

CONVERGENCE RATES FOR LASSERRE’S UPPER-BOUND HIERARCHY ON $\mathbb{R}^d$ WITH GAUSSIAN MEASURE

SANDER GRIBLING, IGOR KLEP, AND POORNENDU KUMAR

ABSTRACT. We establish quantitative convergence rates for Lasserre’s measure-based upper-bound hierarchy on $\mathbb{R}^d$ using the standard Gaussian measure as the reference measure. For any polynomial f attaining its global infimum $f_{\min}$, we prove that the order- r upper bounds satisfy

$$\lambda_r - f_{\min} = O\left(\frac{1}{r}\right),$$

yielding, to the best of our knowledge, the first explicit convergence rate for this hierarchy on a noncompact domain. We further show that the rate is sharp by proving an $\Omega(1/r)$ rate for $f = x^2$, while demonstrating that, without attainment of the infimum, convergence can be arbitrarily slow: for every $m \geq 0$, we construct a polynomial whose upper bounds converge at rate $\Theta(r^{-1/(m+2)})$.

1. INTRODUCTION

Let $\mathbb{R}[x]$ be the space of polynomials and let $\mathbb{R}[x]_r$ be the space of polynomials with degree at most r . Given a polynomial $f \in \mathbb{R}[x]$ that is bounded from below, we consider the global polynomial optimization problem

$$(1) \quad f_{\min} = \inf_{x \in \mathbb{R}^d} f(x).$$

Semidefinite programming (SDP) serves as a cornerstone of polynomial optimization, linking sum-of-squares (SOS) certificates and Positivstellensätze from real algebraic geometry [Put93, BCR98, Mar08, Sce24] to moment problems [Las10, Scm17, Nie23]. This perspective underlies the moment–SOS hierarchies of Lasserre [Las01] (see also Parrilo [Par03] and Laurent’s survey [Lau09]), which replace the intractable nonconvex problem (1) with a hierarchy of solvable semidefinite relaxations. For the moment–SOS hierarchy of lower bounds for $f_{\min}$, quantitative results have been obtained using polynomial-kernel methods [FF21, GdKV26, Mag26], effective Positivstellensatz bounds [Scw05, BM23, BS24, BMP25], and correlative sparsity [KMR25]. Related phenomena concerning slow convergence have recently been investigated in [HLM25]. We refer to [LS26] for a recent overview.

In contrast to lower bounds certified by polynomial nonnegativity, Lasserre [Las11] introduced a complementary hierarchy of *upper* bounds based on measure-theoretic approximations. The idea is to minimize the expectation of f over probability measures possessing a sum-of-squares (SOS) density with respect to a prescribed reference measure. Specifically, let μ be a Borel measure on $\mathbb{R}^d$ with $\text{supp } \mu = \mathbb{R}^d$, whose moment sequence satisfies the Carleman condition [Scm17]. For each $r \geq \lceil \deg(f)/2 \rceil$, the r -th upper bound

2020 *Mathematics Subject Classification.* 90C23, 90C22, 44A60, 42C05.

Key words and phrases. polynomial optimization, Lasserre hierarchy, sum-of-squares, Gaussian measure, convergence rates, semidefinite programming.

is defined by the SDP

$$(2) \quad \begin{aligned} \lambda_r = \inf_{h \in \mathbb{R}[x]_r} \quad & \int_{\mathbb{R}^d} f(x) h^2(x) d\mu(x) \\ \text{s.t.} \quad & \int_{\mathbb{R}^d} h^2(x) d\mu(x) = 1. \end{aligned}$$

Equivalently, one may optimize over SOS densities $\sigma \in \Sigma[x]_{2r}$, where

$$\Sigma[x]_{2r} = \left\{ \sum_j h_j(x)^2 : h_j \in \mathbb{R}[x]_r \right\}.$$

The resulting sequence $(\lambda_r)_r$ is non-increasing and converges to $f_{\min}$ from above [Las11]. We refer the reader to [KMV26] for a noncommutative extension of Lasserre's upper-bound hierarchy.

Existing quantitative analyses for Lasserre's upper-bound hierarchy focus almost exclusively on *compact* sets [DW12, dKLS17, dKHL17, dKL22, SL22]; see [LS26] for an overview. For example, the convergence rate on a convex body with respect to the Lebesgue measure is $O((\log r)^2/r^2)$; the log-factors can be removed for simple sets such as the sphere and hypercube (for the right reference measure). Here, we consider the noncompact setting on $\mathbb{R}^d$ endowed with the standard Gaussian measure.

The strong tail bounds of the Gaussian measure may suggest a similar behaviour to the compact setting, however, we show there are significant differences.

1.1. Main result. Throughout, the ambient dimension is denoted by d . Our main result is an explicit convergence rate for Lasserre's upper-bound hierarchy when μ is the standard Gaussian (probability) measure on $\mathbb{R}^d$,

$$d\mu(x) = (2\pi)^{-d/2} e^{-\|x\|^2/2} dx.$$

For $r \in \mathbb{N}$, we consider Lasserre's upper bound hierarchy (2) applied to this measure.

Theorem 1. *Let $f \in \mathbb{R}[x]$ have degree at most N , and assume that f attains its global infimum on $\mathbb{R}^d$. Then there are constants $C_{f,d,N} > 0$ and $r_0 \in \mathbb{N}$ such that, for all $r \geq r_0$,*

$$0 \leq \lambda_r - f_{\min} \leq \frac{C_{f,d,N}}{r}.$$

We complement this upper bound with examples in Section 5 that illustrate both the rate and the role of the attainment assumption. For the univariate quadratic $f = x^2$, we show that the rate is sharp: the hierarchy converges at the rate $\Theta(1/r)$. A second class of examples, for which the infimum is not attained, shows the convergence can be arbitrarily slow among polynomial rates: for every $m \geq 0$ we exhibit a polynomial f of degree $2m + 4$ for which

$$\lambda_r - f_{\min} = \Theta\left(r^{-1/(m+2)}\right).$$

Figure 1 further illustrates the performance on two examples: a non-centered quadratic $f(x) = (x - 1)^2$, and the Motzkin polynomial $f(x, y) = x^4 y^2 + x^2 y^4 - 3x^2 y^2 + 1$.

1.2. Guide to the paper. Section 2 states the necessary background on Hermite polynomials. Section 3 contains our key technical lemma, establishing the desired localization in the univariate setting. Section 4 extends this to the multivariate setting to establish Theorem 1. Finally, Section 5 studies the hierarchy through two examples: a quadratic example with convergence rate $\Theta(1/r)$, and a family showing arbitrarily slow polynomial convergence when the infimum is not attained.

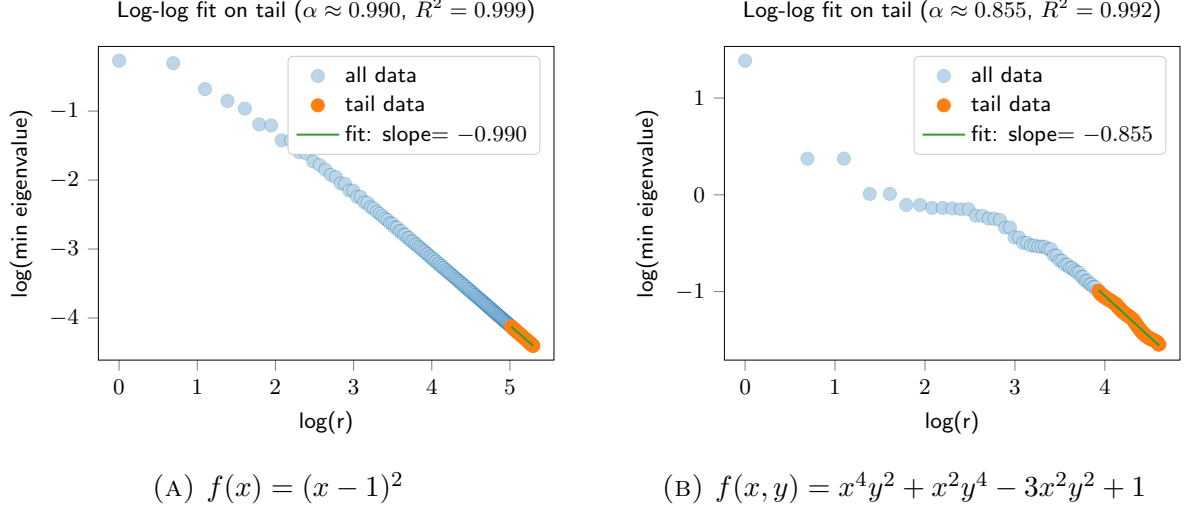


FIGURE 1. Log-log plot of the convergence rate on $(x - 1)^2$ (1a) ($1 \leq r \leq 200$) and the Motzkin polynomial (1b) ($1 \leq r \leq 100$). The power-law fit $\propto r^{-\alpha}$ is based on the last 50 values of r .

