

Convergence rate of the moment-SOS hierarchy for univariate polynomial optimization

Didier Henrion^{*} and Mohab Safey El Din[†]

September 1, 2026

Abstract

We study the convergence rate of the moment-SOS (sum-of-squares) hierarchy for polynomial optimization problems (POPs) on a bounded subset of the real line described by arbitrary polynomial inequalities. We prove that, for every fixed univariate POP, the relaxation error is bounded by $O(1/r^2)$, where r is the relaxation order. In particular, boundary degeneracies in the polynomial description of the feasible set affect the constant but not the convergence exponent. The proof combines the structure of finitely generated univariate quadratic modules with a Chebyshev polynomial construction that approximately recovers the natural generators of the feasible set while controlling the degree of the certificate. We also give an elementary degree-four example for which the relaxation error is exactly $1/(2r(r-1))$, showing that the quadratic rate is optimal. Equivalent reformulations connect this example to a cubic univariate problem and to a bivariate POP whose feasible set has a cusp singularity at the minimizer.

1 Introduction

The moment-SOS hierarchy for POP. A polynomial optimization problem (POP) consists of minimizing globally a given polynomial subject to polynomial inequalities:

$$f^* := \min_x f(x) \text{ s.t. } x \in K := \{x \in \mathbb{R}^n : p_i(x) \geq 0, i = 1, \dots, n_p\} \quad (1)$$

where $f, p_1, \dots, p_{n_p} \in \mathbb{R}[x]$ are given polynomials. It is assumed that K is non-empty and compact, so that a minimizer exists. In full generality, (1) is a difficult non-convex problem, potentially with several minimizers. It includes as particular cases

^{*}LAAS-CNRS, Université de Toulouse, France, and Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia.

[†]Sorbonne Université, CNRS, LIP6, Paris, France.

all combinatorial optimization problems. The moment-SOS (sums of squares) hierarchy, as originally proposed by Lasserre [18], is a computationally efficient approach for polynomial optimization. It consists of constructing a family of convex relaxations of increasing size, generating a sequence of lower bounds on the minimum. Under some conditions, solving a sufficiently large convex relaxation is equivalent to solving the original POP. Let $\Sigma[x] \subset \mathbb{R}[x]$ denote the cone of polynomials that can be written as sums of squares (SOS), and let $p_0(x) \equiv 1$ for notational convenience. The dual SOS formulation of the moment-SOS hierarchy consists of solving the semidefinite optimization problems

$$f_r := \sup_{v \in \mathbb{R}} v \text{ s.t. } f - v \in \mathcal{Q}_r(p) \quad (2)$$

in the truncated quadratic module

$$\mathcal{Q}_r(p) := \left\{ \sum_{i=0}^{n_p} \sigma_i p_i, \sigma_i \in \Sigma[x], \deg(\sigma_i p_i) \leq 2r \right\} \subset \mathbb{R}[x] \quad (3)$$

indexed by $r \in \mathbb{N}$ such that $2r \geq \max\{\deg f, \deg p_1, \dots, \deg p_{n_p}\}$. This problem is called the SOS relaxation of order r . See [12] for an elementary tutorial in the moment-SOS hierarchy for polynomial optimization, [22, 31] for recent overviews, and [20, 27] for the underlying real algebraic geometry.

Relaxation gap and convergence. An SOS relaxation is characterized by its relaxation gap, or error

$$e_r := f^* - f_r \geq 0$$

which is non-negative by construction. In [18], asymptotic convergence was shown:

$$e_r \geq e_{r+1} \geq \dots \geq \lim_{r \rightarrow \infty} e_r = 0$$

by using Putinar's Positivstellensatz [25] which states that any given polynomial strictly positive on K belongs to the quadratic module $\mathcal{Q}_r(p)$ if r is chosen sufficiently large. In [21], generic finite convergence was shown: $e_{r^*} = 0$ for a finite r^* depending on the POP data. Non-finite convergence may occur, i.e. $e_r > 0$ for all finite r , but only for specific POPs whose data $f, p_1, \dots, p_{n_p}$ lie on a low-dimensional manifold. In this paper, we are precisely interested in these instances, and we want to quantify the speed at which the error sequence e_r converges to zero.

Convergence rates. Quantifying the speed of this convergence amounts to bounding the degree of the SOS multipliers needed to certify positivity. All the estimates below concern the quadratic module (3) and the Putinar Positivstellensatz. Sharper rates are typically available for the preordering, which also allows products of the p_i as multipliers in the Schmüdgen Positivstellensatz, but the corresponding relaxations involve 2^{n_p} instead of $n_p + 1$ semidefinite blocks, so that the Putinar-type rates are the practically relevant ones. The first degree bound, exponential in the inverse of the accuracy, was given in [23, Thm. 6] and yields only an inverse-logarithmic rate

[23, Thm. 8]. Whether this exponential dependence could be removed stayed open for more than a decade, until [3] obtained the first general polynomial degree bound, governed by a Łojasiewicz exponent attached to the description of the set; see [5] for refinements and [11] for a recent sharpening based on a lift-and-project construction. Much stronger rates are known on distinguished sets, through the polynomial kernel method and its Christoffel-Darboux reformulation: a rate $O(1/r^2)$ on the sphere for homogeneous f [9], extended to arbitrary f in [7]. On the hypercube $[-1, 1]^n$, [6, Thm. 3, Cor. 15] prove a rate $O(1/r)$, whereas the only available lower bound is of order $1/r^8$ [6, Thm. 4, Cor. 16]; a new family of squared approximation kernels has recently improved the upper rate to $O(\log^3 r/r^2)$ [10], leaving the lower bound as the loose end. This illustrates the fact that the study of the convergence rate of the moment-SOS hierarchy is currently an active research topic, see the recent overviews [19, 28].

Univariate POP. In this paper, we study the convergence rate of the moment-SOS hierarchy in the particular case of a univariate POP, i.e. $n = 1$ in (1). Unless the degree d of f is very large, the univariate POP (1) is not challenging as an optimization problem. Indeed, if $z_1, \dots, z_{d-1}$ denote the roots of the derivative of f , then a global minimizer on K is attained among the finitely many points $(\{z_1, \dots, z_{d-1}\} \cap K) \cup \{a_1, b_1, \dots, a_{m-1}, b_{m-1}\}$, and there is no need for a sophisticated machinery such as the moment-SOS hierarchy. However, studying the worst case convergence rate in this simple univariate setting provides significant insight on the weaknesses of the hierarchy. Indeed, many challenging multivariate POPs with specific degeneracies reduce to univariate POPs, see e.g. [14].

Contributions In the univariate POP setup (1), our main contributions are twofold. First, we prove that there is a constant C (depending on the POP data) such that

$$0 \leq e_r \leq \frac{C}{r^2}$$

for sufficiently large r . It means that the moment-SOS hierarchy cannot converge slower than $O(1/r^2)$. Second, we describe an elementary degree 4 example achieving exactly this convergence rate, i.e. our convergence rate is tight. Our example is a simplification of a well-known example due originally to Stengle [30], recently solved in rational arithmetic [13]. Finally, we relate our example, through equivalent reformulations, to an elementary cubic POP with two constraints, and to a bivariate POP whose feasible set has a cusp singularity at the minimizer [20, Ex. 9.4.6(3)], [4, Ex. A.3], [15], thereby illustrating how multivariate POPs with specific degeneracies reduce to the univariate setting studied here.

2 Natural Generators

Since we assume that the set $K \subset \mathbb{R}$ in (1) is bounded and semi-algebraic, it can be expressed as a finite union of bounded intervals

$$K = [a_1, b_1] \cup [a_2, b_2] \cup \cdots \cup [a_{m-1}, b_{m-1}]$$

for some reals $a_i \leq b_i < a_{i+1}$ and integer $m > 1$. Associated to this representation is the natural choice of generators defined in [16, §2.3], see also [27, §6.1.2]:

$$g_1(x) := x - a_1, \quad g_2(x) := (x - b_1)(x - a_2), \dots, g_m(x) = b_{m-1} - x$$

such that

$$K = \{x \in \mathbb{R} : g_i(x) \geq 0, \quad i = 1, \dots, m\}.$$

Example 1. Let $K = [-1, 0] \cup [1, 2]$. Then

$$g_1(x) = x + 1, \quad g_2(x) = x(x - 1), \quad g_3(x) = 2 - x.$$

For $f(x) = g_1(x)g_2(x) = x^3 - x$ it holds $f^* = \min_K f = 0$. The SOS relaxation (2) of order $r = 2$ yields $e_2 = 6 - 4\sqrt{2} \approx 0.34315$ with

$$f(x) + 6 - 4\sqrt{2} = (x + \sqrt{2} - 2)^2 g_1(x) + (\sqrt{2} - 1)^2 g_2(x).$$

The SOS relaxation (2) of order $r = 3$ yields $e_3 = \frac{1}{8} = 0.125$ with

$$f(x) + \frac{1}{8} = \frac{1}{12}(x^2 - \frac{5}{2}x + 1)^2 g_1(x) + \frac{1}{4}(x + \frac{3}{2})^2 g_2(x) + \frac{1}{12}(x^2 + \frac{1}{2}x - \frac{1}{2})^2 g_3(x).$$

And finally the SOS relaxation (2) of order $r = 4$ yields $e_4 = 0$ with

$$\begin{aligned} f(x) = & x^2(x - 1)^2(\frac{1}{20}(x - \frac{3}{2})^2 + \frac{3}{16})g_1(x) + \\ & (x + 1)^2(\frac{1}{10}(x - 3)^2 + \frac{1}{10})g_2(x) + \frac{1}{20}x^2(x^2 - 1)^2 g_3(x). \end{aligned}$$

In this case the moment-SOS hierarchy converges at the 4th relaxation order:

$$e_2 = 6 - 4\sqrt{2} > e_3 = \frac{1}{8} > e_4 = 0.$$

This example was studied in [17, Ex. 4.4(a)] as a limit case of the parametric problem

$$K_N = [-1, 0] \cup [1, N], \quad N \geq 2.$$

It was shown that $g_1 g_2 \in \mathcal{Q}_{r_N}(g) \setminus \mathcal{Q}_{r_N-1}(g)$ with truncation degree r_N not uniformly bounded in N , i.e. $\lim_{N \rightarrow \infty} r_N = \infty$. In other words, the moment-SOS hierarchy has finite convergence when K_N is bounded, but the convergence can be arbitrarily slow. Note that an affine change of variable shows that the same phenomenon appears with the uniformly bounded set $[0, \frac{1}{N}] \cup [\frac{2}{N}, 1]$. A similar parametric univariate example with arbitrarily slow convergence was studied in [14].

