

Specialized effective Positivstellensätze for improved convergence rates of the moment-SOS hierarchy

Corbinian Schlosser¹ and Matteo Tacchi²

Abstract—Recently a moment-sum-of-squares hierarchy for exit location estimation of stochastic processes has been presented. When restricting to the special case of the unit ball, we show that the solutions approach the optimal value by a super-polynomial rate. To show this result we state a new effective Positivstellensatz on the sphere with quadratic degree bound based on a recent Positivstellensatz for trigonometric polynomials on the hypercube and pair it with a recent effective Positivstellensatz on the unit ball. At the present example, we aim to highlight the effectiveness of specialized Positivstellensätze for the moment-SoS hierarchy and their interplay with problem intrinsic properties.

I. INTRODUCTION

In the last years, there has been increasing interest in leveraging efficient Positivstellensätze to address various classes of (nonlinear) problems, particularly those formulated within the framework of the generalized moment problem (GMP). The versatility of GMPs finds applications across diverse domains, ranging from geometry, where it facilitates volume computation of semialgebraic sets as well as set separation [1], [2], [3], [4], to dynamical systems encompassing optimal control [5], stability analysis [6], [7], [8], [9], [10], [11], as well as partial differential equations [12], and more generally calculus of variations [13]. Moreover, GMPs play a pivotal role in studying stochastic systems and appear in different instances, such as moment methods in parameter estimation or invariant measure computation [14]. Stochastic system analysis gained importance in control theory due to inherent uncertainty in the measurements, the modeling, or the underlying process itself [15]. Recently, GMPs offered computational insights into infinite-time averaging [16], and probability of unsafety [17] to name only a few.

Under mild assumptions [18], the moment-sum-of-squares (SoS) hierarchy provides a convergent method for solving GMPs, but investigation of convergence rates beyond polynomial optimization remains scarce. Notable instances where explicit convergence rates were explored are [19], [20] based on an effective version [21] of Putinar’s Positivstellensatz [22]. Often a large disparity persists between the theoretical convergence bounds and the rates observed in practice. Recent improvements [23] on effective versions of Putinar’s Positivstellensatz strongly reduced this gap. Focusing on

specific semialgebraic sets, such as the unit ball or the hypercube, closed the gap further [24], [25]; and lower bounds have been established [26].

In previous work [27], a framework for obtaining convergence rates from GMPs is investigated. This text can be seen as a follow-up, in which we aim at transferring the improvement in the effective Putinar’s Positivstellensatz for specific semialgebraic sets to GMPs at the example of exit location estimation for stochastic differential equations [28]. This instructive problem models, for instance, the expected position where a safety condition is violated, and has rich properties that can and should be exploited numerically. Compared to [27], we strongly improve the convergence rates for the exit location estimation on the unit ball and point out the interplay between properties of the GMP and the application of an adapted Positivstellensatz.

II. NOTATION

We denote by \mathbb{R} the real numbers, \mathbb{Z} the integers, and \mathbb{N} the natural integers. For $a < b \in \mathbb{N}$, $\llbracket a, b \rrbracket := \{a, a + 1, \dots, b - 1, b\} \subset \mathbb{N}$. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, we set $|\alpha| := \alpha_1 + \dots + \alpha_n$ and $\mathbf{x}^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$. For $n, d \in \mathbb{N}$, we set $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n ; |\alpha| \leq d\}$, and $\mathbb{R}_d[\mathbf{x}] := \{\mathbf{x} \mapsto \sum_{|\alpha| \leq d} c_\alpha \mathbf{x}^\alpha ; (c_\alpha)_{\alpha \in \mathbb{N}_d^n} \in \mathbb{R}^{\mathbb{N}_d^n}\}$ is the space of polynomials of degree at most d , $\mathbb{R}[\mathbf{x}] := \cup_{d \in \mathbb{N}} \mathbb{R}_d[\mathbf{x}]$ is the space of polynomials. We denote the euclidean unit ball $\{x \in \mathbb{R}^n : \|x\|_2 \leq 1\}$ by \mathbf{B} , by $\mathring{\mathbf{B}}$ its interior, and by \mathbf{S} its boundary.

III. PRELIMINARIES

A. Real algebraic Positivstellensätze

We call a set $\mathbf{X} \subset \mathbb{R}^n$ semialgebraic, if there exist $m \in \mathbb{N}$ and polynomials $p_1, \dots, p_m \in \mathbb{R}[\mathbf{x}]$ such that

$$\mathbf{X} = \{\mathbf{x} \in \mathbb{R}[\mathbf{x}] : p_1(\mathbf{x}) \geq 0, \dots, p_m(\mathbf{x}) \geq 0\}. \quad (1)$$

If a set \mathbf{X} is given by (1), we denote it by $\mathbf{X}(\mathbf{p})$.

The moment-SoS hierarchy links polynomial optimization problems with characterizing non-negative polynomials: For $f \in \mathbb{R}[\mathbf{x}]$ and a compact semialgebraic set \mathbf{X} , it holds

$$\begin{aligned} \min_{\mathbf{x}} f(\mathbf{x}) &= \max_{\lambda} \lambda \\ \text{s.t. } \mathbf{x} \in \mathbf{X} &\quad \text{s.t. } f - \lambda \geq 0 \text{ on } \mathbf{X} \end{aligned}$$

and the optimization problem reduces to certification of non-negativity, see to [29], [30] for detailed surveys on polynomial optimization. A natural candidate for non-negative polynomials on $\mathbf{X}(\mathbf{p})$ is the so-called quadratic module.

