POSITIVSTELLENSÄTZE AND MOMENT PROBLEMS WITH UNIVERSAL QUANTIFIERS

XIAOMENG HU, IGOR KLEP, AND JIAWANG NIE

Abstract. This paper studies Positivstellensätze and moment problems for sets K that are given by universal quantifiers. Let $Q \subseteq \mathbb{R}^m$ be a closed set and let $g = (g_1, \ldots, g_s)$ be a tuple of polynomials in two vector variables $\mathtt{x} = (\mathtt{x}_1, \ldots, \mathtt{x}_n)$ and $\mathtt{y} = (\mathtt{y}_1, \ldots, \mathtt{y}_m)$. Then K is described as the set of all points $x \in \mathbb{R}^n$ such that each $g_j(x, y) \geq 0$ for all $y \in Q$. Fix a finite nonnegative Borel measure ν on $\mathbb{R}^m$ with $\mathrm{supp}(\nu) = Q$, and assume it satisfies the multivariate Carleman condition.

The first main result of the paper is a Positivstellensatz with universal quantifiers: if a polynomial $f(\mathtt{x})$ is positive on K, then $f(\mathtt{x})$ belongs to the quadratic module $\mathrm{QM}[g, \nu]$ associated to (g, ν), under the archimedeanness assumption on $\mathrm{QM}[g, \nu]$. Here, $\mathrm{QM}[g, \nu]$ denotes the quadratic module of polynomials in $\mathtt{x}$ that can be represented as

$$\tau_0(\mathtt{x}) + \int \tau_1(\mathtt{x}, y) g_1(\mathtt{x}, y)\, \mathrm{d}\nu(y) + \cdots + \int \tau_s(\mathtt{x}, y) g_s(\mathtt{x}, y)\, \mathrm{d}\nu(y),$$

where each τ_j is a sum of squares polynomial.

Second, necessary and sufficient conditions for a full (or truncated) multisequence to admit a representing measure supported in K are given. In particular, the classical flat extension theorem of Curto and Fialkow is generalized to truncated moment problems on such a set K.

Third, we present applications of the above Positivstellensatz and moment problems in semi-infinite optimization, whose feasible sets are given by infinitely many constraints with universal quantifiers. This results in a new hierarchy of Moment-SOS relaxations. Its convergence is shown under some usual assumptions. The quantifier set Q is allowed to be non-semialgebraic, which makes it possible to solve some optimization problems with non-semialgebraic constraints.

1. Introduction

Positivstellensätze and moment problems are pillars of real algebraic geometry [BCR98, Lau09, Sce09] and are of broad interest in computational and applied mathematics. This paper concerns these two topics when the constraining sets are given by universal quantifiers. Let $\mathtt{x} = (\mathtt{x}_1, \ldots, \mathtt{x}_n)$ and $\mathtt{y} = (\mathtt{y}_1, \ldots, \mathtt{y}_m)$ be tuples of variables. We are interested in subsets K of $\mathbb{R}^n$ that are given by inequalities in x, with y as a universal quantifier (see [Las15]). Let $Q \subseteq \mathbb{R}^m$ be a given closed set. For a tuple $g = (g_1, \ldots, g_s)$ of polynomials in $\mathbb{R}[\mathtt{x}, \mathtt{y}]$, consider the following set given by the universal quantifier y:

$$K = \{x \in \mathbb{R}^n : g_1(x, y) \geq 0, \ldots, g_s(x, y) \geq 0 \ \ \forall y \in Q\}. \tag{1.1}$$

When there is no universal quantifier y, the set K is a classical basic closed semialgebraic set. By Tarski's transfer principle [BCR98], if the quantifier set Q is semialgebraic, then K is semialgebraic. A quantifier-free description for K can be obtained by applying symbolic computations like cylindrical algebraic decompositions (see [BPR]). However, computing a quantifier-free description is typically computationally expensive. In this paper, the set Q is

---

Date: November 30, 2024.

2020 *Mathematics Subject Classification.* 13J30, 44A60, 90C23, 47A57, 90C34.

Key words and phrases. Positivstellensatz, moment problem, polynomial, universal quantifier, semi-infinite optimization, real algebraic geometry.

allowed to be non-semialgebraic, so K may also be non-semialgebraic. For instance, when $Q = \mathbb{Z}^m$ ($\mathbb{Z}$ denotes the set of integers), the set K is defined by countably many constraints.

- For $Q = \mathbb{Z}^1$ and $g(\mathtt{x}, \mathtt{y}) = (\mathtt{x}_1 + \mathtt{y})^2 + \mathtt{y}^2 - \mathtt{x}_2^2 \geq 0$, the set K is given by
$$x_1^2 - x_2^2 \geq 0, \quad 1 \geq \frac{x_2^2}{k^2} - \frac{(x_1 + k)^2}{k^2} \quad \text{for} \quad k = \pm 1, \pm 2, \dots.$$
- For $Q = \mathbb{Z}^1$ and $g(\mathtt{x}, \mathtt{y}) = \mathtt{x}_2 - 2\mathtt{y}\mathtt{x}_1 + \mathtt{y}^2 \geq 0$, the set K is a convex polygon with infinitely many sides.
- For $Q = \mathbb{Z}_+$ (the set of positive integers) and $g(\mathtt{x}, \mathtt{y}) = 4\mathtt{y}^4 - 1 - \mathtt{y}^2(2\mathtt{y}^2 - 1)\mathtt{x}_1^2 - \mathtt{y}^2(2\mathtt{y}^2 + 1)\mathtt{x}_2^2 \geq 0$, the set K is the intersection of infinitely many ellipses:
$$\frac{x_1^2}{2 + k^{-2}} + \frac{x_2^2}{2 - k^{-2}} \leq 1 \quad \text{for} \quad k = 1, 2, \dots.$$

Positivstellensätze concern representations of polynomials that are positive (or nonnegative) on a set K. Equivalently, for a given polynomial $f \in \mathbb{R}[\mathtt{x}]$, what is a test or certificate for $f \geq 0$ on K? When does such a certificate hold necessarily? When K has no universal quantifier y (i.e., the polynomials g_i in (1.1) do not depend on y), Positivstellensätze have been extensively studied, see, e.g., the surveys and books [BCR98, HKL20, Las15, Lau09, Nie23, Sce09] or the following small sample of recent papers [CKS09, EP20, Fri21, GKKS15, LPR20, MNR23, PV99, Rie16, SS24, Scw03] and the references therein. For instance, consider the Putinar certificate

$$f \;=\; \sigma_0 + \sigma_1 g_1 + \cdots + \sigma_s g_s, \tag{1.2}$$

where all σ_i are sum-of-squares (SOS) polynomials in $\mathbb{R}[\mathtt{x}]$. Clearly, if f has a representation of the form (1.2), then $f \geq 0$ on the set K. When the quadratic module of g is archimedean, if $f > 0$ on K, then by Putinar's Positivstellensatz [Put93] a representation of the form (1.2) must hold. A representation more general than (1.2) is given by the Schmüdgen Positivstellensatz [Smü91], which uses the preordering of g. All these classical results assume that K is a basic closed semialgebraic set. However, when K depends on quantifiers as in (1.1), there is little work on Positivstellensätze. This is remedied in the present paper.

Closely related to Positivstellensätze are moment problems. Let $\mathbb{N}$ denote the set of nonnegative integers and $\mathbb{N}^n$ denote the set of nonnegative integer vectors of length n. For a given multisequence $z = (z_\alpha)_{\alpha \in \mathbb{N}^n}$, i.e., z is a vector whose entries are labelled by nonnegative integer vectors in $\mathbb{N}^n$, the moment problem concerns the existence of a nonnegative Borel measure[1] μ on $\mathbb{R}^n$ such that

$$z_\alpha \;=\; \int x^\alpha \, \mathrm{d}\mu(x) \quad \forall\, \alpha \in \mathbb{N}^n. \tag{1.3}$$

In the above, $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ for the multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$. The sequence z is said to be a *moment sequence* if such a Borel measure μ exists, and in this case μ is called

[1] All our measures will be assumed finite.

a representing measure for z. We refer the reader to the surveys [Ber87, Fia16], books [Akh65, Smü17], or papers [BS16, BL20, CMN11, CGIK23, IK17, IKKM22, IKKM23, KW13, Net08, PS01] and the references therein for more details about moment problems.

In many applications, the support of the measure μ is often required to be contained in a set K, i.e., $\mathrm{supp}(\mu) \subseteq K$. Then z is called a *K-moment sequence* and μ is called a *K-representing measure* for z, if (1.3) holds for a Borel measure μ with $\mathrm{supp}(\mu) \subseteq K$. When K is described without quantifiers, this is the classical K-moment problem (see [Fia16, Smü17]). However, there is little work on the K-moment problem when K depends on the quantifier y. This is the second main topic of the present paper.

Positivstellensätze and moment problems with universal quantifiers are useful for solving semi-infinite optimization problems. A typical problem of semi-infinite optimization is

$$\begin{cases} \min\limits_{x\in X} & f(x) \\ \text{s.t.} & g(x,y) \geq 0 \quad \forall y \in Q. \end{cases} \tag{1.4}$$

Here, the constraining function g depends on both x and y as is the case for the g_j in (1.1), and $X \subseteq \mathbb{R}^n$ is another given constraining set for x that does not depend on the quantifier y. The quantifier set Q in (1.4) need not be a basic closed semialgebraic set. Solving this kind of semi-infinite optimization problem is typically a highly challenging task. However, Positivstellensätze and moment problems with universal quantifiers are powerful mathematical tools for solving them. This is the third main topic of our paper.

Contributions. The new contribution of this paper is to solve the three above mentioned major problems.

Our first contribution is a Positivstellensatz for sets K defined with universal quantifiers as in (1.1). If a polynomial $f \in \mathbb{R}[\mathtt{x}]$ has the representation

$$f(\mathtt{x}) = \sigma(\mathtt{x}) + \sum_{j=1}^{s} \int \tau_j(\mathtt{x},y) g_j(\mathtt{x},y)\,\mathrm{d}\nu(y), \tag{1.5}$$

where σ is an SOS polynomial in $\mathtt{x}$, $\tau_1, \ldots, \tau_s$ are SOS polynomials in $(\mathtt{x}, \mathtt{y})$ and ν is a Borel measure on $\mathbb{R}^m$ such that $\mathrm{supp}(\nu) \subseteq Q$, then we clearly have $f \geq 0$ on K. The Positivstellensatz ensures that the reverse implication is essentially true. The set of all polynomials in $\mathbb{R}[\mathtt{x}]$ that can be written as in (1.5) is denoted by $\mathrm{QM}[g,\nu]$. It is called the *quadratic module generated by g and ν*. Assume ν is a Borel measure on $\mathbb{R}^m$ such that $\mathrm{supp}(\nu) = Q$ and ν satisfies the Carleman condition

$$\sum_{d=0}^{\infty} \left(\int y_j^{2d}\,\mathrm{d}\nu(y) \right)^{-\frac{1}{2d}} = \infty \quad \text{for} \quad j = 1, \ldots, m. \tag{1.6}$$

We show in Theorem 3.4 that if $f > 0$ on K and $\mathrm{QM}[g,\nu]$ is archimedean (i.e., $N - \mathtt{x}_1^2 - \cdots - \mathtt{x}_n^2 \in \mathrm{QM}[g,\nu]$ for some positive integer N), then $f \in \mathrm{QM}[g,\nu]$ must hold. This is a generalization of Putinar's Positivstellensatz to sets given by universal quantifiers. Since the truncations of the quadratic module $\mathrm{QM}[g,\nu]$ for given degrees can be represented by semidefinite programs (SDPs), Theorem 3.4 gives rise to a Moment-SOS hierarchy of SDP relaxations to optimize a polynomial over K.

Our second contribution is about K-moment problems for sets K defined by universal quantifiers as in (1.1). A key tool for studying moment problems is the Riesz functional. A multisequence $z = (z_\alpha)_{\alpha\in\mathbb{N}^n}$ gives rise to the linear functional:

$$\mathscr{R}_z : \mathbb{R}[\mathtt{x}] \to \mathbb{R}, \quad \mathtt{x}^\alpha \mapsto z_\alpha.$$

This is equivalent to $\mathscr{R}_z(f) = \sum_\alpha f_\alpha z_\alpha$ for the polynomial $f = \sum_\alpha f_\alpha \mathtt{x}^\alpha$. The functional $\mathscr{R}_z$ is called the *Riesz functional* of z. If μ is a representing measure for z, then

$$\mathscr{R}_z(f) = \int f(x)\,\mathrm{d}\mu(x) \quad \text{for all} \quad f \in \mathbb{R}[\mathtt{x}].$$

If in addition, $\mathrm{supp}(\mu) \subseteq K$, then

$$\mathscr{R}_z(f) \ge 0 \quad \text{for all} \quad f \in \mathbb{R}[\mathtt{x}] : f|_K \ge 0. \tag{1.7}$$

We say the Riesz functional $\mathscr{R}_z$ is *K-positive* if (1.7) holds. The Riesz functional $\mathscr{R}_z$ is simply said to be positive if it is $\mathbb{R}^m$-positive. The K-positivity is necessary for z to have a K-representing measure. For a closed set K, being K-positive is also sufficient. This is a classical result of M. Riesz ($n = 1$) and Haviland ($n > 1$); see the works [Akh65, Ber87, Fia16, Havi36, Riesz, Smü17] for details. When the quadratic module $\mathrm{QM}[g,\nu]$ is archimedean, we show in Section 4 that z is a K-moment sequence if and only if $\mathscr{R}_z(f) \ge 0$ for all $f \in \mathrm{QM}[g,\nu]$. Moreover, we also give concrete conditions for $\mathscr{R}_z \ge 0$ on $\mathrm{QM}[g,\nu]$ in terms of moment and localizing matrices (see Theorem 4.4).

Our third contribution is on semi-infinite optimization. Suppose the constraining set

$$X = \{x \in \mathbb{R}^n : c_{eq}(x) = 0,\, c_{in}(x) \ge 0\},$$

for two tuples c_{eq}, c_{in} of polynomials in $\mathtt{x}$. We refer to (2.1) for the definition of the ideal $\mathrm{Ideal}[c_{eq}]$ generated by c_{eq} and refer to (2.1) for the quadratic module $\mathrm{QM}[c_{in}]$ generated by c_{in}. When $\mathrm{QM}[g,\nu] + \mathrm{Ideal}[c_{eq}] + \mathrm{QM}[c_{in}]$ is archimedean, we show in Section 5 that the semi-infinite optimization problem (1.4) is equivalent to

$$\left\{\begin{array}{cl} \min\limits_{z\in\mathbb{R}^{\mathbb{N}^n}} & \mathscr{R}_z(f) \\ \text{s.t.} & \mathscr{R}_z \ge 0 \quad \text{on} \quad \mathrm{QM}[g,\nu] + \mathrm{Ideal}[c_{eq}] + \mathrm{QM}[c_{in}], \\ & \mathscr{R}_z(1) = 1. \end{array}\right. \tag{1.8}$$

When the ideals and quadratic modules are truncated by degrees, the above produces a hierarchy of Moment-SOS type semidefinite programming relaxations. We prove the convergence property for this hierarchy in Theorem 5.2. Finally, we also discuss how to estimate moments of the measure ν by sampling when the moments are not known explicitly. We remark that the quantifier set Q is allowed to be non-semialgebraic. So this makes it possible to solve some semi-infinite optimization problems with non-semialgebraic constraints.

