HOMOGENIZATION FOR POLYNOMIAL OPTIMIZATION WITH UNBOUNDED SETS

LEI HUANG, JIAWANG NIE, AND YA-XIANG YUAN

ABSTRACT. This paper considers polynomial optimization with unbounded sets. We give a homogenization formulation and propose a hierarchy of Moment-SOS relaxations to solve it. Under the assumptions that the feasible set is closed at infinity and the ideal of homogenized equality constraining polynomials is real radical, we show that this hierarchy of Moment-SOS relaxations has finite convergence, if some optimality conditions (i.e., the linear independence constraint qualification, strict complementarity and second order sufficient conditions) hold at every minimizer, including the one at infinity. Moreover, we prove extended versions of Putinar-Vasilescu type Positivstellensatz for polynomials that are nonnegative on unbounded sets. The classical Moment-SOS hierarchy with denominators is also studied. In particular, we give a positive answer to a conjecture of Mai, Lasserre and Magron in their recent work. Polynomial optimization and Homogenization and Moment-SOS relaxations and Optimality conditions

1. INTRODUCTION

Consider the optimization problem

$$(1.1) \qquad \begin{cases} \min & f(x) \\ s.t. & c_i(x) = 0 \ (i \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}), \end{cases}$$

where $f(x), c_i(x), c_j(x)$ are polynomials in $x := (x_1, \ldots, x_n) \in \mathbb{R}^n$. The $\mathcal{E}$ and $\mathcal{I}$ are disjoint labeling sets for equality and inequality constraining polynomials. Let K denote the feasible set of (1.1) and let $f_{\min}$ denote the optimal value of (1.1). This contains a broad class of important optimization problems, such as stability numbers of graphs [20, 35] and optimization in quantum information theory [8]. We refer to [25, 26, 29] for related work about polynomial optimization.

The Moment-SOS hierarchy proposed by Lasserre [21] is efficient for solving the polynomial optimization (1.1). Under the archimedeanness for constraining polynomials (the set K must be compact for this case; see [21, 28, 46]), it yields a sequence of convergent lower bounds for the minimum $f_{\min}$. Later, it was shown in [41] that the Moment-SOS hierarchy has finite convergence if in addition the linear independence constraint qualification, the strict complementarity and second order sufficient conditions hold at every minimizer.[1] When the set K is compact, we refer to the work [18, 19, 24, 52, 54]. For convex polynomial optimization, the

---

[1]Throughout the paper, for convenience, a minimizer means a global minimizer, unless it is otherwise specified for the meaning.

Moment-SOS hierarchy has finite convergence under the strict convexity or sos-convexity conditions [16, 23]. When the equality constraints give a finite set, the Moment-SOS hierarchy has finite convergence, as shown in [22, 28, 40]. More general introductions to polynomial optimization can be found in the books and surveys [25, 26, 29, 30, 34, 51].

When the feasible set K is unbounded (the archimedeanness fails for this case), the classical Moment-SOS hierarchy typically does not converge. For instance, this is the case if the objective f is a nonnegative polynomial that is not a sum of squares (SOS) and there are no constraints. There exists work on solving polynomial optimization with unbounded sets. Based on Karush-Kuhn-Tucker (KKT) conditions and Lagrange multipliers, there exist tight Moment-SOS relaxations for solving (1.1), such as the work in [5, 36, 39, 43]. In [15], the authors proposed Moment-SOS relaxations based on adding sublevel set constraints. The resulting hierarchy of this type of relaxations is also convergent under the archimedeanness for the new constraints. In the above mentioned work, it is assumed that the optimal value of (1.1) is achieved. However, this is not always the case. For instance, for $f = x_1^2 + (1 - x_1 x_2)^2$ and $K = \mathbb{R}^2$, the optimal value $f_{\min} = 0$ is not achievable at any feasible point. Interestingly, it was shown in [1] that checking whether or not the optimal value is achievable is NP-hard. When the minimum $f_{\min}$ is not achievable, we refer to the work [11, 12, 53] for such polynomial optimization problems.

When the feasible set K is unbounded, Mai, Lasserre and Magron [32] recently proposed a new hierarchy of Moment-SOS relaxations. Instead of solving (1.1) directly, they considered the perturbation $f + \epsilon(1 + \|x\|^2)^{d_0}$ for the objective, where $\epsilon > 0$, and d_0 is the smallest integer such that $2d_0 \geq \deg(f) + 1$. They proved the convergence to a neighborhood of $f_{\min}$ when it is achievable. The convergence is related to Putinar-Vasilescu's Positivstellensatz [47, 48]. The complexity for this new hierarchy is studied in the recent work [31]. For the ideal case $\epsilon = 0$, they conjectured that their hierarchy has finite convergence under some standard optimality conditions at each minimizer.

The polynomial optimization with unbounded sets is typically hard to solve. There are some questions of high interests. For instance, can we get a hierarchy of Moment-SOS relaxations that has finite convergence for almost all cases (i.e., for generic cases)? When the optimal value $f_{\min}$ is not achievable, can we get a hierarchy of Moment-SOS relaxations that has finite convergence?

Contributions. This paper studies how to solve polynomial optimization with unbounded sets. For a polynomial $p(x)$ of degree ℓ, let $\tilde{p}$ denote its homogenization in $\tilde{x} := (x_0, x)$, i.e., $\tilde{p}(\tilde{x}) = x_0^\ell p(x/x_0)$. We consider the following homogenization $\widetilde{K} \subseteq \mathbb{R}^{n+1}$ of K

$$(1.2) \qquad \widetilde{K} := \left\{ \tilde{x} \in \mathbb{R}^{n+1} \,\middle|\, \begin{array}{l} \tilde{c}_i(\tilde{x}) = 0 \ (i \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ \|\tilde{x}\|^2 - 1 = 0, \ x_0 \geq 0 \end{array} \right\}.$$

Denote by $\tilde{c}_{\mathcal{E}}$ (resp., $\tilde{c}_{\mathcal{I}}$) the equality (resp., inequality) constraining polynomial tuple for $\widetilde{K}$. Let $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]_{2k}$ denote the $2k$th degree truncation of the ideal of $\tilde{c}_{\mathcal{E}}$ and let $\mathrm{QM}[\tilde{c}_{\mathcal{I}}]_{2k}$ denote the $2k$th degree truncation of the quadratic module of $\tilde{c}_{\mathcal{I}}$ (see Section 2.2 for the definition). Suppose the degree of f is d. For an order

$k \geq \lceil \frac{d}{2} \rceil$, we consider the kth order relaxation for (1.1):

$$(1.3) \qquad \begin{cases} f_k := \max & \gamma \\ \quad\;\; s.t. & \tilde{f}(\tilde{x}) - \gamma x_0^d \in \mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]_{2k} + \mathrm{QM}[\tilde{c}_{\mathcal{I}}]_{2k}. \end{cases}$$

When the set K is closed at infinity (see Definition 3.1), a scalar γ is a lower bound of f on K if and only if $\tilde{f}(\tilde{x}) - \gamma x_0^d \geq 0$ on $\widetilde{K}$. This is the motivation for considering the relaxation (1.3).

In this paper, we consider the case that the feasible set K of (1.1) is unbounded. Our new contributions are:

I. Our major results are to study conditions for the hierarchy of relaxations (1.3) to have finite convergence. We prove that the finite convergence can happen under some general assumptions, without assuming that f is positive at infinity on K (for such a case, the asymptotic convergence is then guaranteed).

II. Assume that K is closed at infinity, the ideal $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical (see Section 2.2 for the definition). When some optimality conditions (i.e., the linear independence constraint qualification, strict complementarity and second order sufficient conditions) hold at every minimizer of (1.1), including the one at infinity, we show that the hierarchy of (1.3) has finite convergence. In particular, the finite convergence neither assume that the optimal value $f_{\min}$ is achievable nor assume that f is positive at infinity on K. The proof uses some classical results in [41]. To the best of the authors' knowledge, this is the first work that proves the finite convergence for polynomial optimization with unbounded sets.

III. We prove that classical optimality conditions for minimizers of (1.1) are equivalent to those for the homogenized optimization problem (4.1). This shows that the finite convergence of the hierarchy of (1.3) is actually determined by optimality conditions for minimizers of (1.1). When the ideal $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is not real radical, we give a new hierarchy of Moment-SOS relaxations that also has finite convergence, under the same assumptions on optimality conditions. We also study genericity properties of optimality conditions. When the polynomials are generic, we show that optimality conditions hold at every minimizer and there are no minimizers at infinity.

IV. We give extended versions of the Putinar-Vasilescu's Positivstellensatz for polynomials that are nonnegative on unbounded semialgebraic sets. When the linear independence constraint qualification, strict complementarity and second order sufficient conditions hold at every minimizer of the corresponding optimization problem, we prove that a desired SOS type representation with the denominator $\|x\|^{2k}$ or $(1 + \|x\|^2)^k$ exists. The Putinar-Vasilescu's Positivstellensatz gives rise to another hierarchy of Moment-SOS relaxations for solving (1.1). This hierarchy is the same as the one given in [32] for the case $\epsilon = 0$. It was conjectured in [32] that this hierarchy has finite convergence when some standard optimality conditions hold at every minimizer. We prove that this conjecture is true, by applying our finite convergence theory for the hierarchy of (1.3).

We would like to make the following comparisons with prior existing work for solving polynomial optimization with unbounded sets.

- By using KKT conditions and Jacobian representations, a tight hierarchy of Moment-SOS relaxations is given in [39]. By using Lagrange multiplier expressions, an other tight hierarchy of Moment-SOS relaxations is given in [43]. Both methods as in [39, 43] assume that the tuple of constraining polynomials is nonsingular and the optimal value is achievable. In particular, the method as in [39] requires to use defining equations for determinantal varieties. In [43], Lagrange multiplier expressions are required. For general polynomial constraints, it may not be convenient to formulate the Moment-SOS relaxations as in [39, 43]. Moreover, when the optimal value $f_{\min}$ is not achievable, the methods as in [39, 43] are not applicable.
- In the recent work by Jeyakumar et al. [15], a sublevel set constraint $c - f(x) \geq 0$ is posed as a new constraint. When the quadratic module generated for new constraints is archimedean, it produces an asymptotically convergent hierarchy of Moment-SOS relaxations. When the archimedeanness fails, the asymptotic convergence is not guaranteed. Moreover, finding such a c is typically difficult for constraint cases even if it exists and no finite convergence results are known for adding such new constraints.
- When the optimal value $f_{\min}$ is not achievable, there exist convergent hierarchies of SOS relaxations in the work [11, 12, 53], under certain technical assumptions. They are based on gradient tentacles [53] or truncated tangent varieties [11, 12]. Only the asymptotic convergence is shown for these hierarchies, under certain assumptions. There are no finite convergence results when the optimal value $f_{\min}$ is not achievable.
- The recent work [32] of Mai, Lasserre and Magron considers the perturbation $f + \epsilon(1 + \|x\|^2)^{d_0}$, with $2d_0 \geq \deg(f) + 1$, for the objective. When $\epsilon > 0$ is small and $f_{\min}$ is achievable, they give a hierarchy of Moment-SOS relaxations that converges to a neighborhood of the optimal value $f_{\min}$. Finite convergence is not known for this hierarchy. For the case $\epsilon = 0$, the authors conjectured that this hierarchy has finite convergence, under some standard optimality conditions. We give a positive answer to this conjecture in Section 6.

This paper is organized as follows. Section 2 reviews some basics about optimality conditions and polynomial optimization. Section 3 gives the new hierarchy of relaxations and presents some asymptotic convergence results. In Section 4, we prove that the proposed hierarchy of relaxations has finite convergence when some optimality conditions hold at every minimizer, including the one at infinity. In Section 5, we prove extended versions of the Putinar-Vasilescu's Positivstellensatz for polynomials that are nonnegative on unbounded semialgebraic sets. Section 6 studies the Moment-SOS hierarchy with denominators. Section 7 presents some numerical experiments. Section 8 draws conclusions and make some discussions.

2. Preliminaries

Notation. The symbol $\mathbb{N}$ (resp., $\mathbb{R}$, $\mathbb{C}$) denotes the set of nonnegative integers (resp., real numbers, complex numbers). For $x \coloneqq (x_1, \ldots, x_n)$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$, denote $x^\alpha \coloneqq x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| \coloneqq \alpha_1 + \cdots + \alpha_n$. For a degree d, let $[x]_d$ denote the vector of all monomials in x and whose degrees are at most d, ordered in the graded alphabetical ordering, i.e.,

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

Denote the power set

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

Let $\mathbb{R}[x] := \mathbb{R}[x_1, \ldots, x_n]$ denote the ring of polynomials in x with real coefficients, and $\mathbb{R}[x]_d$ is the subset of polynomials in $\mathbb{R}[x]$ with degrees at most d. For a polynomial p, $\deg(p)$ denotes its total degree, and $\tilde{p}$ denotes its homogenization, i.e., $\tilde{p}(\tilde{x}) = x_0^{\deg(p)} p(x/x_0)$ for $\tilde{x} := (x_0, x_1, \ldots, x_n)$. A homogeneous polynomial is said to be a form. A form p is positive definite if $p(x) > 0$ for all nonzero $x \in \mathbb{R}^n$. For a general polynomial $p \in \mathbb{R}[x]$, $p^{(i)}$ denotes the homogeneous part of the ith highest degree for p. For $t \in \mathbb{R}$, $\lceil t \rceil$ denotes the smallest integer greater than or equal to t. For a function p in x, ∇p (resp., $\nabla^2 p$) denotes its gradient (resp., Hessian) with respect to x. If x is a subvector of its variables, then $\nabla_x p$ denotes its gradient with respect to x. If x, y are two subvectors of its variables, then denote the Hessian with respect to $x = (x_{i_1}, \ldots, x_{i_{\ell_1}}), y = (x_{j_1}, \ldots, x_{\ell_2})$ by $\nabla_{x,y}^2 p$, i.e.,

$$\nabla_{x,y}^2 p = \left(\nabla_{x_{i_t} y_{i_k}}^2 p\right)_{t=1,\ldots,\ell_1, k=1,\ldots,\ell_2}.$$

For a matrix A, A^{T} denotes its transpose. A symmetric matrix $X \succeq 0$ if X is positive semidefinite. For a vector v, $\|v\|$ denotes the standard Euclidean norm. For a set $T \subseteq \mathbb{R}^n$, let $cl(T)$ be the closure of T. A property is said to be *generic* for a vector space V if it holds in an open dense set of V.

2.1. **Optimality conditions.** We review some basic theory for nonlinear programming. Let x^* be a local minimizer of (1.1). Denote the label set of active constraints at x^*

$$(2.1) \qquad\qquad J(x^*) := \{i \in \mathcal{E} \cup \mathcal{I} : c_i(x^*) = 0\}.$$

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

$$(2.2) \qquad \boxed{\begin{array}{c} \nabla f(x^*) = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(x^*), \\ \lambda_j \geq 0 \ (j \in \mathcal{I}), \ \lambda_j c_j(x^*) = 0 \ (j \in \mathcal{I}). \end{array}}$$

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

$$\mathscr{L}(x) := f(x) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i c_i(x).$$

The subspace of the gradients of the active constraints is denoted as

$$(2.3) \qquad\qquad V(x^*) := \mathrm{span}\{\nabla c_i(x^*), i \in J(x^*)\}.$$

The orthogonal complement of $V(x^*)$ is denoted as $V(x^*)^\perp$. Under the LICQC, the second order necessary condition (SONC) holds at x^*, i.e.,

$$(2.4) \qquad\qquad v^{\mathrm{T}} \nabla^2 \left(\mathscr{L}(x^*)\right) v \geq 0, \quad \forall v \in V(x^*)^\perp.$$

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

$$(2.5) \qquad\qquad v^{\mathrm{T}} \nabla^2 \left(\mathscr{L}(x^*)\right) v > 0, \quad \forall \, 0 \neq v \in V(x^*)^\perp.$$

If the FOOC, SCC and SOSC hold at x^*, then x^* is a strict local minimizer. We refer to [4, 55] for more details about optimality conditions.

2.2. **Some basics for polynomial optimization.** We review some basics in real algebraic geometry and polynomial optimization. We refer to [25, 29, 51] for more details.

