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

Didier Henrion[1,2], Adrien Le Franc[1], Victor Magron[1]

Draft of March 12, 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 [9] 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 [6, 5, 15]. 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 [14], 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 [4, 17] (see also [10, § 2.3] for a modern exposition) when 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 [12, 13].

[1]CNRS; LAAS; Université de Toulouse, 7 avenue du colonel Roche, F-31400 Toulouse, France.
[2]Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, CZ-16626 Prague, Czechia.

Restricting our attention to POPs on compact sets, several elementary examples without finite convergence are known. A simple example in the univariate case is $\min_{x\in\mathbb{R}} x$ s.t. $-x^2 \geq 0$, see [1, Ex. 1.3.4] or [15, §2.5.2]. A less contrived 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 [2, Prop. 29]. Many results on the speed of convergence of the moment-SOS hierarchy are now available, see, e.g., [1, 19]. 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 [7]. 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

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

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. We estimate the order of finite convergence as a function of ε.

Our contributions can be summarized as follows:

- Finite convergence holds for all $\varepsilon \in [0,1]$ and integer d (Theorem 1);
- There exist steps / threshold values $\varepsilon_d \in [0,1]$ for the exactness of the relaxation of order d (Theorem 2);
- We provide bounds $\underline{\varepsilon}_d \leq \varepsilon_d \leq \overline{\varepsilon}_d$ to estimate the convergence speed of the hierarchy (Theorem 3).

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

2 POP design

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

2.1 Circle and two lines

The following example 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 $l_0(x) = \mathcal{L}_0(\frac{x}{||x||_2})$, where $\mathcal{L}_0$ is a proper lower semicontinuous convex function, introduced in [3]. Consequently, such problems

$\log 1/\varepsilon$	$v_1(\varepsilon)$	$v_2(\varepsilon)$	$v_3(\varepsilon)$	$v_4(\varepsilon)$	$v^*(\varepsilon)$
1	$-6.3212 \cdot 10^{-1}$	0	0	0	0
2	$-8.6466 \cdot 10^{-1}$	0	0	0	0
3	$-9.5021 \cdot 10^{-1}$	$-1.4794 \cdot 10^{-1}$	0	0	0
4	$-9.8168 \cdot 10^{-1}$	$-4.5993 \cdot 10^{-1}$	0	0	0
5	$-9.9326 \cdot 10^{-1}$	$-7.3383 \cdot 10^{-1}$	0	0	0
6	$-9.9752 \cdot 10^{-1}$	$-8.8837 \cdot 10^{-1}$	$-5.1426 \cdot 10^{-2}$	0	0
7	$-9.9909 \cdot 10^{-1}$	$-9.5672 \cdot 10^{-1}$	$-2.9832 \cdot 10^{-1}$	0	0
8	$-9.9966 \cdot 10^{-1}$	$-9.8376 \cdot 10^{-1}$	$-6.0489 \cdot 10^{-1}$	0	0
9	$-9.9988 \cdot 10^{-1}$	$-9.9398 \cdot 10^{-1}$	$-8.2055 \cdot 10^{-1}$	$-1.9947 \cdot 10^{-3}$	0
∞	-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.

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 [11].

It is in this context that we introduced the following toy problem, which turns out to be a bivariate quadratic POP:

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

where $\varepsilon \in [0,1]$ is a given parameter. The non-convex feasible set is the half circle $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1, x_1 \geq 0\}$. The first affine constraint is saturated for $x \in \{(-1+\varepsilon, 0), (0,1)\}$ and the second affine constraint is saturated for $x \in \{(-1+\varepsilon, 0), (0,-1)\}$. 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 constraint saturated for $x \in \{(-1+\varepsilon, 0), (0,1), (0,-1)\}$:

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

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 semidefinite solver SeDuMi, for increasing relaxation orders $d = 1, \ldots, 4$ and different values of $\log 1/\varepsilon$. We report 0 when the bound is less than 10^{-7} in absolute value.