The paper is organized as follows. Notation is fixed and some background on polynomial optimization and moment problems is given in Section 2. Positivstellensätze, moment problems and semi-infinite optimization for sets given by universal quantifiers are respectively presented in Section 3, Section 4, and Section 5. Some computational experiments are presented in Section 6. Finally, in Section 7, we present our conclusions and engage in a detailed discussion of our findings.

2. Preliminaries

2.1. **Notation.** The symbol $\mathbb{R}[\mathtt{x}] = \mathbb{R}[\mathtt{x}_1, \ldots, \mathtt{x}_n]$ denotes the ring of polynomials in $\mathtt{x} = (\mathtt{x}_1, \ldots, \mathtt{x}_n)$ with real coefficients. The symbol $\mathbb{R}_+$ stands for the set of nonnegative real numbers. For a symmetric matrix W, $W \succeq 0$ means that W is positive semidefinite. For a vector u, $\|u\|$ denotes its standard Euclidean norm. The notation I_n denotes the $n \times n$ identity matrix. The superscript ^{T} denotes the transpose of a matrix or vector. The symbol e denotes the vector of all ones, i.e., $e = (1, \ldots, 1)$. We use $\otimes$ to denote the classical Kronecker product.

For $\mathtt{x} = (\mathtt{x}_1, \ldots, \mathtt{x}_n)$ and $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, the notation $\mathtt{x}^\alpha := \mathtt{x}_1^{\alpha_1} \cdots \mathtt{x}_n^{\alpha_n}$ stands for the monomial of $\mathtt{x}$ with power α. We denote the power set

$$\mathbb{N}_d^n = \{\alpha \in \mathbb{N}^n \,|\, \alpha_1 + \cdots + \alpha_n \le d\}.$$

Denote by $\mathbb{R}^{\mathbb{N}_d^n}$ the space of real vectors that are labeled by $\alpha \in \mathbb{N}_d^n$. For a positive integer d, the vector of all monomials in $\mathtt{x}$ of degrees at most d, ordered with respect to the graded lexicographic ordering, is denoted as

$$[\mathtt{x}]_d := \begin{pmatrix} 1 & \mathtt{x}_1 & \cdots & \mathtt{x}_n & \mathtt{x}_1^2 & \mathtt{x}_1\mathtt{x}_2 & \cdots & \mathtt{x}_n^d \end{pmatrix}^T.$$

A polynomial $\sigma \in \mathbb{R}[\mathtt{x}]$ is said to be a sum of squares (SOS) polynomial if $\sigma = \sigma_1^2 + \cdots + \sigma_k^2$ for some $\sigma_1, \ldots, \sigma_k \in \mathbb{R}[\mathtt{x}]$ and $k \in \mathbb{N} \setminus \{0\}$. The symbol $\Sigma^2[\mathtt{x}]$ denotes the cone of SOS polynomials in $\mathtt{x}$. An interesting fact is that SOS polynomials can be represented through semidefinite programming [Las15, Nie23]. Clearly, each SOS polynomial is nonnegative, while not every nonnegative polynomial is SOS. The approximation performance of SOS polynomials is given in [Nie12]. Moreover, SOS polynomials are also very useful in tensor optimization [Nie17, NZ18].

For two sets $S, T \subseteq \mathbb{R}[\mathtt{x}]$, their product and addition are defined as

$$S \cdot T = \{pq : \, p \in S, q \in T\}, \quad S + T = \{p + q : \, p \in S, q \in T\}.$$

In particular, if $S = \{p\}$ is a singleton, then we also use

$$p \cdot T = \{pq : \, q \in T\}, \quad p + T = \{p + q : \, q \in T\}.$$

A polynomial tuple $h = (h_1, \ldots, h_m)$ in $\mathbb{R}[\mathtt{x}]$ generates the ideal

$$\text{Ideal}[h] := h_1 \cdot \mathbb{R}[\mathtt{x}] + \cdots + h_m \cdot \mathbb{R}[\mathtt{x}], \tag{2.1}$$

which is the smallest ideal containing all h_i. For $k \in \mathbb{N}$ and $k \ge \deg(h) := \max\{\deg(h_1), \ldots, \deg(h_m)\}$, the kth truncation of Ideal$[h]$ is

$$\text{Ideal}[h]_k := h_1 \cdot \mathbb{R}[x]_{k-\deg(h_1)} + \cdots + h_m \cdot \mathbb{R}[x]_{k-\deg(h_m)}.$$

A tuple $q = (q_1, \ldots, q_t)$ of polynomials in $\mathbb{R}[\mathtt{x}]$ gives rise to the quadratic module (let $q_0 := 1$)

$$\text{QM}[q] := \left\{ \sum_{i=0}^{t} \sigma_i q_i \,\middle|\, \sigma_i \in \Sigma^2[\mathtt{x}] \right\}. \tag{2.2}$$

For $k \in \mathbb{N}$ with $2k \geq \deg(q)$, the kth truncation of $\mathrm{QM}[q]$ is

$$\mathrm{QM}[q]_{2k} \coloneqq \left\{ \sum_{i=0}^{t} \sigma_i q_i \,\middle|\, \sigma_i \in \Sigma^2[\mathtt{x}], \deg(\sigma_i q_i) \leq 2k \right\}.$$

The quadratic module $\mathrm{QM}[q]$ is said to be *archimedean* if there exists an integer $N > 0$ such that

$$N - \mathtt{x}_1^2 - \cdots - \mathtt{x}_n^2 \in \mathrm{QM}[q].$$

Quadratic modules are basic concepts in polynomial optimization and moment problems. We refer to [HKL20, Las15, Lau09, Nie23, Sce09] for recent work in this area.

Let $Q \subseteq \mathbb{R}^m$ be a closed set and ν be a nonnegative Borel measure on $\mathbb{R}^m$ such that $\mathrm{supp}(\nu) = Q$. We let $L^2(\mathbb{R}^m, \nu)$ denote the Hilbert space of all L^2-integrable functions ϕ on Q, i.e., $\int \phi(y)^2 \,\mathrm{d}\nu(y) < \infty$. The inner product on $L^2(\mathbb{R}^m, \nu)$ is given by

$$\langle \phi, \psi \rangle_{L^2} = \int \phi(y)\psi(y) \,\mathrm{d}\nu(y), \quad \phi, \psi \in L^2(\mathbb{R}^m, \nu).$$

A linear functional ℓ on $\mathbb{R}[\mathtt{y}]$, with $\mathtt{y} = (\mathtt{y}_1, \ldots, \mathtt{y}_m)$, is said to satisfy the *multivariate Carleman condition* if

$$\sum_{d=0}^{\infty} \left(\ell(\mathtt{y}_j^{2d}) \right)^{-\frac{1}{2d}} = \infty \quad \text{for} \quad j = 1, \ldots, m. \tag{2.3}$$

3. Positivstellensätze with universal quantifiers

This section proves a Positivstellensatz for polynomials f positive on a set K given by a universal quantifier as in (1.1). Let $Q \subseteq \mathbb{R}^m$ be a given closed set. We fix a nonnegative Borel measure ν on $\mathbb{R}^m$ satisfying the following assumption.

Assumption 3.1. *The nonnegative Borel measure ν has the support* $\mathrm{supp}(\nu) = Q$ *and it satisfies the multivariate Carleman condition*

$$\sum_{d=0}^{\infty} \left(\int y_j^{2d} \,\mathrm{d}\nu(y) \right)^{-\frac{1}{2d}} = \infty \quad \textit{for} \quad j = 1, \ldots, m. \tag{3.1}$$

A measure ν satisfying (3.1) is known to be *determinate* (i.e., it is uniquely determined by its moments $\int y^\alpha \,\mathrm{d}\nu(y)$) by Nussbaum's theorem [Nuss], and it is *strictly determinate* (i.e., $\mathbb{R}[\mathtt{y}]$ is dense in $L^2(\mathbb{R}^m, \nu)$). See, e.g., [Smü17, Section 14.4] for details and proofs. It is interesting to remark that the Carleman condition (3.1) is automatically satisfied if $Q = \mathrm{supp}(\nu)$ is bounded.

3.1. **Density of SOS polynomials.**

In this subsection we prove the following strengthening of the above-mentioned Nussbaum theorem:

Proposition 3.2. *Let ν be a nonnegative Borel measure satisfying Assumption 3.1. Then SOS polynomials are dense in the cone of nonnegative functions in $L^2(\mathbb{R}^m, \nu)$.*

Proof. Suppose that the conclusion is not true. Then there exists a nonnegative function $\phi \in L^2(\mathbb{R}^m, \nu)$ that is not in the L^2-closure of the convex cone $\Sigma^2[\mathsf{y}]$. By the Hahn-Banach separation theorem (see [Bar02, Theorem III.3.4]), there is a continuous linear functional $\ell : L^2(\mathbb{R}^m, \nu) \to \mathbb{R}$ satisfying

$$\ell\big(\Sigma^2[\mathsf{y}]\big) \subseteq \mathbb{R}_+, \quad \ell(\phi) < 0. \tag{3.2}$$

By adding a small multiple of the linear functional $f \mapsto \int f \,\mathrm{d}\nu$ to ℓ, we can without loss of generality assume there exists $\varepsilon > 0$ such that

$$\ell(\sigma) \geq \varepsilon > 0 \quad \text{ for all } \sigma \in \Sigma^2[\mathsf{y}] \text{ with } \|\sigma\|_{L^2} = 1.$$

The Riesz representation theorem implies there is $h \in L^2(\mathbb{R}^m, \nu)$ such that

$$\ell(f) \;=\; \langle f, h\rangle_{L^2} \;=\; \int fh \,\mathrm{d}\nu$$

for all $f \in L^2(\mathbb{R}^m, \nu)$. Since Assumption 3.1 holds, $\mathbb{R}[\mathsf{y}]$ is dense in $L^2(\mathbb{R}^m, \nu)$ (see, e.g., [Smü17, Theorem 14.2, Section 14.4]). Hence there is a sequence of polynomials $\{p_n\}_{n=1}^{\infty} \subseteq \mathbb{R}[\mathsf{y}]$ that converges to h in the L^2-norm. Applying the Cauchy-Schwartz inequality yields

$$\Big|\langle f, h\rangle_{L^2} - \langle f, p_n\rangle_{L^2}\Big| = \Big|\langle f, h - p_n\rangle_{L^2}\Big| \leq \|f\|_{L^2}\, \|h - p_n\|_{L^2}.$$

Hence, for n large enough, the continuous linear functional

$$\ell_n : f \mapsto \langle f, p_n\rangle_{L^2} \tag{3.3}$$

also satisfies (3.2), i.e., ℓ_n is nonnegative on $\Sigma^2[\mathsf{y}]$ while it is negative at ϕ.

We now adapt the argument in [Smü17, Theorem 14.25] to show that $p_n \geq 0$ on $\operatorname{supp}(\nu)$. The restriction $\ell_n : \mathbb{R}[\mathsf{y}] \to \mathbb{R}$ is a positive linear functional and satisfies the multivariate Carleman condition (see [Smü17, Corollary 14.22]). Hence, by Nussbaum's theorem, ℓ_n is of the form

$$\ell_n(f) = \int f \,\mathrm{d}\tau, \quad \forall f \in \mathbb{R}[\mathsf{y}]$$

for some nonnegative Borel measure τ on $\mathbb{R}^m$. Set

$$M_+ \coloneqq \{y \in \mathbb{R}^m \mid p_n(y) \geq 0\}, \quad M_- \coloneqq \mathbb{R}^m \setminus M_+.$$

Let χ_+, χ_- denote the characteristic functions of M_+, M_- respectively. Then define positive Borel measures

$$\mathrm{d}\nu_+ = \chi_+ \mathrm{d}\nu, \quad \mathrm{d}\nu_- = \chi_- \mathrm{d}\nu, \quad \mathrm{d}\theta_+ = p_n \mathrm{d}\nu_+, \quad \mathrm{d}\theta_- = -p_n \mathrm{d}\nu_-.$$

By definition, $\nu = \nu_+ + \nu_-$, so

$$\int y_j^{2k} \,\mathrm{d}\nu_+(y) \leq \int y_j^{2k} \,\mathrm{d}\nu(y) \quad \text{ for all } j, k \in \mathbb{N},$$

whence ν_+ satisfies the Carleman condition (3.1). Hence, so does θ_+, again by [Smü17, Corollary 14.22]. In particular, the measure θ_+ is determinate. Since $\mathrm{d}\theta_+ - \mathrm{d}\theta_- = p_n \mathrm{d}\nu$, we have

$$\int y^\alpha \,\mathrm{d}\theta_+(y) = \int y^\alpha \,\mathrm{d}\theta_-(y) + \int y^\alpha \,\mathrm{d}\tau(y) = \int y^\alpha \,\mathrm{d}(\theta_- + \tau)(y).$$

Thus, by determinacy, $\theta_+ = \theta_- + \tau$. This yields

$$0 = \theta_+(M_-) \geq \theta_-(M_-) \geq 0,$$

so $\theta_-(M_-) = 0$ and $\theta_- = 0$.

Next, assume, for the sake of contradiction, that $p_n(y_0) < 0$ for some $y_0 \in \mathrm{supp}(\nu)$. Then $-p_n(y) \geq \delta > 0$ for all y in a small ball B around y_0. This yields the contradiction

$$0 = \theta_-(B) = \int (-p_n(y)) \,\mathrm{d}\nu_-(y) = \int (-p_n(y)) \,\mathrm{d}\nu(y) \geq \delta\nu(B) > 0,$$

so $p_n \geq 0$ on $\mathrm{supp}(\nu)$. Finally, this again leads to the contradiction

$$0 > \ell_n(\phi) = \int \phi p_n \,\mathrm{d}\nu \geq 0,$$

which completes the proof. □

3.2. **The Positivstellensatz.** Now we consider the set $K \subseteq \mathbb{R}^n$ as in (1.1). Since K is defined by the universal quantifier y in Q, one can equivalently write K as the intersection

$$K = \bigcap_{y \in Q} \{x \in \mathbb{R}^n : g_1(x, y) \geq 0, \ldots, g_s(x, y) \geq 0\}. \tag{3.4}$$

Clearly, K is closed since each g_i is a polynomial. If the quantifier set Q is semialgebraic, then so is K by Tarski's transfer principle [BCR98]. If Q is not semialgebraic, then K may not be semialgebraic.