Acknowledgments. This work was performed within the project COMPUTE, which is funded within the QuantERA II Programme that has received funding from the EU's H2020 research and innovation programme under the GA No. 101017733 . IK also acknowledges support of the Slovenian Research Agency program P1-0222 and grants J1-50002, N1-0217, J1-60011, J1-50001, J1-3004 and J1-60025.

2. PRELIMINARIES ON HERMITE POLYNOMIALS

Consider the Gaussian probability measure on $\mathbb{R}$,

$$d\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

With respect to μ the orthogonal polynomials are the so-called probabilists' Hermite polynomials $\text{He}_n(x)$ given by

$$\text{He}_n(x) = (-1)^n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} \left(e^{-\frac{x^2}{2}} \right).$$

Then

$$\int_{-\infty}^{\infty} \text{He}_n(x) \text{He}_m(x) d\mu(x) = n! \delta_{nm}$$

so $h_n(x) = \frac{\text{He}_n(x)}{\sqrt{n!}}$, $n \in \mathbb{N}_0$ is an orthonormal basis of $\mathbb{R}[x]$ w.r.t. μ .

The Hermite polynomials satisfy the three-term recurrence relation

$$(3) \quad x \text{He}_m(x) = \text{He}_{m+1}(x) + m \text{He}_{m-1}(x),$$

whence

$$x^2 \text{He}_m(x) = \text{He}_{m+2}(x) + (2m+1) \text{He}_m(x) + m(m-1) \text{He}_{m-2}(x).$$

Equivalently,

$$x h_m(x) = \sqrt{m+1} h_{m+1}(x) + \sqrt{m} h_{m-1}(x),$$

$$x^2 h_m(x) = \sqrt{(m+1)(m+2)} h_{m+2}(x) + (2m+1) h_m(x) + \sqrt{m(m-1)} h_{m-2}(x).$$

For these standard facts, see [Sze39, Section 5.5], with the change of normalization $\text{He}_n(x) = 2^{-n/2} H_n(x/\sqrt{2})$.

3. LOCALIZATION IN THE UNIVARIATE SETTING

Here we prove our main technical lemma, which provides the desired localization in the univariate setting.

Lemma 2. *Let ν be the standard Gaussian probability measure on $\mathbb{R}$. Fix $B > 0$ and $L \in \mathbb{N}$. Then there exists a constant $C = C(B, L) > 0$ such that, for every sufficiently large r and every $a \in \mathbb{R}$ satisfying $|a| \leq Br^{1/4}$, there exists $q \in \mathbb{R}[x]_r$ such that*

$$\int_{\mathbb{R}} q(x)^2 d\nu(x) = 1 \quad \text{and} \quad \int_{\mathbb{R}} |x - a|^{2\ell} q(x)^2 d\nu(x) \leq \frac{C}{r^\ell}, \quad 1 \leq \ell \leq L.$$

Proof. We construct a nonzero polynomial Q of degree at most r such that

$$\|(x - a)^\ell Q\|_{L^2(\nu)} \leq Cr^{-\ell/2} \|Q\|_{L^2(\nu)}, \quad 1 \leq \ell \leq L.$$

Using the Hermite recurrence, we achieve this by choosing the coefficients of Q with a slowly varying amplitude and phases that produce cancellation. Throughout the proof, $C > 0$ may change from line to line and depends only on B , L , and a fixed cutoff function chosen below.

Step 1: Construction of the coefficient sequence.

Let $\{h_k\}_{k \geq 0}$ be the normalized Hermite polynomials, see Section 2. Choose a nonzero real-valued function $\eta \in C_c^{L+2}((1/4, 3/4))$, and set $\eta_k = \eta(k/r)$. Choose constants

$$0 < c_0 < d_0 < d_1 < C_0 < 1$$

such that $\text{supp } \eta \subset [d_0, d_1]$. For sufficiently large r , all indices within distance $L + 2$ of the support of (η_k) lie in $[c_0r, C_0r]$. Thus these indices satisfy $k \asymp r$, with comparison constants independent of r and a . For indices in this interval, define

$$\phi_k = \arccos \left(\frac{a}{\sqrt{k} + \sqrt{k+1}} \right).$$

This is well defined for sufficiently large r , since $|a| \leq Br^{1/4}$ implies

$$\left| \frac{a}{\sqrt{k} + \sqrt{k+1}} \right| \leq Cr^{-1/4} \leq \frac{1}{2}.$$

Assign arbitrary real values to ϕ_k at the remaining nonnegative indices, and define

$$\Phi_0 = 0, \quad \Phi_{k+1} - \Phi_k = \phi_k.$$

Set

$$c_k = \eta_k e^{i\Phi_k}, \quad 0 \leq k \leq r, \quad Q(x) = \sum_{k=0}^r c_k h_k(x),$$

and extend the sequence by setting $c_k = 0$ for $k > r$. By orthonormality,

$$\|Q\|_{L^2(\nu)}^2 = \sum_{k=0}^r \eta_k^2.$$

Moreover, the Riemann-sum convergence

$$\frac{1}{r} \sum_{k=0}^r \eta(k/r)^2 \longrightarrow I_\eta := \int_0^1 \eta(t)^2 dt > 0$$

shows that, for sufficiently large r ,

$$(4) \quad c_\eta r \leq \|Q\|_{L^2(\nu)}^2 \leq C_\eta r,$$

where $c_\eta = I_\eta/2$ and $C_\eta = 2I_\eta$ are positive constants depending only on η , independent of r and a . We next show how the chosen phases produce cancellation under multiplication by $x - a$.

Step 2: Cancellation under multiplication by $x - a$.

For a finitely supported sequence u , define

$$(\mathcal{J}_a u)_j = \sqrt{j} u_{j-1} + \sqrt{j+1} u_{j+1} - a u_j, \quad u_{-1} = 0.$$

The Hermite recurrence gives

$$(x - a) \sum_{j \geq 0} u_j h_j(x) = \sum_{j \geq 0} (\mathcal{J}_a u)_j h_j(x).$$

Consequently,

$$(5) \quad \|(x - a)^\ell Q\|_{L^2(\nu)}^2 = \|\mathcal{J}_a^\ell c\|_{\ell^2}^2.$$

We now remove the oscillatory factor from the coefficients. Writing $u_j = e^{i\Phi_j} v_j$ and $\Delta v_j = v_{j+1} - v_j$, we obtain

$$(6) \quad e^{-i\Phi_j} (\mathcal{J}_a u)_j = b_j v_j - \gamma_j \Delta v_{j-1} + \beta_j \Delta v_j,$$

where

$$\gamma_0 = 0, \quad \gamma_j = \sqrt{j} e^{-i\phi_{j-1}} \quad (j \geq 1),$$

and

$$\beta_j = \sqrt{j+1} e^{i\phi_j}, \quad b_j = \gamma_j + \beta_j - a, \quad j \geq 0.$$

We use the convention $v_{-1} = 0$. The phase was chosen so that the leading terms in b_j cancel. We will use the estimates

$$(7) \quad \begin{aligned} |\Delta^s \gamma_j| + |\Delta^s \beta_j| &\leq C r^{1/2-s}, \quad 0 \leq s \leq L+1, \\ |\Delta^s b_j| &\leq C r^{-1/2-s}, \quad 0 \leq s \leq L. \end{aligned}$$

These estimates are needed only for indices in the fixed neighborhood of the support described above. To verify them, consider the smooth functions

$$\rho(t) = \frac{a}{\sqrt{t} + \sqrt{t+1}}, \quad \phi(t) = \arccos \rho(t),$$

on a slightly larger interval where $t \asymp r$. Differentiation gives

$$|\rho^{(s)}(t)| \leq C |a| r^{-1/2-s} \leq C r^{-1/4-s}, \quad 0 \leq s \leq L+2.$$

Since $|\rho(t)| \leq 1/2$, the chain rule yields

$$|\phi^{(s)}(t)| + \left| \frac{d^s}{dt^s} e^{\pm i\phi(t)} \right| \leq C r^{-1/4-s}, \quad 1 \leq s \leq L+2.$$

Thus the smooth extensions

$$\gamma(t) = \sqrt{t} e^{-i\phi(t-1)}, \quad \beta(t) = \sqrt{t+1} e^{i\phi(t)}$$

satisfy

$$|\gamma^{(s)}(t)| + |\beta^{(s)}(t)| \leq C r^{1/2-s}, \quad 0 \leq s \leq L+1.$$

For the cancellation term, the identity

$$(\sqrt{t} + \sqrt{t+1}) \cos \phi(t) = a$$

gives

$$b(t) := \gamma(t) + \beta(t) - a = \sqrt{t}(e^{-i\phi(t-1)} - e^{-i\phi(t)}) + i(\sqrt{t+1} - \sqrt{t}) \sin \phi(t).$$

The s th derivative of the first term is bounded by $Cr^{-3/4-s}$, using

$$e^{-i\phi(t-1)} - e^{-i\phi(t)} = - \int_0^1 \frac{d}{dt} e^{-i\phi(t-\tau)} d\tau.$$

The s th derivative of the second term is bounded by $Cr^{-1/2-s}$. Hence

$$|b^{(s)}(t)| \leq Cr^{-1/2-s}, \quad 0 \leq s \leq L.$$

The finite-difference bounds (7) now follow from

$$\Delta^s F(j) = \int_{[0,1]^s} F^{(s)}(j + t_1 + \cdots + t_s) dt_1 \cdots dt_s, \quad s \geq 1.$$

We next use these estimates to control the Hermite coefficients of $(x-a)^\ell Q$ for $1 \leq \ell \leq L$.

