

Solving Stengle’s Example in Rational Arithmetic: Exact Values of the Moment-SOS Relaxations

Didier Henrion*

December 21, 2025

Abstract

We revisit Stengle’s classical univariate polynomial optimization example $\min 1 - x^2$ s.t. $(1 - x^2)^3 \geq 0$ whose constraint description is degenerate at the minimizers. We prove that the moment-SOS hierarchy of relaxation order $r \geq 3$ has the exact value $-1/r(r - 2)$. For this we construct in rational arithmetic a dual polynomial sum-of-squares (SOS) certificate and a primal moment sequence representing a finitely atomic measure. The key ingredients are elementary trigonometric properties of Chebyshev and Gegenbauer polynomial, and a Christoffel-Darboux kernel argument.

1 Introduction

Given a polynomial optimization problem (POP), the moment-SOS hierarchy (aka the Lasserre hierarchy [6]) builds a sequence of semidefinite relaxations of increasing size indexed by a relaxation order r . On the primal side, one optimizes a linear functional over truncated moment sequences subject to semidefinite constraints on the moment matrix and localizing matrices. On the dual side, one searches for a sum-of-squares (SOS) decomposition of a shifted polynomial. See [4, 10, 20] for recent general overviews.

Recent work has developed *worst-case upper bounds* on the relaxation error: for broad classes of instances one can guarantee that the gap after r steps is at most $O(r^{-k})$, with the exponent k depending on the geometry of the domain and regularity assumptions; see [16, 7, 3, 15] and references therein. Such results are *sufficiency* statements: they provide a safety guarantee that the hierarchy cannot converge slower than a stated rate, even on the hardest instances in the class.

In contrast, *lower bounds* on the relaxation error show that certain upper bounds are essentially sharp: one constructs explicit instances where the hierarchy converges no faster than a given rate. These hardness results are *necessity* statements: they quantify intrinsic limitations of the method (or of the representation class), and rules out uniformly better guarantees without changing assumptions or algorithms. The earliest quantitative lower bound predates the modern formulation of the moment-SOS hierarchy and is due to Stengle [18]; see the discussions in [14, Chapter 17] and [10, Section 3.5.1]. Stengle provided a univariate example showing a lower bound

*LAAS-CNRS, University of Toulouse, France and Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia.

of $\Omega(1/r^2)$ and an upper bound of $O((\log r)^2/r^2)$. This example is characterized by a degenerate description of the domain. Recently, the paper [2] gives lower bounds for domains described by non-degenerate inequalities, connecting these with quantitative non-stability phenomena [12]. Lower bounds were also constructed for a specific parametric univariate POP in [8], but in this case the moment-SOS hierarchy has always finite (i.e. non-asymptotic) convergence, contrary to the Stengle example.

Contribution. The Stengle example appears as Example 1 in [5], where high-precision semidefinite programming experiments suggest a $\Theta(1/r^2)$ behavior. We solve Stengle's example analytically, in rational arithmetic, and prove that the relaxation error is *exactly* $-1/(r-2)$. The proof is based on convex duality and basic properties of orthogonal polynomials (Chebyshev and Gegenbauer polynomials, and the Christoffel-Darboux kernel) and elementary trigonometric identities, mirroring the role of orthogonal polynomials in upper-bound analyses [16, 7, 15].

2 Stengle's example and its moment-SOS relaxations

2.1 The POP and its degeneracy

Consider the univariate constrained polynomial optimization problem (POP)

$$\min_{x \in \mathbb{R}} f(x) := 1 - x^2 \quad \text{s.t.} \quad g(x) := (1 - x^2)^3 \geq 0. \quad (\text{POP})$$

Since $g(x) \geq 0$ holds if and only if $x \in [-1, 1]$, the problem reduces to

$$f^* := \min_{|x| \leq 1} (1 - x^2) = 0,$$

attained at the boundary points $x^* = \pm 1$.

The POP is ill-posed because the feasible set is described by a single inequality $g(x) \geq 0$ whose gradient vanishes at the minimizers:

$$g'(x) = -6x(1 - x^2)^2, \quad g'(\pm 1) = 0.$$

Thus the active constraint at $x^* = \pm 1$ is *degenerate*: the linearization carries no first-order information. The standard constraint qualifications of nonlinear programming are violated.

For illustration, let us consider a classical interior-point method consisting of minimizing, for a given barrier parameter $\mu > 0$, a log-barrier function

$$h_\mu(x) := f(x) - \mu \log g(x) = 1 - x^2 - 3\mu \log(1 - x^2), \quad x \in (-1, 1).$$

A stationary point satisfies

$$\nabla h_\mu(x) = -2x + \frac{6\mu x}{1 - x^2} = 0,$$

hence either $x = 0$ or $1 - x(\mu)^2 = 3\mu$ which yields

$$x(\mu) = \sqrt{1 - 3\mu} \xrightarrow{\mu \downarrow 0} 1.$$

The dual Lagrange multiplier induced by the log-barrier stationarity is

$$\nabla f(x(\mu)) - \lambda(\mu) \nabla g(x(\mu)) = 0 \quad \text{with} \quad \lambda(\mu) = \frac{\mu}{g(x(\mu))} = \frac{\mu}{(3\mu)^3} = \frac{1}{27\mu^2} \xrightarrow{\mu \downarrow 0} +\infty.$$

Moreover, the curvature of the barrier objective is

$$\nabla^2 h_\mu(x) = -2 + 6\mu \frac{1+x^2}{(1-x^2)^2},$$

and substituting $x(\mu)^2 = 1 - 3\mu$ gives

$$\nabla^2 h_\mu(x(\mu)) = -2 + 6\mu \frac{2-3\mu}{(3\mu)^2} = \frac{4}{3\mu} - 4 \xrightarrow{\mu \downarrow 0} +\infty.$$

Thus the formulation (POP) is a clean toy model for ill-conditioning of barrier methods: the primal iterates converge to the boundary while the dual multiplier and the barrier function Hessian blow up.

2.2 Moment and SOS relaxations

Fix an integer $r \geq 3$. Let $\mathbb{R}[x]_r$ denote the vector space of polynomials of the scalar indeterminate $x \in \mathbb{R}$ of degree at most r . In the monomial basis $b(x) := (1, x, \dots, x^r)^\top$ we identify a polynomial $p(x)$ with its coefficient vector $p \in \mathbb{R}^{r+1}$, i.e. $p(x) = p^\top b(x)$. A truncated moment sequence is a vector $y = (y_0, y_1, \dots, y_{2r})$. The associated Riesz functional ℓ_y acts linearly on $\mathbb{R}[x]_{2r}$ by $\ell_y(x^k) = y_k$ and extension by linearity. The moment matrix $M_r(y)$ of order r is the Hankel matrix indexed by monomials $b(x)$:

$$p^\top M_r(y) p = \ell_y(p(x)^2) \quad \forall p \in \mathbb{R}[x]_r. \quad (1)$$

The localizing matrix for g at order $r-3$ is

$$q^\top M_{r-3}(gy) q = \ell_y(g(x)q(x)^2) \quad \forall q \in \mathbb{R}[x]_{r-3}. \quad (2)$$

The order- r moment relaxation of (POP) is

$$\begin{aligned} & \inf_y \ell_y(f) \\ \text{s.t. } & \ell_y(1) = 1, \quad M_r(y) \succeq 0, \quad M_{r-3}(gy) \succeq 0 \end{aligned} \quad (\text{MOM}_r)$$

and the order- r SOS relaxation of (POP) is

$$\begin{aligned} & \sup_{\varepsilon, p, q} \varepsilon \\ \text{s.t. } & f(x) - \varepsilon = p(x) + g(x)q(x), \quad p \in \Sigma[x]_{2r}, \quad q \in \Sigma[x]_{2(r-3)} \end{aligned} \quad (\text{SOS}_r)$$

where the convex cone $\Sigma[x]_{2d} \subset \mathbb{R}[x]_{2d}$ consists of SOS polynomials of degree at most $2d$. The moment-SOS hierarchy for POP was introduced in this exact form in [6], see e.g. [4, 10, 20] for recent overviews.