For notational convenience, denote

$$g_0 := 1, \quad g := (g_0, g_1, \ldots, g_s).$$

For $f \in \mathbb{R}[\mathtt{x}]$, if there exist SOS polynomials $\tau_0, \tau_1, \ldots, \tau_s \in \Sigma^2[\mathtt{x}, \mathtt{y}]$ such that

$$f(\mathtt{x}) = \sum_{j=0}^{s} \int \tau_j(\mathtt{x}, y) g_j(\mathtt{x}, y) \,\mathrm{d}\nu(y) \tag{3.5}$$

then $f(x) \geq 0$ for all $x \in K$ since $\mathrm{supp}(\nu) = Q$ by Assumption 3.1. The set of all polynomials in $\mathbb{R}[\mathtt{x}]$ that can be represented as in (3.5) is

$$\mathrm{QM}[g, \nu] := \left\{ \sum_{j=0}^{s} \int \tau_j(\mathtt{x}, y) g_j(\mathtt{x}, y) \,\mathrm{d}\nu(y) \,\middle|\, \text{each } \tau_j \in \Sigma^2[\mathtt{x}, \mathtt{y}] \right\}. \tag{3.6}$$

The set $\mathrm{QM}[g,\nu]$ is a convex cone in $\mathbb{R}[\mathtt{x}]$. It is called the *quadratic module* associated to g and ν, since

$$1 \in \mathrm{QM}[g,\nu], \quad \mathrm{QM}[g,\nu] + \mathrm{QM}[g,\nu] \subseteq \mathrm{QM}[g,\nu],$$
$$\Sigma^2[\mathtt{x}] \cdot \mathrm{QM}[g,\nu] \subseteq \mathrm{QM}[g,\nu].$$

Apparently, all polynomials in $\mathrm{QM}[g,\nu]$ are nonnegative on K. The Positivstellensatz concerns the reverse of this implication. We start with the key Proposition 3.3 stating that the positivity domain of $\mathrm{QM}[g,\nu]$ is K.

Proposition 3.3. *Let ν be a nonnegative Borel measure satisfying Assumption 3.1, then we have*

$$K = \{x \in \mathbb{R}^n \mid f(x) \geq 0 \ \forall f \in \mathrm{QM}[g,\nu]\}.$$

Proof. By the definition (3.6), every polynomial in $\mathrm{QM}[g,\nu]$ is nonnegative on K, whence $K \subseteq \{x \in \mathbb{R}^n \mid \forall f \in \mathrm{QM}[g,\nu] : \ f(x) \geq 0\} =: \mathcal{D}$.

To establish the converse inclusion, assume $\hat{x} \notin K$. Then there is a $\hat{y} \in Q$ and a $j \in \{1,\dots,s\}$ such that $g_j(\hat{x},\hat{y}) < 0$. In a small open disk $B_{\varepsilon_1}(\hat{x},\hat{y})$ of radius $\varepsilon_1 > 0$ about $(\hat{x},\hat{y})$ in $\mathbb{R}^{n+m}$, $g_j(x,y) \leq -\lambda$ for some $\lambda > 0$. Consider a continuous function ϕ positive on the open ball $B_{\frac{\varepsilon_1}{2}}(\hat{y}) \subseteq \mathbb{R}^m$ and zero outside of the closed ball $\overline{B}_{\frac{\varepsilon_1}{2}}(\hat{y})$. Clearly,

$$\psi(\mathtt{x}) := \int_Q \phi(y) g_j(\mathtt{x},y) \,\mathrm{d}\nu(y) \in \mathbb{R}[\mathtt{x}]$$

is negative at $\hat{x}$.

By Proposition 3.2, there is a sequence $(\sigma_k)_k$ in $\Sigma^2[\mathtt{y}]$ that converges to ϕ in the L^2-norm. Hence, for each x, as $k \to \infty$, we have

$$\int_Q \sigma_k(y) g_j(x,y) \,\mathrm{d}\nu(y) \longrightarrow \int_Q \phi(y) g_j(x,y) \,\mathrm{d}\nu(y) = \psi(x).$$

In particular, for k large enough,

$$f(\mathtt{x}) := \int_Q \sigma_k(y) g_j(\mathtt{x},y) \,\mathrm{d}\nu(y) \in \mathrm{QM}[g,\nu]$$

is negative at $\hat{x}$. That is, $\hat{x} \notin \mathcal{D}$, whence $\mathcal{D} \subseteq K$ and we are done. □

3.3. **Bounded K.** In Positivstellensätze, we typically require that $f > 0$ on K and the quadratic module associated to K is archimedean. Since K is given by a universal quantifier over $y \in Q$, we form the quadratic module $\mathrm{QM}[g,\nu]$ and we assume it is *archimedean*, i.e., there exists an integer $N > 0$ such that

$$N - \mathtt{x}_1^2 - \cdots - \mathtt{x}_n^2 \ \in \ \mathrm{QM}[g,\nu].$$

Clearly, the archimedeanness of $\mathrm{QM}[g,\nu]$ implies that K is bounded (so it is compact since it is closed). Conversely, if K is bounded, we can generally assume $\mathrm{QM}[g,\nu]$ is archimedean, because one can add the inequality $N - \sum_{i=1}^n x_i^2 \geq 0$ (no y) to the description of the set K as in (1.1). The following is a generalization of the Putinar Positivstellensatz to sets given by universal quantifiers.

Theorem 3.4. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and assume the measure ν satisfies Assumption 3.1. Suppose* $\mathrm{QM}[g,\nu]$ *is archimedean. For a polynomial $f \in \mathbb{R}[\mathtt{x}]$, if $f > 0$ on K, then we have* $f \in \mathrm{QM}[g,\nu]$.

Proof. We shall apply Jacobi's [Jacobi] strengthening of Putinar's Positivstellensatz as presented in [Mar08, Chapter 5]. Consider the archimedean quadratic module $M = \mathrm{QM}[g,\nu]$. By Proposition 3.3, its positivity domain ($\mathcal{K}_M$ in Marshall's notation) is equal to K. Hence, the Jacobi-Putinar Positivstellensatz presented by Marshall in [Mar08, Theorem 5.4.4] implies that every polynomial positive on $\mathcal{K}_M = K$ belongs to the quadratic module $M = \mathrm{QM}[g,\nu]$. □

Theorem 3.4 clearly yields the following two corollaries.

Corollary 3.5. *Let $K, \mathrm{QM}[g,\nu]$ be as in Theorem 3.4 with ν satisfying Assumption 3.1. Then the following are equivalent for $f \in \mathbb{R}[\mathtt{x}]$:*

(i) $f \geq 0$ on K;

(ii) for all $\varepsilon > 0$, $f + \varepsilon \in \mathrm{QM}[g,\nu]$.

Corollary 3.6. *Let $K, \mathrm{QM}[g,\nu]$ be as in Theorem 3.4 with ν satisfying Assumption 3.1. Then the following are equivalent:*

(i) $K = \varnothing$;

(ii) $-1 \in \mathrm{QM}[g,\nu]$.

3.4. **The non-archimedean case.** When the quadratic module $\mathrm{QM}[g,\nu]$ is not archimedean (e.g., this is the case when K is unbounded), the conclusion of Theorem 3.4 may not hold. However, Proposition 3.3 allows us to get a perturbation type Positivstellensatz as in Lasserre-Netzer [LN07], for all (including unbounded) K. For $r \in \mathbb{N}$, denote

$$\Omega_r := \sum_{j=1}^{n} \sum_{k=0}^{r} \frac{\mathtt{x}_j^{2k}}{k!} \in \mathbb{R}[\mathtt{x}].$$

We now have the following Positivstellensatz.

Corollary 3.7. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and with ν satisfying Assumption 3.1. Then the following are equivalent for $f \in \mathbb{R}[\mathtt{x}]$:*

(i) $f \geq 0$ on K;

(ii) for all $\varepsilon > 0$, there exists $r \in \mathbb{N}$ such that $f + \varepsilon\Omega_r \in \mathrm{QM}[g,\nu]$.

Proof. We shall apply a strengthening of the Lasserre-Netzer perturbative Positivstellensatz [LN07, Corollary 3.6] proved in [KMV+] that can handle arbitrary constraints, and is proved as a corollary of more general results on "moment" polynomials, i.e., polynomials in $\mathtt{x}$ and their formal moments with regard to a probability measure.

Consider the constraint set $S = \mathrm{QM}[g,\nu]$. In the notation of [KMV+], $K(S) = K$ and $Q(S) = \mathrm{QM}[g,\nu]$. Now we simply apply [KMV+, Corollary 6.13] (polynomials nonnegative on $K(S)$ are up to a perturbation as in (ii) contained in $Q(S)$) to deduce Corollary 3.7. □

3.5. **Some illustrative examples.** In the following examples, the measure ν is the classical Lebesgue measure. Recall that $g_0 = 1$.

Example 3.8. Consider $f(\mathtt{x}) = -\mathtt{x}_1^3 - \mathtt{x}_2^3 + \frac{1}{9}\mathtt{x}_1^2\mathtt{x}_2 + \frac{1}{9}\mathtt{x}_1\mathtt{x}_2^2 + 8\mathtt{x}_1^2 + 8\mathtt{x}_2^2$ and the set K given as in (3.4) with

$$\begin{pmatrix} g_1(\mathtt{x},\mathtt{y}) \\ g_2(\mathtt{x},\mathtt{y}) \end{pmatrix} := \begin{pmatrix} 1 - \mathtt{x}_1^2\mathtt{y}_1^2 - \mathtt{x}_2^2\mathtt{y}_2^2 \\ \mathtt{x}_1\mathtt{y}_2^2 + \mathtt{x}_2\mathtt{y}_1^2 - 3\mathtt{x}_1\mathtt{x}_2\mathtt{y}_1\mathtt{y}_2 \end{pmatrix}$$

$$\text{and} \quad Q := \left\{ (y_1, y_2) \,\middle|\, \begin{array}{c} y_1 + y_2 \le 1, \\ y_1 \ge 0,\ y_2 \ge 0 \end{array} \right\}.$$

Note that the Lebesgue measure fulfills Assumption 3.1 when Q is compact. A Positivstellensatz certificate for $f \in \mathrm{QM}[g,\nu]$ is

$$f(\mathtt{x}) = \sum_{i=0}^{2} \int_0^1 \int_0^{1-y_2} \tau_i(\mathtt{x},y) g_i(\mathtt{x},y)\, \mathrm{d}y_1 \mathrm{d}y_2, \tag{3.7}$$

where the SOS polynomials $\tau_i(\mathtt{x},\mathtt{y})$ are

$$\begin{aligned} \tau_0 &= (2\mathtt{x}_1^2 - \mathtt{x}_1)^2 + (2\mathtt{x}_2^2 - \mathtt{x}_2)^2 + 5(\mathtt{x}_1 + \mathtt{x}_2)^2, \\ \tau_1 &= 60(\mathtt{x}_1\mathtt{y}_1 - \mathtt{x}_2\mathtt{y}_2)^2, \\ \tau_2 &= 20(\mathtt{x}_1\mathtt{y}_2 - \mathtt{x}_2\mathtt{y}_1)^2 + 4(\mathtt{x}_1^2 + \mathtt{x}_2^2). \end{aligned}$$

One can check the representation (3.7) by a direct evaluation of integrals there.

We remark that a Positivstellensatz certificate for $f \in \mathrm{QM}[g,\nu]$ can be computed numerically by solving a semidefinite program. The following is such an example.

Example 3.9. Consider $f(\mathtt{x}) = \mathtt{x}_1^2\mathtt{x}_2 - \mathtt{x}_1\mathtt{x}_2^2 + \mathtt{x}_1^2 + \mathtt{x}_2$ and the set K given as in (3.4) with

$$\begin{pmatrix} g_1(\mathtt{x},\mathtt{y}) \\ g_2(\mathtt{x},\mathtt{y}) \end{pmatrix} := \begin{pmatrix} \mathtt{x}_1^2\mathtt{y}_2 + \mathtt{x}_2\mathtt{y}_1^2 - \mathtt{x}_1 + \mathtt{x}_2 - \mathtt{y}_1 \\ \mathtt{x}_2\mathtt{y}_2^2 - \mathtt{x}_2^2\mathtt{y}_1 - \mathtt{x}_1\mathtt{y}_2 + \mathtt{x}_2\mathtt{y}_1 \end{pmatrix}$$

$$\text{and} \quad Q = \{(y_1, y_2) : |y_1| + |y_2| \le 1\}.$$

Notice that the Lebesgue measure fulfills Assumption 3.1 since Q is compact. A Positivstellensatz certificate $f \in \mathrm{QM}[g,\nu]$ is

$$f(\mathtt{x}) = \sum_{i=0}^{2} \int_Q \tau_i(\mathtt{x},y) g_i(\mathtt{x},y)\, \mathrm{d}y, \tag{3.8}$$

where $\tau_i(\mathtt{x},\mathtt{y})$ are SOS polynomials. We can represent them as

$$\tau_0(\mathtt{x}) = [\mathtt{x}]_1^T X_0 [\mathtt{x}]_1, \quad \tau_1(\mathtt{x},\mathtt{y}) = [\mathtt{x},\mathtt{y}]_1^T X_1 [\mathtt{x},\mathtt{y}]_1, \quad \tau_2(\mathtt{x},\mathtt{y}) = [\mathtt{x},\mathtt{y}]_1^T X_2 [\mathtt{x},\mathtt{y}]_1,$$

where X_0, X_1, X_2 are symmetric positive semidefinite matrices. By comparing coefficients of monomials of $\mathtt{x}$ in both sides of (3.8), we get a set of linear equations on X_0, X_1, X_2.

The matrices X_0, X_1, X_2 satisfying these conditions can be found by solving a semi-definite program. By using the software `SeDuMi`, we obtained that

$$\tau_0 = \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \end{bmatrix}^T \begin{bmatrix} 0.0288 & 0.0988 & -0.0265 \\ 0.0988 & 0.3385 & -0.0909 \\ -0.0265 & -0.0909 & 0.0244 \end{bmatrix} \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \end{bmatrix},$$

$$\tau_1 = \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \\ \mathtt{y}_1 \\ \mathtt{y}_2 \end{bmatrix}^T \begin{bmatrix} 0.0905 & -0.0988 & 0.0455 & 0.0865 & -0.2965 \\ -0.0988 & 0.1080 & -0.0497 & -0.0945 & 0.3239 \\ 0.0455 & -0.0497 & 0.0229 & 0.0435 & -0.1492 \\ 0.0865 & -0.0945 & 0.0435 & 0.0827 & -0.2835 \\ -0.2965 & 0.3239 & -0.1492 & -0.2835 & 0.9717 \end{bmatrix} \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \\ \mathtt{y}_1 \\ \mathtt{y}_2 \end{bmatrix},$$

$$\tau_2 = \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \\ \mathtt{y}_1 \\ \mathtt{y}_2 \end{bmatrix}^T \begin{bmatrix} 0.3787 & 0.4577 & 0.0895 & 0.5813 & -0.1114 \\ 0.4577 & 1.9459 & -0.0505 & 1.0328 & -0.1879 \\ 0.0895 & -0.0505 & 0.0392 & 0.0998 & -0.0203 \\ 0.5813 & 1.0328 & 0.0998 & 0.9704 & -0.1836 \\ -0.1114 & -0.1879 & -0.0203 & -0.1836 & 0.0348 \end{bmatrix} \begin{bmatrix} 1 \\ \mathtt{x}_1 \\ \mathtt{x}_2 \\ \mathtt{y}_1 \\ \mathtt{y}_2 \end{bmatrix}.$$

The above matrices in the middle are all positive semidefinite. For neatness, only four decimal digits are shown (the errors for matching coefficients are in the order of 10^{-11}).