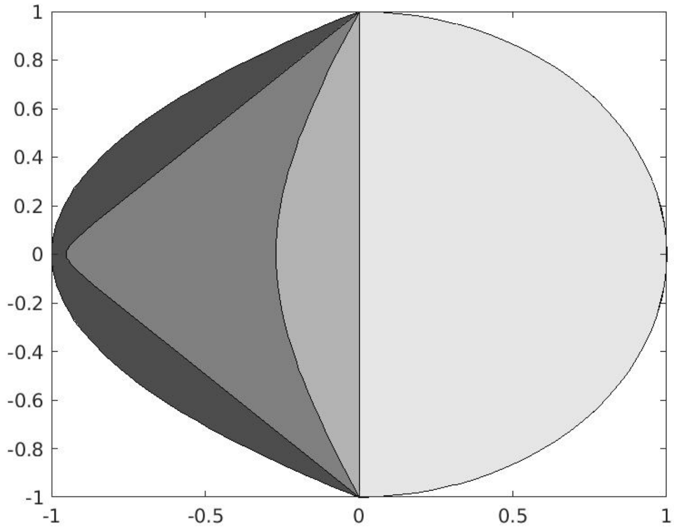

Figure 1: Nested projections on the first degree moments of the moment relaxations for increasing orders $d = 1, \ldots, 4$, from dark to clear gray.

We observe that the hierarchy converges to the optimal value at finite relaxation order, but convergence is slower when ε tends to zero.

On Figure 1 we represent the nested projections on the first degree moments of the moment relaxations for orders $d = 1, \ldots, 4$, 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

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

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 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. 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

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

and the dual SOS problem

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

Thanks to the constraint $1 - x^2 \geq 0$, these are semidefinite optimization problems without duality gap [8]:

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

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

$$x - v^\star(\varepsilon) = q(x) + r(x)(1-x^2) + s(x)(x+(1-\varepsilon)x^2), \tag{6}$$

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., $v_d(\varepsilon) = v^*(\epsilon)$. The steps of the staircase correspond to limit values ε_d such that for all $\varepsilon \geq \epsilon_d$, the relaxation of order d is exact.

Theorem 1 shows that we indeed have a staircase geometry. Theorem 2 shows that the staircase has genuine steps, i.e., the limit values ε_d can be attained. Finally, Theorem 3 gives explicit lower and upper bounds on ε_d as functions of d, to quantify the slope of the staircase.

3.1 Finite convergence

Theorem 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.*

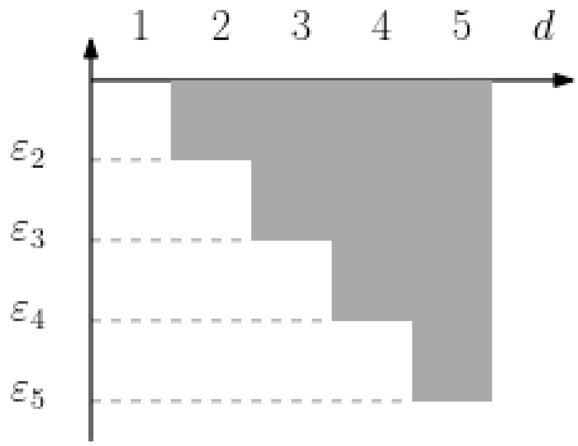

Figure 2: Staircase pattern of the moment-SOS hierarchy: the gray region corresponds to values of relaxation order d and parameter ε for which the relaxation is exact. The steps correspond to threshold values $\varepsilon_2 \geq \varepsilon_3 \geq \cdots \varepsilon_5$ for each degree.

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 $\varepsilon \in (0, 1]$. POP (3) has a unique local and global minimizer $x^* = 0$ over its feasibility set $[0, 1]$.

Both the tangent cone and the linearized cone at x^* are equal to $\mathbb{R}_+$, so that 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:

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

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 [14, Theorem 1.1]. □

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 $v_d(\varepsilon) \in [-1,0]$. So let us denote by $S_d^{[-1,0]}(\varepsilon)$ the feasible set of the SOS relaxation (5) with the additional constraint $v \in [-1,0]$. The supremum in (5) can be equivalently searched over $S_d^{[-1,0]}(\varepsilon)$.

Lemma 1. *$S_d^{[-1,0]}(\varepsilon)$ is compact.*

Proof. The set $S_d^{[-1,0]}(\varepsilon)$ is the preimage of the closed set $\{x\}$ by the linear map $(q,r,s,v) \mapsto v+q(x)+r(x)(1-x^2)+s(x)(x+(1-\varepsilon)x^2)$ defined on a finite dimensional space. It is therefore a closed set.