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 $h := (h_1, \ldots, h_m)$, $\text{Ideal}[h]$ denotes the ideal generated by h, i.e.,

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

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

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

Its real variety is defined as

$$V_{\mathbb{R}}(h) = \{x \in \mathbb{R}^n \mid h_1(x) = \cdots = h_m(x) = 0\}.$$

We say $\text{Ideal}[h]$ is *real radical* if $\text{Ideal}[h] = \text{Ideal}[V_{\mathbb{R}}(h)]$, where $\text{Ideal}[V_{\mathbb{R}}(h)]$ denotes the set of all polynomials vanishing on $V_{\mathbb{R}}(h)$. A polynomial p is said to be a sum of squares (SOS) if $p = p_1^2 + \cdots + p_t^2$ for $p_1, \ldots, p_t \in \mathbb{R}[x]$. The set of all SOS polynomials in x is denoted as $\Sigma[x]$. For an even degree k, denote the truncation

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

For a polynomial tuple $g = (g_1, \ldots, g_\ell)$, the quadratic module generated by g is

$$(2.6) \qquad \text{QM}[g] := \Sigma[x] + g_1 \cdot \Sigma[x] + \cdots + g_\ell \cdot \Sigma[x].$$

Similarly, for an even degree k, the kth degree truncation of $\text{QM}[g]$ is

$$(2.7) \qquad \text{QM}[g]_k = \Sigma[x]_k + g_1 \cdot \Sigma[x]_{k-2\lceil \deg(g_1)/2 \rceil} + \cdots + g_\ell \cdot \Sigma[x]_{k-2\lceil \deg(g_\ell)/2 \rceil}.$$

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 referred to as Putinar's Positivstellensatz.

Theorem 2.1 ([46])**.** *Suppose $\text{Ideal}[h] + \text{QM}[g]$ is archimedean. If a polynomial $p > 0$ on S, then $p \in \text{Ideal}[h] + \text{QM}[g]$.*

For an integer $k > 0$, let $\mathbb{R}^{\mathbb{N}_{2k}^n}$ denote the space of all real vectors that are labeled by $\alpha \in \mathbb{N}_{2k}^n$. Each $y \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ is labeled as

$$y = (y_\alpha)_{\alpha \in \mathbb{N}_{2k}^n}.$$

Such y is called a *truncated multi-sequence* (tms) of degree $2k$. For a polynomial $p = \sum_{|\alpha| \leq 2k} p_\alpha x^\alpha \in \mathbb{R}[x]_{2k}$, define the bilinear operation

$$(2.8) \qquad \langle p, y \rangle = \sum_{|\alpha| \leq 2k} p_\alpha y_\alpha.$$

For given p, $\langle p, y \rangle$ is a linear function in y. For the degree $t := k - \lceil \deg(p)/2 \rceil$, the localizing matrix $L_p^{(k)}[y]$ is a symmetric linear matrix function in y such that

$$(2.9) \qquad q^T \left(L_p^{(k)}[y] \right) q \; = \; \langle p(q^T[x]_t)^2, y \rangle$$

for all real vector $q \in \mathbb{R}^{\mathbb{N}_t^n}$. In particular, if $p = 1$ is the constant one polynomial, then $L_1^{(k)}[y]$ is called a moment matrix, for which we denote as

$$M_k[y] \; := \; L_1^{(k)}[y].$$

Localizing matrices are quite useful for solving polynomial and tensor optimization problems [7, 13, 25, 29, 42, 44].

3. Moment-SOS relaxations for the homogenization

In this section, we give a hierarchy of Moment-SOS relaxations based on the homogenization of (1.1) and study its properties. Let $\tilde{x} := (x_0, x)$. For a polynomial $p \in \mathbb{R}[x]$, let $\tilde{p}(\tilde{x})$ denote the homogenization of p. For the feasible set K as in (1.1), denote the sets

$$(3.1) \qquad \begin{aligned} \widetilde{K}^h &:= \left\{ \tilde{x} \; \middle| \; \begin{array}{l} \tilde{c}_i(\tilde{x}) = 0 \; (i \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 \; (j \in \mathcal{I}), \\ x_0 \geq 0 \end{array} \right\}, \\ \widetilde{K}^c &:= \widetilde{K}^h \cap \{x_0 > 0\}, \\ \widetilde{K} &:= \widetilde{K}^h \cap \{x_0^2 + x^T x = 1\}. \end{aligned}$$

Define the perspective projection map φ as

$$(3.2) \qquad \varphi : \left\{ (x_0, x) \in \mathbb{R}^{n+1} \mid x_0 > 0, \; x_0^2 + x^T x = 1 \right\} \to \mathbb{R}^n, \quad (x_0, x) \mapsto \frac{x}{x_0}.$$

The map φ gives a one-to-one correspondence between $\widetilde{K} \cap \{x_0 > 0\}$ and K.

The set $\widetilde{K}^c$ only depends on the geometry of K, while $\widetilde{K}^h$ and $\widetilde{K}$ depend on the description polynomials for K. Throughout the paper, we assume their description polynomials are as in (1.1). The following is a useful definition for homogenization.

Definition 3.1 ([37]). *The set K is closed at infinity (∞) if $cl(\widetilde{K}^c) = \widetilde{K}^h$.*

The above definition was introduced in [37] for studying the boundary of the cone of polynomials nonnegative on K. The following is a basic property for closedness at ∞.

Lemma 3.2. *For $f \in \mathbb{R}[x]$, we have $f \geq 0$ on K if and only if $\tilde{f} \geq 0$ on $cl(\widetilde{K}^c)$. Moreover, when K is closed at ∞, $f \geq 0$ on K if and only if $\tilde{f} \geq 0$ on $\widetilde{K}$.*

Being closed at ∞ is a generic property for semialgebraic sets (see [10]). We remark that the closedness at ∞ may depend on the choice of description polynomials for K. For instance, the following two sets

$$S_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2^2 \geq 0\}, \quad S_2 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 - x_2^2 \geq 0, x_1 \geq 0\}$$

are the same. However, for their description polynomials, the set S_2 is closed at ∞ while S_1 is not, since $(0, -1, 0) \in \widetilde{S}_1^h \backslash cl(\widetilde{S}_1^c)$. Throughout the paper, when we mention K is closed at ∞, it means that their description polynomials as in (1.1) are used.

In view of the minimum value, the optimization (1.1) is equivalent to

$$(3.3) \qquad \begin{cases} \max & \gamma \\ s.t. & f(x) - \gamma \geq 0 \quad \text{on} \quad K. \end{cases}$$

Let $d := \deg(f)$. When K is closed at ∞, one has $f - \gamma \geq 0$ on K if and only if $\tilde{f}(\tilde{x}) - \gamma x_0^d \geq 0$ on $\widetilde{K}$. Therefore, (3.3) is equivalent to

$$(3.4) \qquad \begin{cases} \max & \gamma \\ s.t. & \tilde{f}(\tilde{x}) - \gamma x_0^d \geq 0 \quad \text{on} \quad \widetilde{K}. \end{cases}$$

For convenience of notation, denote the polynomial sets

$$(3.5) \qquad \tilde{c}_{\mathcal{E}} := \{\tilde{c}_i(\tilde{x})\}_{i \in \mathcal{E}} \cup \{\|\tilde{x}\|^2 - 1\}, \quad \tilde{c}_{\mathcal{I}} := \{\tilde{c}_j(\tilde{x})\}_{j \in \mathcal{I}} \cup \{x_0\}.$$

Note that $\widetilde{K}$ is compact and $\text{Ideal}[\tilde{c}_{\mathcal{E}}] + \text{QM}[\tilde{c}_{\mathcal{I}}]$ is archimedean for all K.

We apply Moment-SOS relaxations to solve (3.4). For an order $k \geq \lceil \frac{d}{2} \rceil$, the kth order SOS relaxation for (3.4) is

$$(3.6) \qquad \begin{cases} \max & \gamma \\ s.t. & \tilde{f}(\tilde{x}) - \gamma x_0^d \in \text{Ideal}[\tilde{c}_{\mathcal{E}}]_{2k} + \text{QM}[\tilde{c}_{\mathcal{I}}]_{2k}. \end{cases}$$

The dual optimization of (3.6) is the kth order moment relaxation

$$(3.7) \qquad \begin{cases} \min & \langle \tilde{f}, y \rangle \\ s.t. & L_p^{(k)}[y] = 0 \ (p \in \tilde{c}_{\mathcal{E}}), \\ & L_q^{(k)}[y] \succeq 0 \ (q \in \tilde{c}_{\mathcal{I}}), \\ & M_k[y] \succeq 0, \\ & \langle x_0^d, y \rangle = 1, \ y \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{cases}$$

Let f_k and f_k' denote the optimal values of (3.6), (3.7) respectively. As k goes to infinity, the sequence of the relaxations (3.6)-(3.7) is called the homogenized hierarchy of Moment-SOS relaxations for solving (1.1).

3.1. **Basic properties.** When K is closed at ∞, the optimal values of (3.3) and (3.4) are the same. This is a basic property of homogenization. The following is implied by Lemma 3.2.

Proposition 3.3. *If K is closed at ∞, then the optimal values of (3.3) and (3.4) are the same.*

When the degrees of f, c_j $(j \in \mathcal{I})$ are all even, the constraint $x_0 \geq 0$ is redundant for homogenization. Consider the optimization problem

$$(3.8) \qquad \begin{cases} \max & \gamma \\ s.t. & \tilde{f}(\tilde{x}) - \gamma x_0^d \geq 0 \quad \text{on} \quad \widetilde{K}^e, \end{cases}$$

where the set $\widetilde{K}^e$ is

$$(3.9) \qquad \widetilde{K}^e := \left\{ \tilde{x} \in \mathbb{R}^{n+1} \ \middle| \ \begin{array}{l} \tilde{c}_i(\tilde{x}) = 0 \ (i \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ x_0^2 + x^T x = 1 \end{array} \right\}.$$

Suppose K is closed at ∞ and $f \geq 0$ on K. By Lemma 3.2, we have $\tilde{f}(\tilde{x}) \geq 0$ for $\tilde{x} = (x_0, x) \in \widetilde{K}^e$ with $x_0 \geq 0$. For $\tilde{x} \in \widetilde{K}^e$ with $x_0 < 0$, we have $\tilde{c}_i(\tilde{x}) = (x_0)^{\deg(c_i)} c_i(\frac{x}{x_0})$ for each $i \in \mathcal{E} \cup \mathcal{I}$. This implies that $\frac{x}{x_0} \in K$ and $\tilde{f}(\tilde{x}) = x_0^d f(\frac{x}{x_0}) \geq 0$ for all $x_0 < 0$, when the degrees of f, c_j $(j \in \mathcal{I})$ are even. Thus, the following holds.

Proposition 3.4. *If the set K is closed at ∞ and the degrees of f, c_j $(j \in \mathcal{I})$ are even, then the optimal values of (3.3) and (3.8) are the same.*

We remark that if one of f and c_j $(j \in \mathcal{I})$ has an odd degree, then (3.3) and (3.8) may not have the same optimal value. For instance, consider the optimization

$$\min \quad x \quad s.t. \quad 1 - x^2 \geq 0, \, x \in \mathbb{R}.$$

The minimum value $f_{\min} = -1$ and $\widetilde{K}^e = \{x_0^2 \geq x^2, \, x_0^2 + x^2 = 1\}$. However, the optimal value of (3.8) is $-\infty$, since there is no scalar γ such that $x - \gamma x_0 \geq 0$ on $\widetilde{K}^e$.

Suppose the optimal value $f_{\min} > -\infty$. A point $x^* \in K$ is a minimizer of (1.1) if and only if $f(x^*) - f_{\min} = 0$, which is equivalent to

$$(3.10) \qquad\qquad \tilde{f}(\tilde{x}^*) - f_{\min} \cdot (\tilde{x}_0^*)^d = 0$$

for the point $\tilde{x}^* := (1 + \|x^*\|^2)^{-\frac{1}{2}}(1, x^*)$. By the perspective projection φ as in (3.2), it is easy to see there exists a one-to-one correspondence between minimizers of (1.1) and feasible points of (3.10) with $\tilde{x} = (x_0, x) \in \widetilde{K}$, $x_0 > 0$. However, the system (3.10) may have feasible points of the form $\tilde{x}^* = (0, u)$. For such a case, it does not give a minimizer of (1.1).

Definition 3.5. *A point x^* is said to be a minimizer at infinity for (1.1) if $\tilde{x}^* = (0, x^*)$ satisfies (3.10) and $\tilde{x}^* \in \widetilde{K}$.*

Recall that for a polynomial p, $p^{(1)}$ denotes the homogeneous part of the highest degree for p. For convenience, denote the set

$$(3.11) \qquad K^{(1)} := \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} c_i^{(1)}(x) = 0 \ (i \in \mathcal{E}), \\ c_j^{(1)}(x) \geq 0 \ (j \in \mathcal{I}), \\ \|x\|^2 - 1 = 0 \end{array} \right\}.$$

The following is a basic property of homogenization.

Theorem 3.6. *Suppose the minimum value $f_{\min} > -\infty$. Then, we have:*
  (i) *A point x^* is a minimizer at ∞ for (1.1) if and only if $f^{(1)}(x^*) = 0$ and $x^* \in K^{(1)}$.*
  (ii) *If the minimum value $f_{\min}$ is not achievable for (1.1), then (1.1) has a minimizer at ∞.*
  (iii) *Suppose K is closed at ∞. Then the form $f^{(1)} \geq 0$ on $K^{(1)}$. In particular, if $f^{(1)} > 0$ on $K^{(1)}$, then there are no minimizers at ∞.*

Proof. (i) For $\tilde{x}^* := (0, x^*)$, note that $\tilde{c}_i(\tilde{x}^*) = c_i^{(1)}(x^*)$ for all i. So, $\tilde{x}^* \in \widetilde{K}$ if and only if $x^* \in K^{(1)}$. Since $f_{\min} > -\infty$, one can see that

$$0 = \tilde{f}(\tilde{x}^*) - f_{\min} \cdot (\tilde{x}_0^*)^d = f^{(1)}(x^*).$$

So the conclusion is true.

(ii) Since the optimal value is not achievable, the set K is unbounded and there is a sequence $\{x^{(k)}\}_{k=1}^{\infty} \subseteq K$ such that $f(x^{(k)}) \to f_{\min}$ and $\|x^{(k)}\| \to \infty$. Let $y^{(k)} := x^{(k)}/\|x^{(k)}\|$ be the normalization. Without loss of generality, one can further assume $y^{(k)} \to y^*$. Clearly, $\|y^*\| = 1$, and

$$f^{(1)}(y^*) = \lim_{k \to \infty} f^{(1)}(y^{(k)}) = \lim_{k \to \infty} f(x^{(k)})/\|x^{(k)}\|^d = 0.$$

Similarly, one can show that

$$c_i^{(1)}(y^*) = 0 \ (i \in \mathcal{E}), \quad c_i^{(1)}(y^*) \geq 0 \ (j \in \mathcal{I}).$$

So $y^* \in K^{(1)}$ and y^* is a minimizer at ∞, by the conclusion in item (i).

(iii) If $K^{(1)} = \emptyset$, the conclusion is clearly true. We consider the case that $K^{(1)} \neq \emptyset$. For each $u \in K^{(1)}$, we have $\tilde{u} := (0, u) \in \widetilde{K}^h$. Since K is closed at ∞, there exists a sequence of $\tilde{u}^{(k)} = (u_0^{(k)}, u^{(k)}) \in \widetilde{K}^c$ such that $\tilde{u}^{(k)} \to \tilde{u}$ and each $u_0^{(k)} > 0$. Note that $u^{(k)}/u_0^{(k)} \in K$ and

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

Letting $k \to \infty$, we get $f^{(1)}(u) \geq 0$ for every $u \in K^{(1)}$, hence $f^{(1)} \geq 0$ on $K^{(1)}$.

When the form $f^{(1)} > 0$ on $K^{(1)}$, there are no minimizers at ∞. This is implied by the item (i). $\qquad\square$

3.2. **Asymptotic convergence.** The asymptotic convergence of the Moment-SOS hierarchy of (3.6)-(3.7) can be shown under the following condition.

Definition 3.7. *A polynomial p is said to be positive at ∞ on K if its highest degree homogeneous part $p^{(1)} > 0$ on $K^{(1)}$.*

This condition has appeared in [6, 33] to study different properties of nonnegative polynomials on unbounded sets. If the objective f is positive at ∞, then f is coercive on K, i.e., for every value $\vartheta \in \mathbb{R}$, the sublevel set $\{x \in K : f(x) \leq \vartheta\}$ is compact. This is shown in the following lemma.