*The work of Corbinian Schlosser was supported by the European Research Council (grant REAL 947908)

¹Corbinian Schlosser is with INRIA - Ecole Normale Supérieure - PSL Research university, Paris, France corbinian.schlosser@inria.fr

²Matteo Tacchi is with Univ. Grenoble Alpes, CNRS, Grenoble INP (Institute of Engineering Univ. Grenoble Alpes), GIPSA-lab, France matteo.tacchi@gipsa-lab.grenoble-inp.fr

Definition 1 (Sum-of-squares and quadratic module). The set $\Sigma[\mathbf{x}]$ of sum-of-squares (SoS) polynomials is defined by

$$\Sigma[\mathbf{x}] := \left\{ \sum_{i=1}^r q_i^2 : r \in \mathbb{N}, q_1, \dots, q_r \in \mathbb{R}[\mathbf{x}] \right\}.$$

For a vector of polynomials $\mathbf{p} = (p_1, \dots, p_m) \in \mathbb{R}[\mathbf{x}]^m$, the quadratic module $\mathcal{Q}(\mathbf{p})$ is defined by

$$\mathcal{Q}(\mathbf{p}) := \{ (\mathbf{p}^\top \ 1) \boldsymbol{\sigma} : \boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_{m+1}) \in \Sigma[\mathbf{x}]^{m+1} \}$$

and for $\ell \in \mathbb{N}$, the truncated quadratic module $\mathcal{Q}_\ell(\mathbf{p})$ is

$$\mathcal{Q}_\ell(\mathbf{p}) := \{ (\mathbf{p}^\top \ 1) \boldsymbol{\sigma} \in \mathcal{Q}(\mathbf{p}) : \deg(p_i \sigma_i) \leq 2\ell \}.$$

It follows $g \geq 0$ on $\mathbf{X}(\mathbf{p})$ for any $g \in \mathcal{Q}(\mathbf{p})$. The celebrated Putinar's Positivstellensatz [22, Theorem 1.3 & Lemma 3.2] states that, under a certain compactness condition on \mathbf{X} (Archimedean condition), any strictly positive polynomial on \mathbf{X} belongs to $\mathcal{Q}(\mathbf{p})$. For g only non-negative on \mathbf{X} , Putinar's Positivstellensatz does not apply to g but to $g + \varepsilon$ for $\varepsilon > 0$. This motivates the following questions:

For given $\varepsilon > 0$, what is the smallest $\ell = \ell(\varepsilon) \in \mathbb{N}$ such that $g + \varepsilon \in \mathcal{Q}_\ell(\mathbf{p})$?

Whether $\ell(0)$ exists, was (partially) answered positively in [31]; namely generically (in g) it holds $\ell(0) < \infty$. However, whether for given g it holds finite convergence, cannot be decided in polynomial time unless $P = NP$, see [32].

Since the existence of $\ell(0)$ is hard to answer, it is reasonable to investigate how $\ell(\varepsilon)$ behaves as $\varepsilon \rightarrow 0$. Bounding $\ell(\varepsilon)$ has strongly improved since [21] to recent [23] for generic semialgebraic sets \mathbf{X} . Here, we use a specialized Positivstellensatz from [24] with even sharper bounds and consider the case $\mathbf{X} = \mathbf{B} := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 \leq 1\} = \mathbf{X}(b)$ for $b(\mathbf{x}) = 1 - \|\mathbf{x}\|_2^2$. Following [24], for $g \in \mathbb{R}[\mathbf{x}]$ we set

$$\text{lb}(g, \ell) := \max\{\lambda \in \mathbb{R} : g - \lambda \in \mathcal{Q}_\ell(b)\} \quad (3)$$

and recall the following Positivstellensatz [24] on \mathbf{B} .

Theorem 2 ([24, Theorem 3]). *Let $g \in \mathbb{R}[\mathbf{x}]$ and $d = \deg(g)$. Consider $\mathbf{B} = \mathbf{X}(b)$ and set $g_{\min} := \min_{\mathbf{x} \in \mathbf{B}} g(\mathbf{x})$, $g_{\max} := \max_{\mathbf{x} \in \mathbf{B}} g(\mathbf{x})$. Then, for $\ell \geq nd$, it holds*

$$g_{\min} - \text{lb}(g, \ell) \leq \frac{c_n(d)}{\ell^2} (g_{\max} - g_{\min}) \quad (4)$$

where $c_n \in \mathbb{R}[x]_n$ is a polynomial in one variable.

We infer a quantitative membership certificate for $\mathcal{Q}_\ell(b)$.

Corollary 3. *A sufficient condition for $g \in \mathbb{R}[\mathbf{x}]$ with $g \geq 0$ on the unit ball \mathbf{B} to be in $\mathcal{Q}_\ell(b)$ is that*

$$\ell^2 \geq c_n(d) \frac{g_{\max} - g_{\min}}{g_{\min}} \quad (5)$$

Proof. Let $g \in \mathbb{R}[\mathbf{x}]$ be positive on \mathbf{B} . Then, $g \in \mathcal{Q}_\ell(b)$ if and only if $\lambda = 0$ is feasible in the optimization problem (3), i.e. $\text{lb}(g, \ell) \geq 0$. Using Theorem 2, a sufficient condition – resulting in (5) after rearranging – for this to happen is that

$$g_{\min} \geq \frac{c_n(d)}{\ell^2} (g_{\max} - g_{\min}).$$

□

B. A Positivstellensatz for trigonometric polynomials

We call a function $f : [0, 1]^n \rightarrow \mathbb{C}$ a trigonometric polynomial of bandwidth d , denoted by $f \in \mathcal{T}_d$, if

$$f(\mathbf{x}) = \sum_{\boldsymbol{\omega} \in \llbracket -d, d \rrbracket^n} f_{\boldsymbol{\omega}} e^{2\pi i \boldsymbol{\omega}^\top \mathbf{x}}$$

for coefficients $f_{\boldsymbol{\omega}} \in \mathbb{C}$ for $\boldsymbol{\omega} \in \llbracket -d, d \rrbracket^n$. By \mathcal{T} we denote the union of all sets \mathcal{T}_d . For $f \in \mathcal{T}$ let

$$\|f\|_F := \sum_{\boldsymbol{\omega} \in \llbracket -d, d \rrbracket^n} |f_{\boldsymbol{\omega}}|. \quad (6)$$

The truncated sum-of-squares of trigonometric polynomials are, for $\ell \in \mathbb{N}$,

$$\Sigma_\ell^\mathcal{T} := \left\{ f \in \mathcal{T} : f = \sum_{i=1}^s f_i^2, s \in \mathbb{N}, f_1, \dots, f_s \in \mathcal{T}_\ell \right\}.$$