Example 3.10. Consider $f(\mathtt{x}) = 4 - \mathtt{x}_1^2 - \mathtt{x}_2^2$ and the set K given as in (3.4) with

$$g_1(\mathtt{x}, \mathtt{y}) \coloneqq 4\mathtt{y}^4 - 1 - \mathtt{y}^2(2\mathtt{y}^2 - 1)\mathtt{x}_1^2 - \mathtt{y}^2(2\mathtt{y}^2 + 1)\mathtt{x}_2^2$$

$$\text{and} \quad Q = \mathbb{Z}_+ = \{1, 2, \ldots\}.$$

We select the measure ν supported on Q such that

$$\nu(\{k\}) = \frac{1}{e} \cdot \frac{1}{k!}, \quad \text{for} \quad k = 1, 2, \ldots.$$

One can directly calculate that

$$\int y^j \mathtt{d}\nu(y) = \frac{1}{e} \sum_{k=1}^{\infty} \frac{k^j}{k!} = \begin{cases} 1 - \frac{1}{e} & \text{if} \quad j = 0, \\ B_j & \text{if} \quad j \geq 1. \end{cases}$$

In the above, B_j denotes the jth Bell number ([FS09, Section II.3] or [Sta12, p. 82]), which counts the number of partitions of the set $[j] = \{1, \ldots, j\}$. It is interesting to remark that $B_j \leq j!$, which can be seen as follows. We assign to each partition $[j] = S_1 \cup \cdots \cup S_k$ a different permutation as follows: sort each S_i in increasing order, and relabel S_i so that the lowest number in S_i is smaller than the lowest number in S_{i+1}. Then each S_i yields a permutation when it is viewed as a cycle, and the product of the disjoint cycles assigned to the S_i's is a permutation of $[j]$. We now claim that ν satisfies the Carleman condition (3.1). Indeed, by Stirling's approximation for factorials (see, e.g., [FS09, Section I.2] or [**?**]), $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, we have

$$\sum_{d=1}^{\infty} \left(\int y^{2d} \mathtt{d}\nu(y) \right)^{-\frac{1}{2d}} \geq \sum_{d=1}^{\infty} \left((2d)!\right)^{-\frac{1}{2d}} \sim \sum_{d=1}^{\infty} \frac{e}{2d} \left(\sqrt{4\pi d}\right)^{-\frac{1}{2d}} = \infty.$$

The set K is the intersection of infinitely many ellipses. The Positivstellensatz certificate $f \in \mathrm{QM}[g,\nu]$ is

$$f(\mathtt{x}) = \int_Q \tau_0(\mathtt{x},\mathtt{y})\mathrm{d}\mathtt{y} + \int_Q \tau_1(\mathtt{x},\mathtt{y})g_1(\mathtt{x},\mathtt{y})\,\mathrm{d}\mathtt{y}$$

where $\tau_0(\mathtt{x},\mathtt{y}), \tau_1(\mathtt{x},\mathtt{y})$ are SOS polynomials. By solving a semidefinite program, we can get

$$\tau_0 \approx 0.1671 + 0.6336\mathtt{y} + 0.0990\mathtt{x}_2^2 + 0.6005\mathtt{y}^2,$$
$$\tau_1 \approx 0.0053 - 0.0100\mathtt{y} + 0.0047\mathtt{y}^2.$$

For neatness, only four decimal digits are shown in the above (the errors for matching coefficients are in the order of 10^{-12}).

In Example 3.10 the set K is compact. Now consider the same $Q = \mathbb{Z}_+$ and the measure ν as in the above example. If $g(\mathtt{x},\mathtt{y}) = \mathtt{x}_2 - 2\mathtt{y}\mathtt{x}_1 + \mathtt{y}^2$, then the set $K \supseteq \{x_1 \le 0, x_2 \ge 0\} \cup \{x_1 \ge 0, x_2 - x_1^2 \ge 0\}$ is an unbounded convex region, and the quadratic module $\mathrm{QM}[g,\nu]$ cannot be archimedean.

4. Moment problems with universal quantifiers

This section considers K-moment problems for the set K given by a universal quantifier as in (1.1). Recall from the introduction that a multisequence $z = (z_\alpha)_{\alpha\in\mathbb{N}^n}$ gives rise to the Riesz functional

$$\mathscr{R}_z : \mathbb{R}[\mathtt{x}] \to \mathbb{R}, \quad \mathtt{x}^\alpha \mapsto z_\alpha. \tag{4.1}$$

Equivalently,

$$\mathscr{R}_z\Big(\sum_\alpha f_\alpha \mathtt{x}^\alpha\Big) = \sum_\alpha f_\alpha z_\alpha,$$

where $f_\alpha \in \mathbb{R}$ are real coefficients. If μ is a representing measure for z, then

$$\mathscr{R}_z(f) = \int f(x)\,\mathrm{d}\mu(x) \quad \text{for all} \quad f \in \mathbb{R}[\mathtt{x}].$$

If $\mathrm{supp}(\mu) \subseteq K$, then z must satisfy

$$\mathscr{R}_z(f) \ \ge\ 0 \quad \text{for all} \quad f \in \mathbb{R}[\mathtt{x}] :\ f|_K \ge 0. \tag{4.2}$$

The multisequence z is said to be K-positive if (4.2) holds. Clearly, being K-positive is a necessary condition for z to have a K-representing measure. When the set K is closed (this is the case if K is given as in (1.1)), being K-positive as in (4.2) is also sufficient for z to be a K-moment sequence. This is a classical result of Riesz and Haviland. The reader is referred to the surveys [Ber87, Fia16] and books [Akh65, Smü17] for more details about classical moment problems.

Using the quadratic module $\mathrm{QM}[g,\nu]$ introduced in (3.6), we have the following characterization of a K-moment sequence.

Theorem 4.1. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and suppose the measure ν satisfies Assumption 3.1. Assume the quadratic module* $\mathrm{QM}[g,\nu]$ *is archimedean. Then the multisequence z is a K-moment sequence if and only if the Riesz functional $\mathscr{R}_z \ge 0$ on* $\mathrm{QM}[g,\nu]$*, i.e., $\mathscr{R}_z(f) \ge 0$ for all $f \in$* $\mathrm{QM}[g,\nu]$.

Proof. (⇒) If $f \in \mathrm{QM}[g,\nu]$, then $f|_K \geq 0$. Hence (4.2) implies $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$. Conversely, if $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$, then $\mathscr{R}_z$ is also nonnegative on each $f \in \mathbb{R}[\mathtt{x}]$ that is nonnegative on K by Theorem 3.4 or Corollary 3.5. The implication (⇐) now follows by the Riesz-Haviland theorem mentioned above. □

As pointed out by one of the referees, an alternate proof of Theorem 4.1 can be given using the Jacobi-Putinar Positivstellensatz, cf. [Mar08, Section 5.6].

Corollary 4.2. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and suppose the measure ν satisfies Assumption 3.1. Then the multisequence z is a K-moment sequence if and only if the Riesz functional $\mathscr{R}_z$ satisfies $\mathscr{R}_z(f) \geq 0$ for all $f \in \mathbb{R}[\mathtt{x}]$ with the following property: for all $\varepsilon > 0$ there exists $r \in \mathbb{N}$ with $f + \varepsilon\Omega_r \in \mathrm{QM}[g,\nu]$.*

Proof. (⇒) is obvious since a polynomial f satisfying the perturbation condition in the statement of the corollary, is nonnegative on K (cf. Proposition 3.3 or Corollary 3.7). For the converse implication (⇐) note that f is nonnegative on K if and only if it satisfies this perturbation condition (again by Corollary 3.7). The conclusion now follows by the Riesz-Haviland theorem. □

Remark 4.3. As pointed out by one of the referees, a strengthening of Theorem 4.1 holds. Namely, assume $K \subseteq \mathbb{R}^n$ is as in (1.1), and the measure ν satisfies Assumption 3.1. Then, if the multisequence z itself satisfies the multivariate Carleman condition, then it is a K-moment sequence iff $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$. Indeed, the proof is essentially the same as that of Theorem 4.1, but at the final step one applies [IKKM22, Theorem 3.16], a far reaching extension of the Nussbaum theorem.

In the sequel, we determine concrete conditions on z for $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$.

4.1. **Moment and localizing matrices.** The multisequence $z = (z_\alpha)_{\alpha\in\mathbb{N}^n}$ gives rise to the infinite matrix

$$\mathrm{H}[z] \,:=\, (z_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}^n}.$$

That is, $\mathrm{H}[z]$ is the matrix labelled by nonnegative integer vectors $\alpha, \beta \in \mathbb{N}^n$ and

$$\mathrm{H}[z]_{\alpha,\beta} \,=\, z_{\alpha+\beta}$$

for all α, β. It is called the *moment matrix* or *multivariate Hankel matrix* of the multisequence z. For a vector $\mathbf{u} = (u_\alpha)_{\alpha\in\mathbb{N}^n}$ with finitely many nonzero entries, we have

$$\mathbf{u}^T\mathrm{H}[z]\mathbf{u} = \mathscr{R}_z\Big(u(\mathtt{x})^2\Big), \quad \text{where} \quad u(\mathtt{x}) = \sum_\alpha u_\alpha \mathtt{x}^\alpha.$$

Hence, if $\mathscr{R}_z \geq 0$ on $\Sigma^2[\mathtt{x}]$, then $\mathrm{H}[z] \succeq 0$. For a degree k, we denote the truncation

$$\mathrm{H}^{(k)}[z] \,:=\, (z_{\alpha+\beta})_{\alpha,\beta\in\mathbb{N}^n_k}.$$

One can easily verify that $\mathrm{H}^{(k)}[z] \succeq 0$ if $\mathscr{R}_z \geq 0$ on $\Sigma^2[\mathtt{x}] \cap \mathbb{R}[\mathtt{x}]_{2k}$.

Next we give localizing matrices for the quadratic module $\mathrm{QM}[g,\nu]$. For a given multisequence z, $\mathscr{R}_z\Big(\int p(\mathtt{x},y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big)$ is a quadratic form in the vector of coefficients of

$p(\mathtt{x},\mathtt{y})$. For convenience, we use $\mathbf{p}$ to denote the vector of coefficients of $p(\mathtt{x},\mathtt{y})$. Let $\mathrm{L}_{\nu,g_j}^{(k,l)}[z]$ be the matrix associated to this quadratic form. Here superscripts k,l denote degree bounds on $\mathtt{x}$ and $\mathtt{y}$, respectively, so that

$$\mathscr{R}_z\Big(\int p(\mathtt{x},y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big) \;=\; \mathbf{p}^T\Big(\mathrm{L}_{\nu,g_j}^{(k,l)}[z]\Big)\mathbf{p},$$

for all $p(\mathtt{x},\mathtt{y})\in\mathbb{R}[\mathtt{x},\mathtt{y}]$ with degrees

$$\deg_{\mathtt{x}}(p(\mathtt{x},\mathtt{y})^2 g_j(\mathtt{x},\mathtt{y}))\le 2k,\quad \deg_{\mathtt{y}}(p(\mathtt{x},\mathtt{y})^2 g_j(\mathtt{x},\mathtt{y}))\le 2l. \tag{4.3}$$

Explicit expressions for $\mathrm{L}_{\nu,g_j}^{(k,l)}[z]$ can be given as follows. For convenience, denote

$$k' \coloneqq k-\lceil \deg_{\mathtt{x}}(g_j(\mathtt{x},\mathtt{y}))/2\rceil,\quad l' \coloneqq l-\lceil \deg_{\mathtt{y}}(g_j(\mathtt{x},\mathtt{y}))/2\rceil. \tag{4.4}$$

Then we can write

$$p(\mathtt{x},\mathtt{y}) = \mathbf{p}^T([\mathtt{x}]_{k'}\otimes[\mathtt{y}]_{l'})$$

where $[\mathtt{x}]_k$ denotes the vector of all monomials in $\mathtt{x}$ of degrees at most k, and likewise for $[\mathtt{y}]_l$. The constraining polynomial $g_j(\mathtt{x},\mathtt{y})$ can be written in the form

$$g_j(\mathtt{x},\mathtt{y}) = \sum_i g_{ji}(\mathtt{x})h_{ji}(\mathtt{y}),$$

for some polynomials $g_{ji}\in\mathbb{R}[\mathtt{x}]$ and $h_{ji}\in\mathbb{R}[\mathtt{y}]$. Then, one can see that

$$\begin{aligned}
&\mathscr{R}_z\Big(\textstyle\int p(\mathtt{x},y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big)\\
=\;& \mathbf{p}^T\Big(\mathscr{R}_z \textstyle\int g_j(\mathtt{x},y)([\mathtt{x}]_{k'}\otimes[y]_{l'})([\mathtt{x}]_{k'}\otimes[y]_{l'})^T\,\mathrm{d}\nu(y)\Big)\mathbf{p}\\
=\;& \mathbf{p}^T\Big(\mathscr{R}_z \textstyle\int g_j(\mathtt{x},y)([\mathtt{x}]_{k'}[\mathtt{x}]_{k'}^T)\otimes[y]_{l'}[y]_{l'}^T\,\mathrm{d}\nu(y)\Big)\mathbf{p}\\
=\;& \mathbf{p}^T\Big(\sum_i\big(\textstyle\int h_{ji}(y)[y]_{l'}[y]_{l'}^T\mathrm{d}\,\nu(y)\big)\otimes\mathscr{R}_z\big(g_{ji}(\mathtt{x})[\mathtt{x}]_{k'}[\mathtt{x}]_{k'}^T\big)\Big)\mathbf{p}.
\end{aligned}$$

(In the above, when $\mathscr{R}_z$ is applied to a matrix, it means that it is applied entrywise, for convenience of notation.) Denote the matrices

$$\mathrm{Y}_{\nu,h_{ji}}^{(l')} \;\coloneqq\; \int h_{ji}(y)[y]_{l'}[y]_{l'}^T\,\mathrm{d}\nu(y),\quad \mathrm{L}_{g_{ji}}^{(k')}[z] \;\coloneqq\; \mathscr{R}_z\Big(g_{ji}(\mathtt{x})[\mathtt{x}]_{k'}[\mathtt{x}]_{k'}^T\Big). \tag{4.5}$$

Then we get the expression

$$\mathrm{L}_{\nu,g_j}^{(k,l)}[z] \;\coloneqq\; \sum_i \mathrm{Y}_{\nu,h_{ji}}^{(l')}\otimes\mathrm{L}_{g_{ji}}^{(k')}[z]. \tag{4.6}$$

Note that k',l' are the degrees defined in (4.4). Observe that $\mathrm{L}_{g_{ji}}^{(k')}[z]$ is the localizing matrix for the polynomial $g_{ji}\in\mathbb{R}[\mathtt{x}]$, and is independent of ν. Similarly, the matrices $\mathrm{Y}_{\nu,h_{ji}}^{(l')}$ are independent of z. In particular, for $g_0=1$, we get

$$\mathrm{L}_{\nu,1}^{(k,l)}[z] \;=\; \Big(\int [y]_l[y]_l^T\mathrm{d}\nu(y)\Big)\otimes\mathrm{H}^{(k)}[z]. \tag{4.7}$$

4.2. **The full moment problem.** We give a full characterization for K-moment sequences when K is defined by universal quantifiers.