Lemma 1 (Weak duality). *For every $r \geq 3$, any feasible y in (MOM_r) and any feasible (ε, p, q) in (SOS_r) satisfy $\ell_y(f) \geq \varepsilon$.*

Proof. If $f - \varepsilon = p + gq$ with p, q SOS, then for any feasible y ,

$$\ell_y(f) - \varepsilon = \ell_y(p) + \ell_y(gq) \geq 0,$$

because $\ell_y(p) \geq 0$ follows from $M_r(y) \succeq 0$ and $\ell_y(gq) \geq 0$ follows from $M_{r-3}(gy) \succeq 0$. \square

2.3 Strong duality and non-attainment at zero

The quadratic module generated by g is the set of polynomials $p + gq$ for p and q SOS. We say that a quadratic module is Archimedean when it contains $R^2 - x^2$ for some $R > 0$. Note that this implies compactness of the set described by the inequality $g(x) \geq 0$. The convergence of the moment-SOS hierarchy, as originally proved in [6], relies on Putinar's Positivstellensatz (Psatz) [11], which in turn relies on the Archimedean property of the quadratic module.

The quadratic module generated by $g(x) = (1 - x^2)^3$ is Archimedean since

$$2 - x^2 = \underbrace{(x^3 - 2x)^2 + (x^2 - 1)^2}_{p(x)} + \underbrace{(1 - x^2)^3}_{q(x)} \underbrace{(1)}_{(1)}. \quad (3)$$

Therefore Putinar's Psatz applies and the moment-SOS hierarchy converges, i.e.

$$\lim_{r \rightarrow \infty} \varepsilon_r^* = f^* = 0.$$

Note however that for our POP the quadratic module and the preordering are the same, since we have only one generator g . Convergence of the moment-SOS hierarchy then also follows from Schmüdgen's Psatz [13] which applies without resorting to the Archimedean property, because the feasibility domain $[-1, 1]$ is compact.

Lemma 2 (Strong duality and attainment). *For every $r \geq 3$, the primal (MOM_r) satisfies Slater's condition, implying strong duality and dual attainment:*

$$\varepsilon_r^* := \inf(\text{MOM}_r) = \min(\text{MOM}_r) = \sup(\text{SOS}_r) = \max(\text{SOS}_r).$$

Proof. Let μ be the normalized Lebesgue measure on $[-1, 1]$ and let y be its moment sequence ($y_k = \int_{-1}^1 x^k d\mu$). Then for any nonzero $p \in \mathbb{R}[x]_r$, $\int p^2 d\mu > 0$, hence $M_r(y) \succ 0$. Similarly, $g(x) = (1 - x^2)^3 > 0$ for all $x \in (-1, 1)$, so for nonzero $q \in \mathbb{R}[x]_{r-3}$, $\int gq^2 d\mu > 0$ and $M_{r-3}(gy) \succ 0$. Thus (MOM_r) is strictly feasible and conic duality yields strong duality and dual attainment.

To prove primal attainment, observe that the feasible set is non-empty and closed. From identity (3), for any feasible y it holds $\ell_y(2 - x^2) = \ell_y(p_0) + \ell_y(g) \geq 0$. Since $\ell_y(1) = 1$, this gives $\ell_y(x^2) \leq 2$. Next, for any integer $k \geq 2$ we have the pointwise inequality on \mathbb{R} : $x^{2k} \leq 2^{k-1}x^2$ whenever $x^2 \leq 2$. To use this within the moment constraints, we encode the bound $x^2 \leq 2$ via the SOS polynomial $(2 - x^2)(x^{k-1})^2$, whose degree is $2k$. Indeed, for every $k \in \{1, \dots, r\}$ the polynomial $(2 - x^2)x^{2(k-1)}$ is in the quadratic module. Hence for every feasible y , $0 \leq \ell_y((2 - x^2)x^{2(k-1)}) = 2y_{2k-2} - y_{2k}$. This yields the recursion $y_{2k} \leq 2y_{2k-2}$ for $k = 1, \dots, r$. Since $y_0 = 1$, we obtain by induction $0 \leq y_{2k} \leq 2^k$ for all $k = 0, 1, \dots, r$. For odd moments, the Cauchy-Schwarz inequality with $M_r(y) \succeq 0$ gives $|y_{2k+1}|^2 = |\ell_y(x^k \cdot x^{k+1})|^2 \leq \ell_y(x^{2k})\ell_y(x^{2k+2}) = y_{2k}y_{2k+2} \leq 2^k 2^{k+1} = 2^{2k+1}$. Thus all coordinates y_0, \dots, y_{2r} are uniformly bounded over the feasible set. The objective $y \mapsto \ell_y(f)$ is linear hence continuous. Minimizing a continuous function over a nonempty compact set attains its minimum. \square

Lemma 3. *Any dual-feasible triple (ε, p, q) in (SOS_r) satisfies $\varepsilon \leq 0$. Moreover, $\varepsilon = 0$ is not feasible in (SOS_r) for any finite r . In particular, $\varepsilon_r^* < 0$ for all $r \geq 3$.*

Proof. Evaluating $f(x) - \varepsilon = p(x) + g(x)q(x)$ at $x = 1$ gives $-\varepsilon = p(1) \geq 0$. Assume by contradiction that $\varepsilon = 0$ is feasible, i.e. $1 - x^2 = p(x) + (1 - x^2)^3q(x)$ for p, q SOS. At $x = 1$, the right-hand side equals $p(1)$, so $p(1) = 0$. Since p is SOS, every real root has even multiplicity;

therefore p vanishes at $x = 1$ with multiplicity at least 2. The term $(1 - x^2)^3 q(x)$ vanishes at $x = 1$ with multiplicity at least 3. Hence the right-hand side vanishes at $x = 1$ with multiplicity at least 2, whereas the left-hand side $1 - x^2$ has a simple zero at $x = 1$, a contradiction. \square

2.4 Main theorem and proof strategy

Theorem 1 (Exact relaxation value). *For every integer $r \geq 3$, the order- r moment-SOS relaxation value is*

$$\varepsilon_r^* = -\frac{1}{r(r-2)}.$$

To prove this result, we will construct analytically:

- in Section 3, a dual-feasible SOS certificate with $\varepsilon = -1/(r(r-2))$;
- in Section 4, a primal-feasible moment sequence y with $\ell_y(1 - x^2) = -1/(r(r-2))$.

The proof of Theorem 1 is then given in Section 5.

3 The SOS side: a pure-square identity

The numerical solution of (SOS_r) suggests that optimal SOS polynomials are perfect squares. Motivated by this observation, we seek an identity of the form

$$r(r-2)(1-x^2) + 1 = A_r(x)^2 + 4(1-x^2)^3 B_r(x)^2, \quad (4)$$

for some $A_r \in \mathbb{R}[x]_r$, $B_r \in \mathbb{R}[x]_{r-3}$, which would yield the order- r certificate

$$1 - x^2 + \frac{1}{r(r-2)} = \underbrace{\frac{A_r(x)^2}{r(r-2)}}_{p(x)} + (1-x^2)^3 \underbrace{\frac{4B_r(x)^2}{r(r-2)}}_{q(x)}, \quad (5)$$

hence $\varepsilon = -1/(r(r-2))$ would be feasible in (SOS_r) .

3.1 Chebyshev and Gegenbauer polynomials

Let T_k denote the Chebyshev polynomial of the first kind, defined by the three-term recurrence

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x), \quad T_0(x) = 1, \quad T_1(x) = x.$$

See [1, Ch. 22, §22.7, Table 22.7.4] for the recurrence and [1, Ch. 22, §22.4, Eq. (22.4.4)] for the initial values.

Let $P_k^{(2)}$ denote the Gegenbauer polynomial of parameter 2, defined by the recurrence

$$(k+1)P_{k+1}^{(2)}(x) = 2(k+2)xP_k^{(2)}(x) - (k+3)P_{k-1}^{(2)}(x), \quad P_0^{(2)}(x) = 1, \quad P_1^{(2)}(x) = 4x.$$

Gegenbauer polynomials are also called Jacobi's ultraspherical polynomials [19, Sec. 4.7], see [19, (4.7.17)] for the recurrence. See also [1, Ch. 22, §22.7, Eq. (22.7.3)] for the recurrence and [1, Ch. 22, §22.4, Eq. (22.4.2)] for the initial values.

Define, for $r \geq 3$,

$$A_r(x) := \frac{r-2}{2}T_r(x) - \frac{r}{2}T_{r-2}(x), \quad B_r(x) := P_{r-3}^{(2)}(x). \quad (6)$$

3.2 Proof of the identity via a Pythagorean trigonometric argument

Lemma 4. *For every $r \geq 3$, the polynomials A_r, B_r defined in (6) satisfy the identity (4).*

Proof. Fix $\theta \in \mathbb{R}$ and $x = \cos \theta$. Recall the classical trigonometric representations [1, Ch. 22, §22.3, Eqs. (22.3.15)–(22.3.16)], [19, (1.12.3)]:

$$T_k(\cos \theta) = \cos(k\theta), \quad U_k(\cos \theta) = \frac{\sin(k+1)\theta}{\sin \theta} \quad (7)$$

where U_k denote the Chebyshev polynomial of the second kind, defined by the recurrence

$$U_{k+1}(x) = 2x U_k(x) - U_{k-1}(x), \quad U_0(x) = 1, \quad U_1(x) = 2x.$$

See [1, Ch. 22, §22.7, Eq. (22.7.5)] for the recurrence and [1, Ch. 22, §22.4, Eq. (22.4.5)] for the initial values.

