LINEAR OPTIMIZATION WITH CONES OF MOMENTS AND NONNEGATIVE POLYNOMIALS

JIAWANG NIE

Abstract. Let \( A \) be a finite subset of \( \mathbb{N}^n \) and \( \mathbb{R}[x]_A \) be the space of real polynomials whose monomial powers are from \( A \). Let \( K \) be a compact basic semialgebraic set of \( \mathbb{R}^n \) such that \( \mathbb{R}[x]_A \) contains a polynomial that is positive on \( K \). Denote by \( \mathcal{P}_A(K) \) the cone of polynomials in \( \mathbb{R}[x]_A \) that are nonnegative on \( K \). The dual cone of \( \mathcal{P}_A(K) \) is \( \mathcal{A}_A(K) \), the set of all \( A \)-truncated moment sequences in \( \mathbb{R}^A \) that admit representing measures supported in \( K \). Our main results are: i) We study the properties of \( \mathcal{P}_A(K) \) and \( \mathcal{A}_A(K) \) (like interiors, closeness, duality, memberships), and construct a convergent hierarchy of semidefinite relaxations for each of them. ii) We propose a semidefinite algorithm for solving linear optimization problems with the cones \( \mathcal{P}_A(K) \) and \( \mathcal{A}_A(K) \), and prove its asymptotic and finite convergence; a stopping criterion is also given. iii) We show how to check whether \( \mathcal{P}_A(K) \) and \( \mathcal{A}_A(K) \) intersect affine subspaces; if they do, we show to get a point in the intersections; if they do not, we prove certificates for the non-intersecting.

1. Introduction

Let \( \mathbb{N} \) (resp., \( \mathbb{R} \)) be the set of nonnegative integers (resp., real numbers), and let \( \mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n] \) be the ring of polynomials in \( x := (x_1, \ldots, x_n) \) and with real coefficients. For \( \alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \), denote \( x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \) and \( |\alpha| := \alpha_1 + \cdots + \alpha_n \). Denote by \( \mathbb{R}[x]_d \) the set of polynomials in \( \mathbb{R}[x] \) with degrees at most \( d \). Let \( K \) be a subset of \( \mathbb{R}^n \). Denote

\[
\mathcal{P}_d(K) = \{ p \in \mathbb{R}[x]_d : p(u) \geq 0 \forall u \in K \},
\]

which is the cone of polynomials in \( \mathbb{R}[x]_d \) that are nonnegative on \( K \). The dual space of \( \mathbb{R}[x]_d \) is the space of truncated moment sequences (tms'), which is denoted as \( \mathcal{M}_{n,d} \). A tms \( y = (y_\alpha) \) in \( \mathcal{M}_{n,d} \) is indexed by \( \alpha \in \mathbb{N}^n \) with \( |\alpha| \leq d \), and it acts on \( \mathbb{R}[x]_d \) as

\[
\langle p, y \rangle := \sum_{|\alpha| \leq d} p_\alpha y_\alpha \quad \text{for all} \quad p = \sum_{|\alpha| \leq d} p_\alpha x^\alpha.
\]

We say that a tms \( y \in \mathcal{M}_{n,d} \) admits a \( K \)-measure \( \mu \) (i.e., \( \mu \) is a nonnegative Borel measure supported in \( K \)) if

\[
y_\alpha = \int_K x^\alpha \text{d}\mu \quad \forall \, \alpha \in \mathbb{N}^n_d = \{ \alpha \in \mathbb{N}^n : |\alpha| \leq d \}.
\]

The measure \( \mu \) satisfying the above is called a \( K \)-representing measure for \( y \). In applications, we are often interested in finitely atomic measures, i.e., their supports

1991 Mathematics Subject Classification. 65K05, 90C22, 90C26.

Key words and phrases. Moment, nonnegative polynomial, representing measure, semidefinite program, sum of squares, truncated moment sequence.

The research was partially supported by the NSF grant DMS-0844775.
are finite sets. Denote by $\delta_u$ the Dirac measure supported at $u$ with unit mass. A measure $\mu$ is called $r$-atomic if $\mu = \lambda_1 \delta_{u_1} + \cdots + \lambda_r \delta_{u_r}$ with each $\lambda_i > 0$ and $u_i \in \mathbb{R}^n$. Let $\text{meas}(y, K)$ denote the set of all $K$-measures admitted by $y$. Denote

$$\mathcal{R}_d(K) = \{ y \in \mathcal{M}_{n,d} : \text{meas}(y, K) \neq \emptyset \}.$$

When $K$ is compact, $\mathcal{R}_d(K)$ is the dual cone of $\mathcal{P}_d(K)$ (cf. Tchakaloff [44] and Laurent [25, Section 5.2]).

Linear optimization problems with cones $\mathcal{P}_d(K)$ and $\mathcal{R}_d(K)$ have wide applications. For instance, the minimum value of a polynomial $f \in \mathbb{R}[x]_d$ on $K$ can be found by maximizing $\gamma$ subject to $f - \gamma \in \mathcal{P}_d(K)$; the corresponding dual problem is minimizing a linear function over the cone $\mathcal{R}_d(K)$ (cf. Lasserre [19]). Generalized problems of moments (GPMs), proposed by Lasserre [22], are optimizing linear moment functionals over the set of measures supported in a given set and satisfying some linear constraints. GPMs are equivalent to linear optimization problems with the cone $\mathcal{R}_d(K)$. Lasserre [22] proposed semidefinite relaxations for solving GPMs. We refer to [12, 25, 27, 28, 33, 37] for moment and polynomial optimization problems. Semidefinite programs are also very useful in representing convex sets and convex hulls, like in [11, 13, 14, 15, 23, 24, 30, 43], and in solving polynomial equations, like in [20, 21, 29].

In this paper, we study optimization problems with more general cones than $\mathcal{R}_d(K)$ and $\mathcal{P}_d(K)$. Let $\mathcal{A} \subseteq \mathbb{N}^n$ be a finite set, and $\mathbb{R}[x]_\mathcal{A} := \text{span}\{x^\alpha : \alpha \in \mathcal{A}\}$. The dual space of $\mathbb{R}[x]_\mathcal{A}$ is $\mathbb{R}\mathcal{A}$, the set of real vectors indexed by elements in $\mathcal{A}$. A vector in $\mathbb{R}\mathcal{A}$ is called an $\mathcal{A}$-truncated moment sequence ($\mathcal{A}$-tms). Define $\text{deg}(\mathcal{A}) := \max\{ |\alpha| : \alpha \in \mathcal{A}\}$. Like before, an $\mathcal{A}$-tms $y \in \mathbb{R}\mathcal{A}$ is said to admit a $K$-measure $\mu$ if $y_\alpha = \int_K x^\alpha \text{d}\mu$ for all $\alpha \in \mathcal{A}$. Such $\mu$ is called a $K$-representing measure for $y$. Denote

$$\mathcal{P}_\mathcal{A}(K) = \{ p \in \mathbb{R}[x]_\mathcal{A} : p(u) \geq 0 \forall u \in K \},$$

$$\mathcal{R}_\mathcal{A}(K) = \{ y \in \mathbb{R}\mathcal{A} : \text{meas}(y, K) \neq \emptyset \}.$$

Indeed, an $\mathcal{A}$-tms $y \in \mathbb{R}\mathcal{A}$ admits a $K$-measure if and only if it admits a $r$-atomic $K$-measure with $r \leq |\mathcal{A}|$ (cf. [33, Proposition 3.3]).

Clearly, if $\mathcal{A} = \mathbb{N}^n_0$, then $\mathcal{P}_\mathcal{A}(K)$ equals $\mathcal{P}_d(K)$, and $\mathcal{R}_\mathcal{A}(K)$ equals $\mathcal{R}_d(K)$. There are other interesting cases of $\mathcal{A}$ in applications.

- **Nonnegative forms and sums of even power of linear forms** When $K$ is the unit sphere $\mathbb{S}^{n-1} = \{ x \in \mathbb{R}^n : ||x||_2 = 1 \}$ and $\mathcal{A} = \{ \alpha \in \mathbb{N}^n : |\alpha| = d \}$ ($d$ is even), $\mathcal{P}_\mathcal{A}(K)$ is the cone of nonnegative forms in $n$ variables and of degree $d$ (we denote it by $P_{n,d}$), and $\mathcal{R}_\mathcal{A}(K)$ is the cone of homogeneous tms’s that admit representing measures supported in $\mathbb{S}^{n-1}$ (cf. [9]). Interestingly, $\mathcal{R}_d(K)$ is equivalent to the cone of sums of even powers (SOEP) of linear forms. A form $f \in \mathbb{R}[x]_d$ is SOEP if there exist $L_1, \ldots, L_r \in \mathbb{R}[x]_1$ such that $f = L_1^d + \cdots + L_r^d$ (cf. [11]). Let $Q_{n,d}$ be the cone of all SOEP forms of degree $d$. Every $f$ can be written as

$$f = \sum_{|\alpha| = d} (\alpha_1, \ldots, \alpha_n) f_\alpha x^\alpha.$$

So, each $f$ can be identified as a tms $\tilde{f} \in \mathbb{R}\mathcal{A}$. It can be shown that $f \in Q_{n,d}$ if and only if $\tilde{f} \in \mathcal{R}_\mathcal{A}(\mathbb{S}^{n-1})$ (cf. [11, 38]).

- **Copositive and completely positive cones** A $n \times n$ symmetric matrix $B$ is copositive if $x^T B x \geq 0$ for all $x \in \mathbb{R}^n_+$ (the nonnegative orthant), and a $n \times n$ symmetric matrix $C$ is completely positive if $C = \sum u_i u_i^T$.
for $u_1, \ldots, u_r \in \mathbb{R}^n$. We refer to [20] for copositive and completely positive matrices. When $K = \{x \in \mathbb{R}^n : |x| = 1\}$ and $A = \{\alpha \in \mathbb{N}^n : |\alpha| = 2\}$, $\mathcal{P}_A(K)$ reduces to the cone of $n \times n$ copositive matrices, and $\mathcal{R}_A(K)$ reduces to the cone of $n \times n$ completely positive matrices (cf. [36]).

**Contributions** Assume $A \subseteq \mathbb{N}^n$ is finite and $K$ is the set
\begin{equation}
K = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \geq 0\},
\end{equation}
defined by two polynomial tuples $h = (h_1, \ldots, h_m)$ and $g = (g_1, \ldots, g_m)$. Assume $K$ is compact and $\mathbb{R}[x]_A$ contains a polynomial that is positive on $K$.

First, we study properties of the cones $\mathcal{P}_A(K)$ and $\mathcal{R}_A(K)$. We characterize their interiors, prove their closeness and dual relationship, i.e., $\mathcal{R}_A(K)$ is the dual cone of $\mathcal{P}_A(K)$. We construct a convergent hierarchy of semidefinite relaxations for each of them. We also show how to check the memberships in $\mathcal{P}_A(K)$ and $\mathcal{R}_A(K)$. This will be in Section 3.

Second, we study how to solve linear optimization problems over the cones $\mathcal{P}_A(K)$ and $\mathcal{R}_A(K)$ with linear constraints. A semidefinite algorithm is proposed for solving these problems. Its asymptotic and finite convergence are proved. A stopping criterion is also given. This will be in Section 4.

Third, we study how to check whether an affine subspace intersects the cone $\mathcal{P}_A(K)$ or $\mathcal{R}_A(K)$. If they intersect, we show how to find a point in the intersection. If they do not, we prove certificates for the non-intersecting. This will be in Section 5.

Some basics in the field will be introduced in Section 2.