Theorem 4.4. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and assume the measure ν satisfies Assumption 3.1. Then, for a multisequence z, we have $\mathscr{R}_z \geq 0$ on* $\mathrm{QM}[g,\nu]$ *if and only if for all $j = 0, 1, \ldots, s$,*

$$\mathrm{L}_{\nu,g_j}^{(k,l)}[z] \succeq 0, \quad k = 1, 2, \ldots, \; l = 1, 2, \ldots. \tag{4.8}$$

Moreover, when $\mathrm{QM}[g,\nu]$ *is archimedean, then z is a K-moment sequence if and only if it satisfies* (4.8).

Proof. Observe that $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$ if and only if

$$\mathscr{R}_z\Big(\int p(\mathtt{x}, y)^2 g_j(\mathtt{x}, y)\, \mathrm{d}\nu(y)\Big) \geq 0$$

for all j and for all $p(\mathtt{x}, \mathtt{y}) \in \mathbb{R}[\mathtt{x}, \mathtt{y}]$. When p is restricted to have degrees as in (4.3), then (4.8) follows from the definition of $\mathrm{L}_{\nu,g_j}^{(k,l)}[z]$ for all k and l. When $\mathrm{QM}[g,\nu]$ is archimedean, the last statement follows from Theorem 4.1. □

When K is given without quantifiers, there is a classical flat extension theorem [CF96, CF05] that recognizes K-moment sequences. Here, we give a similar flat extension theorem for sets K defined with universal quantifiers. Let

$$d_g \coloneqq \max\{1, \deg_{\mathtt{x}}(g)\}. \tag{4.9}$$

Theorem 4.5. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and assume the measure ν satisfies Assumption 3.1. Let z be a multisequence satisfying* (4.8). *If there exists $k \geq d_g$ such that*

$$r \coloneqq \operatorname{rank} \mathrm{H}^{(k-d_g)}[z] \;=\; \operatorname{rank} \mathrm{H}^{(k)}[z],$$

then z admits an r-atomic measure μ supported in K and μ is the unique representing measure for z.

Proof. By the flat extension theorem [CF96, CF05], we know that z admits an r-atomic representing measure, say, μ. Moreover, the μ is the unique representing measure for z. Since z satisfies (4.8), $\mathscr{R}_z \geq 0$ on $\mathrm{QM}[g,\nu]$. Pick an arbitrary $f \in \mathrm{QM}[g,\nu]$, then (4.8) implies that

$$\mathrm{L}_f^{(k)}[z] \succeq 0$$

for all $k = 1, 2, \ldots$. As in [CF96, CF05], we have $\mathrm{supp}(\mu) \subseteq \{x : f(x) \geq 0\}$. As this holds for all $f \in \mathrm{QM}[g,\nu]$, Proposition 3.3 implies that $\mathrm{supp}(\mu) \subseteq K$. □

4.3. **The truncated moment problem.** Now we consider $w = (w_\alpha)_{\alpha \in \mathbb{N}^n_{2d}}$, a truncated multisequence of even degree $2d$. We look for concrete conditions under which w is a K-moment sequence, with a representing measure μ supported in K. As in the above calculations, for w to be a K-moment sequence, it must satisfy

$$\mathscr{R}_w\Big(\int p(\mathtt{x}, y)^2 g_j(\mathtt{x}, y)\, \mathrm{d}\nu(y)\Big) \geq 0 \tag{4.10}$$

for all j and for all $p(\mathtt{x},\mathtt{y}) \in \mathbb{R}[\mathtt{x},\mathtt{y}]$ with the degree

$$\deg_{\mathtt{x}} \left(p(\mathtt{x},\mathtt{y})^2 g_j(\mathtt{x},\mathtt{y})\right) \le 2d.$$

Note that (4.10) is equivalent to

$$\mathrm{L}_{\nu,g_j}^{(k,l)}[w] \succeq 0, \quad k = 1,2,\dots d, \;\; l = 1,2,\dots. \tag{4.11}$$

The following is a generalization of the flat extension theorem in [CF96, CF05].

Theorem 4.6. *Let $K \subseteq \mathbb{R}^n$ be as in* (1.1) *and assume the measure ν satisfies Assumption 3.1. Let $w \in \mathbb{R}^{\mathbb{N}^n_{2d}}$ be a truncated multisequence satisfying* (4.11). *If there exists a positive integer $k \le d - d_g$ such that*

$$r \,:=\, \operatorname{rank} \mathrm{H}^{(k)}[w] \,=\, \operatorname{rank} \mathrm{H}^{(d)}[w], \tag{4.12}$$

then w admits an r-atomic measure μ supported in K and μ is the unique representing measure for w.

Proof. By the flatness condition (4.12), the truncated multisequence w can be extended to a full multisequence $z = (z_\alpha)_{\alpha\in\mathbb{N}^n}$ with $\operatorname{rank} \mathrm{H}[z] = r$ that represents an r-atomic measure μ (see Corollary 1.4 of [Lau05]). Moreover, μ is the unique representing measure for w (and also for z). This can be implied by Theorems 1.2 and 1.6 of [Lau05] (also see [CF96]). It now remains to show that $\operatorname{supp}(\mu) \subseteq K$. For all $a(\mathtt{x}) \in \mathbb{R}[\mathtt{x}]_d$ and $b(\mathtt{y}) \in \mathbb{R}[\mathtt{y}]_l$, it holds that

$$\mathscr{R}_w\Big(a(\mathtt{x})^2 \int b(y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big) = \mathscr{R}_w\Big(\int a(\mathtt{x})^2 b(y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big) \ge 0.$$

Since z is an extension of w, which is represented by μ, we get

$$\mathscr{R}_z\Big(f(\mathtt{x})^2 \int b(y)^2 g_j(\mathtt{x},y)\,\mathrm{d}\nu(y)\Big) \ge 0$$

for all $f \in \mathbb{R}[\mathtt{x}]$. This implies that for each j,

$$\operatorname{supp}(\mu) \subseteq \left\{ x \in \mathbb{R}^n \Big| \int b(y)^2 g_j(x,y)\,\mathrm{d}\nu(y) \ge 0 \right\}.$$

The above is true for all $b \in \mathbb{R}[\mathtt{y}]_l$ and $l = 1,2,\dots$. Hence, as in the proof of Proposition 3.3, we can infer that the intersection over j of the right-hand side sets in the above equation is equal to K. That is, $\operatorname{supp}(\mu) \subseteq K$, which completes the proof. □

We remark that the rank condition (4.12) implies that the truncated multisequence w admits a unique r-atomic representing measure μ, say, $w = \lambda_1[u_1]_{2d} + \cdots + \lambda_r[u_r]_{2d}$, for distinct points $u_1,\dots,u_r$ and positive scalars $\lambda_1,\dots,\lambda_r$, as in [CF96, CF05]. The condition (4.11) ensures that all $u_1,\dots,u_r \in K$. Note that (4.11) requires it to hold for all $l = 1,2,\dots$. If this is not checkable, one can verify $u_i \in K$ by checking nonnegativity of $g(u_i,\mathtt{y})$ on Q. The following is such an example.

Example 4.7. For the set

$$K = \left\{ x \in \mathbb{R}^2 \mid 1 - x^T y \geq 0 \quad \forall y \in \mathbb{R}^2 : y_1^4 + y_2^4 \leq 1 \right\},$$

we consider the truncated multisequence $w \in \mathbb{R}^{\mathbb{N}_4^2}$ given such that

$$\mathrm{H}^{(2)}[w] = \begin{bmatrix} 3 & 0 & \frac{2}{3} & \frac{2}{3} & -\frac{5}{18} & \frac{17}{18} \\ 0 & \frac{2}{3} & -\frac{5}{18} & -\frac{2}{9} & \frac{13}{54} & \frac{7}{108} \\ \frac{2}{3} & -\frac{5}{18} & \frac{17}{18} & \frac{13}{54} & \frac{7}{108} & \frac{8}{27} \\ \frac{2}{3} & -\frac{2}{9} & \frac{13}{54} & \frac{2}{9} & -\frac{23}{162} & \frac{61}{324} \\ -\frac{5}{18} & \frac{13}{54} & \frac{7}{108} & -\frac{23}{162} & \frac{61}{324} & -\frac{17}{648} \\ \frac{17}{18} & \frac{7}{108} & \frac{8}{27} & \frac{61}{324} & -\frac{17}{648} & \frac{209}{648} \end{bmatrix}.$$

One can check that $\operatorname{rank} \mathrm{H}^{(1)}[w] = \operatorname{rank} \mathrm{H}^{(2)}[w] = 3$, so the condition (4.12) of Theorem 4.6 holds. As in [CF05], we obtain $w = [u_1]_4 + [u_2]_4 + [u_3]_4$ for the points

$$u_1 = \left(-\frac{2}{3}, \frac{1}{2} \right), \quad u_2 = \left(\frac{1}{3}, \frac{2}{3} \right), \quad u_3 = \left(\frac{1}{3}, -\frac{1}{2} \right).$$

It is easily seen (e.g., by Hölder's inequality) that these three points belong to the set K.

5. Semi-Infinite Optimization

An important application of Positivstellensätze and moment problems with universal quantifiers is to solve semi-infinite optimization. Consider the semi-infinite program (SIP):

$$\begin{cases} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & g(x,y) \geq 0 \quad \forall y \in Q. \end{cases} \tag{5.1}$$

The constraining function g in (5.1) is the s-dimensional vector of polynomials,

$$g(\mathtt{x}, \mathtt{y}) \coloneqq \big(g_1(\mathtt{x}, \mathtt{y}), \ldots, g_s(\mathtt{x}, \mathtt{y}) \big),$$

$f \in \mathbb{R}[\mathtt{x}]$, and $X \subseteq \mathbb{R}^n$ is another given constraining set that does not depend on $y \in \mathbb{R}^m$. We assume X is given as

$$X = \{ x \in \mathbb{R}^n \mid c_i(x) = 0 \, (i \in \mathcal{I}), \; c_j(x) \geq 0 \, (j \in \mathcal{J}) \}. \tag{5.2}$$

Here, all c_i, c_j are polynomials in $\mathtt{x}$ and $\mathcal{I}$, $\mathcal{J}$ are disjoint finite label sets. For convenience of notation, we denote the polynomial tuples:

$$c_{eq} = (c_i)_{i \in \mathcal{I}}, \quad c_{in} = (c_j)_{j \in \mathcal{J}}.$$

Semi-infinite optimization has broad applications, such as Chebyshev approximation [LS07] and robustness support vector machines [XCM09]. Classical methods for solving semi-infinite optimization include Karush–Kuhn–Tucker multipliers [SS12], discretization methods

[DM17], and Moment-SOS relaxations [HuN23, WG14]. In this section, we show how to use Positivstellensätze and moment problems with universal quantifiers to solve SIPs.

As before, we let ν be a nonnegative Borel measure on $\mathbb{R}^m$ satisfying Assumption 3.1. We assume the moments $\int_Q y^\alpha \, \mathrm{d}\nu(y)$ are available. Then truncations for given degrees of the quadratic module $\mathrm{QM}[g,\nu]$ can be represented by semidefinite programs.

Proposition 5.1. *Let K be as in (1.1) and let ν be a Borel measure satisfying Assumption 3.1. Assume that the quadratic module $\mathrm{QM}[g,\nu]+\mathrm{Ideal}[c_{eq}]+\mathrm{QM}[c_{in}]$ is archimedean. Then the optimal value $f_{\min}$ of (5.1) is equal to the optimal value of the following optimization problem*

$$\left\{\begin{array}{ll} \min & \mathscr{R}_z(f) \\ \text{s.t.} & \mathscr{R}_z \geq 0 \quad \textit{on} \quad \mathrm{QM}[g,\nu]+\mathrm{Ideal}[c_{eq}]+\mathrm{QM}[c_{in}], \\ & \mathscr{R}_z(1)=1, \quad z\in\mathbb{R}^{\mathbb{N}^n}. \end{array}\right. \tag{5.3}$$

Proof. The feasible set of (5.1) is $X\cap K$, where K is as in (1.1). Note that

$$X\cap K = \left\{ x\in\mathbb{R}^n \;\middle|\; \begin{bmatrix} c_{eq}(x) \\ -c_{eq}(x) \\ c_{in}(x) \\ g(x,y) \end{bmatrix} \geq 0 \ \forall y\in Q \right\}.$$

The polynomials c_i can also be viewed as depending on $\mathtt{y}$ trivially. Observe that

$$\mathrm{QM}[(c_{eq},-c_{eq},c_{in},g),\,\nu] \;=\; \mathrm{QM}[g,\nu]+\mathrm{Ideal}[c_{eq}]+\mathrm{QM}[c_{in}].$$

Since $\mathrm{QM}[g,\nu]+\mathrm{Ideal}[c_{eq}]+\mathrm{QM}[c_{in}]$ is archimedean, the set $X\cap K$ is bounded (cf. Proposition 3.3). Thus the optimal value $f_{\min}$ is finite, i.e., $f_{\min}\in\mathbb{R}$. Hence, $f_{\min}$ equals the minimum value of the expectation $\int_{X\cap K} f(x)\,\mathrm{d}\mu(x)$, over all probability measures μ supported in $X\cap K$. When z is a multisequence satisfying the constraints in (5.3), Theorem 4.1 implies that z is the moment sequence of such a probability measure μ. Therefore, $f_{\min}$ is also the minimum value of (5.3). □

Proposition 5.1 can be used to give Moment-SOS type relaxations for solving the semi-infinite optimization (5.1). The full multisequence $z\in\mathbb{R}^{\mathbb{N}^n}$ can be approximated by its truncations

$$w=(z_\alpha)_{\alpha\in\mathbb{N}^n_{2k}},$$

for a degree k. Note that $\mathscr{R}_w\geq 0$ on $\mathrm{QM}[g,\nu]_{2k}$ if and only if

$$\mathrm{L}^{(k,l)}_{\nu,g_j}[w]\succeq 0$$

for all $l=1,2,\ldots$ (cf. Theorem 4.4). The constraining polynomials c_j do not depend on $\mathtt{y}$, so

$$\mathrm{L}^{(k,l)}_{\nu,c_j}[w] \;=\; \Big(\int 1\,\mathrm{d}\nu(y)\Big)\cdot \mathrm{L}^{(k)}_{c_j}[w].$$

In computational practice, we typically scale ν so that $\int 1\,\mathrm{d}\nu(y)=1$, whence $\mathrm{L}^{(k,l)}_{\nu,c_j}[w]=\mathrm{L}^{(k)}_{c_j}[w]$. It is also interesting to note that $\mathscr{R}_z\geq 0$ on $\mathrm{QM}[g,\nu]+\mathrm{Ideal}[c_{eq}]+\mathrm{QM}[c_{in}]$ if and only if $\mathscr{R}_z \geq 0$ on each of the $\mathrm{QM}[g,\nu]$, $\mathrm{Ideal}[c_{eq}]$, $\mathrm{QM}[c_{in}]$. Moreover, $\mathscr{R}_z \geq 0$ on $\mathrm{Ideal}[c_{eq}]$ if and only if $\mathscr{R}_z \equiv 0$ on $\mathrm{Ideal}[c_{eq}]$, since $\mathrm{Ideal}[c_{eq}]$ is a subspace of $\mathbb{R}[\mathtt{x}]$. Note that $\mathscr{R}_z \equiv 0$ on $\mathrm{Ideal}[c_{eq}]_{2k}$ is equivalent to $\mathrm{L}_{c_i}^{(k)}[w] = 0$, for each $i \in \mathcal{I}$.

