

# The Multi-Objective Polynomial Optimization

**Jiawang Nie,<sup>a,\*</sup> Zi Yang<sup>b</sup>**
<sup>a</sup>Department of Mathematics, University of California, San Diego, La Jolla, California 92093; <sup>b</sup>Department of Mathematics and Statistics, University at Albany, State University of New York, Albany, New York 12222

<sup>\*</sup>Corresponding author

**Contact:** njw@math.ucsd.edu,  <https://orcid.org/0000-0003-2637-8274> (JN); zyang8@albany.edu,  <https://orcid.org/0000-0003-1048-9869> (ZY)

**Received:** January 21, 2022

**Revised:** September 6, 2022; June 30, 2023

**Accepted:** November 19, 2023

**Published Online in Articles in Advance:** January 4, 2024

**MSC2020 Subject Classifications:** Primary: 90C23, 90C29, 90C22

<https://doi.org/10.1287/moor.2023.0200>
**Copyright:** © 2024 INFORMS

**Abstract.** The multi-objective optimization is to optimize several objective functions over a common feasible set. Because the objectives usually do not share a common optimizer, people often consider (weakly) Pareto points. This paper studies multi-objective optimization problems that are given by polynomial functions. First, we study the geometry for (weakly) Pareto values and represent Pareto front as the boundary of a convex set. Linear scalarization problems (LSPs) and Chebyshev scalarization problems (CSPs) are typical approaches for getting (weakly) Pareto points. For LSPs, we show how to use tight relaxations to solve them and how to detect existence or nonexistence of proper weights. For CSPs, we show how to solve them by moment relaxations. Moreover, we show how to check whether a given point is a (weakly) Pareto point or not and how to detect existence or nonexistence of (weakly) Pareto points. We also study how to detect unboundedness of polynomial optimization, which is used to detect nonexistence of proper weights or (weakly) Pareto points.

**Funding:** J. Nie is partially supported by the National Science Foundation [Grant DMS-2110780].

**Keywords:** Pareto point • Pareto value • polynomial • scalarization • moment relaxation

## 1. Introduction

The multi-objective optimization problem (MOP) is to optimize several objectives simultaneously over a common feasible set. MOPs have broad applications in economics (Geiger and Sevaux [15]), finance (Chen et al. [6]), medical science (Rosenthal and Borschbach [56], van der Horst et al. [58]), and machine learning (Wang et al. [59]). In this paper, we consider the MOP in the form

$$\begin{cases} \min & f(x) := (f_1(x), \dots, f_m(x)) \\ \text{s.t.} & c_i(x) = 0 \ (i \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}), \end{cases} \quad (1.1)$$

where all functions  $f_i, c_i, c_j$  are polynomials in  $x := (x_1, \dots, x_n) \in \mathbb{R}^n$ . The  $\mathcal{E}$  and  $\mathcal{I}$  are disjoint finite label sets. Let  $K$  denote the feasible set of (1.1). Generally, there does not exist a point such that all  $f_i$ 's are minimized simultaneously. People often look for a point such that some or all of the objectives cannot be further optimized. This leads to the following concepts (see Marler and Arora [43], Miettinen [45], and Jahn et al. [22]).

**Definition 1.1.** A point  $x^* \in K$  is said to be a Pareto point (PP) if there is no  $x \in K$  such that  $f_i(x) \leq f_i(x^*)$  for all  $i = 1, \dots, m$  and  $f_j(x) < f_j(x^*)$  for at least one  $j$ . The point  $x^*$  is said to be a weakly Pareto point (WPP) if there is no  $x \in K$  such that  $f_i(x) < f_i(x^*)$  for all  $i = 1, \dots, m$ .

In the literature, Pareto points (respectively (resp.), weakly Pareto points) are also referenced as Pareto optimizers (resp., weakly Pareto optimizers) or Pareto solutions (resp., weakly Pareto solutions). A vector  $v := (v_1, \dots, v_m)$  is called a *Pareto value* (resp., weakly Pareto value) for (1.1) if there exists a Pareto point (resp., weakly Pareto point)  $x^*$  such that  $v = f(x^*)$ . *Pareto front* is the set of objective values at Pareto points. Every Pareto point is a weakly Pareto point, whereas the converse is not necessarily true. Detecting existence or nonexistence of (weakly) Pareto points is a major task for MOPs. We refer to Bao and Mordukhovich [2], Bao and Mordukhovich [3], Kim et al. [27], Marler and Arora [43], Miettinen [45], and Jahn et al. [22] for related work about existence of PPs and WPPs.

Scalarization is a classical method for finding PPs or WPPs. It transforms a MOP into a single objective optimization problem. A frequently used scalarization is a nonnegative linear combination of objectives.

**Definition 1.2.** The linear scalarization problem (LSP) for the MOP (1.1), with a nonzero weight  $w := (w_1, \dots, w_m) \geq 0$ , is

$$\begin{aligned} & \min w_1 f_1(x) + \dots + w_m f_m(x) \\ & \text{s.t. } x \in K. \end{aligned} \tag{1.2}$$

For the LSP (1.2), the optimization remains unchanged if we normalize the nonzero weight  $w$  such that  $\sum_{i=1}^m w_i = 1$ ,  $w_i \geq 0$ . For neatness of the paper, one can equivalently consider nonzero and nonnegative weights for LSPs. Every minimizer of the LSP (1.2) is a weakly Pareto point for nonzero  $w \geq 0$ , and every minimizer is a Pareto point for  $w > 0$ . Varying weights in (1.2) may give different (weakly) Pareto points. A nonzero weight  $w$  is said to be *proper* if the LSP (1.2) is bounded below. Otherwise, the  $w$  is called *improper*. One wonders whether every Pareto point is a minimizer of (1.2) for some weight  $w$ . However, this is sometimes not the case (see Fleming [14] and Zonts [60]). For instance, Example 4.4 has infinitely many Pareto points, but only two of them can be obtained by solving LSPs. Under some assumptions, LSPs may give all Pareto points (see Emmerich and Deutz [12]).

Another frequently used scalarization is the Chebyshev scalarization. It requires one to use the minimum value of each objective.