### 2. Preliminaries

**Notation** For $t \in \mathbb{R}$, $\lfloor t \rfloor$ (resp., $\lceil t \rceil$) denotes the smallest integer not smaller (resp., the largest integer not greater) than $t$. Denote $|k| := \{1, \ldots, k\}$. For $z \in \mathbb{R}^{n_2}$, denote by $z|_A$ the subvector of $z$ whose indices are in $A$. When $A = \mathbb{N}^n$, we simply denote $z|_A := z|_{\mathbb{N}^n}$. For a set $S \subseteq \mathbb{R}^n$, $|S|$ denotes its cardinality, and $int(S)$ denotes its interior. The superscript $^T$ denotes the transpose of a matrix or vector. For $u \in \mathbb{R}^n$ and $r \geq 0$, denote $\|u\|_2 := \sqrt{u^2}$ and $B(u, r) := \{x \in \mathbb{R}^n : \|x - u\|_2 \leq r\}$. For a polynomial $p \in \mathbb{R}[x]$, $\|p\|_2$ denotes the 2-norm of the coefficient vector of $p$. For a matrix $A$, $\|A\|_F$ denotes its Frobenius norm. If a symmetric matrix $X$ is positive semidefinite (resp., definite), we denote $X \geq 0$ (resp., $X > 0$).

#### 2.1. Riesz functionals, localizing matrices and flatness. Let $A \subseteq \mathbb{N}^n$ be a finite set. An $A$-tms $y \in \mathbb{R}^A$ defines a Riesz functional $\mathcal{L}_y$, acting on $\mathbb{R}[x]_A$ as
\[
\mathcal{L}_y \left( \sum_{\alpha \in A} p_\alpha x^\alpha \right) := \sum_{\alpha \in A} p_\alpha y_\alpha.
\]
We also denote $(p, y) := \mathcal{L}_y(p)$ for convenience. We say that $\mathcal{L}_y$ is $K$-positive if
\[
\mathcal{L}_y(p) \geq 0 \quad \forall p \in \mathcal{P}_A(K),
\]
and $\mathcal{L}_y$ is strictly $K$-positive if
\[
\mathcal{L}_y(p) > 0 \quad \forall p \in \mathcal{P}_A(K) : p|_K \neq 0.
\]
As is well known, $\mathcal{L}_y$ being $K$-positivity is a necessary condition for $y$ to admit a $K$-measure. The reverse is also true if $K$ is compact and $\mathbb{R}[x]_A$ is $K$-full (i.e., there exists $p \in \mathbb{R}[x]_A$ such that $p > 0$ on $K$) (cf. [9] Theorem 2.2).
For $q \in \mathbb{R}[x]_{2k}$, define $L_q^{(k)}(z)$ to be the symmetric matrix, which is linear in $z \in \mathbb{R}^{2k}$, such that
\begin{equation}
\mathcal{L}_z(qp^2) = p^T \left( L_q^{(k)}(z) \right) p \quad \forall p \in \mathbb{R}[x] : \deg(qp^2) \leq 2k.
\end{equation}
(For convenience, we still use $p$ to denote the vector of coefficients of a polynomial $p$, indexed by monomial powers $\alpha \in \mathbb{N}^n$.) The matrix $L_q^{(k)}(z)$ is called the $k$-th \textit{localizing matrix} of $q$ generated by $z$. When $q = 1$, $L_q^{(k)}(z)$ is called a \textit{moment matrix}, and is denoted as $M_k(z)$. The rows and columns of $L_q^{(k)}(z)$ and $M_k(z)$ are indexed by vectors $\alpha \in \mathbb{N}^n$. We refer to [25, 28] for more details about moment and localizing matrices.

Let $K$ be as in (1.1) and $g_0 = 1$. A necessary condition for $z$ to admit a $K$-measure is (cf. [7, 26])
\begin{equation}
L^{(k)}_{h_i}(z) = 0 \quad (i = 1, \ldots, m_1), \quad L^{(k)}_{g_j}(z) \geq 0 \quad (j = 0, 1, \ldots, m_2).
\end{equation}
Define the integer $d_K$ as ($h_i, g_j$ are from (1.1))
\begin{equation}
d_K := \max_{i \in [m_1], j \in [m_2]} \left\{ 1, \lceil \deg(h_i)/2 \rceil, \lceil \deg(g_j)/2 \rceil \right\}.
\end{equation}
In addition to (2.2), if $z$ also satisfies the rank condition
\begin{equation}
\text{rank } M_{k-d_K} = \text{rank } M_k(z),
\end{equation}
then $z$ admits a unique $K$-measure, which is finitely atomic. This is a foundational result of Curto and Fialkow [17]. For convenience, throughout the paper, we simply say $z$ is flat if (2.2) and (2.4) hold for $z$. For a flat $z$, its finitely atomic representing measure can be found by solving eigenvalue problems (cf. Henrion and Lasserre [17]). Flatness is very useful for solving truncated moment problems, as shown by Curto and Fialkow [5, 6, 7]. A nice exposition for flatness can also be found in Laurent [26].

For $z \in \mathbb{R}^{2k}$ and $y \in \mathbb{R}^A$, if $z|_A = y$, we say that $z$ is an \textit{extension} of $y$ and $y$ is a \textit{truncation} of $z$. Clearly, if $z$ is flat and $y = z|_A$, then $y$ admits a $K$-measure. In such case, we say $z$ is a flat extension of $y$. Thus, the existence of a $K$-representing measure for $y$ can be determined by investigating whether $y$ has a flat extension or not. This approach has been exploited in [16, 36].

2.2. Ideals, quadratic modules and positive polynomials. A subset $I \subseteq \mathbb{R}[x]$ is called an \textit{ideal} if $I + I \subseteq I$ and $I \cdot \mathbb{R}[x] \subseteq I$. For a tuple $p = (p_1, \ldots, p_m)$ of polynomials in $\mathbb{R}[x]$, denote by $I(p)$ the ideal generated by $p_1, \ldots, p_m$, which is the set $p_1\mathbb{R}[x] + \cdots + p_m\mathbb{R}[x]$. A polynomial $f \in \mathbb{R}[x]$ is called a \textit{sum of squares} (SOS) if there exist $f_1, \ldots, f_k \in \mathbb{R}[x]$ such that $f = f_1^2 + \cdots + f_k^2$. The cone of all SOS polynomials in $n$ variables and of degree $d$ is denoted by $\Sigma_{n,d}$. We refer to Reznick [40] for survey on SOS polynomials.

Let $h = (h_1, \ldots, h_{m_1})$ and $g = (g_1, \ldots, g_{m_2})$ be as in (1.1). Denote
\begin{equation}
I_{2k}(h) = \left\{ \sum_{i=1}^{m_1} h_i \phi_i \bigg| \text{ each } \deg(h_i \phi_i) \leq 2k \right\},
\end{equation}
\begin{equation}
Q_{2k}(g) = \left\{ \sum_{j=0}^{m_2} g_j \sigma_j \bigg| \text{ each } \deg(\sigma_j g_j) \leq 2k \text{ and } \sigma_j \text{ is SOS} \right\}.
\end{equation}
Lemma 3.1. Characterized as follows. (resp., \( p \) is nonempty and compact, such a measure always exists, as shown in Rogers \[42\].) Let \( \Phi_k(g) \) be as in (1.1). Denote
\[
\Phi_k(g) := \left\{ w \in \mathbb{R}^{2k} \left| L_{g_i}^{(k)}(w) \geq 0, \ j = 0, 1, \ldots, m_2 \right. \right\},
\]
\[
E_k(h) := \left\{ w \in \mathbb{R}^{2k} \left| L_{h_i}^{(k)}(w) = 0, \ i = 1, \ldots, m_1 \right. \right\}.
\]
The set \( I_{2k}(h) + Q_k(g) \) is dual to \( \Phi_k(g) \cap E_k(h) \) (cf. \[25, 28, 36\]), i.e.,
\[
(p, z) \geq 0 \quad \forall p \in I_{2k}(h) + Q_k(g), \forall z \in \Phi_k(g) \cap E_k(h).
\]

3. Properties of \( \mathcal{A}(K) \) and \( \mathcal{P}(K) \)

This section studies properties of the cones \( \mathcal{A}(K) \) and \( \mathcal{P}(K) \).

3.1. Interiors, closedness and duality. Recall that \( \mathbb{R}[x]_A \) is \( K \)-full if there exists \( p \in \mathbb{R}[x]_A \) such that \( p > 0 \) on \( K \). The interiors of \( \mathcal{A}(K) \) and \( \mathcal{P}(K) \) can be characterized as follows.

Lemma 3.1. Let \( K \subseteq \mathbb{R}^n \) be a nonempty compact set. Suppose \( A \subseteq \mathbb{N}^n \) is finite and \( \mathbb{R}[x]_A \) is \( K \)-full. Then we have:

(i) A polynomial \( f \in \mathbb{R}[x]_A \) lies in the interior of \( \mathcal{P}(K) \) if and only if \( f > 0 \) on \( K \).

(ii) An \( A \)-tms \( y \in \mathbb{R}^A \) lies in the interior of \( \mathcal{A}(K) \) if and only if the Riesz functional \( \mathcal{L}_y \) is strictly \( K \)-positive.

Proof. (i) If \( f > 0 \) on \( K \), then \( f \in \text{int}(\mathcal{P}(K)) \). This is because \( f + q > 0 \) on \( K \) for all \( q \in \mathbb{R}[x]_A \) with sufficiently small coefficients (the set \( K \) is compact). Conversely, suppose \( f \in \text{int}(\mathcal{P}(K)) \). Since \( \mathbb{R}[x]_A \) is \( K \)-full, there exists \( p \in \mathbb{R}[x]_A \) with \( p > 0 \) on \( K \). Then \( f - \epsilon p \in \mathcal{P}(K) \) for some \( \epsilon > 0 \). So, \( f \geq \epsilon p > 0 \) on \( K \).

(ii) Let \( \tau \) be a probability measure on \( \mathbb{R}^n \) whose support equals \( K \). (Because \( K \) is nonempty and compact, such a measure always exists, as shown in Rogers \[42\].)

For all \( p \in \mathcal{P}(K) \), \( p|_K \neq 0 \) if and only if \( \int_K p \, d\tau > 0 \). Let
\[
z = \int_K [x]_A p \, d\tau \in \mathbb{R}^A, \quad \mathcal{P}(K, \tau) = \left\{ p \in \mathcal{P}(K) : \int_K p \, d\tau = 1 \right\}.
\]
An \( A \)-tms \( w \) is \( K \)-positive (resp., strictly \( K \)-positive) if and only if \( \mathcal{L}_w(p) \geq 0 \) (resp., \( > 0 \)) for all \( p \in \mathcal{P}(K, \tau) \). So, \( z \) is strictly \( K \)-positive.

\( \Rightarrow \) Suppose \( y \in \text{int}(\mathcal{A}(K)) \). Then \( w := y - \epsilon z \in \mathcal{A}(K) \) for some \( \epsilon > 0 \) and
\[
\mathcal{L}_y(p) = \mathcal{L}_w(p) + \epsilon \mathcal{L}_z(p) \geq \epsilon \mathcal{L}_z(p) > 0
\]
for all \( p \in \mathcal{P}(K, \tau) \). So, \( \mathcal{L}_y \) is strictly \( K \)-positive.