Moreover, the Chebyshev polynomial of the second kind is a special Gegenbauer polynomial:

$$U_k(x) = P_k^{(1)}(x)$$

see [1, Ch. 22, §22.5, Eq. (22.5.34)] or [19, (4.7.2)]. Using the Gegenbauer derivative identity [19, (4.7.14)]:

$$\frac{d}{dx} P_{k+1}^{(1)}(x) = 2 P_k^{(2)}(x),$$

we obtain

$$P_k^{(2)}(x) = \frac{1}{2} \frac{d}{dx} U_{k+1}(x).$$

Combining this with (7) and the chain rule $\frac{d}{dx} = \frac{d\theta}{dx} \frac{d}{d\theta} = -\frac{1}{\sin \theta} \frac{d}{d\theta}$ yields the trigonometric form

$$P_k^{(2)}(\cos \theta) = \frac{\sin(k+2)\theta \cos \theta - (k+2) \cos(k+2)\theta \sin \theta}{2 \sin^3 \theta}. \quad (8)$$

Let $\phi := (r-1)\theta$. From (6) and (7), we obtain

$$A_r(\cos \theta) = \frac{r-2}{2} \cos r\theta - \frac{r}{2} \cos(r-2)\theta. \quad (9)$$

Recall the addition formulas

$$\cos r\theta = \cos(\phi + \theta) = \cos \phi \cos \theta - \sin \phi \sin \theta,$$

$$\cos(r-2)\theta = \cos(\phi - \theta) = \cos \phi \cos \theta + \sin \phi \sin \theta.$$

Substitute these into (9):

$$\begin{aligned} A_r(\cos \theta) &= \frac{r-2}{2} (\cos \phi \cos \theta - \sin \phi \sin \theta) - \frac{r}{2} (\cos \phi \cos \theta + \sin \phi \sin \theta) \\ &= \left(\frac{r-2}{2} - \frac{r}{2} \right) \cos \phi \cos \theta + \left(-\frac{r-2}{2} - \frac{r}{2} \right) \sin \phi \sin \theta \\ &= -\cos \phi \cos \theta - (r-1) \sin \phi \sin \theta. \end{aligned} \quad (10)$$

Define

$$z := (\cos \theta - i(r-1) \sin \theta) e^{i\phi}.$$

Then

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

Moreover, expanding z reveals its real and imaginary parts

$$\Re z = \cos \theta \cos \phi + (r-1) \sin \theta \sin \phi, \quad \Im z = \cos \theta \sin \phi - (r-1) \sin \theta \cos \phi.$$

By (6)-(8)-(10) we get

$$A_r(\cos \theta)^2 = (\Re z)^2, \quad 4 \sin^6 \theta B_r(\cos \theta)^2 = (\Im z)^2.$$

Therefore

$$A_r(\cos \theta)^2 + 4(1 - \cos^2 \theta)^3 B_r(\cos \theta)^2 = (\Re z)^2 + (\Im z)^2 = |z|^2 = 1 + r(r-2) \sin^2 \theta.$$

Since $\sin^2 \theta = 1 - \cos^2 \theta = 1 - x^2$, we obtain (4) at $x = \cos \theta$. Both sides are polynomials in x and agree on infinitely many $x \in [-1, 1]$, hence they agree identically on \mathbb{R} . \square

3.3 Expressions for low relaxation orders

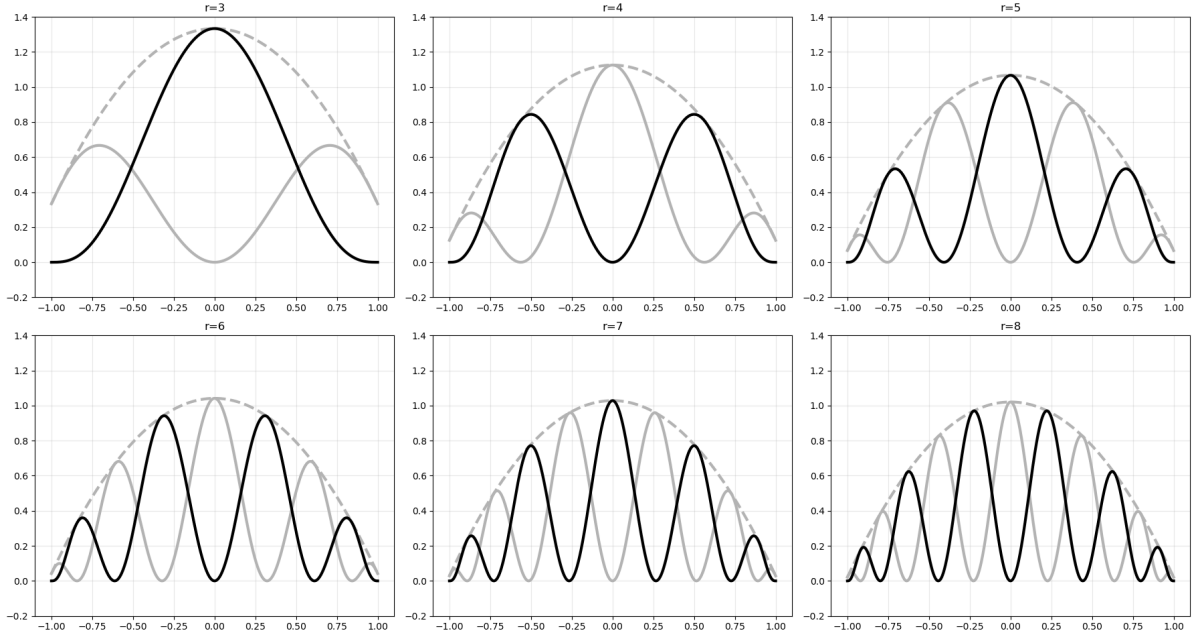


Figure 1: For each relaxation order $r \in \{3, \dots, 8\}$ we represent, over $x \in [-1, 1]$, the SOS identity (5). The dashed gray curve is the shifted objective $1 - x^2 + 1/r(r-2)$, the plain gray curve is the term $A_r(x)^2/r(r-2)$, and the black curve is the term $4(1-x^2)^3 B_r(x)^2/r(r-2)$. The value $-\varepsilon_r = 1/r(r-2)$ can be read from the gray curves at $x = \pm 1$.

For illustration, the polynomials (A_r, B_r) in the monomial basis are:

$$\begin{aligned} A_3(x) &= 2x^3 - 3x, & B_3(x) &= 1, \\ A_4(x) &= 8x^4 - 12x^2 + 3, & B_4(x) &= 4x, \\ A_5(x) &= 24x^5 - 40x^3 + 15x, & B_5(x) &= 12x^2 - 2, \\ A_6(x) &= 64x^6 - 120x^4 + 60x^2 - 5, & B_6(x) &= 32x^3 - 12x, \\ A_7(x) &= 160x^7 - 336x^5 + 210x^3 - 35x, & B_7(x) &= 80x^4 - 48x^2 + 3, \\ A_8(x) &= 384x^8 - 896x^6 + 672x^4 - 168x^2 + 7, & B_8(x) &= 192x^5 - 160x^3 + 24x. \end{aligned}$$

The corresponding SOS identities (5) are visualized on Figure 1.

4 The moment side: atoms, weights, and positivity

In this section we construct a primal-feasible moment sequence y attaining the value $\ell_y(1 - x^2) = -1/(r(r - 2))$.

4.1 Complementarity suggests atomic support on the roots of A_r

The identity (4) implies that if $(\varepsilon_r^*, s_0^*, s_1^*)$ would be an optimal certificate then

$$\varepsilon_r^* = -\frac{1}{r(r-2)}, \quad s_0^*(x) = \frac{1}{r(r-2)} A_r(x)^2, \quad s_1^*(x) = \frac{4}{r(r-2)} B_r(x)^2.$$

Complementarity with a primal optimal y would read

$$\ell_y(s_0^*) = 0, \quad \ell_y(g s_1^*) = 0.$$

Since s_0^* and $g s_1^*$ are nonnegative polynomials, these equalities force $s_0^* = 0$ and $g s_1^* = 0$ on the support of any representing measure of y . In particular, $\text{spt}\mu \subseteq \{x \in \mathbb{R} : A_r(x) = 0\}$ is a natural candidate.

