from [20]. We first check that the so-called constraint qualification holds at each local minimizer, i.e., that the corresponding gradient evaluations of the polynomials involved in the active constraints are linearly independent. Since $g(0) = 1$ and $g_\varepsilon = 0$, there is only one active constraint at the global minimizer $x^* = 0$ so the constraints are qualified.*

We now introduce the Lagrangian

$$L(x, \mu) = x - \mu_1(1 - x^2) - \mu_2(x + (1 - \varepsilon)x^2),$$

which gives the following first-order necessary conditions of optimality:

$$\begin{cases} 1 + 2\mu_1x - \mu_2 - 2\mu_2(1 - \varepsilon)x = 0, \\ \mu_1(1 - x^2) = 0, \\ \mu_2(x + (1 - \varepsilon)x^2) = 0, \\ x \in [0, 1], \mu \geq 0. \end{cases}$$

We deduce that there exists a unique pair of optimal multipliers $(\mu_1, \mu_2) = (0, 1)$. Since moreover, $x + (1 - \varepsilon)x^2 \geq 0$ is the only active constraint at x^* with associated multiplier $\mu_2 > 0$, strict complementarity is satisfied.

Lastly, the Jacobian of the active constraint at x^* is $J(x^*) = 1$, so that its null space in \mathbb{R} is reduced to $\{0\}$. With the second-order derivative $\nabla_x^2 L(x^*, \mu^*) = -2(1 - \varepsilon)$, we check that the second-order necessary condition of optimality

$$x^T \nabla_x^2 L(x^*, \mu^*) x \geq 0, \forall x \in \{0\}$$

is satisfied, and so is the second-order sufficient condition of optimality

$$x^T \nabla_x^2 L(x^*, \mu^*) x > 0, \text{ for all } 0 \neq x \in \{0\}.$$

We conclude that the moment-SOS hierarchy converges in finitely many steps for POP (3), from [20, Theorem 1.1].

Note that finite convergence does not necessarily come together with the existence of maximizers in the SOS hierarchy. A simple example in the univariate case is $\min_{x \in \mathbb{R}} x$ s.t. $-x^2 \geq 0$, see [2, Ex. 1.3.4] or [21, §2.5.2], where finite convergence occurs at the first order but the corresponding SOS optimization problem has no maximizer.

We now prove that, beyond finite convergence, there always exist maximizers in (5). Let $\varepsilon \in [0, 1]$ and $d \in \mathbb{N}^*$. We notice that, as $v_1(\varepsilon) \leq v_d(\varepsilon) \leq v^*(\varepsilon)$, we must have $v_d(\varepsilon) \in [-1, 0]$.

Lemma 1. *The set $\mathcal{M}(S_\varepsilon)_{2d}$ is compact and there always exists a solution (q, r, s, v) maximizing (5).*

Proof. As suggested by an anonymous referee, for $\varepsilon \in [0, 1]$, the set $K(S_\varepsilon)$ has nonempty interior, thus Lemma 4.1.4 and Theorem 10.5.1 from [18] imply that for each $d \in \mathbb{N}$ the truncated quadratic module $\mathcal{M}(S_\varepsilon)_{2d}$ is closed.

We now concentrate on proving that it is bounded. Let (q, r, s, v) be such that $x - v = q(x) + r(x)(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2) \in \mathcal{M}(S_\varepsilon)_{2d}$. By the above remark one has $v \in [-1, 0]$. Let $I = [0.1, 0.9]$, we recall that, for any $n \in \mathbb{N}$, $\|p\|_I = \max_{x \in I} |p(x)|$ is a norm on $\mathbb{R}[x]_n$. Now, for any $\sigma \in \{q, r, s\}$ and its associated constraint $g_\varepsilon \in \{1, 1 - x^2, x + (1 - \varepsilon)x^2\}$, it holds that

$$x - v - \sigma(x)g_\varepsilon(x) \geq 0, \forall x \in I$$

with $g_\varepsilon(x) > 0$ over I . We deduce that

$$\sigma(x) \leq \max_{x \in I, v \in [-1, 0], \varepsilon \in [0, 1]} \frac{x - v}{g_\varepsilon(x)}, \forall x \in I$$

so that, as $\sigma \geq 0$, $\|\sigma\|_I \leq U$ for any upper bound $U > 0$ larger than the right-hand side in the above inequality. In particular, let us set U so that $\|\sigma\|_I \leq U$ for all $\sigma \in \{q, r, s\}$. Since $v \in [-1, 0]$, we conclude that $\mathcal{M}(S_\varepsilon)_{2d}$ is bounded, and thus compact. \square

We now consider the step value

$$\begin{aligned} \varepsilon_d &:= \inf_{\varepsilon, q, r, s} \varepsilon \\ \text{s.t. } &x = q(x) + r(x)(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2) \\ &\varepsilon \in [0, 1], q \in \Sigma[x]_{2d}, r \in \Sigma[x]_{2(d-1)}, s \in \Sigma[x]_{2(d-1)}. \end{aligned} \quad (7)$$

Proposition 2. *The infimum in (7) is attained.*

Proof. Let us assume that (ε, q, r, s) satisfies (7), so in particular $\varepsilon \in [0, 1]$ and $x \in \mathcal{M}(S_\varepsilon)_{2d}$. We deduce from Lemma 1 that the feasible set of (7) is compact, yielding the desired result. \square

Letting $x = 0$ in (7) implies that $q(0) + r(0) = 0$ and since both q and r are SOS this implies $q(0) = r(0) = 0$ and hence both contain the factor x^2 , so the equality constraint can be rewritten as

$$x = \tilde{q}(x)x^2 + \tilde{r}(x)x^2(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2), \quad (8)$$

for $\tilde{q}, \tilde{s} \in \Sigma[x]_{2(d-1)}$, $r \in \Sigma[x]_{2(d-2)}$. Factoring out x and letting $x = 0$ implies that $s(0) = 1$. Let

$$\begin{aligned} \mathcal{M}^*(S_\varepsilon)_{2d} &:= \{(\tilde{q}, \tilde{r}, s) : \tilde{q} \in \Sigma[x]_{2(d-1)}, \tilde{r} \in \Sigma[x]_{2(d-2)}, s \in \Sigma[x]_{2(d-1)}, \\ &\quad x - v^*(\varepsilon) = \tilde{q}(x)x^2 + \tilde{r}(x)x^2(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2)\} \end{aligned}$$

denote the feasible set of the dual SOS problem when the relaxation is exact.