Note also that this arbitrary slow convergence of the moment-SOS hierarchy is a property of the chosen natural generator presentation, not merely of the set and objective. The same set can be described by the single quartic generator:

$$K_N = \{x \in \mathbb{R} : h_N(x) := (x + 1)x(x - 1)(N - x) \geq 0\}$$

and the moment-SOS hierarchy then admits a uniform order-3 certificate for every $N \geq 2$, namely

$$f(x) = \frac{f(x)^2}{N(N^2 - 1)} + \frac{\left(x + \frac{N}{2}\right)^2 + \frac{3N^2 - 4}{4}}{N(N^2 - 1)} h_N(x).$$

Indeed, h_N is the product of the natural generators of the compact set K_N , which has no isolated point, so that $\mathcal{Q}(h_N)$ already contains every polynomial nonnegative on K_N [27, Ex. 6.5.17(3)].

Note that (3) is a finite-dimensional truncation of the quadratic module

$$\mathcal{Q}(p) := \left\{ \sum_{i=0}^{n_p} \sigma_i p_i, \sigma_i \in \Sigma[x] \right\} \subset \mathbb{R}[x].$$

Lemma 1 (Archimedean property). *There exists $R \geq 1$ such that $R^2 - x^2 \in \mathcal{Q}(p)$.*

Proof. Since K is a compact subset of the real line, we can use [20, Thm. 7.1.2] or [27, Prop. 5.5.14], which establishes that $\mathcal{Q}(p)$ is Archimedean. The fact that $R \geq 1$ comes from [20, Cor. 5.2.4]. $\square$

Lemma 2 (Markov-Lukács). *Let $R > 0$. If $q \in \mathbb{R}[x]$ is nonnegative on $[-R, R]$ and has degree d , then*

$$q \in \mathcal{Q}_{\lceil \frac{d}{2} \rceil}(R^2 - x^2).$$

Proof. This is the classical Markov-Lukács theorem, see [24, Sec. 2] for a modern account.

For $q = 0$ the result is trivial. If $d = 0$, then q is a nonnegative constant and the result is immediate. Hence assume $q \neq 0$ and $d \geq 1$. Up to scaling by $1/R$, one can assume that $R = 1$ without loss of generality. Let $\mathbf{G}_m(q)$ be the Goursat transform of q of degree m , i.e. $\mathbf{G}_m(q)(x) = (1+x)^m q\left(\frac{1-x}{1+x}\right)$. Observe that $2^m q = \mathbf{G}_m(\mathbf{G}_m(q))$.

By Goursat's lemma [24, Lemma 1], q is nonnegative on $[-1, 1]$ if and only if $\mathbf{G}_d(q)$ is nonnegative on $[0, +\infty[$. Moreover, it holds that $\deg \mathbf{G}_d(q) \leq d := \deg q$. By [24, Prop. 2], if $\mathbf{G}_d(q)$ is nonnegative on $[0, +\infty[$, then there exist two sums of squares s_0 and s_1 , with $\deg(s_0) \leq d$ and $\deg(s_1) \leq d - 1$, such that $\mathbf{G}_d(q) = s_0 + x s_1$. Now, we perform a Goursat transform of degree d on both sides of this identity. We obtain

$$\begin{aligned} 2^d q &= (1+x)^d s_0 \left(\frac{1-x}{1+x} \right) + (1+x)^d \left(\frac{1-x}{1+x} \right) s_1 \left(\frac{1-x}{1+x} \right) \\ &= \mathbf{G}_d(s_0) + (1-x)(1+x)^{d-1} s_1 \left(\frac{1-x}{1+x} \right). \end{aligned}$$

Note that $\mathbf{G}_d(s_0)$ has degree at most d . Since s_0 (resp. s_1) is a sum of squares, there exist $a_1, \dots, a_u$ (resp. $b_1, \dots, b_v$) such that $s_0 = a_1^2 + \dots + a_u^2$ (resp. $s_1 = b_1^2 + \dots + b_v^2$). Recall that we established that $\deg s_0 \leq d$ and $\deg s_1 \leq d - 1$. We deduce that

$\deg a_i \leq \frac{d}{2}$ for all $1 \leq i \leq u$ and $\deg b_i \leq \frac{d-1}{2}$ for all $1 \leq i \leq v$. When d is even, we deduce that

$$\mathbf{G}_d(s_0) = \sum_{i=1}^u \mathbf{G}_{d/2}(a_i)^2$$

and that

$$\begin{aligned} (1+x)^{d-1} s_1 \left(\frac{1-x}{1+x} \right) &= (1+x)^{d-1} \sum_{i=1}^v b_i^2 \left(\frac{1-x}{1+x} \right) \\ &= (1+x) \sum_{i=1}^v \mathbf{G}_{(d-2)/2}(b_i)^2 \end{aligned}$$

and we conclude that

$$2^d q = \sum_{i=1}^u \mathbf{G}_{d/2}(a_i)^2 + (1-x^2) \sum_{i=1}^v \mathbf{G}_{(d-2)/2}(b_i)^2$$

which establishes that

$$q \in \mathcal{Q}_{d/2}(1-x^2).$$

When d is odd, we deduce that

$$\mathbf{G}_d(s_0) = (1+x) \sum_{i=1}^u \mathbf{G}_{(d-1)/2}(a_i)^2$$

and that

$$(1+x)^{d-1} s_1 \left(\frac{1-x}{1+x} \right) = \sum_{i=1}^v \mathbf{G}_{(d-1)/2}(b_i)^2$$

To summarize, we have

$$2^d q = (1+x) \sum_{i=1}^u \mathbf{G}_{(d-1)/2}(a_i)^{\textcolor{brown}{2}} + (1-x) \sum_{i=1}^v \mathbf{G}_{(d-1)/2}(b_i)^{\textcolor{brown}{2}}. \quad (4)$$

Using

$$1+x = \frac{(1+x)^2}{2} + \frac{1-x^2}{2}, \quad 1-x = \frac{(1-x)^2}{2} + \frac{1-x^2}{2},$$

in (4), and observing that all resulting terms have degree at most $d+1$, we obtain

$$q \in \mathcal{Q}_{(d+1)/2}(1-x^2).$$

□

Lemma 3 (Natural generator representation). *Let $g = (g_1, \dots, g_m)$ be the vector of natural generators of K , and let $g_0 \equiv 1$. Then*

$$\{q \in \mathbb{R}[x] : q \geq 0 \text{ on } K\} = \mathcal{Q}(g) = \left\{ \sum_{i=0}^m \sigma_i g_i, \sigma_i \in \Sigma[x] \right\}.$$

Proof. Every polynomial nonnegative on K belongs to the preordering generated by the natural generators g_i ; this is the saturation result of [16, Thm. 2.2], see also [20, Prop. 2.7.3] and [27, Prop. 6.1.3(a)]. Since K is bounded, the quadratic module $\mathcal{Q}(g)$ is closed under multiplication, and therefore coincides with this preordering, see [26, Cor. 4.4], [20, Thm. 9.3.1], and [27, Prop. 6.5.8] applied to the affine line. □

3 Upper Bound on Convergence Rate

3.1 Statement

We consider the problem (1) for $n = 1$:

$$f^* := \min_x f(x) \text{ s.t. } x \in K := \{x \in \mathbb{R} : p_i(x) \geq 0, i = 1, \dots, n_p\}$$

where f and $p = (p_1, \dots, p_{n_p})$ are in $\mathbb{R}[x]$. Recall that we assume that

$$K = [a_1, b_1] \cup \dots \cup [a_{m-1}, b_{m-1}]$$

is bounded and non-empty.

Theorem 1 (Quadratic convergence rate). *There exist constants $C > 0$ and $r_0 \in \mathbb{N}$ such that*

$$0 \leq e_r \leq \frac{C}{r^2}, \quad r \geq r_0. \quad (5)$$

The proof proceeds by separating the geometry of the feasible set from the degree complexity of the given quadratic module. Let $q := f - f^*$ so that $q \geq 0$ on K . As recalled by Lemma 3, since we are in the univariate case, q belongs to the quadratic module $\mathcal{Q}(g)$ of the natural generators of K . The difficulty is that the natural generators need not themselves belong to the original quadratic module $\mathcal{Q}(p)$.

The proof overcomes this difficulty in two steps.

First, the structure theory of finitely generated univariate quadratic modules shows that, for each natural generator g_j , some odd power belongs to the original module, i.e. $g_i^{d_i} \in \mathcal{Q}(p)$ for d_i odd.

Second, a Chebyshev polynomial construction converts such an odd-power certificate into an approximate certificate for the generator itself: for every $n \geq 1$, $g_i + C_i \varepsilon_n \in \mathcal{Q}_{O(n)}(p)$ with $\varepsilon_n = O(n^{-2})$. Substituting these approximate generators into the fixed natural generator representation gives $q + C \varepsilon_n \in \mathcal{Q}_{O(n)}(p)$.

Finally, choosing n proportional to the relaxation order r yields $e_r = O(1/r^2)$.