As for \mathcal{Q}_ℓ , the set $\Sigma_\ell^\mathcal{T}$ can be represented via positive semidefinite matrices, see [29], [33],

$$\Sigma_\ell^\mathcal{T} = \{ f \in \mathcal{T} : f = \phi^{(\ell)*} Q \phi^{(\ell)}, \text{ for some } Q \succeq 0 \} \quad (7)$$

for $\phi^{(\ell)} := (\phi_{\boldsymbol{\omega}}^{(\ell)})_{\boldsymbol{\omega} \in \llbracket -\ell, \ell \rrbracket^n}$ given by

$$\phi_{\boldsymbol{\omega}}^{(\ell)}(\mathbf{x}) := \frac{1}{(2\ell + 1)^{\frac{n-1}{2}}} e^{-2\pi i \boldsymbol{\omega}^\top \mathbf{x}}.$$

Analogous to (3), for $q \in \mathcal{T}$ and each $\ell \in \mathbb{N}$ we denote

$$\text{lb}_\mathcal{T}(q, \ell) := \max\{c \in \mathbb{R} : q - c \in \Sigma_\ell^\mathcal{T}\}.$$

Setting $q_{\min} := \min_{\mathbf{x} \in [0, 1]^n} q(\mathbf{x})$ we recall a result from [33].

Theorem 4 ([33, Theorem 1]). *Let $q \in \mathcal{T}_{2d}$ be real-valued. For $\ell \geq 3d$ it holds, for q_0 the mean value of q , that*

$$q_{\min} - \text{lb}_\mathcal{T}(q, \ell) \leq \|q - q_0\|_F \left(\left(1 - \frac{6d^2}{\ell^2} \right)^{-d} - 1 \right) \quad (8)$$

Corollary 5. *A sufficient condition for $q \in \mathcal{T}_{2d}$ with $q \geq 0$ on $[0, 1]^n$ to belong to $\Sigma_\ell^\mathcal{T}$ is that*

$$\ell^2 \geq 12d^2 n \max \left\{ 1, \frac{\|q - q_0\|_F}{q_{\min}} \right\}$$

Proof. The function q belongs to $\Sigma_\ell^\mathcal{T}$ if and only if $\text{lb}_\mathcal{T}(q, \ell) \geq 0$. For $\ell \geq 3d$, By Theorem 4, this can be guaranteed when

$$q_{\min} \geq \|q - q_0\|_F \left(\left(1 - \frac{6d^2}{\ell^2} \right)^{-d} - 1 \right). \quad (9)$$

To bound the right-hand side, we use the Bernoulli inequality: $(1+x)^n \geq 1+nx$ for any $x \geq -1$. For $\ell^2 \geq 12d^2 n$ we choose $x = -\frac{6d^2}{\ell^2} > -\frac{1}{2n}$ and get

$$\begin{aligned} \left(1 - \frac{6d^2}{\ell^2} \right)^{-n} - 1 &= (1+x)^{-n} - 1 \leq \frac{1}{1+nx} - 1 \\ &= \frac{-nx}{1+nx} \leq \frac{-nx}{\frac{1}{2}} = \frac{12nd^2}{\ell^2}. \end{aligned}$$

We infer that (9) is satisfied for $\ell^2 \geq 12d^2 n \frac{\|q - q_0\|_F}{q_{\min}}$ (under the assumption $\ell^2 \geq 12d^2 n$). This shows the statement. □

IV. EXIT LOCATION ESTIMATION

We consider the following setting from [28]: Let a stochastic differential equation be given by

$$dX_t = \mathbf{f}_0(X_t) dt + F(X_t) dW_t, \quad X_0 = \mathbf{x}_0 \quad (10)$$

for $\mathbf{f}_0 = (f_{0i})_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $F = (f_{ij})_{i,j} : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times r}$, a deterministic initial condition \mathbf{x}_0 and $(W_t)_{t \geq 0}$ a r -dimensional Brownian motion. For a constraint set $\mathbf{X} \subset \mathbb{R}^n$ and a given function $g : \partial\mathbf{X} \rightarrow \mathbb{R}$, we are interested in the expected exit value given by

$$v^*(\mathbf{x}_0) := \mathbb{E}(g(X_\tau)) \quad (11)$$

where $\tau = \inf\{t \geq 0 ; X_t \in \partial\mathbf{X}\}$ is the first time at which the process $(X_t)_t$ starting at $X_0 = \mathbf{x}_0$ hits $\partial\mathbf{X}$. In this text, we restrict to the case where $\mathbf{X} = \mathbf{B}$ is the unit ball. Furthermore, as in [28], [27], we assume the following

Assumption 1.

1.1 $g, f_{0i}, f_{ij} \in \mathbb{R}[\mathbf{x}]$ for $i = 1, \dots, n$ and $j = 1, \dots, r$.

1.2 $A(\mathbf{x}) := F(\mathbf{x})F(\mathbf{x})^\top$ is positive definite for all $\mathbf{x} \in \mathbf{B}$.

Remark 6. Under the above assumptions there exists a unique solution X_t of (10) for $t \leq \tau$, see [28], [34].

The generator \mathcal{L} of the process X_t is given by the second-order partial differential operator

$$\mathcal{L}v(\mathbf{x}) := - \sum_{i,j=1}^n a_{ij}(\mathbf{x}) \frac{\partial^2 v}{\partial \mathbf{x}_i \partial \mathbf{x}_j}(\mathbf{x}) + \sum_{i=1}^n f_{0i}(\mathbf{x}) \frac{\partial v}{\partial \mathbf{x}_i}(\mathbf{x})$$

for $(a_{ij}(\mathbf{x}))_{i,j=1,\dots,n} = F(\mathbf{x})F(\mathbf{x})^\top$. We refer to [34], [28] for details on the generator of a diffusion process X_t .

Remark 7. The function v^* from (11) is the solution of the elliptic second-order partial differential equation

$$\begin{aligned} \mathcal{L}v &= 0 \text{ on } \mathring{\mathbf{B}} \\ v &= g \text{ on } \partial\mathbf{B} =: \mathbf{S}. \end{aligned} \quad (12)$$

where $\mathring{\mathbf{B}}$ denotes the interior of the unit ball and $\mathbf{S} = \partial\mathbf{B}$ its boundary. Because \mathbf{S} is smooth and g is smooth, it follows from elliptic regularity [35], [28] that a solution $v^* \in C^\infty(\mathbf{B})$ of (12) exists and is unique.