\( \Leftarrow \) Suppose \( \mathcal{L}_y \) is strictly \( K \)-positive. The set \( \mathcal{P}(K, \tau) \) is compact. Let
\[
\epsilon = \min \{ \mathcal{L}_y(p) : p \in \mathcal{P}(K, \tau) \} > 0,
\]
\[
M = \max \{ \|z, p\| : z \in \mathbb{R}^A, \|z\|_2 = 1, p \in \mathcal{P}(K, \tau) \}.
\]
Proposition 3.3. Let $A \subseteq \mathbb{R}$, and $k$ be as in (1.1). Suppose $P_A(K, \tau)$, for each $k \in \mathbb{N}$ and $A \subseteq \mathbb{R}^{2, K}$, by default.) Clearly, $\mathcal{R}_A(K) \subseteq \mathcal{R}_A(k)$ for all $k$. This is because for every $y \in \mathcal{R}_A(K)$, we can always extend $y$ to a tms $z \in \mathcal{R}_2A$ with $z|_A = y$ (cf. [29 Prop. 3.3]). Each $\mathcal{R}_A^k(K)$ is a semidefinite relaxation of $\mathcal{R}_A(K)$, since it is the feasible set of a semidefinite program. Clearly, $\mathcal{R}_A^{k+1}(K) \subseteq \mathcal{R}_A^k(K)$ for all $k$. This results in the hierarchy

$$\mathcal{R}_A^1(K) \supseteq \cdots \supseteq \mathcal{R}_A^k(K) \supseteq \mathcal{R}_A^{k+1}(K) \supseteq \cdots \supseteq \mathcal{R}_A(K).$$

Proposition 3.3. Let $K \neq \emptyset$ be as in (7.7). Suppose $I(h) + Q(g)$ is archimedean, $A \subseteq \mathbb{N}^*$ is finite and $\mathbb{R}[x]_A$ is $K$-full. Then, it holds that

$$\mathcal{R}_A(K) = \bigcap_{k=1}^{\infty} \mathcal{R}_A^k(K).$$
Define the distance \( \delta \). Let \( y \) be such an arbitrary \( p \) of \( (3.6) \). Similarly, there exist \( y \) for all \( p \) with \( \langle p, y \rangle < 0 \). Let \( p_0 \in \mathbb{R}[x] \) be such that \( p_0 > 0 \) on \( K \). Then, for \( \epsilon > 0 \) small, \( p_2 := p_1 + \epsilon p_0 > 0 \) on \( K \) and \( \langle p_2, y \rangle < 0 \). Since \( I(h) + Q(g) \) is archimedean, we have \( p_2 \in Q_{k_1}(g) + I_{2k_1}(h) \) for some \( k_1 \), by Theorem 2.1. If \( y \in \mathcal{P}_A(K) \), then \( y = z|A \) for some \( z \in \Phi_{k_1}(g) \cap E_{k_1}(h) \), and we get

\[
0 > \langle p_2, y \rangle = \langle p_2, z \rangle \geq 0,
\]
a contradiction. The latter inequality is because \( p_2 \in Q_{k_1}(g) + I_{2k_1}(h) \) and \( \Phi_{k_1}(g) \cap E_{k_1}(h) \) is dual to \( Q_{k_1}(g) + I_{2k_1}(h) \). So, \( y \notin \mathcal{P}_A(K) \), and \( (3.4) \) holds.

Proposition 3.3 shows that the semidefinite relaxations \( \mathcal{P}_A(K) \) can approximate \( \mathcal{R}_A(K) \) arbitrarily well. Indeed, we can prove \( \mathcal{P}_A(K) \) converges to \( \mathcal{R}_A(K) \) if we measure their distance by normalization. For \( f \in \mathbb{R}[x]_A \), define

\[
\mathcal{P}_A(K, f) = \{ y \in \mathcal{P}_A(K) : \langle f, y \rangle = 1 \},
\]

\[
\mathcal{R}_A(K, f) = \{ y \in \mathcal{R}_A(K) : \langle f, y \rangle = 1 \}.
\]

Define the distance

\[
(3.5) \quad \delta \left( \mathcal{P}_A(K, f), \mathcal{R}_A(K, f) \right) = \max_{z \in \mathcal{P}_A(K, f)} \min_{y \in \mathcal{R}_A(K, f)} \| z - y \|_2.
\]

**Proposition 3.4.** Let \( K \neq 0 \) be as in (1.1) and \( A \subseteq \mathbb{N}^n \) be finite. Suppose \( I(h) + Q(g) \) is archimedean. If \( f \in \mathbb{R}[x]_A \) is positive on \( K \), then

\[
(3.6) \quad \delta \left( \mathcal{P}_A(K, f), \mathcal{R}_A(K, f) \right) \to 0 \quad \text{as} \quad k \to \infty.
\]

**Proof.** Since \( f > 0 \) on \( K \), there exists \( \epsilon > 0 \) with \( f - \epsilon > 0 \) on \( K \). Since \( I(h) + Q(g) \) is archimedean, by Theorem 2.1 we have \( f - \epsilon \in Q_{N_1}(g) + I_{2N_1}(h) \) for some \( N_1 \). Similarly, there exist \( R > 0 \) and \( N_2 \geq N_1 \) such that for all \( \alpha \in A \)

\[
R \pm x^\alpha \in Q_{N_2}(g) + I_{2N_2}(h).
\]

For all \( y \in \mathcal{P}_A(K, f) \) with \( k \geq N_2 \), there exists \( z \in \Phi_{k}(g) \cap E_{k}(h) \) such that \( z|A = y \). Since \( (f - \epsilon, z) \geq 0 \), we get \( \epsilon z_0 \leq \langle f, z \rangle = \langle f, y \rangle = 1 \) and

\[
0 \leq \langle R \pm x^\alpha, z \rangle = Rz_0 \pm y_\alpha.
\]

(Here \( 0 \) denotes the zero vector in \( \mathbb{N}^n \).) This implies that \( |y_\alpha| \leq R/\epsilon \) for all \( \alpha \in A \). Hence, the sets \( \mathcal{P}_A(K, f) \), with \( k \geq N_2 \), are uniformly bounded.

By (3.3), we know \( \delta \left( \mathcal{P}_A(K, f), \mathcal{R}_A(K, f) \right) \) is monotonically decreasing. Suppose otherwise \( (3.6) \) is not true. Then there exists \( \tau \) such that

\[
\delta \left( \mathcal{P}_A(K, f), \mathcal{R}_A(K, f) \right) \geq \tau > 0
\]

for all \( k \). We can select \( y^k \in \mathcal{P}_A(K, f) \) for all \( k \) such that

\[
\delta \left( y^k, \mathcal{R}_A(K, f) \right) \geq \tau/2.
\]

The sequence \( \{y^k\} \) is bounded, because the sets \( \mathcal{P}_A(K, f) \) \( (k \geq N_2) \) are uniformly bounded, as shown in the above. It has a convergent subsequence, say, \( y^{k_i} \to \hat{y} \) as \( i \to \infty \). Clearly, \( \langle f, \hat{y} \rangle = 1 \) and \( \delta \left( \hat{y}, \mathcal{R}_A(K, f) \right) > 0 \). So, \( \hat{y} \notin \mathcal{R}_A(K, f) \). This
implies $\hat{y} \not\in \mathcal{R}_A(K) = \mathcal{P}_A(K)^*$, by (3.11). So, there exists $p_0 \in \mathcal{P}_A(K)$ such that $\langle p_0, \hat{y} \rangle < 0$. For a small $\epsilon_0 > 0$, we have
$$p_1 := p_0 + \epsilon_0 f > 0 \quad \text{on } K, \quad \langle p_1, \hat{y} \rangle < 0.$$ By Theorem 2.11 we have $p_1 \in Q_{N_3}(g) + I_{2N_3}(h)$ for some $N_3$. So, $\langle p_1, y^{k_i} \rangle \geq 0$ for all $k_i \geq N_3$. This results in
$$\langle p_1, \hat{y} \rangle = \lim_{i \to \infty} \langle p_1, y^{k_i} \rangle \geq 0,$$
which is a contradiction. Thus, $\mathcal{R}_A(K)$ must be true. \hfill \Box

Second, we consider semidefinite relaxations for the cone $\mathcal{R}_A(K)$. Denote
$$\mathcal{D}^k_A(K) = \{ p \in \mathbb{R}[x]_A : p \in Q_k(g) + E_{2k}(h) \}.$$ Clearly, $\mathcal{D}^k_A(K) \subseteq \mathcal{P}_A(K)$ for all $k$. Suppose $K$ is compact and $\mathbb{R}[x]_A$ is $K$-full. If $p$ is in the interior of $\mathcal{P}_A(K)$, then $p > 0$ on $K$ by Lemma 3.11 and $p \in \mathcal{D}^k_A(K)$ for some $k$ by Theorem 2.11 if $I(h) + Q(g)$ is archimedean. So, we get the following proposition.

**Proposition 3.5.** Let $K \neq \emptyset$ be as in (1.7) and $A \subset \mathbb{N}^n$ be finite. Suppose $\mathbb{R}[x]_A$ is $K$-full and $I(h) + Q(g)$ is archimedean. Then, we have
$$\text{int} \ (\mathcal{P}_A(K)) \subseteq \bigcup_{k=1}^{\infty} \mathcal{D}^k_A(K) \subseteq \mathcal{P}_A(K).$$

The second containment inequality in (3.8) generally cannot be changed to an equality. For instance, when $K = B(0,1)$ and $A = \mathbb{N}_0^d$, the Motzkin polynomial $x_1^2 x_2^2 (x_1^2 + x_2^2 - 3x_3^2) + x_3^6 \in \mathcal{P}_A(K)$ but it does not belong to $\mathcal{D}^k_A(K)$ for any $k$ (cf. (33) Example 5.3).

### 3.3. Checking memberships

First, we discuss how to check membership in the cone $\mathcal{R}_A(K)$. Assume $K \subseteq B(0,\rho)$ for some $\rho > 0$, for $K$ as in (1.11). For convenience, let $g_B = (g, \rho^2 - \|x\|^2_2)$. Whether $y \in \mathcal{R}_A(K)$ or not can be checked by solving a sequence of semidefinite optimization problems, as shown in (36).

**Algorithm 3.6.** (36) An algorithm for checking memberships in $\mathcal{R}_A(K)$.

**Input:** $y \in \mathbb{R}^A$ and an even $d > \deg(A)$ (the default is $2[\deg(A)/2]$).

**Output:** A $K$-representing measure for $y$, or an answer that $y \notin \mathcal{R}_A(K)$.

**Procedure:**

**Step 0:** Choose a generic polynomial $R \in \Sigma_{n,d}$ and let $k := d/2$.

**Step 1:** Solve the semidefinite program
$$\min \ (R, w) \quad \text{s.t.} \quad w|_A = y, \ w \in E_k(h) \cap \Phi_k(g_B).$$

If (3.9) is infeasible, output that $y \notin \mathcal{R}_A(K)$ and stop. If (3.9) is feasible, get a minimizer $w^{*,k}$. Let $t := \min\{d_K, \deg(A)\}$.

**Step 2:** Check whether the truncation $w^{*,k}|_{2t}$ is flat or not. If yes, go to Step 3; otherwise, go to Step 4.

**Step 3:** Compute the finitely atomic $K$-representing measure $\mu$ for $w^{*,k}|_{2t}$. Output $\mu$, and stop.

**Step 4:** If $t < k$, let $t := t + 1$ and go to Step 2; otherwise, let $k := k + 1$ and go to Step 1.
Algorithm 3.3 has the following properties (cf. [30]). (1) Suppose $\mathbb{R}[x]_{\mathcal{A}}$ is $K$-full. If $y$ admits no $K$-measures, then $\mathcal{P}_{\mathcal{A}}(K)$ is infeasible for all $k$ big enough. This gives a certificate for $y \notin \mathcal{P}_{\mathcal{A}}(K)$. (2) If $y$ admits a $K$-measure, then, for all generic $R \in \Sigma_{n,a}$, we have: i) for all $t$ big enough, the sequence $\{w^{r,k}|_{2t}\}_k$ is bounded and each of its accumulation points is flat; ii) under some general conditions, $w^{r,k}|_{2t}$ is flat for some $k$ and $t$; iii) the obtained measures are $r$-atomic with $r \leq |A|$. 