We now concentrate on proving that it is bounded. Let $(q,r,s,v) \in S_d^{[-1,0]}(\varepsilon)$, which thus satisfies $x - v = q(x) + r(x)(1-x^2) + s(x)(x+(1-\varepsilon)x^2)$. 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 $\sigma \in \{q,r,s\}$. It follows that

$$S_d^{[-1,0]}(\varepsilon) \subset [-1,0] \times \mathbb{B}_{||\cdot||_I}(0,U)^3 \,.$$

We conclude that $S_d^{[-1,0]}(\varepsilon)$ is bounded, and thus compact. □

We deduce from Lemma 1 that, besides finite convergence of the moment-SOS hierarchy shown in Theorem 1, there always exists a solution (q,r,s,v) maximizing (5).

We now consider the step value

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

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

Proof. A solution of $q(x)+r(x)(1-x^2)+s(x)(x+(1-\varepsilon)x^2)-x=0$ in the variables (ε,q,r,s) is equivalently characterized by a system of $2d+1$ quadratic equations in ε and the coefficients of (q,r,s). It follows that the feasible set of (7) is the preimage of the closed set $\{0\}$ by a continuous quadratic map: it is therefore a closed set.

Now, if (ε,q,r,s) is feasible for (7), then $\varepsilon\in[0,1]$ and $(q,r,s,0)\in F_d(\varepsilon)$. We deduce from Lemma 1 that the feasible set of (7) is bounded, and thus compact. □

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

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

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

$$\begin{aligned} S_d(\varepsilon) \;\; := \;\; &\{q\in\Sigma[x]_{2(d-1)}, r\in\Sigma[x]_{2(d-2)}, s\in\Sigma[x]_{2(d-1)}:\\ &\; x-v^\star(\varepsilon) = q(x)x^2 + r(x)x^2(1-x^2) + s(x)(x+(1-\varepsilon)x^2)\}\end{aligned}$$

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

Lemma 2. *If $S_d(\varepsilon)$ is non-empty then $S_d(\varepsilon+\delta)$ is non-empty for all $\delta\geq 0$.*

Proof. Given $(q_\varepsilon,r_\varepsilon,s_\varepsilon)\in S_d(\varepsilon)$, observe that $(q_\varepsilon+\delta s_\varepsilon x^2, r_\varepsilon, s_\varepsilon)\in S_d(\varepsilon+\delta)$ if $\delta\geq 0$. □

Theorem 2 implies that for every degree d, there is a minimal value of ε, denoted ε_d, such that $S_d(\varepsilon_d)$ 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

3.2.1 First relaxation

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

Proof. See the beginning of the proof of Theorem 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

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

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

3.2.2 Second relaxation

Lemma 5. $\epsilon_2 \leq 1 - \sqrt{3}/2 \approx 0.1340$.

Proof. Denoting $q(x) = q_0 + q_1 x + q_2 x^2$, $r(x) = r_0$, $s(x) = 1 + s_1 x + s_2 x^2$ 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 $s(x) = (1 - \sqrt{3}x)^2$. $\square$

3.2.3 Third relaxation

Lemma 6. $\varepsilon_3 \leq 1 - \dfrac{\sqrt{3 + 12\sqrt{10}\sin\left(\frac{\arctan(3\sqrt{111})}{3} + \frac{\pi}{6}\right)}}{6} \approx 4.7125 \cdot 10^{-3}$.

Proof. Let $q(x) = 0$, $r(x) = r_0 + r_1 + r_2 s^2$, $s(x) = 1 + s_1 x + s_2 x^2 + s_3 x^3 + s_4 x^4$, identifying like powers of x in equation (8) yields the linear system of equations

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

whose solutions can be parametrized with s_3 and s_4 as

$$\left(\begin{array}{c} r_0 \\ r_1 \\ r_2 \\ \hline s_1 \\ s_2 \end{array}\right) = \left(\begin{array}{c} \frac{(1-\varepsilon)^3}{\varepsilon(2-\varepsilon)} + s_3 \\ (1-\varepsilon)s_3 + s_4 \\ (1-\varepsilon)s_4 \\ \hline \frac{-1+\varepsilon}{\varepsilon(2-\varepsilon)} - s_3 \\ \frac{(1-\varepsilon)^2}{\varepsilon(2-\varepsilon)} - s_4 \end{array}\right). \tag{9}$$