The elliptic maximum principle, applied to (12), yields that $v^*(\mathbf{x}_0)$ from (11) is the solution of the following infinite dimensional linear programming problem (LP), see [28],

$$\begin{aligned} v^*(\mathbf{x}_0) &= \max_v v(\mathbf{x}_0) \\ \text{s.t.} \quad &\mathcal{L}v \leq 0 \text{ on } \mathbf{B} \\ &v \leq g \text{ on } \mathbf{S}. \end{aligned} \quad (13)$$

Assumption 1.1 ensures that the LP (13) is an instance of a GMP and thus the moment-SoS hierarchy framework can be applied, leading to a hierarchy of (convex) semidefinite programs (SDPs), see [28]: For each $\ell \in \mathbb{N}$ consider

$$\begin{aligned} v_\ell^*(\mathbf{x}_0) &:= \sup_{v \in \mathbb{R}[\mathbf{x}]} v(\mathbf{x}_0) \\ \text{s.t.} \quad &-\mathcal{L}v \in \mathcal{Q}_\ell(b) \\ &g - v \in \mathcal{Q}_\ell(-b, b) = \Sigma[\mathbf{x}] + b \cdot \mathbb{R}[\mathbf{x}] \end{aligned} \quad (14)$$

As a special case of [28, Theorem 2] it holds $v_\ell^*(\mathbf{x}_0) \rightarrow v^*(\mathbf{x}_0)$ as $\ell \rightarrow \infty$ and by [27, Theorem 4.14], inserting $\mathbf{X} = \mathbf{B}$, we get the following convergence rate

$$v^*(\mathbf{x}_0) - v_\ell^*(\mathbf{x}_0) \in \mathcal{O}\left(\ell^{-\frac{1}{(2.5+s)n}}\right) \text{ for any } s > 0. \quad (15)$$

V. A POSITIVSTELLENSATZ ON THE SPHERE

By using specialized Positivstellensätze we want to improve the rate (15). Theorem 2 is such an effective Positivstellensatz for the unit ball \mathbf{B} . For the sphere, we present and use the effective Positivstellensatz Corollary 8, in which we transfer Theorem 4 to the sphere via spherical coordinates; more precisely, we use the map $\psi : [0, 1]^{n-1} \rightarrow \mathbf{S}$ from (31) in the Appendix.

Corollary 8. Let $p \in \mathbb{R}_d[\mathbf{x}]$ and $q := p \circ \psi$. Then we have

$$p \geq 0 \text{ on } \mathbf{S} \quad \text{if and only if} \quad q \geq 0 \text{ on } [0, 1]^{n-1}.$$

Additionally, assume $p_{\max} \geq p \geq p_{\min} > 0$ on \mathbf{S} . Then it holds $q \in \Sigma_\ell^\tau$ for

$$\ell^2 \geq 48d^2(n-1)(4d+1)^{\frac{n}{2}} \frac{p_{\max}}{p_{\min}}. \quad (16)$$

Proof. Let ψ be the map defined in the appendix. We will use the following of its properties: The map ψ is a trigonometric polynomial of bandwidth 2 and its image is \mathbf{S} . From the first property, we infer that q has bandwidth at most $2d$, because p is of degree d . The first statement follows from the surjectivity of ψ onto the sphere. For the second statement, note that $p_{\max} \geq p \geq p_{\min} > 0$ implies

$$p_{\max} \geq q \geq p_{\min} \text{ on } [0, 1]^{n-1}. \quad (17)$$

By Corollary 5, for $\ell^2 \geq 12(2d)^2 n \max\left\{1, \frac{\|q - q_0\|_F}{p_{\min}}\right\}$ we have $q \in \Sigma_\ell^\tau$; where we used that the bandwidth of q is at most $2d$. It remains to show

$$\|q - q_0\|_F \leq (4d+1)^{\frac{n}{2}} p_{\max}. \quad (18)$$

We write $q(x) = \sum_{\omega \in \mathbb{I}[-2d, 2d]} q_\omega e^{i2\pi\omega^\top x}$ and get by the Cauchy-Schwarz inequality and the Parseval equality that

$$\begin{aligned} \|q - q_0\|_F &= \left(\sum_{\omega} |q_\omega| \right) - |q_0| \leq \sum_{\omega} |q_\omega| \\ &\leq \left(\sum_{\omega} |q_\omega|^2 \right)^{\frac{1}{2}} (4d+1)^{\frac{n}{2}} \\ &= \|q\|_{L^2([0,1]^{n-1})} (4d+1)^{\frac{n}{2}}. \end{aligned}$$

Lastly, we get (18) via $\|q\|_{L^2([0,1]^{n-1})} \leq p_{\max}$ by (17). \square

Remark 9. The Positivstellensatz in [36] also provides a quadratic bound in ℓ ; however, it is restricted to homogenous polynomials and therefore we cannot apply it to (13). We emphasize that the bound on ℓ in [36] does **not** scale with the state dimension, but, compared to (16), unfavorably in the degree of the polynomial d .

VI. SDP FORMULATION AND CONVERGENCE RATE

Corollary 8 states that we can efficiently certify positivity of a polynomial $p \in \mathbb{R}[x]$ on \mathbf{S} via membership of $q := p \circ \psi$ to $\Sigma_\ell^\mathcal{T}$. Together with Corollary 3, this motivates the following hierarchy of SDPs: For each $\ell \in \mathbb{N}$ consider

$$v_{\mathcal{T}\ell}^*(\mathbf{x}_0) := \sup_{\substack{v \in \mathbb{R}[\mathbf{x}]_{2\ell} \\ \text{s.t.} \\ -\mathcal{L}v \in \mathcal{Q}_\ell(b) \\ g \circ \psi - v \circ \psi \in \Sigma_\ell^\mathcal{T}}} v(\mathbf{x}_0) \quad (19)$$

Remark 10. To show that (19) can be stated as an SDP, note that membership of a (trigonometric) polynomial to $\mathcal{Q}_\ell(b)$ respectively $\Sigma_\ell^\mathcal{T}$ is expressed by linear matrix inequalities, as in (7), see [29], [33]. Because the maps $v \mapsto -\mathcal{L}v$ and $v \mapsto g \circ \psi - v \circ \psi$ are affine, the feasible set for (19) corresponds (linearly) to affine sections of cones of positive semidefinite matrices. Since the cost function in (19) is linear in v , this shows that (19) is an SDP for each $\ell \in \mathbb{N}$.