Second, we discuss how to check membership in the cone $\mathcal{P}_{\mathcal{A}}(K)$. Note that a polynomial $f \in \mathbb{R}[x]_{\mathcal{A}}$ belongs to $\mathcal{P}_{\mathcal{A}}(K)$ if and only if its minimum $f_{\min}$ on $K$ is nonnegative. A standard approach for computing $f_{\min}$ is to apply Lasserre’s hierarchy of relaxations ($k = 1, 2, \cdots$):

\begin{equation}
(3.10) \quad f_k = \max_{\gamma} \gamma \quad \text{s.t.} \quad f - \gamma \in I_{2k}(h) + Q_k(g).
\end{equation}

Clearly, if $f_k \geq 0$ for some $k$, then $f \in \mathcal{P}_{\mathcal{A}}(K)$. Suppose $I(h) + Q(g)$ is archimedean. For all $f \in \text{int}(\mathcal{P}_{\mathcal{A}}(K))$, we have $f_k > 0$ for some $k$, by Proposition 3.5. For $f$ lying generically on the boundary of $\mathcal{P}_{\mathcal{A}}(K)$ (e.g., some standard optimality conditions hold), we have $f_k \geq 0$ for some $k$ (cf. [33]). For the remaining non-generic cases, it is possible that $f_k < f_{\min}$ for all $k$ (cf. [33] Examples 5.3, 5.6).

Another method for computing $f_{\min}$ is applying Jacobian SDP relaxation [33]. Its basic idea is to add new polynomial equalities, by using the Jacobian of polynomials $f, h, g_j$. Suppose $\varphi := (\varphi_1, \ldots, \varphi_L) = 0$ is added (cf. [33] Section 2)). Under some generic conditions on $K$ but not on $f$ (cf. Assumption 2.2 of [33]), $f_{\min}$ equals the optimal value of

\begin{equation}
(3.11) \quad \min_{x \in \mathbb{R}^n} \varphi(x) = 0, h(x) = 0, g(x) \geq 0.
\end{equation}

This leads us to consider stronger relaxations ($k = 1, 2, \cdots$):

\begin{equation}
(3.12) \quad f_{\mathrm{jac}} := \max_{\gamma} \gamma \quad \text{s.t.} \quad f - \gamma \in I_{2k}(h) + I_{2k}(\varphi) + Q_k(g).
\end{equation}

An advantage of this approach is that $\{f_{\mathrm{jac}}^k\}$ always have finite convergence to $f_{\min}$ under the archimedeaness (cf. [33] Section 4)). So, we can check whether $f \in \mathcal{P}_{\mathcal{A}}(K)$ or not by solving finitely many semidefinite programs.

4. Linear Optimization Problems

Let $K$ be as in (4.1) and $\mathcal{A} \subseteq \mathbb{N}^n$ be finite. Given $a_1, \ldots, a_m, c \in \mathbb{R}[x]_{\mathcal{A}}$ and $b \in \mathbb{R}^m$, we consider the linear optimization problem

\begin{equation}
(4.1) \quad \begin{cases}
    c_{\min} := \min_{y \in \mathcal{R}_{\mathcal{A}}(K)} \langle c, y \rangle \\
    \text{s.t.} \quad \langle a_i, y \rangle = b_i, \quad i = 1, \ldots, m,
\end{cases}
\end{equation}

The dual problem of (4.1) is

\begin{equation}
(4.2) \quad \begin{cases}
    b^{\max} := \max_{\lambda} b^T \lambda \\
    \text{s.t.} \quad c - \sum_{i=1}^m \lambda_i a_i \in \mathcal{P}_{\mathcal{A}}(K).
\end{cases}
\end{equation}

The cones $\mathcal{R}_{\mathcal{A}}(K)$ and $\mathcal{P}_{\mathcal{A}}(K)$ are typically quite difficult to describe. However, they have nice semidefinite relaxations, as shown in Section 3.2. Indeed, the semidefinite relaxation $\mathcal{P}_{\mathcal{A}}^k(K)$ (resp., $\mathcal{P}_{\mathcal{A}}^k(K)$) converges to $\mathcal{R}_{\mathcal{A}}(K)$ (resp., $\mathcal{P}_{\mathcal{A}}(K)$) as
This gives a convenient way to terminate the loop. It might be possible that 

\[ b^k = \max_{\lambda = (\lambda_1, \ldots, \lambda_m)} b^T \lambda \]

\( s.t. \ c = \sum_{i=1}^{m} \lambda_i a_i \in Q_k(g) + I_2k(h). \)

Note that \( y \) is feasible in (4.3) if and only if \( y \in \mathcal{S}_A(K) \). The integer \( k \) in (4.3) is called a relaxation order. The dual problem of (4.3) is

\[ A \min \{ c|y \} \]

\( s.t. \ \langle a_i, y \rangle = b_i, \ i = 1, \ldots, m \)

\[ y = w|_A, \ w \in \Phi_k(g) \cap E_k(h). \]

Clearly, we have \( c^k \leq c^{\min} \) and \( b^k \leq b^{\max} \) for all \( k \). Let \( (y^{*,k}, w^{*,k}) \) be a minimizer of (4.3), and let \( \lambda^{*,k} \) be a maximizer of (4.4). If \( y^{*,k} \in \mathcal{R}_A(K) \), then \( c^k = c^{\min} \) and \( y^{*,k} \) is a minimizer of (4.1). If \( \lambda^{*,k} \) is a maximizer of (4.2), or an answer for the infeasibility of (4.1). Considering the above, we get the following algorithm.

**Algorithm 4.1.** A semidefinite algorithm for solving (4.1)–(4.2).

**Input:** \( c, a_1, \ldots, a_m \in \mathbb{R}[x]_A, b \in \mathbb{R}^m \) and \( K \) as in (1.1).

**Output:** A minimizer \( y^* \) of (4.1) and a maximizer \( \lambda^* \) of (4.2), or an answer for the infeasibility of (4.1).

**Procedure:**

1. **Step 0:** Let \( k = \lceil \deg(A)/2 \rceil \).
2. **Step 1:** Solve the primal-dual pair (4.3)–(4.4). If (4.3) is infeasible, stop and output that (4.1) is infeasible; otherwise, compute an optimal pair \( (y^{*,k}, w^{*,k}) \) for (4.3) and a maximizer \( \lambda^{*,k} \) for (4.4).
3. **Step 2:** If \( y^{*,k} \in \mathcal{R}_A(K) \), then \( y^{*,k} \) is a minimizer of (4.1); if in addition \( b^k = c^k \), then \( \lambda^{*,k} \) is a maximizer of (4.2) and output \( y^* = y^{*,k}, \lambda^* = \lambda^{*,k} \). Otherwise, let \( k := k + 1 \) and go to Step 1.

**Remark 4.2.** Checking if \( y^{*,k} \in \mathcal{R}_A(K) \) or not is a stopping criterion for Algorithm 4.1. If there exists \( t \geq \deg(A)/2 \) such that \( w^{*,k}|_{2t} \) is flat, then \( y^{*,k} \in \mathcal{R}_A(K) \). This gives a convenient way to terminate the loop. It might be possible that \( y^{*,k} \) belongs to \( \mathcal{R}_A(K) \) while \( w^{*,k}|_{2t} \) is not flat for all \( t \) (cf. Example 4.7). In such cases, we can apply Algorithm 3.6 to check if \( y^{*,k} \in \mathcal{R}_A(K) \) or not.

Feasibility and infeasibility issues of (4.1)–(4.2) are more delicate. They will be studied again in Section 5. In Section 4.1, we prove asymptotic and finite convergence of Algorithm 4.1. In Section 4.2, we present some examples of applying Algorithm 4.1.

### 4.1. Convergence Analysis

First, we prove the asymptotic convergence of Algorithm 4.1. When \( A = N_0^d \) and one of \( c, a_1, \ldots, a_m \) is positive on \( K \), Lasserre [22, Theorem 1] showed that the optimal values \( c^k \to c^{\min} \) as \( k \to \infty \). This is also true for general \( A \). Indeed, we can prove the stronger result that the sequence \( \{ y^{*,k} \} \) produced by Algorithm 4.1 converges to the cone \( \mathcal{R}_A(K) \), under more general assumptions.
Theorem 4.3. Let $K$ be as in (4.2) and $A \subseteq \mathbb{N}^n$ be finite. Suppose (4.1) is feasible, (4.2) has an interior point, $\mathbb{R}[x]_A$ is $K$-full, and $Q(g) + I(h)$ is archimedean. Then, we have:

(i) For all $k$ sufficiently large, (4.4) has an interior point and (4.3) has a minimizing pair $(y^{*,k}, w^{*,k})$.

(ii) The sequence $\{y^{*,k}\}$ is bounded, and each of its accumulation points is a minimizer of (4.3).

(iii) The sequence $\{b^k\}$ converges to the maximum $b^{\text{max}}$ of (4.2).

Proof. (i) Let $\lambda^0$ be an interior point of (4.2). Then $c(\lambda^0) = c - \sum_{i=1}^{m} \lambda^0_i a_i > 0$ on $K$, by Lemma 3.1. The archimedeanness of $I(h) + Q(g)$ implies that $K$ is compact. So, there exist $\epsilon_0 > 0$ and $\theta > 0$ such that

$$c(\lambda) - \epsilon_0 > \epsilon_0 \quad \forall \lambda \in B(\lambda^0, \theta).$$

By Theorem 6 of [31], there exists $N_0 > 0$ such that

$$c(\lambda) - \epsilon_0 \in I_{2N_0}(h) + Q_{N_0}(g) \quad \forall \lambda \in B(\lambda^0, \theta).$$

So, (4.4) has an interior point for all $k \geq N_0$, and the strong duality holds between (4.3) and (4.4). Since (4.1) is feasible, the relaxation (4.3) is also feasible and has a minimizing pair $(y^{*,k}, w^{*,k})$ (cf. [1] Theorem 2.4.1).