The polynomial 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 $r(x) = (c + dx)^2$ for $a, b, c, d \in \mathbb{R}$, i.e., the above positive semidefinite matrices have rank one. So let us express system (9) in terms of the coefficients a, b, c, d:

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

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 $a \approx -6.9296$, $b \approx 6.6375$, $c \approx -3.5866$, $d \approx 6.6219$. □

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

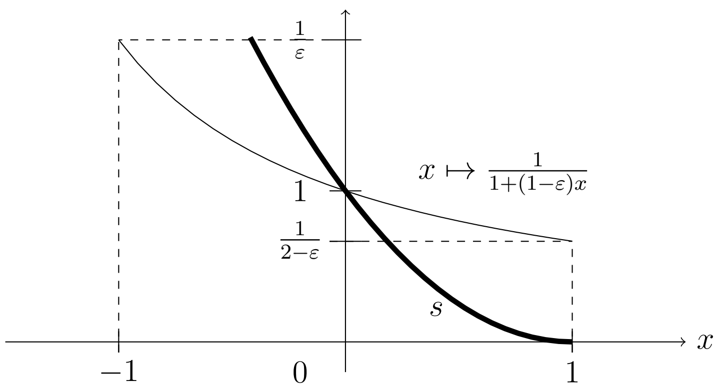

Figure 3: Constraints on SOS polynomial s (in bold).

The upper bound follows from Proposition 1 while the lower bound follows from Proposition 2. 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 $v^\star(\varepsilon) = 0$, namely if for all $x \in \mathbb{R}$

$$x = q(x) + r(x)(1 - x^2) + s(x)(x + (1 - \varepsilon)x^2), \tag{10}$$

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

$$x - s(x)(x + (1 - \varepsilon)x^2) \geq 0, \quad \text{for all } x \in [-1, 1]. \tag{11}$$

Conversely 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 (11). Then it is well known from [4] that there exist $q \in \Sigma[x]_{2d}$ and $r \in \Sigma[x]_{2(d-1)}$ satisfying (10). Therefore we shall restrict our attention to nonnegative polynomials s satisfying (11).

Assume first that $x \neq 0$. From (11), 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 suggested by Pauline Kergus is

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

for appropriate parameters $a \in \mathbb{R}$ and $d \in \mathbb{N}$. In particular one should have $s(-1) = (a + 1)^{2d} \geq 1/\varepsilon$ and $s(1) = (a - 1)^{2d} \leq 1/(2 - \varepsilon)$.

Proposition 1. *Let $\varepsilon \in (0, 1]$, $a \in [1, 2)$ and $d \in \mathbb{N}$ be 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* (11) *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

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

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 on $[-1, \frac{z-2d}{z(a+2d)}]$ resp. $[\frac{z-2d}{z(a+2d)}, 1]$. Therefore f is nondecreasing on $[-1, \frac{z-2d}{z(a+2d)}]$, nonincreasing on $[\frac{z-2d}{z(a+2d)}, a^{-1}]$ and nondecreasing on $[a^{-1}, 1]$. Since $f(0) = 0$ and $0 \in [\frac{z-2d}{z(a+2d)}, a^{-1}]$, 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. $\square$

Remark 1. *If one assumes that $\varepsilon \in \mathbb{Q}$, $\varepsilon > 0$ then there is an algorithm computing an exact SOS decomposition* (10) *with rational coefficients in boolean time*[1] $\tilde{O}\left(\log(\frac{1}{\varepsilon})^4\right)$. *Indeed from Proposition 1 let us select $a = 1$, $s = (x - 1)^{2d}$ with $2d = O\left(\log(\varepsilon^{-1})\right)$, 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 $x \mapsto (1+x^2)^{2d} t\left(\frac{x^2-1}{1+x^2}\right)$ on $\mathbb{R}$. This latter polynomial has coefficients with maximal bit size upper bounded by $\tilde{O}(d)$. As a consequence of [12, Theorem 4.4], there exists an algorithm computing an exact SOS decomposition of $(1+x^2)^{2d} t\left(\frac{x^2-1}{1+x^2}\right)$ with rational coefficients in boolean time $\tilde{O}(d^4)$.*

We now focus on the upper bound estimate. Let $\varepsilon \in (0, 1]$ and $s \in \Sigma[x]_{2d}$ satisfying (11). As discussed above, the following properties hold:

$$\begin{cases} s(0) = 1 \, , \\ s(x) \leq 1 \, , \forall x \in [0, 1] \, , \\ s(x) \geq 0 \, , \forall x \in \mathbb{R} \, . \end{cases} \tag{12}$$

We denote by S_d the set of polynomials of degree $2d$ satisfying (12).