Remark 11. As $\ell \rightarrow \infty$, the overall number of decision variables has order ℓ^{2n} , as for the SDP formulation of (14).

To obtain a convergence rate for $v_{\mathcal{T}\ell}^*$ to v^* as $\ell \rightarrow \infty$, we follow the strategy from [27], that is

- 1) Approximate v^* by well-chosen feasible polynomials.
- 2) Apply effective Positivstellensätze.

For 1), we let $u \in C^\infty(\overline{\mathbf{B}})$ be the unique solution of the following Poisson problem

$$\begin{aligned} \mathcal{L}u &= -1 & \text{in } \mathbf{B} \\ u &= 0 & \text{on } \mathbf{S}. \end{aligned} \quad (20)$$

The existence of a unique smooth solution u is guaranteed by [35]. For $\varepsilon > 0$ we define the function v_ε by

$$v_\varepsilon := v^* + \varepsilon u - \varepsilon \quad (21)$$

We will see that the function v_ε is strictly feasible and for a suited choice of ε and sufficiently close approximation of v_ε by polynomials, the effective Positivstellensätze Corollary 3 and 8 lead to our following main result.

Theorem 12. Under Assumption 1, it holds

$$0 \leq v^*(\mathbf{x}_0) - v_{\mathcal{T}\ell}^*(\mathbf{x}_0) \in \mathcal{O}_{\ell \rightarrow \infty}(\ell^{-s}) \text{ for any } s \geq 0.$$

Proof. Let $\ell \in \mathbb{N}$. We show first that $v_{\mathcal{T}\ell}^*(\mathbf{x}_0)$ is a lower bound for $v^*(\mathbf{x}_0)$. Let $v \in \mathbb{R}[\mathbf{x}]$ be feasible for (19). Hence, we have $-\mathcal{L}v \geq 0$ on \mathbf{B} (by Corollary 8) and $g - v \geq 0$ on \mathbf{S} . For the function $\tilde{v} := v^* - v$ it follows, using (12),

$$\begin{aligned} \mathcal{L}\tilde{v} &= \mathcal{L}v^* - \mathcal{L}v = -\mathcal{L}v \geq 0 \text{ on } \mathbf{B} \\ v &= g - v \geq 0 \text{ on } \mathbf{S}. \end{aligned}$$

By Assumption 1.2 the operator \mathcal{L} is elliptic and the maximum principle for elliptic operators [35, Theorem 3.1] yields for all $x_0 \in \mathbf{B}$ that $v^*(\mathbf{x}_0) - v(\mathbf{x}_0) = \tilde{v}(x_0) \geq 0$ on \mathbf{B} . Because v was an arbitrary feasible point for the optimization problem (19), we conclude $0 \leq v^*(\mathbf{x}_0) - v_{\mathcal{T}\ell}^*(\mathbf{x}_0)$. To show the claimed convergence rate, let $v^*, u \in C^\infty(\mathbf{B})$ be the solution of (12) and (20) respectively. Let $d, k \in \mathbb{N}$. By the

Jackson-inequality [37], there exist $p_d, q_d \in \mathbb{R}_d[\mathbf{x}]$ such that

$$\|v^* - p_d\|_{C^2(\mathbf{B})}, \|u - q_d\|_{C^2(\mathbf{B})} \leq \frac{\gamma_n(k)}{d^k} \quad (22a)$$

for some constant $\gamma_n(k)$ which depends only on n and k but not on d . Additionally, we define the following constant

$$C := \max \left\{ 4, \sup_{\mathbf{x} \in \mathbf{B}} \sum_{i,j=1}^m |a_{ij}(\mathbf{x})| + \sum_{i=1}^m |f_{0i}(\mathbf{x})| \right\}, \quad (22b)$$

for which it holds for all $h \in C^2(\mathbf{B})$

$$|\mathcal{L}h| = \left| \sum_{i,j=1}^m a_{ij} \frac{\partial^2 h}{\partial \mathbf{x}_i \partial \mathbf{x}_j} + \sum_{i=1}^m f_{0i} \frac{\partial h}{\partial \mathbf{x}_i} \right| \leq C \|h\|_{C^2(\mathbf{B})}. \quad (22c)$$

The condition in (22b) of $C \geq 4$ is used later. For large enough $d \in \mathbb{N}$ (i.e. such that $\gamma_n(k)Cd^{-k} < 1$) we set

$$\varepsilon := \frac{2\gamma_n(k)C}{1 - \gamma_n(k)Cd^{-k}} d^{-k} \in \mathcal{O}_{d \rightarrow \infty}(d^{-k}). \quad (22d)$$

and, motivated by (21), we define the candidate v_d by

$$v_d := p_d + \varepsilon q_d - \varepsilon \in \mathbb{R}_d[\mathbf{x}]. \quad (23a)$$

We first show a degree bound on ℓ for the membership of $-\mathcal{L}v_d$ to $\mathcal{Q}_\ell(b)$ using Corollary 3. By (12) and (20), we have

$$\begin{aligned} \mathcal{L}v_d &= \mathcal{L}(v^* + \varepsilon u) - \mathcal{L}((v^* - p_d) + \varepsilon(u - q_d)) \\ &= -\varepsilon - \mathcal{L}(v^* - p_d) - \varepsilon \mathcal{L}(u - q_d). \end{aligned} \quad (23b)$$

By (22c), it holds $|\mathcal{L}(v^* - p_d)| \leq C\|v^* - p_d\|_{C^2(\mathbf{B})}$ and $|\mathcal{L}(u - q_d)| \leq C\|u - q_d\|_{C^2(\mathbf{B})}$. Inserting this with (22a) into (23b) gives, with the choice of ε in (22d),

$$\begin{aligned} \mathcal{L}v_d &\geq -C \frac{\gamma_n(k)}{d^k} \left(1 + 2 \frac{1 + \gamma_n(k)Cd^{-k}}{1 - \gamma_n(k)Cd^{-k}} \right) \\ \mathcal{L}v_d &\leq -C \frac{\gamma_n(k)}{d^k}. \end{aligned}$$