(ii) First, we show that $\{y^{*,k}\}$ is a bounded sequence. Let $c(\lambda^0)$ and $\epsilon_0$ be as in the proof of (i). The set $I_{2N_0}(h) + Q_{N_0}(g)$ is dual to $E_{2N_0}(h) \cap \Phi_{N_0}(g)$. For all $k \geq N_0$, we have $w^{*,k} \in E_{N_0}(h) \cap \Phi_{N_0}(g)$ and

$$0 \leq \langle c(\lambda^0) - \epsilon_0, w^{*,k} \rangle = \langle c(\lambda^0), w^{*,k} \rangle - \epsilon_0 \langle 1, w^{*,k} \rangle,$$

$$\langle c(\lambda^0), w^{*,k} \rangle = \langle c, w^{*,k} \rangle - \sum_{i=1}^{m} \lambda^0_i \langle a_i, y^{*,k} \rangle = \langle c, w^{*,k} \rangle - b^T \lambda^0.$$ 

Since $(c, w^{*,k}) \leq c^{\text{min}}$, it holds that

$$\langle c(\lambda^0), w^{*,k} \rangle \leq T_0 := c^{\text{min}} - b^T \lambda^0.$$

Combining the above, we get (denote by $0$ the zero vector in $\mathbb{N}^n$)

$$0 \leq \langle c(\lambda^0) - \epsilon_0, w^{*,k} \rangle \leq T_0 - \epsilon_0 \langle w^{*,k}, 0 \rangle,$$

$$(w^{*,k})_0 \leq T_1 := T_0 / \epsilon_0.$$ 

Since $I(h) + Q(g)$ is archimedean, there exist $\rho > 0$ and $k_1 \in \mathbb{N}$ such that

$$\rho - \|x\|^2_2 \in I_{2k_1}(h) + Q_{k_1}(g).$$

So, for all $k \geq k_1$, we get

$$0 \leq \langle \rho - \|x\|^2_2, w^{*,k} \rangle = \rho \langle w^{*,k}, 0 \rangle - \sum_{|\alpha|=1} (w^{*,k})_{2\alpha}, \quad \sum_{|\alpha|=1} (w^{*,k})_{2\alpha} \leq \rho T_1.$$ 

For each $t = 1, \ldots, k - k_1$, we have

$$\|x\|^2_2 \rho - \|x\|^2_2 \in I_{2k_1}(h) + Q_{k_1}(g).$$

The membership $w^{*,k} \in \Phi_{k_1}(g) \cap I_{k_1}(h)$ implies that

$$\rho \langle \|x\|^2_2, w^{*,k} \rangle - \|x\|^2_2, w^{*,k} \rangle \geq 0, \quad t = 1, \ldots, k - k_1.$$ 

The above then implies that

$$\langle \|x\|^2_2, w^{*,k} \rangle \leq \rho^t T_1, \quad t = 1, \ldots, k - k_1.$$
Let $z^k := u^{*,k}|_{2k-2k_1}$, then the moment matrix $M_{k-k_1}(z^k) \succeq 0$ and

$$\|z^k\|_2 \leq \|M_{k-k_1}(z^k)\|_F \leq \text{Trace}(M_{k-k_1}(z^k)) = \sum_{i=0}^{k-k_1} \sum_{|\alpha|=i} (w^{*,k})_{2\alpha},$$

$$\sum_{|\alpha|=i} (w^{*,k})_{2\alpha} = \langle \sum_{|\alpha|=i} x^{2\alpha}, z^k \rangle \leq (\|x\|_2^2, z^k) \leq \rho T_1.$$  

The above then implies that

$$\|z^k\|_2 \leq (1 + \rho + \cdots + \rho^{k-k_1}) T_1.$$  

Fix $k_2 > k_1$ such that $u^{*,k}$ is a subvector of the truncation $z^k|_{k_2-k_1}$. From $u^{*,k} = z^k|_A$, we get

$$\|u^{*,k}\|_2 \leq \|z^k\|_2 \leq (1 + \rho + \cdots + \rho^{k_2-k_1}) T_1.$$  

This shows that the sequence $\{u^{*,k}\}$ is bounded.

Second, we show that every accumulation point of $\{y^{*,k}\}$ is a minimizer of (4.1). Let $y^*$ be such an arbitrary one. We can generally further assume $y^{*,k} \to y^*$ as $k \to \infty$. We need to show that $y^*$ is a minimizer of (4.1). Since $K$ is compact, by the archimedeaness of $I(h) + Q(g)$, we can generally assume $K \subseteq B(0, \rho)$ with $\rho < 1$, up to a scaling. In the above, we have shown that

$$\|z^k\|_2 \leq T_1/(1 - \rho).$$

This implies that the sequence $\{z^k\}$ is bounded. Each tms $z^k$ can be extended to a vector in $\mathbb{R}^{N_n}_\infty$ by adding zero entries to the tailing. The set $\mathbb{R}^{N_n}_\infty$ is a Hilbert space, equipped with the inner product

$$\langle u, v \rangle := \sum_{\alpha \in \mathbb{N}^n} u_\alpha v_\alpha, \quad \forall u, v \in \mathbb{R}^{N_n}_\infty.$$  

So, the sequence $\{z^k\}$ is also bounded in $\mathbb{R}^{N_n}_\infty$. By Alaoglu’s Theorem (cf. [3, Theorem V.3.1] or [25, Theorem C.18]), it has a subsequence $\{z^{k_j}\}$ that is convergent in the weak-* topology. That is, there exists $z^* \in \mathbb{R}^{N_n}_\infty$ such that

$$\langle f, z^{k_j} \rangle \to \langle f, z^* \rangle \quad \text{as} \quad j \to \infty$$

for all $f \in \mathbb{R}^{N_n}_\infty$. Clearly, this implies that for each $\alpha \in \mathbb{N}^n$

$$z^{k_j}_\alpha \to (z^*)_\alpha.$$  

Since $z^k|_A = u^{*,k} \to y^*$, we get $z^*|_A = y^*$. Note that $z^{k_j} \in \Phi_{k_j}(g) \cap E_{k_j}(h)$ for all $j$. For each $r = 1, 2, \ldots$, if $k_j \geq 2r$, then (cf. Section 2.1)

$$L_{h_j}^{(r)}(z^{k_j}) \geq 0 (1 \leq i \leq m_1), \quad L_{g_i}^{(r)}(z^{k_j}) \geq 0 (0 \leq i \leq m_2).$$

Hence, (4.5) implies that for all $r = 1, 2, \ldots$

$$L_{h_i}^{(r)}(z^*) \geq 0 (1 \leq i \leq m_1), \quad L_{g_i}^{(r)}(z^*) \geq 0 (0 \leq i \leq m_2).$$

This means that $z^* \in \mathbb{R}^{N_n}_\infty$ is a full moment sequence whose localizing matrices of all orders are positive semidefinite. By Lemma 3.2 of Putinar [39], $z^*$ admits a $K$-measure. Clearly, $\langle a_i, y^* \rangle = b_i$ for all $i$. So, $y^* = z^*|_A$ is feasible for (4.1) and $c^{min} \leq \langle c, y^* \rangle$. Because (4.3) is a relaxation of (4.1) and $w^{*,k}$ is a minimizer of (4.3), it holds that

$$c^{min} \geq \langle c, y^{*,k} \rangle, \quad k = 1, 2, \ldots$$

Hence, we get

$$c^{min} \geq \lim_{k \to \infty} \langle c, y^{*,k} \rangle = \langle c, y^* \rangle.$$
Therefore, \( c_{\min} = (c, y^*) \) and \( y^* \) is a minimizer of \((4.1)\).

(iii) For each \( \epsilon > 0 \), there exists \( \lambda^* \) such that \( c(\lambda^*) \in \mathcal{P}_A(K) \) and
\[
\begin{align*}
0 < b^T \lambda^* - \epsilon < b^T \lambda^* < b^{max}.
\end{align*}
\]
Let \( \lambda^0 \) be as in the proof of item (i), and let \( \lambda(\epsilon) = (1-\epsilon)\lambda^* + \epsilon\lambda^0 \). Then \( c(\lambda(\epsilon)) > 0 \) on \( K \) and
\[
\begin{align*}
0 < b^T \lambda(\epsilon) = (1-\epsilon)b^T \lambda^* + \epsilon b^T \lambda^0 > (1-\epsilon)(b^{max} - \epsilon) + \epsilon b^T \lambda^0.
\end{align*}
\]
By Theorem 2.1 if \( k \) is big enough, then \( c(\lambda(\epsilon)) \in I_{2k}(h) + Q_k(g) \) and
\[
\begin{align*}
0 < b^T \lambda(\epsilon) > (1-\epsilon)(b^{max} - \epsilon) + \epsilon b^T \lambda^0.
\end{align*}
\]
Since \( b^k \leq b^{max} \) for all \( k \), we get \( b^k \to b^{max} \) as \( k \to \infty \). □

Second, we prove the finite convergence of Algorithm 4.1 under a general assumption.

**Assumption 4.4.** Suppose \( \lambda^* \) is a maximizer of \((4.2)\) and \( c^* := c(\lambda^*) \) satisfies:

(i) There exists \( k_1 \in \mathbb{N} \) such that \( c^* \in I_{2k_1}(h) + Q_k(g) \);

(ii) The optimization problem
\[
\begin{align*}
\min \ c^*(x) \quad \text{s.t.} \quad h(x) = 0, \ g(x) \geq 0
\end{align*}
\]
has finitely many KKT points \( u \) with \( c^*(u) = 0 \).

**Theorem 4.5.** Let \( K \) be as in \((4.1)\). Suppose \((4.1)\) is feasible, \((4.2)\) has an interior point, \( \mathbb{R}[x]_A \) is \( K \)-full and Assumption 4.4 holds. If \( w^{*,k} \) is optimal for \((4.3)\), then \( w^{*,k} \) is flat for all \( k \to t \) big enough.

**Proof.** The existence of a minimizer \((y^{*,k}, w^{*,k}) \) is shown in Theorem 4.3. Because \((4.1)\) is feasible and \((4.2)\) has an interior point, \( (4.1) \) has a minimizer \( y^* \) and there is no duality gap between \((4.1)\) and \((4.2)\), i.e.,
\[
0 = \langle c, y^* \rangle - b^T \lambda^* = \langle c^*, y^* \rangle.
\]
Clearly, \( c^* \geq 0 \) on \( K \). Let \( \mu^* \) be a \( K \)-representing measure for \( y^* \). Then, every point in \( \text{supp}(\mu^*) \) is a minimizer of \((4.6)\), and the minimum value is 0. The \( k \)-th Lasserre’s relaxation for \((4.6)\) is (cf. 19)
\[
\begin{align*}
\gamma_k := \max \quad \gamma \quad \text{s.t.} \quad c^* - \gamma \in I_{2k}(h) + Q_k(g).
\end{align*}
\]
Then, \( \gamma_k = 0 \) for all \( k \geq k_1 \). The sequence \( \{\gamma_k\} \) has finite convergence. The relaxation \((4.7)\) achieves its optimal value for all \( k \geq k_1 \), by Assumption 4.4 (i). The dual problem of \((4.7)\) is
\[
\begin{align*}
\min_w \quad \langle c^*, w \rangle \quad \text{s.t.} \quad w \in \Phi_k(g) \cap E_k(h), w_0 = 1.
\end{align*}
\]
By Assumption 4.4, \((4.6)\) has only finitely many critical points on which \( c^* = 0 \). So, Assumption 2.1 in 34 for the problem \((4.6)\) is satisfied. Suppose \( w^{*,k} \) is optimal for \((4.3)\).

If \((w^{*,k})_0 = 0\), then \( \text{vec}(1)^T M_k(w^{*,k}) \text{vec}(1) = 0 \) and \( M_k(w^{*,k}) \text{vec}(1) = 0 \), because \( M_k(w^{*,k}) \geq 0 \). (Here \( \text{vec}(p) \) denotes the coefficient vector of a polynomial \( p \)).

\[\text{Proof}^*\]
This implies that $M_k(w^*, k) vec(x^n) = 0$ for all $|\alpha| \leq k - 1$ (cf. Lemma 5.7). For all $|\alpha| \leq 2k - 2$, we can write $\alpha = \beta + \eta$ with $|\beta|, |\eta| \leq k - 1$, and get

$$(w^*, k)_\alpha = vec(x^\beta)^T M_k(w^*, k) vec(x^n) = 0.$$ 

So, the truncation $w^*, k|_{2k-2}$ is flat.

If $(w^*, k)_0 > 0$, we can scale $w^*, k$ such that $(w^*, k)_0 = 1$. Then $w^*, k$ is a minimizer of (4.3) because $(c^*, w^*, k) = 0$ for all $k \geq k_1$. By Theorem 2.2 of [34], $w^*, k$ has a flat truncation $w^*, k|_{2t}$ for some $t \geq \deg(A)/2$, for all $k$ sufficiently large. $\square$

In generic cases, the conditions in Assumption 1 hold (cf. [32] [35]). Theorem 4.6 implies that Algorithm 4.1 often converges in finitely many steps. This has been observed in the numerical experiments.