4.2 Roots of A_r and their location

A key structural fact is that A_r is (up to scaling) a Jacobi polynomial $P_r^{(\alpha, \beta)}$ with parameters $\alpha = \beta = -\frac{3}{2}$. See e.g. [19, Sec. 2.4] for the definition of Jacobi polynomials, and [19, Chap. 4] for their properties.

Lemma 5 (Jacobi representation). *For every $r \geq 3$,*

$$A_r(x) = \frac{2\sqrt{\pi} \Gamma(r+1)}{\Gamma(r-\frac{1}{2})} P_r^{(-3/2, -3/2)}(x)$$

where Γ is the Gamma function satisfying $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\Gamma(1) = 1$ and $\Gamma(k+1) = k!$ for integer k .

Proof. Fix $\theta \in \mathbb{R}$ and $x = \cos \theta$. First let us establish a derivative identity for A_r . Using the classical trigonometric representations (7), differentiate $T_r(\cos \theta) = \cos(r\theta)$ with respect to θ and use $\frac{dx}{d\theta} = -\sin \theta$ to obtain the standard derivative formula

$$\frac{d}{dx} T_r(x) = r U_{r-1}(x). \tag{11}$$

Using (6), a direct differentiation and a short simplification yield

$$\frac{d}{dx} A_r(x) = r(r-2) T_{r-1}(x). \tag{12}$$

Chebyshev polynomials have the following Jacobi representation:

$$T_k(x) = \frac{k! \sqrt{\pi}}{\Gamma(k + \frac{1}{2})} P_k^{(-1/2, -1/2)}(x), \tag{13}$$

see [1, Ch. 22, §22.5, Eq. (22.5.31)] and [19, Chap. 4]. Applying (13) with $k = r - 1$ in (12) gives

$$\frac{d}{dx} A_r(x) = \frac{r(r-2) \Gamma(r) \sqrt{\pi}}{\Gamma(r - \frac{1}{2})} P_{r-1}^{(-1/2, -1/2)}(x). \tag{14}$$

For Jacobi polynomials one has the derivative relation

$$\frac{d}{dx} P_r^{(\alpha, \beta)}(x) = \frac{1}{2} (r + \alpha + \beta + 1) P_{r-1}^{(\alpha+1, \beta+1)}(x),$$

see [19, Eq. (4.5.5)]. With $\alpha = \beta = -\frac{3}{2}$, it becomes

$$\frac{d}{dx} P_r^{(-3/2, -3/2)}(x) = \frac{1}{2} (r - 2) P_{r-1}^{(-1/2, -1/2)}(x). \quad (15)$$

Comparing (14) with (15), we see that

$$\frac{d}{dx} A_r(x) = \left(\frac{2\sqrt{\pi} \Gamma(r+1)}{\Gamma(r - \frac{1}{2})} \right) \frac{d}{dx} P_r^{(-3/2, -3/2)}(x).$$

Therefore there exists a constant $c_r \in \mathbb{R}$ such that

$$A_r(x) = \frac{2\sqrt{\pi} \Gamma(r+1)}{\Gamma(r - \frac{1}{2})} P_r^{(-3/2, -3/2)}(x) + c_r. \quad (16)$$

Recall the standard normalization

$$P_r^{(-3/2, -3/2)}(1) = \frac{\Gamma(r - \frac{1}{2})}{\Gamma(-\frac{1}{2}) \Gamma(r+1)}$$

see [1, Ch. 22, §22.2, Table 22.2.1]. Evaluating (16) at $x = 1$ gives

$$A_r(1) = \frac{2\sqrt{\pi} \Gamma(r+1)}{\Gamma(r - \frac{1}{2})} \frac{\Gamma(r - \frac{1}{2})}{\Gamma(-\frac{1}{2}) \Gamma(r+1)} + c_r = \frac{2\sqrt{\pi}}{\Gamma(-\frac{1}{2})} + c_r.$$

Using $\Gamma(-\frac{1}{2}) = -2\sqrt{\pi}$, we get $A_r(1) = -1 + c_r$. From (6), it holds $A_r(1) = -1$ and hence $c_r = 0$, and (16) reduces to the desired proportionality. \square

Lemma 6 (Root distribution). *The roots of A_r are symmetric with respect to the origin, and zero is a root if r is odd. The roots are all real and simple. Moreover, $r - 2$ roots lie inside $(-1, 1)$ and 2 roots lie outside $[-1, 1]$.*

Proof. Since $T_k(-x) = (-1)^k T_k(x)$, the definition (6) gives

$$A_r(-x) = \frac{r-2}{2} (-1)^r T_r(x) - \frac{r}{2} (-1)^{r-2} T_{r-2}(x) = (-1)^r A_r(x)$$

which shows that the roots are symmetric with respect to the origin, and also that $A_r(0) = 0$ if r is odd. By Lemma 5, A_r and $P_r^{(-3/2, -3/2)}$ have the same roots. Then we apply [19, Thm. 6.72] with the notations $\alpha = \beta = -\frac{3}{2}$ and $n := r$. First, the excluded cases (6.72.1)–(6.72.3) do not occur: α, β are not negative integers, and $n + \alpha + \beta = n - 3 \geq 0$ for $n \geq 3$. Hence the roots are different from ± 1 and ∞ , and (by the discussion following (6.72.3)) are distinct. Next, compute the integers X, Y, Z from (6.72.5) using Klein's symbol $E(\cdot)$ in (6.72.4). Here $2n + \alpha + \beta + 1 = 2n - 2 > 0$ and $|\alpha| = |\beta| = \frac{3}{2}$, so

$$X = E\left(\frac{1}{2}((2n-2) - \frac{3}{2} - \frac{3}{2} + 1)\right) = E(n-2) = n-3, \quad Y = Z = E\left(\frac{1}{2}(-2n+3)\right) = 0.$$

Moreover,

$$\binom{n+\alpha}{n} = \binom{n-\frac{3}{2}}{n} = \frac{\Gamma(n-\frac{1}{2})}{\Gamma(n+1)\Gamma(-\frac{1}{2})} < 0, \quad \binom{2n+\alpha+\beta}{n} = \binom{2n-3}{n} > 0.$$

Therefore, in (6.72.6) we obtain $N_1 = n - 2$ (both parity cases give the same result), while in (6.72.7)–(6.72.8) we are in the strictly negative branches and get $N_2 = N_3 = 1$. Thus $P_n^{(-3/2, -3/2)}$ has exactly $n - 2$ roots in $(-1, 1)$, one in $(-\infty, -1)$ and one in $(1, \infty)$, and they are all simple. \square

Based on Lemma 6, denoting by $\{x_i\}_{i=1}^r$ the roots of A_r , their pattern is as follows:

$$\begin{aligned} r \text{ even: } & 0 < x_1 < x_2 < \cdots < x_{\frac{r}{2}-1} < 1 < x_{\frac{r}{2}} =: x_{\text{out}}, \\ r \text{ odd: } & 0 = x_1 < x_2 < \cdots < x_{\frac{r-1}{2}} < 1 < x_{\frac{r+1}{2}} =: x_{\text{out}}. \end{aligned}$$

Graphs of polynomials A_r with their root distribution are displayed on Figure 2.

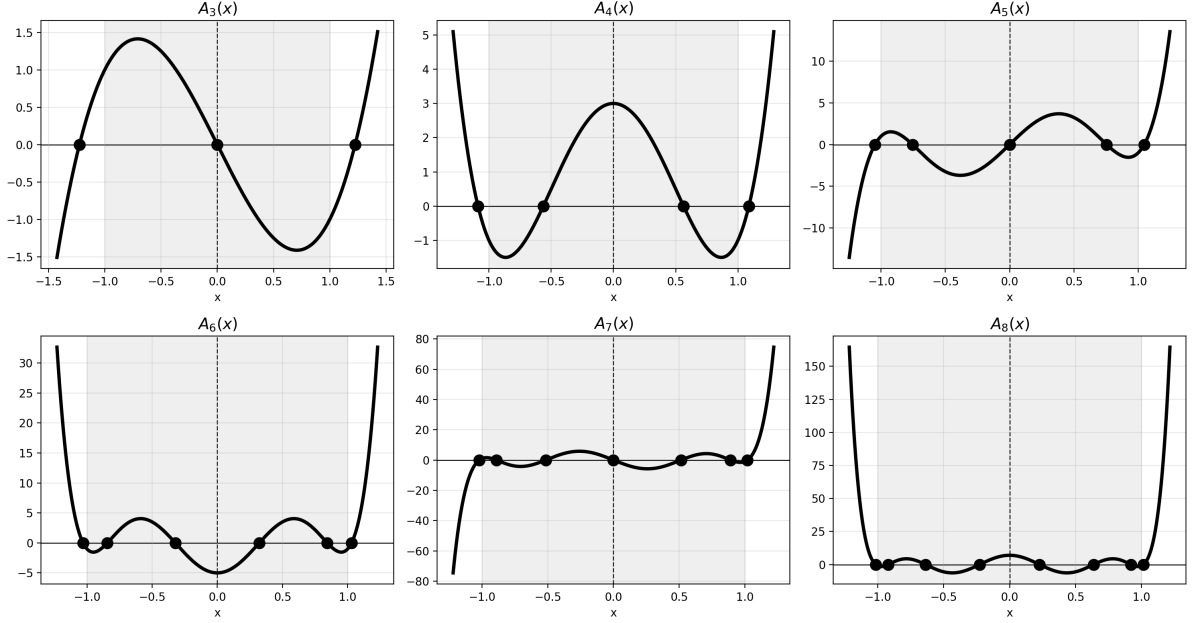


Figure 2: Graphs of polynomials A_r (thick black lines) and root distribution (thick black dots) for each relaxation order $r \in \{3, \dots, 8\}$. We observe that 2 of the r roots are outside the interval $[-1, 1]$ (light gray region).