Step 3: Iterating the cancellation.

Each application of $\mathcal{J}_a$ enlarges the support by at most one index on either side, while Δ^s introduces shifts of at most s indices. Since $\ell + s \leq L + 1$ in the estimates below, all relevant indices lie in the neighborhood specified in Step 1. Thus the coefficient bounds from Step 2 apply throughout the induction.

We claim that

$$(\mathcal{J}_a^\ell c)_j = e^{i\Phi_j} v_j^{(\ell)}, \quad 0 \leq \ell \leq L,$$

where

$$(8) \quad |\Delta^s v_j^{(\ell)}| \leq Cr^{-\ell/2-s}, \quad 0 \leq s \leq L + 1 - \ell.$$

For $\ell = 0$, this follows from $v_j^{(0)} = \eta(j/r)$ and the smoothness of η . Suppose the claim holds for some $\ell < L$. By (6),

$$v_j^{(\ell+1)} = b_j v_j^{(\ell)} - \gamma_j \Delta v_{j-1}^{(\ell)} + \beta_j \Delta v_j^{(\ell)}.$$

Apply the discrete Leibniz rule

$$\Delta^s (fg)_j = \sum_{t=0}^s \binom{s}{t} (\Delta^t f)_{j+s-t} (\Delta^{s-t} g)_j.$$

For $0 \leq s \leq L - \ell$, each term arising from $b_j v_j^{(\ell)}$ is bounded by

$$Cr^{-1/2-t} r^{-\ell/2-(s-t)} = Cr^{-(\ell+1)/2-s}.$$

Likewise, each term arising from either difference term is bounded by

$$Cr^{1/2-t} r^{-\ell/2-(s-t+1)} = Cr^{-(\ell+1)/2-s}.$$

This proves (8) by induction. In particular, taking $s = 0$ gives

$$|(\mathcal{J}_a^\ell c)_j| \leq Cr^{-\ell/2}.$$

By the support observation above, $\mathcal{J}_a^\ell c$ has at most Cr nonzero entries, and hence

$$\|\mathcal{J}_a^\ell c\|_{\ell^2}^2 \leq Cr^{1-\ell}, \quad 1 \leq \ell \leq L.$$

Combining this with (4) and (5), we obtain

$$\frac{\int_{\mathbb{R}} |x - a|^{2\ell} |Q(x)|^2 d\nu(x)}{\int_{\mathbb{R}} |Q(x)|^2 d\nu(x)} \leq Cr^{-\ell}, \quad 1 \leq \ell \leq L.$$

It remains to obtain a real polynomial satisfying these bounds, up to a constant factor.

Step 4: Choosing a real polynomial.

Write $Q = Q_1 + iQ_2$, with $Q_1, Q_2 \in \mathbb{R}[x]_r$. Choose $Q_* \in \{Q_1, Q_2\}$ such that

$$\|Q_*\|_{L^2(\nu)}^2 \geq \frac{1}{2} \|Q\|_{L^2(\nu)}^2.$$

Since $Q_*(x)^2 \leq |Q(x)|^2$, the polynomial $q = \frac{Q_*}{\|Q_*\|_{L^2(\nu)}}$ satisfies

$$\int_{\mathbb{R}} q(x)^2 d\nu(x) = 1, \quad \int_{\mathbb{R}} |x - a|^{2\ell} q(x)^2 d\nu(x) \leq Cr^{-\ell}, \quad 1 \leq \ell \leq L.$$

This completes the proof. $\square$

4. PROOF OF THEOREM 1

Let $f \in \mathbb{R}[x]$ have degree at most N . Assume f attains its global infimum on $\mathbb{R}^d$ at $x^* = (a_1, \dots, a_d)$. We construct a sequence of degree- r polynomials q such that $\int_{\mathbb{R}^d} q(x)^2 d\mu(x) = 1$ and

$$\int_{\mathbb{R}^d} (f(x) - f_{\min}) q(x)^2 d\mu(x) \leq C_{f,d,N}/r$$

for some constant $C_{f,d,N}$.

To do so, we use Lemma 2 with $B = 1$ and $L = \lceil N/2 \rceil$. For each $j \in [d]$, it provides a sequence of degree- $\lfloor r/d \rfloor$ univariate polynomials q_j such that, for r large enough, we have

$$\int_{\mathbb{R}} q_j(x_j)^2 d\nu(x_j) = 1 \quad \text{and} \quad \int_{\mathbb{R}} |x_j - a_j|^{2\ell} q_j(x_j)^2 d\nu(x_j) \leq \frac{C}{r^\ell}, \quad 1 \leq \ell \leq L.$$

Here $C > 1$ is a constant and $d\nu$ is the univariate Gaussian measure. In particular, this bounds the even moments of $x_j - a_j$. We first extend this bound to odd moments. Let $1 \leq k = 2\ell - 1 \leq 2L$. Then $2\ell/k = 1 + 1/k > 1$ and therefore Hölder's inequality yields

$$\begin{aligned} \int_{\mathbb{R}} |x_j - a_j|^k q_j(x_j)^2 d\nu(x_j) &\leq \left(\int_{\mathbb{R}} |x_j - a_j|^{2\ell} q_j(x_j)^2 d\nu(x_j) \right)^{k/2\ell} \left(\int_{\mathbb{R}} q_j(x_j)^2 d\nu(x_j) \right)^{1-k/2\ell} \\ &\leq (Cr^{-\ell})^{k/2\ell} \leq Cr^{-k/2}. \end{aligned}$$

We then set

$$q(x) = \prod_{j \in [d]} q_j(x_j),$$

which is a degree- r polynomial since $d\lfloor r/d \rfloor \leq r$. Since the multivariate Gaussian measure equals the product of the univariate measures, we obtain

$$\int_{\mathbb{R}^d} q(x)^2 d\mu(x) = 1 \quad \text{and} \quad \int_{\mathbb{R}^d} |x - a|^\alpha q(x)^2 d\mu(x) \leq Cr^{-|\alpha|/2}, \quad 1 \leq |\alpha| \leq 2L.$$

To conclude the proof, it remains to observe that since x^* is a global minimizer of f , the Taylor expansion of $f - f_{\min}$ around x^* has no constant or linear terms:

$$f - f_{\min} = \sum_{\alpha \in \mathbb{N}^d: 2 \leq |\alpha| \leq N} c_{\alpha} (x - x^*)^{\alpha}.$$

We therefore have

$$\lambda_r - f_{\min} \leq \sum_{\alpha \in \mathbb{N}^d: 2 \leq |\alpha| \leq N} |c_{\alpha}| C r^{-|\alpha|/2} \leq \left(C \sum_{\alpha \in \mathbb{N}^d: 2 \leq |\alpha| \leq N} |c_{\alpha}| \right) r^{-1}$$

which is of the desired form for $C_{f,d,N} = C \sum_{\alpha \in \mathbb{N}^d: 2 \leq |\alpha| \leq N} |c_{\alpha}|$.

5. EXAMPLES

To provide context and examine the sharpness and limitations of Theorem 1, we first recall the dual SDP formulation of Lasserre's upper hierarchy (2). For this we need to introduce moment matrices and localized moment matrices. We refer to [Lau09, Nie23, HKL20] for background on moment and localizing matrices and their role in polynomial optimization.

Given a measure μ on $\mathbb{R}^d$ and $r \in \mathbb{N}$, the truncated commutative *moment matrix* $\mathbf{M}_r(\mu)$ associated with μ is the real symmetric matrix with rows and columns indexed by a basis (e.g. the canonical basis x^{α}) of $\mathbb{R}[x]_r$ and with entries:

$$\mathbf{M}_r(\mu)(\alpha, \beta) = \int_{\mathbb{R}^d} x^{\alpha+\beta} d\mu, \quad \alpha, \beta \in \mathbb{N}_r^n,$$

where $\mathbb{N}_r^n = \{\alpha \in (\mathbb{N} \cup \{0\})^n \mid \sum_i \alpha_i \leq r\}$. This matrix is a multivariate version of a (univariate) Hankel matrix.

Similarly, for all $f \in \mathbb{R}[x]$, the truncated *localizing moment matrix* $\mathbf{M}_r(f\mu)$ is the real symmetric matrix with rows and columns indexed by a basis (e.g., the canonical one) of $\mathbb{R}[x]_r$ and with entries

$$\mathbf{M}_r(f\mu)(\alpha, \beta) = \int_{\mathbb{R}^d} f(x) x^{\alpha+\beta} d\mu, \quad \alpha, \beta \in \mathbb{N}_r^n.$$

Note that $\mathbf{M}_r(1\mu) = \mathbf{M}_r(\mu)$.