**Definition 1.3.** The Chebyshev scalarization problem (CSP) for the MOP (1.1), with a nonzero weight  $w = (w_1, \dots, w_m) \geq 0$ , is

$$\begin{aligned} & \min \max_{1 \leq i \leq m} w_i (f_i(x) - f_i^*) \\ & \text{s.t. } x \in K, \end{aligned} \tag{1.3}$$

where the minimum value  $f_i^* := \min_{x \in K} f_i(x) > -\infty$ .

Every minimizer of the CSP (1.3) is a weakly Pareto point. Interestingly, every weakly Pareto point is the minimizer of a CSP for some weight (see Koski and Silvennoinen [28] and Miettinen [45]). However, the minimizer of a CSP may not be a Pareto point. There also exist other scalarization methods, such as the  $\epsilon$ -constraint method (Anagnostopoulos et al. [1], Matsatsinis and Delias [44]) and the lexicographic method (Clayton et al. [8], Jones and Tamiz [26]). We refer to Cho et al. [7], Donoso and Fabregat [11], Marler and Arora [43], Miettinen [45], and Ruiz-Canales and Rufián-Lizana [57] for different scalarizations.

There exists important work for MOPs given by polynomials. When all functions are linear, a semidefinite programming method is given to obtain the set of Pareto points in Blanco et al. [5]. When the functions are convex polynomials, Moment- sum of squares (SOS) relaxation methods are given to compute (weakly) Pareto points in Jiao et al. [25], Jiao et al. [24], Jiao and Lee [23], Lee et al. [38], and Lee and Jiao [37], as well as some useful conditions for existence of (weakly) Pareto points. Because the Pareto front is an image set of polynomial functions, semidefinite relaxations can be used to approximate the Pareto front, as in the works Magron et al. [40] and Magron et al. [41].

When the functions are nonconvex polynomials, nonemptiness and boundedness of Pareto solution sets are shown in Liu et al. [39], under certain regularity conditions. When the objectives are polynomials and  $K$  is the entire space  $\mathbb{R}^n$ , some novel conditions are shown for existence of (weakly) Pareto points in Kim et al. [27]. The following questions are of great interest for studying MOPs:

- What is a convenient description for the set of (weakly) Pareto values? How can we represent the Pareto front in a geometrically clean way?
- For an LSP, how can we solve it efficiently for a Pareto point? When the constraint  $K$  is unbounded, how can we find a proper weight such that the LSP is bounded? How can we detect nonexistence of proper weights?
- For a CSP, how can we solve it efficiently for a weakly Pareto point? How do we get the global minimum value for each objective? If some minimum value is  $-\infty$ , how can we get a weakly Pareto point?
- For a given point, how can we detect whether it is a (weakly) Pareto point? How can we get a (weakly) Pareto point if LSPs/CSPs fail to give one? How do we detect nonexistence of (weakly) Pareto points?

## 1.1. Contributions

The above questions are the major topics of this paper. When MOPs are given by polynomials, there are special properties for them. The following are our major contributions.

We study the convex geometry for (weakly) Pareto values. The epigraph set—that is, the set  $\mathcal{U}$  as in (3.1)—is useful for (weakly) Pareto values. We give a characterization for the Pareto front. When the objectives are convex, we show that the set of weakly Pareto values can be expressed in terms of the boundary of a convex set. When the

MOP is given by SOS convex polynomials, we show that  $\mathcal{U}$  can be given by semidefinite representations. This is shown in Section 3.

For solving LSPs and CSPs, or detecting nonexistence of (weakly) Pareto points, we often need to detect whether an optimization problem is unbounded. There exist few works for detecting unboundedness in nonconvex optimization. We give a convex relaxation method for detecting unboundedness in polynomial optimization under some genericity assumptions. To the best of the authors' knowledge, this is the first work that can achieve this goal. The results are in the appendix.

We discuss how to solve LSPs in Section 4. Under a genericity assumption, we give a tight relaxation method for solving LSPs and obtaining Pareto points. When the feasible set  $K$  is unbounded, we show how to find proper weights such that the LSP is bounded below. We also show how to detect that the LSP is unbounded below for all weights—that is, how to detect nonexistence of proper weights.

Section 5 studies how to solve CSPs. We first apply the tight relaxation method to compute global minimum values  $f_1^*, \dots, f_m^*$  for the individual objectives. After that, we formulate the CSP equivalently as a polynomial optimization problem and then solve it by using Moment-SOS relaxations.

Section 6 discusses how to detect whether a given point is a (weakly) Pareto point or not. This can be done by solving certain polynomial optimization. We also show how to detect existence or nonexistence of (weakly) Pareto points. This requires one to solve some moment feasibility problems.

We make some conclusions and propose some open questions in Section 7. Section 2 reviews some basic results for optimization with polynomials and moments.

## 2. Preliminary

### 2.1. Notation

The symbol  $\mathbb{N}$  (resp.,  $\mathbb{R}$ ,  $\mathbb{C}$ ) denotes the set of nonnegative integral (resp., real, complex) numbers. The  $\mathbb{R}_+^n$  stands for the nonnegative orthant—that is, the set of nonnegative vectors. For each label  $i$ , the  $e_i$  denotes the vector of all zeros excepts its  $i$ th entry being one, whereas  $e$  denotes the vector of all ones. For an integer  $k > 0$ , denote  $[k] := \{1, 2, \dots, k\}$ . For  $t \in \mathbb{R}$ ,  $\lceil t \rceil$  denotes the smallest integer greater than or equal to  $t$ . Denote by  $\mathbb{R}[x] := \mathbb{R}[x_1, \dots, x_n]$  the ring of polynomials in  $x := (x_1, \dots, x_n)$  with real coefficients. The  $\mathbb{R}[x]_d$  stands for the set of polynomials in  $\mathbb{R}[x]$  with degrees at most  $d$ . For a polynomial  $p$ ,  $\deg(p)$  denotes its total degree,  $\tilde{p}$  denotes its homogenization, and  $p^{hom}$  denotes the homogeneous part of the highest degree. For  $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ , we denote  $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  and  $|\alpha| := \alpha_1 + \cdots + \alpha_n$ . The power set of degree  $d$  is

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

The vector of monomials in  $x$  and up to degree  $d$  is

$$[x]_d := [1 \ x_1 \ \cdots \ x_n \ x_1^2 \ x_1x_2 \ \cdots \ x_n^d]^T.$$

The superscript  $T$  denotes the transpose of a matrix/vector. The  $I_N$  stands for the  $N$ -by- $N$  identity matrix. By writing  $X \succeq 0$  (resp.,  $X > 0$ ), we mean that  $X$  is a symmetric positive semidefinite (resp., positive definite) matrix. For a set  $T$ ,  $\text{conv}(T)$  denotes its convex hull,  $cl(T)$  denotes its closure, and  $int(T)$  denotes its interior, under the Euclidean topology. The cardinality of  $T$  is  $|T|$ . For a vector  $u$ , the  $\|u\|$  denotes its standard Euclidean norm. For a function  $h$  in  $x$ , the  $\nabla h$  denotes its gradient vector in  $x$ . All computational results are shown with four decimal digits.