4.3 The atomic measure and its basic moment identities

With $\{x_i\}_{i=1}^r$ denoting the roots of A_r , define the finitely atomic measure

$$\mu(dx) := \frac{1}{(r(r-2))^2} \sum_{i=1}^r \frac{1}{(1-x_i^2)^2} \delta_{x_i}(dx). \quad (17)$$

Let $y = \{y_k\}_{k=0}^{2r}$ be its moment sequence: $y_k = \int x^k \mu(dx)$.

Lemma 7 (Two residue identities). *It holds*

$$\sum_{i=1}^r \frac{1}{(1-x_i^2)^2} = (r(r-2))^2, \quad (18)$$

and

$$\sum_{i=1}^r \frac{1}{1-x_i^2} = -r(r-2). \quad (19)$$

Proof. The proof uses the classical residue formulas of complex analysis [17, Chap. 3, Sec. 2]. Let F be a holomorphic function on a punctured neighborhood $\{z \in \mathbb{C} : 0 < |z-a| < r\}$ and

admit a Laurent expansion

$$F(z) = \sum_{k=-m}^{\infty} f_k(z-a)^k.$$

The *residue* of F at a is the coefficient of $(z-a)^{-1}$ in this expansion, i.e.

$$\operatorname{Res}_{z=a} F(z) := f_{-1}.$$

Equivalently, if a is a pole of order n of F , then by [17, Ch. 3, Thm. 1.4],

$$\operatorname{Res}_{z=a} F(z) = \lim_{z \rightarrow a} \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}} \left((z-a)^n F(z) \right). \quad (20)$$

In particular, for a simple pole ($n=1$),

$$\operatorname{Res}_{z=a} F(z) = \lim_{z \rightarrow a} (z-a) F(z). \quad (21)$$

1) Proof of (18). Let

$$F(z) := \frac{A'_r(z)}{A_r(z)} \frac{1}{(1-z^2)^2}, \quad G(z) := \frac{A'_r(z)}{A_r(z)}.$$

Let $\mathcal{P} \subset \mathbb{R}$ denote the poles of F , which are the (simple) zeros x_i of A_r and the points ± 1 , which are poles of order 2.

Let $\Gamma_R = \{z \in \mathbb{C} : |z| = R\}$ with $R > 0$ large enough so that all poles lie inside. By the residue formula [17, Ch. 3, §2],

$$\int_{\Gamma_R} F(z) dz = 2\pi i \sum_{a \in \mathcal{P}} \operatorname{Res}_{z=a} F(z).$$

As $z \rightarrow \infty$, $G(z) = O(1/z)$ and $(1-z^2)^{-2} = O(1/z^4)$, hence $F(z) = O(1/z^5)$. Thus $\int_{\Gamma_R} F(z) dz \rightarrow 0$ as $R \rightarrow \infty$, and we obtain

$$\sum_{a \in \mathcal{P}} \operatorname{Res}_{z=a} F(z) = 0. \quad (22)$$

Each x_i is a simple zero of A_r , so $A'_r(x_i) \neq 0$ and F has a simple pole at x_i . Using (21):

$$\operatorname{Res}_{z=x_i} F(z) = \lim_{z \rightarrow x_i} (z-x_i) F(z) = \frac{1}{(1-x_i^2)^2}.$$

At $z=1$, F has a pole of order 2, hence from (20):

$$\operatorname{Res}_{z=1} F(z) = \lim_{z \rightarrow 1} \frac{d}{dz} \left((z-1)^2 F(z) \right) = \frac{d}{dz} \left(\frac{G(z)}{(1+z)^2} \right) \Big|_{z=1} = \frac{G'(1) - G(1)}{4}.$$

Similarly,

$$\operatorname{Res}_{z=-1} F(z) = \lim_{z \rightarrow -1} \frac{d}{dz} \left((z+1)^2 F(z) \right) = \frac{d}{dz} \left(\frac{G(z)}{(1-z)^2} \right) \Big|_{z=-1} = \frac{G'(-1) + G(-1)}{4}.$$

Now $A_r(1) = -1$, $A_r(-1) = -(-1)^r$ and from (12) we have $A'_r(1) = r(r-2)$, $A'_r(-1) = r(r-2)(-1)^{r-1}$ which implies

$$G(1) = -r(r-2), \quad G(-1) = r(r-2).$$

Moreover $G'(z) = A_r''(z)/A_r(z) - (A_r'(z)/A_r(z))^2$. Using (11) and (12) we have $A_r'(x) = r(r-2)T_{r-1}(x)$, $T_{r-1}'(x) = (r-1)U_{r-2}(x)$, and $U_{r-2}(1) = r-1$, $U_{r-2}(-1) = (-1)^{r-2}(r-1)$. So we get $A_r''(1) = r(r-2)(r-1)^2$, $A_r''(-1) = r(r-2)(-1)^r(r-1)^2$, hence

$$G'(1) = -r(r-2)(r-1)^2 - r^2(r-2)^2, \quad G'(-1) = -r(r-2)(r-1)^2 - r^2(r-2)^2.$$

Therefore $G'(1) - G(1) = -2r^2(r-2)^2$ and $G'(-1) + G(-1) = -2r^2(r-2)^2$, so

$$\operatorname{Res}_{z=1} F(z) = \operatorname{Res}_{z=-1} F(z) = -\frac{r^2(r-2)^2}{2}.$$

Substituting the residues into (22) yields

$$\sum_{a \in \mathcal{P}} \operatorname{Res}_{z=a} F(z) = \sum_{i=1}^r \frac{1}{(1-x_i^2)^2} - \frac{r^2(r-2)^2}{2} - \frac{r^2(r-2)^2}{2} = 0,$$

which is (18).

2) Proof of (19). Define

$$H(z) := \frac{A_r'(z)}{A_r(z)} \frac{1}{1-z^2} = G(z) \frac{1}{(1-z)(1+z)}.$$

The poles are simple at x_i and at ± 1 . As $z \rightarrow \infty$, $H(z) = O(1/z^3)$, so the same contour argument gives

$$\sum_{a \in \mathcal{P}} \operatorname{Res}_{z=a} H(z) = 0.$$

By (21):

$$\operatorname{Res}_{z=x_i} H(z) = \lim_{z \rightarrow x_i} (z - x_i) H(z) = \frac{1}{1 - x_i^2}.$$

Also,

$$\operatorname{Res}_{z=1} H(z) = \lim_{z \rightarrow 1} (z - 1) H(z) = \lim_{z \rightarrow 1} G(z) \frac{z - 1}{(1 - z)(1 + z)} = -\frac{G(1)}{2} = \frac{r(r-2)}{2},$$

and

$$\operatorname{Res}_{z=-1} H(z) = \lim_{z \rightarrow -1} (z + 1) H(z) = \lim_{z \rightarrow -1} G(z) \frac{z + 1}{(1 - z)(1 + z)} = \frac{G(-1)}{2} = \frac{r(r-2)}{2}.$$

Thus

$$\sum_{a \in \mathcal{P}} \operatorname{Res}_{z=a} H(z) = \sum_{i=1}^r \frac{1}{1 - x_i^2} + \frac{r(r-2)}{2} + \frac{r(r-2)}{2} = 0,$$

which is (19). □

Lemma 8 (Probability and objective value). *The measure μ in (17) satisfies*

$$y_0 = \int d\mu(x) = 1, \quad \ell_y(1 - x^2) = \int (1 - x^2) d\mu(x) = -\frac{1}{r(r-2)}.$$