4.2. Some examples. Semidefinite relaxations 13 and 14 can be solved by GloptiPoly 3 IN.

Example 4.6. Let $K$ be the simplex $\Delta_n = \{ x \in \mathbb{R}_+^n : x_1 + \cdots + x_n = 1 \}$ and $A = \{ \alpha \in \mathbb{N}_n : |\alpha| = 2 \}$. Then $\mathcal{P}_A(\Delta_n)$ is the cone of $n \times n$ copositive matrices (denoted as $\mathcal{C}_n(n)$), and $\mathcal{R}_A(\Delta_n)$ is the cone of $n \times n$ completely positive matrices (denoted as $\mathcal{C}_p(n)$). The simplex $\Delta_n$ is defined by the tuples $h = (x_1 + \cdots + x_{n-1} = 0,$ $\ldots,$ $x_{n-1} + x_n = 0)$. We want to know what is the maximum $\lambda$ such that $c - \lambda a_1 \in \mathcal{C}_0(6)$.

(i) Let $c = (x_1 + \cdots + x_6)^2$ and $a_1 = x_1 x_2 - x_2 x_3 + x_3 x_4 - x_4 x_5 + x_5 x_6 + x_6 x_1$. We formulate this problem in the form (4.2) and then solve it by Algorithm 4.1. For $k = 2$, $y^*, k \in \mathcal{C}_p(6)$ (because it admits the measure $4h(1/2, 1/4, 0, 0, 0, 0, 1/4)$) and $\lambda^*, k = 4$. Since $c^k = b^k$ for $k = 2$, we know the maximum $\lambda$ in the above is 4.

(ii) Consider the matrix

$$C = \begin{bmatrix}
* & 1 & 2 & 3 & 4 \\
1 & * & 1 & 2 & 3 \\
2 & 1 & * & 1 & 2 \\
3 & 2 & 1 & * & 1 \\
4 & 3 & 2 & 1 & *
\end{bmatrix}$$

where each * means the entry is not given. We want to know what is the smallest trace of $C$ when $C \in \mathcal{C}_p(5)$. We formulate this problem in the form (4.2) and then solve it by Algorithm 4.1. For $k = 2$, $y^*, k \in \mathcal{R}_A(\Delta_5)$ (verified by Algorithm 3.6). So, the minimum trace of $C$ is $\mathcal{C}_p(5) = 20.8172$ with the diagonal entries $C_{11}, \ldots, C_{55}$ being $6.0317$, $3.9688$, $0.8162$, $3.9688$, $6.0317$ respectively. $\square$

Example 4.7. Let $K = B(0, 1)$ be the unit ball in $\mathbb{R}^2$ and $A = \mathbb{N}_0^2$. We want to know what is the maximum $\lambda_1 + \lambda_2$ such that

$$x_1^4 x_2^3 + 6x_1^2 x_2^5 + 4x_1 x_2^4 + x_2^6 - \lambda_1 (x_1^2 x_2^2 + x_1 x_2^3) - \lambda_2 (x_1^2 x_2^2 + x_2^3) \in \mathcal{P}_A(K).$$

We formulate this problem in the form (4.2) and then solve it by Algorithm 4.1. When $k = 3$, $y^*, k \in \mathcal{R}_A(\Delta_k)$ (verified by Algorithm 3.6), and $\lambda^*, k = (4, 2)$. Since $c^k = b^k$ for $k = 2$, the optimal $(\lambda_1, \lambda_2)$ in the above is $(4, 2)$. $\square$

Example 4.8. Let $K = S^2$ and $A = \{ \alpha \in \mathbb{N}^3 : |\alpha| = 6 \}$. Then, $\mathcal{P}_A(K) = P_{3, 6}$ and $\mathcal{R}_A(K)$ is the cone of sextic tms’ admitting measures supported on $S^2$.

(i) The form $c = x_1^6 + x_2^6 + x_3^6$ is in the interior of $P_{3, 6}$. Let

$$a_1 = x_1^2 x_2^4 + x_2^2 x_3^4 + x_3^2 x_1^4, a_2 = x_1^3 x_2^3 + x_2^3 x_3^3 + x_3^3 x_1^3, a_3 = x_1^5 x_2 + x_2^5 x_3 + x_3^5 x_1.$$

2Throughout the paper, only four decimal digits are shown for numerical results.
We want to know what is the maximum \( \lambda_1 + \lambda_2 + \lambda_3 \) such that
\[
c - \lambda_1 a_1 - \lambda_2 a_2 - \lambda_3 a_3 \in P_{3,6}.
\]
We formulate this problem in the form (4.2) and then solve it by Algorithm 4.1.

When \( k = 3 \), \( y^{*3} \in \mathcal{A}(K) \) (it admits the measure \( 9\delta_{(1,1,1)/\sqrt{3}} \)), and \( \lambda^{*,3} = (-1.4404, 2.2190, 0.2214) \). Since \( c^k = b^k \) for \( k = 2 \), we know \( \lambda^{*,3} \) is also optimal for the above.

(ii) We want to know what is the minimum value of \( \int (x_1^6 + x_2^6 + x_3^6) \mu \) for all measures \( \mu \) supported on \( S^2 \) such that
\[
\int x_1^3x_2^3d\mu = \int x_2^3x_3^3d\mu = \int x_3^3x_1^3d\mu,
\]
\[
\int x_1^2x_2^2x_3^2d\mu = 1, \quad \int (x_1^4x_2^2 + x_2^4x_3^2 + x_3^4x_1^2)d\mu = 3.
\]
We formulate this problem in the form (4.2) and then solve it by Algorithm 4.1.

When \( k = 3 \), \( y^{*,3} \in \mathcal{A}(K) \) because it admits the measure
\[
\frac{27}{4} \left( \delta_{(1,1,1)/\sqrt{3}} + \delta_{(-1,1,1)/\sqrt{3}} + \delta_{(1,-1,1)/\sqrt{3}} + \delta_{(-1,1,-1)/\sqrt{3}} \right).
\]
So, the minimum of \( \int (x_1^6 + x_2^6 + x_3^6) \mu \) for \( \mu \) satisfying the above is 3.

If a linear optimization problem with cone \( \mathcal{A}(K) \) is given in the form (4.2), it can also be equivalently formulated in the form (4.1). For instance, given \( z_0, \ldots, z_m \in \mathbb{R}^4 \) and \( \ell = (\ell_1, \ldots, \ell_m) \in \mathbb{R}^m \), consider the problem
\[
\max \quad \ell_1\lambda_1 + \cdots + \ell_m\lambda_m
\]
\[
s.t. \quad z_0 - \lambda_1z_1 - \cdots - \lambda_mz_m \in \mathcal{A}(K).
\]
Let \( \{p_1, \ldots, p_r\} \) be a basis of the orthogonal complement of \( \text{span}\{z_1, \ldots, z_m\} \). Then, \( y = z_0 - \lambda_1z_1 - \cdots - \lambda_mz_m \) for some \( \lambda_1, \ldots, \lambda_m \) if and only if
\[
p_1^Ty = p_1^Tz_0, \ldots, p_m^Ty = p_m^Tz_0.
\]
We can consider each \( p_i \) as a polynomial in \( \mathbb{R}[x]_A \). Let \( Z = [z_1 \cdots z_m] \). Assume \( \text{rank}(Z) = m \). (Otherwise, choose \( Z \) to be a basis of \( \text{span}\{z_1, \ldots, z_m\} \).) If \( y = z_0 - Z\lambda \), then
\[
\lambda = (Z^TZ)^{-1}Z^T(z_0 - y).
\]
Let \( p_0 \) be a polynomial in \( \mathbb{R}[x]_A \) such that
\[
\langle p_0, y \rangle = \ell^T(Z^TZ)^{-1}Z^Ty = \ell^T\lambda.
\]
Then (4.9) is equivalent to
\[
\min \quad \langle p_0, y \rangle
\]
\[
s.t. \quad \langle p_i, y \rangle = p_i^Tz_0, \quad i = 1, \ldots, m,
\]
\[
y \in \mathcal{A}(K).
\]
If \( y^* \) is a minimizer of (4.10), then
\[
\lambda^* = (Z^TZ)^{-1}Z^T(z_0 - y^*)
\]
is a maximizer of (4.9). Similarly, every linear optimization problem with cone \( \mathcal{A}(K) \), which is given in the form (4.1), can also be equivalently formulated in the form (4.2).
Example 4.9. Let $K = S^{n-1}$ and $\mathcal{A} = \{\alpha \in \mathbb{N}^n : |\alpha| = d\}$ $(d$ is even). Then $\mathbb{R}_A(K)$ is equivalent to $Q_{n,d}$, the cone of sums of $d$-th power of real linear forms in $n$ variables (cf. [39 Sec. 6.2]).

(i) The sextic form $(x_1^2 + x_2^2 + x_3^2)^3$ belongs to $Q_{3,4}$ (cf. [41]). We want to know what is the maximum $\lambda$ such that

$$(x_1^2 + x_2^2 + x_3^2)^3 - \lambda(x_1^6 + x_2^6 + x_3^6) \in Q_{3,6}.$$ 

The problem is equivalent to finding the biggest $\lambda$ such that $z_0 - \lambda z_1 \in \mathbb{R}_A(K)$, where $z_0, z_1$ are tms’ whose entries are zeros except $$(z_0)_{(0,0,0)} = (z_0)_{(0,0,0)} = (z_0)_{(0,0,0)} = 1, \quad (z_0)_{(2,2,2)} = 1/15,$$ $$(z_0)_{(4,2,0)} = (z_0)_{(2,4,0)} = (z_0)_{(0,4,2)} = (z_0)_{(2,0,2)} = (z_0)_{(2,0,2)} = 1/5,$$ $$(z_1)_{(0,0,0)} = (z_1)_{(0,0,0)} = (z_1)_{(0,0,0)} = 1.$$ 

We formulate this problem in the form (4.10) and solve it by Algorithm 4.1.

For $k = 4$, $y^{*,4} \in \mathbb{R}_A(K)$ (verified by Algorithm 4.1), and $\lambda^{*,4} = 2/3$. Since $c^k = b^k$ for $k = 4$, the maximum $\lambda$ in the above is $2/3$, which confirms the result of Reznick [41 p. 146].

(ii) We are interested to know what is the maximum $\lambda_1 + \lambda_2$ such that

$$(x_1^2 + x_2^2 + x_3^2)^3 - \lambda_1(x_1^3x_2^3 + x_2^3x_3^3 + x_3^3x_1^3) - \lambda_2 x_1^7x_2^5x_3^3 \in Q_{3,6}.$$ 

The problem is equivalent to

$$\max \lambda_1 + \lambda_2 \quad s.t. \quad z_0 - \lambda_1 z_1 - \lambda_2 z_2 \in \mathbb{R}_A(K)$$

where $z_0$ is same as in (i) and $z_1, z_2$ are tms’ whose entries are zeros except $$(z_1)_{(3,3,0)} = (z_1)_{(3,3,0)} = (z_1)_{(0,3,3)} = 1/20, \quad (z_2)_{(2,2,2)} = 1/90.$$ 

We formulate this problem in the form (4.10) and solve it by Algorithm 4.1. For $k = 3$, $y^{*,3} \in \mathbb{R}_A(K)$ (verified by Algorithm 4.1), and $\lambda^{*,3} = (2, 6)$. Since $c^k = b^k$ for $k = 3$, we know the optimal $\lambda$ in the above is $(2, 6)$. The SOEP decomposition of the polynomial $(x_1^2 + x_2^2 + x_3^2)^3 - 2(x_1^3x_2^3 + x_2^3x_3^3 + x_3^3x_1^3) - 6x_1^6x_2^5x_3^3$ is

$$7 \sum_{1 \leq i < j \leq 3} (x_i - x_j)^6 + \sum_{1 \leq i < j \leq 3} \left( \frac{1}{\sqrt{10}} - \frac{\sqrt{2}}{5} \right)^{1/3} x_i + \left( \frac{1}{\sqrt{10}} + \frac{\sqrt{2}}{5} \right)^{1/3} x_j.$$

5. Feasibility and Infeasibility

A basic question in linear optimization is to check whether a cone intersects an affine subspace. For the cones $\mathbb{R}_A(K)$ and $\mathbb{P}_A(K)$, this question is about checking whether the optimization problem (4.1) or (4.2) is feasible. If they are feasible, we are interested in getting a feasible point; if they are infeasible, we want certificates for the infeasibility.

