

FINITE CONVERGENCE OF MOMENT-SOS RELAXATIONS WITH NON-REAL RADICAL IDEALS

LEI HUANG, JIAWANG NIE, AND YA-XIANG YUAN

ABSTRACT. We consider the linear conic optimization problem with the cone of nonnegative polynomials. Its dual optimization problem is the generalized moment problem. Moment-SOS relaxations are powerful for solving them. This paper studies finite convergence of the Moment-SOS hierarchy when the constraining set is defined by equations whose ideal may not be real radical. Under the archimedeaness, we show that the Moment-SOS hierarchy has finite convergence if some classical optimality conditions hold at every minimizer of the optimal nonnegative polynomial for the linear conic optimization problem. When the archimedeaness fails (this is the case for unbounded sets), we propose a homogenized Moment-SOS hierarchy and prove its finite convergence under similar assumptions. Furthermore, we also prove the finite convergence of the Moment-SOS hierarchy with denominators. In particular, this paper resolves a conjecture posed in the earlier work.

1. INTRODUCTION

Let $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$ denote the ring of polynomials in $x := (x_1, \dots, x_n)$ with real coefficients, and let $\mathbb{R}[x]_d$ be the set of polynomials in $\mathbb{R}[x]$ with degrees at most d . For a nonnegative integer vector power $\alpha := (\alpha_1, \dots, \alpha_n)$, the α th moment of a Borel measure μ on \mathbb{R}^n is the integral $\int x^\alpha d\mu$ for the monomial $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of μ , for which we denote $\text{supp}(\mu)$, is the smallest closed set $T \subseteq \mathbb{R}^n$ such that $\mu(\mathbb{R}^n \setminus T) = 0$. For the degree $d > 0$, denote the power set $(|\alpha| := \alpha_1 + \cdots + \alpha_n)$

$$\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| \leq d\}.$$

Let $\mathbb{R}^{\mathbb{N}_d^n}$ denote the space of real vectors y that are labeled by $\alpha \in \mathbb{N}_d^n$, i.e.,

$$y = (y_\alpha)_{\alpha \in \mathbb{N}_d^n}.$$

Such a vector y is called a *truncated multi-sequence* (tms) of degree d . For a closed set $K \subseteq \mathbb{R}^n$, let $\mathcal{R}_d(K)$ denote the moment cone

$$\mathcal{R}_d(K) := \left\{ y \in \mathbb{R}^{\mathbb{N}_d^n} \mid \begin{array}{l} \exists \text{ a Borel measure } \mu, \text{ supp}(\mu) \subseteq K, \\ y_\alpha = \int x^\alpha d\mu \text{ for each } \alpha \in \mathbb{N}_d^n \end{array} \right\}.$$

The moment cone $\mathcal{R}_d(K)$ is always convex. It is closed if K is compact, but it is typically not closed if K is noncompact. We define the bilinear operation $\langle \cdot, \cdot \rangle$ between $\mathbb{R}[x]_d$ and $\mathbb{R}^{\mathbb{N}_d^n}$:

$$\langle p, y \rangle = \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha y_\alpha \quad \text{for} \quad p = \sum_{\alpha \in \mathbb{N}_d^n} p_\alpha x^\alpha.$$

2020 *Mathematics Subject Classification.* 90C23,65K05,90C22.

Key words and phrases. Moment-SOS relaxations, finite convergence, nonnegative polynomial, real radical, optimality condition.

We study the generalized moment problem (GMP), which is also called the generalized problem of moments in some literature (see [23])

$$(1.1) \quad \begin{cases} \min & \langle f, y \rangle \\ \text{s.t.} & \langle a_i, y \rangle = b_i \quad (1 \leq i \leq m_1), \\ & \langle a_i, y \rangle \geq b_i \quad (m_1 < i \leq m), \\ & y \in \mathcal{R}_d(K). \end{cases}$$

In the above, f, a_1, \dots, a_m are given polynomials in $\mathbb{R}[x]_d$, $m_1 \leq m$ are integers and $b_1, \dots, b_m \in \mathbb{R}$ are given constants. For convenience, we denote the vector $b = (b_1, \dots, b_m)$. The GMP has broad applications in optimal transport [33], distributionally robust optimization [43, 46], computational geometry and optimal control [15], different ranks for matrices and tensors [10, 12, 20], and tensor optimization [44, 40, 45]. More applications can be found in [15, 23, 28, 41].

The d th degree cone of polynomials that are nonnegative on K is

$$\mathcal{P}_d(K) := \{p \in \mathbb{R}[x]_d : p(x) \geq 0 \text{ for all } x \in K\}.$$

Note that $\mathcal{P}_d(K)$ is a closed convex cone of $\mathbb{R}[x]_d$. The nonnegative polynomial cone $\mathcal{P}_d(K)$ and the moment cone $\mathcal{R}_d(K)$ are dual to each other, under the above bilinear operation $\langle \cdot, \cdot \rangle$, i.e.,

$$\langle p, y \rangle \geq 0 \text{ for all } p \in \mathcal{P}_d(K) \text{ and for all } y \in \mathcal{R}_d(K).$$

It can be shown that the dual optimization of (1.1) is the following linear conic optimization problem:

$$(1.2) \quad \begin{cases} \max & b_1\theta_1 + \dots + b_m\theta_m \\ \text{s.t.} & f - \sum_{i=1}^m \theta_i a_i \in \mathcal{P}_d(K), \\ & \theta_i \geq 0 \quad (m_1 < i \leq m), \\ & \theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m. \end{cases}$$

We refer to [23, 41] for more details about the dual relationship.

Moment-SOS relaxations are introduced by Lasserre [23] for solving the pair (1.1)-(1.2). Under the archimedeaness, the Moment-SOS hierarchy has asymptotic convergence. When the set K is a simplex or sphere, the convergence rates of the Moment-SOS hierarchy are studied in [19]. We also refer to the survey by De Klerk and Laurent [9] for the convergence rate analysis.

The classical polynomial optimization problem (see [21]) can be viewed as a special case of the GMP. There are interesting results about the finite convergence of the Moment-SOS hierarchy for solving polynomial optimization. Under the archimedeaness, Nie [39] proved that the Moment-SOS hierarchy has finite convergence if some classical optimality conditions hold at every minimizer (throughout this paper, a minimizer means a global minimizer, unless its meaning is otherwise specified). For convex polynomial optimization, the Moment-SOS hierarchy has finite convergence under the strict convexity or SOS-convexity conditions [8, 24]. When the equality constraints give a finite set, the Moment-SOS hierarchy also has finite convergence, as shown in [22, 27, 38]. For asymptotic convergence rates of the Moment-SOS hierarchy for solving polynomial optimization, we refer to [1, 11, 30, 42, 50].

Contributions. Despite rich work for polynomial optimization, there are relatively less results about finite convergence of the Moment-SOS hierarchy for solving generalized moment problems. A partial reason for this is that the classical optimality conditions for nonlinear programming [3] do not generalize conveniently for (1.1)-(1.2). Consequently, the finite convergence theory in [39] does not generalize directly.

Assume that K is the semialgebraic set given as

$$K := \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} c_j(x) = 0 \ (j \in \mathcal{E}), \\ c_j(x) \geq 0 \ (j \in \mathcal{I}) \end{array} \right\},$$

where c_j ($j \in \mathcal{E} \cup \mathcal{I}$) are given polynomials. The \mathcal{E} and \mathcal{I} are disjoint labeling sets for equality and inequality constraining polynomials respectively. Suppose $\theta^* = (\theta_1^*, \dots, \theta_m^*)$ is a maximizer of (1.2). Consider the polynomial optimization problem

$$(1.3) \quad \begin{cases} \min & f_{\theta^*}(x) := f(x) - \sum_{i=1}^m \theta_i^* a_i(x) \\ \text{s.t.} & c_j(x) = 0 \ (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{cases}$$

It is worthy to note that the polynomial $f_{\theta^*}(x)$ is nonnegative on K and it also typically has a zero $u \in K$. For convenience, we denote the polynomial tuples

$$c_{eq} = (c_j(x))_{j \in \mathcal{E}}, \quad c_{in} = (c_j(x))_{j \in \mathcal{I}}.$$

When $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is archimedean (see Section 2.2 for the notation), if the linear independence constraint qualification, strict complementarity and second order sufficient conditions hold at each minimizer of (1.3), then there exists a polynomial $\sigma \in \text{QM}[c_{in}]$ such that $f_{\theta^*} - \sigma \equiv 0$ on $V_{\mathbb{R}}(c_{eq})$. If $\text{Ideal}[c_{eq}]$ is a real radical ideal (see Section 2.2 for the definition), then $f_{\theta^*}(x) - \sigma \in \text{Ideal}[c_{eq}]$, or equivalently, $f_{\theta^*}(x) \in \text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$. However, if $\text{Ideal}[c_{eq}]$ is not real radical, we typically do not have $f_{\theta^*}(x) \in \text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$. This is a major difficulty for studying finite convergence of the Moment-SOS hierarchy when solving generalized moment problems.

We remark that the equality constraint $c_j(x) = 0$ in (1.3) can be equivalently replaced by two inequalities $c_j(x) \geq 0$, $-c_j(x) \geq 0$. However, by doing this, the linear independence constraint qualification will fail for this reformulation and forbid us to get the finite convergence results.

In this paper, we study the Moment-SOS hierarchy for solving (1.1)-(1.2) *without* the assumption that $\text{Ideal}[c_{eq}]$ is real radical. Indeed, we can prove its finite convergence under the archimedeaness and the above mentioned optimality conditions. Our major contributions are as follows.

- First, we consider the case that $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is archimedean but $\text{Ideal}[c_{eq}]$ is not necessarily real radical. We prove that the Moment-SOS hierarchy for solving (1.1)-(1.2) has finite convergence if the linear independence constraint qualification, strict complementarity and second order sufficient conditions hold at every minimizer of (1.3). In particular, we are also able to show that every minimizer of the moment relaxation for (1.1) must satisfy the flat truncation, for all sufficient high relaxation orders. This also gives a certificate for checking finite convergence of the

Moment-SOS hierarchy in computational practice. Applications in polynomial optimization and the semialgebraic super resolution problem are also discussed.

- Second, we consider the case that $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is not archimedean. This is typically the case when K is an unbounded set and the Moment-SOS hierarchy generally does not converge. For this case, we propose a homogenized Moment-SOS hierarchy introduced in [17, 18] and prove its finite convergence under similar optimality conditions for the homogenization of the optimization (1.3). Consequently, we are also able to show the finite convergence of the Moment-SOS hierarchy with denominators, motivated by the Putinar-Vasilescu type Positivstellensatz [29, 48]. In particular, our results resolve Conjecture 8.2 in the earlier work [17].

This paper is organized as follows. Section 2 reviews some backgrounds for polynomial optimization and moment problems. In Section 3, we prove the finite convergence of the Moment-SOS hierarchy under the archimedeaness and optimality conditions, without the assumption of real radicalness. In Section 4, we prove the finite convergence of the homogenized Moment-SOS hierarchy and the one with denominators. In Section 5, we study the semialgebraic super resolution problem. We make some conclusions and pose a new conjecture in Section 6.

2. PRELIMINARIES

Notation. The symbol \mathbb{N} (resp., \mathbb{R}) denotes the set of nonnegative integers (resp., real numbers). For $x = (x_1, \dots, x_n)$ and $\alpha = (\alpha_1, \dots, \alpha_n)$, denote

$$x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}, \quad |\alpha| := \alpha_1 + \cdots + \alpha_n.$$

For $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer greater than or equal to t . For a positive integer m , denote $[m] := \{1, \dots, m\}$. For a matrix A , A^T denotes its transpose. A symmetric matrix $X \succeq 0$ if X is positive semidefinite. For a set $S \subseteq \mathbb{R}^n$, $cl(S)$ and $int(S)$ denote its closure and interior in the Euclidean topology. Denote by $\deg(p)$ the total degree of the polynomial p . For a degree d , let $[x]_d$ denote the vector of all monomials in x with degrees $\leq d$, ordered in the graded alphabetical ordering, i.e.,

$$[x]_d^T = [1, x_1, x_2, \dots, x_1^2, x_1x_2, \dots, x_1^d, x_1^{d-1}x_2, \dots, x_n^d].$$

For a polynomial p , p^{hom} denotes its homogeneous terms of the highest degree and \tilde{p} denotes its homogenization, i.e., $\tilde{p}(\tilde{x}) = x_0^{\deg(p)} p(x/x_0)$ for $\tilde{x} := (x_0, x_1, \dots, x_n)$. A homogeneous polynomial p is said to be a form and p is positive definite if $p(x) > 0$ for all nonzero $x \in \mathbb{R}^n$.

2.1. Optimality conditions. We first review some classical optimality conditions in nonlinear programming [3]. Consider the nonlinear optimization problem

$$(2.1) \quad \begin{cases} \min & f(x) \\ \text{s.t.} & c_j(x) = 0 \quad (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \quad (j \in \mathcal{I}), \end{cases}$$

where f, c_j ($j \in \mathcal{E} \cup \mathcal{I}$) are second order continuous differentiable. Suppose x^* is a local minimizer of (2.1). Denote the label set of active constraints at x^*

$$(2.2) \quad J(x^*) := \{j \in \mathcal{E} \cup \mathcal{I} : c_j(x^*) = 0\}.$$

If the gradient vector set $\{\nabla c_j(x^*)\}_{j \in J(x^*)}$ is linearly independent, the linear independence constraint qualification condition (LICQC) is said to hold at x^* . When the LICQC holds at x^* , there exist Lagrange multipliers λ_j ($j \in \mathcal{E} \cup \mathcal{I}$) such that

$$(2.3) \quad \boxed{\begin{aligned} \nabla f(x^*) &= \sum_{j \in \mathcal{E} \cup \mathcal{I}} \lambda_j \nabla c_j(x^*), \\ \lambda_j &\geq 0, \lambda_j c_j(x^*) = 0 \quad (j \in \mathcal{I}) \end{aligned}}$$

Moreover, if $\lambda_j + c_j(x^*) > 0$ for every $j \in \mathcal{I}$, then the strict complementarity condition (SCC) is said to hold at x^* . For the above Lagrange multipliers, the Lagrange function is

$$\mathcal{L}(x) := f(x) - \sum_{j \in \mathcal{E} \cup \mathcal{I}} \lambda_j c_j(x).$$

Under the LICQC, the second order necessary condition (SONC) must hold at x^* , i.e., the following holds (∇^2 denotes the Hessian operator)

$$(2.4) \quad v^T \nabla^2(\mathcal{L}(x^*))v \geq 0$$

for every $v \in \mathbb{R}^n$ satisfying

$$(2.5) \quad v^T \nabla c_j(x^*) = 0 \quad \text{for all } j \in J(x^*).$$

The second order sufficient condition (SOSC) is said to hold at x^* if

$$(2.6) \quad v^T \nabla^2(\mathcal{L}(x^*))v > 0,$$

for every nonzero v satisfying (2.5).