Proposition 2 implies that for every degree d , there is a minimal value of ε , denoted ε_d , such that $\mathcal{M}^*(S_\varepsilon)_{2d}$ is non-empty. In the sequel, we provide lower and upper bounds estimates on ε_d . Section 3.2 is dedicated to lower relaxation orders, namely $d \in \{1, 2, 3\}$, while Section 3.3 focuses on arbitrary high relaxation orders.

3.2 Lower relaxations

Lemma 2. *If $\mathcal{M}^*(S_\varepsilon)_{2d}$ is non-empty then $\mathcal{M}^*(S_{\varepsilon+\delta})_{2d}$ is non-empty for all $\delta \geq 0$.*

Proof. Given $(q, r, s) \in \mathcal{M}^*(S_\varepsilon)_{2d}$, observe that $(q + \delta s x^2, r, s) \in \mathcal{M}^*(S_{\varepsilon+\delta})_{2d}$ if $\delta \geq 0$. \square

3.2.1 First relaxation

Lemma 3. *When $\varepsilon = 0$, the first relaxation is exact.*

Proof. See the beginning of the proof of Proposition 1. □

Lemma 4. *When $\varepsilon > 0$, the first relaxation is never exact.*

Proof. The first relaxation always yields the strict lower bound $v_1 = \varepsilon - 1 < v^*(\varepsilon) = 0$. Indeed equation (6) holds for $q = 0$, $r = 1 - \varepsilon$, $s = 1$ and we can use the moment dual to prove that this is optimal: the first moment relaxation for (3) writes

$$\begin{aligned} v_1(\varepsilon) = \min_{y \in \mathbb{R}^2} \quad & y_1 \\ \text{s.t.} \quad & \begin{pmatrix} 1 & y_1 \\ y_1 & y_2 \end{pmatrix} \succeq 0 \\ & 1 - y_2 \geq 0 \\ & y_1 + (1 - \varepsilon)y_2 \geq 0, \end{aligned}$$

and admits $y = (\varepsilon - 1, 1)$ as a feasible point for any $\varepsilon \in [0, 1]$. It follows that $v_1(\varepsilon) \leq \varepsilon - 1$. □

3.2.2 Second relaxation

Lemma 5. $\varepsilon_2 = 1 - \sqrt{3}/2 \approx 0.1340$.

Proof. We first prove that $\varepsilon_2 \leq 1 - \sqrt{3}/2$. Denoting $\tilde{q}(x) = q_0 + q_1x + q_2x^2$, $\tilde{r}(x) = r_0$, $s(x) = 1 + s_1x + s_2x^2$ and identifying like powers of x in equation (8) yields the linear system of equations

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 - \varepsilon & 1 \\ -1 & 0 & 1 - \varepsilon \end{pmatrix} \begin{pmatrix} r_0 \\ s_1 \\ s_2 \end{pmatrix} + \begin{pmatrix} q_0 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -1 + \varepsilon \\ 0 \\ 0 \end{pmatrix}.$$

Letting $q_0 = q_1 = q_2 = 0$, it holds

$$\begin{pmatrix} r_0 \\ s_1 \\ s_2 \end{pmatrix} = \frac{1}{\varepsilon(2 - \varepsilon)} \begin{pmatrix} (1 - \varepsilon)^3 \\ -1 + \varepsilon \\ (1 - \varepsilon)^2 \end{pmatrix}.$$

The polynomial s is SOS if and only if $4s_2 \geq s_1^2$ i.e. $-4\varepsilon^2 + 8\varepsilon - 1 = -4(\varepsilon - 1 - \sqrt{3}/2)(\varepsilon - 1 + \sqrt{3}/2) \geq 0$. It follows that the second relaxation is exact if $\varepsilon \geq 1 - \sqrt{3}/2$. When $\varepsilon = 1 - \sqrt{3}/2$, it holds $r(x) = 3\sqrt{3}/2$ and $s(x) = (1 - \sqrt{3}x)^2$. This implies that $\varepsilon_2 \leq 1 - \sqrt{3}/2$ by Lemma 2.

Next we prove that $\varepsilon_2 \geq 1 - \sqrt{3}/2$. Let $\varepsilon \in (0, 1]$. The second-order moment relaxation for (3) is given by

$$\begin{aligned} v_2(\varepsilon) = \min_{y \in \mathbb{R}^5} \quad & y_1 \\ \text{s.t.} \quad & M_2(y) \succeq 0, M_1(gy) \succeq 0, M_1(g_\varepsilon y) \succeq 0, y_0 = 1, \end{aligned} \tag{9}$$

where $M_2(y)$ is the second-order moment matrix:

$$M_1(gy) = \begin{pmatrix} y_0 - y_2 & y_1 - y_3 \\ y_1 - y_3 & y_2 - y_4 \end{pmatrix} \quad \text{and} \quad M_1(g_\varepsilon y) = \begin{pmatrix} y_1 + (1 - \varepsilon)y_2 & y_2 + (1 - \varepsilon)y_3 \\ y_2 + (1 - \varepsilon)y_3 & y_3 + (1 - \varepsilon)y_4 \end{pmatrix}. \quad (10)$$

From [11], due to the ball constraint in (3), there exists a solution y^* to the moment relaxation (9) and strong duality holds between (9) and its Lagrangian dual, the SOS relaxation. Moreover, from Lemma 1, there exists a solution (t^*, X^*) to the SOS relaxation. Thus, introducing the Lagrangian

$$L(y, t, X) = y_1 - t(y_0 - 1) - \langle M_2(y), X_0 \rangle - \langle M_1(gy), X_1 \rangle - \langle M_1(g_\varepsilon y), X_2 \rangle, \quad (11)$$

an optimal primal-dual solution (y^*, t^*, X^*) satisfies the necessary KKT conditions

$$\begin{cases} y^* \text{ is primal feasible,} \\ (t^*, X^*) \text{ is dual feasible,} \\ \nabla_y L(y^*, t^*, X^*) = 0, \\ M_2(y)X_0 = 0, M_1(gy)X_1 = 0, M_1(g_\varepsilon y)X_2 = 0, \end{cases} \quad (12)$$

which are in fact also sufficient optimality conditions since the primal problem is convex.