Suppose $\deg(c_i) \leq 2k$ for each i. Let $\mathscr{V}_{c_i}^{(2k)}[w]$ denote the vector such that

$$\mathscr{R}_w(c_i(\mathtt{x})u(\mathtt{x})) \;=\; \left(\mathscr{V}_{c_i}^{(2k)}[w]\right)^T \mathbf{u} \tag{5.4}$$

for all $u(\mathtt{x}) \in \mathbb{R}[\mathtt{x}]_{2k-\deg(c_i)}$. The $\mathscr{V}_{c_i}^{(2k)}[w]$ is called the *localizing vector* of the polynomial c_i, generated by the truncated multisequence w. It is important to observe that $\mathscr{V}_{c_i}^{(2k)}[w] = 0$ if w has a representing measure supported on $c_i(x) = 0$.

To get a finite dimensional optimization problem, we choose a finite value for l, e.g., $l = k$. Recall that

$$\mathscr{R}_w(f) = \sum_\alpha f_\alpha w_\alpha \quad \text{for} \quad f(\mathtt{x}) = \sum_\alpha f_\alpha \mathtt{x}^\alpha.$$

In particular, $w_0 = \mathscr{R}_w(1) = 1$. Therefore, the kth order truncation of (5.3) is

$$\left\{\begin{array}{rl} \gamma_k := \min & \sum\limits_\alpha f_\alpha w_\alpha \\ \text{s.t.} & \mathscr{V}_{c_i}^{(2k)}[w] = 0\,(i \in \mathcal{I}), \\ & \mathrm{L}_{c_j}^{(k)}[w] \succeq 0\,(j \in \mathcal{J}), \\ & \mathrm{L}_{\nu,g_j}^{(k,k)}[w] \succeq 0\,(j = 0,1,\ldots,s), \\ & w_0 = 1, \quad w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{array}\right. \tag{5.5}$$

Note $g_0 = 1$ in the above. For each given k, (5.5) is a semidefinite program. The length of the moment vector w is $\binom{n+2k}{2k}$. The vector $\mathscr{V}_{c_i}^{(2k)}[w]$ has length $\binom{n+2k-\deg(c_i)}{2k-\deg(c_i)}$. The matrix $\mathrm{L}_{c_j}^{(k)}[w]$ has length $\binom{n+k-\lceil\deg(c_i)\rceil}{k-\lceil\deg(c_i)\rceil}$. The length of $\mathrm{L}_{\nu,g_j}^{(k,k)}[w]$ is $\binom{m+k}{k}\cdot\binom{n+k}{k}$. The following is the convergence property of the moment relaxations (5.5).

Theorem 5.2. *Let K be as in* (1.1) *and suppose the measure ν satisfies Assumption 3.1. Assume* $\mathrm{QM}[g,\nu] + \mathrm{Ideal}[c_{eq}] + \mathrm{QM}[c_{in}]$ *is archimedean. Then the sequence $(\gamma_k)_k$ of* (5.5) *is monotonically increasing and*

$$\gamma_k \;\to\; f_{\min} \qquad \textit{as} \qquad k \to \infty.$$

Proof. Clearly, the sequence γ_k is monotonically increasing and $\gamma_k \leq f_{\min}$ for all k. For all $\varepsilon > 0$, the polynomial $f(\mathtt{x}) - f_{\min} + \varepsilon > 0$ on $X \cap K$, so

$$f(\mathtt{x}) - f_{\min} + \varepsilon \in \mathrm{QM}[g,\nu]_{2k} + \mathrm{Ideal}[c_{eq}]_{2k} + \mathrm{QM}[c_{in}]_{2k},$$

for k large enough, by Theorem 3.4. For each truncated multisequence w that is feasible in (5.5), we have

$$\mathscr{R}_w(f(\mathtt{x}) - (f_{\min} - \varepsilon)1) \geq 0.$$

This implies that

$$\mathscr{R}_w(f) \geq (f_{\min} - \varepsilon)\mathscr{R}_w(1) = f_{\min} - \varepsilon.$$

So the optimal value $\gamma_k \geq f_{\min} - \varepsilon$. Since $\varepsilon > 0$ can be arbitrarily small, the limit of γ_k must be $f_{\min}$. □

5.1. **Sampling.** In the expression for the localizing matrix $\mathrm{L}_{\nu,g_j}^{(k,l)}[w]$ in (5.5), we need the matrix $Y_{ji}^{(l')}$, which then requires the moments $\int y^\alpha \,\mathrm{d}\nu(y)$, for the chosen measure ν with $\mathrm{supp}(\nu) = Q$. If Q is a well-known and understood set (e.g., a box $[-1,1]^n$, a simplex, a unit ball or a sphere), the moments can be given by explicit formulas, such as for the uniformly distributed probability measure. If Q is not such a convenient set, the moments $\int_Q y^\alpha \,\mathrm{d}\nu(y)$ may not be readily available. However, this issue can be fixed by sampling.

For a given degree l, the moment vector $\int [y]_{2l} \,\mathrm{d}\nu(y)$ can always be written as a sample average, i.e., there exist points $u_1, \ldots, u_N \in Q$ such that

$$\int [y]_{2l} \,\mathrm{d}\nu(y) \;=\; \frac{1}{N}\sum_{i=1}^{N} [u_i]_{2l}.$$

This is guaranteed by Caratheodory's theorem [Bar02, Theorem I.2.3]. To get such sample points u_i can be tricky for some Q. In our computation, we assume they are available from the description of Q. Properties for them to satisfy are discussed in Theorem 5.3. Interestingly, the above sample average is actually the moment sequence of a certain measure whose support equals $\mathrm{supp}(\nu) = Q$ if the sample size N is large enough. We refer to Remark 5.4 for how large the sample size N should be. For a given degree d, consider the cone of all possible moment sequences

$$\mathcal{P}_d := \left\{ \int [y]_d \,\mathrm{d}\mu(y) \,\middle|\, \mu \text{ is a Borel measure on } \mathbb{R}^m,\ \mathrm{supp}(\mu) \subseteq Q \right\}. \tag{5.6}$$

Denote the relative interior of $\mathcal{P}_d$ by $\mathrm{relint}(\mathcal{P}_d)$.

Theorem 5.3. *Let Q be a closed set and let $\mathcal{P}_d$ be as above. For every $\xi \in \mathrm{relint}(\mathcal{P}_d)$, there exists a measure ν on $\mathbb{R}^m$ such that*

$$\xi \;=\; \int [y]_d \,\mathrm{d}\nu(y), \quad \mathrm{supp}(\nu) = Q. \tag{5.7}$$

Moreover, for points $u_1, \ldots, u_D \in Q$, if $\dim \mathrm{Span}\{[u_1]_d, \ldots, [u_D]_d\} = \dim \mathcal{P}_d$, then the sample average

$$A(u_1, \ldots, u_D) := \frac{1}{D}([u_1]_d + \cdots + [u_D]_d)$$

belongs to the relative interior $\mathrm{relint}(\mathcal{P}_d)$.

Proof. Consider the subcone

$$\mathcal{P}_d' := \left\{ \int [y]_d \,\mathrm{d}\mu(y) \,\middle|\, \mu \text{ is a Borel measure on } \mathbb{R}^m,\ \mathrm{supp}(\mu) = Q \right\}.$$

We show that $\mathcal{P}_d'$ is contained in the relative interior of $\mathcal{P}_d$. Let T denote the embedding Euclidean space of $[y]_d$ over all possible $y \in \mathbb{R}^m$. Then $\mathcal{P}_d' \subseteq \mathcal{P}_d$ are both convex cones in T. Let ℓ be any linear functional such that

$$\ell \geq 0 \quad \text{on} \quad \mathcal{P}_d, \qquad \ell(\eta) = 0 \quad \text{for some} \quad \eta \in \mathcal{P}_d'.$$

Since $\ell \geq 0$ on $\mathcal{P}_d$ and $[y]_d \in \mathcal{P}_d$ for all $y \in Q$, it is evident that the polynomial $p(y) := \ell([y]_d)$ is nonnegative on Q. Let μ be the Borel measure such that $\eta = \int [y]_d \, \mathrm{d}\mu(y)$ and $\mathrm{supp}(\mu) = Q$. Then

$$0 = \ell(\eta) = \int \ell([y]_d) \, \mathrm{d}\mu(y) = \int p(y) \, \mathrm{d}\mu(y).$$

Since $p(y) \geq 0$ on Q and $\mathrm{supp}(\mu) = Q$, the above implies that $p(y) \equiv 0$ on Q, i.e., $\ell \equiv 0$ on $\mathcal{P}_d$. This shows that every supporting hyperplane of $\mathcal{P}_d$ passing through any point of $\mathcal{P}'_d$ must also contain $\mathcal{P}_d$ entirely. So $\mathcal{P}'_d$ lies in the relative interior of $\mathcal{P}_d$. We remark that $\mathcal{P}'_d$ is dense in $\mathcal{P}_d$. To see this, fix a measure ν such that $\mathrm{supp}(\nu) = Q$. Then for every $\xi \in \mathcal{P}_d$ and each integer $k > 0$, we have $\xi + \frac{1}{k} \int [y]_d \, \mathrm{d}\nu(y) \in \mathcal{P}'_d$ and it converges to ξ as k goes to infinity. Since $\mathcal{P}'_d$ is dense in $\mathcal{P}_d$, they have the same relative interior.

So, every $\xi \in \mathrm{relint}(\mathcal{P}_d)$ is the expectation of $[y]_d$ for a certain measure ν whose support equals Q. Let ℓ be a linear functional such that $\ell \geq 0$ on $\mathcal{P}_d$. If $\ell(A(u_1, \ldots, u_D)) = 0$, then

$$\mathrm{Span}\{[u_1]_d, \ldots, [u_D]_d\} \subseteq \ker \ell.$$

If $\dim \mathrm{Span}\{[u_1]_d, \ldots, [u_D]_d\} = \dim \mathcal{P}_d$, then

$$\mathcal{P}_d \subseteq \ker \ell.$$

This implies that $\ell \equiv 0$ on $\mathcal{P}_d$. The above is true for every linear functional $\ell \geq 0$ on $\mathcal{P}_d$. Therefore, $A(u_1, \ldots, u_D)$ lies in the relative interior of $\mathcal{P}_d$. □

Remark 5.4. Note that the measure ν in (5.7) automatically satisfies the Carleman condition (3.1) if Q is bounded. However, for unbounded Q, we are not sure if (3.1) still holds. For the case of unbounded Q, we may apply the homogenization trick to transform to bounded sets. We refer to the work [HNY23a, HNY23b] for how to do this. To summarize, to formulate the localizing matrix $\mathrm{L}_{\nu,g_j}^{(k,l)}[w]$ in (4.6), we can select sample points $u_1, \ldots, u_N \in Q$ such that $\mathrm{Span}\{[u_1]_{2l}, \ldots, [u_N]_{2l}\}$ has maximum dimension, and then let

$$Y_{h_{ji}}^{(l')} = \frac{1}{N} \sum_{t=1}^{N} h_{ji}(u_t)[u_t]_{l'}[u_t]_{l'}^T. \tag{5.8}$$

6. Numerical Experiments

This section reports numerical examples to show the hierarchy of moment relaxations (5.5) for solving the SIP (5.1). The computations are implemented in MATLAB R2023b on a laptop equipped with a 10th Generation Intel® Core™ i7-10510U processor and 16GB memory. The moment relaxations are implemented by the software `Gloptipoly` [HLL09], which calls the software `SeDuMi` [Str01] to solve the corresponding semidefinite programs. For the SIP (5.1), we use x^* and f^* to denote the global minimizer and the global minimum value respectively. The relaxation order is labelled by k. For each k, we use $w^{(k)}$ to denote the minimizer of (5.5). The minimum value of (5.5) is denoted as γ_k, which is a lower bound for the SIP (5.1).

The flat extension condition (4.12) can also be used to get minimizers. However, this works only if the moment relaxation (5.5) is tight for solving the SIP. When (4.12) fails, a practical way to get an approximate minimizer is to let

$$\hat{x}_k := (w_{e_1}^{(k)}, \ldots, w_{e_n}^{(k)}).$$

The feasibility of the computed point $\hat{x}_k$ is measured as the function value

$$\delta_{k,j} \ := \ \min_{y \in Q} g_j(\hat{x}_k, y), \quad j = 1, \ldots, s.$$

Then $\hat{x}_k$ satisfies the inequality constraint in (5.1) if and only if

$$\delta_k := \min_{j \in [s]} \delta_{k,j} \ \geq \ 0.$$

For each example, if the measure ν is not specified, we set it to be the normalized Lebesgue measure so that $\nu(Q) = 1$. The consumed computational time is denoted as `time`. For neatness of the presentation, all computational results are displayed with four decimal digits.

Example 6.1. (i) Consider the following SIP from [CG85, WY15]:

$$(6.1) \qquad \begin{cases} \min\limits_{x \in \mathbb{R}^2} & \frac{1}{3}x_1^2 + x_2^2 + \frac{1}{2}x_1 \\ \text{s.t.} & -\left(1 - x_1^2 y^2\right)^2 + x_1 y^2 + x_2^2 - x_2 \geq 0 \quad \forall y \in Q, \end{cases}$$

where $Q = [0, 1]$. Computational results for Problem (6.1) are shown in Table 1. The true

k	$\hat{x}_k$	γ_k	δ_k	time(s)
3	$(-0.8433, -0.6041)$	0.1803	-0.0310	0.4503
4	$(-0.7847, -0.6140)$	0.1899	-0.0090	0.4991
5	$(-0.7650, -0.6164)$	0.1926	-0.0036	0.5749
6	$(-0.7574, -0.6173)$	0.1935	-0.0017	0.9063
7	$(-0.7541, -0.6176)$	0.1940	$-9.46 \cdot 10^{-4}$	1.8849

Table 1. Computational results for SIP (6.1).

minimizer is $x^* \approx (-0.7500, -0.6180)$ with the minimum value $f^* \approx 0.1945$.