2.2. Some basics for polynomial optimization. Here we review some basics for polynomial optimization and moment problems. A subset $I \subseteq \mathbb{R}[x]$ is called an ideal of $\mathbb{R}[x]$ if $I \cdot \mathbb{R}[x] \subseteq \mathbb{R}[x]$, $I + I \subseteq I$. For a polynomial tuple $h := (h_1, \dots, h_t)$, $\text{Ideal}[h]$ denotes the ideal generated by h , i.e.,

$$\text{Ideal}[h] = h_1 \cdot \mathbb{R}[x] + \dots + h_t \cdot \mathbb{R}[x].$$

For a degree k , the k th degree truncation of $\text{Ideal}[h]$ is

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

Its real variety is

$$V_{\mathbb{R}}(h) := \{x \in \mathbb{R}^n \mid h_1(x) = \dots = h_t(x) = 0\}.$$

For a set $T \subseteq \mathbb{R}^n$, its vanishing ideal is

$$\text{Ideal}[T] = \{p \in \mathbb{R}[x] : p(u) = 0 \text{ for all } u \in T\}.$$

Clearly, if $T = V_{\mathbb{R}}(h)$, then $\text{Ideal}[h] \subseteq \text{Ideal}[T]$. However, the reverse inclusion may not hold. The ideal $\text{Ideal}[h]$ is said to be *real radical* if $\text{Ideal}[h] = \text{Ideal}[T]$ for $T = V_{\mathbb{R}}(h)$.

A polynomial p is said to be a sum of squares (SOS) if $p = p_1^2 + \dots + p_s^2$ for $p_1, \dots, p_s \in \mathbb{R}[x]$. The set of all SOS polynomials in x is denoted as $\Sigma[x]$. For a degree k , denote the truncation

$$\Sigma[x]_k := \Sigma[x] \cap \mathbb{R}[x]_k.$$

The notion of SOS polynomials plays a core role in optimization. Interesting, this notion can also be defined for matrix polynomials (see [16, 34]). Clearly, an SOS polynomial is nonnegative everywhere, while the reverse is typically not true. The

quality of SOS approximations is discussed in [11, 36]. For a polynomial tuple $g = (g_1, \dots, g_\ell)$, the quadratic module generated by g is

$$(2.7) \quad \text{QM}[g] := \Sigma[x] + g_1 \cdot \Sigma[x] + \dots + g_\ell \cdot \Sigma[x].$$

Similarly, the k th degree truncation of $\text{QM}[g]$ is

$$(2.8) \quad \text{QM}[g]_k = \Sigma[x]_k + g_1 \cdot \Sigma[x]_{k-\deg(g_1)} + \dots + g_\ell \cdot \Sigma[x]_{k-\deg(g_\ell)}.$$

The sum $\text{Ideal}[h] + \text{QM}[g]$ is said to be archimedean if there exists $R > 0$ such that $R - \|x\|^2 \in \text{Ideal}[h] + \text{QM}[g]$. If it is archimedean, then the set

$$S := \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \geq 0\}$$

must be compact. Clearly, if $p \in \text{Ideal}[h] + \text{QM}[g]$, then $p \geq 0$ on S while the converse is not always true. However, if p is positive on S and $\text{Ideal}[h] + \text{QM}[g]$ is archimedean, we have $p \in \text{Ideal}[h] + \text{QM}[g]$. This conclusion is often referred as Putinar's Positivstellensatz.

Theorem 2.1 ([47]). *Suppose $\text{Ideal}[h] + \text{QM}[g]$ is archimedean. If a polynomial $p > 0$ on S , then $p \in \text{Ideal}[h] + \text{QM}[g]$.*

In the above theorem, if $p \geq 0$ on S , $\text{Ideal}[h]$ is real radical and the LICQC, SCC and SOSC hold at each zero of p on S , then we also have $p \in \text{Ideal}[h] + \text{QM}[g]$. This is shown in [39].

For a polynomial $q \in \mathbb{R}[x]_{2k}$ and a tms $w \in \mathbb{R}^{\mathbb{N}_{2k}^n}$, the k th order *localizing matrix* of q for w is the symmetric matrix $L_q^{(k)}[w]$ such that

$$(2.9) \quad \text{vec}(p)^T \left(L_q^{(k)}[w] \right) \text{vec}(p) = \langle qp^2, w \rangle,$$

for every polynomial $p \in \mathbb{R}[x]_s$ with $2s \leq 2k - \deg(q)$. In the above, $\text{vec}(p)$ denotes the coefficient vector of p . In particular, for the constant one polynomial $q = 1$, the localizing matrix $L_q^{(k)}[w]$ becomes the k th order *moment matrix* $M_k[w] := L_1^{(k)}[w]$. For instance, when $n = 2$, we have

$$M_2[w] = \begin{bmatrix} w_{00} & w_{10} & w_{01} & w_{20} & w_{11} & w_{02} \\ w_{10} & w_{20} & w_{11} & w_{30} & w_{21} & w_{12} \\ w_{01} & w_{11} & w_{02} & w_{21} & w_{12} & w_{03} \\ w_{20} & w_{30} & w_{21} & w_{40} & w_{31} & w_{22} \\ w_{11} & w_{21} & w_{12} & w_{31} & w_{22} & w_{13} \\ w_{02} & w_{12} & w_{03} & w_{22} & w_{13} & w_{04} \end{bmatrix},$$

$$L_{x_1-x_2^2}^{(2)}[w] = \begin{bmatrix} w_{10} - w_{02} & w_{20} - w_{12} & w_{11} - w_{03} \\ w_{20} - w_{12} & w_{30} - w_{22} & w_{21} - w_{13} \\ w_{11} - w_{03} & w_{21} - w_{13} & w_{12} - w_{04} \end{bmatrix}.$$

For the polynomial q , let $\mathcal{V}_q^{(2k)}[z]$ denote the vector such that

$$(2.10) \quad \langle qp, w \rangle = \left(\mathcal{V}_q^{(2k)}[w] \right)^T \text{vec}(p),$$

for every polynomial $p \in \mathbb{R}[x]_s$ with $s \leq 2k - \deg(q)$. The $\mathcal{V}_q^{(2k)}[w]$ is called the *localizing vector* of q and w . We refer to [41, Chap. 2] for more details about quadratic modules and localizing matrices/vectors. They are basic tools for solving polynomial optimization and moment problems. We refer to the books and surveys [15, 25, 28, 32, 49] for more comprehensive introductions.

3. FINITE CONVERGENCE OF THE MOMENT-SOS HIERARCHY

In this section, we study the Moment-SOS hierarchy for solving the optimization problems (1.1)-(1.2). Assume K is the basic closed semialgebraic set

$$(3.1) \quad K := \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} c_j(x) = 0 \ (j \in \mathcal{E}), \\ c_j(x) \geq 0 \ (j \in \mathcal{I}) \end{array} \right\},$$

where c_j ($j \in \mathcal{E} \cup \mathcal{I}$) are given polynomials in x . The \mathcal{E} and \mathcal{I} are disjoint labeling sets. For convenience, we denote the polynomial tuples

$$c_{eq} = (c_j(x))_{j \in \mathcal{E}}, \quad c_{in} = (c_j(x))_{j \in \mathcal{I}}.$$

This section studies the finite convergence theory for the Moment-SOS hierarchy, without the assumption that $\text{Ideal}[c_{eq}]$ is real radical.

3.1. The Moment-SOS hierarchy. Denote the degrees

$$d_K = \max_{j \in \mathcal{E} \cup \mathcal{I}} \{ \lceil \deg(c_j)/2 \rceil \}, \quad d_0 = \max \left\{ \lceil \deg(f)/2 \rceil, d_K, \max_{1 \leq i \leq m} \lceil \deg(a_i)/2 \rceil \right\}.$$

For a degree $k \geq d_0$, the k th moment relaxation for (1.1) is

$$(3.2) \quad \left\{ \begin{array}{l} \min \quad \langle f, w \rangle \\ s.t. \quad \langle a_i, w \rangle = b_i \ (1 \leq i \leq m_1), \\ \quad \quad \langle a_i, w \rangle \geq b_i \ (m_1 < i \leq m), \\ \quad \quad \mathcal{V}_{c_j}^{(2k)}[w] = 0 \ (j \in \mathcal{E}), \\ \quad \quad L_{c_j}^{(k)}[w] \geq 0 \ (j \in \mathcal{I}), \\ \quad \quad M_k[w] \succeq 0, \ w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{array} \right.$$

The dual optimization of (3.2) is the k th order SOS relaxation for (1.2):

$$(3.3) \quad \left\{ \begin{array}{l} \max \quad b_1 \theta_1 + \cdots + b_m \theta_m \\ s.t. \quad f - \sum_{i=1}^m \theta_i a_i \in \text{Ideal}[c_{eq}]_{2k} + \text{QM}[c_{in}]_{2k}, \\ \quad \quad \theta \in \mathbb{R}^m, \ \theta_{m_1+1} \geq 0, \dots, \theta_m \geq 0. \end{array} \right.$$

As k goes to infinity, the sequence of (3.2)-(3.3) is called the Moment-SOS hierarchy for solving (1.1)-(1.2).

Let ϕ^* , ϑ^* denote the optimal value of (1.1), (1.2) respectively, and let ϕ_k , ϑ_k denote the optimal value of (3.2), (3.3) respectively. By the weak duality, it holds that

$$\vartheta_k \leq \vartheta^*, \quad \phi_k \leq \phi^*, \quad \vartheta_k \leq \phi_k, \quad \vartheta^* \leq \phi^*.$$

The hierarchy of (3.2)-(3.3) is said to have finite convergence (or be *tight*) if there exists $k_0 \in \mathbb{N}$ such that $\vartheta_k = \phi_k = \vartheta^* = \phi^*$ for all $k \geq k_0$.

In computational practice, a convenient criterion to check finite convergence is the flat truncation [5, 37]. Suppose $w^* \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ is a minimizer of (3.2). The tms w^* is said to have a flat truncation if there exists a degree $t \in [d_0, k]$ such that

$$(3.4) \quad \text{rank } M_{t-d_K}[w^*|_{2t}] = \text{rank } M_t[w^*|_{2t}].$$

When it holds, we must have $\phi_k = \phi^*$ and $w^*|_d$ is the optimal solution to (1.1). We refer to [15, 26, 28, 41] for more details.

3.2. The finite convergence theory. We now prove the finite convergence of the Moment-SOS hierarchy (3.2)–(3.3) under the archimedeaness and optimality conditions. This extends the finite convergence results in [39] for polynomial optimization to the generalized moment problem.

For convenience of notation, we denote the polynomial

$$f_{\theta}(x) := f(x) - \sum_{i=1}^m \theta_i a_i(x).$$

Suppose $\theta^* \in \mathbb{R}^m$ is a maximizer of (1.2). Consider the polynomial optimization problem

$$(3.5) \quad \begin{cases} \min & f_{\theta^*}(x) \\ \text{s.t.} & c_j(x) = 0 \ (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{cases}$$

Generally, the optimal value of (3.5) is zero and each of its minimizers is a zero of $f_{\theta^*}(x)$ in the set K . This is guaranteed under the following assumption.

Assumption 3.1. Suppose $\theta^* \in \mathbb{R}^m$ is a maximizer of (1.2), y^* is a minimizer of (1.1) with $y^* \neq 0$, and the Slater condition holds for (1.2), i.e., there exists a feasible point $\bar{\theta} \in \mathbb{R}^m$ of (1.2) such that $f_{\bar{\theta}} \in \text{int}(\mathcal{P}_d(K))$.

Under Assumption 3.1, the strong duality holds between (1.2) and (1.1), so

$$0 = \langle f, y^* \rangle - b^T \theta^* = \langle f_{\theta^*}, y^* \rangle = \int f_{\theta^*}(x) d\mu,$$

for every representing measure μ of y^* , with $\text{supp}(\mu) \subseteq K$. Note that $f_{\theta^*} \geq 0$ on K . If $y^* \neq 0$, then $\text{supp}(\mu) \neq \emptyset$ and each point in $\text{supp}(\mu)$ must be a zero of f_{θ^*} in the set K . The following is our major conclusion in this paper.