As a first remark, we observe that $\bar{y} = (1, 0, 0, 0, 0)$ is always feasible in (9), hence $v_2(\varepsilon) \leq 0$. Thus, since $v^*(\varepsilon) = 0$, the moment-SOS relaxation is exact if and only if \bar{y} is a solution of (9). We now reformulate as follows: the second relaxation is exact if and only if there exists a X such that $(\bar{y}, 0, X)$ satisfies the KKT conditions (12).

Let us consider (12). Clearly \bar{y} is primal feasible: $M_2(\bar{y})X_0 = 0$ and $M_1(g\bar{y})X_1 = 0$ is equivalent to $X_0^{1,1} = X_0^{1,2} = X_0^{1,3} = 0$ and $X_1^{1,1} = X_1^{1,2} = 0$. Expanding the condition $\nabla_y L(\bar{y}, 0, X) = 0$ and introducing $\eta = 1 - \varepsilon$, we conclude that $(\bar{y}, 0, X)$ satisfies the KKT conditions (12) if and only if

$$\begin{cases} X_2^{1,1} = 1 \\ X_0^{2,2} + X_1^{2,2} + \eta X_2^{1,1} + 2X_2^{1,2} = 0 \\ 2X_0^{2,3} + 2\eta X_2^{1,2} + X_2^{2,2} = 0 \\ X_0^{3,3} - X_1^{2,2} + \eta X_2^{2,2} = 0 \end{cases} \quad \begin{pmatrix} X_0^{2,2} & X_0^{2,3} \\ X_0^{2,3} & X_0^{3,3} \end{pmatrix} \succeq 0, X_1^{2,2} \geq 0, X_2 \succeq 0, \quad (13)$$

which can be further simplified, observing that $X_1^{2,2} = X_0^{3,3} + \eta X_2^{2,2} \geq 0$, to obtain that $(\bar{y}, 0, X)$ satisfies the KKT conditions (12) if and only if

$$\begin{cases} X_2^{1,1} = 1 \\ X_0^{2,2} + X_0^{3,3} + \eta X_2^{2,2} + \eta X_2^{1,1} + 2X_2^{1,2} = 0 \\ 2X_0^{2,3} + 2\eta X_2^{1,2} + X_2^{2,2} = 0 \end{cases} \quad \begin{pmatrix} X_0^{2,2} & X_0^{2,3} \\ X_0^{2,3} & X_0^{3,3} \end{pmatrix} \succeq 0, X_2 \succeq 0. \quad (14)$$

We introduce the linear operator $\mathcal{A} : \mathbb{S}_2 \times \mathbb{S}_2 \rightarrow \mathbb{R}^3$ and the cone $\mathcal{K} = \mathbb{S}_2^+ \times \mathbb{S}_2^+$ which let us write the system (14) as $\mathcal{A}(Z) = b$ and $Z \in \mathcal{K}$. We also introduce the system $\mathcal{A}^*(u) \in \mathcal{K}^*$ and $\langle u, b \rangle < 0$, which gives

$$\begin{pmatrix} u_2 & u_3 \\ u_3 & u_2 \end{pmatrix} \succeq 0, \begin{pmatrix} u_1 + \eta u_2 & u_2 + \eta u_3 \\ u_2 + \eta u_3 & \eta u_2 + u_3 \end{pmatrix} \succeq 0, u_1 < 0. \quad (15)$$

From Farkas lemma for cones, either (14) is limit-feasible or (15) has a solution, but not both [6, Lemma 4.5.6].

Let \mathcal{U}_η denote the set defined by (15), and let us investigate under which conditions it is empty when $\eta \in [0, 1)$. First, we observe that if $u \in \mathcal{U}_\eta$, the polynomial

$$(\eta^2 - 1)u_2^2 - \eta^2 u_3^2 - \eta u_2 u_3 + u_1(\eta u_2 + u_3) = \det \begin{pmatrix} u_1 + \eta u_2 & u_2 + \eta u_3 \\ u_2 + \eta u_3 & \eta u_2 + u_3 \end{pmatrix} \quad (16)$$

must have all its terms except $-\eta u_2 u_3$ nonpositive: we deduce that $u_3 \leq 0$. Then, we obtain

$$u \in \mathcal{U}_\eta \iff \begin{cases} u_1 < 0, u_2 \geq 0, u_3 \leq 0, \\ u_2 \geq -u_3, \\ u_1 + \eta u_2 \geq 0, \\ \eta u_2 + u_3 \geq 0, \\ (u_1 + \eta u_2)(\eta u_2 + u_3) - (u_2 + \eta u_3)^2 \geq 0. \end{cases} \quad (17)$$

As \mathcal{U}_η is a cone, it is nonempty if and only if its intersection with the ℓ_1 -ball is nonempty: we can thus enforce $-u_1 + u_2 - u_3 = 1$ in (17) to claim that

$$\mathcal{U}_\eta \neq \emptyset \iff \exists v \in \mathbb{R}^2, \begin{cases} v_2 \geq 0, v_3 \geq 0, \\ v_2 + v_3 - 1 < 0, \\ (1 + \eta)v_2 + v_3 - 1 \geq 0, \\ \eta v_2 - v_3 \geq 0, \\ ((1 + \eta)v_2 + v_3 - 1)(\eta v_2 - v_3) - (v_2 - \eta v_3)^2 \geq 0. \end{cases} \quad (18)$$