3.2 Proof

Let

$$a = \operatorname{arcosh} 3 = \log(3 + 2\sqrt{2}) = 2 \log(1 + \sqrt{2}) = 2 \operatorname{arsinh} 1 \approx 1.76275$$

and

$$\varepsilon_n = \frac{\cosh(a/n) - 1}{2} = \sinh^2\left(\frac{a}{2n}\right) = \frac{1}{4} \left(\left(1 + \sqrt{2}\right)^{1/n} - \left(\sqrt{2} - 1\right)^{1/n} \right)^2,$$

for every integer $n \geq 1$: in particular $\varepsilon_1 = 1$ and $\varepsilon_2 = \frac{\sqrt{2}-1}{2}$. Moreover

$$\lim_{n \rightarrow \infty} n^2 \varepsilon_n = \frac{a^2}{4} = \log^2(1 + \sqrt{2}) \approx 0.77682.$$

Lemma 4. For every $n \geq 1$, it holds $0 < \varepsilon_n \leq \frac{1}{n^2}$.

Proof. Because $a > 0$, it holds

$$0 < \cosh(a/n) - 1 = \sum_{k \geq 1} \frac{a^{2k}}{(2k)!n^{2k}} \leq \frac{1}{n^2} \sum_{k \geq 1} \frac{a^{2k}}{(2k)!} = \frac{\cosh a - 1}{n^2} = \frac{2}{n^2}.$$

□

Let T_n be the Chebyshev polynomial of the first kind. Recall that these are defined by the recurrence $T_{n+1}(t) = 2tT_n(t) - T_{n-1}(t)$, initialized with $T_0(t) = 1$ and $T_1(t) = t$. We also have $T_n(\cos \theta) = \cos(n\theta)$. Define

$$V_n(t) := \frac{1 - T_n(1 - 2t)}{2}.$$

Since $V_n(0) = 0$, the quotient

$$W_n(t) := \frac{V_n(t)}{t}$$

is a polynomial of degree $n - 1$.

Lemma 5 (Odd-multiplicity desingularization). *Let $d \geq 1$ be odd and define*

$$q_{n,d}(t) := \varepsilon_n + t(1 - V_n(t)^{d-1}). \quad (6)$$

Then

$$q_{n,d}(t) \geq 0 \quad \text{for every } t \leq 1, \quad (7)$$

and

$$t + \varepsilon_n = q_{n,d}(t) + t^d W_n(t)^{d-1}. \quad (8)$$

Moreover,

$$\deg q_{n,d} \leq 1 + n(d - 1), \quad \deg(t^d W_n(t)^{d-1}) \leq 1 + n(d - 1). \quad (9)$$

Proof. Since $V_n(t) = tW_n(t)$, it holds that $t^d W_n(t)^{d-1} = tV_n(t)^{d-1}$ and (8) follows immediately from (6). It remains to prove (7), i.e. $q_{n,d}(t) \geq 0$ for every $t \leq 1$.

First let $0 \leq t \leq 1$. Then $1 - 2t \in [-1, 1]$, so $-1 \leq T_n(1 - 2t) \leq 1$ and hence $0 \leq V_n(t) \leq 1$, which implies $1 - V_n(t)^{d-1} \geq 0$, and therefore $q_{n,d}(t) \geq \varepsilon_n > 0$.

Second, let $t < 0$ and set $z := 1 - 2t > 1$.

For $z > 1$, $T_n(z) = \cosh(n \operatorname{arcosh} z) > 1$, so that

$$V_n(t) = (1 - T_n(z))/2 < 0.$$

We investigate the two complementary cases $V_n(t) \leq -1$ and $-1 < V_n(t) < 0$.

If $V_n(t) \leq -1$, then $V_n(t)^{d-1} \geq 1$ because $d - 1$ is even, so that $1 - V_n(t)^{d-1} \leq 0$. As $t < 0$, $t(1 - V_n(t)^{d-1}) \geq 0$, hence again $q_{n,d}(t) \geq \varepsilon_n$.

Suppose finally that $-1 < V_n(t) < 0$. Then, since $d - 1$ is even, it holds that

$$0 < V_n(t)^{d-1} < 1 \quad \text{and} \quad T_n(z) = 1 - 2V_n(t) \leq 3 = \cosh a.$$

Because $\cosh$ is increasing on $[0, \infty)$, $\operatorname{arcosh} z \leq \frac{a}{n}$. Thus $z \leq \cosh(a/n)$, and consequently

$$|t| = (z - 1)/2 \leq (\cosh(a/n) - 1)/2 := \varepsilon_n.$$

Therefore $q_{n,d}(t) = \varepsilon_n - |t|(1 - V_n(t)^{d-1}) \geq \varepsilon_n - |t| \geq 0$. This ends the proof of (7).

Finally, note that, by definition of V_n , it holds that $\deg V_n = \deg T_n = n$ and $\deg W_n = (\deg V_n) - 1 = n - 1$, which establishes (9). $\square$

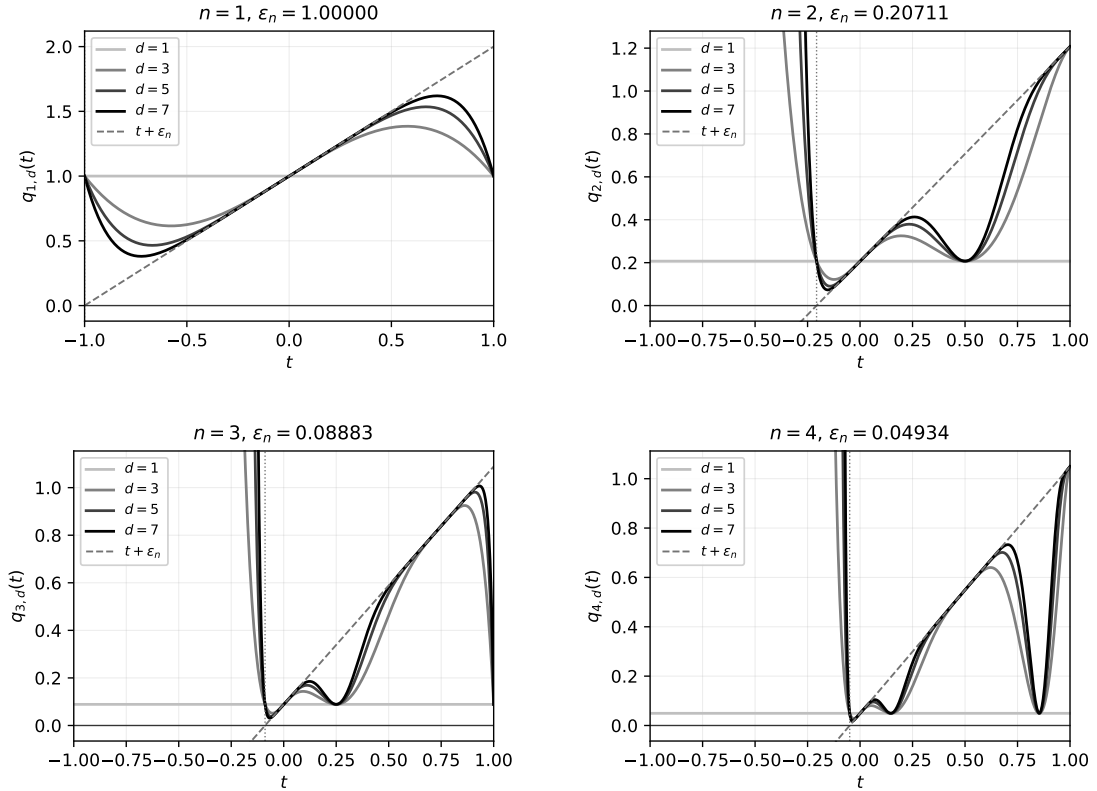


Figure 1: Graphs of the polynomials $q_{n,d}$ of Lemma 5 on $(-1, 1)$, for $n = 1, 2, 3, 4$ and odd exponents $d = 1, 3, 5, 7$. On $[0, 1]$ each graph oscillates between the constant floor ε_n , attained where $V_n(t) = 1$, and the line $t + \varepsilon_n$ (dashed), attained where $V_n(t) = 0$. On the interval $(-\varepsilon_n, 0)$, delimited by the dotted vertical line at $t = -\varepsilon_n$, the graphs dip below ε_n while remaining bounded from below by $t + \varepsilon_n \geq 0$, in accordance with the nonnegativity certificate (7). For $t < -\varepsilon_n$ one has $V_n(t) < -1$, so the polynomials increase very steeply and leave the plotted range. The degenerate case $d = 1$ yields the constant $q_{n,1} \equiv \varepsilon_n$.

Recall that $p = (p_1, \dots, p_{n_p})$ is the vector of polynomials whose simultaneous non-negativity defines

$$K = [a_1, b_1] \cup \dots \cup [a_{m-1}, b_{m-1}].$$

Recall that we assume K be bounded and non-empty.

Lemma 6 (Odd powers of the natural generators). *For every $i = 1, \dots, m$, there exists an odd integer $d_i \geq 1$ such that*

$$g_i^{d_i} \in \mathcal{Q}(p).$$

Proof. We repeatedly use the generalized natural-generator description of a finitely generated univariate quadratic module with nonempty bounded positivity set. By [2, Cor. 4.6], together with [2, Defs. 4.1 and 4.4] (which we can apply because we assume that K is bounded and non-empty), every generalized natural generator belongs to $\mathcal{Q}(p)$, see also [1, Thm. 2.26 and Cor. 2.28].

At the left boundary, [2, Cor. 4.6] establishes that there exists an odd integer $d_1 \geq 1$ such that $g_1^{d_1}(x) = (x - a_1)^{d_1}$ belongs to $\mathcal{Q}(p)$.