Lemma 3.8. *If f is positive at ∞ on K, then f is coercive on K.*

Proof. When K is compact, the conclusion is clearly true. Consider the case that K is unbounded. Suppose otherwise f was not coercive on K, then there exist a value ϑ and a sequence $\{u^{(k)}\}_{k=1}^{\infty} \subseteq K$ such that

$$\|u^{(k)}\| \to \infty, \quad f(u^{(k)}) \leq \vartheta \text{ for all } k.$$

Let $\tilde{u}^{(k)} := (1, u^{(k)})/\sqrt{1 + \|u_k\|^2}$, then $\tilde{u}^{(k)} \in \widetilde{K}$ and $\|\tilde{u}^{(k)}\| = 1$ for all k. Without loss of generality, we can further assume that $\tilde{u}^{(k)} \to \tilde{u}^* := (0, u^*) \in \widetilde{K}$ as $k \to \infty$. Note that $\|u^*\| = 1$, and

$$f^{(1)}(u^*) = \tilde{f}(\tilde{u}^*) = \lim_{k \to \infty} \tilde{f}(\tilde{u}^{(k)}) = \lim_{k \to \infty} \frac{f(u^{(k)})}{\sqrt{1 + \|u^{(k)}\|^2}^d} \leq 0.$$

One can similarly show that $c_i^{(1)}(u^*) = 0$ for every $i \in \mathcal{E}$ and $c_j^{(1)}(u^*) \geq 0$ for every $j \in \mathcal{I}$. Hence, we get $u^* \in K^{(1)}$ and $f^{(1)}(u^*) \leq 0$, a contradiction to the positivity of f at ∞. So f must be coercive on K. $\qquad\square$

We remark that if f is coercive on K, it may not be positive at ∞ on K. For instance, the polynomial $f = x_1^4 + x_2^2$ is not positive at ∞ for $K = \mathbb{R}^2$, but it is coercive. It is typically a hard question to check coercivity, as shown in [1]. Coercivity of polynomials is also studied in [2, 15]. The coercivity is sufficient but not necessary for the optimal value to be achievable.

When f is positive at ∞ on K, the hierarchy of relaxations (3.6) has asymptotic convergence. This is shown as follows.

Theorem 3.9. *If f is positive at ∞ on K, then the optimization (1.1) achieves the minimum value and $f_k \to f_{\min}$ as $k \to \infty$.*

Proof. Since f is positive at ∞ on K, Lemma 3.8 implies that f is coercive on K, so (1.1) must achieve its minimum value $f_{\min} > -\infty$ and it has minimizers. For every scalar $\gamma < f_{\min}$, we show that $\tilde{f}(\tilde{u}) - \gamma u_0^d > 0$ for all $\tilde{u} := (u_0, u) \in \widetilde{K}$. If $u_0 = 0$, then $u \in K^{(1)}$ and $\tilde{f}(\tilde{u}) = f^{(1)}(u) > 0$, since f is positive at ∞ on K. If $u_0 > 0$, then $u/u_0 \in K$ and

$$\tilde{f}(\tilde{u}) - \gamma(u_0)^d = u_0^d(f(u/u_0) - \gamma) > 0.$$

Note that $\mathrm{Ideal}[\tilde{c}_\mathcal{E}] + \mathrm{QM}[\tilde{c}_\mathcal{I}]$ is archimedean, due to the sphere constraint $x_0^2 + x^T x = 1$. By Theorem 2.1, we have $\tilde{f}(\tilde{x}) - \gamma x_0^d \in \mathrm{Ideal}[\tilde{c}_\mathcal{E}] + \mathrm{QM}[\tilde{c}_\mathcal{I}]$. Thus, when k is sufficiently large, we get $\tilde{f}(\tilde{x}) - \gamma x_0^d \in \mathrm{Ideal}[\tilde{c}_\mathcal{E}]_{2k} + \mathrm{QM}[\tilde{c}_\mathcal{I}]_{2k}$. This is true for every $\gamma < f_{\min}$. On the another hand, if γ is feasible for (3.6), we must have $\tilde{f}(\tilde{x}) - \gamma x_0^d \geq 0$ on $\widetilde{K}$, which implies that $f - \gamma \geq 0$ on K and hence $\gamma \leq f_{\min}$. This shows that $f_k \to f_{\min}$ as $k \to \infty$. $\qquad\square$

The following is an example for the hierarchy of (3.6)-(3.7).

Example 3.10. Consider the optimization problem

$$\begin{cases} \min & x_1 + x_2 \\ s.t. & x_1^3 + x_2 + 1 \geq 0,\ x_2^3 - x_1 + 1 \geq 0. \end{cases}$$

The feasible set K is unbounded. One can check that the minimum value and the unique minimizer are respectively

$$f_{\min} = -1 - \frac{2\sqrt{3}}{9}, \quad x^* = \left(-\frac{\sqrt{3}}{3}, -1 + \frac{\sqrt{3}}{9}\right).$$

Note that $f^{(1)} = x_1 + x_2$ and

$$K^{(1)} = \{x_1^3 \geq 0, x_2^3 \geq 0, x_1^2 + x_2^2 = 1\}.$$

The form $f^{(1)}$ is positive on $K^{(1)}$. The hierarchy of (3.6)-(3.7) has the asymptotic convergence. Interestingly, it also has finite convergence. This can be implied by Theorem 4.3.

4. Optimality conditions and finite convergence

Optimality conditions are closely related to finite convergence of the classical Moment-SOS hierarchy in [21]. Under the archimedeanness for constraining polynomials, the Moment-SOS hierarchy has finite convergence when the linear independence constraint qualification, strict complementarity and second order sufficient conditions hold at every minimizer. This is shown in [41]. When the feasible set K is unbounded, the above conclusion may not hold. However, we can prove similar conclusions for the homogenized hierarchy of relaxations (3.6)-(3.7).

Recall that f_k is the optimal value of the relaxation (3.6) for the relaxation order k. The hierarchy of (3.6)-(3.7) is said to have finite convergence or be tight if $f_k = f_{\min}$ for all k big enough. To guarantee the finite convergence, we assume that K is closed at ∞, $\mathrm{Ideal}[\tilde{c}_\mathcal{E}]$ is real radical, the LICQC, SCC and SOSC hold at every minimizer, including the one at infinity. However, the set K is not assumed to be bounded and we do not assume the optimal value of (1.1) is achievable.

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

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

Since there is the sphere constraint, the above optimization must have minimizers.

Lemma 4.1. *If K is closed at ∞ and $f_{\min} > -\infty$, the minimum value of (4.1) is 0.*

Proof. Since K is closed at ∞, we know $F(\tilde{x}) \geq 0$ on $\widetilde{K}$, by Lemma 3.2. If $f_{\min}$ is achievable at a minimizer x^* of (1.1), then $\tilde{x}^* := (1 + \|x^*\|^2)^{-\frac{1}{2}}(1, x^*) \in \widetilde{K}$ and $F(\tilde{x}^*) = 0$. If $f_{\min}$ is not achievable, then (1.1) has a minimizer at ∞, say u^*, by Theorem 3.6. For the point $\tilde{u}^* := (0, u^*)$, we have $F(\tilde{u}^*) = 0$ and $\tilde{u}^* \in \widetilde{K}$. So the minimum value of (4.1) is 0. $\qquad\square$

Since the minimum value is 0, a point $\tilde{u}$ is a minimizer of (4.1) if and only if $F(\tilde{u}) = 0$ and $\tilde{u} \in \widetilde{K}$. Suppose $\tilde{u} = (u_0, u)$ is a minimizer of (4.1). If $u_0 > 0$, then u/u_0 is a minimizer of (1.1). If $u_0 = 0$, then u is a minimizer at ∞ for (1.1). For the case $u_0 > 0$, we call $\tilde{u}$ a regular minimizer of (4.1). For the case $u_0 = 0$, we call $\tilde{u}$ a minimizer at infinity of (4.1). Recall the optimality conditions LICQC, SCC and SOSC as in Subsection 2.1. Note that $\widetilde{K}$ is always compact. The following is the finite convergence result based on optimality conditions for the homogenized optimization problem (4.1), which follows from [41].

Lemma 4.2. *Assume K is closed at ∞ and $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical. If the LICQC, SCC and SOSC hold at every minimizer of (4.1), then the hierarchy of relaxations (3.6)-(3.7) is tight, i.e., there exists $k_0 \in \mathbb{N}$ such that*

$$
f_k = f_k' = f_{\min} \quad \text{for all } k \geq k_0.
$$

Proof. Note that $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}] + \mathrm{QM}[\tilde{c}_{\mathcal{I}}]$ is archimedean. By Theorem 1.1 of [41], there exists $\sigma \in \mathrm{QM}[\tilde{c}_{\mathcal{I}}]$ such that

$$
\tilde{f} - f_{\min} x_0^d \equiv \sigma \mod \mathrm{Ideal}[V_{\mathbb{R}}(\tilde{c}_{\mathcal{E}})],
$$

Since the ideal $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical, there exists $\phi \in \mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ such that

$$
\tilde{f} - f_{\min} x_0^d = \phi + \sigma.
$$

This implies that $f_k = f_k' = f_{\min}$ for all k big enough. $\qquad\square$

In Lemma 4.2, optimality conditions are stated for the homogenized optimization problem (4.1). Interestingly, the optimality conditions at regular minimizers for (4.1) are equivalent to those for (1.1). The equivalence will be shown in the following subsections (see Theorem 4.4). Therefore, the finite convergence result can be stated under the optimality conditions of (1.1). For a minimizer at infinity x^*, we say the LICQC, SCC and SOSC hold at x^* if they hold for (4.1) at $(0, x^*)$. The following is the main result about the finite convergence.

Theorem 4.3. *Assume K is closed at ∞ and the ideal $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), including the one at infinity, then the hierarchy of relaxations (3.6)-(3.7) is tight, i.e., $f_k = f_k' = f_{min}$ for all k big enough.*

Proof. Since K is closed at ∞, we know $\tilde{f} - f_{\min}x_0^d \geq 0$ on $\widetilde{K}$, by Lemma 3.2. Let $\tilde{x}^* = (x_0^*, x^*) \in \widetilde{K}$ be a minimizer of (4.1). If $x_0^* > 0$, then $\frac{x^*}{x_0^*}$ is a minimizer of (1.1). Since the LICQC, SCC and SOSC hold at $\frac{x^*}{x_0^*}$, these optimality conditions also hold at $\tilde{x}^*$ for (4.1), by Theorem 4.4. If $x_0^* = 0$, then x^* is a minimizer at infinity. By the assumptions, these optimality conditions also hold at $\tilde{x}^*$. Hence, the LICQC, SCC and SOSC hold at every minimizer of (4.1). By Lemma 4.2, we have $f_k = f_k' = f_{\min}$ for all k big enough. $\qquad\square$

In Theorem 4.3, the minimum value $f_{\min}$ is not assumed to be achievable for (1.1). When there are no equality constraints, the ideal $\text{Ideal}[\tilde{c}_{\mathcal{E}}] = \text{Ideal}[\|\tilde{x}\|^2 - 1]$ is real radical, regardless of inequality constraints.

For an exposition for Theorem 4.3, we consider Example 3.10. At the unique minimizer $x^* = (-\frac{\sqrt{3}}{3}, -1 + \frac{\sqrt{3}}{9})$, the first constraint $x_1^3 + x_2 + 1 \geq 0$ is active while the second one is not. Hence, we have that the Lagrange multipliers $\lambda_1 = 1$, $\lambda_2 = 0$. The Hessian of the Lagrange function at x^* is $\begin{bmatrix} 2\sqrt{3} & 0 \\ 0 & 0 \end{bmatrix}$. One can see that the LICQC, SCC and SOSC all hold at x^*. One can check that there are no minimizers at infinity and $\text{Ideal}[\tilde{c}_{\mathcal{E}}] = \text{Ideal}[\|\tilde{x}\|^2 - 1]$ is real radical. By Theorem 4.3, the hierarchy of (3.6)-(3.7) is tight for this optimization problem. In fact, we have $f_k = f_{\min}$ for all $k \geq 3$.

In the following, we investigate when the optimality conditions hold for (4.1).

4.1. **Optimality conditions for regular minimizers.** An important property of homogenization is that it preserves optimality conditions for regular minimizers.

Theorem 4.4. *Suppose K is closed at ∞. Then, a point $x^* \in K$ is a minimizer of (1.1) if and only if $\tilde{x}^* := (1, x^*)/\sqrt{1 + \|x^*\|^2}$ is a minimizer of (4.1). Moreover, we have:*

(i) *The LICQC holds for (1.1) at x^* if and only if it holds for (4.1) at $\tilde{x}^*$.*
(ii) *Suppose the LICQC holds for (1.1) at x^*. Then, the SCC holds for (1.1) at x^* if and only if it holds for (4.1) at $\tilde{x}^*$.*
(iii) *Suppose the LICQC holds for (1.1) at x^*. Then, the SOSC holds for (1.1) at x^* if and only if it holds for (4.1) at $\tilde{x}^*$.*

Proof. Without loss of generality, one can assume $f_{\min} = 0$ up to shifting a constant in f. By Lemma 3.2, when K is closed at ∞, $f(x) \geq 0$ on K if and only if $\tilde{f}(\tilde{x}) \geq 0$ on $\widetilde{K}$, which is the feasible set of (4.1). Note that x^* is a minimizer of (1.1) if and only if $f(x^*) = 0$. Since $\tilde{f}(\tilde{x}^*) = (1 + \|x^*\|^2)^{-d/2}f(x^*)$, $\tilde{f}(\tilde{x}^*) = 0$ if and only if $f(x^*) = 0$. Hence, x^* is a minimizer of (1.1) if and only if $\tilde{x}^*$ is a minimizer of (4.1). Since the constraint $x_0 \geq 0$ is not active at $\tilde{x}^*$, the label set of active constraints is

$$J(x^*) := \{i \in \mathcal{E} \cup \mathcal{I} : c_i(x^*) = 0\}.$$

By the Euler's identity for homogeneous polynomials, we have that for all $\tilde{x}$

$$x^{\mathrm{T}}\nabla_x \tilde{c}_i(\tilde{x}) + x_0 \nabla_{x_0}\tilde{c}_i(\tilde{x}) = \deg(c_i) \cdot \tilde{c}_i(\tilde{x}).$$

The above implies that

$$(4.2) \qquad (x^*)^{\mathrm{T}}\nabla_x \tilde{c}_i(\tilde{x}^*) + \nabla_{x_0}\tilde{c}_i(\tilde{x}^*) = 0$$

for all $i \in J(x^*)$. Also note that

$$(4.3) \qquad \nabla_x \tilde{c}_i(\tilde{x}^*) = (\sqrt{1 + \|x^*\|^2})^{1-\deg(c_i)}\nabla c_i(x^*).$$

(i) "$\Leftarrow$": Suppose the LICQC holds for (4.1) at $\tilde{x}^*$, i.e., the gradients

$$\begin{bmatrix} \nabla_{x_0}\tilde{c}_i(\tilde{x}^*) \\ \nabla_x \tilde{c}_i(\tilde{x}^*) \end{bmatrix} (i \in J(x^*)), \quad \begin{bmatrix} 1 \\ x^* \end{bmatrix}$$

are linearly independent. Then (4.2) implies that $\nabla_x \tilde{c}_i(\tilde{x}^*)$ $(i \in J(x^*))$ are linearly independent. By (4.3), we know the gradients $\nabla c_i(x^*)$ $(i \in J(x^*))$ are linearly independent, i.e., the LICQC holds for (1.1) at x^*.

"$\Rightarrow$": Assume the LICQC holds for (1.1) at x^*, i.e., the gradients $\nabla c_i(x^*)$ $(i \in J(x^*))$ are linearly independent. Suppose there are scalars μ_i, μ_0 such that

$$(4.4) \qquad \sum_{i \in J(x^*)} \mu_i \begin{bmatrix} \nabla_{x_0}\tilde{c}_i(\tilde{x}^*) \\ \nabla_x \tilde{c}_i(\tilde{x}^*) \end{bmatrix} + \mu_0 \begin{bmatrix} 1 \\ x^* \end{bmatrix} = 0.$$

Note that $\tilde{c}_i(\tilde{x}^*) = 0$ for all $i \in J(x^*)$. Premultiplying $(\tilde{x}^*)^{\mathrm{T}}$ in (4.4) results in

$$0 = \sum_{i \in J(x^*)} \mu_i \deg(c_i) \cdot \tilde{c}_i(\tilde{x}^*) + \mu_0 \sqrt{1 + \|x^*\|^2} = \mu_0 \sqrt{1 + \|x^*\|^2}.$$

Hence, we get $\mu_0 = 0$. By (4.3), the linear independence of $\nabla c_i(x^*)$ $(i \in J(x^*))$ implies that $\mu_i = 0$ for all $i \in J(x^*)$. So, the LICQC holds for (4.1) at $\tilde{x}^*$.