For $d \in \mathbb{N}$ large enough, such that $\gamma_n(k)Cd^{-k} \leq \frac{1}{2}$, this simplifies to

$$-7 \frac{\gamma_n(k)C}{d^k} \leq \mathcal{L}v_d \leq -\frac{\gamma_n(k)C}{d^k}. \quad (23c)$$

To bound the degree of $\mathcal{L}v_d$ let $\deg A$ and $\deg \mathbf{f}_0$ denote the maximum degree of the polynomials a_{ij} for $i, j \in \llbracket 1, n \rrbracket$ and f_{0i} for $i \in \llbracket 1, n \rrbracket$ respectively. We then have

$$\deg(\mathcal{L}v_d) \leq \underbrace{\max\{d - 2 + \deg A, d - 1 + \deg \mathbf{f}_0\}}_{=: \hat{d} \in \mathcal{O}(d)} \quad (23d)$$

Corollary 3 applied to $-\mathcal{L}v_d$, with lower and upper bounds taken from (23c) and degree (23d), yields $-\mathcal{L}v_d \in \mathcal{Q}_\ell(b)$ for

$$\ell^2 \geq c_n(\hat{d}) \frac{7 \frac{\gamma_n(k)C}{d^k} - \frac{\gamma_n(k)C}{d^k}}{\frac{\gamma_n(k)C}{d^k}} = 6c_n(\hat{d}) \quad (24)$$

where c_n is the univariate polynomial from Theorem 2. Next, we treat the membership of $g - v_d$ to $\Sigma_\ell^\mathcal{T}$. We want to apply

Corollary 8 and therefore we first bound $g - v_d$ from above and below on \mathbf{S} . We have, again by (12) and (20),

$$\begin{aligned} g - v_d &= g - p_d - \varepsilon q_d + \varepsilon \\ &= g - v^* - \varepsilon u + \varepsilon + (v^* - p_d) + (u - q_d) \\ &\geq \varepsilon - \|v^* - p_d\|_{C^2} - \|u - q_d\|_{C^2} \\ (22a) \quad &\geq \varepsilon - \frac{2\gamma_n(k)}{d^k}. \end{aligned}$$

For d large enough, such that $\gamma_n(k)Cd^{-k} > \frac{1}{2}$, we get by the choice of ε , see (22d), and from $C \geq 4$, see (22b), that $\varepsilon \geq \frac{4\gamma_n(k)}{d^k}$. An analog computation for an upper bound of $g - v_d$ on \mathbf{S} then gives $\frac{2\gamma_n(k)}{d^k} \leq g - v_d \leq \frac{6\gamma_n(k)}{d^k}$. Now we can apply Corollary 8 to get $g \circ \psi - v_d \circ \psi \in \Sigma_\ell^T$ for

$$\ell^2 \geq 48d^2(n-1)(4d+1)^{\frac{n}{2}} \frac{6\gamma_n(k)}{2\gamma_n(k)} = 144d^2(n-1)(4d+1)^{\frac{n}{2}}.$$

Together with (23d) we get that v_d is feasible (for large enough d such that $\gamma_n(k)Cd^{-k} \leq \frac{1}{2}$), for $\ell \in \mathbb{N}$ with

$$\ell \geq \ell_d := \max\{6c_n(\hat{d}), 144d^2(n-1)(4d+1)^{\frac{n}{2}}\}^{\frac{1}{2}}. \quad (25)$$

Finally, we have a look at how fast the cost of v_d in (19), namely $v_d(\mathbf{x}_0)$, approaches the optimal cost $v^*(\mathbf{x}_0)$. We have

$$\begin{aligned} v^*(\mathbf{x}_0) - v_d(\mathbf{x}_0) &= (v^* - p_d + \varepsilon q_d - \varepsilon)(\mathbf{x}_0) \\ &\leq \|v^* - p_d\|_{C^2} + \varepsilon(q_d(\mathbf{x}_0) + 1) \\ &\leq \frac{\gamma_n(k)}{d^k} + \varepsilon \left(u(\mathbf{x}_0) + \frac{\gamma_n(k)}{d^k} + 1 \right) \\ &\in \mathcal{O}_{d \rightarrow \infty}(d^{-k}) \end{aligned} \quad (26)$$

since $\varepsilon \in \mathcal{O}(d^{-k})$ by (22d). To conclude the statement, for $\ell \in \mathbb{N}$ large enough we need an upper bound d_ℓ on d such that $\ell_d \leq \ell$. Let us first analyze c_n : By [24, eq. (39)] we get

$$c_n(d) = 2(n+1)^2 d^3 \Gamma_n(\mathbf{B})_d \quad (27a)$$

where [24, Section 4.1] gives

$$\begin{aligned} \Gamma_n(\mathbf{B})_d^2 &\leq \max_{0 \leq k \leq d} \left(1 + \frac{2k}{n-1} \right) \binom{k+n-2}{k} \\ &\leq \left(1 + \frac{2d}{n-1} \right) \binom{d+n-2}{d} \\ &\leq \left(1 + \frac{2d}{n-1} \right) (d+1)^{n-2}. \end{aligned} \quad (27b)$$

Hence, combining (27a) and (27b) yields

$$c_n(d) \leq 2(n+1)^2 d^3 \left(1 + \frac{2d}{n-1} \right)^{1/2} (d+1)^{n/2-1}$$

and with (23d) we deduce the following asymptotic behavior:

$$c_n(\hat{d}) \in \mathcal{O}_{d \rightarrow \infty}(d^{(n+5)/2}) \quad (27c)$$

For the second term in the expression in (25), we have $144d^2(n-1)(4d+1)^{n/2} \in \mathcal{O}_{d \rightarrow \infty}(d^{(n+4)/2})$ which we combine with (25) and (27c) to get

$$\ell_d \in \mathcal{O}_{d \rightarrow \infty}(d^{(n+5)/4})$$

By inverting this asymptotic estimate of ℓ_d , it holds

$$d_\ell \in \mathcal{O}_{\ell \rightarrow \infty}\left(\ell^{\frac{4}{n+5}}\right) \quad (28)$$

and, for ℓ large enough such that $\varepsilon > 0$ in (22d) and $\gamma_n(k)Cd_\ell^k \leq \frac{1}{2}$ (this was used in (23c)), the polynomial v_{d_ℓ} is feasible for (19) for ℓ . Formulating (26) in terms of ℓ gives, using the estimate (28), the convergence rate

$$v^*(\mathbf{x}_0) - v_{d_\ell}(\mathbf{x}_0) \in \mathcal{O}\left(\ell^{-\frac{4k}{n+5}}\right). \quad (29)$$

Because $k \in \mathbb{N}$ was arbitrary the statement follows. \square

Remark 13. The convergence rate in Theorem 12 is not independent of the dimension n . In the proof, for each $s \geq 0$ and $n \in \mathbb{N}$, we find constants $M(n, s)$ such that $v^*(\mathbf{x}_0) - v_{\mathcal{T}_\ell}^*(\mathbf{x}_0) \leq M(n, s)\ell^{-s}$ holds asymptotically as $\ell \rightarrow \infty$.