### 2.2. Positive Polynomials

A subset  $I \subseteq \mathbb{R}[x]$  is an ideal if  $I \cdot \mathbb{R}[x] \subseteq I$  and  $I + I \subseteq I$ . For a tuple  $p = (p_1, \dots, p_k)$  of polynomials in  $\mathbb{R}[x]$ ,  $\text{Ideal}(p)$  denotes the smallest ideal containing all  $p_i$ , which is the set  $p_1 \cdot \mathbb{R}[x] + \cdots + p_k \cdot \mathbb{R}[x]$ . In computation, we often need to work with the *truncation* of degree  $2k$ :

$$\text{Ideal}[p]_{2k} := p_1 \cdot \mathbb{R}[x]_{2k-\deg(p_1)} + \cdots + p_k \cdot \mathbb{R}[x]_{2k-\deg(p_k)}.$$

A polynomial  $\sigma$  is said to be a sum of squares if  $\sigma = s_1^2 + \cdots + s_k^2$  for some polynomials  $s_1, \dots, s_k$ . Checking whether a polynomial is SOS can be done by solving a semidefinite program (SDP) (Lasserre [29]). If a polynomial is SOS, then it is nonnegative everywhere. The set of all SOS polynomials in  $x$  is denoted by  $\Sigma[x]$ , and its  $d$ th truncation is  $\Sigma[x]_d := \Sigma[x] \cap \mathbb{R}[x]_d$ . For a tuple  $q = (q_1, \dots, q_t)$  of polynomials, its *quadratic module* is

$$Qmod[q] := \Sigma[x] + q_1 \cdot \Sigma[x] + \cdots + q_t \cdot \Sigma[x].$$

The truncation of degree  $2k$  for  $Qmod[q]$  is

$$Qmod[q]_{2k} := \Sigma[x]_{2k} + q_1 \cdot \Sigma[x]_{2k-\deg(g_1)} + \cdots + q_t \cdot \Sigma[x]_{2k-\deg(g_t)}.$$

A subset  $A \subseteq \mathbb{R}[x]$  is said to be *archimedean* if there exists  $\sigma \in A$  such that  $\sigma(x) \geq 0$  defines a compact set in  $\mathbb{R}^n$ . If  $Ideal[p] + Qmod[q]$  is archimedean, then the set  $T := \{x \in \mathbb{R}^n : p(x) = 0, q(x) \geq 0\}$  must be compact. The reverse is not necessarily true. However, if  $T$  is compact, the archimedeaness can be met by adding a redundant ball condition. When  $Ideal[p] + Qmod[q]$  is archimedean, every polynomial that is positive on  $T$  must belong to  $Ideal[p] + Qmod[q]$ . This conclusion is referenced as Putinar's Positivstellensatz (Putinar [55]). Furthermore, if a polynomial is nonnegative on  $T$ , then it also belongs to  $Ideal[p] + Qmod[q]$ , under some standard optimality conditions on its minimizers (see Nie [49]).

### 2.3. Localizing and Moment Matrices

Denote by  $\mathbb{R}^{\mathbb{N}_d^n}$  the space of real sequences labeled by  $\alpha \in \mathbb{N}_d^n$ . A vector  $y := (y_\alpha)_{\alpha \in \mathbb{N}_d^n}$  is called a *truncated multisequence* (tms) of degree  $d$ . It gives a linear functional on  $\mathbb{R}[x]_d$  such as

$$\left\langle \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha x^\alpha, y \right\rangle := \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha y_\alpha, \quad (2.1)$$

where each  $f_\alpha$  is a coefficient. The tms  $y$  is said to *admit a Borel measure*  $\mu$  if  $y_\alpha = \int x^\alpha d\mu$  for all  $\alpha \in \mathbb{N}_d^n$ . If it exists, such  $\mu$  is called a *representing measure* for  $y$ , and  $y$  is said to admit the measure  $\mu$ . The support of  $\mu$  is denoted as  $\text{supp}(\mu)$ . If the cardinality  $|\text{supp}(\mu)|$  is finite, the measure  $\mu$  is called *finitely atomic*. It is called *r-atomic* if  $|\text{supp}(\mu)| = r$ .

In optimization, the support of  $\mu$  is often constrained in a set  $K$ . For a degree  $d$ , denote the moment cone

$$\mathcal{R}_d(K) := \left\{ y \in \mathbb{R}^{\mathbb{N}_d^n} : \exists \mu, y = \int [x]_d d\mu, \text{supp}(\mu) \subseteq K \right\}. \quad (2.2)$$

The dual cone of  $\mathcal{R}_d(K)$  is the nonnegative polynomial cone

$$\mathcal{P}_d(K) := \{p \in \mathbb{R}[x]_d : p(x) \geq 0 \ \forall x \in K\}. \quad (2.3)$$

The dual cone of  $\mathcal{P}_d(K)$  is the closure of  $\mathcal{R}_d(K)$ . When  $K$  is compact, the moment cone  $\mathcal{R}_d(K)$  is closed. We refer to Lasserre [32] and Laurent [35] for more details about moment cones.