Proof. The mass is

$$\int d\mu(x) = \frac{1}{(r(r-2))^2} \sum_{i=1}^r \frac{1}{(1 - x_i^2)^2} = 1$$

by (18). Next,

$$\int (1 - x^2) d\mu(x) = \frac{1}{(r(r-2))^2} \sum_{i=1}^r \frac{1 - x_i^2}{(1 - x_i^2)^2} = \frac{1}{(r(r-2))^2} \sum_{i=1}^r \frac{1}{1 - x_i^2} = -\frac{1}{r(r-2)}$$

by (19). □

4.4 Positivity of the moment matrix

Lemma 9. *Let μ be the measure (17) with moment vector y . Then its moment matrix of order r is positive semidefinite:*

$$M_r(y) \succeq 0.$$

Proof. Given any polynomial $p \in \mathbb{R}[x]_r$, using the definition (1) of the moment matrix:

$$p^\top M_r(y) p = \ell_y(p^2) = \int p(x)^2 d\mu(x).$$

Since μ is a nonnegative measure, it holds

$$\int p(x)^2 d\mu(x) \geq 0,$$

which implies $M_r(y) \succeq 0$. □

4.5 Positivity of the localizing matrix

Lemma 10. *Let μ be the measure (17) with moment vector y . Then its localizing moment matrix of order $r - 3$ is positive semidefinite:*

$$M_{r-3}(gy) \succeq 0.$$

Proof. Set $n := r - 3$ and write $b(x) := (1, x, \dots, x^n)^\top$. By (2), $M_n(gy) \succeq 0$ is equivalent to

$$\int (1 - x^2)^3 p(x)^2 d\mu(x) \geq 0 \quad \forall p \in \mathbb{R}[x]_n.$$

Using (17) it holds

$$\int (1 - x^2)^3 p(x)^2 d\mu(x) = \sum_{i=1}^r \frac{1 - x_i^2}{(r(r-2))^2} p(x_i)^2.$$

Form Lemma 6 we know that among the r roots of A_r , exactly $r - 2$ satisfy $|x_i| < 1$ and two are $\pm x_{\text{out}}$ with $x_{\text{out}} > 1$. Define the positive weights

$$w_i := \frac{1 - x_i^2}{(r(r-2))^2} > 0 \quad (|x_i| < 1), \quad w_{\text{out}} := \frac{x_{\text{out}}^2 - 1}{(r(r-2))^2} > 0,$$

and the matrices

$$H := \sum_{|x_i| < 1} w_i b(x_i) b(x_i)^\top, \quad V := \sqrt{w_{\text{out}}} [b(x_{\text{out}}) \ b(-x_{\text{out}})].$$

Then the localizing matrix decomposes as

$$M_n(gy) = H - VV^\top. \tag{23}$$

Since the inner measure $\sum_{|x_i| < 1} w_i \delta_{x_i}$ has $n + 1$ distinct nodes in $(-1, 1)$ with positive weights, its moment matrix is positive definite, i.e. $H \succ 0$. Therefore, by a rank-two Schur complement, we have the equivalence

$$M_n(gy) = H - VV^\top \succeq 0 \iff \begin{pmatrix} H & V \\ V^\top & I_2 \end{pmatrix} \succeq 0 \iff G := V^\top H^{-1} V \preceq I_2. \tag{24}$$

Define the Christoffel-Darboux kernel of the inner measure

$$K_n(x, y) := b(x)^\top H^{-1} b(y),$$

see e.g. [9, Chap. 2]. Then

$$G = w_{\text{out}} \begin{pmatrix} K_n(x_{\text{out}}, x_{\text{out}}) & K_n(x_{\text{out}}, -x_{\text{out}}) \\ K_n(x_{\text{out}}, -x_{\text{out}}) & K_n(x_{\text{out}}, x_{\text{out}}) \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix},$$

with $\alpha := w_{\text{out}} K_n(x_{\text{out}}, x_{\text{out}})$ and $\beta := w_{\text{out}} K_n(x_{\text{out}}, -x_{\text{out}})$. Hence

$$\lambda_+ = \alpha + \beta \text{ with eigenvector } u_+ = (1, 1), \quad \lambda_- = \alpha - \beta \text{ with eigenvector } u_- = (1, -1).$$

Let $q(x) := B_r(x) = P_n^{(2)}(x)$ and let q be its coefficient vector in the basis b , i.e. $q(x) = q^\top b(x)$. We claim that

$$M_n(gy)q = 0. \quad (25)$$

Indeed, it suffices to show $\ell_y(gx^k q) = 0$ for $k = 0, \dots, n$. Using (17),

$$\ell_y(gx^k q) = \frac{1}{(r(r-2))^2} \sum_{i=1}^r (1-x_i^2) x_i^k q(x_i),$$

so we have to prove the discrete orthogonality relation

$$\sum_{i=1}^r (1-x_i^2) x_i^k B_r(x_i) = 0, \quad k = 0, \dots, n. \quad (26)$$

Recall the two polynomial identities derived from the trigonometric formulas (7), (8) and (10):

$$A_r(x) = -\left(x T_{r-1}(x) + (r-1)(1-x^2) U_{r-2}(x)\right), \quad (27)$$

$$2(1-x^2) B_r(x) = x U_{r-2}(x) - (r-1) T_{r-1}(x). \quad (28)$$

Let x_i be such that $A_r(x_i) = 0$. Since $x_i \neq \pm 1$, (27) gives

$$x_i T_{r-1}(x_i) = -(r-1)(1-x_i^2) U_{r-2}(x_i), \quad \text{hence} \quad U_{r-2}(x_i) = -\frac{x_i}{(r-1)(1-x_i^2)} T_{r-1}(x_i).$$

Substitute this into (28) at $x = x_i$:

$$\begin{aligned} 2(1-x_i^2) B_r(x_i) &= x_i U_{r-2}(x_i) - (r-1) T_{r-1}(x_i) = -\left(\frac{x_i^2}{(r-1)(1-x_i^2)} + (r-1)\right) T_{r-1}(x_i) \\ &= -\frac{x_i^2 + (r-1)^2(1-x_i^2)}{(r-1)(1-x_i^2)} T_{r-1}(x_i) = -\frac{1+r(r-2)(1-x_i^2)}{(r-1)(1-x_i^2)} T_{r-1}(x_i), \end{aligned}$$

where we used the algebraic identity $x^2 + (r-1)^2(1-x^2) = 1 + r(r-2)(1-x^2)$. Multiplying by $(r-1)(1-x_i^2)$ yields

$$2(r-1)(1-x_i^2)^2 B_r(x_i) = -(1+r(r-2)(1-x_i^2)) T_{r-1}(x_i). \quad (29)$$

Now evaluate the SOS identity (4) at $x = x_i$. Since $A_r(x_i) = 0$, it reduces to

$$1 + r(r-2)(1-x_i^2) = 4(1-x_i^2)^3 B_r(x_i)^2. \quad (30)$$

Insert (30) into (29):

$$2(r-1)(1-x_i^2)^2 B_r(x_i) = -4(1-x_i^2)^3 B_r(x_i)^2 T_{r-1}(x_i).$$

We now justify division by $(1-x_i^2)^2 B_r(x_i)$. First, $x_i \neq \pm 1$ so $(1-x_i^2)^2 \neq 0$. Second, the Gegenbauer polynomial $B_r = P_{r-3}^{(2)}$ has all its real zeros in $(-1, 1)$, and if $|x_i| < 1$ then (30) gives $1+r(r-2)(1-x_i^2) > 1$, so $B_r(x_i) \neq 0$ as well. Hence $B_r(x_i) \neq 0$ for every root x_i of A_r . Dividing by $(1-x_i^2)^2 B_r(x_i)$ gives

$$(1-x_i^2) T_{r-1}(x_i) B_r(x_i) = -\frac{r-1}{2}.$$

Using the derivative identity (12):

$$A'_r(x) = r(r-2) T_{r-1}(x),$$

one gets the pointwise relation

$$(1-x_i^2) B_r(x_i) = -\frac{r(r-1)(r-2)}{2} \frac{1}{A'_r(x_i)}. \quad (31)$$

Next, for each $k \leq n = r-3$, consider the meromorphic function $F_k(z) := z^k/A_r(z)$. Its only finite poles are the simple zeros x_i of A_r , and as in the proof of Lemma 7, by the simple-pole residue formula [17, Chap. 3, Thm. 1.4],

$$\operatorname{Res}_{z=x_i} F_k(z) = \frac{x_i^k}{A'_r(x_i)}.$$

Since $\deg A_r = r$ and $k \leq r-3$, we have $F_k(z) = O(1/z^3)$ as $z \rightarrow \infty$; hence the contour integral over $|z| = R$ vanishes as $R \rightarrow \infty$. By the residue theorem [17, Chap. 3, Sec. 2],

$$0 = \sum_{i=1}^r \operatorname{Res}_{z=x_i} F_k(z) = \sum_{i=1}^r \frac{x_i^k}{A'_r(x_i)}.$$

Multiplying by $-\frac{r(r-1)(r-2)}{2}$ and using (31) yields (26), hence (25) holds.