We now introduce $\bar{v}_2(\eta) = \frac{\eta(1+2\eta)}{2(\eta+1)}$ and $\bar{v}_3(\eta) = 1 - \bar{v}_2(\eta)$ which satisfy, for any $\eta \in (\sqrt{3}/2, 1)$:

$$\begin{cases} \bar{v}_2(\eta) > 0, \bar{v}_3(\eta) > 0, \\ \bar{v}_2(\eta) + \bar{v}_3(\eta) - 1 = 0, \\ (1 + \eta)\bar{v}_2(\eta) + \bar{v}_3(\eta) - 1 > 0, \\ \eta\bar{v}_2(\eta) - \bar{v}_3(\eta) > 0, \\ ((1 + \eta)\bar{v}_2(\eta) + \bar{v}_3(\eta) - 1)(\eta\bar{v}_2(\eta) - \bar{v}_3(\eta)) - (\bar{v}_2(\eta) - \eta\bar{v}_3(\eta))^2 = \frac{\eta^2(4\eta^2-3)}{4(\eta+1)} > 0. \end{cases} \quad (19)$$

We deduce that, for any $\eta \in (\sqrt{3}/2, 1)$, there exists a small enough perturbation $\xi_\eta > 0$ such that $(\bar{v}_2(\eta) - \xi_\eta, \bar{v}_3(\eta) - \xi_\eta)$ is a solution of the system in (18). It follows that $\mathcal{U}_\eta \neq \emptyset$, which implies that (14) is infeasible, and therefore that the second order relaxation is not exact.

We conclude that $\varepsilon_2 \geq 1 - \sqrt{3}/2$, yielding the desired result. \square

3.2.3 Third relaxation

Lemma 6. $\varepsilon_3 \leq 1 - \frac{\sqrt{3+12\sqrt{10} \sin\left(\frac{\arctan(3\sqrt{111})}{3} + \frac{\pi}{6}\right)}}{6} \approx 4.712527 \cdot 10^{-3}$.

Proof. Let $\tilde{q}(x) = 0$, $\tilde{r}(x) = r_0 + r_1x + r_2x^2$, $s(x) = 1 + s_1x + s_2x^2 + s_3x^3 + s_4x^4$, identifying like powers of x in equation (8) yields the linear system of equations

$$\left(\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1-\varepsilon & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1-\varepsilon & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1-\varepsilon & 1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1-\varepsilon \end{array} \right) \begin{pmatrix} r_0 \\ r_1 \\ r_2 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} -1 + \varepsilon \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

whose solutions can be parametrized with s_3 and s_4 as

$$\begin{pmatrix} r_0 \\ r_1 \\ r_2 \\ s_1 \\ s_2 \end{pmatrix} = \begin{pmatrix} \frac{(1-\varepsilon)^3}{\varepsilon(2-\varepsilon)} + s_3 \\ (1-\varepsilon)s_3 + s_4 \\ \frac{(1-\varepsilon)s_4}{\frac{-1+\varepsilon}{\varepsilon(2-\varepsilon)} - s_3} \\ \frac{(1-\varepsilon)^2}{\varepsilon(2-\varepsilon)} - s_4 \end{pmatrix}. \quad (20)$$

The polynomial \tilde{r} is SOS if and only if

$$\begin{pmatrix} \frac{(1-\varepsilon)^3}{\varepsilon(2-\varepsilon)} + s_3 & \frac{1-\varepsilon}{2}s_3 + \frac{1}{2}s_4 \\ \frac{1-\varepsilon}{2}s_3 + \frac{1}{2}s_4 & (1-\varepsilon)s_4 \end{pmatrix} \succeq 0,$$

and the polynomial s is SOS if and only if

$$\begin{pmatrix} 1 & \frac{-1+\varepsilon}{2\varepsilon(2-\varepsilon)} - \frac{1}{2}s_3 & z \\ \frac{-1+\varepsilon}{2\varepsilon(2-\varepsilon)} - \frac{1}{2}s_3 & \frac{(1-\varepsilon)^2}{\varepsilon(2-\varepsilon)} - s_4 - 2z & \frac{1}{2}s_3 \\ z & \frac{1}{2}s_3 & s_4 \end{pmatrix} \succeq 0,$$

for some $z \in \mathbb{R}$.

Assume that $s(x) = (1 + ax + bx^2)^2$ and $\tilde{r}(x) = (c + dx)^2$ for $a, b, c, d \in \mathbb{R}$, i.e., the above positive semidefinite matrices have rank one. So let us express system (20) in terms of the coefficients a, b, c, d :

$$\begin{pmatrix} c^2 \\ 2cd \\ d^2 \\ \frac{2a}{2b + a^2} \end{pmatrix} = \begin{pmatrix} \frac{(1-\varepsilon)^3}{\varepsilon(2-\varepsilon)} + 2ab \\ (1-\varepsilon)2ab + b^2 \\ \frac{(1-\varepsilon)b^2}{\frac{-1+\varepsilon}{\varepsilon(2-\varepsilon)} - 2ab} \\ \frac{(1-\varepsilon)^2}{\varepsilon(2-\varepsilon)} - b^2 \end{pmatrix}.$$

Clearing up denominators, this is a system of 5 polynomial equations in 5 unknowns a, b, c, d, ε . Using Maple's Groebner basis engine we can eliminate a, b, c, d and obtain the following polynomial

$$(64\varepsilon^6 - 384\varepsilon^5 + 944\varepsilon^4 - 1216\varepsilon^3 + 812\varepsilon^2 - 216\varepsilon + 1)(4\varepsilon^2 - 8\varepsilon + 1)^5(\varepsilon - 1)^8,$$

vanishing at each solution. The first factor is a degree 6 polynomial with 4 real roots. The Galois group of this polynomial allows the roots to be expressed with radicals. The smallest positive real root is

$$1 - \sqrt{\frac{1}{12} + \frac{5}{3}\alpha^{-\frac{1}{3}} + \frac{1}{6}\alpha^{\frac{1}{3}}}, \quad \alpha = -1 + 3i\sqrt{111},$$

which can be expressed with trigonometric functions as in the statement of the lemma. For this value of ε we have $a \approx -6.9296$, $b \approx 6.6375$, $c \approx -3.5866$, $d \approx 6.6219$. The desired result follows from Lemma 2. \square

3.3 Higher relaxations

The aim of this section is to provide quantitative enclosures of ε_d for arbitrary relaxation orders d . Our main result is as follows.