With this notation introduced, the dual SDP to (2) is

$$(9) \quad \begin{aligned} \lambda_r = \sup_{\lambda \in \mathbb{R}} \quad & \lambda \\ \text{s.t.} \quad & \mathbf{M}_r(f\mu) \succeq \lambda \mathbf{M}_r(\mu), \end{aligned}$$

and is a generalized eigenvalue problem. Further, if one considers the two matrices $\mathbf{M}_r(f\mu)$ and $\mathbf{M}_r(\mu)$ indexed by an orthonormal basis of $\mathbb{R}[x]_r$ w.r.t. μ , then $\mathbf{M}_r(\mu) = I$ and (9) is in fact equivalent to an eigenvalue problem.

Connections between polynomial optimization, moment matrices, Jacobi matrices, and extremal zeros of orthogonal polynomials have been exploited in [dKL20] in the context of Lasserre's measure-based upper bounds.

5.1. An example with convergence rate $\frac{1}{r}$.

Example 3. Let $d = 1$ and consider the Gaussian probability measure on $\mathbb{R}$,

$$d\mu(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

We will consider (9) for $f = x^2$ with respect to the orthonormal basis $\{h_n\}_{n \in \mathbb{N}_0}$ of normalized Hermite polynomials (see Section 2) so that

$$\mathbf{M}_r(\mu) = I.$$

Using the three-term recurrence relation of the Hermite polynomials (3), the localizing matrix $\mathbf{M}_r(x^2 \mu)$ with rows and columns indexed by $0, \dots, r$ is easily seen to be

$$\mathbf{M}_r(x^2 \mu)(n, m) = \begin{cases} 2m + 1 & \text{if } n = m \\ \sqrt{(m+1)(m+2)} & \text{if } n = m + 2 \\ \sqrt{m(m-1)} & \text{if } n = m - 2 \\ 0 & \text{otherwise} \end{cases},$$

for $n, m = 0, 1, \dots, r$; that is,

$$\mathbf{M}_r(x^2 \mu) = \begin{pmatrix} 1 & 0 & \sqrt{2} & 0 & 0 & \dots & 0 \\ 0 & 3 & 0 & \sqrt{6} & 0 & \dots & 0 \\ \sqrt{2} & 0 & 5 & 0 & \sqrt{12} & \dots & 0 \\ 0 & \sqrt{6} & 0 & 7 & 0 & \ddots & \vdots \\ 0 & 0 & \sqrt{12} & 0 & 9 & \ddots & \sqrt{r(r-1)} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \dots & \sqrt{r(r-1)} & 0 & 2r+1 \end{pmatrix}$$

The solution λ_r to (2) (or, equivalently, (9)) is the smallest eigenvalue of $\mathbf{M}_r(x^2 \mu)$.

In the spectral theory of orthogonal polynomials, the infinite matrix representing the operator of multiplication by x is called the Jacobi matrix J (see, e.g., [Sim15, Chapter 4]). For the normalized Hermite polynomials, using (3), J is a tridiagonal matrix, whose only non-zero entries are

$$J_{m,m+1} = J_{m+1,m} = \sqrt{m+1} \quad \text{for } 0 \leq m.$$

Since our matrix represents multiplication by x^2 , the infinite-dimensional version of $\mathbf{M}_r(x^2 \mu)$ is exactly the squared Jacobi matrix J^2 . Let J_r denote the principal $(r+1) \times (r+1)$ submatrix of J indexed by $0, \dots, r$. Let $\mathbf{e}_j \in \mathbb{R}^{r+1}$, $0 \leq j \leq r$, denote the corresponding coordinate vectors. Evaluating $(J_r^2)_{n,m} = \sum_{k=0}^r (J_r)_{n,k} (J_r)_{k,m}$ and comparing this to $\mathbf{M}_r(x^2 \mu)$, we obtain

$$\mathbf{M}_r(x^2 \mu) = J_r^2 + (r+1)E_{r,r},$$

where $E_{r,r} = \mathbf{e}_r \mathbf{e}_r^T$ is the rank-one projection onto the boundary coordinate.

The eigenvalues of J_r correspond precisely to the roots γ of the $(r+1)$ -th Hermite polynomial (see, e.g., [Gau04, Theorem 1.32]), which are distributed symmetrically around zero. Squaring these roots yields the eigenvalues of J_r^2 , whence every non-zero eigenvalue possesses a multiplicity of exactly 2. To analyze the asymptotics of the minimal eigenvalue

$$\lambda_r = \lambda_{\min}(\mathbf{M}_r(x^2 \mu)),$$

it suffices to consider odd r as the sequence λ_r is monotone by definition. Because $\mathbf{M}_r(x^2 \mu)$ is a rank-one positive-semidefinite perturbation of J_r^2 , its eigenvalues interlace with those of J_r^2 via Cauchy's interlacing theorem [HJ12, Theorem 4.3.1].

Since r is odd, the polynomial $\text{He}_{r+1}(x)$ contains no root at the origin. Thus, the minimum eigenvalue of J_r^2 is $\gamma_{\min}^2 > 0$ with multiplicity 2. Because the positive perturbation $(r+1)E_{r,r}$ has rank one, it cannot increase the minimal eigenvalue of J_r^2 . Hence

$$\lambda_r = \gamma_{\min}^2.$$

By classical asymptotic limits for the zeros of Hermite polynomials, the smallest positive zero scales as $\gamma_{\min} = \Theta(1/\sqrt{r})$ as $r \rightarrow \infty$ (e.g., [Sze39, Theorem 6.32]). Squaring this relation delivers the desired asymptotic rate:

$$\lambda_r = \Theta\left(\frac{1}{r}\right).$$

This shows that even for a simple univariate quadratic objective function, the Lasserre upper bound hierarchy (2) yields a sharp convergence rate of $\frac{1}{r}$.

5.2. A family with arbitrarily slow polynomial convergence rates. If one drops the assumption of f having a minimizer (i.e., if f is merely bounded from below without attaining its infimum), then the conclusion of Theorem 1 may fail to hold.

Example 4. For an integer $m \geq 0$, consider the polynomial

$$f_m(x, y) = x^2 + (1 + y^2)^m(xy - 1)^2.$$

One can verify that $f_m(x, y) > 0$ for all $x, y \in \mathbb{R}$ and for $\varepsilon > 0$ we have $f_m(\varepsilon, 1/\varepsilon) = \varepsilon^2$. Hence, the infimum of f_m is equal to 0, but it is not attained. We show

$$(10) \quad \lambda_r^{(m)} = \Theta\left(r^{-1/(m+2)}\right), \quad r \rightarrow \infty.$$

To establish the rate (10) we require two lemmas.

Lemma 5. For $r \geq 0$ and $a \in \mathbb{R}$, define

$$\theta_r(a) = \inf_{\substack{q \in \mathbb{R}[x]_r \\ q \neq 0}} \frac{\int_{\mathbb{R}} (x-a)^2 q(x)^2 d\nu(x)}{\int_{\mathbb{R}} q(x)^2 d\nu(x)}.$$

Then, for all $a \in \mathbb{R}$, we have

$$\theta_r(a) \geq \frac{1}{4r+2}.$$

Proof. Let $q \in \mathbb{R}[x]_r$ and define

$$\psi(x) = (2\pi)^{-1/4} q(x) e^{-x^2/4}.$$

Then

$$\|\psi\|_{L^2(dx)}^2 = \int_{\mathbb{R}} q(x)^2 d\nu(x),$$

and

$$\|(x-a)\psi\|_{L^2(dx)}^2 = \int_{\mathbb{R}} (x-a)^2 q(x)^2 d\nu(x).$$

Let

$$\psi_k(x) = (2\pi)^{-1/4} h_k(x) e^{-x^2/4}.$$

The functions ψ_k are the normalized Hermite functions. They are eigenfunctions of the harmonic oscillator Hamiltonian $H = -\frac{d^2}{dx^2} + \frac{x^2}{4}$, satisfying

$$\left(-\frac{d^2}{dx^2} + \frac{x^2}{4}\right) \psi_k = \left(k + \frac{1}{2}\right) \psi_k.$$

Since q has degree at most r , ψ belongs to $\text{span}\{\psi_0, \dots, \psi_r\}$. Hence

$$\|\psi'\|_2^2 + \frac{1}{4}\|x\psi\|_2^2 = \langle H\psi, \psi \rangle \leq \left(r + \frac{1}{2}\right) \|\psi\|_2^2.$$