Theorem 3.2. *Suppose $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is archimedean and Assumption 3.1 holds. If the LICQC, SCC and SOSC hold at every minimizer of (3.5), then we have:*

- (i) *The hierarchy of (3.2)–(3.3) has finite convergence, i.e., $\vartheta_k = \phi_k = \vartheta^* = \phi^*$ for all k big enough.*
- (ii) *Every minimizer of the moment relaxation (3.2) must have a flat truncation, when k is sufficiently large.*

Proof. (i) Note that the minimum value of (3.5) is nonnegative. Since the LICQC, SCC and SOSC hold at every minimizer of (3.5), by Theorem 1.1 of [39], there exists $\sigma_0 \in \text{QM}[c_{in}]$ such that

$$f_{\theta^*} \equiv \sigma_0 \pmod{\text{Ideal}[V_{\mathbb{R}}(c_{eq})]}.$$

This means that $\hat{f} := f_{\theta^*} - \sigma_0$ vanishes identically on the real variety $V_{\mathbb{R}}(c_{eq})$. By the Real Nullstellensatz (see [41]), there exist $k_1 \in \mathbb{N}$, $\sigma_1 \in \Sigma[x]$ such that $\hat{f}^{2k_1} + \sigma_1 \in \text{Ideal}[c_{eq}]$. The archimedeaness of $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ implies K is compact. By the Slater condition, there exist a feasible point $\bar{\theta} \in \mathbb{R}^m$ for (1.2) and a scalar $\delta > 0$ such that $f_{\bar{\theta}} \geq \delta > 0$ on K . By Theorem 2.1, we have that

$$(3.6) \quad f_{\bar{\theta}} - \delta \in \text{Ideal}[c_{eq}] + \text{QM}[c_{in}].$$

For $\gamma \geq \frac{1}{2k_1}(1 - \frac{1}{2k_1})^{2k_1-1}$, the polynomial $\omega(t) := 1 + t + \gamma t^{2k_1}$ is SOS (cf. [38]), so

$$(3.7) \quad \delta \cdot \omega(t/\delta) = \delta + t + \frac{1}{2k_1}(1 - \frac{1}{2k_1})^{2k_1-1} \delta^{1-2k_1} t^{2k_1} \in \Sigma[t]_{2k_1}.$$

Combining with (3.6) and (3.7), we have

$$(3.8) \quad s(t, x) := f_{\bar{\theta}} - \delta + (\delta + t + \eta t^{2k_1}) \in \text{Ideal}[c_{eq}] + \text{QM}[c_{in}] + \Sigma[t],$$

for every scalar

$$\eta \geq \frac{1}{2k_1} \left(1 - \frac{1}{2k_1}\right)^{2k_1-1} \delta^{1-2k_1}.$$

For $0 < \lambda \leq 1$, replacing t by $(\lambda^{-1} - 1)\hat{f}$ in $s(t, x)$, we then get

$$\lambda f_{\bar{\theta}} + (1 - \lambda)\hat{f} = \lambda s((\lambda^{-1} - 1)\hat{f}, x) - \lambda^{1-2k_1}(1 - \lambda)^{2k_1} \eta \hat{f}^{2k_1}.$$

Note that

$$\begin{aligned} \lambda f_{\bar{\theta}} + (1 - \lambda)\hat{f} &= \lambda f_{\bar{\theta}} + (1 - \lambda)f_{\theta^*} - (1 - \lambda)\sigma_0 \\ &= f_{(1-\lambda)\theta^* + \lambda\bar{\theta}} - (1 - \lambda)\sigma_0. \end{aligned}$$

Thus, we get

$$f_{(1-\lambda)\theta^* + \lambda\bar{\theta}} = (1 - \lambda)\sigma_0 + \lambda s((\lambda^{-1} - 1)\hat{f}, x) - \lambda^{1-2k_1}(1 - \lambda)^{2k_1} \eta \hat{f}^{2k_1}.$$

We select η sufficiently large so that (3.8) holds. Then there exists $k_0 \in \mathbb{N}$ such that

$$(3.9) \quad f_{(1-\lambda)\theta^* + \lambda\bar{\theta}} \in \text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0}.$$

Since (3.9) holds for all $0 < \lambda \leq 1$ and k_0 is independent of λ , we know $(1 - \lambda)\theta^* + \lambda\bar{\theta}$ is feasible for (3.3) at the relaxation order k_0 , for all $0 < \lambda \leq 1$. As $\lambda \rightarrow 0$, we get

$$(3.10) \quad b^T((1 - \lambda)\theta^* + \lambda\bar{\theta}) = (1 - \lambda)b^T\theta^* + \lambda b^T\bar{\theta} \rightarrow b^T\theta^*.$$

This implies that $\vartheta_{k_0} = b^T\theta^* = \vartheta^*$. Since the Slater condition holds for (1.2), we know $\vartheta^* = \phi^* = \vartheta_{k_0} = \phi_{k_0}$, by the strong duality. Hence, the hierarchy (3.2)–(3.3) has finite convergence.

(ii) Suppose $w^{(k)}$ is a minimizer of (3.2) at the relaxation order k . First, we prove that $w_0^{(k)} \neq 0$. Suppose otherwise $w_0^{(k)} = 0$. The constraint $M_k[w^{(k)}] \succeq 0$ implies that $w_\alpha^{(k)} = 0$ for all $|\alpha| \leq k$. By Lemma 5.7 [28], we have that $M_k[w^{(k)}] \text{vec}(x^\alpha) = 0$ for all $|\alpha| \leq k - 1$. For all $|\alpha| \leq 2k - 2$, we can write $\alpha = \beta + \eta$ with $|\beta|, |\eta| \leq k - 1$ and the following holds

$$w_\alpha^{(k)} = \text{vec}(x^\beta)^T M_k[w^{(k)}] \text{vec}(x^\eta) = 0.$$

Clearly, the truncation $w^{(k)}|_{2k-2}$ is flat. It implies that the minimizer of (1.1) is $y^* = 0$, which is a contradiction.

Thus, we know that $w_0^{(k)} > 0$. Up to scaling, we can assume $w_0^{(k)} = 1$. Since the Slater condition holds for (1.2), we have $\vartheta^* = \phi^*$ and (1.1) has a minimizer, saying, y^* . Suppose μ is a K -representing measure of y^* . Since $y^* \neq 0$, we know that $\text{supp}(\mu) \neq \emptyset$. It holds that

$$\begin{aligned} \int f_{\theta^*} d\mu &= \langle f - \sum_{i=1}^m \theta_i^* a_i, y^* \rangle = \langle f, y^* \rangle - \sum_{i=1}^{m_1} \theta_i^* b_i - \sum_{i=m_1+1}^m \theta_i^* \langle a_i, y^* \rangle \\ &\leq \langle f, y^* \rangle - \sum_{i=1}^m \theta_i^* b_i \\ &= 0. \end{aligned}$$

Note that $f_{\theta^*}(x) \geq 0$ on K and $\text{supp}(\mu) \neq \emptyset$. Hence, the optimal value of (3.5) is 0. The moment relaxation for polynomial optimization (3.5) at the order k is

$$(3.11) \quad \begin{cases} \min & \langle f - \sum_{i=1}^m \theta_i^* a_i, w \rangle \\ \text{s.t.} & \langle 1, w \rangle = 1, \\ & \mathcal{Y}_{c_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \\ & L_{c_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \\ & M_k[w] \succeq 0, \quad w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases}$$

By item (i), we know that $\vartheta^* = \phi^* = \vartheta_k = \phi_k$ for $k \geq k_0$ and the following holds

$$\begin{aligned} \langle f_{\theta^*}, w^{(k)} \rangle &= \langle f, w^{(k)} \rangle - \sum_{i=1}^{m_1} \theta_i^* b_i - \sum_{i=m_1+1}^m \theta_i^* \langle a_i, w^{(k)} \rangle \\ &\leq \langle f, w^{(k)} \rangle - \sum_{i=1}^m \theta_i^* b_i = 0. \end{aligned}$$

Since $f_{\theta^*} \in \text{cl}(\text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0})$, we have $\langle f_{\theta^*}, w^{(k)} \rangle \geq 0$. Thus,

$$\langle f_{\theta^*}, w^{(k)} \rangle = 0$$

and $w^{(k)}$ is a minimizer of (3.11) for all $k \geq k_0$. Let

$$Q := \text{Ideal}[c_{eq}, f_{\theta^*}] + \text{QM}[c_{in}].$$

Note that Q is archimedean and the intersection $J := Q \cap -Q$ is an ideal. Since the LICQC, SCC, and SOSC hold at every minimizer of (3.5), the boundary Hessian conditions hold at each zero of f_{θ^*} in K (see [39]). By Theorem 2.3 of [31], the coordinate ring $\frac{\mathbb{R}[x]}{J}$ has dimension 0. Let

$$S := \left\{ x \in \mathbb{R}^n \mid \begin{array}{l} f_{\theta^*}(x) = 0, \\ c_j(x) = 0 \quad (j \in \mathcal{E}), \\ c_j(x) \geq 0 \quad (j \in \mathcal{I}) \end{array} \right\}.$$

Clearly, S is the set of all minimizers of (3.5). The SOSC implies that each minimizer of (3.5) is an isolated minimizer. Note that K is compact. Hence, the set S is finite and the vanishing ideal $I(S)$ is zero dimensional. Let $\{h_1, \dots, h_r\}$ be a Gröbner basis of $I(S)$ with respect to a total degree ordering. Since $h_t \equiv 0$ ($t = 1, \dots, r$) on S and J is zero dimensional, by [32, Corollary 7.4.2], there exist $\ell \in \mathbb{N}$, $\phi_0, \phi_j \in \mathbb{R}[x]$, and $\psi_0, \psi_j \in \Sigma[x]$ such that

$$h_t^{2\ell} + \sum_{j \in \mathcal{E}} \phi_j c_j + \phi_0 f_{\theta^*} + \sum_{j \in \mathcal{I}} \psi_j c_j + \psi_0 = 0.$$

When $2k$ is bigger than the degrees of all above polynomials, we have

$$(3.12) \quad \langle h_t^{2\ell}, w^{(k)} \rangle + \sum_{j \in \mathcal{E}} \langle \phi_j c_j, w^{(k)} \rangle + \langle \phi_0 f_{\theta^*}, w^{(k)} \rangle + \sum_{j \in \mathcal{I}} \langle \psi_j c_j, w^{(k)} \rangle + \langle \psi_0, w^{(k)} \rangle = 0.$$

From item (i), we know that $\hat{f}^{2k_1} + \sigma_1 \in \text{Ideal}[c_{eq}]$ and the following holds

$$f_{\theta^*} + \epsilon = \sigma_0 + \epsilon \cdot \omega(\hat{f}/\epsilon) - \gamma \epsilon^{1-2k_1} \hat{f}^{2k_1}.$$

Thus, when k_0 is sufficiently large, we have that for all $\epsilon > 0$,

$$f_{\theta^*} + \epsilon \in \text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0}.$$

Equivalently, there exist $\phi_j^\epsilon \in \mathbb{R}[x]$, and $\psi_0^\epsilon, \psi_j^\epsilon \in \Sigma[x]_{2k_0}$ such that

$$\deg(\phi_j^\epsilon c_j) \leq 2k_0 \quad (j \in \mathcal{E}), \quad \deg(\psi_j^\epsilon c_j) \leq 2k_0 \quad (j \in \mathcal{I}),$$

$$(3.13) \quad f_{\theta^*} + \epsilon = \sum_{j \in \mathcal{E}} \phi_j^\epsilon c_j + \sum_{j \in \mathcal{I}} \psi_j^\epsilon c_j + \psi_0^\epsilon.$$

Applying the bi-linear operation $\langle \cdot, w^{(k)} \rangle$ to (3.13) for $k > k_0$, we get

$$(3.14) \quad \epsilon = \sum_{j \in \mathcal{E}} \langle \phi_j^\epsilon c_j, w^{(k)} \rangle + \sum_{j \in \mathcal{I}} \langle \psi_j^\epsilon c_j, w^{(k)} \rangle + \langle \psi_0^\epsilon, w^{(k)} \rangle.$$

Since $\phi_i^\epsilon \in \mathbb{R}[x]$, and $\psi_0^\epsilon, \psi_j^\epsilon$ are SOS, the constraints

$$\mathcal{V}_{c_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \quad L_{c_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \quad M_k[w] \succeq 0$$

imply that

$$\langle \phi_j^\epsilon c_j, w^{(k)} \rangle = 0 \quad (j \in \mathcal{E}), \quad \langle \psi_j^\epsilon c_j, w^{(k)} \rangle \geq 0 \quad (j \in \mathcal{I}), \quad \langle \psi_0^\epsilon, w^{(k)} \rangle \geq 0.$$

Hence, we have that

$$(3.15) \quad \lim_{\epsilon \rightarrow 0} \langle \psi_j^\epsilon c_j, w^{(k)} \rangle = 0, \quad \lim_{\epsilon \rightarrow 0} \langle \psi_0^\epsilon, w^{(k)} \rangle = 0.$$

By Lemma 2.5 of [37], we can get for k big enough,

$$(3.16) \quad \lim_{\epsilon \rightarrow 0} \langle \psi_j^\epsilon c_j \phi_0, w^{(k)} \rangle = 0, \quad \lim_{\epsilon \rightarrow 0} \langle \psi_0^\epsilon \phi_0, w^{(k)} \rangle = 0.$$

Combining the equations (3.13) and (3.16), it holds that

$$(3.17) \quad \begin{aligned} & \langle \phi_0 f_{\theta^*}, w^{(k)} \rangle \\ &= \lim_{\epsilon \rightarrow 0} \left(\sum_{j \in \mathcal{E}} \langle \phi_j^\epsilon c_j \phi_0, w^{(k)} \rangle + \sum_{j \in \mathcal{I}} \langle \psi_j^\epsilon c_j \phi_0, w^{(k)} \rangle + \langle \psi_0^\epsilon \phi_0, w^{(k)} \rangle \right) \\ &= 0. \end{aligned}$$

Hence, we can get

$$\langle \phi_0 f_{\theta^*}, w^{(k)} \rangle = 0.$$

It follows from (3.12) that

$$(3.18) \quad \langle h_t^{2\ell}, w^{(k)} \rangle + \sum_{j \in \mathcal{E}} \langle \phi_j c_j, w^{(k)} \rangle + \sum_{j \in \mathcal{I}} \langle \psi_j c_j, w^{(k)} \rangle + \langle \psi_0, w^{(k)} \rangle = 0.$$

Since $h_t^{2\ell}, \psi_j$ ($j \in \mathcal{I}$), ψ_0 are SOS, we can easily verify that

$$\langle h_t^{2\ell}, w^{(k)} \rangle \geq 0, \quad \langle \phi_j c_j, w^{(k)} \rangle = 0 \quad (j \in \mathcal{E}),$$

$$\langle \psi_0, w^{(k)} \rangle \geq 0, \quad \langle \psi_j c_j, w^{(k)} \rangle \geq 0 \quad (j \in \mathcal{I}).$$

Hence, $\langle h_t^{2\ell}, w^{(k)} \rangle = 0$ and $h_t \in \ker M_k[w^{(k)}]$ (cf. [22, 28, 41]).