(ii) Consider the following SIP:

$$(6.2) \qquad \begin{cases} \min\limits_{x \in X} & (x_1 - x_2)(x_1 - 1) + (x_2 - x_1)(x_2 - 1) + (x_1 - 1)(x_2 - 1) + x_1^3 + x_2^3 \\ \text{s.t.} & x_1 x_2 y_1 y_2 - (x_1 x_2 + x_2^2 + 0.01)(y_1 y_3 + y_2 + 1) - x_2^2 y_2 y_3 \geq 0 \quad \forall y \in Q, \end{cases}$$

where the sets are

$$X = [-10, 10]^2 \cap \left\{(x_1, x_2) : x_1 x_2 + x_1 + 1 \geq 0\right\},$$
$$Q = \left\{y \in \mathbb{R}^3 : y_1 \geq 0, \ y_2 \geq 0, \ y_3 \geq 0, 1 - y_1 - y_2 - y_3 \geq 0\right\}.$$

As in [Las21], we have the moment formula:

$$\int_Q y_1^{\alpha_1} y_2^{\alpha_2} y_3^{\alpha_3} \, \mathrm{d}y = \frac{6\alpha_1!\alpha_2!\alpha_3!}{(|\alpha| + 3)!}.$$

Computational results for Problem (6.2) are shown in Table 2. The true minimizer is $x^* \approx (0.3705, -0.0371)$ with the minimum value $f^* \approx 0.8697$.

k	$\hat{x}_k$	γ_k	δ_k	time(s)
2	$(0.3643, -0.0327)$	0.8624	-0.0017	0.4783
3	$(0.3661, -0.0340)$	0.8645	-0.0012	0.6968
4	$(0.3671, -0.0347)$	0.8658	$-9.28 \cdot 10^{-4}$	3.1540
5	$(0.3677, -0.0351)$	0.8665	$-7.71 \cdot 10^{-4}$	66.2828

Table 2. Computational results for SIP (6.2).

Example 6.2. (i) Consider the following SIP:

$$\left\{ \begin{array}{ll} \min\limits_{x \in \mathbb{R}^3} & -x_1^2(100 - x_1 - x_2) + x_2^2 + 2x_3^2 \\ \text{s.t.} & \begin{pmatrix} x_1 y_1^2 - x_1 x_2 y_1 y_2 - x_2 x_3 y_2^3 + 0.1 \\ x_3^2(y_1^2 - y_2^2) + x_2^2 y_1 y_2 + x_1 y_2 + 0.1 \end{pmatrix} \geq 0 \quad \forall y \in Q, \end{array} \right. \tag{6.3}$$

where $Q = \{(y_1, y_2) : y_1^4 + y_2^4 = 1\}$. We use the sampling as in (5.8) to get moments for ν. Computational results for for Problem (6.3) are shown in Table 3.

k	$\hat{x}_k$	γ_k	$(\delta_{k,1}, \delta_{k,2})$	time(s)
2	$(-0.1253, -0.0078, -0.0000)$	-1.5728	$(-0.0254, -0.0253)$	0.4407
3	$(-0.1128, -0.0062, -0.0000)$	-1.2747	$(-0.0129, -0.0128)$	0.7351
4	$(-0.1016, -0.0051, -0.0000)$	-1.0340	$(-0.0016, -0.0016)$	4.3025
5	$(-0.1009, -0.0049, -0.0000)$	-1.0247	$(-9.52, -9.03) \cdot 10^{-4}$	104.4387

Table 3. Computational results for SIP (6.3).

The true minimizer is $x^* \approx (-0.1000, -0.0018, 0.0000)$, with minimum value $f^* \approx -1.0010$.

(ii) Consider the following SIP:

$$\left\{ \begin{array}{ll} \min\limits_{x \in X} & \big(-\sum\limits_{i=1}^{4} x_i^4 \big) + x_1^3 x_2^2 + x_2^2 x_3^3 + x_3 x_4^4 - x_1^2 x_2^2 + x_1 x_2 x_3 x_4 + x_1 x_3 \\ \text{s.t.} & y^T \begin{bmatrix} x_1^2 - x_2 x_3 & x_1 + x_2 x_4 & x_3^2 - x_1 x_2 \\ x_1 + x_2 x_4 & x_1 - x_4^2 & 1 - e^T x \\ x_3^2 - x_1 x_2 & 1 - e^T x & x_1 x_2 + x_3 x_4 \end{bmatrix} y \geq 0 \quad \forall y \in Q, \end{array} \right. \tag{6.4}$$

where $X = \{x \in \mathbb{R}^4 : 4 - x^T x \geq 0\}$ and

$$Q = \left\{ y \in \mathbb{R}^3 : \, 1 - y^T y = 0, \, y \geq 0 \right\}.$$

We apply the sampling as in (5.8) to get moments of ν. Computational results for Problem (6.4) are shown in Table 4.

k	$\hat{x}_k$	γ_k	δ_k	time(s)
3	$(1.3903, -0.0000, -0.7310, 0.0000)$	-16.1250	$-3.12 \cdot 10^{-6}$	2.0730
4	$(0.1252, 0.0000, -1.9961, 0.0000)$	-16.1250	$2.62 \cdot 10^{-6}$	192.2331

TABLE 4. Computational results for SIP (6.4).

The true minimizer is $x^* \approx (0.1252, 0.0000, -1.9961, 0.0000)$, and the minimum value is $f^* \approx -16.1250$.

The following are examples where the quantifier set Q is not semialgebraic. For such Q, we typically need sampling to get moments of ν. We refer to Remark 5.4 for this issue. Generally, we pick sample points $u_1, \ldots, u_N \in Q$ such that $\mathrm{Span}\{[u_1]_{2l}, \ldots, [u_N]_{2l}\}$ has maximum dimension.

Example 6.3. (i) Consider the following SIP

$$\begin{cases} \min\limits_{x \in X} & -x_1x_2x_3 + x_1^3 + x_2^2 + x_3 \\ \text{s.t.} & (x_1x_2+1)y_2^4 + (e^Tx)y_1^2y_2 + (x_1 + x_2x_3)y_1^3 - 0.1 \geq 0 \quad \forall y \in Q, \end{cases} \tag{6.5}$$

where the sets

$$X = \left\{ x \in \mathbb{R}^3 \,\middle|\, \begin{array}{l} 5 - x^Tx \geq 0, \\ x_1x_2 - x_3 \geq 0 \end{array} \right\}, \quad Q = \left\{ y \in \mathbb{R}^2 \,\middle|\, \begin{array}{l} 4 - 3^{y_1^2} - 3^{y_2^2} \geq 0, \\ 3^{y_1} - 3^{y_2} - 1 \geq 0 \end{array} \right\}.$$

We apply the sampling as in (5.8) to get moments of ν. Computational results for Problem (6.5) are shown in Table 5.

k	$\hat{x}_k$	γ_k	δ_k	time(s)
2	$(-2.0115, -0.6861, -0.6951)$	-7.4037	-2.7956	0.4233
3	$(0.4350, -0.3706, -2.1618)$	-2.2907	$-7.89 \cdot 10^{-4}$	0.5196
4	$(0.4350, -0.3706, -2.1618)$	-2.2907	$-7.71 \cdot 10^{-4}$	1.4434
5	$(0.4350, -0.3707, -2.1618)$	-2.2907	$-6.25 \cdot 10^{-4}$	19.9535

TABLE 5. Computational results for SIP (6.5).

The true minimizer is $x^* \approx (0.4353, -0.3710, -2.1617)$, and the minimum value is $f^* \approx -2.2907$. They are estimated by applying the 6th order Taylor expansion of the exponential function.

(ii) Consider the following SIP:

$$\begin{cases} \min\limits_{x \in X} & x_1^3 - x_3^3 + x_1x_2^2 + (x_2 + x_3^2)^2 \\ \text{s.t.} & x_1x_3y_2y_3 + x_1x_2y_3 + x_2x_3y_1 + (x_1 + 2x_2 + x_3)(y_1y_2 + 2y_3) \geq 0 \quad \forall y \in Q, \end{cases} \tag{6.6}$$

where $X = [-1.5, 1.5]^3$ and

$$Q = \left\{ y \in \mathbb{R}^3 \,\middle|\, \begin{array}{l} 2 - y^T y \geq 0, \\ 2^{y_3} - 2^{y_1} - 2^{y_2} \geq 0 \end{array} \right\}.$$

We apply the sampling as in (5.8) to get moments of ν. Computational results are shown in Table 6.

k	$\hat{x}_k$	γ_k	δ_k	time(s)
3	$(-1.5000, 1.5000, -0.0000)$	-4.5000	$-6.76 \cdot 10^{-6}$	0.7619
4	$(-1.5000, 1.5000, -0.0000)$	-4.5000	$-7.99 \cdot 10^{-6}$	14.9845

Table 6. Computational results for SIP (6.6).

The true minimizer is $x^* \approx (-1.5000, 1.5000, 0.0000)$ and $f^* \approx -4.5000$. They are estimated by applying the 6th order Taylor expansion of exponential functions.

(iii) Consider the following SIP:

$$(6.7) \qquad \begin{cases} \min & x_1^3 + x_2^3 \\ \text{s.t.} & 4y^4 - 1 - y^2(2y^2 - 1)x_1^2 - y^2(2y^2 + 1)x_2^2 \geq 0 \qquad \forall y \in Q, \end{cases}$$

where $Q = \mathbb{Z}_+$. We select the same measure ν as in Example 3.10. The computational results are shown in Table 7. For this SIP, the true minimum value $f^* \approx -2.8284$ and the true minimizer $x^* \approx (-1.4142, 0.0000)$. We remark that there are numerical issues for solving the moment relaxation (5.5) when the relaxation order $k \geq 7$.

k	$\hat{x}_k$	γ_k	time(s)
2	$(-1.4561, -0.0000)$	-3.0874	0.7859
3	$(-1.4322, -0.0000)$	-2.9375	0.9379
4	$(-1.4246, -0.0000)$	-2.8915	0.8905
5	$(-1.4211, -0.0000)$	-2.8701	1.0247
6	$(-1.4191, -0.0000)$	-2.8581	1.4835
7	$(-1.4179, -0.0002)$	-2.8506	2.4642
8	$(-1.4156, -0.0002)$	-2.8424	5.8728

Table 7. Computational results for the SIP (6.7).

The following are examples where the quantifier set Q is a union of several closed sets, say,

$$Q \;=\; Q_1 \cup \cdots \cup Q_l, \quad \text{for each} \quad Q_i \subseteq \mathbb{R}^m.$$

For each i, let ν be a measure on $\mathbb{R}^m$ such that $\text{supp}(\nu_i) = Q_i$. Then $\nu \coloneqq \nu_1 + \cdots + \nu_l$ is a measure such that $\text{supp}(\nu) = Q$. By the definition, one can see that

$$\text{QM}[g, \nu] \;=\; \text{QM}[g, \nu_1] + \cdots + \text{QM}[g, \nu_l].$$

Example 6.4. (i) Consider the following SIP:

$$
(6.8) \qquad \begin{cases} \min\limits_{x\in\mathbb{R}^3} & (x_1^2+1.8x_3^2)^2 + x_1x_2x_3 + x_1^3 - 2x_2^3 - 4x_3 \\ \text{s.t.} & \begin{pmatrix} x_1x_2y_1y_2 - x_2x_3(y_1+y_3^2) - 0.01 \\ x_3^2y_1^2 - x_2^2y_2y_3 + x_1^2(y_1+y_3-0.1) \end{pmatrix} \geq 0 \quad \forall y \in Q, \end{cases}
$$

where $Q = Q_1 \cup Q_2$ is the union of the following two sets:

$$
\begin{array}{rcl} Q_1 & = & \left\{ y \in \mathbb{R}^3 : (y_1-1)^2 + (y_2-1)^2 + (y_3-1)^2 \leq 1 \right\}, \\ Q_2 & = & \{ y \in \mathbb{R}^3 : (y_1-1)^2 + y_2^2 + (y_3-1)^2 \leq 1 \}. \end{array}
$$

We apply the sampling as in (5.8) to get moments of ν. Computational results are shown in Table 8.

k	$\hat{x}_k$	γ_k	$(\delta_{k,1}, \delta_{k,2})$	time(s)
2	$(1.8793, 2.2691, -0.2007)$	-3.7910	$(-4.4703, -6.6000)$	0.5968
3	$(0.0047, -0.0231, 0.6758)$	-2.0274	$(-4.44, -0.58) \cdot 10^{-3}$	1.1363
4	$(0.0047, -0.0230, 0.6758)$	-2.0274	$(-4.42, -0.58) \cdot 10^{-3}$	48.7642

TABLE 8. Computational results for SIP (6.8).

(ii) Consider the following SIP:

$$
(6.9) \qquad \begin{cases} \min\limits_{x\in X} & (x_1^2 - x_2)^2 - 3x_1x_2^2 + 3x_1^3 \\ \text{s.t.} & -x_1x_2(y_1^2+2y_3^2) + x_2^2(y_1 - y_2y_3) + 2y_1y_3 - e^Tx - 1.4 \geq 0 \quad \forall y \in Q, \end{cases}
$$

where $X = \{x \in \mathbb{R}^2 : 8 - x^Tx \geq 0\}$ and

$$
Q = \left\{ y \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} 10 - y^Ty \geq 0, \\ |y_1| + |y_2| + |y_3| - 1 \geq 0 \end{array} \right\}.
$$

Note that Q is a union of 8 basic closed semialgebraic sets, that is,

$$
Q \quad = \bigcup_{s_1,s_2,s_3\in\{-1,1\}} Q_{s_1,s_2,s_3} := \left\{ y \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} 10 - y^Ty \geq 0, \\ s_1y_1 \geq 0, s_2y_2 \geq 0, s_3y_3 \geq 0, \\ s_1y_1 + s_2y_2 + s_3y_3 - 1 \geq 0 \end{array} \right\}.
$$

We apply the sampling as in (5.8) to get moments of ν. Computational results are shown in Table 9.

We remark that the flat extension condition (4.12) can also be used to get minimizers when it holds. This happens only if the moment relaxation is tight for solving the SIP. See the following example.

k	$\hat{x}_k$	γ_k	δ_k	time(s)
2	$(-2.4791, 0.7284)$	-12.4135	$-6.46 \cdot 10^{-4}$	0.4755
3	$(-2.4791, 0.7284)$	-12.4135	$-6.33 \cdot 10^{-4}$	0.7434
4	$(-2.4791, 0.7284)$	-12.4135	$-6.36 \cdot 10^{-4}$	2.8389
5	$(-2.4779, 0.7272)$	-12.4099	$4.01 \cdot 10^{-4}$	53.6131

TABLE 9. Computational results for SIP (6.9).