(ii) By the item (i), the LICQC holds at $\tilde{x}^*$ for (4.1). Since $\tilde{x}^*$ is a minimizer of (4.1), there exist Lagrange multipliers λ_i $(i \in J(x^*))$ and λ_0 such that

$$(4.5) \qquad \begin{bmatrix} \nabla_{x_0} F(\tilde{x}^*) \\ \nabla_x F(\tilde{x}^*) \end{bmatrix} = \sum_{i \in J(x^*)} \lambda_i \begin{bmatrix} \nabla_{x_0}\tilde{c}_i(\tilde{x}^*) \\ \nabla_x \tilde{c}_i(\tilde{x}^*) \end{bmatrix} + \lambda_0 \begin{bmatrix} 1 \\ x^* \end{bmatrix}.$$

By the Euler's identity for homogeneous polynomials, we have

$$(\tilde{x}^*)^{\mathrm{T}} \nabla_{\tilde{x}} F(\tilde{x}^*) = 0, \quad (\tilde{x}^*)^{\mathrm{T}} \nabla_{\tilde{x}} \tilde{c}_i(\tilde{x}^*) = 0 \ (i \in J(x^*)).$$

The above and (4.5) imply that $\lambda_0 = 0$. Let $\hat{\lambda}_i = \lambda_i(\sqrt{1 + \|x^*\|^2})^{d - \deg(c_i)}$, then

$$(4.6) \qquad \nabla f(x^*) = \sum_{i \in J(x^*)} \hat{\lambda}_i \nabla c_i(x^*).$$

When the LICQC holds, the Lagrange multipliers are unique for both (1.1) and (4.1).

"$\Rightarrow$": Suppose the SCC holds for (1.1) at x^*. If otherwise the SCC fails to hold for (4.1) at $\tilde{x}^*$, say, $\lambda_{j_0} = 0$ for some $j_0 \in J(x^*) \cap \mathcal{I}$, then as in the above we can get

$$\nabla f(x^*) = \sum_{i \in J(x^*) \backslash \{j_0\}} \hat{\lambda}_i \nabla c_i(x^*).$$

Since the Lagrange multipliers are unique, the above implies that the SCC fails for (1.1) at x^*, which is a contradiction. Therefore, the SCC holds for (4.1) at $\tilde{x}^*$.

"$\Leftarrow$": Suppose the SCC holds for (4.1) at $\tilde{x}^*$, then $\lambda_i > 0$ for all $i \in J(x^*) \cap \mathcal{I}$ in (4.5). So, $\hat{\lambda}_i > 0$ for each $i \in J(x^*) \cap \mathcal{I}$ in (4.6). Since the Lagrange multipliers are unique, this means that the SCC holds for (1.1) at x^*.

(iii) Note that the LICQC holds for both (1.1) and (4.1).

"⇒": Suppose the SOSC holds (1.1) at x^*. Consider $\tilde{y} := (y_0, y)$ in the tangent space of active constraints of (4.1) at $\tilde{x}^*$, i.e.,

$$(4.7) \qquad \boxed{\begin{array}{c} y_0 \nabla_{x_0} \tilde{c}_i(\tilde{x}^*) + y^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = 0 \; (i \in J(x^*)), \\ y_0 + y^{\mathrm{T}} x^* = 0 \end{array}}.$$

The equation (4.2) implies that $\nabla_{x_0} \tilde{c}_i(\tilde{x}^*) = -(x^*)^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*)$ for every $i \in J(x^*)$. So (4.7) is equivalent to

$$(4.8) \qquad \boxed{\begin{array}{rl} (y - y_0 x^*)^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) &= 0 \; (i \in J(x^*)), \\ y_0 + y^{\mathrm{T}} x^* &= 0 \end{array}}.$$

Let $\lambda_i, \hat{\lambda}_i$ be Lagrange multipliers as in (4.5)-(4.6) for the proof of item (ii). Note that $\lambda_0 = 0$ and we shift f as $f_{\min} = 0$, so we have

$$(4.9) \qquad (y - y_0 x^*)^{\mathrm{T}} \nabla_x \tilde{f}(\tilde{x}^*) = \sum_{i \in J(x^*)} \lambda_i (y - y_0 x^*)^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = 0.$$

For the Lagrange function

$$(4.10) \qquad L(\tilde{x}) := \tilde{f}(\tilde{x}) - \sum_{i \in J(x^*)} \lambda_i \tilde{c}_i(\tilde{x}),$$

its Hessian at $\tilde{x}^*$ is

$$\nabla^2_{\tilde{x}\tilde{x}} L(\tilde{x}^*) = \left[\begin{array}{cc} \nabla^2_{x_0, x_0} L(\tilde{x}^*) & \nabla^2_{x, x_0} L(\tilde{x}^*) \\ \nabla^2_{x, x_0} L(\tilde{x}^*)^{\mathrm{T}} & \nabla^2_{x, x} L(\tilde{x}^*) \end{array} \right].$$

We express $\nabla^2_{\tilde{x}\tilde{x}} L(\tilde{x}^*)$ in terms of $\frac{\partial^2 L(\tilde{x}^*)}{\partial x_i x_j}$ $(1 \le i \le n, \ 1 \le j \le n)$. Let

$$\eta_0 := \frac{1}{\sqrt{1 + \|x^*\|^2}}, \quad \eta := \frac{x^*}{\sqrt{1 + \|x^*\|^2}}.$$

Similar to (4.2), it holds that

$$\boxed{\begin{array}{c} \eta_0 \nabla_{x_0} \tilde{f}(\tilde{x}^*) + \eta^{\mathrm{T}} \nabla_x \tilde{f}(\tilde{x}^*) = 0, \\ \eta_0 \nabla_{x_0} \tilde{c}_i(\tilde{x}^*) + \eta^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = 0 \; (i \in J(x^*)) \end{array}}.$$

Again, by Euler's identity, we similarly have

$$(4.11) \qquad (d-1) \nabla_{x_j} \tilde{f}(\tilde{x}^*) = \left[\nabla^2_{x_j, x_0} \tilde{f}(\tilde{x}^*) \quad \nabla^2_{x_j, x} \tilde{f}(\tilde{x}^*) \right] \left[\begin{array}{c} \eta_0 \\ \eta \end{array} \right],$$

$$(\deg(c_i) - 1) \nabla_{x_j} \tilde{c}_i(\tilde{x}^*) = \left[\nabla^2_{x_j, x_0} \tilde{c}_i(\tilde{x}^*) \quad \nabla^2_{x_j, x} \tilde{c}_i(\tilde{x}^*) \right] \left[\begin{array}{c} \eta_0 \\ \eta \end{array} \right],$$

for all $i \in J(x^*)$ and $j = 0, \ldots, n$. For convenience, denote

$$v^* := d \nabla_x \tilde{f}(\tilde{x}^*) - \sum_{i \in J(x^*)} \deg(c_i) \lambda_i \nabla_x \tilde{c}_i(\tilde{x}^*),$$

$$\gamma := d \nabla_{x_0} \tilde{f}(\tilde{x}^*) - \sum_{i \in J(x^*)} \deg(c_i) \lambda_i \nabla_{x_0} \tilde{c}_i(\tilde{x}^*).$$

With (4.5) and (4.11), we have

$$(4.12)$$
$$\eta_0 \nabla^2_{x,x_0} L(\tilde{x}^*) = (d-1)\nabla_x \tilde{f}(\tilde{x}^*) - \sum_{i \in J(x^*)} (\deg(c_i) - 1)\lambda_i \nabla_x \tilde{c}_i(\tilde{x}^*) - \nabla^2_{x,x} L(\tilde{x}^*)\eta$$
$$= v^* - \nabla^2_{x,x} L(\tilde{x}^*)\eta,$$

$$(4.13) \qquad \eta_0 \gamma = -d\nabla_x \tilde{f}(\tilde{x}^*)^{\mathrm{T}}\eta + \sum_{i \in J(x^*)} \deg(c_i)\lambda_i \nabla_x \tilde{c}_i(\tilde{x}^*)^{\mathrm{T}}\eta = -\eta^{\mathrm{T}} v^*.$$

Hence, the following holds

$$(4.14) \quad (\eta_0)^2 \nabla^2_{x_0,x_0} L(\tilde{x}^*) = \eta_0 \gamma - \eta_0 \nabla^2_{x_0,x} L(\tilde{x}^*)\eta = -2\eta^{\mathrm{T}} v^* + \eta^{\mathrm{T}} \nabla^2_{x,x} L(\tilde{x}^*)\eta.$$

Consider the new Lagrange function

$$(4.15) \qquad \hat{L}(x) := f(x) - \sum_{i \in J(x^*)} \hat{\lambda}_i c_i(x).$$

Since the SOSC holds for (1.1) at x^*, we have

$$y^{\mathrm{T}} \nabla^2 \hat{L}(x^*) y > 0, \quad \text{for each } y \neq 0, \ y^{\mathrm{T}} \nabla c_i(x^*) = 0 \ (i \in J(x^*)).$$

The above is equivalent to

$$y^{\mathrm{T}} \nabla^2_{xx} L(\tilde{x}^*) y > 0, \quad \text{for each } y \neq 0, \ y^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = 0 \ (i \in J(x^*)).$$

Using (4.12)-(4.14), one can verify that for $\tilde{y} = (y_0, y)$ satisfying (4.8),

$$\tilde{y}^{\mathrm{T}} \nabla^2_{\tilde{x}\tilde{x}} L(\tilde{x}^*) \tilde{y} = y^{\mathrm{T}} \nabla^2_{xx} L(\tilde{x}^*) y + 2y_0 y^{\mathrm{T}} \nabla^2_{x,x_0} L(\tilde{x}^*) + y_0^2 \nabla^2_{x_0,x_0} L(\tilde{x}^*)$$
$$(4.16) \qquad = (y - y_0 x^*)^{\mathrm{T}} \nabla^2_{x,x} L(\tilde{x}^*)(y - y_0 x^*) + \frac{2y_0(y - y_0 x^*)^{\mathrm{T}} v^*}{\eta_0}$$
$$= (y - y_0 x^*)^{\mathrm{T}} \nabla^2_{x,x} L(\tilde{x}^*)(y - y_0 x^*) \geq 0.$$

In the third equality above, $(y - y_0 x^*)^{\mathrm{T}} v^* = 0$ follows from (4.8)-(4.9). Moreover, if $\tilde{y}^{\mathrm{T}} \nabla^2_{\tilde{x}\tilde{x}} L(\tilde{x}^*) \tilde{y} = 0$, then $y - y_0 x^* = 0$, by the SOSC for (1.1) at x^*. Since $y_0 + y^{\mathrm{T}} x^* = 0$ as in (4.8), we get

$$0 = (y - y_0 x^*)^{\mathrm{T}} x^* = -y_0(1 + \|x^*\|^2).$$

So, $y_0 = 0$ and hence $y = 0$, i.e., $\tilde{y} = 0$. This shows that the SOSC holds for (4.1) at $\tilde{x}^*$.

"$\Leftarrow$": When the SOSC holds for (4.1) at $\tilde{x}^*$, we show that it also holds for (1.1) at x^*. Suppose otherwise the SOSC fails to hold for (1.1) at x^*, then there exists $u \neq 0$ such that

$$u^{\mathrm{T}} \nabla^2 \hat{L}(x^*) u = 0, \ u^{\mathrm{T}} \nabla c_i(x^*) = 0 \ (i \in J(x^*)).$$

Let $\tilde{\ell} := (\ell_0, \ell)$, where $\ell_0 = -u^{\mathrm{T}} x^*/(1 + \|x^*\|^2)$ and $\ell = u + \ell_0 x^*$. For the case $\ell_0 = 0$, $\ell = u$ is nonzero. So, $\tilde{\ell}$ is a nonzero vector. Note that

$$(\ell - \ell_0 x^*)^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = u^{\mathrm{T}} \nabla_x \tilde{c}_i(\tilde{x}^*) = 0 \ (i \in J(x^*)),$$
$$\ell^{\mathrm{T}} x^* + \ell_0 = u^{\mathrm{T}} x^* + \ell_0(1 + \|x^*\|^2) = 0.$$

By equation (4.8), we know $\tilde{\ell}$ lies in the tangent space of active constraints of (4.1) at $\tilde{x}^*$. Moreover, similar as in (4.16), the following holds

$$\tilde{\ell}^{\mathrm{T}} \nabla_{\tilde{x}\tilde{x}}^2 L(\tilde{x}^*) \tilde{\ell} = (\ell - \ell_0 x^*)^{\mathrm{T}} \nabla_{xx}^2 L(\tilde{x}^*)(\ell - \ell_0 x^*)$$
$$= u^{\mathrm{T}} \nabla_{\tilde{x}\tilde{x}}^2 L(\tilde{x}^*) u = u^{\mathrm{T}} \nabla^2 \hat{L}(x^*) u = 0.$$

This is a contradiction to that the SOSC holds at $\tilde{x}^*$ for (4.1). So the SOSC must hold for (1.1) at x^*. $\qquad\square$

The LICQC, SCC and SOSC all hold at every local minimizer of (1.1) for generic polynomial optimization problems. This is a major conclusion in [41]. By Theorem 4.4, we know these optimality conditions all hold at every regular minimizer of (4.1) for generic polynomials.

4.2. **Optimality conditions for minimizers at infinity.** We consider minimizers at infinity. For a polynomial p, recall that $p^{(1)}$ denotes the homogeneous part of the highest degree for p. Suppose K is closed at ∞ and $\tilde{x}^* = (0, x^*)$ is a minimizer at ∞ for (4.1). By Theorem 3.6 and Lemma 4.1, we know $f^{(1)}(x^*) = 0$ and x^* is a minimizer for

(4.17)
$$\begin{cases} \min & f^{(1)}(x) \\ s.t. & c_i^{(1)}(x) = 0 \ (i \in \mathcal{E}), \\ & c_j^{(1)}(x) \geq 0 \ (j \in \mathcal{I}), \\ & \|x\|^2 - 1 = 0. \end{cases}$$

Denote the active label set

(4.18)
$$J_1(x^*) := \left\{ i \in \mathcal{E} \cup \mathcal{I} \mid c_i^{(1)}(x^*) = 0 \right\}.$$

By the Fritz-John condition (see [4]), there exist scalars μ_0, μ_i, $\bar{\mu}$ such that

$$\mu_0 \nabla f^{(1)}(x^*) = \sum_{i \in J_1(x^*)} \mu_i \nabla c_i^{(1)}(x^*) + \bar{\mu} x^*,$$

while μ_0, μ_i, $\bar{\mu}$ are not all zero. Note that $\|x^*\| = 1$. Since $f^{(1)}(x^*) = 0$ and $c_i^{(1)}(x^*) = 0$ for each $i \in J_1(x^*)$, premultiplying $(x^*)^{\mathrm{T}}$ in the above results in $\bar{\mu} = 0$, by the Euler's identity. So the above is the same as

(4.19)
$$\mu_0 \nabla f^{(1)}(x^*) = \sum_{i \in J(x^*)} \mu_i \nabla c_i^{(1)}(x^*).$$

First, we show that there are no minimizers at infinity for (1.1) if the polynomials are generic. For this, we review some basic theory for resultants and discriminants. We refer to [9, 37, 57]. Let $p_1, \ldots, p_n$ be forms in $x := (x_1, \ldots, x_n)$. The resultant $\mathrm{Res}(p_1, \ldots, p_n)$ is a polynomial in the coefficients of $p_1, \ldots, p_n$ such that

$$\mathrm{Res}(p_1, \ldots, p_n) = 0 \Leftrightarrow \exists \, 0 \neq u \in \mathbb{C}^n, \ p_1(u) = \cdots = p_n(u) = 0.$$

For $m \leq n$, the discriminant $\Delta(p_1, \ldots, p_m)$ is a polynomial in the coefficients of $p_1, \ldots, p_m$ such that $\Delta(p_1, \ldots, p_m) = 0$ if and only if there exists $0 \neq u \in \mathbb{C}^n$ satisfying

$$p_1(u) = \cdots = p_m(u) = 0, \ \mathrm{rank} \left[\nabla p_1(u) \ \cdots \ \nabla p_m(u) \right] < m.$$

Both $\mathrm{Res}(p_1, \ldots, p_n)$ and $\Delta(p_1, \ldots, p_m)$ are homogeneous polynomials in the coefficients of p_i. The following is about nonexistence of minimizers at infinity for generic polynomial optimization problems.