Finally, we show that $w^{(k)}|_{2k-2}$ is flat. Since S is finite, we know that the quotient space $\mathbb{R}[x]/I(S)$ is finite-dimensional (see Theorem 2.6 of [28]). Let $\{q_1, \dots, q_T\}$ be a basis of $\mathbb{R}[x]/I(S)$. For every $\alpha \in \mathbb{N}^n$, we have that

$$x^\alpha = \sum_{i=1}^T \beta_i q_i + \sum_{t=1}^r p_t h_t, \quad \deg(p_t h_t) \leq |\alpha|,$$

where $\beta_i \in \mathbb{R}, p_t \in \mathbb{R}[x]$. Since $h_t \in \ker M_k[w^{(k)}]$, if $|\alpha| \leq k-1$, we have

$$p_t h_t \in \ker M_k[w^{(k)}], \quad x^\alpha - \sum_{i=1}^T \beta_i q_i \in \ker M_k[w^{(k)}].$$

Let $d_q := \max_{i \in [T]} \deg(q_i)$. If $d_q + 1 \leq |\alpha| \leq k - 1$, then each α th column of $M_k[w^{(k)}]$ is a linear combination of its first d_q columns. When $k - 1 - d_K \geq d_q$, we have

$$\text{rank } M_{k-1-d_K}[w^{(k)}] = \text{rank } M_{k-1}[w^{(k)}].$$

This implies that $w^{(k)}$ has a flat truncation when k is big enough. \square

We would like to make the following remarks:

- i) Theorem 3.2 gives sufficient conditions for the Moment-SOS hierarchy to have finite convergence for solving the GMP. It is typically hard to check these optimality conditions a priori or computationally, since the maximizer θ^* of (1.2) is usually not known. However, there is no need to check them in the computational practice of Moment-SOS relaxations. As in Conjecture 6.1, we conjecture that these optimality conditions are satisfied when the defining polynomials of the GMP have generic coefficients.
- ii) In Theorem 3.2, we assume that the optimal value of (1.2) is achievable. By the classical linear conic optimization theory (see [2]), if the Slater condition holds for (1.1), the dual optimization (1.2) has a maximizer. The maximizer of (3.2) may not be unique. When (1.2) has more than one maximizer, Theorem 3.2 only requires Assumption 3.1 to hold for one maximizer θ^* of (1.2), not necessarily all. This is because the proof of Theorem 3.2 only uses optimality conditions of (3.5) for θ^* . We refer to Example 3.4 for such an example.
- iii) In Theorem 3.2, we assume that the dual optimization (1.2) has an interior point, i.e., there exists a feasible point $\bar{\theta} \in \mathbb{R}^m$ such that $f_{\bar{\theta}} > 0$ on K . Then, the strong duality holds for (1.1)–(1.2). There also exist other type conditions, such as closedness of the image of the primal cone under the primal affine map (see [2, Theorem 7.2]), which can also guarantee the strong duality. In our proof, the existence of an interior point for (1.2) is quite important, more than just guaranteeing the strong duality. For instance, it is used to get the inclusion relation (3.7).

3.3. The special case of polynomial optimization. Consider the polynomial optimization problem

$$(3.19) \quad \begin{cases} \min & f(x) \\ \text{s.t.} & c_j(x) = 0 \ (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}), \end{cases}$$

for given polynomials f, c_j . Let f_{min} denote the optimal value of (3.19). It is interesting to note that (3.19) is equivalent to the moment optimization

$$(3.20) \quad \begin{cases} \min & \langle f, y \rangle \\ \text{s.t.} & \langle 1, y \rangle = 1, \\ & y \in \mathcal{R}_d(K). \end{cases}$$

In the above, the equality $\langle 1, y \rangle = 1$ means that every representing measure μ for y must have unit mass (i.e., $\mu(K) = 1$). The dual optimization of (3.20) is

$$(3.21) \quad \begin{cases} \max & \theta \\ \text{s.t.} & f - \theta \cdot 1 \in \mathcal{P}_d(K), \\ & \theta \in \mathbb{R}. \end{cases}$$

The optimal value θ^* of (3.21) is the minimum value f_{\min} of (3.19). The k th order moment relaxation for (3.20) is

$$(3.22) \quad \begin{cases} \min & \langle f, w \rangle \\ \text{s.t.} & \langle 1, w \rangle = 1, \\ & \mathcal{V}_{c_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \\ & L_{c_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \\ & M_k[w] \succeq 0, \quad w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases}$$

When Theorem 3.2 is applied to (3.20)-(3.21), we can get the following theorem.

Theorem 3.3. *Suppose $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is archimedean. If the LICQC, SCC, SOSC hold at every minimizer of (3.19), then every minimizer of the moment relaxation (3.22) must have a flat truncation when k is sufficiently large.*

When $\text{Ideal}[c_{eq}]$ is a real radical ideal, the above result appeared in [41, Section 5.4]. However, the real radicalness of $\text{Ideal}[c_{eq}]$ is not required in Theorem 3.3. This improves the results in the earlier work [37, 41]. When there exists a equality constraint $c_j(x) = 0$ for the set K , we can equivalently replace it by two inequalities $c_j(x) \geq 0$, $-c_j(x) \geq 0$. However, this reformulation will make the linear independence constraint qualification fail at each minimizer of (3.19), and we can not get the finite convergence by using this formulation.

Remark: In Theorem 3.2, we assume that the LICQC, SCC and SOSC hold at every minimizer of (3.5), which depends on the maximizer θ^* of (1.2). When the GMP specializes to the polynomial optimization problem (3.19), the optimization (3.5) reads

$$(3.23) \quad \begin{cases} \min & f(x) - f_{\min} \\ \text{s.t.} & c_j(x) = 0 \quad (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \quad (j \in \mathcal{I}). \end{cases}$$

Clearly, (3.23) is equivalent to (3.19), since their objectives differ only by the constant f_{\min} . Thus, the LICQ, SCC, SOSC hold at every minimizer of (3.23) if and only if they hold at every minimizer of (3.19), which is independent of the maximizers of (1.2).

3.4. Some exposition examples. As shown in Theorem 3.2, if the LICQC, SCC, SOSC hold at every minimizer of (3.5), not only the Moment-SOS hierarchy has finite convergence, but also the flat truncation is satisfied. This is observed in the following example.

Example 3.4. For polynomials

$$\begin{aligned} a_1 &= x_1^2 x_2^4 + x_2^2 x_3^4 + x_3^2 x_1^4, \quad a_2 = x_1^3 x_2^3 + x_2^3 x_3^3 + x_3^3 x_1^3, \\ a_3 &= x_1^5 x_2 + x_2^5 x_3 + x_3^5 x_1, \end{aligned}$$

we consider the GMP

$$(3.24) \quad \begin{cases} \min & \langle x_1^6 + x_2^6 + x_3^6, y \rangle \\ \text{s.t.} & \langle a_i, y \rangle = 1, \quad i = 1, 2, 3, \\ & y \in \mathcal{R}_6(K), \end{cases}$$

where K is the unit sphere $\{x \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 - 1 = 0\}$. The optimal value $\phi^* = 1$. The dual problem of (3.24) is

$$(3.25) \quad \begin{cases} \max & \theta_1 + \theta_2 + \theta_3 \\ \text{s.t.} & x_1^6 + x_2^6 + x_3^6 - \theta_1 a_1 - \theta_2 a_2 - \theta_3 a_3 \in \mathcal{P}_6(K). \end{cases}$$

For $\theta = (0, 0, 0)$, we have $f_\theta(x) = x_1^6 + x_2^6 + x_3^6 \in \text{int}(\mathcal{P}_6(K))$, i.e., the Slater condition holds. We can verify that $\{\theta_1 + \theta_2 + \theta_3 = 1, \theta_i \geq 0 \ (i = 1, 2, 3)\}$ are all maximizers of (3.25), which is infinite. For one maximizer $\theta^* = (0, 1, 0)$, the maximum problem (3.5) reads

$$(3.26) \quad \begin{cases} \min & x_1^6 + x_2^6 + x_3^6 - (x_1^3 x_2^3 + x_2^3 x_3^3 + x_3^3 x_1^3) \\ \text{s.t.} & x_1^2 + x_2^2 + x_3^2 - 1 = 0. \end{cases}$$

All minimizers of (3.26) are

$$\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{-1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}, \frac{-1}{\sqrt{3}}\right).$$

The unit sphere is smooth, so the constraint qualification condition holds at every feasible point. There is no inequality constraint, so strict complementarity is automatically satisfied. We can verify that the second order sufficient condition holds at both minimizers. For instance, at the minimizer $u = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$,

$$\nabla_x^2 L(u) = 3 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^T, \quad G(u)^\perp = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^\perp.$$

Clearly, the SOS is satisfied at u . By Theorem 3.2, the hierarchy of (3.2)-(3.3) has finite convergence. A numerical experiment by GloptiPoly 3 [14] indicates that $\vartheta_3 = \phi_3 = 1$.

We would like to remark that if the ideal $\text{Ideal}[c_{eq}]$ is real radical, the finite convergence of the Moment-SOS hierarchy can be implied by [39]. However, when $\text{Ideal}[c_{eq}]$ is not real radical, the finite convergence still holds. This is shown in Theorem 3.2. The following is such an example.

Example 3.5. We consider the GMP

$$(3.27) \quad \begin{cases} \min & \langle x_1 + \cdots + x_6, y \rangle \\ \text{s.t.} & \langle x_1^3 + \cdots + x_6^3, y \rangle \geq 1, \\ & \langle x_1^4 + \cdots + x_6^4, y \rangle = 1, \\ & y \in \mathcal{R}_4(K), \end{cases}$$

where the set K is

$$K = \{x \in \mathbb{R}^6 : x_1^2 + \cdots + x_6^2 - 2 = 0, 2(x_1^2 + x_2^2 + x_3^2) - x_4 = 0, \\ x_4 - 2(x_5^2 + x_6^2) = 0, x_1 \geq 0, \dots, x_6 \geq 0\}.$$

The optimal value $\phi^* = \frac{2+2\sqrt{2}}{3}$. One can see that $x_4 - 1 \in \text{Ideal}[V_{\mathbb{R}}(c_{eq})]$. However, $x_4 - 1 \notin \text{Ideal}[c_{eq}]$. Suppose otherwise it were true, then there exist $h_1, h_2, h_3 \in \mathbb{R}[x]$ such that

$$x_4 - 1 = h_1 \cdot (x_1^2 + \cdots + x_6^2 - 2) + h_2 \cdot (2(x_1^2 + x_2^2 + x_3^2) - x_4) \\ + h_3 \cdot (x_4 - 2(x_5^2 + x_6^2)).$$

Substitute x for $(\frac{\sqrt{3}\mathbf{i}}{3}, \frac{\sqrt{3}\mathbf{i}}{3}, \frac{\sqrt{3}\mathbf{i}}{3}, -2, \frac{\sqrt{2}\mathbf{i}}{2}, \frac{\sqrt{2}\mathbf{i}}{2})$ (\mathbf{i} is the imaginary unit) in the above identity, we get the contradiction $-3 = 0$. Thus, the ideal $\text{Ideal}[c_{eq}]$ is not real radical. The dual problem of (3.27) is

$$(3.28) \quad \begin{cases} \max & \theta_1 + \theta_2 \\ \text{s.t.} & \sum_{i=1}^6 x_i - \theta_1 \sum_{i=1}^6 x_i^3 - \theta_2 \sum_{i=1}^6 x_i^4 \in \mathcal{P}_4(K), \\ & \theta \in \mathbb{R}^2, \theta_1 \geq 0. \end{cases}$$

For $\theta = (0, 0)$, we have $f_\theta(x) = \sum_{i=1}^6 x_i \in \text{int}(\mathcal{P}_4(K))$, so the Slater condition holds.

One can verify that the maximizer of (3.28) is $\theta^* = (0, \frac{2+2\sqrt{2}}{3})$. The corresponding optimization problem (3.5) reads

$$(3.29) \quad \begin{cases} \min & \sum_{i=1}^6 x_i - \frac{2+2\sqrt{2}}{3} \sum_{i=1}^6 x_i^4 \\ \text{s.t.} & x_1^2 + \dots + x_6^2 - 2 = 0, \\ & 2(x_1^2 + x_2^2 + x_3^2) - x_4 = 0, \\ & x_4 - 2(x_5^2 + x_6^2) = 0, \\ & x_1 \geq 0, \dots, x_6 \geq 0. \end{cases}$$

The set of minimizers for (3.29) consists of the points:

$$\begin{aligned} & (\frac{\sqrt{2}}{2}, 0, 0, 1, \frac{\sqrt{2}}{2}, 0), (0, \frac{\sqrt{2}}{2}, 0, 1, \frac{\sqrt{2}}{2}, 0), (0, 0, \frac{\sqrt{2}}{2}, 1, \frac{\sqrt{2}}{2}, 0), \\ & (\frac{\sqrt{2}}{2}, 0, 0, 1, 0, \frac{\sqrt{2}}{2}), (0, \frac{\sqrt{2}}{2}, 0, 1, 0, \frac{\sqrt{2}}{2}), (0, 0, \frac{\sqrt{2}}{2}, 1, 0, \frac{\sqrt{2}}{2}). \end{aligned}$$

The LICQC, SCC, SOSC hold at all these minimizers. A numerical experiment by GloptiPoly 3 indicates that $\vartheta_3 = \phi_3 = \phi^*$.

We remark that in Theorems 3.2 and 3.3, none of the optimality conditions there can be dropped. We refer to [39] for such counterexamples.