Example 6.5. Consider the following SIP:

$$\left\{ \begin{array}{ll} \min\limits_{x \in X} & -x_1^2 x_2^2 \\ \text{s.t.} & (x_1 + x_2) y_2^2 - x_1 x_2 (y_1 y_2 + 1) \geq 0 \quad \forall y \in Q, \end{array} \right. \tag{6.10}$$

where $X = \{x \in \mathbb{R}^2 : 1 - x^T x \geq 0\}$ and

$$Q = \{y \in \mathbb{R}^2 : |y_1| + |y_2| \leq 1\}.$$

For the relaxation order $k = 2$, we get the optimal w^* such that

$$\mathrm{H}^{(2)}[w^*] = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \\ -\frac{1}{2} & 0 & 0 & -\frac{1}{4} & \frac{1}{4} & -\frac{1}{4} \\ \frac{1}{2} & 0 & 0 & \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

The flat extension (4.12) holds. Indeed, we can get $w = \frac{1}{2}([u_1^*]_4 + [u_2^*]_4)$ for points

$$u_1^* = \Big(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\Big), \quad u_2^* = \Big(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\Big).$$

They are both minimizers for this SIP.

7. CONCLUSIONS AND DISCUSSIONS

We study Positivstellensätze and moment problems for sets that are given by universal quantifiers. For the set K as in (1.1) given by a universal quantifier $y \in Q$, we discuss representation of polynomials that are positive on K. Let ν be a measure satisfying the Carleman condition (3.1). When the quadratic module $\mathrm{QM}[g, \nu]$ is archimedean, we show in Theorem 3.4 that a polynomial $f(\mathtt{x})$ positive on K as in (1.1) must be in $\mathrm{QM}[g, \nu]$. For the non-archimedean case, we give a similar result in Corollary 3.7. We also study K-moment problems for the set K. Necessary and sufficient conditions for a full (or truncated) multisequence to admit a representing measure supported in K are given. In particular, the classical flat extension theorem is generalized for truncated moment problems with such a set K. These results are presented in Theorems 4.1, 4.4 and 4.6, respectively. These new

Positivstellensätze and moment problems can be applied to solve semi-infinite optimization (SIP). For the SIP (5.1), a hierarchy of moment relaxations (5.5) is proposed to solve it. Its convergence is shown in Theorem 5.2. Various examples for semi-infinite optimization are demonstrated in Section 6.

Our work leads to many intriguing questions to explore in the future. For instance, without assuming archimedeanness, is there a preordering version of Theorem 3.4? Equivalently, does there exist a clean algebraic reformulation of the compactness (or emptiness) of the set K given with a universal quantifier as in (1.1)? Is there an analog of the Krivine-Stengle Positivstellensatz for such sets K? It would also be interesting to establish the universal Positivstellensätze for matrix-valued polynomials and matrix-valued constraints. In Theorem 4.6, the condition (4.11) is assumed to hold for all $l = 1, 2, \ldots$. If it holds for only finitely many l, the conclusion of Theorem 4.6 may not hold. It would be interesting to find a *finite* set of conditions for a truncated multisequence to admit a representing measure supported in K. Finally, in their previous joint work, the second and third author [KN20] gave Positivstellensätze and solvability criteria for moment problems for sets given with *existential* quantifiers. A major future task will be to give a common extension of the results from [KN20] and the present paper, that is, Positivstellensätze and moment problems for sets given with a combination of universal and existential quantifiers.

Acknowledgment. We thank the anonymous referees for their valuable comments and suggestions. Xiaomeng Hu and Jiawang Nie are partially supported by the NSF grant DMS-2110780. Igor Klep is supported by the Slovenian Research Agency program P1-0222 and grants J1-50002, J1-2453, N1-0217, J1-3004, and was partially supported by the Marsden Fund Council of the Royal Society of New Zealand. Igor's work was partly performed within the project COMPUTE, funded within the QuantERA II Programme that has received funding from the EU's H2020 research and innovation programme under the GA No 101017733 .

References

[Akh65] N.I. Akhiezer, *The classical moment problem and some related questions in analysis* (Translated by N. Kemmer), Hafner Publishing Co., New York, 1965

[Bar02] A. Barvinok, *A course in convexity*, Graduate studies in Mathematics 54, AMS, 2002.

[BPR] S. Basu, R. Pollack and M.-F. Roy, *Algorithms in Real Algebraic Geometry*, Springer-Verlag, Berlin, 2006.

[Ber87] C. Berg, *The multidimensional moment problem and semigroups*, Moments in mathematics (San Antonio, Tex., 1987), 110–124, *Proc. Sympos. Appl. Math.* 37, AMS Short Course Lecture Notes, AMS, Providence, RI, 1987.

[BS16] C. Berg, R. Szwarc, *On the order of indeterminate moment problems*, Adv. Math. 250 (2014) 105–143.

[BL20] G. Blekherman, L. Fialkow, *The core variety and representing measures in the truncated moment problem*, J. Operator Theory 84 (2020) 185–209.

[BCR98] J. Bochnack, M. Coste and M.-F. Roy, *Real algebraic geometry*, Ergebnisse der Mathematik und ihrer Grenzgebiete 3, Springer, 1998.

[CKS09] J. Cimprič, S. Kuhlmann, and C. Scheiderer, *Sums of squares and moment problems in equivariant situations*, Trans. Amer. Math. Soc. 361 (2009) 735–765.

[CMN11] J. Cimprič, M. Marshall, and T. Netzer, *On the real multidimensional rational K-moment problem*, Trans. Amer. Math. Soc. 363 (2011) 5773–5788.

[CG85] I. Coope and G. Watson, A projected Lagrangian algorithm for semi-infinite programming, *Math. Program.* 32 (1985) 337–356.

[CF96] R. Curto and L. Fialkow, *Solution of the truncated complex moment problem for flat data*, Mem. Amer. Math. Soc. 119 (1996), no. 568.

[CF05] R. Curto and L. Fialkow, *Truncated K-moment problems in several variables*, J. Oper. Theory 54 (2005), 189–226.

[CGIK23] R.E. Curto, M. Ghasemi, M. Infusino, and S. Kuhlmann, *The truncated moment problem for unital commutative $\mathbb{R}$-algebras*, J. Operator Theory 90 (2023) 223–261.

[DM17] H. Djelassi and A. Mitsos, *A hybrid discretization algorithm with guaranteed feasibility for the global solution of semi-infinite programs*, J. Global Optim. 68 (2017) 227–253.

[EP20] P. Escorcielo and D. Perrucci, *A version of Putinar's Positivstellensatz for cylinders*, J. Pure Appl. Algebra 224 (2020), 106448, 16 pp.

[Fia16] L. Fialkow, *The truncated K-moment problem: a survey*, Operator theory: the state of the art, Theta Ser. Adv. Math. 18, 25–51, Theta, Bucharest, 2016.

[FS09] P. Flajolet, R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, 2009.

[Fri21] T. Fritz, *A generalization of Strassen's Positivstellensatz*, Comm. Algebra 49 (2021) 482–499.

[GKKS15] A. Gala-Jaskórzyńska, K. Kurdyka, K. Kuta, and S. Spodzieja, *Positivstellensatz for homogeneous semialgebraic sets*, Arch. Math. (Basel) 105 (2015) 405–412.

[Havi36] E.K. Haviland, *On the momentum problem for distribution functions in more than one dimension II*, Amer. J. Math., 58:164–168, 1936.

[HKL20] D. Henrion, M. Korda, and J. Lasserre, *The Moment-SOS Hierarchy*, World Scientific, Singapore, 2020.

[HLL09] D. Henrion, J. Lasserre, J. Löfberg, *GloptiPoly 3: moments, optimization and semidefinite programming*, Optim. Methods Softw., **24**, pp. 761–779, 2009.

[HuN23] X. Hu and J. Nie, *Polynomial Optimization Relaxations for Generalized Semi-Infinite Programs*, Preprint, 2023. `arxiv.org/abs/2303.14308`

[HNY23a] L. Huang, J. Nie and Y.-X. Yuan, *Homogenization for polynomial optimization with unbounded sets*, Math. Program. 200(1) (2023) 105–145.

[HNY23b] L. Huang, J. Nie and Y.-X. Yuan, *Generalized truncated moment problems with unbounded sets*, *J. Sci. Comput.* 95(1) (2023), art. 15.

[IK17] M. Infusino, S. Kuhlmann, *Infinite dimensional moment problem: open questions and applications*, in: Ordered algebraic structures and related topics, 187–201, Contemp. Math. 697, AMS, 2017.

[IKKM22] M. Infusino, S. Kuhlmann, T. Kuna, P. Michalski, *Projective limit techniques for the infinite dimensional moment problem*, Integral Equations Operator Theory 94 (2022), no. 2, Paper No. 12, 44 pp.

[IKKM23] M. Infusino, S. Kuhlmann, T. Kuna, P. Michalski, *An Intrinsic Characterization of Moment Functionals in the Compact Case*, International Mathematics Research Notices 2023(3), pp. 2281–2303.

[Jacobi] T. Jacobi, *A representation theorem for certain partially ordered commutative rings*, Math. Z. 237 (2001) 259–273.

[KW13] D.P. Kimsey, H.J. Woerdeman, *The truncated matrix-valued K-moment problem on $\mathbb{R}^d$, $\mathbb{C}^d$, and $\mathbb{T}^d$*, Trans. Amer. Math. Soc. 365 (2013) 5393–5430.

[KMV+] I. Klep, V. Magron and J. Volčič, Sums Of Squares Certificates For Polynomial Moment Inequalities, *Preprint*, 2023. `arxiv.org/abs/2306.05761`

[KN20] I. Klep, J. Nie, *A Matrix Positivstellensatz with Lifting Polynomials*, SIAM J. Optim. 30 (2020) 240–261.

[Las15] J.B. Lasserre, *Tractable approximations of sets defined with quantifiers*, Math. Program. 151 (2015) 507–527.

[Las21] J.B. Lasserre, *Simple formula for integration of polynomials on a simplex*, BIT 61(2) (2021) 523–533.

[Las15] J.B. Lasserre, *Introduction to polynomial and semi-algebraic optimization*, Cambridge University Press, Cambridge, 2015.

[LN07] J.B. Lasserre and T. Netzer: *SOS approximations of nonnegative polynomials via simple high degree perturbations*, Math. Z. 256 (2007) 1432–1823.

[Lau05] M. Laurent, *Revisiting two theorems of Curto and Fialkow on moment matrices*, *Proc. Amer. Math. Soc.* 133(10) (2005) 2965–2976.

[Lau09] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, Emerging Applications of Algebraic Geometry of IMA Volumes in Mathematics and its Applications, Vol. 149, pp. 157–270, Springer, 2009.

[LPR20] H. Lombardi, D. Perrucci, M.-F. Roy, *An elementary recursive bound for effective Positivstellensatz and Hilbert's 17th problem*, Mem. Amer. Math. Soc. 263 (2020), no. 1277, v+125 pp.

[LS07] M. López and G. Still, *Semi-infinite programming*, European J. Oper. Res. 180 (2007) 491–518.

[Mar03] M. Marshall, *Approximating positive polynomials using sums of squares*, Canad. Math. Bull. 46, no. 3(2003): 400–418.

[Mar08] M. Marshall, *Positive polynomials and sums of squares*, AMS, 2008.

[MNR23] A. Müller-Hermes, I. Nechita, D. Reeb, *A refinement of Reznick's Positivstellensatz with applications to quantum information theory*, Quantum 7 (2023) 1001, 29pp.

[Net08] T. Netzer, *An elementary proof of Schmüdgen's theorem on the moment problem of closed semi-algebraic sets*, Proc. Amer. Math. Soc. 136 (2008) 529–537.

[Nie12] J. Nie, *Sum of squares methods for minimizing polynomial forms over spheres and hypersurfaces*, *Front. Math. China* 7 (2012) 321–346.

[Nie17] J. Nie, *Symmetric tensor nuclear norms*, SIAM J. Appl. Algebra Geometry 1(1) (2017) 599–625.

[Nie23] J. Nie, *Moment and Polynomial Optimization*, SIAM, Philadelphia, PA, 2023.

[NZ18] J. Nie and X. Zhang, *Real eigenvalues of nonsymmetric tensors*, Comp. Optim. Appl. 70(1) (2018) 1–32.

[Nuss] A.E. Nussbaum, *Quasi-analytic vectors*, Ark. Mat. 6 (1965) 179–191.

[PS01] V. Powers and C. Scheiderer, *The moment problem for non-compact semialgebraic sets*, Adv. Geom. 1 (2001) 71–88.

[Put93] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Ind. Univ. Math. J. 42 (1993) 969–984.

[PV99] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, Ann. of Math. (2) 149 (1999) 1087–1107.

[Rie16] C. Riener, *Symmetric semi-algebraic sets and non-negativity of symmetric polynomials*, J. Pure Appl. Algebra 220 (2016) 2809–2815.

[Riesz] M. Riesz, *Sur le problème de moments: Troisième note*, Ark. Mat. Astron. Fys. 17 (1923).

[Rob55] H. Robbins, *A remark on Stirling's formula*, Am. Math. Mon. 62 (1955) 26–29.

[Sce09] C. Scheiderer, *Positivity and sums of squares: A guide to recent results*, Emerging Applications of Algebraic Geometry (M. Putinar, S. Sullivant, eds.), IMA Volumes Math. Appl. 149, Springer, 2009, pp. 271–324.

[Smü91] K. Schmüdgen, *The K-moment problem for compact semialgebraic sets*, Math. Ann. 289 (1991), 203–206.

[Smü17] K. Schmüdgen, *The moment problem*, Graduate Texts in Mathematics, 277. Springer, Cham, 2017.

[SS24] K. Schmüdgen, M. Schötz, *Positivstellensätze for semirings*, Math. Ann. 389 (2024) 947–985.

[Scw03] M. Schweighofer, *Iterated rings of bounded elements and generalizations of Schmüdgen's Positivstellensatz*, J. Reine Angew. Math. 554 (2003) 19–45.

[Sta12] R.P. Stanley, *Enumerative combinatorics*, Vol. 1. 2nd ed., Cambridge Studies in Advanced Mathematics 49. Cambridge, 2012.

[SS12] O. Stein and P. Steuermann, *The adaptive convexification algorithm for semi-infinite programming with arbitrary index sets*, Math. Program. 136 (2012) 183–207.

[Str01] J. Sturm, *Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones*, Optim. Methods Softw. 11/12 (1999) 625–653.

[WG14] L. Wang and F. Guo, *Semidefinite relaxations for semi-infinite polynomial programming*, Comput. Optim. Appl.. 58 (2014) 133–159.

[WY15] S. Wang and Y. Yuan, *Feasible method for semi-infinite programs.* SIAM J. Optim. 25 (2015) 2537–2560.

[XCM09] H. Xu, C. Caramanis and S. Mannor, *Robustness and Regularization of Support Vector Machines*, *J. Mach. Learn. Res.* 10 (2009) 1485–1510.

Department of Mathematics, University of California San Diego
Email address: x8hu@ucsd.edu

Faculty of Mathematics and Physics, University of Ljubljana & Faculty of Mathematics, Science and Information Technology, University of Primorska, Koper & Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia
Email address: igor.klep@fmf.uni-lj.si

Department of Mathematics, University of California San Diego
Email address: njw@math.ucsd.edu