Consider a polynomial  $q \in \mathbb{R}[x]_{2k}$  with  $\deg(q) \leq 2k$ . The  $k$ th *localizing matrix* of  $q$ , generated by a tms  $z \in \mathbb{R}^{\mathbb{N}_{2k}^n}$ , is the symmetric matrix  $L_q^{(k)}[z]$  such that

$$\text{vec}(a_1)^T (L_q^{(k)}[z]) \text{vec}(a_2) = \langle qa_1 a_2, z \rangle, \quad (2.4)$$

for all  $a_1, a_2 \in \mathbb{R}[x]_{k-\lceil \deg(q)/2 \rceil}$ . (The  $\text{vec}(a_i)$  denotes the coefficient vector of  $a_i$ .) When  $q=1$ ,  $L_q^{(k)}[z]$  is called a *moment matrix*, and we denote

$$M_k[z] := L_1^{(k)}[z].$$

The columns and rows of  $L_q^{(k)}[z]$ , as well as  $M_k[z]$ , are labeled by  $\alpha \in \mathbb{N}^n$  with  $2|\alpha| \leq 2k - \deg(q)$ .

Each  $y \in \mathcal{R}_d(K)$  can be extended to a tms  $z \in \mathcal{R}_{2t}(K)$  such that  $y = z|_d$ , where  $d \leq 2t$  and  $z|_d$  denotes the truncation of  $z$  with degree  $d$ :

$$z|_d := (z_\alpha)_{\alpha \in \mathbb{N}_d^n}. \quad (2.5)$$

When  $K$  is the feasible set of (1.1), a necessary condition for  $z \in \mathcal{R}_{2t}(K)$  is

$$L_{c_i}^{(t)}[z] = 0 (i \in \mathcal{E}), \quad L_{c_j}^{(t)}[z] \succeq 0 (j \in \mathcal{I}),$$

whereas they may not be sufficient (see Lasserre [32] and Laurent [35]). However, if  $z$  further satisfies

$$\text{rank } M_{t-d_c}[z] = \text{rank } M_t[z], \quad (2.6)$$

then  $z$  admits an  $r$ -atomic measure supported in  $K$ , with  $r = \text{rank } M_t[z]$ . The above integer  $d_c$  is the degree

$$d_c := \max\{\lceil \deg(c_i)/2 \rceil : i \in \mathcal{E} \cup \mathcal{I}\}. \quad (2.7)$$

This Condition (2.6) is called flat extension (see Curto and Fialkow [9], Curto and Fialkow [10], Henrion and Lasserre [18], and Laurent [34]). To get optimizers in computation, the flat truncation is more frequently used (see Nie [48]).

Moment and localizing matrices are important tools for solving polynomial optimization (Fan et al. [13], Henrion and Lasserre [18], Lasserre [29], and Nie [46]). They are also useful in tensor decompositions (Nie [51], Nie and Zhang [53]). We refer to Lasserre [32], Lasserre [33], Laurent [35], and Laurent [36] for the books and surveys about polynomial and moment optimization.

### 3. Geometry of Pareto Values

Recall that a vector  $v := (v_1, \dots, v_m)$  is a *Pareto value* (PV) if there exists a Pareto point  $x^*$  such that  $v = f(x^*)$ . Similarly,  $v$  is called a weakly Pareto value (WPV) if  $v = f(p)$  for a weakly Pareto point  $p$ . PVs and WPVs are closely related to the epigraph set

$$\mathcal{U} := \{u = (u_1, \dots, u_m) \mid u_i \geq f_i(x), \text{ for some } x \in K\}. \quad (3.1)$$

The image of the set  $K$  under the objective vector  $f = (f_1, \dots, f_m)$  is

$$f(K) := \{(f_1(x), \dots, f_m(x)) : x \in K\}.$$

Then,  $\mathcal{U} = f(K) + \mathbb{R}_+^m$  and its convex hull  $\text{conv}(\mathcal{U}) = \text{conv}(f(K)) + \mathbb{R}_+^m$ . If  $K$  is convex and each objective  $f_i$  is convex, the set  $\mathcal{U}$  is also convex. The converse is not necessarily true. When  $\mathcal{U}$  is convex, every Pareto point is a minimizer of some LSP (see Emmerich and Deutz [12]). In this section, we study the geometry of PVs and discuss how to characterize PVs and WPVs through the set  $\mathcal{U}$ .

#### 3.1. Supporting Hyperplanes

For a nonzero vector  $w \in \mathbb{R}^m$  and  $b \in \mathbb{R}$ , the set

$$H = \{u \in \mathbb{R}^m : w^T u = b\},$$

is a supporting hyperplane for  $\mathcal{U}$  if  $b = \inf_{u \in \mathcal{U}} w^T u$ . The  $w$  is the normal of  $H$ . In particular, if there exists  $v \in \mathcal{U}$  such that  $w^T u \geq w^T v$  for all  $u \in \mathcal{U}$ , then  $H$  is called a supporting hyperplane through  $v$ . Because  $\mathcal{U}$  contains  $f(x) + \mathbb{R}_+^n$ , the normal  $w$  must be nonnegative for  $H$  to be a supporting hyperplane.

In MOP, people often use different orderings to define various minimizers. We refer to Marler and Arora [43], Miettinen [45], and Jahn et al. [22] for general orderings in MOP. Here, we introduce the convenient lexicographical ordering, up to permutations. Let  $\pi$  be a permutation of  $(1, \dots, m)$ . For a set  $T \subseteq \mathbb{R}^m$ , construct the following chain of nesting subsets

$$T = T_0 \supseteq T_1 \supseteq \dots \supseteq T_m,$$

such that: for each  $k = 1, \dots, m$ ,  $T_k$  is the subset of vectors in  $T_{k-1}$  whose  $\pi(k)$ th entry is the smallest. If  $T_m \neq \emptyset$ , then each  $v \in T_m$  is called a  $\pi$ -minimal point of  $T$ . For  $u, v \in T_m$ , all the entries of  $u, v$  must be the same, so  $u = v$ , and, hence,  $T_m$  consists of a single point, if it is nonempty. In particular, if  $T$  is compact, then  $T_m \neq \emptyset$ , and it consists of a single point.

PVs and WPVs are characterized in the following. Some of these results may already exist in the literature. For the convenience of readers, we summarize them together and give direct proofs.