4. THE MOMENT-SOS HIERARCHY FOR GMPs WITH UNBOUNDED SETS

When the feasible set K is unbounded, the Moment-SOS hierarchy (3.2)–(3.3) may not converge. To fix this issue, a new type Moment-SOS hierarchy based on homogenization is proposed in [17]. Its finite convergence is shown if the ideal of equality constraining polynomials is real radical, provided the LICQC, SCC, SOSC hold at every minimizer. In this section, we prove the same conclusion without the real radicalness assumption. In particular, this resolves Conjecture 8.2 raised in [17] when it specializes to polynomial optimization. We also prove similar results for the finite convergence of the Moment-SOS hierarchy with denominators, induced by Putinar-Vasilescu's Positivstellensatz [29, 30, 48].

We begin with Moment-SOS relaxations for solving generalized moment problems with unbounded sets.

4.1. Homogenization of GMPs. This subsection generalizes the homogenization trick given in [17, 18] to solve the GMP with unbounded sets and proves its finite convergence without the real radicalness assumption.

For a polynomial $p(x)$ of degree ℓ , let \tilde{p} denote its homogenization in $\tilde{x} := (x_0, x)$, i.e., $\tilde{p}(\tilde{x}) = x_0^\ell p(x/x_0)$. For the set K as in (3.1), its homogenization \tilde{K} is the set

$$(4.1) \quad \tilde{K} := \left\{ \tilde{x} \in \mathbb{R}^{n+1} \left| \begin{array}{l} \tilde{c}_j(\tilde{x}) = 0 \ (j \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ \|\tilde{x}\|^2 - 1 = 0, \ x_0 \geq 0 \end{array} \right. \right\}.$$

Note that \tilde{K} is always compact. The set K is said to be *closed at infinity* if

$$(4.2) \quad \tilde{K} = cl(\tilde{K} \cap \{x_0 > 0\}).$$

When K is closed at infinity, a polynomial $f \geq 0$ on K if and only if its homogenization $\hat{f} \geq 0$ on \tilde{K} ; see [18, 35]. Denote the polynomial tuples

$$(4.3) \quad \tilde{c}_{eq} := \{\tilde{c}_j(\tilde{x})\}_{j \in \mathcal{E}} \cup \{\|\tilde{x}\|^2 - 1\}, \quad \tilde{c}_{in} := \{\tilde{c}_j(\tilde{x})\}_{j \in \mathcal{I}} \cup \{x_0\}.$$

Recall that d is the degree as in (1.1)-(1.2). Denote the degree- d homogenization

$$(4.4) \quad \hat{f} := x_0^d f(x/x_0), \quad \hat{a}_i := x_0^d a_i(x/x_0) \ (i = 1 \dots, m).$$

The homogenization of the generalized moment problem (1.1) is

$$(4.5) \quad \begin{cases} \min & \langle \hat{f}, z \rangle \\ \text{s.t.} & \langle \hat{a}_i, z \rangle = b_i \ (1 \leq i \leq m_1), \\ & \langle \hat{a}_i, z \rangle \geq b_i \ (m_1 < i \leq m), \\ & z \in \mathcal{H}_d(\tilde{K}). \end{cases}$$

Similarly, the homogenized optimization of (1.2) is

$$(4.6) \quad \begin{cases} \max & \sum_{i=1}^m b_i \theta_i \\ \text{s.t.} & \hat{f} - \sum_{i=1}^m \theta_i \hat{a}_i \in \mathcal{P}_d(\tilde{K}), \\ & \theta \in \mathbb{R}^m, \ \theta_i \geq 0 \ (m_1 < i \leq m). \end{cases}$$

Let $\hat{\phi}^*$, $\hat{\vartheta}^*$ denote the optimal values of (4.5), (4.6) respectively. By the weak duality, $\hat{\vartheta}^* \leq \hat{\phi}^*$. Denote

$$\hat{f}_\theta := \hat{f} - \sum_{i=1}^m \theta_i \hat{a}_i,$$

$$H = \left\{ f - \sum_{i=1}^m \theta_i a_i : \theta \in \mathbb{R}^m, \ \theta_i \geq 0 \ \text{for} \ m_1 < i \leq m \right\}.$$

The following is the relationship between the optimal values of the optimization (1.1)–(1.2) and the homogenized optimization (4.5)–(4.6).

Proposition 4.1. *Suppose K is closed at infinity. Then, we have*

$$\hat{\vartheta}^* = \vartheta^*, \quad \hat{\phi}^* \leq \phi^*.$$

Furthermore, if $H \cap \text{int}(\mathcal{P}_d(K)) \neq \emptyset$, then

$$\hat{\vartheta}^* = \vartheta^* = \hat{\phi}^* = \phi^*.$$

Proof. Since K is closed at infinity, the polynomial $f - \sum_{i=1}^m \theta_i a_i \geq 0$ on K if and only if $\hat{f}_\theta \geq 0$ on \tilde{K} . So, $\hat{\vartheta}^* = \vartheta^*$.

Suppose the tms $y \in \mathbb{R}^{\mathbb{N}_d^n}$ is a feasible point of (1.1). Then there exist $u_t \in K$, $\lambda_t \geq 0$ ($t = 1, \dots, r$) such that (see [41, Theorem 8.1.1])

$$(4.7) \quad y = \lambda_1[u_1]_d + \dots + \lambda_r[u_r]_d.$$

For each $t = 1, \dots, r$, let

$$\tilde{\lambda}_t = \lambda_t \|(1, u_t)\|^d, \quad (\tau_t, v_t) = (1, u_t) / \sqrt{1 + \|u_t\|^2}.$$

Then we can see that

$$\tilde{c}_j(\tau_t, v_t) = \tau_t^{\deg(c_j)} c_j\left(\frac{v_t}{\tau_t}\right) = \tau_t^{\deg(c_j)} c_j(u_t) = 0 \quad (j \in \mathcal{E}),$$

$$\tilde{c}_j(\tau_t, v_t) = \tau_t^{\deg(c_j)} c_j\left(\frac{v_t}{\tau_t}\right) = \tau_t^{\deg(c_j)} c_j(u_t) \geq 0 \quad (j \in \mathcal{I}).$$

Hence, we know $(\tau_t, v_t) \in \tilde{K}$. Let

$$(4.8) \quad z = \tilde{\lambda}_1[(\tau_1, v_1)]_d + \dots + \tilde{\lambda}_r[(\tau_r, v_r)]_d.$$

A direct computation shows that

$$\begin{aligned} \langle \hat{f}, z \rangle &= \sum_{t=1}^r \tilde{\lambda}_t \langle \hat{f}, [(\tau_t, v_t)]_d \rangle = \sum_{t=1}^r \tilde{\lambda}_t \hat{f}(\tau_t, v_t) \\ &= \sum_{t=1}^r \tilde{\lambda}_t \tau_t^d f(u_t) = \sum_{t=1}^r \lambda_t f(u_t) = \langle f, y \rangle. \end{aligned}$$

Similarly, we also have

$$\langle \hat{a}_i, z \rangle = \langle a_i, y \rangle = b_i \quad (1 \leq i \leq m_1),$$

$$\langle \hat{a}_i, z \rangle = \langle a_i, y \rangle \geq b_i \quad (m_1 < i \leq m).$$

This shows that z is a feasible point of (4.5) and $\langle \hat{f}, z \rangle = \langle f, y \rangle$. The above holds for every feasible point y of (1.1), so $\hat{\phi}^* \leq \phi^*$.

When $H \cap \text{int}(\mathcal{P}_d(K)) \neq \emptyset$, the strong duality holds between (1.2) and (1.1), i.e., $\phi^* = \vartheta^*$, so,

$$\vartheta^* = \phi^* \geq \hat{\phi}^* \geq \hat{\vartheta}^* = \vartheta^*.$$

Therefore, all these values must be equal to each other. \square

Remark 4.2. In the proof of Proposition 4.1, we have seen that if y as in (4.7) is feasible for (1.1), then z as in (4.8) is feasible for (4.5). The reverse of this is also true under some conditions. Suppose

$$z = \tilde{\lambda}_1[(\tau_1, v_1)]_d + \dots + \tilde{\lambda}_r[(\tau_r, v_r)]_d,$$

is feasible for (4.5), with $\tilde{\lambda}_t \geq 0$, $(\tau_t, v_t) \in \tilde{K}$ and all $\tau_t > 0$. Let

$$\lambda_t = \tilde{\lambda}_t \tau_t^d, \quad u_t = v_t / \tau_t.$$

Then $y = \lambda_1[u_1]_d + \dots + \lambda_r[u_r]_d$ is a feasible point of (1.1) and $\langle \hat{f}, z \rangle = \langle f, y \rangle$. Therefore, if the above z is a minimizer of (4.5), then y is also a minimizer of (1.1).

For a degree $k \geq d_0$, the k th order moment relaxation of (4.5) is

$$(4.9) \quad \left\{ \begin{array}{l} \min \quad \langle \hat{f}, w \rangle \\ \text{s.t.} \quad \langle \hat{a}_i, w \rangle = b_i \quad (1 \leq i \leq m_1), \\ \quad \quad \langle \hat{a}_i, w \rangle \geq b_i \quad (m_1 < i \leq m), \\ \quad \quad \mathcal{V}_{1-\|\tilde{x}\|^2}^{(2k)}[w] = 0, \quad \mathcal{V}_{\tilde{c}_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \\ \quad \quad L_{x_0}^{(k)}[w] \succeq 0, \quad L_{\tilde{c}_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \\ \quad \quad M_k[w] \succeq 0, \quad w \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{array} \right.$$

The k th order SOS relaxation of (4.6) is

$$(4.10) \quad \left\{ \begin{array}{l} \max \quad b_1\theta_1 + \cdots + b_m\theta_m \\ \text{s.t.} \quad \hat{f} - \sum_{i=1}^m \theta_i \hat{a}_i \in \text{Ideal}[\tilde{c}_{eq}]_{2k} + \text{QM}[\tilde{c}_{in}]_{2k}, \\ \quad \quad \theta \in \mathbb{R}^m, \quad \theta_{m_1+1} \geq 0, \dots, \theta_m \geq 0. \end{array} \right.$$

Denote the optimal values of (4.9), (4.10) by $\hat{\phi}_k, \hat{\vartheta}_k$ respectively.

Theorem 4.3. *Suppose K is closed at infinity, $H \cap \text{int}(\mathcal{P}_d(K)) \neq \emptyset$ and the minimizer y^* of (1.1) satisfies $y^* \neq 0$. Assume $\theta^* \in \mathbb{R}^m$ is a maximizer of (1.2). If the LICQC, SCC, SOS hold at every minimizer of the optimization ($\tilde{x} := (x_0, x)$)*

$$(4.11) \quad \left\{ \begin{array}{l} \min \quad \hat{f}(\tilde{x}) - \sum_{i=1}^m \theta_i^* \hat{a}_i(\tilde{x}) \\ \text{s.t.} \quad \tilde{c}_j(\tilde{x}) = 0 \quad (j \in \mathcal{E}), \\ \quad \quad \tilde{c}_j(\tilde{x}) \geq 0 \quad (j \in \mathcal{I}). \end{array} \right.$$

Then, we have:

- (i) *The Moment-SOS hierarchy (4.9)–(4.10) has finite convergence, i.e., for all k sufficiently large,*

$$\hat{\vartheta}_k = \hat{\phi}_k = \vartheta^* = \phi^*.$$

- (ii) *Every minimizer of the moment relaxation (4.9) must have a flat truncation, when k is sufficiently large.*

Proof. It follows from Proposition 4.1 that $\theta^* \in \mathbb{R}^m$ is also a maximizer of (4.6) and $\hat{\vartheta}^* = \vartheta^* = \hat{\phi}^* = \phi^*$. By Theorem 3.4 [18], the condition $H \cap \text{int}(\mathcal{P}_d(K)) \neq \emptyset$ implies that there exists $\bar{\theta} \in \mathbb{R}^m$ with $\bar{\theta}_{m_1+1} \geq 0, \dots, \bar{\theta}_m \geq 0$ such that $\hat{f}_{\bar{\theta}} > 0$ on \tilde{K} , i.e., the Slater condition holds for optimization (4.6). Thus, Assumption 3.1 holds for optimization (4.6). Note that $\text{Ideal}[\tilde{c}_{eq}] + \text{QM}[\tilde{c}_{in}]$ is archimedean. The conclusions follow from Theorem 3.2. \square

Remark: To verify that K is closed at infinity, we need to check the condition (4.2), which can be quite tricky and depends on the choice of describing polynomials for K (see [35]). However, being closed at infinity is a generic property for semialgebraic sets (see [13]). Most frequently appearing unbounded semialgebraic sets are closed at infinity. For instance, the nonnegative orthant \mathbb{R}_+^n , the exterior ball $\{x^T x \geq 1\}$, and the exterior box $\{x_1^2 \geq 1, \dots, x_n^2 \geq 1\}$ are all closed at infinity.

The following is an illustrated example for the GMP with unbounded sets, where the ideal $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical.