Theorem 1. *For all $d \in \mathbb{N}$, one has $(1 + 2d(4e)^{2d})^{-1} \leq \varepsilon_{d+1} \leq 4^{-d}$.*

The upper bound follows from Proposition 3 while the lower bound follows from Proposition 4. The remaining part of this section is dedicated to proving both propositions.

We start with a preliminary discussion to provide insights to the reader. Notice that if (6) holds with $v^*(\varepsilon) = 0$, namely if for all $x \in \mathbb{R}$

$$x = q(x) + r(x)(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2), \quad (21)$$

for some $q \in \Sigma[x]_{2d}$, $r, s \in \Sigma[x]_{2(d-1)}$, then s satisfies the following inequality

$$x - s(x)(x + (1 - \varepsilon)x^2) \geq 0, \quad \text{for all } x \in [-1, 1]. \quad (22)$$

Let us recall at this point that all nonnegative univariate polynomials with real coefficients are sums of squares of polynomials with real coefficients. Assume that there is a nonnegative polynomial $s \in \mathbb{R}[x]_{2(d-1)}$, or equivalently $s \in \Sigma[x]_{2(d-1)}$, satisfying (22). Then it is well known from [5] that there exist $q \in \Sigma[x]_{2d}$ and $r \in \Sigma[x]_{2(d-1)}$ satisfying (21). Therefore we shall restrict our attention to nonnegative polynomials s satisfying (22).

Assume first that $x \neq 0$. From (22), if $x \in (0, 1]$ then $s(x) \leq 1/(1 + (1 - \varepsilon)x)$, and if $x \in [-1, 0)$ then $s(x) \geq 1/(1 + (1 - \varepsilon)x)$. By continuity of s , this implies that $s(0) = 1$. If $x \in [0, 1]$ then $1 + (1 - \varepsilon)x \in [1, 2 - \varepsilon]$ and $1/(1 + (1 - \varepsilon)x) \in [1/(2 - \varepsilon), 1]$. If $x \in [-1, 0]$ then $1 + (1 - \varepsilon)x \in [\varepsilon, 1]$ and $1/(1 + (1 - \varepsilon)x) \in [1, 1/\varepsilon]$. It follows that s should be above resp. below the hyperbola $x \mapsto 1/(1 + (1 - \varepsilon)x)$ for x negative resp. positive, and in particular $s(-1) \geq 1/\varepsilon$, see Figure 3.

A particular choice is

$$s(x) = (ax - 1)^{2d}$$

for appropriate parameters $a \in \mathbb{R}$ and $d \in \mathbb{N}$. An anonymous referee noticed that this polynomial is surprisingly similar to the one appearing in [22, Eq. (11)]. In particular one should have $s(-1) = (a + 1)^{2d} \geq 1/\varepsilon$ and $s(1) = (a - 1)^{2d} \leq 1/(2 - \varepsilon)$.

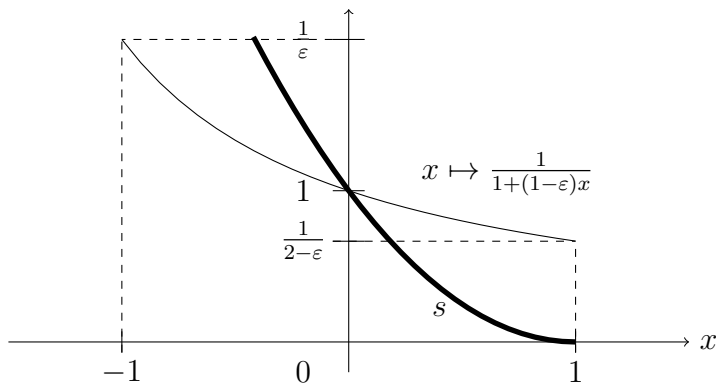


Figure 3: Constraints on SOS polynomial s (in bold).

Proposition 3. *Let $\varepsilon \in (0, 1]$, $a \in [1, 2)$ and $d \in \mathbb{N}$ be the smallest integer greater than $\frac{\log(\varepsilon^{-1})}{2\log(1+a)}$ and $\frac{\log((2-\varepsilon)^{-1})}{2\log(a-1)}$. Then the polynomial $s = (ax - 1)^{2d}$ satisfies (22) and the relaxation (5) of order $d + 1$ is exact.*

Proof. Let $z := 1 - \varepsilon$. To prove the claim, it is enough to show that s should be above resp. below the hyperbola $x \mapsto 1/(1 + zx)$ on the interval $[-1, 0]$ resp. $[0, 1]$, or equivalently that the polynomial function $x \mapsto f(x) = (ax - 1)^{2d}(1 + zx) - 1$ is above resp. below the x -axis on $[-1, 0]$ resp. $[0, 1]$. The derivative of f is given by

$$\begin{aligned} f'(x) &= 2da(ax - 1)^{2d-1}(1 + zx) + z(ax - 1)^{2d} \\ &= (ax - 1)^{2d-1}[za(2d + 1)x + 2da - z]. \end{aligned}$$

The first factor is below resp. above the x -axis on $[-1, a^{-1}]$ resp. $[a^{-1}, 1]$. The second factor is below resp. above the x -axis for $x \leq \frac{z-2da}{za(2d+1)}$ resp. $[\frac{z-2da}{za(2d+1)}, 1]$. Since $f(0) = 0$ and $0 \in [\frac{z-2da}{za(2d+1)}, a^{-1}]$, one just needs to verify the sign of the values of f at -1 and 1 . One has $f(-1) = (1 + a)^{2d}(1 - z) - 1 = \varepsilon(1 + a)^{2d} - 1$ and $f(1) = (a - 1)^{2d}(1 + z) - 1 = (a - 1)^{2d}(2 - \varepsilon) - 1$, thus by assumption one has $f(-1) \geq 0$ and $f(1) \leq 0$, which yields the desired claim. \square

Proposition 3 immediately implies the upper bound estimate from Theorem 1 for the choice $a = 1$ where $d = \lceil \frac{\log(\varepsilon^{-1})}{2\log 2} \rceil$, yielding that $\varepsilon_d \leq 4^{-d+1}$ and thus $\varepsilon_{d+1} \leq 4^{-d}$.