**Proposition 3.1.** *Let  $\mathcal{U}$  be as in (3.1). For each  $v \in f(K)$ , we have:*

- i. *The vector  $v$  is a WPV if and only if  $v$  lies on the boundary of  $\mathcal{U}$ . Moreover, if  $v$  is an extreme point of  $\text{conv}(\mathcal{U})$ , then  $v$  is a PV.*
- ii. *Assume  $\mathcal{U}$  is convex. If  $v$  is a WPV, then there exists a supporting hyperplane for  $\mathcal{U}$  through  $v$  whose normal is nonnegative—that is, there exists  $0 \neq w \geq 0$  such that  $w^T u \geq w^T v$  for all  $u \in \mathcal{U}$ .*
- iii. *Suppose  $H = \{u : w^T u = w^T v\}$  is a supporting hyperplane for  $\mathcal{U}$  through  $v$ , with a normal vector  $0 \neq w \geq 0$ . If  $w > 0$ , then  $v$  is a PV. For  $w$  with a zero entry, if  $u \in f(K)$  is a  $\pi$ -minimal point of  $H \cap \mathcal{U}$ , then  $u$  is a PV. If  $u \in f(K)$  is an extreme point of  $H \cap \mathcal{U}$ , then  $u$  is also a PV.*

**Proof.** (i) If  $v$  lies on the boundary of  $\mathcal{U}$ , then there is no  $p \in K$  such that  $f(p) < v$ , so  $v$  is a WPV. If  $v$  is an interior point of  $\mathcal{U}$ , then there exist  $p \in K$  and  $q \geq 0$  such that  $f(p) + q < v$ , which denies that  $v$  is a WPV. This shows that  $v$  is a WPV if and only if  $v$  lie on the boundary of  $\mathcal{U}$ .

Next, suppose  $v$  is an extreme point of  $\text{conv}(\mathcal{U})$ . Suppose otherwise that  $v$  is not a PV; then, there exists  $p \in K$  such that  $f(p) \leq v, f(p) \neq v$ . This means that  $v = f(p) + q$ , for some  $0 \neq q \in \mathbb{R}_+^m$ . Hence,  $v = [f(p) + (f(p) + 2q)]/2$ , which implies  $v$  is not an extreme point of  $\text{conv}(\mathcal{U})$ , a contradiction. So  $v$  is a PV.

(ii) If  $v$  is a WPV, then  $v$  lies on the boundary of  $\mathcal{U}$ . Because  $\mathcal{U}$  is convex, there is a supporting hyperplane for  $\mathcal{U}$  through  $v$ —that is, there exists  $w \neq 0$  such that  $w^T u \geq w^T v$  for all  $u \in \mathcal{U}$ . The set  $\mathcal{U}$  contains  $v + \mathbb{R}_+^m$ , so  $w \geq 0$ .

(iii) For the case  $w > 0$ , the conclusion is obvious. When  $w$  has zero entries, let  $I = \{i \in [m] : w_i > 0\}$ . To prove  $u := (u_1, \dots, u_m)$  is a PV, suppose  $p \in K$  is such that  $f(p) \leq u$ . Because  $u \in H \cap \mathcal{U}$ ,  $w^T f(p) \leq w^T u = w^T v$ . Also, note that  $w^T f(p) \geq w^T u$  because  $H$  is a supporting hyperplane. So, we must have  $w^T f(p) = w^T v$  and  $f_i(p) = u_i$  for all  $i \in I$ . Write that  $u = f(p) + q$ , for some  $q \in \mathbb{R}_+^m$ . Note that  $q_i = 0$  for all  $i \in I$ . Because  $u$  is a  $\pi$ -minimal point of  $H \cap \mathcal{U}$  and  $f(p) \leq u$ , the vector  $f(p)$  is also a  $\pi$ -minimal point of  $H \cap \mathcal{U}$ . Hence,  $u = f(p)$ , by the  $\pi$ -minimality. This means that  $u$  is a PV.

When  $u$  is an extreme point of  $H \cap \mathcal{U}$ , we can prove that  $u$  is a PV in the same way as for the item (i).  $\square$

We have the following remarks for Proposition 3.1.

- Not every WPV lies on the boundary of  $\text{conv}(\mathcal{U})$ . For instance, consider

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

For each  $t \in (0, 1)$ , the point  $(t, \sqrt{1-t^2})$  is a WPP (also a PP), but it does not lie on the boundary of  $\text{conv}(\mathcal{U})$ .

• If  $\mathcal{U}$  is not convex, there may not exist a supporting hyperplane through a WPV. For instance, in the above MOP, for every  $t \in (0, 1)$ , there is no supporting hyperplane for  $\text{conv}(\mathcal{U})$  through  $(t, \sqrt{1-t^2})$ .

• For the item (iii) of Proposition 3.1, if  $w$  has a zero entry, then  $v$  may not be a Pareto value. For instance, consider the unconstrained MOP

$$\min (x_1, x_2^2).$$

For  $w = (0, 1)$  and  $v = (0, 0)$ , the equation  $w^T u = 0$  gives a supporting hyperplane through  $(0, 0)$ , but  $(0, 0)$  is not a Pareto value.

- If  $v$  is a PV, it may not be an extreme point of  $\mathcal{U}$  or  $H \cap \mathcal{U}$ . For instance, consider the MOP

$$\begin{cases} \min & (x_1, x_2) \\ \text{s.t.} & x_1 \geq 0, x_2 \geq 0, x_1 + x_2 = 1. \end{cases}$$

The set  $\mathcal{U} = \{x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \geq 1\}$ . Clearly, for every  $t \in (0, 1)$ , the vector  $(t, 1-t)$  is a PV, but it is not an extreme point of  $\mathcal{U}$ . The hyperplane  $H = \{x_1 + x_2 = 1\}$  supports  $\mathcal{U}$  at  $(t, 1-t)$ . However,  $(t, 1-t)$  is not an extreme point of the intersection  $H \cap \mathcal{U}$ , for every  $t \in (0, 1)$ .