5.1. Finding feasible points. First, we discuss how to check whether (4.1) is feasible. Suppose $a_1, \ldots, a_m \in \mathbb{R}[x]_A$ and $b \in \mathbb{R}^m$ are given as in (4.1), while the objective $c$ is not necessarily given. Generally, we can assume $\mathbb{R}[x]_A$ is $K$-full (otherwise, add vectors in $\mathbb{N}^n$ to $A$, to make it $K$-full). Choose $c \in \mathbb{R}[x]_A$ such that $c > 0$ on $K$. Consider the resulting optimization problems (4.1) and (4.2). For such $c$, the dual problem (4.2) has an interior point. We can apply Algorithm 4.1 to solve (4.1)–(4.2). If (4.1) is feasible, we can get a feasible point of (4.1).
Example 5.1. Let $\mathcal{A} = \mathbb{N}_0^3$ and $K = [-1, 1]^3$ be the unit cube, which is defined by $h = (0)$ and $g = (1 - x_1^2, 1 - x_2^2, 1 - x_3^2)$. We want to know whether there exists a measure $\mu$ supported on $[-1, 1]^3$ such that
\[
\int_{[-1, 1]^3} (x_1 x_2 + x_2 x_3 + x_3 x_1) d\mu = 0, \quad \int_{[-1, 1]^3} (x_1^2 x_2^2 + x_2^2 x_3^2 + x_3^2 x_1^2) d\mu = 1,
\]
\[
\int_{[-1, 1]^3} (x_1^3 x_2^2 + x_2^3 x_3^2 + x_3^3 x_1^2) d\mu = 1.
\]
Let $a_1, a_2, a_3$ be the polynomials inside the above integrals respectively. This problem is equivalent to whether there exists $y \in \mathcal{P}_\mathcal{A}([-1, 1]^3)$ satisfying
\[
\langle a_1, y \rangle = 0, \quad \langle a_2, y \rangle = 1, \quad \langle a_3, y \rangle = 1.
\]
Choose $c = \sum_{0 \leq |\alpha| \leq 3} x^{2\alpha}$. For $k = 3$, $y^{*,3}$ admits the measure $\frac{1}{2} \delta_{(0,1,-1)} + \frac{1}{5} \delta_{(1,1,1)}$, which satisfies the above. □

Second, we discuss how to check whether (4.2) is feasible. Suppose $c, a_1, \ldots, a_m \in \mathbb{R}[x, 1, \mathcal{A}]$ are given, while $b$ is not necessarily. Let $k = \lceil \deg(\mathcal{A})/2 \rceil$. Solve the semidefinite feasibility problem
\[
(5.1) \quad c - \lambda_1 a_1 - \cdots - \lambda_m a_m \in Q_k(g) + I_{2k}(h).
\]
If (5.1) is feasible, we can get a feasible point of (4.2); if not, let $k := k + 1$ and solve (5.1) again. Repeat this process. If the affine subspace $c + \text{span}\{a_1, \ldots, a_m\}$ intersects the interior of $\mathcal{P}_\mathcal{A}(K)$, we can always find a feasible point of (5.1) by solving (5.1). This is implied by Proposition 3.5 under the archimedeaness. If $c + \text{span}\{a_1, \ldots, a_m\}$ intersects a generic point of the boundary of $\mathcal{P}_\mathcal{A}(K)$, we can also get a feasible point of (4.2) by solving (5.1) (cf. [35]). In the remaining cases, it is still an open question to find a feasible point of (4.2) by using semidefinite relaxations, in the author’s best knowledge.

Example 5.2. We want to find $\lambda_1, \lambda_2$ such that $c - \lambda_1 a_1 - \lambda_2 a_2 \in P_{3,6}$, where $c, a_1, a_2$ are given as
\[
c = x_1^3 (x_1^2 + x_2^2 x_3^2 - x_2^2 (x_2^2 + x_3^2)), \quad a_1 = x_2^2 (x_1^2 + x_2^2 x_3^2 - x_3^2 (x_2^2 + x_1^2)),
\]
\[
a_2 = x_3^2 (x_1^2 + x_2^2 x_3^2 - x_2^2 (x_1^2 + x_2^2))
\]
For $k = 4$, (5.1) is feasible with $(\lambda_1, \lambda_2) = (-1, -1)$. □

5.2. Infeasibility certificates. First, we give a certificate for the infeasibility of (4.1). Suppose $a_1, \ldots, a_m \in \mathbb{R}[x, 1, \mathcal{A}]$ and $b \in \mathbb{R}^m$ are given, while $c$ is not necessarily.

Lemma 5.3. Let $K$ be as in (1.1). Then, we have:
\begin{enumerate}
\item[(i)] The problem (4.1) is infeasible if (4.3) is infeasible for some order $k$; (4.3) is infeasible if there exist $\lambda$ and $k$ such that
\[
b^T \lambda < 0, \quad \lambda_1 a_1 + \cdots + \lambda_m a_m \in Q_k(g) + I_{2k}(h).
\]
\item[(ii)] Suppose $I(h) + Q(g)$ is archimedean and there exists $a \in \text{span}\{a_1, \ldots, a_m\}$ such that $a > 0$ on $K$. If (4.1) is infeasible, then (5.3) holds for some $\lambda, k$ and (4.3) is infeasible.
\end{enumerate}
Example 5.4. Let $\mu$ be a measure with $(4.3)$ is infeasible. This is equivalent to checking whether there exists $y$ such that $(4.1)$ is contained in that of $(4.3)$. Clearly, if $(4.3)$ is infeasible, then $(4.1)$ is also infeasible.

If $(5.2)$ holds for some $\lambda$, then $(4.3)$ must be infeasible, because any feasible $y$ for $(4.3)$ results in the contradiction

$$0 > b^T \lambda = \sum_{i=1}^{m} \lambda_i(a_i, y) = \langle \sum_{i=1}^{m} \lambda_i a_i, y \rangle \geq 0.$$ 

(ii) Suppose $(4.1)$ is infeasible. Consider the optimization problem

$$(5.3) \quad \max_{y \in \mathbb{A}} 0 \quad s.t. \quad \langle a_i, y \rangle = b_i (i = 1, \ldots, m), \quad y \in \mathcal{R}_A(K).$$

Its dual problem is

$$(5.4) \quad \min_{\lambda \in \mathbb{R}^m} b^T \lambda \quad s.t. \quad \lambda a_1 + \cdots + \lambda_m a_m \in \mathcal{P}_A(K).$$

By the assumption, $\mathcal{R}_A(K)$ and $\mathcal{P}_A(K)$ are closed convex cones (cf. Proposition 5.2), and $(5.4)$ has an interior point. So, the strong duality holds and $(5.4)$ must be unbounded from below (cf. [1, Theorem 2.4.i]), i.e., there exists $\hat{\lambda}$ satisfying

$$b^T \hat{\lambda} < 0, \quad \hat{\lambda} a_1 + \cdots + \hat{\lambda}_m a_m > 0 \text{ on } K.$$ 

By the assumption, there exists $\hat{\lambda}$ such that $\lambda a_1 + \cdots + \lambda_m a_m > 0$ on $K$. For $\epsilon > 0$ small, $\lambda := \hat{\lambda} + \epsilon \hat{\lambda}$ satisfies $(5.2)$ for some $k$, by Theorem 2.1. By item (i), we know $(4.3)$ is infeasible. \(\square\)

Here is an example for the infeasibility certificate $(5.2)$.

Example 5.4. Let $\mathcal{A} = \mathbb{N}_0^2$ and $K = S^1$ be the unit circle in $\mathbb{R}^2$, defined by $h = (x_1^2 + x_2^2 - 1)$ and $g = (0)$ as in $(1.1)$. We want to know whether there exists a measure $\mu$ supported on $S^1$ such that

$$\int_{S^1} x_1^2 x_2^2 d\mu = 1, \quad \int_{S^1} (x_1^4 + x_2^4) d\mu = 1, \quad \int_{S^1} (x_1^6 + x_2^6) d\mu = 1.$$ 

This is equivalent to checking whether there exists $y \in \mathcal{R}_A(S^1)$ satisfying

$$\langle a_1, y \rangle = 1, \quad \langle a_2, y \rangle = 1, \quad \langle a_3, y \rangle = 1,$$

with $a_1 = x_1^2 x_2^2$, $a_2 = x_1^4 + x_2^4$, $a_3 = x_1^6 + x_2^6$. Indeed, such an $A$-tms $y$ does not exist, because $(5.2)$ is satisfied for $\lambda = (-3, 1, 1)$: $\lambda_1 + \lambda_2 + \lambda_3 < 0$ and

$$-3a_1 + a_2 + a_3 = 2(x_1^2 - x_2^2)^2 + (x_1^4 - x_1^2 x_2^2 + x_2^4) h \in I_6(h) + Q_3(g).$$

By Lemma 5.3 the above measure $\mu$ does not exist. \(\square\)

Second, we give a certificate for the infeasibility of $(4.2)$. Suppose $c, a_1, \ldots, a_m \in \mathbb{R}[x]_A$ are given, while $b$ is not necessarily.

Lemma 5.5. Let $K$ be compact and $c, a_1, \ldots, a_m \in \mathbb{R}[x]_A$ be given.

(i) Problem $(4.2)$ is infeasible if there exists $y$ satisfying

$$(5.5) \quad c^T y < 0, \quad \langle a_i, y \rangle = 0 (i = 1, \ldots, m), \quad y \in \mathcal{R}_A(K).$$

(ii) Suppose there does not exist $0 \neq a \in \text{span}\{a_1, \ldots, a_m\}$ such that $a \geq 0$ on $K$. If $(4.2)$ is infeasible, then there exists $y$ satisfying $(5.5)$.
Proof. (i) Suppose (5.5) holds. If (4.2) has a feasible \( \lambda \), then we get

\[
0 \leq \langle c(\lambda), y \rangle = \langle c, y \rangle - \lambda_1 \langle a_1, y \rangle - \cdots - \lambda_m \langle a_m, y \rangle = \langle c, y \rangle < 0,
\]
a contradiction. So (4.2) must be infeasible if (5.5) is satisfied.

(ii) Without loss of generality, we can assume \( a_1, \ldots, a_m \) are linearly independent in the quotient space \( \mathbb{R}[x]/I(K) \) (i.e., the space of polynomial functions defined on \( K \), cf. [3]). We show that there exists \( T > 0 \) such that

(5.6) \( \mathcal{P}_A(K) \cap (c + \text{span}\{a_1, \ldots, a_m\}) \subseteq B(0, T) \).