Theorem 4.5. *Suppose K is closed at ∞ and $\mathcal{E} = \{1, \ldots, m_1\}$, $\mathcal{I} = \{m_1 + 1, \ldots, m_2\}$, $m_1 \leq n - 1$. If the polynomials $f \in \mathbb{R}[x]_{d_0}$ and $c_i \in \mathbb{R}[x]_{d_i}$ $(i \in \mathcal{E} \cup \mathcal{I})$ satisfy the conditions:*

   (i) *For all $m_1 + 1 \leq j_1 < \cdots < j_{n-m_1} \leq m_2$*

$$\mathrm{Res}(c_1^{(1)}, \ldots, c_{m_1}^{(1)}, c_{j_1}^{(1)}, \ldots, c_{j_{n-m_1}}^{(1)}) \neq 0;$$

   (ii) *For all $m_1 + 1 \leq j_1 < \cdots < j_r \leq m_2$ with $0 \leq r \leq n - m_1 - 1$*

$$\Delta(f^{(1)}, c_1^{(1)}, \ldots, c_{m_1}^{(1)}, c_{j_1}^{(1)}, \ldots, c_{j_r}^{(1)}) \neq 0,$$

then (1.1) has no minimizers at ∞.

Proof. Since K is closed at ∞, the optimal value of (4.17) is 0 if there is a minimizer at infinity x^*. Let $j_1, \ldots, j_r$ be the labels of active inequality constraints for (4.17) at x^*. The item (a) implies that $r < n - m_1$. This is because if otherwise $r \geq n - m_1$, then

$$\mathrm{Res}(c_1^{(1)}, \ldots, c_{m_1}^{(1)}, c_{j_1}^{(1)}, \ldots, c_{j_{n-m_1}}^{(1)}) = 0.$$

The Fritz-John condition (4.19) implies that

$$\Delta(f^{(1)}, c_1^{(1)}, \ldots, c_{m_1}^{(1)}, c_{j_1}^{(1)}, \ldots, c_{j_r}^{(1)}) = 0.$$

The conditions in (i)-(ii) deny existence of minimizers at infinity.   □

For special polynomial optimization problems, there may exist minimizers at infinity. For instance, this is the case if the optimal value is not achievable for (1.1) (see Theorem 3.6). In the following, we study optimality conditions for minimizers at infinity. Suppose $\tilde{x}^* \coloneqq (0, x^*)$ is a minimizer at infinity for (4.1). Recall that $d = \deg(f)$. The KKT equation for (4.1) at $\tilde{x}^*$ is in the form

$$(4.20) \quad \begin{bmatrix} f^{(2)}(x^*) - d \cdot f_{\min} \cdot 0^{d-1} \\ \nabla f^{(1)}(x^*) \end{bmatrix} = \sum_{i \in J_1(x^*)} \lambda_i \begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} + \lambda_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \bar{\lambda} \begin{bmatrix} 0 \\ x^* \end{bmatrix}.$$

Theorem 4.6. *Suppose K is closed at ∞ and $\tilde{x}^* \coloneqq (0, x^*)$ is a minimizer at infinity for (4.1). Let $J_1(x^*)$ be as in (4.18). Then, we have:*

   (i) *The LICQC holds for (4.1) at $\tilde{x}^*$ if and only if the gradients*

$$\nabla c_i^{(1)}(x^*) \ (i \in J_1(x^*))$$

   are linearly independent.

   (ii) *Suppose the LICQC holds for (4.1) at $\tilde{x}^*$. Then, in (4.20), we have*

$$(4.21) \quad \lambda_0 = f^{(2)}(x^*) - d \cdot f_{\min} \cdot 0^{d-1} - \sum_{i \in J_1(x^*)} \lambda_i c_i^{(2)}(x^*), \quad \bar{\lambda} = 0.$$

   Moreover, the SCC holds for (4.1) at $\tilde{x}^*$ if and only if

$$(4.22) \quad \lambda_0 > 0, \ \lambda_i > 0 \ (i \in J_1(x^*) \cap \mathcal{I}).$$

(iii) *Suppose the LICQC holds for (4.1) at $\tilde{x}^*$. Let λ_i be Lagrange multipliers as in (4.20). Then, the SOSC holds for (4.1) at $\tilde{x}^*$ if and only if for every nonzero y satisfying*

(4.23)
$$\boxed{\begin{array}{c} y^{\mathrm{T}}\nabla c_i^{(1)}(x^*) = 0 \ (i \in J_1(x^*)), \\ y^{\mathrm{T}}x^* = 0 \end{array}},$$

we have $y^{\mathrm{T}}\nabla^2 L_1(x^*)y > 0$, where

$$L_1(x) = f^{(1)}(x) - \sum_{i \in J_1(x^*)} \lambda_i c_i^{(1)}(x).$$

Proof. (i) The constraint $x_0 \geq 0$ is active (4.1) at $\tilde{x}^*$. The LICQC at $\tilde{x}^*$ for (4.1) requires the linear independence of the gradients

$$\begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} \ (i \in J_1(x^*)), \ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \begin{bmatrix} 0 \\ x^* \end{bmatrix}.$$

"$\Rightarrow$": Suppose the LICQC holds at $\tilde{x}^*$ for (4.1). If there is a linear combination such that $\sum_{i \in J_1(x^*)} \mu_i \nabla c_i^{(1)}(x^*) = 0$, then

$$\sum_{i \in J_1(x^*)} \mu_i \begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} + \mu_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 0$$

for $\mu_0 = -\sum_{i \in J_1(x^*)} \mu_i c_i^{(2)}(x^*)$. The LICQC holds at $\tilde{x}^*$ for (4.1) implies that all $\mu_i = 0$. Hence, the gradients $\nabla c_i^{(1)}(x^*) \ (i \in J_1(x^*))$ are linearly independent.

"$\Leftarrow$": Suppose the gradients $\nabla c_i^{(1)}(x^*) \ (i \in J_1(x^*))$ are linearly independent. Consider a linear combination such that

(4.24)
$$\sum_{i \in J_1(x^*)} \mu_i \begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} + \mu_0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \bar{\mu} \begin{bmatrix} 0 \\ x^* \end{bmatrix} = 0.$$

Since $c_i^{(1)}(x^*) = 0$ for each $i \in J_1(x^*)$, we have $(x^*)^{\mathrm{T}}\nabla c_i^{(1)}(x^*) = 0$. Premultiplying $(\tilde{x}^*)^{\mathrm{T}}$ in the above results in $\bar{\mu}\|x^*\|^2 = 0$, so $\bar{\mu} = 0$. The linear independence of $\nabla c_i^{(1)}(x^*) \ (i \in J(x^*))$ implies that $\mu_i = 0$ for all $i \in J_1(x^*)$. Finally, we get $\mu_0 = 0$. So, the LICQC holds at $\tilde{x}^*$ for (4.1).

(ii) Note that $\tilde{x}^* = (0, x^*)$ is a minimizer of (4.1). Since the LICQC holds at $\tilde{x}^*$, there exist scalars $\bar{\lambda}, \lambda_0, \lambda_i \ (i \in J_1(x^*))$ satisfying (4.20). Premultiplying $(\tilde{x}^*)^{\mathrm{T}}$ in (4.20), by the Euler's identity, we can express $\lambda_0, \bar{\lambda} = 0$ as in (4.21). The SCC for (4.1) at $\tilde{x}^*$ is equivalent to (4.22).

(iii) Since the LICQC holds, there exist Lagrange multipliers λ_0, λ_i as in (4.20). The Lagrange function for (4.1) is

$$L(\tilde{x}) := F(\tilde{x}) - \sum_{i \in J_1(x^*)} \lambda_i \tilde{c}_i(x) - \lambda_0 x_0.$$

Its Hessian expression at $\tilde{x}^*$ is

$$\nabla^2_{\tilde{x}, \tilde{x}} L(\tilde{x}^*) = \begin{bmatrix} L_3(x^*) & \nabla L_2(x^*)^{\mathrm{T}} \\ \nabla L_2(x^*) & \nabla^2 L_1(x^*) \end{bmatrix},$$

where

$$L_2(x) = f^{(2)}(x) - \sum_{i \in J_1(x^*)} \lambda_i c_i^{(2)}(x),$$

$$L_3(x) = 2f^{(3)}(x) - d(d-1) \cdot f_{\min} \cdot 0^{d-2} - \sum_{i \in J_1(x^*)} 2\lambda_i c_i^{(3)}(x).$$

Consider $\tilde{y} := (y_0, y)$ in the tangent space of active constraints of (4.1) at $\tilde{x}^*$, i.e.,

(4.25)
$$\boxed{\begin{array}{c} y_0 c_i^{(2)}(x^*) + y^{\mathrm{T}} \nabla c_i^{(1)}(x^*) = 0 \ (i \in J(x^*)), \\ y^{\mathrm{T}} x^* = 0, \ y_0 = 0 \end{array}}.$$

Note that (4.25) is equivalent to (4.23). The SOSC for (4.1) at $\tilde{x}^*$ requires: for every $\tilde{y} \neq 0$ satisfying (4.25), it holds that

$$\tilde{y}^{\mathrm{T}} \nabla_{\tilde{x}\tilde{x}}^2 L(\tilde{y}^*)\tilde{y} = y^{\mathrm{T}} \nabla^2 L_1(x^*)y > 0.$$

The above is equivalent to that $y^{\mathrm{T}} \nabla^2 L_1(x^*)y > 0$ for every $y \neq 0$ satisfying (4.23). So the conclusion of item (iii) holds. □

4.3. **The even degree case.** When f and c_j $(j \in \mathcal{I})$ all have even degrees, the homogenization (3.4) is equivalent to (3.8); see Proposition 3.4. This means that the constraint $x_0 \geq 0$ is redundant for (4.1). Typically, the SCC fails to hold for minimizers at infinity. For this case, Theorem 4.3 is not applicable for showing the finite convergence for the hierarchy of (3.6)-(3.7). However, the same conclusion like in Theorem 4.3 holds under optimality conditions without the constraint $x_0 \geq 0$.

The kth order SOS relaxation for (3.8) is

(4.26)
$$\begin{cases} \max & \gamma \\ \text{s.t.} & \tilde{f}(\tilde{x}) - \gamma x_0^d \in \mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]_{2k} + \mathrm{QM}[\tilde{c}_{\mathcal{I}^*}]_{2k}. \end{cases}$$

Its dual optimization is the kth order moment relaxation

(4.27)
$$\begin{cases} \min & \langle \tilde{f}, y \rangle \\ \text{s.t.} & L_p^{(k)}[y] = 0 \ (p \in \tilde{c}_{\mathcal{E}}), \\ & L_q^{(k)}[y] \succeq 0 \ (q \in \tilde{c}_{\mathcal{I}^*}), \\ & M_k[y] \succeq 0, \\ & \langle x_0^d, y \rangle = 1, \ y \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{cases}$$

In the above, the polynomial tuple $\tilde{c}_{\mathcal{I}^*}$ is

(4.28)
$$\tilde{c}_{\mathcal{I}^*} := \{\tilde{c}_j(\tilde{x})\}_{j \in \mathcal{I}}.$$

Let f_k^e and $f_k^{e,\prime}$ denote the optimal values of (4.26), (4.27) respectively. To study their convergence, we consider the optimization problem

(4.29)
$$\begin{cases} \min & F(\tilde{x}) = \tilde{f}(\tilde{x}) - f_{\min} \cdot x_0^d \\ \text{s.t.} & \tilde{c}_i(\tilde{x}) = 0 \ (i \in \mathcal{E}), \\ & \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ & x_0^2 + \|x\|^2 - 1 = 0. \end{cases}$$

For the even degree case, the LICQC, SCC and SOSC are said to hold at a minimizer at infinity x^* if they hold for (4.29) at $(0, x^*)$. The following is the convergence for the hierarchy of relaxations (4.26)-(4.27).

Theorem 4.7. *Assume K is closed at ∞, $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical, f and c_j $(j \in \mathcal{I})$ all have even degrees. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), including the one at infinity, then the hierarchy of relaxations (4.26)-(4.27) is tight, i.e., there exists k_0 such that*

$$f_k^e = f_k^{e,\prime} = f_{\min} \quad \text{for all } k \geq k_0.$$

Moreover, the hierarchy of relaxations (3.6)-(3.7) is also tight.

Proof. The proof is almost the same as for Theorem 4.3. Under the given assumptions, there exist $\sigma \in \mathrm{QM}[\tilde{c}_{\mathcal{I}^*}]$, $\phi \in \mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ such that

$$\tilde{f} - f_{\min} x_0^d = \phi + \sigma.$$

So, the hierarchy of (4.26)-(4.27) is tight. Since $\mathrm{QM}[\tilde{c}_{\mathcal{I}^*}] \subseteq \mathrm{QM}[\tilde{c}_{\mathcal{I}}]$, the hierarchy of (3.6)-(3.7) is also tight. $\qquad\square$

In the following, we discuss optimality conditions for minimizers at infinity of (4.29). Suppose $\tilde{x}^* = (0, x^*)$ is a minimizer at infinity for (4.29). Let $J_1(x^*)$ be as in (4.18). Since d is even, the KKT equation for (4.29) at $\tilde{x}^*$ is

$$(4.30) \qquad \begin{bmatrix} f^{(2)}(x^*) \\ \nabla f^{(1)}(x^*) \end{bmatrix} = \sum_{i \in J_1(x^*)} \lambda_i \begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} + \bar{\lambda} \begin{bmatrix} 0 \\ x^* \end{bmatrix}.$$

The following is a similar version of Theorem 4.6.

Theorem 4.8. *Suppose K is closed at ∞ and $\tilde{x}^* = (0, x^*)$ is a minimizer at ∞ for (4.29). Assume f and c_j $(j \in \mathcal{I})$ all have even degrees. Let $d = \deg(f) > 1$ and let $J_1(x^*)$ be as in (4.18). Then, we have:*

(i) *The LICQC holds for (4.29) at $\tilde{x}^*$ if and only if the gradients*

$$\begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix} \quad \left(i \in J_1(x^*)\right)$$

are linearly independent.

(ii) *Suppose the LICQC holds for (4.29) at $\tilde{x}^*$. Then, $\bar{\lambda} = 0$ and (4.30) is reduced to*

$$(4.31) \qquad \begin{bmatrix} f^{(2)}(x^*) \\ \nabla f^{(1)}(x^*) \end{bmatrix} = \sum_{i \in J_1(x^*)} \lambda_i \begin{bmatrix} c_i^{(2)}(x^*) \\ \nabla c_i^{(1)}(x^*) \end{bmatrix}.$$

The SCC holds for (4.29) at $\tilde{x}^*$ if and only if $\lambda_i > 0$ for $i \in J_1(x^*) \cap \mathcal{I}$.

(iii) *Suppose the LICQC holds for (4.29) at $\tilde{x}^*$. Let λ_i be Lagrange multipliers as in the item (ii) and let H be the matrix*

$$(4.32) \qquad H = \begin{bmatrix} 2f^{(3)}(x^*) - d(d-1)f_{\min}0^{d-2} & \nabla f^{(2)}(x^*)^{\mathrm{T}} \\ \nabla f^{(2)}(x^*) & \nabla^2 f^{(1)}(x^*) \end{bmatrix}$$
$$- \sum_{i \in J_1(x^*)} \lambda_i \begin{bmatrix} 2c_i^{(3)}(x^*) & \nabla c_i^{(2)}(x^*)^{\mathrm{T}} \\ \nabla c_i^{(2)}(x^*) & \nabla^2 c_i^{(1)}(x^*) \end{bmatrix}.$$

Then, the SOSC holds for (4.29) at $\tilde{x}^*$ if and only if $\tilde{y}^{\mathrm{T}} H \tilde{y} > 0$ for every nonzero $\tilde{y} = (y_0, y)$ satisfying

$$(4.33) \qquad \boxed{\begin{aligned} y_0 c_i^{(2)}(x^*) + y^{\mathrm{T}} \nabla c_i^{(1)}(x^*) = 0 \ (i \in J_1(x^*)), \\ y^{\mathrm{T}} x^* = 0 \end{aligned}}.$$

Proof. The proofs for the item (i) and (ii) are exactly the same as for (i), (ii) in Theorem 4.6. So they are omitted for cleanness. The item (iii) can be shown as follows. The matrix H is the Hessian at $\tilde{x}^*$ of the Lagrange function

$$L(\tilde{x}) = \tilde{f}(\tilde{x}) - \sum_{i \in J_1(x^*)} \lambda_i \tilde{c}_i(x).$$

A vector $\tilde{y} = (y_0, y)$ lies in the tangent space of active constraints of (4.29) at $\tilde{x}^*$ if and only if it satisfies (4.33). Thus, the SOSC holds for (4.29) at $\tilde{x}^*$ if and only if $\tilde{y}^{\mathrm{T}} H \tilde{y} > 0$ for every nonzero $\tilde{y}$ satisfying (4.33). $\qquad\square$

5. Extensions of Putinar-Vasilescu's Positivstellensatz

In this section, we generalize the Putinar-Vasilescu's Positivstellensatz [47, 48] to polynomials that are nonnegative on unbounded semialgebraic sets. Under some assumptions on optimality conditions, we prove the conclusions of the Putinar-Vasilescu's Positivstellensatz.