Now combine (23) and (25):

$$(H - VV^\top)q = 0 \implies Hq = VV^\top q.$$

Left-multiplying by $V^\top H^{-1}$ gives

$$V^\top q = (V^\top H^{-1}V)(V^\top q) = G(V^\top q),$$

so 1 is an eigenvalue of G with eigenvector $V^\top q \neq 0$. Moreover, since $q(-x) = (-1)^n q(x)$,

$$V^\top q = \sqrt{w_{\text{out}}} (q(x_{\text{out}}), q(-x_{\text{out}}))^\top = \sqrt{w_{\text{out}}} q(x_{\text{out}}) (1, (-1)^n)^\top,$$

so the eigenvector is precisely $u_{(-1)^n} = (1, (-1)^n)$ and therefore

$$\lambda_{(-1)^n} = 1. \quad (32)$$

It remains to show $\max(\lambda_+, \lambda_-) = 1$. For this we use only the sign of β . Let $\{\pi_k\}$ be the monic orthogonal polynomials for the inner product $\langle f, g \rangle_{\text{inn}} := \sum_{|x_i| < 1} w_i f(x_i) g(x_i)$ and $h_k := \langle \pi_k, \pi_k \rangle_{\text{inn}} > 0$. The Christoffel-Darboux formula [9, §3.1.1] gives, for $x \neq y$,

$$K_n(x, y) = \frac{1}{h_n} \frac{\pi_{n+1}(x)\pi_n(y) - \pi_n(x)\pi_{n+1}(y)}{x - y}.$$

Since ν_{inn} is symmetric, $\pi_k(-x) = (-1)^k \pi_k(x)$, so with $(x, y) = (x_{\text{out}}, -x_{\text{out}})$,

$$K_n(x_{\text{out}}, -x_{\text{out}}) = \frac{(-1)^n}{x_{\text{out}} h_n} \pi_n(x_{\text{out}}) \pi_{n+1}(x_{\text{out}}).$$

All zeros of π_k lie in $(-1, 1)$, hence $\pi_k(x_{\text{out}}) > 0$ for $x_{\text{out}} > 1$, so $K_n(x_{\text{out}}, -x_{\text{out}}) \neq 0$ and $\text{sign}(K_n(x_{\text{out}}, -x_{\text{out}})) = (-1)^n$. As $w_{\text{out}} > 0$, we obtain

$$\text{sign}(\beta) = (-1)^n,$$

so the eigenvalue associated with $u_{(-1)^n}$ is exactly $\alpha + |\beta| = \max(\lambda_+, \lambda_-) = 1$. Together with (32), this yields $\max(\lambda_+, \lambda_-) = 1$ hence $G \preceq I_2$. Finally, (24) implies $M_n(gy) = M_{r-3}(gy) \succeq 0$. \square

4.6 Further properties of the moments

Lemma 11. *For the moment sequence y of the measure μ in (17), it holds $y_k = 0$ for odd k and $1 + k(y_2 - 1) \leq y_{2k} \leq y_{2(k+1)}$ for all $k \geq 0$.*

Proof. The roots of A_r are symmetric: if x_i is a root then so is $-x_i$. The weights in (17) depend only on x_i^2 , hence the atom at x_i and the atom at $-x_i$ have the same weight. Therefore, for any odd integer k ,

$$y_k = \int x^k d\mu(x) = \int (-x)^k d\mu(x) = - \int x^k d\mu(x),$$

hence $y_k = 0$.

Fix an integer $k \geq 1$ and set $u(x) := x^2 \geq 0$. Then $y_{2k} = \int u(x)^k d\mu(x)$. By convexity of $u \mapsto u^k$ on $[0, \infty)$, its tangent inequality at $u = 1$ reads

$$u^k(x) \geq 1 + k(u(x) - 1) \quad \forall x.$$

Integrating this inequality with respect to μ gives

$$y_{2k} = \int u(x)^k d\mu(x) \geq \int (1 + k(u(x) - 1)) d\mu(x) = 1 + k \left(\int u(x) d\mu(x) - \int 1 d\mu(x) \right) = 1 + k(y_2 - 1),$$

where we used Lemma 8 and $\int d\mu = 1$. The same Lemma 8 yields

$$y_2 = \int x^2 d\mu(x) = 1 - \int (1 - x^2) d\mu(x) = 1 + \frac{1}{r(r-2)} > 1,$$

hence $y_{2k} > 1$ for all $k \geq 1$.

To prove that the even moments are nondecreasing, define two nondecreasing functions on $[0, \infty)$:

$$f(u) := u^{k-1}, \quad g(u) := u - 1.$$

Consider the product measure $\mu \otimes \mu$ on $\mathbb{R} \times \mathbb{R}$. Since f and g are nondecreasing, for all $x, x' \in \mathbb{R}$ we have the pointwise inequality

$$(f(u(x)) - f(u(x')))(g(u(x)) - g(u(x'))) \geq 0.$$

Integrating with respect to $\mu(x)\mu(x')$ gives

$$\begin{aligned} 0 &\leq \iint (f(u(x)) - f(u(x')))(g(u(x)) - g(u(x'))) d\mu(x) d\mu(x') \\ &= 2 \int f(u(x))g(u(x)) d\mu(x) - 2 \left(\int f(u(x)) d\mu(x) \right) \left(\int g(u(x)) d\mu(x) \right). \end{aligned}$$

Hence

$$\int f(u(x))g(u(x)) d\mu(x) \geq \left(\int f(u(x)) d\mu(x) \right) \left(\int g(u(x)) d\mu(x) \right). \quad (33)$$

Substituting $f(u) = u^{k-1}$ and $g(u) = u - 1$ in (33) yields

$$\int u(x)^{k-1}(u(x) - 1) d\mu(x) \geq \left(\int u(x)^{k-1} d\mu(x) \right) \left(\int (u(x) - 1) d\mu(x) \right).$$

Now

$$\int u(x)^{k-1}(u(x) - 1) d\mu(x) = \int u(x)^k d\mu(x) - \int u(x)^{k-1} d\mu(x) = y_{2k} - y_{2(k-1)},$$

and

$$\int (u(x) - 1) d\mu(x) = \int x^2 d\mu(x) - \int 1 d\mu(x) = y_2 - 1.$$

Therefore we obtain the moment gap bound:

$$y_{2k} - y_{2(k-1)} \geq y_{2(k-1)}(y_2 - 1).$$

Since $y_{2(k-1)} \geq 0$ and $y_2 - 1 = \frac{1}{r(r-2)} > 0$, it follows that $y_{2k} \geq y_{2(k-1)}$ for all $k \geq 1$. \square

Lemma 12. *For the moment sequence y of the measure μ in (17), it holds $\lim_{r \rightarrow \infty} y_{2r} = 1$.*

Proof. By Lemma 11, $y_{2r} \geq 1$. For the upper bound, note that μ is supported on $|x| \leq x_{\text{out}}$, hence

$$y_{2r} = \int x^{2r} d\mu \leq x_{\text{out}}^{2r}.$$

It remains to show $x_{\text{out}}^{2r} \rightarrow 1$. Writing $x_{\text{out}} = \cosh t_r$ for some $t_r > 0$ and using $T_k(\cosh t) = \cosh(kt)$, the equation $A_r(x_{\text{out}}) = 0$ becomes

$$(r-2) \cosh(rt_r) = r \cosh((r-2)t_r).$$

A Taylor expansion of \cosh for small t_r implies $t_r = \Theta(1/r)$ and therefore $x_{\text{out}} - 1 = \cosh t_r - 1 = \Theta(t_r^2) = \Theta(1/r^2)$. Consequently,

$$\log(x_{\text{out}}^{2r}) = 2r \log(1 + O(1/r^2)) = O(1/r) \rightarrow 0,$$

so $x_{\text{out}}^{2r} \rightarrow 1$ and hence $y_{2r} \rightarrow 1$. \square

4.7 Rational expressions of the moments

The optimal moments are actual *rational* numbers. To see why, define a polynomial $P_r(u) \in \mathbb{Q}[u]$ by extracting the even part of A_r :

$$A_r(x) = \begin{cases} P_r(x^2), & r \text{ even}, \\ x P_r(x^2), & r \text{ odd}. \end{cases}$$

Then the nonzero squared nodes $u_i > 0$ are precisely the roots of $P_r(u)$. The crucial point is that the moment sum

$$y_{2p} := \int x^{2p} d\mu_r(x) = w_0 0^p + 2 \sum_{u_i > 0} w_i u_i^p, \quad p \geq 0$$

is a *symmetric rational function* of the roots of an explicit polynomial with rational coefficients.