Example 4.4. We consider the GMP

$$(4.12) \quad \begin{cases} \min & \langle x_1^4 + (2x_2 - 1)^4 + (2x_3 - 1)^4, y \rangle \\ \text{s.t.} & \langle x_2^2 + x_3^2, y \rangle = 1, \\ & \langle x_1^2 x_2 + x_2^2 x_3 + x_3^2 x_1, y \rangle \geq \frac{1}{5}, \\ & y \in \mathcal{R}_4(K), \end{cases}$$

with the unbounded set from [4]:

$$K = \{x \in \mathbb{R}^3 : x_1 x_2 x_3 = 0, x_3(x_1^2 + x_2^2 + x_3^2 + x_2) = 0, x_2(x_2 + x_3) = 0\}.$$

The optimal values $\phi^* = \hat{\phi}^* = 32$. One can check that $x_1 x_3 \in \text{Ideal}[V_{\mathbb{R}}(\tilde{c}_{eq})]$ but $x_1 x_3 \notin \text{Ideal}[\tilde{c}_{eq}]$. Suppose otherwise it were true, then there exist $h_1, h_2, h_3, h_4 \in \mathbb{R}[\hat{x}]$ such that

$$\begin{aligned} x_1 x_3 &= h_1 \cdot x_1 x_2 x_3 + h_2 \cdot x_3(x_1^2 + x_2^2 + x_3^2 + x_2 x_0) \\ &\quad + h_3 \cdot x_2(x_2 + x_3) + h_4 \cdot (x_0^2 + x_1^2 + x_2^2 + x_3^2 - 1). \end{aligned}$$

Substituting (x_0, x_2) for $(1, 0)$ in the above, we get the representation

$$x_1 x_3 = \left(h_2(1, x_1, 0, x_3)x_3 + h_4(1, x_1, 0, x_3) \right) (x_1^2 + x_3^2),$$

which can never hold. So, $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical. The dual optimization problem of (4.12) is

$$(4.13) \quad \begin{cases} \max & \theta_1 + \frac{1}{5}\theta_2 \\ \text{s.t.} & x_1^4 + (2x_2 - 1)^4 + (2x_3 - 1)^4 - \theta_1(x_2^2 + x_3^2) \\ & \quad - \theta_2(x_1^2 x_2 + x_2^2 x_3 + x_3^2 x_1) \in \mathcal{P}_4(K), \\ & \theta \in \mathbb{R}^2, \theta_2 \geq 0. \end{cases}$$

For $\theta = (0, 0)$, we have $f_{\theta} = x_1^4 + (2x_2 - 1)^4 + (2x_3 - 1)^4 \in \text{int}(\mathcal{P}_4(K))$, so $H \cap \text{int}(\mathcal{P}_4(K)) \neq \emptyset$. A maximizer of (4.13) is $\theta^* = (32, 0)$ and the optimization problem (4.11) reads

$$\begin{cases} \min & x_1^4 + (2x_2 - x_0)^4 + (2x_3 - x_0)^4 - 32(x_2^2 + x_3^2)x_0^2 \\ \text{s.t.} & x_1 x_2 x_3 = 0, x_2(x_2 + x_3) = 0, \\ & x_3(x_1^2 + x_2^2 + x_3^2 + x_2 x_0) = 0, \\ & x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1, \\ & x_0 \geq 0. \end{cases}$$

Its unique minimizer is $(\frac{\sqrt{6}}{3}, 0, \frac{-\sqrt{6}}{6}, \frac{\sqrt{6}}{6})$, and we can verify that the LICQC, SCC, SOS all hold at it. By Theorem 4.3, the hierarchy (4.10)-(4.9) has finite convergence. A numerical experiment by GloptiPoly 3 indicates that $\hat{\nu}_2 = \hat{\phi}_2 = 32$.

4.2. Polynomial optimization with unbounded sets. Consider the polynomial optimization problem

$$(4.14) \quad \begin{cases} \min & f(x) \\ \text{s.t.} & c_j(x) = 0 \quad (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \quad (j \in \mathcal{I}), \end{cases}$$

for given polynomials $f, c_j \in \mathbb{R}[x]$. We consider the case that the feasible set K is unbounded. Let f_{min} denote the optimal value of (4.14). Suppose the degree of f is d . A homogenized Moment-SOS hierarchy is proposed in [17] for solving (4.14) when K is unbounded.

Recall the notations $\tilde{c}_{eq}, \tilde{c}_{in}$ as in (4.3). For a degree $k \geq \lceil \frac{d}{2} \rceil$, the k th order homogenized SOS relaxation of (4.14) is

$$(4.15) \quad \begin{cases} \max & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{x}) - \gamma x_0^d \in \text{Ideal}[\tilde{c}_{eq}]_{2k} + \text{QM}[\tilde{c}_{in}]_{2k}. \end{cases}$$

Its dual optimization is the k th order moment relaxation

$$(4.16) \quad \begin{cases} \min & \langle \tilde{f}, y \rangle \\ \text{s.t.} & \langle x_0^d, w \rangle = 1, \\ & \mathcal{V}_{1-\|\tilde{x}\|^2}^{(2k)}[w] = 0, \mathcal{V}_{\tilde{c}_j}^{(2k)}[w] = 0 \ (j \in \mathcal{E}), \\ & L_{x_0}^{(k)}[w] \succeq 0, L_{\tilde{c}_j}^{(k)}[w] \succeq 0 \ (j \in \mathcal{I}), \\ & M_k[w] \succeq 0, w \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{cases}$$

Let f_k and f'_k denote the optimal values of (4.15), (4.16) respectively. We remark that (4.15)–(4.16) is a special case of (4.9)–(4.10) for (4.14).

We consider the homogenized optimization problem (note $\tilde{x} := (x_0, x)$)

$$(4.17) \quad \begin{cases} \min & F(\tilde{x}) := \tilde{f}(\tilde{x}) - f_{min} \cdot x_0^d \\ \text{s.t.} & \tilde{c}_j(\tilde{x}) = 0 \ (j \in \mathcal{E}), \\ & x_0^2 + \|x\|^2 - 1 = 0, \\ & \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ & x_0 \geq 0. \end{cases}$$

Assume the optimal value $f_{min} > -\infty$ and K is closed at infinity. Note that the optimal value of (4.17) is 0, and a feasible point $x^* \in K$ is a minimizer of (4.14) if and only if the point $\tilde{x}^* := (1 + \|x^*\|^2)^{-\frac{1}{2}}(1, x^*)$ is a minimizer of (4.17).

In the work [17], the Moment-SOS hierarchy (4.15)–(4.16) is shown to have finite convergence if $\text{Ideal}[\tilde{c}_{eq}]$ is real radical, provided the LICQC, SCC and SOSC hold at every minimizer of (4.17). We remark that these optimality conditions are essentially equivalent for (4.14) and (4.17). Here, we prove the same conclusion holds even if $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical. This resolves Conjecture 8.2 of the work [17].

Theorem 4.5. *Suppose K is closed at infinity and $f \in \mathbb{R} + \text{int}(\mathcal{P}_d(K))$. If the LICQC, SCC and SOSC hold at every minimizer of (4.17), then the Moment-SOS hierarchy (4.15)–(4.16) has finite convergence, i.e., $f_k = f'_k = f_{min}$ for all k sufficiently large. Furthermore, every minimizer of the moment relaxation (4.16) must have a flat truncation, when k is sufficiently large.*

Proof. The optimization (4.14) is equivalent to the following GMP

$$\begin{cases} \min & \langle f, y \rangle \\ \text{s.t.} & \langle 1, y \rangle = 1, \\ & w \in \mathcal{R}_d(K). \end{cases}$$

The assumption $f \in \mathbb{R} + \text{int}(\mathcal{P}_d(K))$ implies the minimum value $f_{min} > -\infty$. For the corresponding optimization problem (1.2), the optimal value is f_{min} and it is achievable. The conclusion then follows directly from Theorem 4.3. \square

The following is an example for the polynomial optimization with unbounded sets, where the ideal $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical.

Example 4.6. We consider the polynomial optimization

$$(4.18) \quad \begin{cases} \min & x_1^6 + x_2^6 + 3x_1^2x_2^2 - x_1^4(x_2^2 + 1) - x_2^4(x_1^2 + 1) - (x_1^2 + x_2^2) \\ \text{s.t.} & x_1^3 + x_1x_2^4 + x_1 = 0, \ x_2 \geq 0. \end{cases}$$

The optimal value $f_{min} = -1$. One can check that $x_1 \in \text{Ideal}[V_{\mathbb{R}}(\tilde{c}_{eq})]$ but $x_1 \notin \text{Ideal}[\tilde{c}_{eq}]$. Suppose otherwise it were true, then there exist $h_1, h_2 \in \mathbb{R}[\tilde{x}]$ such that

$$x_1 = h_1 \cdot (x_1^3 x_0^2 + x_1 x_2^4 + x_1 x_0^4) + h_2 \cdot (x_1^2 + x_2^2 + x_0^2 - 1).$$

Substituting x for $(0, 1, 0)$ in the above, we get the contradiction $1 = 0$. So, the ideal $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical. The optimization problem (4.17) reads

$$\begin{cases} \min & x_0^6 + x_1^6 + x_2^6 + 3x_1^2 x_2^2 x_0^2 - x_1^4 (x_2^2 + x_0^2) \\ & \quad - x_2^4 (x_1^2 + x_0^2) - x_0^4 (x_1^2 + x_2^2) \\ \text{s.t.} & x_1^3 x_0^2 + x_1 x_2^4 + x_1 x_0^4 = 0, \\ & x_0^2 + x_1^2 + x_2^2 = 1, \\ & x_0 \geq 0, x_2 \geq 0. \end{cases}$$

Its unique minimizer is $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ and we can verify that the LICQC, SCC, SOSC hold at it. By Theorem 4.5, the hierarchy (4.15)-(4.16) has finite convergence. A numerical experiment by GloptiPoly 3 indicates that $f_3 = f'_3 = -1$.

4.3. The Moment-SOS hierarchy with denominators. Let K be the feasible set of (4.14). The Putinar-Vasilescu's Positivstellensatz states that for a polynomial f , if the highest degree homogeneous part f^{hom} is a positive definite form and $f > 0$ on K , then

$$(1 + \|x\|^2)^k f \in \text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$$

for some power $k \in \mathbb{N}$; see [48]. This motivates the Moment-SOS hierarchy with denominators to solve (4.14):

$$(4.19) \quad \begin{cases} \max & \gamma \\ \text{s.t.} & \theta^k (f - \gamma) \in \text{Ideal}[c_{eq}]_{2k} + \text{QM}[c_{in}]_{2k}, \end{cases}$$

where $\theta(x) = 1 + \|x\|^2$. Denote by f_k^{den} the optimal value of (4.19). It was conjectured in [29] that the hierarchy of (4.19) has finite convergence under the optimality condition assumptions. This conjecture was shown to hold in [17] when $\text{Ideal}[\tilde{c}_{eq}]$ is real radical. We remark that when the degrees of f, c_j ($j \in \mathcal{I}$) are all even, the constraint $x_0 \geq 0$ is redundant to define the optimization (4.17) and the SCC will fail for the minimizers of (4.17) with $x_0 = 0$ (cf. [17]). In this case, we consider the alternative optimization

$$(4.20) \quad \begin{cases} \min & F(\tilde{x}) := \tilde{f}(\tilde{x}) - f_{min} \cdot x_0^d \\ \text{s.t.} & \tilde{c}_j(\tilde{x}) = 0 \ (j \in \mathcal{E}), \\ & x_0^2 + \|x\|^2 - 1 = 0, \\ & \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}). \end{cases}$$

In the following, we show that finite convergence also holds even if $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical. This resolves the conjecture made in [29, Section 4.2] about the finite convergence of (4.19) under optimality condition assumptions.

Theorem 4.7. *Let K be the feasible set of (4.14).*

- (i) *Suppose K is closed at infinity, the degrees of f, c_j ($j \in \mathcal{I}$) are all even, and $f \in \mathbb{R} + \text{int}(\mathcal{P}_d(K))$. If the LICQC, SCC and SOSC hold at every minimizer of (4.20), then $f_k^{den} = f_{min}$ for all k big enough.*
- (ii) *Suppose f^{hom} is a positive definite form. If the LICQC, SCC and SOSC hold at every minimizer of (4.17), then $f_k^{den} = f_{min}$ for all k big enough.*

Proof. (i) Let $d_f := d/2$, $d_i := \deg(c_j)/2$ ($j \in \mathcal{I}$). Denote $\tilde{c}_{ie} := \{\tilde{c}_j(\tilde{x})\}_{j \in \mathcal{I}}$. For a degree $k \in \mathbb{N}$, consider the following alternative SOS relaxations

$$(4.21) \quad \begin{cases} \max & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{x}) - \gamma x_0^d \in \text{Ideal}[\tilde{c}_{eq}]_{2k} + \text{QM}[\tilde{c}_{ie}]_{2k}. \end{cases}$$

Similar to Theorem 4.5, we know that the hierarchy (4.21) has finite convergence, under the given assumptions of item (i). This means that there exists an integer k_0 such that for all $\gamma < f_{min}$, we have that

$$(4.22) \quad \tilde{f} - \gamma x_0^d \in \text{Ideal}[\tilde{c}_{eq}]_{2k_0} + \text{QM}[\tilde{c}_{ie}]_{2k_0}.$$

Equivalently, there exist $h_0, h_j \in \mathbb{R}[\tilde{x}]_{2k_0}$, $\sigma_0, \sigma_j \in \Sigma[\tilde{x}]_{2k_0}$ such that

$$\begin{aligned} \deg(h_j c_j) &\leq 2k_0, \quad \deg(\sigma_j c_j) \leq 2k_0, \\ \tilde{f} - \gamma x_0^d &= \sum_{j \in \mathcal{E}} h_j \tilde{c}_j + \sum_{j \in \mathcal{I}} \sigma_j \tilde{c}_j + \sigma_0 + h_0 \cdot (\|\tilde{x}\|^2 - 1). \end{aligned}$$

The above induces that

$$(4.23) \quad \begin{aligned} 2(\tilde{f} - \gamma x_0^d) &= \sum_{j \in \mathcal{E}} (h_j \tilde{c}_j + h_j(-\tilde{x})\tilde{c}_j(-\tilde{x})) + \sum_{j \in \mathcal{I}} (\sigma_j + \sigma_j(-\tilde{x}))\tilde{c}_j \\ &\quad + \sigma_0 + \sigma_0(-\tilde{x}) + (h_0 + h_0(-x)) \cdot (\|\tilde{x}\|^2 - 1). \end{aligned}$$

if we substitute \tilde{x} for $\frac{(1,x)}{\sqrt{\theta}}$ in the above, then

$$\begin{aligned} 2 \cdot \frac{f(x) - \gamma}{\theta^{d_f}} &= \sum_{j \in \mathcal{E}} \left[h_j \left(\frac{(1,x)}{\sqrt{\theta}} \right) + (-1)^{\deg(c_j)} h_j \left(\frac{(-1,-x)}{\sqrt{\theta}} \right) \right] \frac{c_j(x)}{\sqrt{\theta}^{\deg(c_j)}} \\ &\quad + \sum_{j \in \mathcal{I}} \left[\sigma_j \left(\frac{(1,x)}{\sqrt{\theta}} \right) + \sigma_j \left(\frac{(-1,-x)}{\sqrt{\theta}} \right) \right] \frac{c_j(x)}{\theta^{d_j}} \\ &\quad + \left[\sigma_0 \left(\frac{(1,x)}{\sqrt{\theta}} \right) + \sigma_0 \left(\frac{(-1,-x)}{\sqrt{\theta}} \right) \right]. \end{aligned}$$

Note that the odd degree terms in $\sigma_j \left(\frac{(1,x)}{\sqrt{\theta}} \right) + \sigma_j \left(\frac{(-1,-x)}{\sqrt{\theta}} \right)$ are cancelled. The above implies that there exists $k_1 \in \mathbb{N}$ such that

$$\theta^k (f - \gamma) \in \text{Ideal}[c_{eq}]_{2k+2d_f} + \text{QM}[c_{in}]_{2k+2d_f}$$

for all $k \geq k_1$, and k_1 is independent of γ . Hence, $f_k^{den} = f_{min}$ for all $k \geq k_1$.