For a polynomial tuple $g := (g_1, \ldots, g_m)$, consider the semialgebraic set

$$(5.1) \qquad S := \{x \in \mathbb{R}^n \mid g_j(x) \geq 0, \ j = 1, \ldots, m\}.$$

Recall that $f^{(1)}$ denotes the homogeneous part of the highest degree for f. The following is the classical Putinar-Vasilescu's Positivstellensatz.

Theorem 5.1. *([47, Theorems 1,2]) Let $f, g_1, \ldots, g_m \in \mathbb{R}[x]$ and S be as in (5.1). Then, we have:*

  (i) *Suppose $f, g_1, \ldots, g_m$ are homogeneous polynomials of even degrees. If $f > 0$ on $S \backslash \{0\}$, then $\|x\|^{2k} f \in QM[g]$ for some power $k \in \mathbb{N}$.*
  (ii) *If the form $f^{(1)}$ is positive definite in $\mathbb{R}^n$ and $f > 0$ on S, then $(1 + \|x\|^2)^k f \in QM[g]$ for some power $k \in \mathbb{N}$.*

For the case that $S = \mathbb{R}^n$, if f is a positive definite form, then $\|x\|^{2k} f$ is SOS for some power k. This conclusion is referred to as the Reznick's Positivstellensatz and it is shown in [49].

First, we generalize the item (i) of Theorem 5.1. Consider the normalized optimization problem

$$(5.2) \qquad \begin{cases} \min & f(x) \\ s.t. & g_j(x) \geq 0, \ j = 1, \ldots, m, \\ & \|x\|^2 - 1 = 0. \end{cases}$$

Theorem 5.2. *Let S be the set as in (5.1). Suppose $f, g_1, \ldots, g_m$ are homogeneous polynomials of even degrees such that $f \geq 0$ on S. If the LICQC, SCC and SOSC hold at every minimizer of (5.2), then $\|x\|^{2k} f \in QM[g]$ for some $k \in \mathbb{N}$.*

Proof. By Theorem 1.1 of [41], since $f_{\min} \geq 0$, there exist polynomials $\sigma_i \in \Sigma[x]$ and $h \in \mathbb{R}[x]$ such that (let $g_0 := 1$)

$$f = \sum_{i=0}^{m} \sigma_i g_i + h \cdot (\|x\|^2 - 1).$$

Let $2d_0 := \deg(f)$ and $2d_i := \deg(g_i)$. In the above identity, if we substitute x_i for $\frac{x_i}{\|x\|}$, then we get

$$\frac{f(x)}{\|x\|^{2d_0}} = \sum_{i=0}^{m} \sigma_i \left(\frac{x}{\|x\|} \right) \frac{g_i(x)}{\|x\|^{2d_i}}.$$

Furthermore, we can also get

$$\frac{2f(x)}{\|x\|^{2d_0}} = \sum_{i=0}^{m} \left[\sigma_i\left(\frac{x}{\|x\|}\right) + \sigma_i\left(\frac{-x}{\|x\|}\right)\right] \frac{g_i(x)}{\|x\|^{2d_i}}.$$

Note that the odd degree terms in $\sigma_i\left(\frac{x}{\|x\|}\right) + \sigma_i\left(\frac{-x}{\|x\|}\right)$ are cancelled. The above implies that $\|x\|^{2k}f \in \mathrm{QM}[g]$ when k is large enough. $\square$

Second, we generalize the item (ii) of Theorem 5.1. Consider the optimization

$$(5.3) \qquad \begin{cases} \min & f(x) \\ s.t. & g_j(x) \geq 0, \ j = 1, \ldots, m. \end{cases}$$

For convenience, we still let $f_{\min}$ denote the minimum value of (5.3). The homogenized optimization problem is

$$(5.4) \qquad \begin{cases} \min & \tilde{f}(\tilde{x}) - f_{\min} \cdot (x_0)^d \\ s.t. & \tilde{g}_1(\tilde{x}) \geq 0, \ldots, \tilde{g}_m(\tilde{x}) \geq 0, \\ & \|\tilde{x}\|^2 - 1 = 0. \end{cases}$$

Theorem 5.3. *Let S be as in (5.1). Suppose $f \geq 0$ on S, then we have:*
  (i) *Suppose S is closed at ∞ and the degrees of $f, g_1, \ldots, g_m$ are even. If the LICQC, SCC and SOSC hold at every minimizer of (5.3), including the one at infinity, then $(1 + \|x\|^2)^k f \in \mathrm{QM}[g]$ for some $k \in \mathbb{N}$.*
  (ii) *If the form $f^{(1)}$ is positive definite in $\mathbb{R}^n$ and the LICQC, SCC and SOSC hold at every minimizer of (5.3), then $(1 + \|x\|^2)^k f \in \mathrm{QM}[g]$ for some $k \in \mathbb{N}$.*

Proof. (i) It follows from Theorem 4.4 that the LICQC, SCC and SOSC hold at every regular minimizer of (5.4). By Theorem 4.7, since $f_{\min} \geq 0$, there exist polynomials $\sigma_i \in \Sigma[\tilde{x}]$ and $h \in \mathbb{R}[\tilde{x}]$ such that (note $g_0 = 1$)

$$(5.5) \qquad \tilde{f} = \sum_{i=0}^{m} \sigma_i \tilde{g}_i + h \cdot (\|\tilde{x}\|^2 - 1).$$

As in the proof of Theorem 5.2, we can similarly show that

$$\|\tilde{x}\|^{2k} \tilde{f} \in \mathrm{QM}[\tilde{g}_1, \ldots, \tilde{g}_m],$$

for some $k \in \mathbb{N}$. Substituting x_0 for 1, we get $(1 + \|x\|^2)^k f \in \mathrm{QM}[g]$.

(ii) For each i, let $\theta_i := 2\lceil \frac{\deg(g_i)}{2} \rceil - \deg(g_i)$. We consider the following homogenized optimization problem

$$(5.6) \qquad \begin{cases} \min_{\tilde{x} \in \mathbb{R}^{n+1}} & \tilde{f}(\tilde{x}) - f_{\min} \cdot (x_0)^d \\ s.t. & x_0^{\theta_1} \tilde{g}_1(\tilde{x}) \geq 0, \ldots, x_0^{\theta_m} \tilde{g}_m(\tilde{x}) \geq 0, \\ & \|\tilde{x}\|^2 - 1 = 0. \end{cases}$$

Note that $f_{\min} \geq 0$ and the degree of f is even. Let $\tilde{u} = (u_0, u)$ be a feasible point of (5.6). If $u_0 = 0$, then $\|u\| = 1$ and $\tilde{f}(\tilde{u}) - f_{\min} \cdot (u_0)^d = f^{(1)}(u) > 0$, since $f^{(1)}$ is a positive definite form. If $u_0 \neq 0$, then u/u_0 is feasible for (5.3) and

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

Hence, the optimal value of (5.6) is zero. By similar arguments as for Theorem 4.4, the LICQC, SCC and SOSC hold at every regular minimizer of (5.6). Since $f^{(1)}$ is a positive definite form, (5.6) has no minimizers at infinity. Note that each

$x_0^{\theta_i} \tilde{g}_i$ is a polynomial of even degree. As for part (i), we can similarly show that $(1 + \|x\|^2)^k f \in \mathrm{QM}[g]$ for some $k \in \mathbb{N}$. $\qquad\square$

When one of the LICQC, SCC and SOSC fails to hold, the conclusion $(1 + \|x\|^2)^k f \in \mathrm{QM}[g]$ may not hold. We refer to [32] for such examples. When there is an equality constraint in (1.1), we cannot simply replace $c_i(x) = 0$ by two inequalities $c_i(x) \geq 0$ and $-c_i(x) \geq 0$, since the LICQC always fails for the latter case. For the case of equality constrains, we need to assume that $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical. For convenience of notation, denote the polynomial tuples

$$c_{\mathcal{E}} := \big\{c_i(x)\big\}_{i \in \mathcal{E}}, \quad c_{\mathcal{I}} := \big\{c_j(x)\big\}_{j \in \mathcal{I}}.$$

The following is the Positivstellensatz for sets with equality constraints.

Theorem 5.4. *Let K be the feasible set of (1.1). Assume that $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical and $f \geq 0$ on K, then we have:*
  (i) *Suppose K is closed at ∞ and the degrees of f, c_i ($i \in \mathcal{I}$) are all even. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), including the one at infinity, then $(1 + \|x\|^2)^k f \in \mathrm{Ideal}[c_{\mathcal{E}}] + \mathrm{QM}[c_{\mathcal{I}}]$ for some $k \in \mathbb{N}$.*
  (ii) *Suppose the form $f^{(1)}$ is positive definite in $\mathbb{R}^n$. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), then $(1 + \|x\|^2)^k f \in \mathrm{Ideal}[c_{\mathcal{E}}] + \mathrm{QM}[c_{\mathcal{I}}]$ for some $k \in \mathbb{N}$.*

Proof. (i) Since the degrees of f, c_i ($i \in \mathcal{I}$) are all even, we consider the optimization problem (4.29). The conclusion can be shown in the same way as for item (i) of Theorem 5.3.

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

$$(5.7) \qquad \begin{cases} \min & \tilde{f}(\tilde{x}) - f_{\min} \cdot (x_0)^d \\ s.t. & \tilde{c}_i(\tilde{x}) = 0 \ (i \in \mathcal{E}), \\ & x_0^{\theta_j} \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ & \|\tilde{x}\|^2 - 1 = 0. \end{cases}$$

Since $f^{(1)}$ is a positive definite form, the optimization (5.7) has no minimizers at infinity. By similar arguments as for Theorem 4.4, the optimality conditions LICQC, SCC and SOSC hold at every regular minimizer of (5.7). The conclusion can be shown in the same way as for item (ii) of Theorem 5.3. $\qquad\square$

6. THE MOMENT-SOS HIERARCHY WITH DENOMINATORS

The Putinar-Vasilescu's Positivstellensatz motivates Moment-SOS relaxations with denominators for solving polynomial optimization. In view of Theorem 5.4, we consider the hierarchy of relaxations ($d = \deg(f)$)

$$(6.1) \qquad \begin{cases} \max & \gamma \\ s.t. & (1 + \|x\|^2)^{k - \lceil \frac{d}{2} \rceil}(f - \gamma) \in \mathrm{Ideal}[c_{\mathcal{E}}]_{2k} + \mathrm{QM}[c_{\mathcal{I}}]_{2k}, \end{cases}$$

for the order $k \geq \lceil \frac{d}{2} \rceil$. The relaxation (6.1) is essentially expressing $f - \gamma$ in terms of sums of squares of rational polynomials, with the denominator a power of $1 + \|x\|^2$. Let f_k^{den} denote the optimal value of (6.1) for the relaxation order k. The following is the comparisons between the relaxations (3.6) (or (4.26) for the even degree case) and (6.1):

- The polynomials in (6.1) are only in the variable x, while the polynomials in (3.6) and (4.26) are in both x and x_0.
- The relaxation (3.6) uses the constraint $x_0 \geq 0$, while (6.1) and (4.26) do not use it. So (3.6) is the strongest among them.
- When the degrees of f and c_i ($i \in \mathcal{I}$) are all even, the relaxation (4.26) is equivalent to (6.1). This can be observed as in the proof of Theorem 5.2. For the special case that f is an even degree form and K is the unit sphere, the equivalence is shown in [17].

We remark that (6.1) is the relaxation given in [32] for the parameter $\epsilon = 0$. It is conjectured in [32, Sec. 4.2] that the hierarchy of (6.1) has finite convergence, under some optimality conditions. We prove this conjecture is true under the assumptions of Theorem 5.4. Recall that for the even degree case, the LICQC, SCC and SOSC are said to hold at a minimizer at infinity x^* if they hold for (4.29) at $(0, x^*)$.

Theorem 6.1. *Let K be the feasible set of (1.1). Assume $\mathrm{Ideal}[\tilde{c}_{\mathcal{E}}]$ is real radical.*

(i) *Suppose K is closed at ∞ and the degrees of f, c_i ($i \in \mathcal{I}$) are even. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), including the one at infinity, the hierarchy of (6.1) has finite convergence, i.e., $f_k^{den} = f_{\min}$ for all k big enough.*

(ii) *Suppose $f^{(1)}$ is a positive definite form in $\mathbb{R}^n$. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), then $f_k^{den} = f_{\min}$ for all k big enough.*

Proof. The conclusions follow from Theorem 5.4. For the item (ii), there are no minimizers at infinity if $f^{(1)}$ is a positive definite form in $\mathbb{R}^n$. $\square$

In Theorem 6.1(i), the degrees of f, c_i ($i \in \mathcal{I}$) are assumed to be even, but the degrees of equality constraining polynomials c_i ($i \in \mathcal{E}$) can be either odd or even.

A special case of (1.1) is that there are no constraints. Then the resulting version of the relaxation (6.1) is

$$(6.2) \qquad \begin{cases} \max & \gamma \\ \mathrm{s.t.} & (1 + \|x\|^2)^{k - \lceil \frac{d}{2} \rceil}(f - \gamma) \in \Sigma[x]_{2k}, \end{cases}$$

for the order $k \geq \lceil \frac{d}{2} \rceil$. The degree d must be even for $f_{\min} > -\infty$, when there are no constraints. When $f^{(1)}$ is a positive definite form, the asymptotic convergence of (6.2) can be shown by Reznick's Positivstellensatz [49, Sec. 7]. We have the following theorem about the finite convergence.

Theorem 6.2. *Suppose $K = \mathbb{R}^n$ and $f_{\min} > -\infty$. Let f_k^{den} denote the optimal value of (6.2) for the order k.*

(i) *If the SOSC holds at every minimizer of (1.1), including the one at infinity, then $f_k^{den} = f_{\min}$ for all k big enough.*

(ii) *If $n \leq 2$, then $f_k^{den} = f_{\min}$ for all k big enough.*

Proof. The item (i) follows from Theorem 5.4 (i). It is shown by Scheiderer [50] that if $p(u)$ is a nonnegative form in $u \in \mathbb{R}^n$ with $n \leq 3$, then $\|u\|^{2N} p(u)$ is SOS when N is big enough. This implies that

$$(1 + \|x\|^2)^{k - \lceil \frac{d}{2} \rceil}(f - f_{\min}) \in \Sigma[x]_{2k}$$

when k is big enough, for the case $n \leq 2$. The item (ii) follows from this conclusion. $\square$

7. Numerical examples

This section presents some examples on the relaxations (3.6) for solving the optimization problem (1.1). The computation is implemented in MATLAB R2019a, on a Dell Desktop with CPU@2.90GHz and RAM 32.0G. The relaxations (3.6) − (3.7) are solved by the software `GloptiPoly 3` [14], which calls the SDP package `SeDuMi` [56]. For neatness, only four decimal digits are displayed for computational results. The feasible sets of all examples are unbounded.

7.1. **The case with minimizers.** A convenient criterion for obtaining minimizers is the flat extension or truncation. Denote the degree

$$(7.1) \qquad d_K := \max\{\lceil \deg(f)/2 \rceil, \lceil \deg(c_i)/2 \rceil (i \in \mathcal{E} \cup \mathcal{I})\}.$$

Suppose y^* is a minimizer of (3.7) for the relaxation order k. If there exists an integer $t \in [d_K, k]$ such that

$$(7.2) \qquad \operatorname{rank} M_t[y^*] = \operatorname{rank} M_{t-d_K}[y^*],$$

then we can get one or several minimizers for (4.1) (see [3, 13, 27, 38]). When (7.2) holds, we have the decomposition

$$y^*|_{2t} = a_1 \Big[\binom{\tau_1}{v_1} \Big]_{2t} + \cdots + a_r \Big[\binom{\tau_r}{v_r} \Big]_{2t}$$

for positive scalars $a_i > 0$ and distinct points $(\tau_i, v_i) \in \widetilde{K}$. Denote two label sets

$$I_1 = \{i : \tau_i > 0\}, \quad I_2 = \{i : \tau_i = 0\}.$$

For each $i \in I_1$, let $u_i = v_i/\tau_i$. Then $u_i \in K$ for each $i \in I_1$ and $v_i \in K^{(1)}$ for each $i \in I_2$. Let $\nu_i = a_i(\tau_i)^d$ for each $i \in I_1$ and $\nu_i = a_i$ for each $i \in I_2$.