Let $C_r \in \mathbb{Q}^{m \times m}$ denote the companion matrix of the monic polynomial proportional to P_r (here $m = \deg P_r = \lfloor r/2 \rfloor$), so that the eigenvalues of C_r are exactly the roots (u_i) of P_r . Since $u = 1$ is not a root of P_r , the matrix $I - C_r$ is invertible, and the spectral mapping theorem yields

$$\text{tr}(C_r^p (I - C_r)^{-2}) = \sum_{u_i > 0} \frac{u_i^p}{(1 - u_i)^2}.$$

Therefore, for every $p \geq 1$,

$$y_{2p} = \frac{2}{r^2(r-2)^2} \text{tr}(C_r^p (I - C_r)^{-2}). \quad (34)$$

4.8 Moment vectors for low relaxation orders

Using the trace formula (34) we can compute the first moment vectors.

Order $r = 3$.

$$(y_0, y_2, y_4, y_6) = \left(1, \frac{4}{3}, 2, 3\right).$$

Order $r = 4$.

$$(y_0, y_2, y_4, y_6, y_8) = \left(1, \frac{9}{8}, \frac{21}{16}, \frac{99}{64}, \frac{117}{64}\right).$$

Order $r = 5$.

$$(y_0, y_2, y_4, y_6, y_8, y_{10}) = \left(1, \frac{16}{15}, \frac{52}{45}, \frac{34}{27}, \frac{223}{162}, \frac{1465}{972}\right).$$

Order $r = 6$.

$$(y_0, y_2, y_4, y_6, y_8, y_{10}, y_{12}) = \left(1, \frac{25}{24}, \frac{35}{32}, \frac{295}{256}, \frac{7475}{6144}, \frac{21075}{16384}, \frac{178425}{131072}\right).$$

Order $r = 7$.

$$(y_0, y_2, y_4, y_6, y_8, y_{10}, y_{12}, y_{14}) = \left(1, \frac{36}{35}, \frac{186}{175}, \frac{963}{875}, \frac{2853}{2500}, \frac{29613}{25000}, \frac{1230417}{1000000}, \frac{12788307}{10000000}\right).$$

Order $r = 8$.

$$(y_0, y_2, y_4, y_6, y_8, y_{10}, y_{12}, y_{14}, y_{16}) = \left(1, \frac{49}{48}, \frac{301}{288}, \frac{3703}{3456}, \frac{22799}{20736}, \frac{561883}{497664}, \frac{3463859}{2985984}, \frac{42726971}{35831808}, \frac{65904559}{53747712}\right).$$

5 Proof of Theorem 1

From Lemma 4, for every $r \geq 3$, the polynomials A_r, B_r defined in (6) satisfy the identity (4). Thus the triple

$$\varepsilon := -\frac{1}{r(r-2)}, \quad p(x) := \frac{A_r(x)^2}{r(r-2)}, \quad q(x) := \frac{4B_r(x)^2}{r(r-2)}$$

is feasible for (SOS_r) . Therefore

$$\sup(\text{SOS}_r) \geq -\frac{1}{r(r-2)}. \quad (35)$$

Let μ be the atomic measure (17) supported on the roots of A_r and let y be its moment vector. By Lemma 8, $\ell_y(1) = y_0 = 1$. By Lemma 9, $M_r(y) \succeq 0$. By Lemma 10, $M_{r-3}(gy) \succeq 0$. Hence y is feasible for (MOM_r) . Again by Lemma 8,

$$\ell_y(f) = -\frac{1}{r(r-2)},$$

so

$$\inf(\text{MOM}_r) \leq -\frac{1}{r(r-2)}. \quad (36)$$

By the weak duality Lemma 1), we have

$$\inf(\text{MOM}_r) \geq \sup(\text{SOS}_r).$$

Combining this inequality with (35) and (36) yields

$$\inf(\text{MOM}_r) = \sup(\text{SOS}_r) = -\frac{1}{r(r-2)}.$$

By Lemma 2, strong duality and attainment hold on both sides, so this common value is precisely the relaxation optimum ε_r^* . This proves Theorem 1. \square

6 Conclusion

In this paper we show that trigonometric properties of orthogonal polynomials can be exploited to construct in rational arithmetic an analytic solution of the semidefinite relaxations of Stengle's example [18], thereby settling the question of the exact convergence rate of the moment-SOS hierarchy.

It would be interesting to investigate whether similar techniques could be used to solve analytically the other low-dimensional POP examples used as challenging benchmarks for high-precision semidefinite solvers [5].

Key to the analytic solution of problem (SOS_r) is the pure square form of the polynomials p and q , expressed as a quadratic algebraic equation (4) in polynomials A_r and B_r . The rank-one structure of the Gram matrices of the SOS polynomials was also exploited in [8] to derive tight upper and lower bounds on the values of the moment-SOS hierarchy for a parametric POP problem. It is currently not well understood for which class of POP does optimality of a mom-SOS relaxation imply that the Gram matrices of the SOS dual multipliers are rank-one.

Acknowledgments

The author is grateful to Michal Kočvara for sharing the slides of his talk [5] and for insightful discussions. The author acknowledges the use of AI for assistance with brainstorming ideas, mathematical development, coding and drafting the manuscript. The final content, analysis and conclusions remain the sole responsibility of the author.

References

- [1] M. Abramowitz, I. A. Stegun. Handbook of mathematical functions with formulas, graphs, and mathematical tables. U.S. Government Printing Office, Applied Mathematics Series, Vol. 55, 1964. Reprinted by Dover Publications, 1972.
- [2] L. Baldi, L. Slot. Degree bounds for Putinar’s Positivstellensatz on the hypercube. *SIAM Journal on Applied Algebra and Geometry* 8(1):1–25, 2024.
- [3] S. Gribling, E. de Klerk, J. Vera. Revisiting the convergence rate of the Lasserre hierarchy for polynomial optimization over the hypercube. *arXiv:2505.00544*, 2025.
- [4] D. Henrion, M. Korda, J. B. Lasserre. The moment-SOS hierarchy. World Scientific, 2020.
- [5] M. Kočvara. Conjectures in polynomial optimization through high precision semidefinite programming (joint work with L. Baldi). SPOT seminar, Toulouse, 13 October 2025.
- [6] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization* 11(3):796–817, 2001.
- [7] M. Laurent, L. Slot. An overview of convergence rates for sum of squares hierarchies in polynomial optimization. *arXiv:2408.04417*, 2024.
- [8] D. Henrion, A. Le Franc, V. Magron. Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem. *SIAM Journal on Applied Algebra and Geometry* 9(1), 2025.
- [9] J. B. Lasserre, E. Pauwels, M. Putinar. The Christoffel-Darboux kernel for data analysis. Cambridge University Press, 2022.
- [10] J. Nie. Moment and polynomial optimization, SIAM, 2023.
- [11] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal* 42(3):969–984, 1993.
- [12] C. Scheiderer. Non-existence of degree bounds for weighted sums of squares representations. *Journal of Complexity* 21:823–844, 2005.
- [13] K. Schmüdgen. The K -moment problem for compact semi-algebraic sets. *Mathematische Annalen* 289:203–206, 1991.
- [14] K. Schmüdgen. The moment problem. Springer, 2017.
- [15] C. Schlosser, M. Tacchi-Bénard, A. Lazarev. Convergence rates for the moment-SoS hierarchy. *Numerical Algebra, Control and Optimization* 16:105–156, 2026.
- [16] L. Slot. Sum-of-squares hierarchies for polynomial optimization and the Christoffel-Darboux kernel. *SIAM Journal on Optimization* 32(4), 2022.

- [17] E. M. Stein, R. Shakarchi. Complex analysis. Princeton University Press, 2003.
- [18] G. Stengle. Complexity estimates for the Schmüdgen Positivstellensatz. *Journal of Complexity* 12:167-174, 1996.
- [19] G. Szegő. Orthogonal polynomials. AMS Colloquium Publications, Vol. 23, 4th ed., 1975.
- [20] T. Theobald. Real algebraic geometry and optimization. AMS, 2024