In particular,

$$(11) \quad \|\psi'\|_2^2 \leq \left(r + \frac{1}{2}\right) \|\psi\|_2^2.$$

On the other hand, integration by parts gives

$$\|\psi\|_2^2 = -2 \int_{\mathbb{R}} (x - a) \psi'(x) \psi(x) dx.$$

Thus, by the Cauchy–Schwarz inequality,

$$\|\psi\|_2^2 \leq 2\|(x - a)\psi\|_2 \|\psi'\|_2.$$

Combining this estimate with (11), we get

$$\|(x - a)\psi\|_2^2 \geq \frac{\|\psi\|_2^4}{4\|\psi'\|_2^2} \geq \frac{1}{4r + 2} \|\psi\|_2^2.$$

Returning to q , this says

$$\int_{\mathbb{R}} (x - a)^2 q(x)^2 d\nu(x) \geq \frac{1}{4r + 2} \int_{\mathbb{R}} q(x)^2 d\nu(x).$$

Taking the infimum over $q \in \mathbb{R}[x]_r$ proves the claim. $\square$

Lemma 6. Fix $m \in \mathbb{N}_0$ and $B > 0$. Then there exists a constant $C = C(B, m) > 0$ such that, for every sufficiently large r and every $a \in \mathbb{R}$ satisfying

$$|a| \leq Br^{1/4},$$

there exists a polynomial $q \in \mathbb{R}[x]_r$ with $\int_{\mathbb{R}} q(x)^2 d\nu(x) = 1$, such that

$$\begin{aligned} \int_{\mathbb{R}} (x - a)^2 q(x)^2 d\nu(x) &\leq \frac{C}{r}, \\ \int_{\mathbb{R}} (1 + x^2)^{m+1} q(x)^2 d\nu(x) &\leq C(1 + a^2)^{m+1}, \\ \int_{\mathbb{R}} (1 + x^2)^m (x - a)^2 q(x)^2 d\nu(x) &\leq \frac{C}{r} (1 + a^2)^m. \end{aligned}$$

Proof. Apply Lemma 2 with $L = m + 1$. Then

$$\int_{\mathbb{R}} |x - a|^{2j} q(x)^2 d\nu(x) \leq \frac{C}{r^j}, \quad 1 \leq j \leq m + 1.$$

Since $1 + x^2 \leq 2(1 + a^2) + 2(x - a)^2$, we obtain

$$(1 + x^2)^{m+1} \leq C_m \left((1 + a^2)^{m+1} + |x - a|^{2m+2} \right).$$

Therefore,

$$\int_{\mathbb{R}} (1 + x^2)^{m+1} q(x)^2 d\nu(x) \leq C(1 + a^2)^{m+1}.$$

Similarly,

$$(1 + x^2)^m (x - a)^2 \leq C_m \left((1 + a^2)^m (x - a)^2 + |x - a|^{2m+2} \right).$$

Hence

$$\int_{\mathbb{R}} (1 + x^2)^m (x - a)^2 q(x)^2 d\nu(x) \leq \frac{C}{r} (1 + a^2)^m + \frac{C}{r^{m+1}} \leq \frac{C}{r} (1 + a^2)^m.$$

The estimate for $\int (x - a)^2 q^2 d\nu$ is the case $j = 1$. □

Proof of the rate in Example 4. We first prove the lower bound of (10). Set

$$A(y) = (1 + y^2)^m, \quad D(y) = 1 + y^2 A(y).$$

Completing the square in x , we obtain

$$f_m(x, y) = D(y) \left(x - \frac{yA(y)}{D(y)} \right)^2 + \frac{A(y)}{D(y)}.$$

Hence, by the lower estimate of Lemma 5, for every $h \in \mathbb{R}[x, y]_r$,

$$\int_{\mathbb{R}} f_m(x, y) h(x, y)^2 d\nu(x) \geq \left(\frac{D(y)}{4r + 2} + \frac{A(y)}{D(y)} \right) \int_{\mathbb{R}} h(x, y)^2 d\nu(x).$$

Let $R = 4r + 2$ and $z = 1 + y^2$. Then $A(y) = z^m$ and $D(y) = 1 + (z - 1)z^m$. If $z \leq R^{1/(m+2)}$, then

$$\frac{A(y)}{D(y)} \geq \frac{1}{2z} \geq \frac{1}{2} R^{-1/(m+2)}.$$

If $z \geq R^{1/(m+2)}$, then, for sufficiently large R ,

$$\frac{D(y)}{R} \geq \frac{(z - 1)z^m}{R} \geq \frac{z^{m+1}}{2R} \geq \frac{1}{2} R^{-1/(m+2)}.$$

Therefore,

$$(12) \quad \lambda_r^{(m)} \geq c_m r^{-1/(m+2)}.$$

For the upper bound, let

$$n = \left\lfloor \frac{r}{2} \right\rfloor, \quad Y = n^{1/(2m+4)}, \quad a = Y^{-1}.$$

Then $a = n^{-1/(2m+4)} \leq 1$ and $Y = n^{1/(2m+4)} \leq n^{1/4}$. Lemma 6, applied with $r = n$ first at the center a and then at the center Y , yields polynomials $q \in \mathbb{R}[x]_n$ and $v \in \mathbb{R}[y]_n$ such that

$$(13) \quad \begin{aligned} \int_{\mathbb{R}} q(x)^2 d\nu(x) &= \int_{\mathbb{R}} v(y)^2 d\nu(y) = 1, \\ \int_{\mathbb{R}} (x - a)^2 q(x)^2 d\nu(x) &\leq \frac{C}{n}, \\ \int_{\mathbb{R}} (1 + y^2)^{m+1} v(y)^2 d\nu(y) &\leq C_m (1 + Y^2)^{m+1}, \\ \int_{\mathbb{R}} (1 + y^2)^m (y - Y)^2 v(y)^2 d\nu(y) &\leq \frac{C_m}{n} (1 + Y^2)^m. \end{aligned}$$

Set $h(x, y) = q(x)v(y)$. Then $\deg h \leq 2n \leq r$ and $\int_{\mathbb{R}^2} h(x, y)^2 d\nu(x) d\nu(y) = 1$. Moreover, from $x^2 \leq 2a^2 + 2(x - a)^2$, we deduce

$$\int_{\mathbb{R}^2} x^2 h^2 d\nu(x) d\nu(y) \leq 2a^2 + \frac{C}{n}.$$

Since $aY = 1$, we have $xy - 1 = y(x - a) + a(y - Y)$, and therefore

$$(xy - 1)^2 \leq 2y^2(x - a)^2 + 2a^2(y - Y)^2.$$

Using Fubini's theorem together with the estimates (13) above, we obtain

$$\int_{\mathbb{R}^2} (1 + y^2)^m (xy - 1)^2 h^2 d\nu(x) d\nu(y) \leq \frac{C_m}{n} (1 + Y^2)^{m+1} + \frac{C_m a^2}{n} (1 + Y^2)^m.$$

Consequently,

$$\lambda_r^{(m)} \leq 2a^2 + \frac{C}{n} + \frac{C_m}{n} (1 + Y^2)^{m+1} + \frac{C_m a^2}{n} (1 + Y^2)^m.$$

Since $Y = n^{1/(2m+4)}$ and $a = Y^{-1}$, each term on the right-hand side is bounded by $C_m n^{-1/(m+2)}$. Thus

$$(14) \quad \lambda_r^{(m)} \leq C_m r^{-1/(m+2)}.$$

Combining the (12) and (14) gives

$$\lambda_r^{(m)} = \Theta(r^{-1/(m+2)}). \quad \square$$

6. CONCLUSION AND DISCUSSION

We have established an explicit convergence rate for Lasserre's measure-based upper-bound hierarchy on the noncompact domain $\mathbb{R}^d$ when the reference measure is Gaussian. More precisely, if $f \in \mathbb{R}[x]$ attains its global infimum, then

$$\lambda_r - f_{\min} = O\left(\frac{1}{r}\right).$$

To the best of our knowledge, this is the first quantitative convergence rate for this hierarchy in a noncompact setting.

The example $f = x^2$ shows that the upper bound cannot be improved for even the simplest polynomial objective:

in that case,

$$\lambda_r = \Theta(1/r).$$

The family of polynomials in Section 5 shows, on the other hand, that attainment of the infimum is a necessary assumption. If the infimum is not attained, the convergence may occur at rates

$$\Theta(r^{-1/(m+2)}), \quad m \geq 0,$$

and hence may be arbitrarily slow among polynomial rates. Consequently, some form of control on the location of near-minimizers is indispensable for uniform quantitative results of the type proved in Theorem 1. It would be interesting to identify conditions weaker than attainment that still guarantee a prescribed convergence rate.

6.1. On the use of LLMs. We acknowledge the use of ChatGPT Astra during the derivation of our main bounds. While our initial argument, preserved in Appendix A, was formulated without automated assistance, discussions with Astra prompted us to recognize that our own Lemma 2 could be leveraged more broadly to remove the log-factor. The final proof structure presented in the main text was rigorously checked and written by the authors.