Lemma 7.1. *Suppose y^* is a minimizer of (3.7) and the rank condition (7.2) is satisfied. Let each u_i, v_i, ν_i be as above. Then the set $I_1 \neq \emptyset$ and each u_i ($i \in I_1$) is a minimizer for (1.1), and each v_i ($i \in I_2$) is a minimizer at infinity for (1.1).*

Proof. Note that $u_i \in K$ for each $i \in I_1$ and $v_i \in K^{(1)}$ for each $i \in I_2$, the constraint $\langle x_0^d, y \rangle = 1$ in (3.7) implies that $\sum_{i \in I_1} \nu_i = 1$. So at least one $\nu_i > 0$ and $I_1 \neq \emptyset$. Since v_i ($i \in I_2$) is a feasible point of $K^{(1)}$, by item (iii) of Theorem 3.6, we have $f^{(1)}(v_i) \geq 0$ for $i \in I_2$ and hence

$$f_{\min} = \langle \tilde{f}, y^* \rangle = \sum_{i \in I_1} \nu_i f(u_i) + \sum_{i \in I_2} \nu_i f^{(1)}(v_i)$$

$$\geq \sum_{i \in I_1} \nu_i f(u_i) \geq \sum_{i \in I_1} \nu_i f_{\min} = f_{\min}.$$

Therefore, $f(u_i) = f_{\min}$ for $i \in I_1$, and $f(v_i) = 0$ for $i \in I_2$. $\qquad\square$

We refer to [3, 13, 27, 38] for flat extensions and truncations. The procedure of extracting minimizers by using (7.2) is implemented in `GloptiPoly 3` [14]. The rank condition (7.2) is a sufficient and almost necessary condition for checking convergence of the Moment-SOS hierarchy [38].

Example 7.2. (i) Consider the optimization (the variable $x_0 = 1$):

$$\min_{x \in \mathbb{R}^4} \sum_{i=0}^{4} \prod_{i \neq j \in \{0,1,\ldots,4\}} (x_i - x_j) + 0.1 \Big(\sum_{i=1}^{4} x_i^4 \Big).$$

The first part of the objective is a nonnegative but non-SOS polynomial [49]. For the order $k = 3$, we get $f_3 \approx 0.0763$ and the minimizer $0.5757 \cdot (1,1,1,1)$. There are no minimizers at infinity.

(ii) Consider the optimization:

$$\begin{cases} \min_{x \in \mathbb{R}^2} & x_1^2 x_2 + x_2^2 x_1 - 3 x_1 x_2 \\ \text{s.t.} & x_1 \geq 0, \ x_2 \geq 0. \end{cases}$$

Up to a constant, the objective becomes the dehomogenization of the Motzkin form [49] if each x_i is changed to x_i^2. For the order $k = 3$, we get $f_3 \approx -1.0000$ and a minimizer $(1.0000, 1.0000)$. We also get two minimizers at infinity. They are $(1.0000, 0.0000)$ and $(0.0000, 1.0000)$.

(iii) Consider the optimization:

$$\begin{cases} \min_{x \in \mathbb{R}^2} & x_1^2 x_2 + x_2^2 + x_1 - 3 x_1 x_2 \\ \text{s.t.} & x_1 \geq 0, \ x_2 \geq 0. \end{cases}$$

The objective becomes the dehomogenization of the Choi-Lam form [49] if each x_i is changed to x_i^2. For the order $k = 2$, we get $f_2 \approx 2.9586 \times 10^{-8}$ and two minimizers: $(1.0000, 1.0000)$, $(0.0000, 0.0000)$. We also get two minimizers at infinity. They are $(1.0000, 0.0000)$ and $(0.0000, 1.0000)$.

(iv) Consider the optimization:

$$\begin{cases} \min_{x \in \mathbb{R}^2} & x_1^3 + x_2^3 + 3 x_1 x_2 - x_1^2 (x_2 + 1) - x_2^2 (x_1 + 1) - (x_1 + x_2) \\ \text{s.t.} & x_1 \geq 0, \ x_2 \geq 0. \end{cases}$$

Up to a constant, the objective becomes the dehomogenization of the Robinson form [49] if each x_i is changed to x_i^2. For the order $k = 2$, we get $f_2 \approx -1.0000$ and three minimizers:

$$(1.0000, 1.0000), \quad (0.0000, 1.0000), \quad (1.0000, 0.0000).$$

We also get one minimizer at infinity: $(0.7071, 0.7071)$.

Example 7.3. (i) Consider the constrained optimization:

$$\begin{cases} \min_{x \in \mathbb{R}^2} & x_1^6 + x_2^6 + 1 + 3 x_1^2 x_2^2 - x_1^2 \left(x_2^4 + 1 \right) \\ & \qquad - x_2^2 \left(1 + x_1^4 \right) - \left(x_1^4 + x_2^4 \right) \\ \text{s.t.} & x_1 + x_2 + 1 = 0 \end{cases}$$

This problem is a variation of Example 5.2 in [39]. For the order $k = 3$, we get $f_3 \approx 5.4436 \times 10^{-6}$ and we get two minimizers: $(-1.0000, 0.0000), (0.0000, -1.0000)$. We also get two minimizers at infinity:

$$(0.7068, -0.7074), \quad (-0.7074, 0.7068).$$

(ii) Consider the constrained optimization:

$$\begin{cases} \min\limits_{x\in\mathbb{R}^2} & x_1^2 + x_2^2 \\ \quad s.t. & x_2^2 - 1 \geq 0, \\ & x_1^2 - 2x_1x_2 - 1 \geq 0, \\ & x_1^2 + 2x_1x_2 - 1 \geq 0. \end{cases}$$

This example is from [39]. The minimum value $f_{\min} \approx 6.8284$ and the minimizers are $(\pm(1 + \sqrt{2}), \pm 1)$. There are no minimizers at infinity. For the order $k = 3$, we get $f_3 \approx 6.8284$ and four minimizers: $(\pm 2.4142, \pm 1.0000)$.

(iii) Consider the constrained optimization:

$$\begin{cases} \min\limits_{x\in\mathbb{R}^3} & x_1^2(x_1 - 1)^2 + x_2^2(x_2 - 1)^2 + x_3^2(x_3 - 1)^2 \\ & \quad + 2x_1x_2x_3(x_1 + x_2 + x_3 - 2) \\ & \quad + (x_1 - 1)^2 + (x_2 - 1)^2 + (x_3 - 1)^2 \\ \quad s.t. & x_1 - 2x_2^2 \geq 0, \ x_2 - x_3 \geq 0. \end{cases}$$

The sum of the first four terms of the objective is a nonnegative polynomial [49]. For $k = 2$, we get $f_2 \approx 0.4708$ and the minimizer $(0.6979, 0.6980, 0.6978)$.

(iv) Consider the constrained optimization:

$$\begin{cases} \min\limits_{x\in\mathbb{R}^2} & 2x_1^3 + 2x_2^3 + 4x_1x_2 - x_1(x_2^2 + 1) \\ & \quad + x_2(1 + x_1^2) + x_1^2 + x_2^2 \\ \quad s.t. & x_1 \geq 1, \ x_2 \geq 1. \end{cases}$$

For $k = 2$, we get $f_2 \approx 2.0000$ and the minimizer $(1.0000, 1.0000)$.

7.2. The case with no minimizers. When the minimum value $f_{\min}$ of (1.1) is not achievable, i.e., (1.1) has no minimizers, then the moment relaxation (3.7) can not have a minimizer y^* satisfying (7.2). This is implied by Lemma 7.1. Indeed, the moment relaxation (3.7) typically does not achieve its optimal value, i.e., it does not have optimizers either. For such cases, there often exist numerical issues for solving Moment-SOS relaxations (3.6)-(3.7), although the finite convergence is guaranteed under some assumptions on minimizers at infinity.

For instance, consider the unconstrained optimization with the objective $f = x_1^4 + (x_1x_2 - 1)^2$. Clearly, $f_{\min} = 0$ is not achievable and $f_k = f'_k = 0$ for all $k \geq 2$. The moment relaxation (3.7) does not have an optimizer. The objective in (3.7) is $y_{040} + y_{022} - 2y_{211} + y_{400}$. For $k \geq 2$, the moment matrix constraint $M_k[y] \succeq 0$ implies that

$$\begin{pmatrix} y_{400} & y_{310} & y_{301} & y_{220} & y_{211} & y_{202} \\ y_{310} & y_{220} & y_{211} & y_{130} & y_{121} & y_{112} \\ y_{301} & y_{211} & y_{202} & y_{121} & y_{112} & y_{103} \\ y_{220} & y_{130} & y_{121} & y_{040} & y_{031} & y_{022} \\ y_{211} & y_{121} & y_{112} & y_{031} & y_{022} & y_{013} \\ y_{202} & y_{112} & y_{103} & y_{022} & y_{013} & y_{004} \end{pmatrix} \succeq 0.$$

Note that $y_{400} = 1$. When y is feasible for (3.7), we can get

$$y_{040} \geq 0, \quad y_{022} \geq 0, \quad y_{022} - 2y_{211} + 1 \geq 0.$$

If $y_{040} > 0$, then $\langle \tilde{f}, y \rangle > 0$. If $y_{040} = 0$, then

$$y_{022} = 0, \quad y_{211} = 0, \quad y_{022} - 2y_{211} + 1 = 1 > 0.$$

The objective of (3.7) is positive for all feasible y, so it does not achieve the optimal value. There are numerical troubles for for solving the Moment-SOS relaxations.

When $f_{\min}$ is not achievable, a more numerically well-posed problem is to compute minimizers at infinity. We apply the Moment-SOS relaxations to solve the optimization problem (4.17) for one or several minimizers x^* at infinity. When the optimality conditions hold at minimizers of (4.17), its Moment-SOS hierarchy has finite convergence, so its minimizers can be obtained (see [38, 41]). For $f_{\min} > -\infty$, it is necessary that $f^{(1)}(x^*) = 0$. This is implied by Theorem 3.6.

Example 7.4. (i) Consider the optimization:

$$\min_{x \in \mathbb{R}^2} \quad x_2^2 + (2x_2^2 + 2x_1x_2 + 1)^2.$$

The minimum value $f_{\min} = 0$ is not achievable. The minimizers at infinity are $(\pm 1, 0)$, $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$, $(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. For the order $k = 3$, we get $f_3 \approx 9.8893 \times 10^{-9}$ and four minimizers at infinity:

$$(-1.0000, 0.0001), (1.0000, -0.0001), (-0.7071, 0.7071), (0.7071, -0.7071).$$

(ii) Consider the optimization:

$$\min_{x \in \mathbb{R}^3} \quad x_1^2 + (1 - x_1x_2)^2 + L(x),$$

where $L = x_1^4x_2^2 + x_2^4x_3^2 + x_3^4x_1^2 - 3x_1^2x_2^2x_3^2$ is the Choi-Lam form. Clearly, $f_{\min} \geq 0$. For the sequence of $x^{(k)} = (\frac{1}{n}, n, 0)$, we have $f(x^{(k)}) = \frac{2}{n^2}$. As $k \to 0$, we have $f(x^{(k)}) \to 0$, which implies $f_{\min} = 0$. However, $f_{\min} = 0$ is not achievable, since the polynomial $x_1^2 + (1 - x_1x_2)^2$ has no real zeros. The objective f is not an SOS, since $f^{(1)} = L(x)$ is not an SOS. The minimizers at infinity are

$$\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1), (\pm 1, 0, 0), (0, \pm 1, 0), (0, 0, \pm 1).$$

For $k = 5$, we get $f_5 \approx -1.6413 \times 10^{-8}$ and all the minimizers at infinity.
(iii) Consider the optimization:

$$\min_{x \in \mathbb{R}^3} \quad \epsilon((x_3^2 + x_1x_3 + 1)^2 + x_3^6) + R(x),$$

where $\epsilon > 0$ and R is the dehomogenized Robinson polynomial

$$R(x) = 1 + x_2^6 + x_3^6 + 3x_2^2x_3^2 - (x_2^2 + x_3^2) - x_2^4(1 + x_3^2) - x_3^4(1 + x_2^2).$$

Clearly, $f_{\min} \geq 0$, since the objective is a sum of two nonnegative polynomials. For the sequence of $x^{(k)} = (-\frac{1+n^2}{n}, 1, \frac{1}{n})$, we have $f(x^{(k)}) = \frac{1+\epsilon-2n^2+n^4}{n^6}$. As $k \to 0$, we have $f(x^{(k)}) \to 0$. So, $f_{\min} = 0$. However, $f_{\min}$ is not achievable, since the polynomial $(x_3^2 + x_1x_3 + 1)^2 + x_3^2$ has no real zeros. For $\epsilon > 0$ small enough, f is not an SOS. The minimizers at infinity are $(\pm 1, 0, 0)$. Let $\epsilon = 1$, for the order $k = 3$, we get $f_3 \approx 3.1810 \times 10^{-10}$ and two minimizers at infinity: $(\pm 1.0000, 0.0000, 0.0000)$.
(iv) Consider the optimization:

$$\begin{cases} \min_{x \in \mathbb{R}^5} & (x_1 + x_2 + x_3 + x_4x_5)^2 - 4\Big(x_1x_2 + x_2x_3 + x_3(x_4x_5 - 1) \\ & \quad + x_4x_5 - 1 + x_1\Big) + (x_1 - 1)^2 + x_4^2 \\ s.t. & x_1 \geq 0, x_2 - x_1 \geq 0, x_3 - x_2 \geq 0, \\ & x_4 - x_3 \geq 0, x_5 - x_4 \geq 0, x_4x_5 \geq 1. \end{cases}$$

The minimum value $f_{\min} = 0$ is not achievable. The minimizer at infinity is $(0, 0, 0, 0, 1)$. For the order $k = 2$, we get $f_2 \approx 6.2486 \times 10^{-9}$ and the minimizer at infinity $(0.0000, 0.0000, 0.0000, 0.0000, 1.0000)$.

When minimizers at infinity are obtained, we can apply the procedure in [45] to determine the value $f_{\min}$. With these minimizers at infinity, one can drop some connected components of the tangency variety that do not converge to minimizers at infinity. This can save computational expense. For instance, we consider the polynomial $f = (x_1 x_2 - 1)^2 + x_2^2$ and $K = \mathbb{R}^2$. Clearly, the minimum value $f_{\min}$ is not attainable. By the procedure in [45], the tangency variety $\Gamma\left(f, \mathbb{R}^2\right)$ is given by the equation:

$$2\left(-x_1^3 x_2 + x_1 x_2^3 + x_1^2 - x_1 x_2 - x_2^2\right) = 0.$$

For $R > 0$ large enough, the set $\Gamma\left(f, \mathbb{R}^2\right) \cap \{x_1^2 + x_2^2 \geq R\}$ has eight connected components:

$$
\begin{aligned}
\Gamma_{\pm 1} : \quad & x_1 := t, \quad & x_2 := -t + \tfrac{1}{2}t^{-1} + \tfrac{5}{8}t^{-3} + \cdots, \\
\Gamma_{\pm 2} : \quad & x_1 := t, \quad & x_2 := t + \tfrac{1}{2}t^{-1} + \tfrac{3}{8}t^{-3} + \cdots, \\
\Gamma_{\pm 3} : \quad & x_1 := t, \quad & x_2 := t^{-1} - t^{-3} + \cdots, \\
\Gamma_{\pm 4} : \quad & x_1 := t^{-1}, \quad & x_2 := t + t^{-1} - t^{-3} + \cdots,
\end{aligned}
$$

for a parameter $t \to +\infty$ or $t \to -\infty$. As in [45], one can compute the asymptotic value of f in each component $\Gamma_{\pm i}$. Substituting the parametrization in f, we get

$$
\begin{aligned}
f|_{\Gamma_{\pm 1}} &= t^4 + 4t^2 + 2 - \tfrac{23}{8}t^{-2} + \cdots, \quad & f|_{\Gamma_{\pm 2}} &= t^4 + 2 + \tfrac{5}{8}t^{-2} + \cdots \\
f|_{\Gamma_{\pm 3}} &= t^{-2} - t^{-4} + t^{-6} + \cdots, \quad & f|_{\Gamma_{\pm 4}} &= t^2 + 2 - t^{-2} - t^{-4} + \cdots.
\end{aligned}
$$

The asymptotic values on these components are

$$\lambda_{\pm 1} = \lambda_{\pm 2} = \lambda_{\pm 4} = +\infty, \quad \lambda_{\pm 3} = 0.$$

Therefore, the minimum value $f_{\min} = \min\limits_{k=1,2,3,4} \lambda_{\pm k} = 0$. For this example, the minimizers at infinity are $(\pm 1, 0)$ and $(0, \pm 1)$. In the computation, if minimizers at infinity are obtained, we do not need to consider the components $\Gamma_{\pm 1}$, $\Gamma_{\pm 2}$, since their normalizations do not converge to minimizers at infinity. We refer to [11, 12, 53] for related work for the case that $f_{\min}$ is not achievable.