Remark 14. The reason why we overcome the polynomial convergence rates from [27] for the problem (14) is threefold. First, the optimal solution v^* of (13) is smooth (we used this most efficiently in (26)). Second, we use specialized Positivstellensätze with polynomial scaling in the degree d . Third, the solution v^* satisfies the inequality constraints in (13) with equality. This affected the terms $\frac{p_{\max} - p_{\min}}{p_{\min}}$ and $\frac{p_{\max}}{p_{\min}}$ when we applied Corollary 3 and 8. Namely, for the candidate functions v_{d_ℓ} the corresponding terms are in $\mathcal{O}(1)$ as $\ell \rightarrow \infty$ even though the denominator tended to zero.

VII. CONCLUSION

We continue the pathway of recent research and interest in convergence rates in the moment-SoS hierarchy for generalized moment problems. We focus on the example from [28] of computing exit location estimations for stochastic processes, and, by restricting to the unit ball, we improve our recent polynomial rate in [27] to a super-polynomial convergence rate. Such a convergence rate is remarkable for the moment-SoS hierarchy because it lies in between the recent polynomial convergence rates for static polynomial optimization on the hypercube respectively the sphere [26], [24], [36], [23], [33], and the exponential rate in [33] which is obtained under additional regularity assumptions.

The main ingredients for our improvement compared to recent work [27] are: First, we restrict to a problem setting that allows us to apply specialized Positivstellensätze. More precisely, we apply a Positivstellensatz on the unit ball from [24] (see Theorem 2) and we present a new Positivstellensatz on the sphere, see Corollary 8. Secondly, we exploit two strong properties of the problem at hand and their interaction with the chosen Positivstellensätze, see also Remark 14.

Consequently, the convergence rate from Theorem 12 does not need to transfer to moment-SoS hierarchies for other problems. However, it highlights the importance of adapted and specialized Positivstellensätze and the exploitation of certain problem intrinsic properties in the moment-SoS hierarchy.

Possible continuations include applying the same line of reasoning to different problems with GMP formulations or obtaining explicit bounds on the optimal value, here v^* , based on computed optimal values in the hierarchy, here $v_{\mathcal{T}_\ell}^*$.

APPENDIX

Let $\hat{\psi} = (\hat{\psi}_1, \dots, \hat{\psi}_n) : \mathbb{R}^{n-1} \rightarrow \mathbb{R}^n$ be defined as $\hat{\psi}_i(\theta_1, \dots, \theta_{n-1}) = \cos(\pi\theta_i) \prod_{j=1}^{i-1} \sin(\pi\theta_j)$ for $1 \leq i \leq n-2$ and $\hat{\psi}_{n-1}(\theta_1, \dots, \theta_{n-1}) = \prod_{j=1}^{n-2} \sin(\pi\theta_j) \cos(2\pi\theta_{n-1})$ and $\hat{\psi}_n(\theta_1, \dots, \theta_{n-1}) = \prod_{j=1}^{n-2} \sin(\pi\theta_j) \sin(2\pi\theta_{n-1})$. The map $\hat{\psi}$ restricted to $[0, 1]^{n-2} \times [0, 1)$ is called spherical coordinates because it is injective and it holds, see for instance [38],

$$\hat{\psi}([0, 1]^{n-2} \times [0, 1)) = \mathbf{S}. \quad (30)$$

The function $\hat{\psi}$ is 2-periodic in each variable; therefore we consider

$$\psi(\theta) := \hat{\psi}(2\theta). \quad (31)$$

for which each component is a trigonometric polynomial of bandwidth 2 (because the highest frequency appears in the last two components ψ_{n-1} and ψ_n and is 2). Further, it holds

$$\psi([0, 1]^{n-1}) = \mathbf{S}. \quad (32)$$

This follows from (30), i.e. $\psi([0, \frac{1}{2}]^{n-1}) = \mathbf{S}$, and the identities $\sin(2\pi(\frac{1}{2} + \mathbf{x})) = -\sin(2\pi\mathbf{x})$, $\cos(2\pi(\frac{1}{2} + \mathbf{x})) = -\cos(2\pi\mathbf{x})$ and $\sin(4\pi(\frac{1}{2} + \mathbf{x})) = \sin(4\pi\mathbf{x})$, $\cos(4\pi(\frac{1}{2} + \mathbf{x})) = \cos(4\pi\mathbf{x})$ applied to each of the coordinates of ψ .

ACKNOWLEDGMENT

The authors are grateful to Didier Henrion, Francis Bach, and Alessandro Rudi for their feedback on this work.