### 3.2. A Convex Representation

When the feasible set  $K$  is bounded, there always exist supporting hyperplanes for  $\mathcal{U}$ . When  $K$  is unbounded, they may or may not exist. For, given  $v = (v_1, \dots, v_m) \in f(K)$ , how do we determine whether there is a supporting hyperplane through it? For this purpose, we consider the linear optimization in  $w_0 \in \mathbb{R}$  and  $w = (w_1, \dots, w_m) \in \mathbb{R}^m$ :

$$\begin{cases} \omega^* := \max & w_0 \\ \text{s.t.} & 1 - e^T w = 0, w_i \geq w_0 \ (i \in [m]), \\ & \sum_{i=1}^m w_i (f_i(x) - v_i) \geq 0 \text{ on } K. \end{cases} \quad (3.2)$$

Clearly, there is a supporting hyperplane through  $v$  if and only if the optimal value  $\omega^* \geq 0$ . Let  $d$  be the maximum degree of objectives  $f_i$ . The third constraint in (3.2) is equivalent to the membership

$$\sum_{i=1}^m w_i (f_i(x) - v_i) \in \mathcal{P}_d(K),$$

where  $\mathcal{P}_d(K)$  is the nonnegative polynomial cone as in (2.3). The dual cone of  $\mathcal{P}_d(K)$  is the closure  $\text{cl}(\mathcal{R}_d(K))$ , where  $\mathcal{R}_d(K)$  is the moment cone as in (2.2). The dual optimization of (3.2) can be shown to be

$$\begin{cases} \min & t \\ \text{s.t.} & t - \langle f_i - v_i, y \rangle \geq 0 \ (i \in [m]), \\ & 1 = mt - \sum_{i=1}^m \langle f_i - v_i, y \rangle, \ y \in \text{cl}(\mathcal{R}_d(K)). \end{cases} \quad (3.3)$$

In the above, the vector  $y$  is a tms labeled as

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

If (3.3) has a feasible point with  $t < 0$ , then there are no nonnegative supporting hyperplanes through  $v$ . Because each  $v_i$  is a scalar, one can see that

$$\langle f_i - v_i, y \rangle = \langle f_i, y \rangle - v_i \langle 1, y \rangle = \langle f_i, y \rangle - v_i y_0.$$

When  $t < 0$  is feasible for (3.3), there also exists a feasible  $y \in \mathcal{R}_d(K)$  with  $y_0 > 0$ . One can scale such  $(t, y)$  so that  $y_0 = 1$ . Hence, the existence of  $t < 0$  in (3.3) is equivalent to

$$\begin{cases} \tau = mt' - \sum_{i=1}^m (\langle f_i, y \rangle - v_i), \\ t' \geq \langle f_i, y \rangle - v_i, i = 1, \dots, m, \\ \tau > 0 > t', y_0 = 1, y \in \mathcal{R}_d(K). \end{cases}$$

The above is then equivalent to that

$$\begin{cases} v_i > \langle f_i, y \rangle, i = 1, \dots, m, \\ y_0 = 1, y \in \mathcal{R}_d(K). \end{cases}$$

We define the set  $\mathcal{V}$  containing all such  $v$ :

$$\mathcal{V} := \left\{ v \left| \begin{array}{l} v = (v_1, \dots, v_m) \\ v_i > \langle f_i, y \rangle, i = 1, \dots, m, \\ y_0 = 1, y \in \mathcal{R}_d(K) \end{array} \right. \right\}. \quad (3.4)$$

**Theorem 3.2.** Assume  $K$  has nonempty interior. Then, the interior of the convex hull  $\text{conv}(\mathcal{U})$  is the set  $\mathcal{V}$  as in (3.4). Moreover, when  $\mathcal{U}$  is convex, a vector  $v \in f(K)$  is a weakly Pareto value if and only if  $v$  belongs to the boundary of the closure  $\text{cl}(\mathcal{V})$ .

**Proof.** Because  $K$  has nonempty interior, the cone  $\mathcal{R}_d(K)$  has nonempty interior. Hence, the strong duality holds between (3.2) and (3.3) because (3.3) has strictly feasible points. This is because one can select  $y$  from the interior of  $\mathcal{R}_d(K)$ , choose  $t$  sufficiently large to satisfy all the inequalities, and then scale such  $(t, y)$  for the equality to hold.

A point  $v$  lies in the interior of  $\text{conv}(\mathcal{U})$  if and only if there is no supporting hyperplane for  $\mathcal{U}$  through it. The normal of every supporting hyperplane for  $\mathcal{U}$  is nonnegative. Thus,  $v$  lies in the interior of  $\text{conv}(\mathcal{U})$  if and only if the optimal value  $\omega^*$  of (3.2) is negative or it is infeasible. By the strong duality between (3.2) and (3.3), this is equivalent to that  $v$  belongs to  $\mathcal{V}$ .

When  $\mathcal{U}$  is convex—that is,  $\text{conv}(\mathcal{U}) = \mathcal{U}$ —a vector  $v \in f(K)$  is a WPV if and only if  $v$  lies on the boundary of  $\mathcal{U}$ , by Proposition 3.1. This is equivalent to that  $v$  lies on the boundary of  $\text{cl}(\mathcal{V})$  because the interior of  $\mathcal{U}$  is  $\mathcal{V}$ .  $\square$

A computational efficient description for the moment cone  $\mathcal{R}_d(K)$  is usually not available. However, when the polynomials are SOS-convex, there exists a semidefinite representation for the set  $\mathcal{V}$  in (3.4). Recall that a polynomial  $p \in \mathbb{R}[x]$  is SOS-convex (see Helton and Nie [17]) if  $\nabla^2 p = Q(x)^T Q(x)$  for some matrix polynomial  $Q(x)$ .

**Theorem 3.3.** Assume  $\mathcal{E} = \emptyset$  and  $K$  has nonempty interior. If all  $f_i$  and  $-c_j$  ( $j \in \mathcal{I}$ ) are SOS-convex polynomials, then the interior of  $\mathcal{U}$  is equal to

$$\mathcal{V}_1 := \left\{ (v_1, \dots, v_m) \left| \begin{array}{l} \langle c_j, y \rangle \geq 0 (j \in \mathcal{I}), \\ v_i > \langle f_i, y \rangle (i \in [m]), \\ M_{d_0}[y] \succeq 0, y_0 = 1, \\ y \in \mathbb{R}^{\mathbb{N}_{2d_0}^n} \end{array} \right. \right\}, \quad (3.5)$$

where  $d_0 := \max\{\lceil d/2 \rceil, \lceil \deg(c_j)/2 \rceil (j \in \mathcal{I})\}$ . Moreover, a vector  $v \in f(K)$  is a weakly Pareto value if and only if it lies on the boundary of  $\text{cl}(\mathcal{V}_1)$ .