Remark 2. *If one assumes that $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$ then there is an algorithm computing an exact SOS decomposition (21) with rational coefficients in $\tilde{O}(\log(\frac{1}{\varepsilon})^4)$ boolean operations¹. Indeed from Proposition 3 let us select $a = 1$, $s = (x - 1)^{2d}$ with $2d = O(\log(\varepsilon^{-1}))$, so that $t : x \mapsto x - s(x)(x + (1 - \varepsilon)x^2)$ is nonnegative on $[-1, 1]$. This nonnegativity condition is equivalent to nonnegativity of $x \mapsto (1 + x^2)^{2d}t(\frac{x^2-1}{1+x^2})$ on \mathbb{R} . This latter polynomial has coefficients with maximal bit size upper bounded by $\tilde{O}(d)$. As a consequence of [16, Theorem 24], there exists an algorithm computing*

¹ $\tilde{O}(\cdot)$ stands for the variant of the big- O notation ignoring logarithmic factors.

an exact SOS decomposition of $(1+x^2)^{2d}t\left(\frac{x^2-1}{1+x^2}\right)$ with rational coefficients in $\tilde{O}(d^4)$ boolean operations.

We now focus on the lower bound estimate. Let $\varepsilon \in (0, 1]$ and $s \in \Sigma[x]_{2d}$ satisfying (22). As discussed above, the following properties hold:

$$\begin{cases} s(0) = 1, \\ s(x) \leq 1, \forall x \in [0, 1], \\ s(x) \geq 0, \forall x \in \mathbb{R}. \end{cases} \quad (23)$$

We denote by S_d the set of polynomials of degree $2d$ satisfying (23).

Proposition 4. For any $d \in \mathbb{N}$, $\max_{s \in S_d} s(-1)$ is upper bounded by $1 + 2d(4e)^{2d}$.

Proof. From [25, Lemma 4.1], one has $|s_k| \leq (4e)^{2d} \max_{0 \leq j \leq 2d} s\left(\frac{j}{2d}\right) \leq (4e)^{2d}$, for all $k \in \{0, \dots, 2d\}$ as s is upper bounded by 1 on $[0, 1]$. Thus, $s(-1) = \sum_{k=0}^{2d} (-1)^k s_k \leq 1 + 2d(4e)^{2d}$. \square

Proposition 4 immediately implies the lower bound estimate from Theorem 1 since any $s \in \Sigma[x]_{2d}$ satisfying (22) must in particular satisfy $s(-1) \geq 1/\varepsilon$.

We end this section by additional numerical experiments suggested by an anonymous referee. The optimal value of the quasi-convex SOS program (7) can be approximated as closely as desired by applying a suitable bisection scheme on ε . Doing so, we provide interval enclosures of the exact value of ε_d , for integer values of d between 2 and 10. The corresponding results are displayed in Figure 4. The dotted curves correspond to the lower and upper bounds derived in Theorem 1.

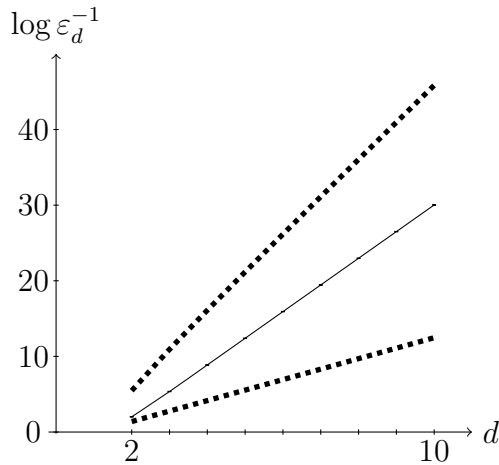


Figure 4: Values of $\log \varepsilon_d^{-1}$ obtained by solving (7) for values of $d \in \{2, \dots, 10\}$.

In particular we retrieve an interval enclosure of $\varepsilon_2 \in [0.1339740, 0.1339750]$ whose exact value, given in Lemma 5, is $1 - \sqrt{3}/2 \approx 0.1339745$. We could also numerically certify that $\varepsilon_3 \in [4.712522 \cdot 10^{-3}, 4.7125518 \cdot 10^{-3}]$, which shows that the theoretical upper bound $4.712527 \cdot 10^{-3}$ provided in Lemma 6 is actually very close to the true value of ε_3 . The function $d \mapsto \log \varepsilon_d^{-1}$ has an approximate linear growth rate of 3.5, between the lower and upper bounds from Theorem 1. The Julia source code implementing these experiments is available online.²

4 Discussions and further remarks

In this note we studied a specific SOS decomposition of a parametrized univariate polynomial to design an elementary POP for which the moment-SOS hierarchy shows finite yet arbitrarily slow convergence when the parameter tends to a limit.

As suggested by an anonymous referee, we end this section with a discussion about the finite convergence behavior of the moment-SOS hierarchy, related to the saturation and stability properties of the considered quadratic module. By definition of the moment-SOS hierarchy, if the quadratic module $\mathcal{M}(S)$ is saturated then the moment-SOS hierarchy has finite convergence, and $\sup = \max$ is attained in the SOS part, when we reach the finite convergence. If $\mathcal{M}(S)$ is stable, then the moment-SOS hierarchy stabilizes. These results can be found in, e.g., [18, Chapter 10].

Saturated and stable quadratic module The univariate case with nonempty interior is essentially the unique case, where these two properties can happen simultaneously, and we provide such an example in our study. In particular, the quadratic module $\mathcal{M}(S_\varepsilon)$ of problem (3) is both saturated and stable, for $\varepsilon > 0$. Saturation has been shown in the proof of Proposition 1. In addition, one can show that $\mathcal{M}(S_\varepsilon)$ is stable by choosing a so-called *natural description* for $K(S_\varepsilon) = [0, 1]$ with the generator $h = x(1 - x)$. For more details about natural descriptions see [18], especially Proposition 2.7.3 and its proof. This notion of stability depends only on the quadratic module, and not on the particular finite system of generators. Let us consider a polynomial $p \in \mathcal{M}(h)$ of degree d . Since $\mathcal{M}(S_\varepsilon)$ is saturated, p is nonnegative on $K(\{h\})$, so admits by [5] a representation $p = \sigma_0 + \sigma_1 h$ with SOS polynomials σ_0 of degree $2\lceil d/2 \rceil$ and σ_1 of degree $2\lceil d/2 \rceil - 2$. So every polynomial of degree d in $\mathcal{M}(\{h\})$ can be represented in $\mathcal{M}(\{h\})_{2k}$ with $k = \lceil d/2 \rceil$, thus $\mathcal{M}(h) = \mathcal{M}(S_\varepsilon)$ is stable. Moreover, if we fix a univariate quadratic module as above (stable+saturated) and let the objective function vary (keeping the degree bounded) we cannot observe that the order of finite convergence goes to infinity, as this would contradict the stability of the quadratic module.