(ii) For each $j \in \mathcal{I}$, let $\theta_j := 2 \lceil \frac{\deg(c_j)}{2} \rceil - \deg(c_j)$. Consider the optimization problem

$$(4.24) \quad \begin{cases} \min & \tilde{f}(\tilde{x}) - f_{min} \cdot (x_0)^d \\ \text{s.t.} & \tilde{c}_j(\tilde{x}) = 0 \quad (j \in \mathcal{E}), \\ & \|\tilde{x}\|^2 - 1 = 0, \\ & x_0^{\theta_j} \tilde{c}_j(\tilde{x}) \geq 0 \quad (j \in \mathcal{I}). \end{cases}$$

Note that the degree of f must be even. Let $\tilde{u} = (u_0, u)$ be a feasible point of (4.24). If $u_0 = 0$, then $\|u\| = 1$ and

$$\tilde{f}(\tilde{u}) - f_{min} \cdot (u_0)^d = f^{hom}(u) > 0,$$

since f^{hom} is positive definite. If $u_0 \neq 0$, then u/u_0 is feasible for (4.14) and

$$\tilde{f}(\tilde{u}) - f_{min} \cdot (u_0)^d = u_0^d (f(u/u_0) - f_{min}) \geq 0.$$

Hence, the optimal value of (4.24) is zero. As for Theorem 4.4 of [17], the LICQC, SCC and SOS hold at every minimizer of (4.24). The conclusion then follows from item (i). \square

Remark: Theorem 4.5 (Theorem 4.7, respectively) assumes that the LICQC, SCC and SOSC hold at every minimizer of (4.17) (the optimization (4.20), respectively). The objective function depends on the minimum value f_{\min} . Suppose $\tilde{x}^* = (x_0^*, x^*)$ is a minimizer of (4.17). If $x_0^* > 0$, it was shown in [17] that x^*/x_0^* is a minimizer of (4.14), and the LICQC, SCC, SOSC hold at \tilde{x}^* of (4.17) if and only if they hold at x^*/x_0^* of (4.14). For this case, these optimality conditions are independent of the minimum value f_{\min} . If $x_0^* = 0$, then the LICQC, SCC and SOSC at \tilde{x}^* for (4.17) depend on f_{\min} . The same comments hold for the optimization (4.20). We refer to [17] for the relationship between their optimality conditions.

The following is an example to illustrate the finite convergence of (4.19).

Example 4.8. We consider the polynomial optimization

$$(4.25) \quad \begin{cases} \min & x_1^2(x_1 - 1)^2 + (x_2 - 1)^2(x_2 - 2)^2 + (x_3 + 1)^2x_3^2 \\ & + 2x_1(x_2 - 1)(x_3 + 1)(x_1 + x_2 + x_3 - 2) \\ \text{s.t.} & x_1^3 - x_2^3 - x_3^3 - 1 = 0, \\ & (x_1^4 + 1)(x_1 - 1) + (x_1^2 - x_1)(x_1x_2^2 - 2x_2) = 0. \end{cases}$$

The optimal value $f_{\min} = 0$. One can verify that $x_2 + x_3 \in \text{Ideal}[V_{\mathbb{R}}(\tilde{c}_{eq})]$ but $x_2 + x_3 \notin \text{Ideal}[\tilde{c}_{eq}]$. Suppose otherwise it were true, then there exist $h_1, h_2, h_3 \in \mathbb{R}[\tilde{x}]$ such that

$$\begin{aligned} x_2 + x_3 &= h_1 \cdot (x_1^3 - x_2^3 - x_3^3 - x_0^3) + h_2 \cdot (x_1^2 + x_2^2 + x_3^2 + x_0^2 - 1) \\ &+ h_3 \cdot ((x_1^4 + x_0^4)(x_1 - x_0) + (x_1^2 - x_1x_0)(x_1x_2^2 - 2x_2x_0^2)), \end{aligned}$$

Substituting (x_0, x_1) for $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ in the above, we can get

$$x_2 + x_3 = -h_1\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, x_2, x_3\right)(x_2^3 + x_3^3) + h_2\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, x_2, x_3\right)(x_2^2 + x_3^2),$$

which can never hold. So, $\text{Ideal}[\tilde{c}_{eq}]$ is not real radical. The homogenized optimization problem (4.17) reads

$$(4.26) \quad \begin{cases} \min & x_1^2(x_1 - x_0)^2 + (x_2 - x_0)^2(x_2 - 2x_0)^2 + (x_3 + x_0)^2x_3^2 \\ & + 2x_1(x_2 - x_0)(x_3 + x_0)(x_1 + x_2 + x_3 - 2x_0) \\ \text{s.t.} & x_1^3 - x_2^3 - x_3^3 - x_0^3 = 0, \\ & (x_1^4 + x_0^4)(x_1 - x_0) + (x_1^2 - x_0x_1)(x_1x_2^2 - 2x_2x_0^2) = 0, \\ & x_0^2 + x_1^2 + x_2^2 + x_3^2 = 1, \\ & x_0 \geq 0. \end{cases}$$

Its unique minimizer is $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$, and we can verify that the LICQC, SCC, SOSC hold at it. By Theorem 4.7, the hierarchy (4.19) has finite convergence. A numerical experiment by GloptiPoly 3 indicates that $f_3^{\text{den}} = 0$.

5. THE GMP FOR THE SUPER RESOLUTION

GMPs have broad applications. This section gives the application on the super resolution problem for semialgebraic sets. The super resolution problem aims to reconstruct high-dimensional sparse vectors from the observation of a low-pass filter [6, 7, 15]. This problem can be formulated as the GMP:

$$(5.1) \quad \begin{cases} \min & \langle 1, y \rangle + \langle 1, z \rangle \\ \text{s.t.} & \langle a_i, y \rangle - \langle a_i, z \rangle = b_i \quad (1 \leq i \leq m), \\ & y, z \in \mathcal{B}_d(K). \end{cases}$$

In the above, $b = (b_1, \dots, b_m)$ is a given vector, K is the semialgebraic set as in (3.1) and $\{a_i\}_{i=1}^m \subseteq \mathbb{R}[x]_d$ is a set of linearly independent polynomial functions on K . The dual optimization of (5.1) is

$$(5.2) \quad \begin{cases} \max & b_1\theta_1 + \dots + b_m\theta_m \\ \text{s.t.} & 1 - \sum_{i=1}^m \theta_i a_i \in \mathcal{P}_d(K), \\ & 1 + \sum_{i=1}^m \theta_i a_i \in \mathcal{P}_d(K), \\ & \theta = (\theta_1, \dots, \theta_m) \in \mathbb{R}^m. \end{cases}$$

For a degree $k \geq d$, the k th moment relaxation for (5.1) is

$$(5.3) \quad \begin{cases} \min & \langle 1, v \rangle + \langle 1, w \rangle \\ \text{s.t.} & \langle a_i, v \rangle - \langle a_i, w \rangle = b_i \quad (1 \leq i \leq m), \\ & \mathcal{V}_{c_j}^{(2k)}[v] = 0, \mathcal{V}_{c_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \\ & L_{c_j}^{(k)}[v] \succeq 0, L_{c_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \\ & M_k[v] \succeq 0, M_k[w] \succeq 0, v, w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases}$$

The dual optimization of (5.3) is the k th order SOS relaxation for (5.2):

$$(5.4) \quad \begin{cases} \max & b_1\theta_1 + \dots + b_m\theta_m \\ \text{s.t.} & 1 - \sum_{i=1}^m \theta_i a_i \in \text{Ideal}[c_{eq}]_{2k} + \text{QM}[c_{in}]_{2k}, \\ & 1 + \sum_{i=1}^m \theta_i a_i \in \text{Ideal}[c_{eq}]_{2k} + \text{QM}[c_{in}]_{2k}, \\ & \theta \in \mathbb{R}^m. \end{cases}$$

Let ϕ^* , ϑ^* denote the optimal value of (5.1), (5.2) respectively, and let ϕ_k , ϑ_k denote the optimal value of (5.3), (5.4) respectively. It was observed in [15, Section 5.2.2] that the Moment-SOS hierarchy of (5.3)–(5.4) often has finite convergence. In the following, we study conditions for its finite convergence.

Suppose $\theta^* = (\theta_1^*, \dots, \theta_m^*)$ is a maximizer of (5.2). Consider the optimization problems:

$$(5.5) \quad \begin{cases} \min & 1 - \sum_{i=1}^m \theta_i^* a_i(x) \\ \text{s.t.} & c_j(x) = 0 \quad (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \quad (j \in \mathcal{I}), \end{cases}$$

and

$$(5.6) \quad \begin{cases} \min & 1 + \sum_{i=1}^m \theta_i^* a_i(x) \\ \text{s.t.} & c_j(x) = 0 \quad (j \in \mathcal{E}), \\ & c_j(x) \geq 0 \quad (j \in \mathcal{I}). \end{cases}$$

Note that the objective polynomials in (5.5) and (5.6) are both nonnegative on the set K .

Theorem 5.1. *Suppose $\text{Ideal}[c_{eq}] + \text{QM}[c_{in}]$ is archimedean and $\theta^* = (\theta_1^*, \dots, \theta_m^*)$ is a maximizer of (5.2). If the LICQC, SCC and SOS hold at every minimizer of (5.5) and (5.6), then we have:*

- (i) *The hierarchy of (5.3)–(5.4) has finite convergence, i.e., $\vartheta_k = \phi_k = \vartheta^* = \phi^*$ for all k big enough.*

- (ii) *Every minimizer of the moment relaxation (5.3) must have a flat truncation, when k is sufficiently large.*

Proof. (i) Note that for $\theta = (0, \dots, 0)$, we have

$$1 - \sum_{i=1}^m \theta_i a_i(x) = 1 + \sum_{i=1}^m \theta_i a_i(x) = 1 \in \text{int}(\mathcal{P}_d(K)).$$

Thus, the Slater condition holds for the pair (5.1)–(5.2). By following the same proof of Theorem 3.2 (i), we can show that for all $\epsilon > 0$, there exists $k_0 \in \mathbb{N}$, independent of ϵ , such that

$$\begin{aligned} 1 - \sum_{i=1}^m \theta_i^* a_i(x) + \epsilon &\in \text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0}, \\ 1 + \sum_{i=1}^m \theta_i^* a_i(x) + \epsilon &\in \text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0}. \end{aligned}$$

Thus, we know $\frac{1}{1+\epsilon} \cdot \theta^*$ is feasible for (5.4) at the relaxation order k_0 , for all $\epsilon > 0$. As $\epsilon \rightarrow 0$, we get

$$(5.7) \quad b^\top \left(\frac{1}{1+\epsilon} \cdot \theta^* \right) = \frac{1}{1+\epsilon} b^\top \theta^* \rightarrow b^\top \theta^*.$$

This implies that $\vartheta_{k_0} = b^\top \theta^* = \vartheta^*$ and then $\vartheta^* = \phi^* = \vartheta_{k_0} = \phi_{k_0}$. Hence, the Moment-SOS hierarchy of (5.3)–(5.4) has finite convergence.

(ii) Suppose $(v^{(k)}, w^{(k)})$ is a minimizer of (5.3) for the relaxation order k . The k th order moment relaxations for polynomial optimization problems (5.5) and (5.6) are respectively

$$(5.8) \quad \begin{cases} \min & \langle 1 - \sum_{i=1}^m \theta_i^* a_i, v \rangle \\ \text{s.t.} & \langle 1, v \rangle = 1, \\ & \mathcal{V}_{c_j}^{(2k)}[v] = 0 \quad (j \in \mathcal{E}), \\ & L_{c_j}^{(k)}[v] \succeq 0 \quad (j \in \mathcal{I}), \\ & M_k[v] \succeq 0, \quad v \in \mathbb{R}^{\mathbb{N}_{2k}^n}, \end{cases}$$

and

$$(5.9) \quad \begin{cases} \min & \langle 1 + \sum_{i=1}^m \theta_i^* a_i, w \rangle \\ \text{s.t.} & \langle 1, w \rangle = 1, \\ & \mathcal{V}_{c_j}^{(2k)}[w] = 0 \quad (j \in \mathcal{E}), \\ & L_{c_j}^{(k)}[w] \succeq 0 \quad (j \in \mathcal{I}), \\ & M_k[w] \succeq 0, \quad w \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases}$$

It follows from (i) that for all $k \geq k_0$,

$$\langle 1, v^{(k)} \rangle + \langle 1, w^{(k)} \rangle = \sum_{i=1}^m b_i \theta_i^* = \sum_{i=1}^m (\langle a_i, v^{(k)} \rangle - \langle a_i, w^{(k)} \rangle) \theta_i^*.$$

Hence, we get

$$(5.10) \quad \langle 1 - \sum_{i=1}^m \theta_i^* a_i, v^{(k)} \rangle + \langle 1 - \sum_{i=1}^m \theta_i^* a_i, w^{(k)} \rangle = 0.$$

In the item (i), we have shown that

$$\left\{1 - \sum_{i=1}^m \theta_i^* a_i, 1 + \sum_{i=1}^m \theta_i^* a_i\right\} \subseteq \text{cl}(\text{Ideal}[c_{eq}]_{2k_0} + \text{QM}[c_{in}]_{2k_0}).$$

Since $v^{(k)}, w^{(k)}$ are feasible for (5.8), (5.9) respectively, we have

$$\left\langle 1 - \sum_{i=1}^m \theta_i^* a_i, v^{(k)} \right\rangle \geq 0, \quad \left\langle 1 + \sum_{i=1}^m \theta_i^* a_i, w^{(k)} \right\rangle \geq 0.$$

Combining with (5.10), we get

$$\left\langle 1 - \sum_{i=1}^m \theta_i^* a_i, v^{(k)} \right\rangle = 0, \quad \left\langle 1 + \sum_{i=1}^m \theta_i^* a_i, w^{(k)} \right\rangle = 0.$$

We show that $v^{(k)}, w^{(k)}$ both have flat truncations, in four cases.