For each gap (b_i, a_{i+1}) , $i = 1, \dots, m - 2$, using again [2, Cor. 4.6], we deduce that there exist odd integers $\beta_i, \alpha_{i+1} \geq 1$ such that $h_{i+1}(x) := (x - b_i)^{\beta_i}(x - a_{i+1})^{\alpha_{i+1}} \in \mathcal{Q}(p)$. Set $d_{i+1} := \max\{\beta_i, \alpha_{i+1}\}$. Since β_i and α_{i+1} are odd, d_{i+1} is odd, and both $d_{i+1} - \beta_i$ and $d_{i+1} - \alpha_{i+1}$ are nonnegative even integers. Consequently,

$$\begin{aligned} g_{i+1}(x)^{d_{i+1}} &= ((x - b_i)(x - a_{i+1}))^{d_{i+1}} \\ &= h_{i+1}(x) ((x - b_i)^{(d_{i+1} - \beta_i)/2}(x - a_{i+1})^{(d_{i+1} - \alpha_{i+1})/2})^2. \end{aligned}$$

A quadratic module is closed under multiplication by polynomial squares. Hence $g_{i+1}^{d_{i+1}} \in \mathcal{Q}(p)$, for $i = 1, \dots, m - 2$.

Finally, at the right boundary, [2, Cor. 4.6] shows that there exists an odd integer $d_m \geq 1$ such that $g_m^{d_m}(x) = (b_{m-1} - x)^{d_m} \in \mathcal{Q}(p)$. $\square$

Example 2. Consider $p_1(x) = (1 - x)^3(1 + x)$, whose positivity set is $K = [-1, 1]$. The natural generators are $g_1(x) = 1 + x$, $g_2(x) = 1 - x$. It holds

$$g_1 = \frac{1}{8}g_2^2((g_2 + 1)^2 + 3) + \frac{1}{8}p_1 \text{ which lies in } \mathcal{Q}(p)$$

so $d_1 = 1$. For the other natural generator,

$$g_2^3 = \frac{1}{2}g_2^4 + \frac{1}{2}p_1 \text{ which also lies in } \mathcal{Q}(p)$$

so $d_2 = 3$. Moreover, $g_2 \notin \mathcal{Q}(p)$. Indeed, suppose that $g_2 = \sigma_0 + \sigma_1 p_1$, $\sigma_0, \sigma_1 \in \Sigma[x]$. Evaluating at $x = 1$ gives $\sigma_0(1) = 0$. Since σ_0 is a sum of squares, its zero at $x = 1$ has even multiplicity, and hence $\sigma'_0(1) = 0$. Since p_1 has a zero of multiplicity 3 at $x = 1$, one also has $p_1(1) = p'_1(1) = 0$. Differentiating the assumed identity at $x = 1$ would therefore give $g'_2(1) = 0$, contrary to $g'_2(1) = -1$. Thus $g_2 \notin \mathcal{Q}(p)$, and the smallest possible odd exponent is $d_2 = 3$. The failure of g_2 to lie in $\mathcal{Q}(p)$ is an instance of the criterion [27, Prop. 6.5.16]: on a non-singular affine curve, a finitely generated preordering with compact positivity set is saturated if and only if some generator vanishes to order exactly one at every boundary point, which fails here because p_1 vanishes at order 3.

Example 3. Consider $p_1(x) = -x^2$, whose positivity set is the singleton $K = \{0\}$. The natural generators are $g_1(x) = x$, $g_2(x) = -x$. It holds that

$$g_1^3 = x^3 = \left(\frac{x(x+1)}{2}\right)^2 + \left(\frac{x-1}{2}\right)^2 (-x^2) \in \mathcal{Q}(p)$$

so $d_1 = 3$. Similarly,

$$g_2^3 = -x^3 = \left(\frac{x(x-1)}{2}\right)^2 + \left(\frac{x+1}{2}\right)^2 (-x^2) \in \mathcal{Q}(p)$$

so $d_2 = 3$.

Neither x nor $-x$ belongs to $\mathcal{Q}(-x^2)$. For instance, if $x = \sigma_0 - \sigma_1 x^2$, $\sigma_0, \sigma_1 \in \Sigma[x]$, then evaluation at $x = 0$ gives $\sigma_0(0) = 0$. Since σ_0 is a sum of squares, this implies $\sigma_0'(0) = 0$. Differentiating the identity at $x = 0$ would yield $1 = 0$, a contradiction. The same argument applies to $-x$. Again [27, Prop. 6.5.16] predicts the failure: at an isolated point of the positivity set, saturation requires two generators of order one whose product is nonpositive nearby, whereas here p_1 vanishes at the second order at zero.

Lemma 7 (Approximate recovery of a natural generator). *Let $g \in \mathbb{R}[x]$ be such that*

$$g^d \in \mathcal{Q}(p) \tag{10}$$

for some odd integer $d \geq 1$. Then there exist constants $A > 0$, $B \geq 0$, $C > 0$ depending only on g , p , and d , such that

$$g + C\varepsilon_n \in \mathcal{Q}_{An+B}(p)$$

for every $n \geq 1$.

Proof. Let $I = [-R, R]$, where $R \geq 1$ is fixed by the Archimedean certificate of Lemma 1. Since I is compact and g is a polynomial, there exists $C > 0$ such that

$$g(x) \leq C \quad \text{for every } x \in I. \tag{11}$$

We apply Lemma 5 with $t = g(x)/C$. Multiplying (8) by C gives

$$g + C\varepsilon_n = Cq_{n,d}(g/C) + C^{1-d}g^dW_n(g/C)^{d-1}. \tag{12}$$

By (11), $g(x)/C \leq 1$ for every $x \in I$. Then, we deduce by (7) that

$$Cq_{n,d}(g(x)/C) \geq 0 \quad \text{for every } x \in I.$$

Hence, by Lemma 2 (which we can apply since $R \geq 1$),

$$Cq_{n,d}(g/C) \in \mathcal{Q}(R^2 - x^2) \subset \mathcal{Q}(p) \quad \text{using again Lemma 1.}$$

For the second term, $d - 1$ is even, so we can write

$$C^{1-d}W_n(g/C)^{d-1} = (C^{(1-d)/2}W_n(g/C)^{(d-1)/2})^2.$$

Thus (10) and the closure of $\mathcal{Q}(p)$ under multiplication by squares imply that

$$C^{1-d}g^dW_n(g/C)^{d-1} \in \mathcal{Q}(p). \quad (13)$$

Equations (12)-(13) prove that $g + C\varepsilon_n \in \mathcal{Q}(p)$.

It remains to bound the degree of the resulting certificate in $\mathcal{Q}(p)$.

By (9), it holds that

$$\deg q_{n,d}(g/C) \leq (\deg g)(1 + n(d-1)).$$

Since $Cq_{n,d}(g/C)$ is nonnegative on $[-R, R]$ and $R \geq 1$, Lemma 2 yields a representation

$$Cq_{n,d}(g/C) \in \mathcal{Q}_{s_n}(R^2 - x^2), \quad s_n := \left\lceil \frac{(\deg g)(1 + n(d-1))}{2} \right\rceil.$$

Thus there exist $\sigma_{0,n}, \sigma_{1,n} \in \Sigma[x]$ such that

$$Cq_{n,d}(g/C) = \sigma_{0,n} + \sigma_{1,n}(R^2 - x^2),$$

with

$$\deg \sigma_{0,n} \leq 2s_n, \quad \deg \sigma_{1,n} \leq 2s_n - 2.$$

Fix once and for all an Archimedean certificate

$$R^2 - x^2 = \sum_{i=0}^{n_p} \tau_i p_i, \quad \tau_i \in \Sigma[x],$$

and choose $r_R \in \mathbb{N}$ such that

$$R^2 - x^2 \in \mathcal{Q}_{r_R}(p).$$

Equivalently,

$$\deg(\tau_i p_i) \leq 2r_R, \quad i = 0, \dots, n_p.$$

Substituting this fixed certificate into the preceding representation gives

$$Cq_{n,d}(g/C) = \sigma_{0,n} + \sum_{i=0}^{n_p} (\sigma_{1,n} \tau_i) p_i.$$

Since a product of sums of squares is again a sum of squares, this is a certificate in $\mathcal{Q}(p)$. Moreover,

$$\deg((\sigma_{1,n} \tau_i) p_i) \leq 2s_n - 2 + 2r_R \leq 2(s_n + r_R).$$

Consequently,

$$Cq_{n,d}(g/C) \in \mathcal{Q}_{s_n + r_R}(p).$$

Since

$$s_n = \left\lceil \frac{\deg g + \deg g(d-1)n}{2} \right\rceil \leq \left\lceil \frac{\deg g}{2} \right\rceil + n \left\lceil \frac{\deg g(d-1)}{2} \right\rceil,$$

the order of this first certificate is bounded by an affine function of n .

We now consider the second term in (12). Fix once and for all a certificate

$$g^d = \sum_{i=0}^{n_p} \lambda_i p_i, \quad \lambda_i \in \Sigma[x],$$

and choose $r_g \in \mathbb{N}$ such that

$$g^d \in \mathcal{Q}_{r_g}(p).$$

Because $d - 1$ is even, the polynomial

$$U_n(x) := C^{(1-d)/2} W_n(g(x)/C)^{(d-1)/2}$$

is well defined, and

$$C^{1-d} g^d W_n(g/C)^{d-1} = U_n^2 g^d.$$

Multiplying the fixed certificate of g^d by U_n^2 yields

$$C^{1-d} g^d W_n(g/C)^{d-1} = \sum_{i=0}^{n_p} (U_n^2 \lambda_i) p_i.$$

Since

$$\deg U_n^2 = D(n-1)(d-1),$$

we obtain

$$C^{1-d} g^d W_n(g/C)^{d-1} \in \mathcal{Q}_{r_g + \frac{D(d-1)}{2}(n-1)}(p).$$

Thus the order of the second certificate is also bounded by an affine function of n .

Combining the certificates of the two terms in (12), there exist integers $A \geq 1$, $B \geq 0$, depending only on g , p , and d , such that $g + C\varepsilon_n \in \mathcal{Q}_{An+B}(p)$ for every $n \geq 1$. $\square$

We can now prove Theorem 1.

Proof of Theorem 1. Define the objective gap

$$q(x) := f(x) - f^*. \tag{14}$$

Then, it holds that $q \geq 0$ on K , and by Lemma 3 there exist fixed $\sigma_i \in \Sigma[x]$ such that $q = \sum_{i=0}^m \sigma_i g_i$ where the g_i 's for $i \geq 1$ are the natural generators associated to K for $1 \leq i \leq m$ and $g_0 := 1$.