Neither saturated nor stable quadratic module In [27], Stengle focuses on finding similar SOS decompositions associated to the minimization of the polynomial $1 - x^2 + \varepsilon$ (parametrized by $\varepsilon > 0$) on $K(S) = [-1, 1]$, where the latter set is encoded as the super-level set of $(1 - x^2)^3$. Stengle proves in [27, Theorem 4] that finite

²<https://homepages.laas.fr/vmagron/files/slowuniv.jl>

convergence is reached after a number of iterations proportional to the reciprocal square root of the parameter, i.e. $1/\sqrt{\varepsilon}$. In this case the semialgebraic set is fixed and the associated quadratic module is neither saturated nor stable. The lack of saturation is due to the fact that $1 - x^2$ is nonnegative on $K(S)$ but it does not belong to the quadratic module $\mathcal{M}(\{(1 - x^2)^3\})$. The lack of stability is due to the fact that the family of polynomials $(1 - x^2 + \varepsilon)_{\varepsilon > 0}$ is contained in the module but it has arbitrary large SOS representations.

Saturated not stable quadratic module As suggested by an anonymous referee the same phenomenon can happen with a saturated quadratic module that is not stable. To observe the order of finite convergence going to infinity while keeping the semialgebraic set fixed, with an associated quadratic module being saturated, it is necessary to go one dimension higher. For instance, Example 3.7 in [19] is a candidate. Indeed, let us consider the preordering, i.e., the quadratic module generated by all the products of $S = \{p_1, p_2, p_3, p_4\}$ as in the example, and as objective functions the linear polynomials $l_a = a_0 + a_1 t_1 + a_2 t_2 \in \mathbb{R}[t_1, t_2]$ parametrized by $a \in \mathbb{R}$ such that l_a is nonnegative on $K(S)$. In this case the preordering is saturated, and thus the moment-SOS hierarchy has finite convergence for every objective polynomial. But the order of finite convergence is going to infinity, as the family of linear polynomials has representations of unbounded degree by [19, Theorem 3.5]. Moreover, since the semialgebraic set has nonempty interior, $\sup = \max$ in the SOS hierarchy at every order.

Extension to the multivariate case In our case Theorem 1 shows that finite convergence is reached after a number of iterations proportional to the parameter bit size inverse $\log(1/\varepsilon)$. Overall the results from [27] imply that for univariate POPs the minimal relaxation order required for finite convergence could be exponential in the bit size of the coefficients involved in the objective function. We also point out [23, Theorem 7] where the authors consider SOS representations associated to the minimization of $1 + x + \varepsilon$ on $K(S) = [-1, 1]$, where the latter set is encoded as the super-level set of an odd power of $1 - x^2$. If the power is greater than 3 then they provide upper bounds on the minimal order that are proportional to the reciprocal of the parameter, i.e. $1/\varepsilon$. Our complementary study shows that the minimal order can be linear in the bit size of the coefficients involved in the constraints. Further research could focus on studying quantitatively higher-dimensional cases where the finite convergence order goes to infinity while keeping the semialgebraic set fixed; as suggested by an anonymous referee [19, Example 3.7] could be a candidate. The same behavior happens for the following (contrived) multivariate problem built from the univariate one from our paper, e.g., when minimizing the polynomial $x_1 + \dots + x_n$ on the set of constraints $1 - x_i^2 \geq 0, x_i + (1 - \varepsilon)x_i^2 \geq 0, i = 1, \dots, n$. Similar results hold for the bivariate polynomial optimization problem (2) from Section 2.2, since any SOS representation associated to (3) provides one for (2). Indeed suppose that there exists $d \in \mathbb{N}$ such that $x_1 \in \mathcal{M}(S_\varepsilon)_{2d}$, that is, $x_1 = q + r(1 - x_1^2) + s(x_1 + (1 - \varepsilon)x_1^2)$ for some $q \in \Sigma[x_1]_{2d}$, and $r, s \in \Sigma[x_1]_{2(d-1)}$. Then $x_1 = q + rx_1^2 + (r - (1 - \varepsilon)s)(1 - x_1^2 - x_1^2) + s(1 - \varepsilon + x_1 - (1 - \varepsilon)x_1^2)$, so $x_1 \in \mathcal{M}(S_\varepsilon^2)_{2d}$

with $S_\varepsilon^2 = \{\pm 1 - x_1^2 - x_2^2, 1 - \varepsilon + x_1 - (1 - \varepsilon)x_2^2\}$.

Influence of the number of connected components and solver precision

As in [27], the description of the set of constraints plays a crucial role in the behavior of the moment-SOS hierarchy. For $\varepsilon \in (0, 1]$, our considered set of constraints is the interval $[0, 1]$. When this interval is described by either the super-level set of the single quadratic polynomial $x(1 - x)$ or by the intersection of the ones of x and $1 - x$, the hierarchy immediately converges at order 1. The present work shows that a slight modification of this description, by means of only two quadratic polynomials, can have significant impact on the efficiency of the moment-SOS hierarchy. What happens here is that the interval defined by the first constraint almost intersects one side of the second constraint (as ε goes to zero). One might think that the convergence behavior tends to be bad when the semialgebraic sets defined by constraints almost intersect. Further research efforts should be pursued to investigate how bad the convergence behavior could be when the number of connected components of the feasible set grows.