Case I: $(v^{(k)})_0 = (w^{(k)})_0 = 0$. Then, for all $|\alpha| \leq 2k - 2$, we have $(v^{(k)})_\alpha = 0, (w^{(k)})_\alpha = 0$. This can be implied by Lemma 5.7 of [28]. Thus, the truncations $v^{(k)}|_{2k-2}$ and $w^{(k)}|_{2k-2}$ are flat.

Case II: $(v^{(k)})_0 = 0, (w^{(k)})_0 > 0$. As for case I, the truncation $v^{(k)}|_{2k-2}$ is flat. Since $(w^{(k)})_0 > 0$ and $\left\langle 1 + \sum_{i=1}^m \theta_i^* a_i, w^{(k)} \right\rangle = 0$, the normalization $w^{(k)}/w_0^{(k)}$ is a minimizer of (5.9) for all $k \geq k_0$. It follows from Theorem 3.3 that $w^{(k)}$ has a flat truncation, when k is sufficiently large.

Case III: $(v^{(k)})_0 > 0, (w^{(k)})_0 = 0$. The proof is the same as for Case II.

Case IV: $(v^{(k)})_0 > 0, (w^{(k)})_0 > 0$. Then, $v^{(k)}/(v^{(k)})_0$ is a minimizer of (5.8) and $w^{(k)}/(w^{(k)})_0$ is a minimizer of (5.9) for all $k \geq k_0$. By Theorem 3.3, we know that both $v^{(k)}$ and $w^{(k)}$ have flat truncations for sufficiently large k . □

6. CONCLUSIONS AND DISCUSSIONS

In this paper, we prove the finite convergence of the Moment-SOS hierarchy for solving generalized moment problems under the archimedeaness and optimality conditions, but without real radicalness of the equality constraint ideal. This improves the finite convergence theory in the earlier work. When the constraint set K is unbounded (the archimedeaness fails in this case), we propose a homogenized Moment-SOS hierarchy and prove similar finite convergence results. The applications of these results in polynomial optimization with unbounded sets are also discussed.

There still exist many interesting problems for future work. One of them is about finite convergence of the Moment-SOS hierarchy for generic GMPs. To be more specific, will the finite convergence always hold except a zero-measure set in the space of input polynomials? We have shown the finite convergence under the archimedeaness and optimality conditions for (3.5). To the best of authors' knowledge, it is an open question whether or not the LICQC, SCC, SOSC hold at every minimizer of (3.5), when the polynomials f, a_i, b_i, c_j have generic coefficients. Note that $\theta^* = (\theta_1^*, \dots, \theta_m^*)$ is an optimizer of (1.2) and it is not generic but determined by f, a_i, b_i, c_j . Therefore, we pose the following conjecture.

Conjecture 6.1. Let d_0 and d_i be positive degrees, for $i \in [m] \cup \mathcal{E} \cup \mathcal{I}$. Consider $f \in \mathbb{R}[x]_{d_0}$, $a_i \in \mathbb{R}[x]_{d_i}$ ($i \in [m]$), $b \in \mathbb{R}^m$, $c_j \in \mathbb{R}[x]_{d_j}$ ($j \in \mathcal{E} \cup \mathcal{I}$). Then there exists a finite set of polynomials $\varphi_1, \dots, \varphi_L$, which are in the coefficients of polynomials f , $a = (a_i)_{i \in [m]}$, b , $c = (c_j)_{j \in \mathcal{E} \cup \mathcal{I}}$ such that if

$$\varphi_k(f, a, b, c) \neq 0, \quad k = 1 \dots, L,$$

then the Moment-SOS hierarchy (3.2)–(3.3) has finite convergence.

Acknowledgements. The authors would like to thank the editors and anonymous referees for their fruitful suggestions on the paper.

REFERENCES

- [1] L. Baldi and B. Mourrain, *On the effective Putinar's Positivstellensatz and moment approximation*, Math. Program., 200 (2023), pp. 71–103.
- [2] A. Barvinok, *A Course in Convexity*, American Mathematical Society, Providence, 2002.
- [3] D. Bertsekas, *Nonlinear Programming, 2nd edn*, Athena Scientific, Belmont, 1995.
- [4] D. Brake, J. Hauenstein, and A. Liddell, *Validating the completeness of the real solution set of a system of polynomial equations*, in Proceedings of the the ISSAC '16, ACM, New York, 2016, pp. 143–150.
- [5] R. Curto and L. Fialkow, *Truncated K -moment problems in several variables*, J. Oper. Theory 54 (2005), pp. 189–226.
- [6] Y. De Castro and F. Gamboa, *Exact reconstruction using Beurling minimal extrapolation*, J. Math. Anal. Appl., 395 (2012), pp. 336–354.
- [7] Y. De Castro, F. Gamboa, D. Henrion, and J. Lasserre, *Exact solutions to super resolution on semi-algebraic domains in higher dimensions*, IEEE Trans. Inform. Theory, 63 (2017), pp. 621–630.
- [8] E. de Klerk and M. Laurent, *On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems*, SIAM J. Optim., 21 (2011), pp. 824–832.
- [9] E. de Klerk and M. Laurent, *A survey of semidefinite programming approaches to the generalized problem of moments and their error analysis*, in World Women in Mathematics 2018: Proceedings of the First World Meeting for Women in Mathematics (WM), C. Araujo, G. Benkart, C. E. Praeger, and B. Tanbay, eds., Springer, Cham, 2019, pp. 17–56.
- [10] M. Dressler, J. Nie, and Z. Yang, *Separability of Hermitian tensors and PSD decompositions*, Linear Multilinear Algebra, 70 (2022), pp. 6581–6608.
- [11] K. Fang and H. Fawzi, *The sum-of-squares hierarchy on the sphere and applications in quantum information theory*, Math. Program., 190 (2021), pp. 331–360.
- [12] S. Gribling, M. Laurent, and A. Steenkamp, *Bounding the separable rank via polynomial optimization*, Linear Algebra Appl., 648 (2022), pp. 1–55.
- [13] F. Guo, L. Wang, G. Zhou: *Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities*. J. Global Optim., 58 (2014), pp. 261–284.
- [14] D. Henrion, J. Lasserre, and J. Loefberg, *Gloptipoly 3: moments, optimization and semidefinite programming*, Optim. Methods Softw., 24 (2009), pp. 761–779.
- [15] D. Henrion, M. Korda, and J. Lasserre, *The Moment-SOS Hierarchy: Lectures In Probability, Statistics, Computational Geometry, Control And Nonlinear Pdes*, World Scientific, Cambridge, 2020.
- [16] C. Hillar and J. Nie, *An elementary and constructive proof of Hilbert's 17th Problem for matrices*, Proc. Amer. Math. Soc., 136 (2008), pp. 73–76.
- [17] L. Huang, J. Nie, and Y. Yuan, *Homogenization for polynomial optimization with unbounded sets*, Math. Program., 200 (2023), pp. 105–145.
- [18] L. Huang, J. Nie, and Y. Yuan, *Generalized truncated moment problems with unbounded sets*, J. Sci. Comput., 95 (2023).
- [19] F. Kirschner and E. de Klerk, *Convergence rates of RLT and Lasserre-type hierarchies for the generalized moment problem over the simplex and the sphere*, Optim. Lett., 16 (2022), pp. 2191–2208.

- [20] M. Korda, M. Laurent, V. Magron, and A. Steenkamp, *Exploiting ideal-sparsity in the generalized moment problem with application to matrix factorization ranks*, Math. Program., (2023).
- [21] J. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.
- [22] J. Lasserre, M. Laurent, and P. Rostalski, *Semidefinite characterization and computation of zero-dimensional real radical ideals*, Found. Comput. Math., 8 (2008), pp. 607–647.
- [23] J. Lasserre, *A semidefinite programming approach to the generalized problem of moments*, Math. Program., 112 (2008), pp. 65–92.
- [24] J. Lasserre, *Convexity in semi-algebraic geometry and polynomial optimization*, SIAM J. Optim., 19 (2009), pp. 1995–2014.
- [25] J. Lasserre, *An Introduction to Polynomial and Semi-algebraic Optimization*, Cambridge University Press, Cambridge, 2015.
- [26] M. Laurent, *Revisiting two theorems of Curto and Fialkow on moment matrices*, Proc. Am. Math. Soc., 133 (2005), pp. 2965–2976.
- [27] M. Laurent, *Semidefinite representations for finite varieties*, Math. Program., 109 (2007), pp. 1–26.
- [28] M. Laurent, *Sums of squares, moment matrices and optimization over polynomials*, in Emerging applications of algebraic geometry of IMA Volumes in Mathematics and its Applications, IMA Vol. Math. Appl. 149, Springer, New York, 2009, pp. 157–270.
- [29] N. Mai, J. Lasserre, and V. Magron, *Positivity certificates and polynomial optimization on non-compact semialgebraic sets*, Math. Program., 194 (2022), pp. 443–485.
- [30] N. Mai and V. Magron, *On the complexity of Putinar–Vasilescu’s Positivstellensatz*, J. Complexity, 72 (2022).
- [31] M. Marshall, *Representation of non-negative polynomials with finitely many zeros*, Ann. Faculte Sci. Toulouse, 15 (2006), pp. 599–609.
- [32] M. Marshall, *Positive Polynomials and Sums of Squares*, American Mathematical Society, Providence, 2008.
- [33] O. Mula, and A. Nouy, *Moment-SoS Methods for Optimal Transport Problems*, preprint, [arXiv:2211.10742](https://arxiv.org/abs/2211.10742), 2022.
- [34] J. Nie, *Polynomial matrix inequality and semidefinite representation*, Math. Oper. Res., 36 (2011), pp. 398–415.
- [35] J. Nie, *Discriminants and nonnegative polynomials*, J. Symb. Comput., 47 (2012), pp. 167–191.
- [36] J. Nie, *Sum of squares methods for minimizing polynomial forms over spheres and hypersurfaces*, Front. Math. China, 7 (2012), pp. 321–346.
- [37] J. Nie, *Certifying convergence of Lasserre’s hierarchy via flat truncation*, Math. Program., 142 (2013), pp. 485–510.
- [38] J. Nie, *Polynomial optimization with real varieties*, SIAM J. Optim., 23 (2013), pp. 1634–1646.
- [39] J. Nie, *Optimality conditions and finite convergence of Lasserre’s hierarchy*, Math. Program., 146 (2014), pp. 97–121.
- [40] J. Nie, *Symmetric tensor nuclear norms*, SIAM J. Appl. Algebra Geom., 1 (2017), pp. 599–625.
- [41] J. Nie, *Moment and Polynomial Optimization*, SIAM, Philadelphia, 2023.
- [42] J. Nie and M. Schweighofer, *On the complexity of Putinar’s Positivstellensatz*, J. Complexity 23 (2007), pp. 135–150.
- [43] J. Nie, L. Yang, S. Zhong, and G. Zhou, *Distributionally robust optimization with moment ambiguity sets*, J. Sci. Comput. 94 (2022).
- [44] J. Nie and X. Zhang, *Positive maps and separable matrices*, SIAM J. Optim. 26 (2016), pp. 1236–1256.
- [45] J. Nie and X. Zhang, *Real eigenvalues of nonsymmetric tensors*, Comp. Opt. Appl., 70 (2018), pp. 1–32.
- [46] J. Nie and S. Zhong, *Distributionally robust optimization with polynomial robust constraints*, preprint, [arXiv:2308.15591](https://arxiv.org/abs/2308.15591), 2023.
- [47] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, Indiana Univ. Math. J., 42 (1993), pp. 969–984.

- [48] M. Putinar and F. Vasilescu, *Positive polynomials on semi-algebraic sets*, C. R. Acad. Sci. Ser. I Math., 328 (1999), pp. 585–589.
- [49] C. Scheiderer, *Positivity and sums of squares: A guide to recent results*, in Emerging applications of algebraic geometry of IMA Volumes in Mathematics and its Applications, IMA Vol. Math. Appl. 149, Springer, New York, 2009, pp. 271–324.
- [50] L. Slot and M. Laurent, *Sum-of-squares hierarchies for binary polynomial optimization*, Math. Program., 197 (2023), pp. 621–660.

LEI HUANG, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SAN DIEGO, 9500 GILMAN DRIVE, LA JOLLA, CA, USA, 92093.
Email address: leh010@ucsd.edu

JIawang NIE, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SAN DIEGO, 9500 GILMAN DRIVE, LA JOLLA, CA, USA, 92093.
Email address: njw@math.ucsd.edu

YA-XIANG YUAN, INSTITUTE OF COMPUTATIONAL MATHEMATICS AND SCIENTIFIC/ENGINEERING COMPUTING, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES, BEIJING, CHINA, 100049.
Email address: yyx@lsec.cc.ac.cn