For every $i = 1, \dots, m$, Lemma 6 establishes that there exists an odd integer d_i such that $g_i^{d_i} \in \mathcal{Q}(p)$. Applying Lemma 7 to each g_i , we deduce that there exist fixed constants $A_i > 0$, $B_i \geq 0$, $C_i > 0$ such that

$$g_i + C_i \varepsilon_n \in \mathcal{Q}_{A_i n + B_i}(p). \tag{15}$$

Now, observe that

$$\sum_{i=0}^m \sigma_i (g_i + C_i \varepsilon_n) = q + \sigma \varepsilon_n$$

where the polynomial $\sigma := \sum_{i=0}^m C_i \sigma_i \in \Sigma[x]$ is fixed (since all C_i 's and σ_i 's are fixed). Also, since all σ_i are fixed, the certificate degrees in (15) imply that there exist constants $A > 0$ and $B' \geq 0$ such that

$$q + \varepsilon_n \sigma \in \mathcal{Q}_{An+B'}(p). \quad (16)$$

Let $C' := \max_{x \in [-R, R]} \sigma(x)$. Since σ is a sum of SOS polynomials with positive coefficients, it holds that $C' \geq 0$. Moreover $C' - \sigma(x) \geq 0$ for every $x \in [-R, R]$ by construction. By Lemmas 1 and 2,

$$C' - \sigma \in \mathcal{Q}(p). \quad (17)$$

This is a fixed polynomial with a fixed certificate. Multiplying (17) by the positive scalar ε_n and adding the result to (16) yields

$$q + C' \varepsilon_n \in \mathcal{Q}_{An+B}(p) \quad (18)$$

for some $B \geq B'$.

Now let r be sufficiently large and choose $n(r) := \lfloor \frac{r-B}{A} \rfloor$. Then $n(r) \geq 1$ and $An(r) + B \leq r$. Hence (18) implies $q + C' \varepsilon_{n(r)} \in \mathcal{Q}_r(p)$. Recalling (14), we obtain $f - (f^* - C' \varepsilon_{n(r)}) \in \mathcal{Q}_r(p)$. By the definition (2) of f_r as a supremum, $f_r \geq f^* - C' \varepsilon_{n(r)}$. Therefore

$$0 \leq f^* - f_r \leq C' \varepsilon_{n(r)} \leq \frac{C'}{n(r)^2}, \quad (19)$$

where the last inequality follows from Lemma 4.

If $r \geq 2(A + B)$, then

$$n(r) \geq \frac{r-B}{A} - 1 = \frac{r-A-B}{A} \geq \frac{r}{2A}.$$

Thus (19) gives

$$0 \leq f^* - f_r \leq \frac{4C'A^2}{r^2}.$$

Letting

$$C := 4C'A^2$$

and increasing r if necessary proves (5). $\square$

4 Basic Example with Worst Convergence Rate

Consider

$$f^* := \min_{x \in \mathbb{R}} f(x) \text{ s.t. } p(x) \geq 0, \text{ with } f(x) := 1-x, \quad p(x) := (1-x)^3(1+x). \quad (\text{POP})$$

The feasible set is $K = [-1, 1]$ and $f^* = 0$, attained at $x = 1$, where the constraint is cubically degenerate: p has a zero of multiplicity three at the minimizer.

For $r \geq 2$, the order- r relaxation error of (POP) is

$$e_r := f^* - \sup\{v \in \mathbb{R} : 1 - x - v \in \mathcal{Q}_r(p)\}. \quad (\text{SOS}_r)$$

Membership in the truncated quadratic module $\mathcal{Q}_r(p)$ certifies nonnegativity on K , so $e_r \geq f^* = 0$; and the value 0 is never attained: if $1 - x = \sigma_0 + p\sigma_1$ then $\sigma_0(1) = 0$, so σ_0 vanishes with even multiplicity at least two while $p\sigma_1$ vanishes with odd multiplicity at least three whereas $1 - x$ has a simple zero there. Hence $e_r > 0$ at every order r .

Theorem 2. *For every $r \geq 2$, it holds that*

$$e_r = \frac{1}{2r(r-1)}.$$

The proof occupies the rest of this section: one identity gives the lower bound, three short lemmas and one Cauchy-Schwarz inequality give the matching upper bound.

Let T_r resp. U_r be the Chebyshev polynomial of first resp. second kind, i.e. $T_r(\cos \theta) = \cos r\theta$ and $U_r(\cos \theta) \sin \theta = \sin(r+1)\theta$.

For a polynomial $f \in \mathbb{R}[x]$, f' denotes its first order derivative.

Recall that both T_r and U_r have degree r and that

$$T'_r = rU_{r-1}, \quad T_r = U_r - xU_{r-1} \quad \text{and} \quad U_r = 2xU_{r-1} - U_{r-2}.$$

Throughout, fix $r \geq 2$ and set

$$A_r(x) := rT_{r-1}(x) - (r-1)T_r(x), \quad B_r(x) := \frac{rU_{r-2}(x) - (r-1)U_{r-1}(x)}{1-x}.$$

Lemma 8. *It holds that A_r and B_r are polynomials with integer coefficients with $\deg A_r = r$, $\deg B_r = r-2$ and*

$$A_r(x)^2 + p(x)B_r(x)^2 = C_r(x) := 1 + 2r(r-1)(1-x). \quad (20)$$

Consequently $1 - x + \frac{1}{2r(r-1)} \in \mathcal{Q}_r(p)$ and $e_r \leq \frac{1}{2r(r-1)}$.

Proof. Let $x := \cos \theta$ and $z := (r - (r-1)e^{i\theta})e^{i(r-1)\theta}$. Then

$$|z|^2 = r^2 + (r-1)^2 - 2r(r-1)\cos \theta = 1 + 2r(r-1)(1-x) = C_r(x).$$

The real and imaginary parts of z are

$$\operatorname{Re} z = r \cos(r-1)\theta - (r-1) \cos r\theta = A_r(x)$$

and

$$\operatorname{Im} z = r \sin(r-1)\theta - (r-1) \sin r\theta = \sin \theta [rU_{r-2}(x) - (r-1)U_{r-1}(x)].$$

The bracket vanishes at $x = 1$, because $rU_{r-2}(1) - (r-1)U_{r-1}(1) = r(r-1) - (r-1)r = 0$; dividing by the integer linear factor $1-x$ gives the integer coefficient polynomial $B_r(x)$ of degree $r-2$, and

$$(\operatorname{Im} z)^2 = (1-x^2)(1-x)^2 B_r^2(x) = p(x) B_r^2(x).$$

Then (20) writes

$$(\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 = |z|^2$$

on $[-1, 1]$, hence as polynomials. Since $\deg A_r = r$ (its leading coefficient is $-(r-1)2^{r-1}$), dividing (20) by $2r(r-1)$ exhibits $1-x + \frac{1}{2r(r-1)}$ as an element of $\mathcal{Q}_r(p)$ with admissible degrees.

The fact that $\deg A_r = r$ and $\deg B_r = r-2$ are immediate from their definitions and $\deg T_r = r$ and $\deg U_r = r$. □

The table below displays the polynomials of identity (20) for small values of r .

r	$A_r(x)$	$B_r(x)$	$C_r(x)$
2	$-2x^2 + 2x + 1$	2	$5 - 4x$
3	$-8x^3 + 6x^2 + 6x - 3$	$8x + 2$	$13 - 12x$
4	$-24x^4 + 16x^3 + 24x^2 - 12x - 3$	$24x^2 + 8x - 4$	$25 - 24x$
5	$-64x^5 + 40x^4 + 80x^3 - 40x^2 - 20x + 5$	$64x^3 + 24x^2 - 24x - 4$	$41 - 40x$

Lemma 9. *It holds that A_r has r simple real roots distributed as follows:*

$$-1 < x_1 < \cdots < x_{r-1} < 1 < x_r.$$

Proof. Let $z_k := \cos \frac{k\pi}{r-1}$ for $k = 0, \dots, r-1$, so that $1 = z_0 > \cdots > z_{r-1} = -1$. Then $T_{r-1}(z_k) = (-1)^k$ and $T_r(z_k) = \cos(k\pi + \frac{k\pi}{r-1}) = (-1)^k z_k$, so

$$A_r(z_k) = (-1)^k \underbrace{(r - (r-1)z_k)}_{\geq 1}.$$

The signs alternate, giving at least $r-1$ roots in $(-1, 1)$. Since $\deg A_r = r$ with negative leading coefficient and $A_r(1) = 1 > 0$, there is a further root in $(1, \infty)$. This accounts for $r = \deg A_r$ roots, all simple. □

Lemma 10. *With $c := r(r-1)(2r-1)$,*

$$C_r(x) A'_r(x) + r(r-1) A_r(x) = c(1-x)^2 B_r(x). \quad (21)$$

Proof. From $T'_r = rU_{r-1}$, we deduce that

$$A'_r = r(r-1)(U_{r-2} - U_{r-1}).$$

So, after dividing by $r(r-1)$ and using $(1-x)B_r = rU_{r-2} - (r-1)U_{r-1}$, (21) is equivalent, to

$$C_r(x) (U_{r-2}(x) - U_{r-1}(x)) + A_r(x) = (2r-1)(1-x)[rU_{r-2}(x) - (r-1)U_{r-1}(x)].$$