**Proof.** Clearly, if  $v$  belongs to  $\mathcal{V}$  as in (3.4) for some  $y \in \mathcal{R}_d(K)$ , then it must belong to  $\mathcal{V}_1$ . Conversely, if  $(v, y)$  satisfies (3.5), then let  $\hat{x} := (y_{e_1}, \dots, y_{e_n})$  and  $\hat{y} := [\hat{x}]_d$ . Under the SOS convexity assumption, the Jensen's inequality (see Lasserre [31]) implies that

$$\langle f_i, y \rangle \geq f_i(\hat{x}) = \langle f_i, \hat{y} \rangle, \quad 0 \geq \langle -c_j, y \rangle \geq -c_j(\hat{x}) = \langle -c_j, \hat{y} \rangle.$$

So, we have  $\hat{x} \in K$  and  $\hat{y} \in \mathcal{R}_d(K)$ ; hence,  $v$  belongs to  $\mathcal{V}_1$ . The conclusion then follows from Theorem 3.2.  $\square$

**Example 3.4.** Consider the SOS-convex polynomials

$$f_1 = (x_1 - x_2)^4 + (x_2 - x_3)^4, \quad f_2 = \sum_{i=1}^3 x_i^4 + x_1^2 x_2^2 + x_1^2 x_3^2 + x_2^2 x_3^2,$$

and the ball constraint  $1 \geq \|x\|^2$ . One can verify that

$$\begin{aligned} \nabla^2 f_1 &= 12 \begin{bmatrix} x_1 - x_2 & 0 \\ x_2 - x_1 & x_2 - x_3 \\ 0 & x_3 - x_2 \end{bmatrix} \begin{bmatrix} x_1 - x_2 & 0 \\ x_2 - x_1 & x_2 - x_3 \\ 0 & x_3 - x_2 \end{bmatrix}^T, \\ \nabla^2 f_2 &= 4 \begin{bmatrix} x_1 & x_1 & 0 \\ x_2 & 0 & x_2 \\ 0 & x_3 & x_3 \end{bmatrix} \begin{bmatrix} x_1 & x_1 & 0 \\ x_2 & 0 & x_2 \\ 0 & x_3 & x_3 \end{bmatrix}^T + \sum_{i=1}^3 A_i A_i^T, \end{aligned}$$

where each  $A_i$  is the diagonal matrix with the diagonal vector  $\sqrt{2}x_i(e + (\sqrt{2} - 1)e_i)$ . Note that  $y_{000} = 1$ . The inequalities in the set  $\mathcal{V}_1$  as in (3.5) are

$$\begin{aligned} 1 - y_{200} - y_{020} - y_{002} &\geq 0, \\ v_1 &> \sum_{i=0}^4 \binom{4}{i} (-1)^i (y_{(4-i)e_1+ie_2} + y_{(4-i)e_2+ie_3}), \\ v_2 &> \sum_{i=1}^3 y_{4e_i} + y_{220} + y_{022} + y_{202}. \end{aligned}$$

The moment matrix inequality  $M_2[y] \succeq 0$  reads as

$$\begin{bmatrix} y_{000} & y_{100} & y_{010} & y_{001} & y_{200} & y_{110} & y_{101} & y_{020} & y_{011} & y_{002} \\ y_{100} & y_{200} & y_{110} & y_{101} & y_{300} & y_{210} & y_{201} & y_{120} & y_{111} & y_{102} \\ y_{010} & y_{110} & y_{020} & y_{011} & y_{210} & y_{120} & y_{111} & y_{030} & y_{021} & y_{012} \\ y_{001} & y_{101} & y_{011} & y_{002} & y_{201} & y_{111} & y_{102} & y_{021} & y_{012} & y_{003} \\ y_{200} & y_{300} & y_{210} & y_{201} & y_{400} & y_{310} & y_{301} & y_{220} & y_{211} & y_{202} \\ y_{110} & y_{210} & y_{120} & y_{111} & y_{310} & y_{220} & y_{211} & y_{130} & y_{121} & y_{112} \\ y_{101} & y_{201} & y_{111} & y_{102} & y_{301} & y_{211} & y_{202} & y_{121} & y_{112} & y_{103} \\ y_{020} & y_{120} & y_{030} & y_{021} & y_{220} & y_{130} & y_{121} & y_{040} & y_{031} & y_{022} \\ y_{011} & y_{111} & y_{021} & y_{012} & y_{211} & y_{121} & y_{112} & y_{031} & y_{022} & y_{013} \\ y_{002} & y_{102} & y_{012} & y_{003} & y_{202} & y_{112} & y_{103} & y_{022} & y_{013} & y_{004} \end{bmatrix} \succeq 0.$$

We would like to remark that the Pareto front can be expressed as an image set of polynomial functions. Thus, semidefinite relaxations can be used to approximate the Pareto front. We refer to Magron et al. [40] and Magron et al. [41] for related work on this technique. In contrast, our work expresses the Pareto front in terms of the boundary of sets  $cl(\mathcal{V})$  in (3.4) or  $cl(\mathcal{V}_1)$  in (3.5). In comparison, the expression for the Pareto front via  $cl(\mathcal{V})$  or  $cl(\mathcal{V}_1)$  in our work is exact, but more for theoretical interest, whereas the expression in Magron et al. [40] is approximate, but more for computational interest.

## 4. The Linear Scalarization

This section discusses how to solve linear scalarization problems, how to choose proper weights, and how to detect nonexistence of proper weights. For a weight  $w := (w_1, \dots, w_m)$ , denote the weighted sum

$$f_w(x) := w_1 f_1(x) + \dots + w_m f_m(x).$$

We consider the LSP

$$\min \quad f_w(x) \quad \text{s.t.} \quad x \in K. \quad (4.1)$$

Recall that  $w \neq 0$  is a proper weight if (4.1) is bounded below. Equivalently,  $w$  is a proper weight if and only if  $w$  is the normal of a supporting hyperplane for the set  $\mathcal{U}$  as in (3.1).

### 4.1. Tight Relaxations for LSPs

The Moment-SOS hierarchy of semidefinite relaxations Lasserre [29] can be applied to solve (4.1). When the feasible set  $K$  is unbounded, the Moment-SOS hierarchy may not converge. Here, we apply the tight relaxation method in Nie [52] to solve (4.1).