8. Conclusions and discussions

This paper gives a Moment-SOS hierarchy for polynomial optimization with unbounded sets, based on homogenization. We prove this hierarchy has finite convergence under some assumptions on optimality conditions about minimizers. We also extend the Putinar-Vasilescu type Positivstellensatz to polynomials that are nonnegative on unbounded sets. The classical Moment-SOS hierarchy with denominators is also studied. Moreover, we give a positive answer to a conjecture of Mai, Lasserre and Magron in their recent work [32].

To prove the convergence of the hierarchy of (3.6)-(3.7), we made the assumption that the ideal Ideal$[\tilde{c}_{\mathcal{E}}]$ is real radical. We remark that Ideal$[\tilde{c}_{\mathcal{E}}]$ being real radical is a general condition (see [58]). When Ideal$[\tilde{c}_{\mathcal{E}}]$ is not real radical, we do not know if the hierarchy of (3.6)-(3.7) still has the finite convergence, under the remaining assumptions as in Theorem 4.3. However, we have the following result when Ideal$[\tilde{c}_{\mathcal{E}}]$ is not real radical.

Theorem 8.1. *Assume K is closed at ∞. If the LICQC, SCC and SOSC hold at every minimizer of (1.1), including the one at infinity, then there exists an integer $k_0 > 0$ such that*

$$(8.1) \qquad x_0^{2\ell}(\tilde{f} - (f_{\min} - \epsilon) \cdot x_0^d) \in Ideal[\tilde{c}_{\mathcal{E}}] + QM[\tilde{c}_{\mathcal{I}}]$$

for every $\epsilon > 0$ and for all $\ell \geq k_0$.

Proof. As in Theorem 4.3, there exists $\sigma \in QM[\tilde{c}_{\mathcal{I}}]$ such that

$$\tilde{f} - f_{\min}x_0^d \equiv \sigma \mod Ideal[V_{\mathbb{R}}(\tilde{c}_{\mathcal{E}})].$$

Let $\bar{f} = \tilde{f} - f_{\min} \cdot x_0^d - \sigma$. Then $\bar{f}$ vanishes identically on the variety $V_{\mathbb{R}}(\tilde{c}_{\mathcal{E}})$. By the Real Nullstellensatz (see [29]), there exist $k_1 \in \mathbb{N}$, $\sigma_1 \in \Sigma[\tilde{x}]$ such that $\bar{f}^{2k_1} + \sigma_1 \in Ideal[\tilde{c}_{\mathcal{E}}]$. Let $\omega > 0$ be big enough such that the univariate polynomial $s(t) := 1 + t + \omega t^{2k_1}$ is SOS (see [40]). For each $\epsilon > 0$, we get

$$(\epsilon x_0^d)^{2k_1} s\left(\frac{\bar{f}}{\epsilon x_0^d}\right) = (\epsilon x_0^d)^{2k_1 - 1}(\bar{f} + \epsilon x_0^d) + \omega \bar{f}^{2k_1}.$$

This implies that

$$x_0^{(2k_1-1)d}\left(\tilde{f} - (f_{\min} - \epsilon)x_0^d - \sigma\right) = (\epsilon x_0^d)^{2k_1} s\left(\frac{\bar{f}}{\epsilon x_0^d}\right) - \omega \bar{f}^{2k_1}.$$

Since $x_0 \in \tilde{c}_{\mathcal{I}}$ and σ_1 is SOS, we have

$$(\epsilon x_0^d)^{2k_1} s\left(\frac{\bar{f}}{\epsilon x_0^d}\right) - \omega \bar{f}^{2k_1} \in \Sigma[\tilde{x}] + Ideal[\tilde{c}_{\mathcal{E}}] \subseteq Ideal[\tilde{c}_{\mathcal{E}}] + QM[\tilde{c}_{\mathcal{I}}],$$

$$x_0\left((\epsilon x_0^d)^{2k_1} s\left(\frac{\bar{f}}{\epsilon x_0^d}\right) - \omega \bar{f}^{2k_1}\right) \in x_0 \Sigma[\tilde{x}] + Ideal[\tilde{c}_{\mathcal{E}}] \subseteq Ideal[\tilde{c}_{\mathcal{E}}] + QM[\tilde{c}_{\mathcal{I}}].$$

Let $k_0 = \lceil \frac{(2k_1-1)d}{2} \rceil$. Then, we have

$$x_0^{2k_0}(\tilde{f} - (f_{\min} - \epsilon)x_0^d) = x_0^{2k_0}\sigma + x_0^{2k_0 - (2k_1-1)d}\left((\epsilon x_0^d)^{2k_1} s\left(\frac{\bar{f}}{\epsilon x_0^d}\right) - \omega \bar{f}^{2k_1}\right).$$

It implies that

$$x_0^{2k_0}(\tilde{f} - (f_{\min} - \epsilon)x_0^d) \in Ideal[\tilde{c}_{\mathcal{E}}] + QM[\tilde{c}_{\mathcal{I}}].$$

Note that k_0 is independent of $\epsilon > 0$. The above implies that (8.1) holds for all $\epsilon > 0$ and for all $\ell \geq k_0$. $\qquad\square$

For a degree $\ell \geq k_0$, (8.1) motivates the hierarchy of the following relaxations

$$(8.2) \qquad \begin{cases} \max & \gamma \\ s.t. & x_0^{2\ell}(\tilde{f}(\tilde{x}) - \gamma x_0^d) \in Ideal[\tilde{c}_{\mathcal{E}}]_{2k} + QM[\tilde{c}_{\mathcal{I}}]_{2k}. \end{cases}$$

Theorem 8.1 shows that the optimal value of relaxations (8.2) has the finite convergence to $f_{\min}$, even if $Ideal[\tilde{c}_{\mathcal{I}}]$ is not real radical, under the remaining assumptions. Because of this result, we make the following conjecture.

Conjecture 8.2. *When the ideal $Ideal[\tilde{c}_{\mathcal{E}}]$ is not real radical, the hierarchy of relaxations (3.6)-(3.7) is also tight, i.e., $f_k = f'_k = f_{\min}$, for all k big enough, under the other assumptions of Theorem 4.3.*

When the minimum value $f_{\min}$ is not achievable, the Moment-SOS hierarchy of relaxations (3.6)-(3.7) also has finite convergence under some assumptions on optimality conditions for minimizers at infinity. However, there are numerical issues for solving the hierarchy, since there are no optimizers satisfying flat truncation for the moment relaxations. It is an interesting future work to get numerically stable Moment-SOS relaxations for computing $f_{\min}$ when it is not achievable.

References

[1] Ahmadi, A.A., Zhang, J.: On the complexity of testing attainment of the optimal value in nonlinear optimization. *Math. Program.* 184, 221–241 (2020)

[2] Bajbar, T., Stein, O.: Coercive polynomials and their newton polytopes. *SIAM J. Optim.* 25(3), 1542–1570 (2015)

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

[4] Bertsekas, D.: Nonlinear Programming, 2nd edn. Athena Scientific, Belmont (1995)

[5] Demmel, J., Nie, J., Powers, V.: Representations of positive polynomials on non-compact semialgebraic sets via KKT ideals. *J. Pure Appl. Algebra* 209(1), 189–200 (2007)

[6] Dickinson, P., Povh, J.: A new approximation hierarchy for polynomial conic optimization. *Comput. Optim. Appl.* 73, 37–67 (2019)

[7] Fan, J., Nie, J., Zhou, A.: Tensor eigenvalue complementarity problems. *Math. Program.* 170 (2), 507–539 (2018)

[8] Fang, K., Fawzi, H.: The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Math. Program.* 190, 331-–360 (2021)

[9] Gelfand, I., Kapranov, M., Zelevinsky, A.: Discriminants, resultants, and multidimensional determinants. Mathematics: Theory & applications, Birkhäuser (1994)

[10] Guo, F., Wang, L., Zhou, G.: Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities. *J. Global Optim.* 58(2), 261–284 (2014)

[11] Ha, H., Pham, T.: Solving polynomial optimization problems via the truncated tangency variety and sums of squares. *J. Pure Appl. Algebra* 213(11), 2167–2176 (2009)

[12] Ha, H., Pham, T.: Global optimization of polynomials using the truncated tangency variety and sums of squares. *SIAM J. Optim.* 19(2), 941–951 (2008)

[13] Henrion, D., Lasserre, J.: Detecting global optimality and extracting solutions in Gloptipoly. In *Positive polynomials in control*, 293.C310, Lecture Notes in Control and Inform. Sci., 312, Springer (2005)

[14] Henrion, D., Lasserre, J., Loefberg, J.: Gloptipoly 3: moments, optimization and semidefinite programming. *Optim. Methods Softw.* 24(4-5), 761–779 (2009)

[15] Jeyakumar, V., Lasserre, J., Li, G.: On polynomial optimization over non-compact semialgebraic sets. *J. Optim. Theory Appl.* 163(3), 707–718 (2014)

[16] De Klerk, E., Laurent, M.: On the Lasserre hierarchy of semidefinite programming relaxations of convex polynomial optimization problems. *SIAM J. Optim.* 21, 824–832 (2011)

[17] De Klerk E., Laurent M., Parrilo P.: On the equivalence of algebraic approaches to the minimization of forms on the simplex. In: *Henrion D., Garulli A. (eds) Positive Polynomials in Control. Lecture Notes in Control and Information Science,* vol 312. Springer (2005)

[18] De Klerk, E., Lasserre, J., Laurent, M., Sun, Z.: Bound-constrained polynomial optimization using only elementary calculations. *Math. Oper. Res.* 42(3), 834–853 (2017)

[19] De Klerk, E., Laurent, M.: Convergence analysis of a Lasserre hierarchy of upper bounds for polynomial minimization on the sphere. *Math. Program.* 193(2), 665–685 (2022)

[20] De Klerk, E., Pasechnik, D.: Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.* 12(4), 875–892 (2002)

[21] Lasserre, J.: Global optimization with polynomials and the problem of moments. *SIAM J. Optim.* 11(3), 796–817 (2001)

[22] Lasserre, J., Laurent, M., Rostalski, P.: Semidefinite characterization and computation of zero-dimensional real radical ideals. *Found. Comput. Math.* 8, 607–647 (2008)

[23] Lasserre, J.: Convexity in semi-algebraic geometry and polynomial optimization. *SIAM J. Optim.* 19, 1995–2014 (2009)

[24] Lasserre, J.: A new look at nonnegativity on closed sets and polynomial optimization. *SIAM J. Optim.* 21, 864–885 (2011)

[25] Lasserre, J.: An Introduction to Polynomial and Semi-algebraic Optimization. Cambridge University Press, Cambridge (2015)

[26] Lasserre,J.: The Moment-SOS hierarchy. In: Sirakov, B., Ney de Souza., P., Viana, M. (eds.) *Proceedings of the International Congress of Mathematicians (ICM 2018)*, vol 3, pp. 3761–3784, World Scientific (2019)

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

[28] Laurent, M.: Semidefinite representations for finite varieties. *Math. Program.* 109, 1–26 (2007)

[29] Laurent, M.: Sums of squares, moment matrices and optimization over polynomials. In: *Emerging applications of algebraic geometry of IMA Volumes in Mathematics and its Applications*, 149: 157–270. Springer (2009)

[30] Laurent, M.: Optimization over polynomials: Selected topics. In: Jang, S., Kim, Y., Lee, D.W., Yie, I. (eds.) *Proceedings of the International Congress of Mathematicians (ICM 2014)*, pp. 843–869 (2014)

[31] Mai, N., Magron, V.: On the complexity of Putinar-Vasilescu's Positivstellensatz. arXiv preprint, `arXiv:2104.11606` (2021)

[32] Mai, N., Lasserre, J., Magron, V.: Positivity certificates and polynomial optimization on non-compact semialgebraic sets. *Math. Program.* (2021)

[33] Marshall, M.: Optimization of polynomial functions. *Can. Math. Bull.* 46(3), 400–418 (2003)

[34] Marshall, M.: Positive Polynomials and Sums of Squares. American Mathematical Society, Providence (2008)

[35] Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Turán. *Can. J. Math.* 17, 533–540 (1965)

[36] Nie, J., Demmel, J., Sturmfels, B.: Minimizing polynomials via sum of squares over the gradient ideal. *Math. Program.* 106(3), 587–606 (2006)

[37] Nie, J.: Discriminants and nonnegative polynomials. *J. Symb. Comput.* 47(2), 167–191, (2012)

[38] Nie, J.: Certifying convergence of Lasserre's hierarchy via flat truncation. *Math. Program.* 142(1-2), 485–510 (2013)

[39] Nie, J.: An exact Jacobian SDP relaxation for polynomial optimization. *Math. Program.* 137(1-2), 225–255 (2013)

[40] Nie, J.: Polynomial optimization with real varieties. *SIAM J. Optim.* 23(3), 1634–1646 (2013)

[41] Nie, J.: Optimality conditions and finite convergence of Lasserre's hierarchy. *Math. Program.* 146(1-2), 97–121 (2014)

[42] Nie, J.: The hierarchy of local minimums in polynomial optimization. *Math. Program.* 151 (2) 555–583 (2015)

[43] Nie, J.: Tight relaxations for polynomial optimization and Lagrange multiplier expressions. *Math. Program.* 178(1–2), 1–37 (2019)

[44] Nie, J., Yang, Z., Zhang, X.: A complete semidefinite algorithm for detecting copositive matrices and tensors. *SIAM J. Optim.* 28(4), 2902–2921 (2018)

[45] Pham, T.: Tangencies and polynomial optimization. arXiv preprint, `arXiv:1902.06041` (2019)

[46] Putinar, M.: Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* 42(3), 969–984 (1993)

[47] Putinar, M., Vasilescu, F: Positive polynomials on semi-algebraic sets. *C. R. Acad. Sci. Ser. I Math.* 328(7), 585–589 (1999)

[48] Putinar, M., Vasilescu, F: Solving moment problems by dimensional extension. *C. R. Acad. Sci. Ser. I Math.* 328(6), 495–499 (1999)

[49] Reznick, B.: Some concrete aspects of Hilbert's 17th problem. *Contemp. Math.* 253, 251–272 (2000)

[50] Scheiderer, C.: Sums of squares on real algebraic surfaces. *Manuscr. Math.* 119, 395–410 (2006)

[51] Scheiderer, C.: Positivity and sums of squares: A guide to recent results. In: *Emerging applications of algebraic geometry of IMA Volumes in Mathematics and its Applications*, 149, pp. 271–324. Springer (2009)

[52] Schweighofer, M.: Optimization of polynomials on compact semialgebraic sets. *SIAM J. Optim.* 15 (3), 805–825 (2005)

[53] Schweighofer, M.: Global optimization of polynomials using gradient tentacles and sums of squares. *SIAM J. Optim.* 17(3), 920–942 (2006)

[54] Slot, L., Laurent, M.: Improved convergence analysis of Lasserre's measure-based upper bounds for polynomial minimization on compact sets. *Math. Program.* 193(2), 831–871 (2022)

[55] Sun, W., Yuan, Y.: Optimization Theory and Methods: Nonlinear Programming. Springer, New York (2006)

[56] Sturm, J.F.: SeDuMi 1.02: a matlab toolbox for optimization over symmetric cones. *Optim. Methods Softw.* 11(1-4), 625–653 (1999)

[57] Sturmfels, B.: Solving systems of polynomial equations. In: *CBMS Regional Conference Series in Mathematics.* American Mathematical Society, Providence (2002)

[58] Yu, J.: Do most polynomials generate a prime ideal? *J. Algebra* 459, 468–474 (2016)

LEI HUANG, INSTITUTE OF COMPUTATIONAL MATHEMATICS AND SCIENTIFIC/ENGINEERING COMPUTING, ACADEMY OF MATHEMATICS AND SYSTEMS SCIENCE, CHINESE ACADEMY OF SCIENCES, AND SCHOOL OF MATHEMATICAL SCIENCES, UNIVERSITY OF CHINESE ACADEMY OF SCIENCES, BEIJING, CHINA, 100190.
Email address: `huanglei@lsec.cc.ac.cn`

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, 100190.
Email address: `yyx@lsec.cc.ac.cn`