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

Proof. From [18, Lemma 4.1], one has $|s_k| \leq (4e)^{2d} \max\limits_{0 \leq j \leq 2d} s\left(\dfrac{j}{2d}\right) \leq (4e)^{2d}$, for all $k \in \{0, \dots, 2d\}$ as s is upper bounded by 1 on $[0, 1]$. Thus, $s(-1) = \displaystyle\sum_{k=0}^{2d} (-1)^k s_k \leq 1 + 2d(4e)^{2d}$. $\square$

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

Proposition 2 immediately implies the upper bound estimate from Theorem 3 since any $s \in \Sigma[x]_{2d}$ satisfying (11) must in particular satisfy $s(-1) \geq 1/\varepsilon$.

4 Concluding 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.

In [16], the authors focus on finding similar SOS decompositions associated to the minimization of the polynomial $1 + x + \varepsilon$ (parametrized by $\varepsilon > 0$) on $[-1, 1]$, where the latter set is encoded as the super-level set of a power p of $1 - x^2$. They prove in [16, Theorem 7] that finite convergence is reached after a number of iterations proportional to the power p and the reciprocal $1/\varepsilon$ of the parameter. In our case Theorem 3 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 [16] 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. Our complementary study shows that the minimal order could be linear in the bit size of the coefficients involved in the constraints.

As in [16], 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. In addition to slow convergence behaviors, complementary studies such as [20] 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.

References

[1] L. Baldi. Représentations effectives en géométrie algébrique réelle et optimisation polynomiale. PhD thesis, Inria Univ. Côte d'Azur, 2022.

[2] L. Baldi, L. Slot. Degree bounds for Putinar's Positivstellensatz on the hypercube. SIAM J. Applied Algebra and Geometry 8(1):1-25, 2024.

[3] J.P. Chancelier, M. De Lara. Hidden Convexity in the l_0 Pseudonorm. J. Convex Analysis 28(1):203-236, 2021.

[4] M. Fekete. Proof of three propositions of Paley. Bull. Amer. Math. Soc. 41:138–144, 1935

[5] 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.

[6] D. Henrion, M. Korda and J. B. Lasserre. The moment-SOS hierarchy. World Scientific, 2020.

[7] D. Henrion, J. B. Lasserre. Solving global Optimization Problems over Polynomials with GloptiPoly 2.1. LNCS 2861:43-58, Springer, 2003.

[8] C. Josz, D. Henrion. Strong duality in Lasserre's hierarchy for polynomial optimization. Optim. Letters 1(10):3-10, 2016.

[9] J. B. Lasserre. Global optimization with polynomials and the problem of moments. SIAM J. Optimization 11(3):796–817, 2001.

[10] J. B. Lasserre. Moments, positive polynomials and their applications (Vol. 1). World Scientific, 2009.

[11] A. Le Franc, J.P. Chancelier, M. De Lara. The Capra-subdifferential of the l_0 pseudonorm. Optimization 73(4):1229-1251, 2022.

[12] V. Magron, M. Safey El Din, and M. Schweighofer. Algorithms for weighted sum of squares decomposition of non-negative univariate polynomials. J. Symbol. Comput. 93:200-220, 2019.

[13] V. Magron, M. Safey El Din, M. Schweighofer and 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.

[14] J. Nie. Optimality conditions and finite convergence of Lasserre's hierarchy. Math. Prog. 146:97-121, 2014.

[15] J. Nie. Moments and polynomial optimization. SIAM, 2023.

[16] V. Powers and B. Reznick. Polynomials that are positive on an interval. Trans. Amer. Math. Soc., 352(10):4677-4692, 2000.

[17] G. Pólya and G. Szegö. Problems and theorems in analysis II, Springer, 1976.

[18] A. A. Sherstov. Making polynomials robust to noise. Proceedings of the forty-fourth annual ACM symposium on Theory of computing, 747-758, 2012.

[19] L. Slot. Asymptotic analysis of semidefinite bounds for polynomial optimization and independent sets in geometric hypergraphs. PhD thesis, CWI Amsterdam and Tilburg Univ., 2022.

[20] 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.