By $T_r = U_r - xU_{r-1}$ and $U_r = 2xU_{r-1} - U_{r-2}$, it holds that

$$\begin{aligned} A_r(x) &= r(U_{r-1}(x) - xU_{r-2}(x)) - (r-1)(U_r(x) - xU_{r-1}(x)) \\ &= (r - (r-1)x)U_{r-1}(x) + (r-1-rx)U_{r-2}(x), \end{aligned}$$

so, in the left-hand side, the coefficient of U_{r-2} is

$$C_r(x) + r - 1 - rx = r(1-x) + 2r(r-1)(1-x) = r(2r-1)(1-x)$$

and the coefficient of U_{r-1} is

$$-C_r(x) + r - (r-1)x = (r-1)(1-x) - 2r(r-1)(1-x) = -(r-1)(2r-1)(1-x).$$

This is the right-hand side. $\square$

Lemma 11. *At every root x_i of A_r it holds*

$$p(x_i) B_r(x_i) A'_r(x_i) = c(1-x_i)^2 > 0.$$

Proof. Differentiating (20) gives $2A_r A'_r + p' B_r^2 + 2p B_r B'_r = -2r(r-1)$. At a common root of A_r and B_r the left side would vanish, a contradiction; so $B_r(x_i) \neq 0$, and by (20) and Lemma 9, $C_r(x_i) = p(x_i) B_r(x_i)^2 \neq 0$. We apply Lemma 10 by evaluating (21) at x_i and multiplying by $p(x_i) B_r(x_i) / C_r(x_i)$, we obtain

$$p(x_i) B_r(x_i) A'_r(x_i) = c(1-x_i)^2 \frac{p(x_i) B_r(x_i)^2}{C_r(x_i)} = c(1-x_i)^2,$$

which is positive because $x_i \neq 1$ (see Lemma 9). $\square$

Proof of Theorem 2. By Lemma 11, the positive numbers

$$\widehat{w}_i := \frac{1}{c(1-x_i)^2}, \quad i = 1, \dots, r,$$

satisfy

$$\widehat{w}_i p(x_i) B_r(x_i) = \frac{1}{A'_r(x_i)}.$$

Indeed, $x_i \neq 1$ by Lemma 9 and $c > 0$, so each $\widehat{w}_i$ is well defined and positive; and Lemma 11 asserts $p(x_i) B_r(x_i) A'_r(x_i) = c(1-x_i)^2 \neq 0$, so that the three factors $p(x_i)$, $B_r(x_i)$, $A'_r(x_i)$ are all non-zero and dividing that identity by $c(1-x_i)^2 A'_r(x_i)$ gives the claim. Let

$$w_i := \frac{\widehat{w}_i}{\sum_{j=1}^r \widehat{w}_j}$$

and define the linear functional

$$\ell(q) := \sum_{i=1}^r w_i q(x_i), \quad q \in \mathbb{R}[x]_{2r}.$$

Since the weights w_i are positive and sum to one, we have $\ell(1) = 1$ and

$$\ell(q^2) = \sum_{i=1}^r w_i q(x_i)^2 \geq 0$$

for all $q \in \mathbb{R}[x]$ such that $\deg q \leq r$.

We next prove that $\ell(pq^2) \geq 0$ for every $q \in \mathbb{R}_{r-2}[x]$ (recall that p has degree 4). Let $q \in \mathbb{R}[x]$ satisfy $\deg q \leq r - 2$. The Lagrange interpolation formula at the r simple roots $x_1, \dots, x_r$ of A_r reads

$$q(x) = \sum_{i=1}^r q(x_i) \frac{A_r(x)}{A'_r(x_i)(x - x_i)}.$$

Since the roots of A_r are simple, each quotient $A_r(x)/(x - x_i)$ is a polynomial of degree $r - 1$ whose leading coefficient is that of A_r , namely $\alpha_r := -(r - 1)2^{r-1} \neq 0$. The coefficient of x^{r-1} in the right-hand side is therefore $\alpha_r \sum_{i=1}^r q(x_i)/A'_r(x_i)$, whereas it vanishes in the left-hand side because $\deg q \leq r - 2$. Dividing by α_r we obtain

$$\sum_{i=1}^r \frac{q(x_i)}{A'_r(x_i)} = 0.$$

Let $S := \sum_{j=1}^r \widehat{w}_j$, a positive real number, so that $w_i = \widehat{w}_i/S$ for every i . Dividing by S the identity $\widehat{w}_i p(x_i) B_r(x_i) = 1/A'_r(x_i)$ established above yields

$$w_i p(x_i) B_r(x_i) = \frac{1}{S A'_r(x_i)}, \quad i = 1, \dots, r,$$

so that the weighted sum against B_r is, up to the factor $1/S$, the sum of the Lagrange coefficients just computed:

$$\sum_{i=1}^r w_i p(x_i) B_r(x_i) q(x_i) = \frac{1}{S} \sum_{i=1}^r \frac{q(x_i)}{A'_r(x_i)} = 0 \quad (22)$$

for every $q \in \mathbb{R}[x]$ with $\deg q \leq r - 2$. By Lemma 9, $-1 < x_i < 1$ for $1 \leq i \leq r - 1$ and $x_r > 1$. Then, it holds that $p(x_i) > 0$, $i = 1, \dots, r - 1$, whereas $p(x_r) < 0$. Set

$$\lambda := -w_r p(x_r) > 0. \quad (23)$$

We introduce the weighted scalar product on the values at the interior nodes $x_1, \dots, x_{r-1}$:

$$\langle u, v \rangle_{\text{int}} := \sum_{i=1}^{r-1} w_i p(x_i) u(x_i) v(x_i).$$

All the weights $w_i p(x_i)$, $i = 1, \dots, r - 1$, are positive. Isolating the exterior node x_r in (22) gives

$$\langle B_r, q \rangle_{\text{int}} = \lambda B_r(x_r) q(x_r). \quad (24)$$

Since $\deg B_r = r - 2$, we may choose $q = B_r$ in (24). We obtain

$$\langle B_r, B_r \rangle_{\text{int}} = \lambda B_r(x_r)^2. \quad (25)$$

Let now $q \in \mathbb{R}[x]_{r-2}$ be arbitrary. The Cauchy-Schwarz inequality for the weighted scalar product gives

$$\langle B_r, q \rangle_{\text{int}}^2 \leq \langle B_r, B_r \rangle_{\text{int}} \langle q, q \rangle_{\text{int}}.$$

Using (24) and (25), this becomes

$$\lambda^2 B_r(x_r)^2 q(x_r)^2 \leq \lambda B_r(x_r)^2 \sum_{i=1}^{r-1} w_i p(x_i) q(x_i)^2.$$

By Lemma 11, $B_r(x_r) \neq 0$. Since also $\lambda > 0$, division by $\lambda B_r(x_r)^2$ yields

$$\lambda q(x_r)^2 \leq \sum_{i=1}^{r-1} w_i p(x_i) q(x_i)^2.$$

Recalling (23), we conclude that

$$\ell(pq^2) = \sum_{i=1}^r w_i p(x_i) q(x_i)^2 = \sum_{i=1}^{r-1} w_i p(x_i) q(x_i)^2 - \lambda q(x_r)^2 \geq 0.$$

In particular, taking $q = B_r$ in (25) gives

$$\ell(pB_r^2) = \sum_{i=1}^{r-1} w_i p(x_i) B_r(x_i)^2 - \lambda B_r(x_r)^2 = 0. \quad (26)$$

Applying the linear functional ℓ to the identity (20) yields

$$\ell(A_r^2) + \ell(pB_r^2) = \ell(C_r) = 1 + 2r(r-1)\ell(1-x) = 0$$

since $\ell(A_r^2) = \sum_{i=1}^r w_i A_r(x_i)^2 = 0$ because x_i are the roots of A_r , and $\ell(pB_r^2) = 0$ by (26). Therefore

$$\ell(1-x) = -\frac{1}{2r(r-1)}. \quad (27)$$

We have thus shown that $\ell(q^2) \geq 0$ for every $q \in \mathbb{R}[x]_r$, and $\ell(pq^2) \geq 0$ for every $q \in \mathbb{R}[x]_{r-2}$, i.e. ℓ is nonnegative on the truncated quadratic module $\mathcal{Q}_r(p)$ and normalized by $\ell(1) = 1$. Let $v \in \mathbb{R}$ be any feasible value for the SOS problem (SOS_r). Since ℓ is nonnegative on $\mathcal{Q}_r(p)$, it follows that $0 \leq \ell(1-x-v) = \ell(1-x) - v\ell(1) = \ell(1-x) - v$ and hence $v \leq \ell(1-x)$ for every feasible value v . Taking the supremum over all feasible v and recalling (27) yields

$$e_r \geq -\ell(1-x) = \frac{1}{2r(r-1)}.$$

On the other hand, Lemma 8 gives the reverse inequality, and consequently

$$e_r = \frac{1}{2r(r-1)}.$$

□

Remark 1 (Comparison with the Stengle problem). *The classical symmetric Stengle problem is $\min_{x \in \mathbb{R}} \{1 - x^2 : (1 - x^2)^3 \geq 0\}$. Its feasible set is again $[-1, 1]$ and in [13] it is shown that its exact order- r moment-SOS error is $\frac{1}{r(r-2)}$ for $r \geq 3$. The structural reason for $e_r > 0$ at every order is that $\mathcal{Q}((1 - x^2)^3)$ is not stable. Stengle's original estimate, recorded as [27, Ex. 6.6.6(3)], states that there is $c > 0$ with $\deg \sigma_0 > c \varepsilon^{-1/2}$ in every identity $1 - x^2 + \varepsilon = \sigma_0 + \sigma_1(1 - x^2)^3$, $\sigma_0, \sigma_1 \in \Sigma[x]$; the exact value $\frac{1}{r(r-2)}$ pins down the constant in this $\varepsilon^{-1/2}$ law. The proof of Theorem 2 follows the same broad primal-dual architecture as the exact analysis of [13], but the asymmetric constraint and the linear objective function makes the analysis simpler.*

5 Univariate and Bivariate Cusp Problems

In this section we show that the extremal example (POP) is, up to an affine change of variables, equivalent to a univariate problem with split constraints, which is in turn an exact univariate reduction of a bivariate cusp problem studied in the literature [20, Ex. 9.4.6(3)], [4, Ex. A.3], [15]. As a consequence, the exact $\Theta(1/r^2)$ convergence rate of Theorem 2 transfers to all these formulations.