REFERENCES

- [1] M. Korda, J.-B. Lasserre, A. Lazarev, V. Magron, and S. Naldi, "Urysohn in action: separating semialgebraic sets by polynomials," *arXiv preprint arXiv:2207.00570*, 2022.
- [2] D. Henrion, J. B. Lasserre, and C. Savorgnan, "Approximate volume and integration for basic semialgebraic sets," *SIAM Review*, vol. 51, pp. 722–743, 2009.
- [3] J. B. Lasserre, "Computing Gaussian and exponential measures of semi-algebraic sets," *Advances in Applied Mathematics*, vol. 91, pp. 137–163, 2017.
- [4] M. Tacchi, T. Weisser, J. B. Lasserre, and D. Henrion, "Exploiting sparsity for semi-algebraic set volume computation," *Foundations of Computational Mathematics*, pp. 1–49, 2022.
- [5] J. B. Lasserre, D. Henrion, C. Prieur, and E. Trélat, "Nonlinear optimal control via occupation measures and LMI-relaxations," *SIAM journal on control and optimization*, vol. 47, no. 4, pp. 1643–1666, 2008.
- [6] M. Korda, "Computing controlled invariant sets from data using convex optimization," *SIAM Journal on Control and Optimization*, vol. 58, no. 5, pp. 2871–2899, 2020.
- [7] A. Oustry, M. Tacchi, and D. Henrion, "Inner approximations of the maximal positively invariant set for polynomial dynamical systems," *IEEE Control Systems Letters*, vol. 3, no. 3, pp. 733–738, 2019.
- [8] M. Korda, D. Henrion, and C. N. Jones, "Convex computation of the maximum controlled invariant set for polynomial control systems," *SIAM Journal on Control and Optimization*, vol. 52, no. 5, pp. 2944–2969, 2014.
- [9] M. Jones and M. M. Peet, "Converse Lyapunov functions and converging inner approximations to maximal regions of attraction of nonlinear systems," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 5312–5319.
- [10] C. Schlosser and M. Korda, "Converging outer approximations to global attractors using semidefinite programming," *Automatica*, vol. 134, p. 109900, 2021.
- [11] D. Goluskin, "Bounding extrema over global attractors using polynomial optimisation," *Nonlinearity*, vol. 33, no. 9, p. 4878, 2020.
- [12] S. Marx, T. Weisser, D. Henrion, and J. B. Lasserre, "A moment approach for entropy solutions to nonlinear hyperbolic PDEs," *Mathematical Control and Related Fields*, vol. 10, no. 1, pp. 113–140, 2020.
- [13] D. Henrion, M. Korda, M. Kružik, and R. Rios-Zertuche, "Occupation measure relaxations in variational problems: the role of convexity," *arXiv preprint arXiv:2303.02434*, 2023.
- [14] M. Korda, D. Henrion, and I. Mezić, "Convex computation of extremal invariant measures of nonlinear dynamical systems and Markov processes," *Journal of Nonlinear Science*, vol. 31, no. 1, pp. 1–26, 2021.
- [15] J. C. Doyle, "Structured uncertainty in control system design," in *1985 24th IEEE Conference on Decision and Control*. IEEE, 1985, pp. 260–265.
- [16] G. Fantuzzi, D. Goluskin, D. Huang, and S. I. Chernyshenko, "Bounds for deterministic and stochastic dynamical systems using sum-of-squares optimization," *SIAM Journal on Applied Dynamical Systems*, vol. 15, no. 4, pp. 1962–1988, 2016.
- [17] J. Miller, M. Tacchi, D. Henrion, and M. Szaier, "Unsafe probabilities and risk contours for stochastic processes using convex optimization," *arXiv preprint arXiv:2401.00815*, 2023.
- [18] M. Tacchi, "Convergence of Lasserres hierarchy: the general case," *Optimization Letters*, vol. 16, pp. 1015–1033, 2022.
- [19] M. Korda, D. Henrion, and C. N. Jones, "Convergence rates of moment-sum-of-squares hierarchies for optimal control problems," *Systems & Control Letters*, vol. 100, pp. 1–5, 2017.
- [20] M. Korda and D. Henrion, "Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets," *Optimization Letters*, vol. 12, no. 3, pp. 435–442, 2018.
- [21] J. Nie and M. Schweighofer, "On the complexity of putinar's positivstellensatz," *Journal of Complexity*, vol. 23, no. 1, pp. 135–150, 2007.
- [22] M. Putinar, "Positive polynomials on compact semi-algebraic sets," *Indiana University Mathematics Journal*, vol. 42, no. 3, pp. 969–984, 1993.
- [23] L. Baldi and B. Mourrain, "On the effective putinar's positivstellensatz and moment approximation," *Mathematical Programming*, vol. 200, no. 1, pp. 71–103, 2023.
- [24] L. Slot, "Sum-of-squares hierarchies for polynomial optimization and the christoffel-darboux kernel," *SIAM Journal on Optimization*, vol. 32, no. 4, pp. 2612–2635, 2022.
- [25] M. Laurent and L. Slot, "An effective version of schmüdgen's positivstellensatz for the hypercube," *Optimization Letters*, vol. 17, no. 3, pp. 515–530, 2023.
- [26] L. Baldi and L. Slot, "Degree bounds for putinar's positivstellensatz on the hypercube," *SIAM Journal on Applied Algebra and Geometry*, vol. 8, no. 1, pp. 1–25, 2024.
- [27] C. Schlosser, M. Tacchi, and A. Lazarev, "Convergence rates for the moment-sos hierarchy," *arXiv preprint arXiv:2402.00436*, 2024.
- [28] D. Henrion, M. Junca, and M. Velasco, "Moment-SOS hierarchy and exit location of stochastic processes," *HAL preprint: hal-03110452*, 2023.
- [29] J. B. Lasserre, *An introduction to polynomial and semi-algebraic optimization*. Cambridge University Press, 2015, vol. 52.
- [30] P. A. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*. California Institute of Technology, 2000.
- [31] J. Nie, "Optimality conditions and finite convergence of lasserres hierarchy," *Mathematical programming*, vol. 146, pp. 97–121, 2014.
- [32] L. F. Vargas, "On the hardness of deciding the finite convergence of lasserre hierarchies," *arXiv preprint arXiv:2401.12613*, 2024.
- [33] F. Bach and A. Rudi, "Exponential convergence of sum-of-squares hierarchies for trigonometric polynomials," *SIAM Journal on Optimization*, vol. 33, no. 3, pp. 2137–2159, 2023.
- [34] B. Oksendal, *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
- [35] D. Gilbarg, N. S. Trudinger, D. Gilbarg, and N. Trudinger, *Elliptic partial differential equations of second order*. Springer, 1977, vol. 224.
- [36] K. Fang and H. Fawzi, "The Sum-of-Squares hierarchy on the sphere and applications in quantum information theory," *Mathematical Programming*, vol. 190, pp. 331–360, 2021.
- [37] T. Bagby, L. Bos, and N. Levenberg, "Multivariate simultaneous approximation," *Constructive approximation*, vol. 18, no. 4, pp. 569–577, 2002.
- [38] L. Blumenson, "A derivation of n-dimensional spherical coordinates," *The American Mathematical Monthly*, vol. 67, no. 1, pp. 63–66, 1960.