REFERENCES

- [AS64] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, 1964.
- [BM23] L. Baldi and B. Mourrain, On the effective Putinar’s Positivstellensatz and moment approximation, *Math. Program.* 200 (2023), 71–103.
- [BMP25] L. Baldi, B. Mourrain, and A. Parusiński, On Łojasiewicz inequalities and the effective Putinar’s Positivstellensatz, *J. Algebra* 662 (2025), 741–767.
- [BS24] L. Baldi and L. Slot, Degree bounds for Putinar’s Positivstellensatz on the hypercube, *SIAM J. Appl. Algebra Geom.* 8 (2024), 1–25.
- [BCR98] J. Bochnak, M. Coste, and M.-F. Roy, *Real Algebraic Geometry*, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, 3. Folge, vol. 36, Springer, 1998.
- [DW12] A. Doherty and S. Wehner, Convergence of SDP hierarchies for polynomial optimization on the hypersphere, preprint arXiv:1210.5048, 2012.
- [FF21] K. Fang and H. Fawzi, The sum-of-squares hierarchy on the sphere and applications in quantum information theory, *Math. Program.* 190 (2021), 331–360.
- [Gau04] W. Gautschi, *Orthogonal Polynomials: Computation and Approximation*, Oxford University Press, 2004.
- [GdKV26] S. Gribling, E. de Klerk, and J. Vera, Revisiting the convergence rate of the Lasserre hierarchy for polynomial optimization over the hypercube, *Optimization* (2026), 1–28, published online. doi: 10.1080/02331934.2026.2641624.
- [HKL20] D. Henrion, M. Korda, and J.-B. Lasserre, *The Moment-SOS Hierarchy: Lectures in Probability, Statistics, Computational Geometry, Control and Nonlinear PDEs*, Series on Optimization and its Applications, vol. 4, World Scientific, 2020.
- [HLM25] D. Henrion, A. Le Franc, and V. Magron, Slow convergence of the moment-SOS hierarchy for an elementary polynomial optimization problem, *SIAM J. Appl. Algebra Geom.* 9 (2025), 261–278.
- [HJ12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed., Cambridge University Press, 2012.
- [KMV26] I. Klep, V. Magron, and J. Volčič, Complete upper bound hierarchies for spectral minimum in noncommutative polynomial optimization, *J. Math. Anal. Appl.* 563 (2026), article 130794.
- [dKHL17] E. de Klerk, R. Hess, and M. Laurent, Improved convergence rates for Lasserre-type hierarchies of upper bounds for box-constrained polynomial optimization, *SIAM J. Optim.* 27 (2017), 347–367.
- [dKL20] E. de Klerk and M. Laurent, Worst-case examples for Lasserre’s measure-based hierarchy for polynomial optimization on the hypercube, *Math. Oper. Res.* 45 (2020), 86–98.
- [dKL22] E. de Klerk and M. Laurent, Convergence analysis of a Lasserre hierarchy of upper bounds for polynomial minimization on the sphere, *Math. Program.* 193 (2022), 665–685.
- [dKLS17] E. de Klerk, M. Laurent, and Z. Sun, Convergence analysis for Lasserre’s measure-based hierarchy of upper bounds for polynomial optimization, *Math. Program.* 162 (2017), 363–392.
- [KMR25] M. Korda, V. Magron, and R. Ríos-Zertuche, Convergence rates for sums-of-squares hierarchies with correlative sparsity, *Math. Program.* 209 (2025), 435–473.
- [Las01] J. B. Lasserre, Global optimization with polynomials and the problem of moments, *SIAM J. Optim.* 11 (2001), 796–817.
- [Las10] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press Optimization Series, vol. 1, Imperial College Press, 2010.
- [Las11] J. B. Lasserre, A new look at nonnegativity on closed sets and polynomial optimization, *SIAM J. Optim.* 21 (2011), 864–885.
- [Lau09] M. Laurent, Sums of squares, moment matrices and optimization over polynomials, in *Emerging Applications of Algebraic Geometry*, M. Putinar and S. Sullivant (eds.), The IMA Volumes in Mathematics and its Applications, vol. 149, Springer, 2009, pp. 157–270.
- [LS26] M. Laurent and L. Slot, An overview of convergence rates for sum of squares hierarchies in polynomial optimization, in *Recent Developments in Industrial and Applied Mathematics*, ICIAM2023 Springer Series, vol. 1, Springer, 2026, pp. 149–176.
- [Mag26] V. Magron, Convergence rates for polynomial optimization on set products, *SIAM J. Optim.* 36 (2026), 1867–1886.

- [Mar08] M. Marshall, *Positive Polynomials and Sums of Squares*, Mathematical Surveys and Monographs, vol. 146, American Mathematical Society, 2008.
- [MH03] J. C. Mason and D. C. Handscomb, *Chebyshev Polynomials*, Chapman & Hall/CRC, 2003.
- [Nie23] J. Nie, *Moment and Polynomial Optimization*, MOS-SIAM Series on Optimization, vol. 31, Society for Industrial and Applied Mathematics, 2023.
- [Par03] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Math. Program.* 96 (2003), 293–320.
- [Put93] M. Putinar, Positive polynomials on compact semi-algebraic sets, *Indiana Univ. Math. J.* 42 (1993), 969–984.
- [Riv74] T. J. Rivlin, *The Chebyshev Polynomials*, Pure and Applied Mathematics, Wiley-Interscience, 1974.
- [SV14] S. Sachdeva and N. K. Vishnoi, Faster algorithms via approximation theory, *Foundations and Trends in Theoretical Computer Science* 9 (2014), 125–210.
- [Sce24] C. Scheiderer, *A Course in Real Algebraic Geometry: Positivity and Sums of Squares*, Graduate Texts in Mathematics, vol. 303, Springer, 2024.
- [Scm17] K. Schmüdgen, *The Moment Problem*, Graduate Texts in Mathematics, vol. 277, Springer, 2017.
- [Scw05] M. Schweighofer, Optimization of polynomials on compact semialgebraic sets, *SIAM J. Optim.* 15 (2005), 805–825.
- [Sim15] B. Simon, *Operator Theory: A Comprehensive Course in Analysis, Part 4*, American Mathematical Society, Providence, RI, 2015.
- [SL22] L. Slot and M. Laurent, Improved convergence analysis of Lasserre's measure-based upper bounds for polynomial minimization on compact sets, *Math. Program.* 193 (2022), 831–871.
- [Sze39] G. Szegő, *Orthogonal Polynomials*, American Mathematical Society, 1939.

APPENDIX A. GAUSSIAN LOCALIZATION APPROACH

Here we give an alternative, perhaps more intuitive, argument that yields an upper bound of $O(\frac{\log r}{r})$ for the upper bound hierarchy with respect to the standard Gaussian measure. In a nutshell, we introduce a Gaussian localization centered at a global minimizer x^* of f . We then approximate the Gaussian localization by a sum-of-squares density. Our proof uses some results from approximation theory and Gaussian tail bounds, which we state first.

For a polynomial $p(x) = \sum_{\mu} p_{\mu} x^{\mu}$, we define the ℓ^1 -norm of its coefficients by

$$\|p\|_1 := \sum_{\mu} |p_{\mu}|.$$

We use the following polynomial approximation result for the exponential function to construct the “polynomialized” Gaussian.

Theorem 7 ([SV14, Theorem 4.1]). *For every $b > 0$ and $0 < \delta \leq 1$, there exists a polynomial $S_{b,\delta}$ that satisfies*

$$\sup_{t \in [0,b]} |e^{-t} - S_{b,\delta}(t)| \leq \delta,$$

and has degree $O(\sqrt{\max\{b, \log(1/\delta)\}} \cdot \log(1/\delta))$.

We need the following elementary growth estimate, see, e.g. [Riv74, Section 2.7, (2.37)].

Lemma 8. *Let $p \in \mathbb{R}[u]$ have degree at most s , and suppose that*

$$\|p\|_{L^{\infty}[0,1]} \leq B.$$

Then, for every $u \geq 1$,

$$|p(u)| \leq B T_s(2u - 1) \leq B(4u)^s,$$

where T_s is the Chebyshev polynomial of the first kind; see [MH03].

We now state the two tail estimates used in this argument. The first controls polynomial moments of a rescaled Gaussian, while the second controls the polynomial approximation outside its localization region. This estimate follows by switching to polar coordinates and using the asymptotics of the incomplete gamma function [AS64, Eq. (6.5.32)].

Lemma 9. *Let $\mathbf{a} \in \mathbb{Z}_+^d$. There exists a constant $C_{d,|\mathbf{a}|} > 0$ such that, whenever $\beta R^2 \geq 1$,*

$$(15) \quad \int_{\|x\| > R} |x^{\mathbf{a}}| e^{-\frac{\beta}{2}\|x\|^2} dx \leq C_{d,|\mathbf{a}|} \beta^{-1} R^{d+|\mathbf{a}|-2} e^{-\frac{\beta R^2}{2}}.$$

Consequently, if $|\mathbf{a}| \leq N$ and $R \geq 1$, then

$$\int_{\|x\| > R} |x^{\mathbf{a}}| e^{-\frac{\beta}{2}\|x\|^2} dx \leq C_{d,N} \beta^{-1} R^{d+N-2} e^{-\frac{\beta R^2}{2}}.$$

Using the Chebyshev growth estimate we obtain exponential tail control for the resulting polynomialized Gaussian, uniformly over bounded shifts.