5.1 Affine change of variables

The affine substitution $x \mapsto 1 - 2x$ maps the interval $[0, 1]$ onto $[-1, 1]$ and transforms (POP) into

$$\min_{x \in \mathbb{R}} x \quad \text{s.t.} \quad x^3(1 - x) \geq 0, \quad (\text{POP}')$$

whose feasible set is $[0, 1]$ and whose value is 0, attained at $x = 0$, where the constraint has a zero of multiplicity three. Indeed, substituting $x \mapsto 1 - 2x$ in (POP) gives objective $2x$ and constraint $(2x)^3(2 - 2x) = 16x^3(1 - x)$, and neither the positive scaling of the constraint nor the positive scaling of the objective affects membership in the truncated quadratic module at a given order. Consequently, denoting by e'_r the order- r relaxation error of (POP'), Theorem 2 yields

$$e'_r = \frac{e_r}{2} = \frac{1}{4r(r-1)}, \quad r \geq 2.$$

Remark 2. *The constraint polynomial $x^3(1 - x)$ of (POP') is the subject of [27, Exerc. 5.5.5, p. 200], where the reader is asked to prove that in any sequence of representations $x + \frac{1}{n} = \sigma_{0,n} + \sigma_{1,n}x^3(1 - x)$ with $\sigma_{0,n}, \sigma_{1,n} \in \Sigma[x]$ one necessarily has $\deg \sigma_{0,n}, \deg \sigma_{1,n} \rightarrow \infty$. The value $e'_r = \frac{1}{4r(r-1)}$ makes this quantitative: such a representation exists within $\mathcal{Q}_r(x^3(1 - x))$ if and only if $4r(r-1) \geq n$, so the smallest admissible order is $r_n = \lceil \frac{1+\sqrt{1+n}}{2} \rceil$ and the certificate degrees grow exactly like $\Theta(\sqrt{n})$. The same exercise, part (a), shows that $\mathcal{Q}_r(x^3(1 - x))$ is a closed cone, so the supremum in (SOS_r) is attained; here this is also witnessed explicitly by Lemma 8.*

5.2 Splitting the constraint

The constraint polynomial of (POP') is the product of the two polynomials x^3 and $1 - x$. Splitting it accordingly yields the univariate problem

$$\min_{x \in \mathbb{R}} x \quad \text{s.t.} \quad x^3 \geq 0, \quad 1 - x \geq 0, \quad (\text{POP}'')$$

with the same feasible set $[0, 1]$ and the same value 0. Although the truncated quadratic modules of the two descriptions differ at each fixed order, the underlying quadratic modules coincide, and the truncations agree up to a unit shift of the order.

Lemma 12 (Product versus split generators). *It holds that*

$$\mathcal{Q}(x^3(1 - x)) = \mathcal{Q}(x^3, 1 - x)$$

and, for every $r \geq 2$,

$$\mathcal{Q}_r(x^3(1 - x)) \subseteq \mathcal{Q}_{r+1}(x^3, 1 - x), \quad \mathcal{Q}_r(x^3, 1 - x) \subseteq \mathcal{Q}_{r+1}(x^3(1 - x)).$$

Proof. The three polynomial identities

$$\begin{aligned} x^3 &= x^4 + x^3(1 - x), \\ 1 - x &= (1 - x)^2 \left((x + \frac{1}{2})^2 + \frac{3}{4} \right) + x^3(1 - x), \\ x^3(1 - x) &= x^4(1 - x) + (1 - x)^2 x^3, \end{aligned} \quad (28)$$

are verified by expansion. The first two show that both generators x^3 and $1 - x$ belong to $\mathcal{Q}(x^3(1 - x))$, hence $\mathcal{Q}(x^3, 1 - x) \subseteq \mathcal{Q}(x^3(1 - x))$; the third shows that the product generator $x^3(1 - x)$ belongs to $\mathcal{Q}(x^3, 1 - x)$, hence the reverse inclusion. This proves the equality of the quadratic modules.

For the truncated statements, consider an element $\sigma_0 + \sigma_1 \cdot x^3(1 - x) \in \mathcal{Q}_r(x^3(1 - x))$, so that

$$\deg \sigma_0 \leq 2r \quad \text{and} \quad \deg \sigma_1 \leq 2r - 4.$$

The third identity of (28) gives

$$\sigma_1 \cdot x^3(1 - x) = (\sigma_1 \cdot (1 - x)^2)x^3 + (\sigma_1 \cdot x^4)(1 - x),$$

where $\sigma_1 \cdot (1 - x)^2$ and $\sigma_1 \cdot x^4$ are SOS of degrees at most $2r - 2$ and $2r$, so that

$$\deg(\sigma_1 \cdot (1 - x)^2 x^3) \leq 2r + 1 \quad \text{and} \quad \deg(\sigma_1 \cdot x^4(1 - x)) \leq 2r + 1.$$

Hence $\sigma_0 + \sigma_1 \cdot x^3(1 - x) \in \mathcal{Q}_{r+1}(x^3, 1 - x)$.

Conversely, consider $\sigma_0 + \sigma_1 \cdot x^3 + \sigma_2 \cdot (1 - x) \in \mathcal{Q}_r(x^3, 1 - x)$, with $\deg \sigma_1 \leq 2r - 4$ and $\deg \sigma_2 \leq 2r - 2$. By the first two identities of (28),

$$\sigma_1 \cdot x^3 + \sigma_2 \cdot (1 - x) = \sigma_1 \cdot x^4 + \sigma_2 \cdot (1 - x)^2 \left((x + \frac{1}{2})^2 + \frac{3}{4} \right) + (\sigma_1 + \sigma_2) x^3(1 - x),$$

where the first two terms are SOS of degrees at most $2r$ and $2r + 2$, and the multiplier $\sigma_1 + \sigma_2$ is SOS of degree at most $2r - 2$, so that $\deg((\sigma_1 + \sigma_2)x^3(1 - x)) \leq 2r + 2$. Hence the element belongs to $\mathcal{Q}_{r+1}(x^3(1 - x))$. $\square$

Denoting by e_r'' the order- r relaxation error of (POP''), Lemma 12 immediately yields a two-sided estimate.

Corollary 1 (Convergence rate of the split formulation). *For every $r \geq 3$, it holds*

$$\frac{1}{4r(r+1)} = e'_{r+1} \leq e_r'' \leq e'_{r-1} = \frac{1}{4(r-1)(r-2)}.$$

In particular

$$\lim_{r \rightarrow \infty} r^2 e_r'' = \lim_{r \rightarrow \infty} r^2 e'_r = \frac{1}{4},$$

so the moment-SOS hierarchies of (POP') and (POP'') share the same exact asymptotic convergence rate $\Theta(1/r^2)$.

Proof. By Lemma 12, every certificate for (POP'') at order r is a certificate for (POP') at order $r+1$, so $e'_{r+1} \leq e_r''$, and symmetrically $e_r'' \leq e'_{r-1}$. The limits follow from $e'_r = \frac{1}{4r(r-1)}$. $\square$

Thus splitting the quartic constraint of (POP') into a cubic and a linear constraint changes the relaxation values only marginally: the cost of separating the generators is at most one relaxation order.

5.3 Reduction of the bivariate cusp problem

Consider now the bivariate cusp problem

$$\min_{(x_1, x_2) \in \mathbb{R}^2} x_1 \quad \text{s.t.} \quad x_1^3 - x_2^2 \geq 0, \quad 1 - x_1 \geq 0, \quad (\text{POP}^{\text{cusp}})$$

whose feasible set K is delimited by the cusp arcs $x_2 = \pm x_1^{3/2}$ and the vertical segment $x_1 = 1$. This problem appears in [20, Ex. 9.4.6(3)] and [4, Ex. A.3] as an elementary instance for which the Łojasiewicz exponent of the constraint description equals 3, and it is used as a benchmark for high precision semidefinite solvers in [15]. Its value is 0, attained at the cusp point $(0, 0)$, where neither constraint qualification nor strict complementarity holds.

The univariate problem (POP'') is precisely the restriction of (POP^{cusp}) to the axis $x_2 = 0$, with $x := x_1$. The following lemma shows that this restriction is exact order by order for the SOS hierarchy.

Lemma 13 (Exact univariate reduction). *For every $r \geq 2$, the order- r SOS relaxation values of (POP^{cusp}) and (POP'') coincide, i.e. denoting by e_r^{cusp} the order- r relaxation error of (POP^{cusp}), it holds*

$$e_r^{\text{cusp}} = e_r''.$$

Proof. If $x_1 - v = \sigma_0 + \sigma_1 \cdot (x_1^3 - x_2^2) + \sigma_2 \cdot (1 - x_1)$ is a certificate of order r for (POP^{cusp}), with $\sigma_i \in \Sigma[x_1, x_2]$ of admissible degrees, then restricting to $x_2 = 0$ gives

$$x_1 - v = \sigma_0(x_1, 0) + \sigma_1(x_1, 0)x_1^3 + \sigma_2(x_1, 0)(1 - x_1),$$

where the restricted multipliers are univariate SOS of the same degrees, i.e. a certificate of order r for (POP'').

Conversely, if $x - v = \sigma_0 + \sigma_1 \cdot x^3 + \sigma_2 \cdot (1 - x)$ is a certificate of order r for (POP''), then substituting $x := x_1$ and using the identity $x_1^3 = (x_1^3 - x_2^2) + x_2^2$ gives

$$x_1 - v = (\sigma_0(x_1) + x_2^2 \sigma_1(x_1)) + \sigma_1(x_1)(x_1^3 - x_2^2) + \sigma_2(x_1)(1 - x_1),$$

where $\sigma_0(x_1) + x_2^2 \sigma_1(x_1)$ is a bivariate SOS polynomial of degree at most

$$\max\{\deg \sigma_0, \deg \sigma_1 + 2\} \leq 2r,$$

since $\deg \sigma_1 \leq 2r - 4$. This is a certificate of order r for (POP^{cusp}). The two relaxation values therefore coincide at every order $r \geq 2$. $\square$

Combining Corollary 1 and Lemma 13, the bivariate cusp problem (POP^{cusp}) inherits the exact asymptotic convergence rate

$$\lim_{r \rightarrow \infty} r^2 e_r^{\text{cusp}} = \frac{1}{4}$$

of the extremal example (POP). This is consistent with the general Łojasiewicz-based bounds of [3, 5]: the cusp description has Łojasiewicz exponent $L = 3$, yet the actual convergence rate does not deteriorate below the universal univariate exponent 2 of Theorem 1. Figure 2 displays the outer approximations of K computed by the moment relaxations of orders $r = 2, 3, 4$, illustrating how the relaxations overshoot the cusp point along the negative x_1 axis by the amount e_r^{cusp} .