In addition to slow convergence behaviors, complementary studies such as [29] have emphasized that inappropriate constraint descriptions may provide wrong relaxation bounds when relying on standard double-precision semidefinite solvers. A possible remedy consists of using instead a multiple precision semidefinite solver, which, again, might slow down significantly the required computation.

Further investigation should focus on means to appropriately describe POP constraint sets, in order to obtain sound numerical solutions at lower relaxation orders and with lower solver accuracy.

Acknowledgements

The authors thank Pauline Kergus for suggesting the polynomial from Section 3.3 and the anonymous referees for their very insightful feedback. The authors are also grateful to Michal Kočvara and Nando Leijenhorst for their advice regarding the use of the high-precision SDP solvers `Loraine` and `ClusteredLowRankSolver`, respectively. This work benefited from the `FastOPF2` grant funded by RTE, the `EPOQCS` grant funded by the `LabEx CIMI (ANR-11-LABX-0040)`, the `HORIZON-MSCA-2023-DN-JD` of the European Commission under the Grant Agreement No 101120296 (`TENORS`), the `AI Interdisciplinary Institute ANITI` funding, through the French “Investing for the Future `PIA3`” program under the Grant agreement n° `ANR-19-PI3A-0004` as well as the `National Research Foundation, Prime Minister’s Office, Singapore` under its `Campus for Research Excellence and Technological Enterprise (CREATE)` programme. This work was performed within the project `COMPUTE`, funded within the `QuantERA II Programme` that has received funding from the `EU’s H2020 research and innovation programme` under the `GA No 101017733`.

References

- [1] E. D. Andersen, K. D. Andersen. The MOSEK interior point optimizer for linear programming: an implementation of the homogeneous algorithm. In H. Frenk, K. Roos, T. Terlaky, S. Zhang (Editors). *High Performance Optimization*, 197-232. Springer, 2000.
- [2] L. Baldi. *Représentations effectives en géométrie algébrique réelle et optimisation polynomiale*. PhD thesis, Inria Univ. Côte d’Azur, 2022.
- [3] L. Baldi, L. Slot. Degree bounds for Putinar’s Positivstellensatz on the hypercube. *SIAM J. Applied Algebra and Geometry* 8(1):1-25, 2024.
- [4] J.P. Chancelier, M. De Lara. Hidden convexity in the l_0 pseudonorm. *J. Convex Analysis* 28(1):203-236, 2021.
- [5] M. Fekete. Proof of three propositions of Paley. *Bull. Amer. Math. Soc.* 41:138–144, 1935.
- [6] B. Gärtner, J. Matousek. *Approximation algorithms and semidefinite programming*. Springer Science & Business Media, 2012.
- [7] S. Habibi, M. Kočvara, M. Stingl. Loraine - an interior-point solver for low-rank semidefinite programming. *Optimization Methods and Software*, 1-31, 2023.
- [8] D. Henrion. Moments for polynomial optimization - An illustrated tutorial. Lecture notes of a course given for the programme Recent Trends in Computer Algebra, Institut Henri Poincaré, Paris, 2023.
- [9] D. Henrion, M. Korda, J. B. Lasserre. *The moment-SOS hierarchy*. World Scientific, 2020.
- [10] D. Henrion, J. B. Lasserre. *Solving global optimization problems over polynomials with GloptiPoly 2.1*. LNCS 2861:43-58, Springer, 2003.
- [11] C. Josz, D. Henrion. Strong duality in Lasserre’s hierarchy for polynomial optimization. *Optim. Letters* 1(10):3-10, 2016.
- [12] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optimization* 11(3):796–817, 2001.
- [13] J. B. Lasserre. *Moments, positive polynomials and their applications*. World Scientific, 2009.
- [14] A. Le Franc, J.P. Chancelier, M. De Lara. The Capra-subdifferential of the l_0 pseudonorm. *Optimization* 73(4):1229-1251, 2022.
- [15] N. Leijenhurst, D. de Laat. Solving clustered low-rank semidefinite programs arising from polynomial optimization. *Math. Prog. Comput.* 16:503–534, 2024.

- [16] V. Magron, M. Safey El Din, M. Schweighofer. Algorithms for weighted sum of squares decomposition of non-negative univariate polynomials. *J. Symbol. Comput.* 93:200-220, 2019.
- [17] V. Magron, M. Safey El Din, M. Schweighofer, T. H. Vu. Exact SOHS decompositions of trigonometric univariate polynomials with Gaussian coefficients. *Proc. Int. Symp. on Symbolic and Algebraic Comput. (ISSAC)*, 325-332, 2022.
- [18] M. Marshall. *Positive polynomials and sums of squares*. American Mathematical Society, 2008.
- [19] T. Netzer, D. Plaumann, M. Schweighofer. Exposed faces of semidefinitely representable sets. *SIAM Journal on Optimization*, 20(4):1944-1955, 2010.
- [20] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. *Math. Prog.* 146:97-121, 2014.
- [21] J. Nie. *Moments and polynomial optimization*. SIAM, 2023.
- [22] J. Nie and M. Schweighofer. On the complexity of Putinar’s Positivstellensatz. *Journal of Complexity*, 23(1):135-150, 2007.
- [23] V. Powers and B. Reznick. Polynomials that are positive on an interval. *Trans. Amer. Math. Soc.*, 352(10):4677-4692, 2000.
- [24] G. Pólya and G. Szegő. *Problems and theorems in analysis II*, Springer, 1976.
- [25] A. A. Sherstov. Making polynomials robust to noise. *Proceedings of the ACM Symposium on Theory of Computing*, 747-758, 2012.
- [26] L. Slot. Asymptotic analysis of semidefinite bounds for polynomial optimization and independent sets in geometric hypergraphs. PhD thesis, CWI Amsterdam and Tilburg Univ., 2022.
- [27] G. Stengle. Complexity estimates for the Schmüdgen Positivstellensatz, *Journal of Complexity*, 12:167-174, 1996.
- [28] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625-653, 1999.
- [29] H. Waki, M. Nakata and M. Muramatsu. Strange behaviors of interior-point methods for solving semidefinite programming problems in polynomial optimization. *Computational Optimization and Applications*, 53:823-844, 2012.