(The left above intersection might be empty.) Suppose otherwise such \( T \) does not exist, then there exists a sequence \( \{\lambda^k\} \) such that \( \|\lambda^k\|_2 \to \infty \) and \( c(\lambda^k) \in \mathcal{P}_A(K) \) for all \( k \). The sequence \( \{\lambda^k/\|\lambda^k\|_2\} \) is bounded. We can generally assume \( \lambda^k/\|\lambda^k\|_2 \to \lambda^* \neq 0 \). Clearly, \( c(\lambda^k)/\|\lambda^k\|_2 \in \mathcal{P}_A(K) \) for all \( k \). So,

\[
c(\lambda^k)/\|\lambda^k\|_2 \to a^*:=- (\lambda^*_1 a_1 + \cdots + \lambda^*_m a_m) \in \mathcal{P}_A(K).
\]

Since \( a_1, \ldots, a_m \) are linearly independent in \( \mathbb{R}[x]/I(K) \) and \( \lambda^* \neq 0 \), we know \( a^*_k \neq 0 \) and \( a^*_k |_K \geq 0 \). This contradicts the given assumption. So (5.6) must be satisfied for some \( T > 0 \). Let

\[
\mathcal{C}_1 = \{ p \in \mathcal{P}_A(K) \mid \|p\|_2 \leq T \}, \quad \mathcal{C}_2 = c + \text{span}\{a_1, \ldots, a_m\}.
\]

By (5.5), (4.2) is infeasible if and only if \( \mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset \). Because \( K \) is compact, the set \( \mathcal{C}_1 \) is compact convex, and \( \mathcal{C}_2 \) is closed convex. By the strict convex set separation theorem, they do not intersect if and only if there exists \( y \in \mathbb{R}^A \) and \( \tau \in \mathbb{R} \) such that

\[
\langle p, y \rangle > \tau \quad \forall p \in \mathcal{C}_1,
\]

\[
\langle p, y \rangle < \tau \quad \forall p \in \mathcal{C}_2.
\]

The first above inequality implies \( \tau < 0 \) and \( y \in \mathcal{P}_A(K) \), and the second one implies \( c^Ty < 0 \) and \( \langle a_i, y \rangle = 0 \) for all \( i \). Thus, this \( y \) satisfies (5.5). \( \square \)

Clearly, if there exists \( a \in \text{span}\{a_1, \ldots, a_m\} \) such that \( a > 0 \) on \( K \), then (4.2) must be feasible. Therefore, for (4.2) to be infeasible, none of polynomials in \( \text{span}\{a_1, \ldots, a_m\} \) can be positive on \( K \). So, the assumption in Lemma 5.5 (ii) is almost necessary for (4.2) to be infeasible. Indeed, it cannot be removed from the lemma, as shown in the below.

Example 5.6. Let \( K = S^1 \) and \( A = \{|a| = 2\} \). Choose \( c, a_1 \) such that \( c(\lambda) = x_1 x_2 - \lambda x_1^2 \). Clearly, \( c(\lambda) \notin \mathcal{P}_A(S^1) \) for all \( \lambda \). For all \( y \in \mathcal{P}_A(S^1) \), if \( \langle a, y \rangle = 0 \), then \( \langle c, y \rangle = 0 \). This is because of the Cauchy-Schwartz inequality: for all measure \( \mu \),

\[
\left| \int x_1 x_2 d\mu \right| \leq \left( \int x_1^2 d\mu \right)^{1/2} \left( \int x_2^2 d\mu \right)^{1/2}.
\]

So, there is no \( y \) satisfying (5.5). \( \square \)

The certificate (5.5) can be checked by solving the feasibility problem:

(5.7) \[ c^Ty = -1, \quad \langle a_i, y \rangle = 0 \ (i = 1, \ldots, m), \quad y \in \mathcal{P}_A(K). \]
Example 5.7. Let $K = S^2$ and $A = \{ \alpha \in \mathbb{N}^n : |\alpha| = 6 \}$. Then $\mathcal{P}_A(K)$ equals $P_{3,6}$, the cone of nonnegative ternary sextic forms. We want to know whether there exist $\lambda_1, \lambda_2, \lambda_3$ such that
\[
\begin{align*}
\sum_{c} -c \lambda_1 x_1^2 x_2^2 (x_1^2 + x_2^2 - 4x_3^2) + x_3^6 - \lambda_1 x_1^2 x_2^2 - \lambda_2 x_2^2 x_3^2 - \lambda_3 x_1^2 x_3^2 \in P_{3,6}.
\end{align*}
\]
Indeed, there are no $\lambda_1, \lambda_2, \lambda_3$ satisfying the above. To get a certificate for this, solve the feasibility problem (5.7). It has a feasible $y$ that admits the finitely atomic measure $2 \delta_{(1,1,1)/\sqrt{\pi}} + \delta_{(-1,1,1)/\sqrt{\pi}} + \delta_{(1,-1,1)/\sqrt{\pi}} + \delta_{(1,1,-1)/\sqrt{\pi}}$. 

References

[1] A. Ben-Tal and A. Nemirovski. Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.
[2] A. Berman and N. Shaked-Monderer. Completely Positive Matrices, World Scientific, 2003.
[3] J. B. Conway. A course in Functional Analysis. Springer-Verlag, 1990 2nd edition.
[4] D. Cox, J. Little and D. O’Shea. Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra. Third edition. Undergraduate Texts in Mathematics. Springer, New York, 1997.
[5] J. Curto and L. Fialkow. Solution of the truncated complex moment problem for flat data. Memoirs of the American Mathematical Society, 119(1996), No. 568, Amer. Math. Soc., Providence, RI, 1996.
[6] R. Curto and L. Fialkow. Flat extensions of positive moment matrices: recursively generated relations. Memoirs of the American Mathematical Society No. 648, AMS, Providence, 1998.
[7] R. Curto and L. Fialkow. Truncated K-moment problems in several variables. Journal of Operator Theory, 54(2005), pp. 189-226.
[8] M. Diehl. Copositive Programming - a Survey. In: M. Diehl, F. Glineur, E. Jarlebring, W. Michiels (Eds.), Recent Advances in Optimization and its Applications in Engineering, Springer 2010, pp. 3-20.
[9] L. Fialkow and J. Nie. The truncated moment problem via homogenization and flat extensions. J. Functional Analysis 263(2012), no. 6, pp. 1682–1700.
[10] J. Gouveia, P. Parrilo and R. Thomas. Theta bodies for polynomial ideals. SIAM Journal of Optimization 20 (2010), no. 4, 2097–2118.
[11] J. Gouveia, and T. Netzer. Positive Polynomials and Projections of Spectrahedra. SIAM Journal on Optimization 21 (3), 960-976, 2011.
[12] J. Gouveia, M. Laurent, P. Parrilo and R. Thomas. A new semidefinite programming relaxation for cycles in binary matroids and cuts in graphs. Mathematical Programming, 112(1-2), 203–225, 2012.
[13] J. Gouveia and R. Thomas. Convex Hulls of Algebraic Sets. Handbook of Semidefinite, Cone and Polynomial Optimization, 113-138, International Series in Operations Research & Management Science, Vol. 166, M. Anjos and J. Lasserre (eds), Springer, 2012.
[14] J.W. Helton and J. Nie. Semidefinite representation of convex sets. Mathematical Programming 122 (2010), no. 1, Ser. A, 21-64.
[15] J.W. Helton and J. Nie. Sufficient and necessary conditions for semidefinite representability of convex hulls and sets. SIAM J. Optim. 20 (2009), no. 2, 759-791.
[16] J.W. Helton and J. Nie. A semidefinite approach for truncated K-moment problems. Foundations of Computational Mathematics 12 (2012), no. 6, 851-881.
[17] D. Henrion and J. Lasserre. Detecting global optimality and extracting solutions in GloptiPoly. Positive polynomials in control, 293-310, Lecture Notes in Control and Inform. Sci., 312, Springer, Berlin, 2005.
[18] D. Henrion, J. Lasserre and J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. Optimization Methods and Software 24(2009), no. 4-5, pp. 761-779.
[19] J. B. Lasserre. Global optimization with polynomials and the problem of moments. SIAM J. Optim. 11(2001), no.3, pp. 786-817.
[20] J.B. Lasserre, M. Laurent, and F. Rostalski. Semidefinite characterization and computation of zero-dimensional real radical ideals. Found. Comput. Math. 8, pp. 607-647, 2008.
[21] J.B. Lasserre, M. Laurent, and P. Rostalski. A prolongation-projection algorithm for computing the finite real variety of an ideal. *Theoretical Computer Science* 410(27-29):2685-2700, 2009.

[22] J.B. Lasserre. A semidefinite programming approach to the generalized problem of moments. *Mathematical Programming* 112 (2008), pp. 65–92.

[23] J.B. Lasserre. Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* 19 (2008), no. 4, 1995-2014.

[24] J.B. Lasserre. Convex sets with semidefinite representation. *Mathematical Programming* 120 (2009), no. 2, Ser. A, 457-477.

[25] J.B. Lasserre. *Moments, Positive Polynomials and Their Applications*, Imperial College Press, 2009.

[26] M. Laurent. Revisiting two theorems of Curto and Fialkow on moment matrices. *Proceedings of the American Mathematical Society* 133(2005), no. 10, pp. 2965–2976.

[27] M. Laurent. Semidefinite representations for finite varieties. *Mathematical Programming*, 109 (2007), 1–26.

[28] M. Laurent. Sums of squares, moment matrices and optimization over polynomials, Emerging Applications of Algebraic Geometry, Vol. 149 of IMA Volumes in Mathematics and its Applications, M. Putinar and S. Sullivant (eds), Springer, pages 157-270, 2009.

[29] M. Laurent and P. Rostalski. The Approach of Moments for Polynomial Equations. *Handbook on Semidefinite, Cone and Polynomial Optimization* (eds. M. Anjos and J.B. Lasserre), Volume 166, pp. 25–60, International Series in Operations Research & Management Science, Springer, 2012.

[30] T. Netzer, D. Plaumann and M. Schweighofer. Exposed faces of semidefinite representable sets. *SIAM Journal on Optimization* 20(4): 1944-1955, 2010.

[31] J. Nie and M. Schweighofer. On the complexity of putinar’s positivstellensatz. *Journal of Complexity* 23(2007), pp. 135-150.

[32] J. Nie and K. Ranestad. Algebraic degree of polynomial optimization. *SIAM Journal on Optimization*, Vol. 20, No. 1, pp. 485–502, 2009.

[33] J. Nie. An exact Jacobian SDP relaxation for polynomial optimization. *Mathematical Programming*, 137 (2013), Ser. A, pp. 225-255.

[34] J. Nie. Certifying convergence of Lasserre’s hierarchy via flat truncation. *Mathematical Programming*, Ser. A, to appear.

[35] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. *Mathematical Programming*, Ser. A, to appear.

[36] J. Nie. The A-truncated K-moment problem. *Preprint*, 2012.

[37] P. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Ser. B*, Vol. 96, No.2, pp. 293-320, 2003.

[38] P. A. Parrilo and B. Sturmfels. Minimizing polynomial functions. In S. Basu and L. Gonzalez-Vega, editors, *Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science*, volume 60 of DIMACS Series in Discrete Mathematics and Computer Science, pages 83-99. AMS, 2003.

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

[40] B. Reznick. Some concrete polynomials of Hilbert’s 17th problem. In *Contemp. Math.*, Vol. 253, pp. 251-272. American Mathematical Society, 2000.

[41] B. Reznick. Sums of even powers of real linear forms. *Mem. Amer. Math. Soc.* Vol. 96, No. 463, 1992.

[42] C.A. Rogers. Probability measures on compact sets. *Proc. London Math. Soc.* (3) 52 (1986), no. 2, 328348.

[43] C. Scheiderer. Semidefinite representation for convex hulls of real algebraic curves. *Preprint*, 2012.

[44] V. Tchakaloff. Formules de cubatures mécanique à coefficients non négatifs. *Bull. Sci. Math.* (2) 81(1957), pp. 125-134.

Department of Mathematics, University of California, San Diego

E-mail address: njw@math.ucsd.edu