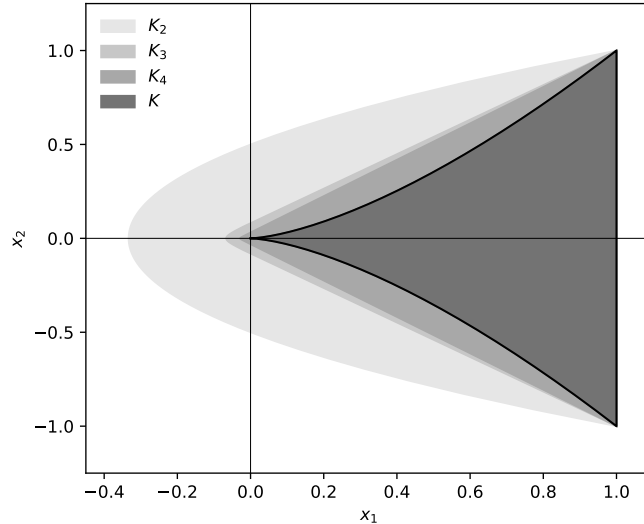


Figure 2: Outer approximations of the cusp feasible set K (dark gray) computed by the moment relaxations of orders $r = 2, 3, 4$ (nested regions, from light to darker). The minimum of x_1 over the order- r relaxation equals $-e_r^{\text{cusp}}$, i.e. $-\frac{1}{3}$, $-\frac{1}{15}$ and $-\frac{1}{35}$ for $r = 2, 3, 4$ respectively.

Remark 3 (Exact relaxation values). *The two-sided estimate of Corollary 1 can be refined to an exact value. The substitution $x = 1 - t^2$ maps the univariate cusp problem (POP'') to the symmetric Stengle problem $\min_t \{1 - t^2 : (1 - t^2)^3 \geq 0\}$ recalled in Remark 1, and a parity argument on the extremal Chebyshev-Gegenbauer certificates of [13] shows that the order- r relaxation of (POP'') matches the order- $(2r - 1)$ relaxation of the Stengle problem. Together with the exact Stengle value $\frac{1}{s(s-2)}$ at order s from [13], this yields*

$$e_r^{\text{cusp}} = e_r'' = \frac{1}{(2r - 1)(2r - 3)}, \quad r \geq 2,$$

which indeed satisfies the bounds of Corollary 1.

6 Conclusion

We have established a universal quadratic upper bound on the convergence rate of the moment-SOS hierarchy for polynomial optimization on a bounded subset of the real line. More precisely, for every fixed univariate POP of the form (1), there exists a constant $C > 0$ such that $0 \leq e_r \leq \frac{C}{r^2}$ for all sufficiently large relaxation orders r , where e_r is the error at the relaxation order r . We have also shown that this quadratic rate is optimal, and achieved by the elementary degree-four example (POP) for which $e_r = 1/(2r(r - 1))$, $r \geq 2$.

The main difficulty is that the natural generators of the feasible set need not belong to the quadratic module generated by the original constraints. We overcome this difficulty by combining the univariate structure theorem for quadratic modules with a Chebyshev polynomial construction. Odd powers of the natural generators belong to the original quadratic module, and the Chebyshev construction approximately recovers the generators themselves with an error of order $1/r^2$ and with certificate degree growing linearly in r .

A natural open question is whether the quadratic exponent remains universal for bivariate polynomial optimization. At present, we do not know any bivariate POP for which the convergence of the moment-SOS hierarchy is slower than $O(1/r^2)$. It would therefore be desirable to prove that no such example exists, namely that, under the usual Archimedean assumptions and for fixed problem data, every bivariate POP satisfies $e_r = O(1/r^2)$. Such a result would show that the deterioration predicted by general dimension-dependent and Łojasiewicz-based estimates is not intrinsic in dimension two, and would constitute a first step toward identifying the precise role of dimension in the worst-case convergence rate of the moment-SOS hierarchy.

Acknowledgement

The mathematical writing was done with the help of ChatGPT 5.6 and Claude Fable 5. This work benefited from feedback by Victor Magron. The first author is

grateful to Juan Vera Lizcano for figuring out, during the Oberwolfach workshop on conic linear optimization for computer-assisted proofs [8], the reformulation of the univariate cusp problem as a Stengle problem.

References

- [1] D. Augustin. The Membership Problem for Quadratic Modules with Focus on the One Dimensional Case, Ph.D. thesis, Univ. Regensburg, 2008.
- [2] D. Augustin. The membership problem for finitely generated quadratic modules in the univariate case, *J. Pure Appl. Alg.* 216(10):2204–2212, 2012.
- [3] L. Baldi, B. Mourrain. On the effective Putinar’s Positivstellensatz and moment approximation, *Math. Prog.* 200:71–103, 2023.
- [4] L. Baldi, B. Mourrain. Exact moment representation in polynomial optimization, *J. Symb. Comp.* 129:102403, 2025.
- [5] L. Baldi, B. Mourrain, A. Parusiński. On Łojasiewicz inequalities and the effective Putinar’s Positivstellensatz, *J. Alg.* 662, 2025.
- [6] L. Baldi, L. Slot. Degree bounds for Putinar’s Positivstellensatz on the hypercube. *SIAM J. Appl. Alg. Geom.* 8(1):1–25, 2024.
- [7] A. T. Blomenhofer and M. Laurent. Moment-SOS and spectral hierarchies for polynomial optimization on the sphere and quantum de Finetti theorems. *SIAM J. Optim.* 36(1):204–232, 2026.
- [8] E. de Klerk, F. Vallentin, A. Wiegele (organizers). Conic linear optimization for computer-assisted proofs, Oberwolfach Report 21/2026, Mathematisches Forschungsinstitut Oberwolfach, 2026.
- [9] K. Fang, H. Fawzi. The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Math Prog.* 190:331–360, 2021.
- [10] S. Gribling, E. de Klerk, J. C. Vera Lizcano. Squared polynomial approximation kernels for the hypercube: improved error bounds and implications for Lasserre hierarchies. *arXiv:2605.31496*, 2026.
- [11] O. Heijmans-Kuryatnikova, J. C. Vera Lizcano, L. F. Zuluaga. Degree bounds for Positivstellensätze of general semialgebraic sets, *arXiv:2605.15821*, 2026.
- [12] D. Henrion. Moments for polynomial optimization – An illustrated tutorial. Lecture notes for the programme Recent Trends in Computer Algebra, Institut Henri Poincaré, Paris, hal-05665923, 2023.
- [13] D. Henrion. Solving Stengle’s example in rational arithmetic: exact values of the moment-SOS relaxations, *arXiv:2512.19141*, to appear in *Open J. Math. Optim.*, 2026.

- [14] D. Henrion, A. Le Franc, V. Magron. Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem. *SIAM J. Appl. Alg. Geom.* 9(1):261–278, 2025.
- [15] M. Kočvara. Conjectures in polynomial optimization through high precision semidefinite programming (joint work with L. Baldi), SPOT optimization seminar, Toulouse, October 2025.
- [16] S. Kuhlmann, M. Marshall. Positivity, sums of squares and the multi-dimensional moment problem, *Trans. AMS* 354(11):4285–4301, 2002.
- [17] S. Kuhlmann, M. Marshall, N. Schwartz. Positivity, sums of squares and the multi-dimensional moment problem II. *Adv. Geom.* 5(4):583–606, 2005.
- [18] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* 11(3):796–817, 2001.
- [19] M. Laurent, L. Slot. An overview of convergence rates for sum of squares hierarchies in polynomial optimization. In S. Oishi, H. Okamoto, K. Hayami (Editors). *Recent Developments in Industrial and Applied Mathematics: Conference Proceedings ICIAM2023*, ICIAM2023 Springer Series, 1:149–176, 2026.
- [20] M. Marshall. *Positive Polynomials and Sums of Squares*. AMS, 2008.
- [21] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. *Math. Prog. Ser. A* 146(1-2):97–121, 2014.
- [22] J. Nie. *Moment and Polynomial Optimization*. SIAM, 2023.
- [23] J. Nie, M. Schweighofer. On the complexity of Putinar’s Positivstellensatz. *J. Complexity* 23(1):135–150, 2007.
- [24] V. Powers and B. Reznick. Polynomials that are positive on an interval, *Trans. AMS* 352(10):4677–4692, 2000.
- [25] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* 42(3):969–984, 1993.
- [26] C. Scheiderer. Distinguished representations of non-negative polynomials, *J. Alg.* 289(2):558–573, 2005.
- [27] C. Scheiderer. *A Course in Real Algebraic Geometry: Positivity and Sums of Squares*. Springer, 2024.
- [28] C. Schlosser, M. Tacchi-Bénard, A. Lazarev. Convergence rates for the moment-SoS hierarchy. *Num. Alg. Contr. Opt.* 16:105-156, 2026.
- [29] L. Slot. Sum-of-squares hierarchies for polynomial optimization and the Christoffel–Darboux kernel, *SIAM J. Opt.* 32(4):2612–2635, 2022.
- [30] G. Stengle. Complexity estimates for the Schmüdgen Positivstellensatz, *J. Complexity* 12(2):167–174, 1996.

- [31] T. Theobald. Real Algebraic Geometry and Optimization. AMS, 2024.