Lemma 10. *Let $g \in \mathbb{R}[x]$ have degree at most N , and let $B > 0$. Suppose*

$$q(x) = p\left(\frac{\|x - a\|^2}{R^2}\right), \quad \deg p = s, \quad \|p\|_{L^\infty[0,1]} \leq 2,$$

where $\|a\| \leq B$. Then there are constants $K \geq 1$ and $c, C > 0$, depending only on d, N, B , such that whenever $R^2 \geq K(s+1)$,

$$\int_{\|x-a\| > R} |g(x)| q(x)^2 d\mu(x) \leq C \|g\|_1 e^{-cR^2},$$

where μ denotes the standard Gaussian probability measure.

Proof. Put $y = x - a$ and $r = \|y\|$. By squaring the triangle inequality $\|y\| \leq \|y+a\| + \|a\|$, we have $\|y+a\|^2 \geq \frac{1}{2}\|y\|^2 - \|a\|^2$, whence

$$e^{-\|y+a\|^2/2} \leq e^{B^2/2} e^{-\|y\|^2/4}$$

since $\|a\| \leq B$. We also have an upper bound on g in terms of the coefficient norm:

$$|g(y+a)| \leq A_{d,N,B} \|g\|_1 (1+r)^N,$$

for a suitable constant $A_{d,N,B}$.

For $r > R$, Lemma 8 and the inequality $\log(4u) \leq u \log 4$ for $u \geq 1$ give

$$q(y+a)^2 \leq 4(4r^2/R^2)^{2s} \leq 4 \exp\left(\frac{2s \log 4}{R^2} r^2\right) \leq 4e^{r^2/8},$$

provided $R^2 \geq 16(\log 4)(s+1)$. Consequently,

$$\begin{aligned} \int_{\|x-a\| > R} |g(x)| q(x)^2 d\mu(x) &\leq C_{d,N,B} \|g\|_1 \int_{\|y\| > R} (1+\|y\|)^N e^{-\|y\|^2/8} dy \\ &\leq C_{d,N,B} \|g\|_1 e^{-R^2/16} \int_{\mathbb{R}^d} (1+\|y\|)^N e^{-\|y\|^2/16} dy. \end{aligned}$$

The last integral is finite and depends only on d, N . Absorbing it into the constant proves the claim with $K = 16 \log 4$ and $c = 1/16$. $\square$

A.1. Approximating localized Gaussians with polynomial densities. Let $x^* \in \mathbb{R}^d$ be a global minimizer of f , and write $f_{\min} = f(x^*)$. By definition of the r -th Lasserre upper bound λ_r in (2), for every nonzero test polynomial $q \in \mathbb{R}[x]$ with $\deg(q) \leq r$,

$$(16) \quad 0 \leq \lambda_r - f_{\min} \leq \frac{\int_{\mathbb{R}^d} (f(x) - f_{\min}) q(x)^2 d\mu(x)}{\int_{\mathbb{R}^d} q(x)^2 d\mu(x)}.$$

The argument proceeds by constructing a sequence of sum-of-squares weights $q(x)^2$ that approximate a narrow Gaussian kernel centered near x^* , and bounding the numerator and denominator in (16). The narrow Gaussian kernel that we use is as follows. Fix $\alpha > 0$, and define

$$\gamma = 1 + \alpha^2, \quad \beta = 1 + \frac{1}{\alpha^2}.$$

We then consider the Gaussian localization integral

$$(17) \quad \begin{aligned} I_\alpha(f) &= \int_{\mathbb{R}^d} f(x) \exp\left(-\frac{\|x - \gamma x^*\|^2}{2\alpha^2} - \frac{\|x\|^2}{2}\right) dx. \\ &= e^{-\frac{\gamma}{2}\|x^*\|^2} \int_{\mathbb{R}^d} f(x) e^{-\frac{\beta}{2}\|x - x^*\|^2} dx. \\ &= e^{-\frac{\gamma}{2}\|x^*\|^2} \int_{\mathbb{R}^d} f(x + x^*) e^{-\frac{\beta}{2}\|x\|^2} dx. \end{aligned}$$

Here, the first equality uses the following completion-of-squares identity

$$(18) \quad \frac{\|x - \gamma x^*\|^2}{2\alpha^2} + \frac{\|x\|^2}{2} = \frac{\beta}{2}\|x - x^*\|^2 + \frac{\gamma}{2}\|x^*\|^2,$$

and the second equality uses the change of variables $x \mapsto x + x^*$. One can show that, as $\alpha \rightarrow 0$, we have $I_\alpha(f)/I_\alpha(1) \rightarrow f(x^*)$.

Step 1: Construction and feasibility of the test polynomial.

Fix constants $A > \frac{d}{2} + 1$ and $K \geq 1$ (to be chosen sufficiently large). For $r \in \mathbb{N}$, set

$$\delta = r^{-A}, \quad L = \log(1/\delta) = A \log r, \quad R = \sqrt{K}r,$$

and define the parameters

$$\alpha = C_{\deg} \frac{R\sqrt{L}}{r} = C_{\deg} \sqrt{AK \frac{\log r}{r}}, \quad \gamma = 1 + \alpha^2, \quad \beta = 1 + \frac{1}{\alpha^2},$$

where $C_{\deg} > 0$ is a constant to be fixed shortly. Note that $\alpha \rightarrow 0$ and $\beta = 1 + \alpha^{-2} = \Theta(r/\log r)$ as $r \rightarrow \infty$. Consider the target Gaussian localization profile

$$G_\alpha(x) = \exp\left(-\frac{\beta - 1}{4}\|x - \gamma x^*\|^2\right).$$

On the Euclidean ball $B(\gamma x^*, R)$, the exponent satisfies $\frac{\beta - 1}{4}\|x - \gamma x^*\|^2 \in [0, b]$, where

$$b = \frac{\beta - 1}{4} R^2 = \frac{R^2}{4\alpha^2} = \frac{r^2}{4C_{\deg}^2 L}.$$

For all sufficiently large r , we have $b \geq L$. Applying Theorem 7 with interval $[0, b]$ and precision δ , there exists a univariate polynomial $S_{b,\delta}$ such that

$$\sup_{t \in [0, b]} |e^{-t} - S_{b,\delta}(t)| \leq \delta,$$

with degree bounded by $s = \deg(S_{b,\delta}) \leq C\sqrt{bL} = \frac{C}{2C_{\deg}}r$. We define our test polynomial by

$$q(x) = S_{b,\delta} \left(\frac{\beta - 1}{4} \|x - \gamma x^*\|^2 \right).$$

Choosing $C_{\deg} \geq C$ ensures that $\deg(q) = 2s \leq r$, so $q \in \mathbb{R}[x]_r$ is feasible for the r -th relaxation. Moreover, $|q(\gamma x^*) - 1| = |S_{b,\delta}(0) - 1| \leq \delta < 1$ for large r , so $q \neq 0$.

Step 2: Comparison between $q^2 d\mu$ and the localized Gaussian $G_\alpha^2 d\mu$.

On the ball $B(\gamma x^*, R)$, we have $|G_\alpha(x) - q(x)| \leq \delta$. Since $0 < G_\alpha(x) \leq 1$, it follows that $|q(x)| \leq 1 + \delta \leq 2$, and hence

$$|q(x)^2 - G_\alpha(x)^2| \leq |q(x) - G_\alpha(x)| \cdot |q(x) + G_\alpha(x)| \leq 3\delta.$$

Therefore, for any polynomial $g \in \mathbb{R}[x]$ of degree at most N ,

$$\begin{aligned} \left| \int_{\mathbb{R}^d} g q^2 d\mu - \int_{\mathbb{R}^d} g G_\alpha^2 d\mu \right| &\leq 3\delta \int_{\|x - \gamma x^*\| \leq R} |g(x)| d\mu(x) \\ &\quad + \int_{\|x - \gamma x^*\| > R} |g(x)| q(x)^2 d\mu(x) \\ &\quad + \int_{\|x - \gamma x^*\| > R} |g(x)| G_\alpha(x)^2 d\mu(x). \end{aligned} \tag{19}$$

We bound the three terms on the right-hand side of (19) as follows:

- (i) *Ball approximation term:* Since polynomials are integrable against the Gaussian measure,

$$3\delta \int_{\|x - \gamma x^*\| \leq R} |g(x)| d\mu(x) \leq 3\delta \|g\|_{L^1(\mu)} \leq C_{d,N} \|g\|_1 \delta.$$

- (ii) *Polynomialized tail:* Set $p(t) = S_{b,\delta}(bt)$ for $t \in [0, 1]$. Then $\|p\|_{L^\infty([0,1])} \leq 1 + \delta \leq 2$ and $q(x) = p(\|x - \gamma x^*\|^2 / R^2)$. Since $\gamma = 1 + \alpha^2 \rightarrow 1$, the shift vector satisfies $\|\gamma x^*\| \leq 2\|x^*\|$. By choosing $K \geq 1$ sufficiently large so that $R^2 = Kr \geq K'(s+1)$, Lemma 10 implies

$$\int_{\|x - \gamma x^*\| > R} |g(x)| q(x)^2 d\mu(x) \leq C_{d,N,x^*} \|g\|_1 e^{-cR^2}.$$

- (iii) *Gaussian tail:* By Equation (18), the Gaussian weight G_α satisfies

$$G_\alpha(x)^2 d\mu(x) = (2\pi)^{-d/2} e^{-\frac{\gamma}{2}\|x^*\|^2} e^{-\frac{\beta}{2}\|x - x^*\|^2} dx.$$

Under the substitution $u = x - x^*$, the condition $\|x - \gamma x^*\| > R$ translates to $\|u - \alpha^2 x^*\| > R$. Since $\alpha^2 \|x^*\| \rightarrow 0$, for all sufficiently large r this implies $\|u\| > R/2$. Applying Lemma 9 to the polynomial $u \mapsto g(u + x^*)$, with radius $R/2$ and scaling parameter β yields

$$\int_{\|x - \gamma x^*\| > R} |g(x)| G_\alpha(x)^2 d\mu(x) \leq C_{d,N,x^*} \|g\|_1 \beta^{-1} R^{d+N-2} e^{-\frac{1}{8}\beta R^2}.$$

Combining (i)–(iii), there exists a constant $C_g > 0$, depending on d, N, x^* , and $\|g\|_1$, such that

$$\left| \int_{\mathbb{R}^d} g q^2 d\mu - \int_{\mathbb{R}^d} g G_\alpha^2 d\mu \right| \leq C_g \left(\delta + e^{-cR^2} + e^{-c'\beta R^2} \right) = \mathcal{E}_g(r). \tag{20}$$

Here, we have absorbed polynomial factors into the exponential. Indeed, since $\beta R^2 \rightarrow \infty$, after decreasing the exponential constant if necessary, for all sufficiently large r ,

$$\beta^{-1} R^{d+N-2} e^{-\beta R^2/8} \leq C e^{-c' \beta R^2}.$$

Step 3: Normalization and bounding the denominator.

Applying the estimate (20) with $g = 1$ gives

$$\left| \int_{\mathbb{R}^d} q(x)^2 d\mu(x) - J_0 \right| \leq \mathcal{E}_1(r),$$

where $J_0 = \int_{\mathbb{R}^d} G_\alpha(x)^2 d\mu(x) = (2\pi)^{-d/2} e^{-\frac{\gamma}{2} \|x^*\|^2} \int_{\mathbb{R}^d} e^{-\frac{\beta}{2} \|u\|^2} du = e^{-\frac{\gamma}{2} \|x^*\|^2} \beta^{-d/2}$. We examine the relative error $\mathcal{E}_1(r)/J_0$. Since $\gamma \rightarrow 1$, the factor $e^{\frac{\gamma}{2} \|x^*\|^2}$ is bounded by a constant for large r . Hence, there exists a constant $C_0 > 0$ such that

$$\frac{\mathcal{E}_1(r)}{J_0} \leq C_0 \beta^{d/2} \left(\delta + e^{-cR^2} + e^{-c' \beta R^2} \right).$$

Since $\beta = \Theta(r/\log r)$, we estimate the three terms:

$$\begin{aligned} \beta^{d/2} \delta &\leq C \left(\frac{r}{\log r} \right)^{d/2} r^{-A} = o(1) \quad \text{since } A > \frac{d}{2}, \\ \beta^{d/2} e^{-cR^2} &\leq C \left(\frac{r}{\log r} \right)^{d/2} e^{-cKr} = o(1), \\ \beta^{d/2} e^{-c' \beta R^2} &\leq C \left(\frac{r}{\log r} \right)^{d/2} \exp \left(-c'' \frac{r^2}{\log r} \right) = o(1). \end{aligned}$$

Thus $\mathcal{E}_1(r)/J_0 = o(1)$. In particular, for all sufficiently large r , $\mathcal{E}_1(r) \leq \frac{1}{2} J_0$, which establishes the denominator lower bound

$$(21) \quad \int_{\mathbb{R}^d} q(x)^2 d\mu(x) \geq \frac{1}{2} J_0 = \frac{1}{2} e^{-\frac{\gamma}{2} \|x^*\|^2} \beta^{-d/2}.$$

Step 4: Bounding the numerator.

Now apply (20) to the nonnegative polynomial $g(x) = f(x) - f_{\min}$. Setting

$$J_F := \int_{\mathbb{R}^d} (f(x) - f_{\min}) G_\alpha(x)^2 d\mu(x),$$

we have $\int_{\mathbb{R}^d} (f - f_{\min}) q^2 d\mu \leq J_F + \mathcal{E}_{f-f_{\min}}(r)$. Using Equation (18) and substituting $u = x - x^*$,

$$J_F = (2\pi)^{-d/2} e^{-\frac{\gamma}{2} \|x^*\|^2} \int_{\mathbb{R}^d} F(u) e^{-\frac{\beta}{2} \|u\|^2} du,$$

where $F(u) := f(x^* + u) - f(x^*)$. Because x^* is a global minimizer of f , we have

$$F(0) = 0 \quad \text{and} \quad \nabla F(0) = \nabla f(x^*) = 0.$$

Consequently, the Taylor expansion of F at the origin contains neither a constant nor linear terms, i.e.,

$$F(u) = \frac{1}{2} u^T \nabla^2 f(x^*) u + \sum_{3 \leq |\nu| \leq N} \frac{D^\nu f(x^*)}{\nu!} u^\nu.$$

By symmetry of the Gaussian measure $e^{-\frac{\beta}{2}\|u\|^2} du$, all odd-degree monomials integrate to zero. For every even monomial of degree $|\nu| = 2k \geq 2$,

$$\frac{\int_{\mathbb{R}^d} u^\nu e^{-\frac{\beta}{2}\|u\|^2} du}{\int_{\mathbb{R}^d} e^{-\frac{\beta}{2}\|u\|^2} du} = O(\beta^{-k}) \leq O(\beta^{-1}).$$

Summing over the finitely many terms in F , we obtain

$$(22) \quad \frac{J_F}{J_0} = \frac{\int_{\mathbb{R}^d} F(u) e^{-\frac{\beta}{2}\|u\|^2} du}{\int_{\mathbb{R}^d} e^{-\frac{\beta}{2}\|u\|^2} du} \leq \frac{A_{f,d}}{\beta},$$

for a positive constant $A_{f,d}$ depending only on f and d .

Step 5: Combining the estimates.

Combining the Rayleigh quotient (16) with the denominator bound (21) and numerator estimate gives

$$(23) \quad \lambda_r - f_{\min} \leq \frac{J_F + \mathcal{E}_{f-f_{\min}}(r)}{\frac{1}{2}J_0} = 2\frac{J_F}{J_0} + 2\frac{\mathcal{E}_{f-f_{\min}}(r)}{J_0}.$$

By (22) and the definition of $\beta = 1 + \alpha^{-2}$, the leading term satisfies

$$2\frac{J_F}{J_0} \leq \frac{2A_{f,d}}{\beta} \leq 2A_{f,d}\alpha^2 = 2A_{f,d}C_{\deg}^2 AK \frac{\log r}{r}.$$

For the remaining error term, since $A > \frac{d}{2} + 1$, the same asymptotics as in Step 3 show that

$$\frac{\mathcal{E}_{f-f_{\min}}(r)}{J_0} \leq C\beta^{d/2} \left(r^{-A} + e^{-cKr} + e^{-c'' \frac{r^2}{\log r}} \right) = O\left(\left(\frac{r}{\log r} \right)^{d/2} r^{-A} \right) = o\left(\frac{\log r}{r} \right).$$

Substituting these bounds into (23) yields

$$\lambda_r - f_{\min} \leq C_{f,d,N} \frac{\log r}{r}$$

for all sufficiently large r .

DEPARTMENT OF ECONOMETRICS & OPERATIONS RESEARCH, TILBURG UNIVERSITY, NETHERLANDS

Email address: s.j.gribling@tilburguniversity.edu

UNIVERSITY OF LJUBLJANA, FACULTY OF MATHEMATICS AND PHYSICS, JADRANSKA 21, 1000 LJUBLJANA & UNIVERSITY OF PRIMORSKA, FACULTY OF MATHEMATICS, NATURAL SCIENCES AND INFORMATION TECHNOLOGIES, GLAGOLJAŠKA ULICA 8, 6000 KOPER, SLOVENIA.

Email address: igor.klep@fmf.uni-lj.si

UNIVERSITY OF LJUBLJANA, FACULTY OF MATHEMATICS AND PHYSICS, JADRANSKA 21, 1000 LJUBLJANA

Email address: poornendu.kumar@fmf.uni-lj.si