The Karush-Kuhn-Tucker (KKT) conditions for (4.1) are

$$\nabla f_w(u) = \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i \nabla c_i(u), \quad \lambda_j \geq 0, \quad \lambda_j c_j(u) = 0 \quad (j \in \mathcal{I}),$$

where the  $\lambda_j$ 's are Lagrange multipliers. For convenience, we write such that

$$\begin{aligned} \mathcal{E} \cup \mathcal{I} &= \{1, \dots, s\}, \quad c := (c_1(x), \dots, c_s(x)), \\ c_{eq} &:= (c_i)_{i \in \mathcal{E}}, \quad c_{in} := (c_j)_{j \in \mathcal{I}}. \end{aligned}$$

The KKT conditions imply that

$$\underbrace{\begin{bmatrix} \nabla c_1(x) & \nabla c_2(x) & \dots & \nabla c_s(x) \\ c_1(x) & 0 & \dots & 0 \\ 0 & c_2(x) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_s(x) \end{bmatrix}}_{C(x)} \underbrace{\begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_s \\ \lambda \end{bmatrix}}_{\lambda} = \begin{bmatrix} \nabla f_w(x) \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.2)$$

The polynomial tuple  $c$  is said to be *nonsingular* if the matrix  $C(x)$  as above has full column rank for all complex  $x \in \mathbb{C}^n$  (see Nie [52]). When  $c$  is nonsingular, there exists a matrix polynomial  $L(x)$  such that  $L(x)C(x) = I_s$ . Then,

$$\lambda = L(x) \begin{bmatrix} \nabla f_w(x) \\ 0 \end{bmatrix}.$$

For each  $i = 1, \dots, s$ , let  $\lambda_i(x) := (L(x)_{:,1:n} \nabla f_w(x))_i$  be the  $i$ th entry polynomial. Denote the polynomial sets

$$\Phi := \{c_i\}_{i \in \mathcal{E}} \cup \{\lambda_j(x)c_j\}_{j \in \mathcal{I}} \cup \left\{ \nabla f_w - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i(x)c_i \right\}, \quad (4.3)$$

$$\Psi := \{c_j, \lambda_j(x)\}_{j \in \mathcal{I}}. \quad (4.4)$$

(If  $p$  is a vector of polynomials, then  $\{p\}$  denotes the set of entries of  $p$ .) If its minimum value is achieved at a KKT point, then (4.1) is equivalent to

$$\begin{cases} \min & f_w(x) \\ \text{s.t.} & p(x) = 0 \quad (p \in \Phi), \\ & q(x) \geq 0 \quad (q \in \Psi). \end{cases} \quad (4.5)$$

Let  $k_0 := \max\{\lceil \deg(p)/2 \rceil : p \in \Phi \cup \Psi\}$ . For an integer  $k \geq k_0$ , the  $k$ th-order moment relaxation is

$$\left\{ \begin{array}{l} \min \quad \langle f_w, y \rangle \\ \text{s.t.} \quad L_p^{(k)}[y] = 0 \ (p \in \Phi), \\ \quad L_q^{(k)}[y] \succeq 0 \ (q \in \Psi), \\ \quad M_k[y] \succeq 0, \\ \quad y_0 = 1, \ y \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{array} \right. \quad (4.6)$$

For  $k = k_0, k_0 + 1, \dots$ , the Relaxation (4.6) is a semidefinite program. The following is the algorithm for solving (4.5).

**Algorithm 4.1** Formulate the sets  $\Phi, \Psi$  as in (4.3) and (4.4). Let  $k := k_0$ .

**Step 1.** Solve the Relaxation (4.6) for a minimizer  $y^*$  and let  $t := k_0$ .

**Step 2.** If  $y^*$  satisfies the rank condition

$$\text{rank } M_t[y^*] = \text{rank } M_{t-k_0}[y^*], \quad (4.7)$$

then extract  $r := \text{rank } M_t[y^*]$  minimizers for (4.5).

**Step 3.** If (4.7) fails to hold and  $t < k$ , let  $t := t + 1$  and then go to Step 2; otherwise, let  $k := k + 1$  and go to Step 1.

The Rank Condition (4.7) is called *flat truncation*. It is a sufficient (and almost necessary) condition for checking convergence of the Moment-SOS hierarchy (Nie [48]). Algorithm 4.1 can be implemented in **GloptiPoly 3** (Henrion et al. [19]). The following is the convergence property for the hierarchy of Relaxations (4.6), which follows from Nie et al. [54, theorem 4.4].

**Theorem 4.2.** Assume  $c$  is nonsingular and the LSP (4.1) has a minimizer for the weight  $w$ . Then, for all  $k$  large enough, the optimal value of the Relaxation (4.6) is equal to that of (4.1). Moreover, under either one of the following conditions

i. The set  $\text{Ideal}[\Phi] + \text{Qmod}[\Psi]$  is archimedean, or

ii. The real zero set of polynomials in  $\Phi$  is finite,

if each minimizer of (4.1) is an isolated critical point, then all minimizers of the Relaxation (4.6) must satisfy (4.7), when  $k$  is big enough. Therefore, Algorithm 4.1 must terminate within finitely many loops.

**Example 4.3.** Consider the objectives

$$\begin{aligned} f_1 &= \sum_{i=1}^5 x_i^4 + x_1^2 x_2 + x_1 x_2^2 - 3x_1 x_2 x_3 + x_3 x_4 x_5 + x_3^3, \\ f_2 &= \sum_{i=1}^5 x_i^2 - x_1 x_2^2 - x_2 x_3^2 + x_3 x_4^2 + x_4 x_5^2, \end{aligned}$$

and the constraint  $x_1^2 + \dots + x_5^2 \geq 1$ . The feasible set is unbounded. A list of some weights and the corresponding Pareto points are given in Table 1.<sup>1</sup>

It is worthy to note that

$$\text{Ideal}[c_{eq}] \subseteq \text{Ideal}[\Phi], \quad \text{Qmod}[c_{in}] \subseteq \text{Qmod}[\Psi].$$

Hence, if  $\text{Ideal}[c_{in}] + \text{Qmod}[c_{in}]$  is archimedean, then the condition (i) in Theorem 4.2 holds. Therefore, if the archimedeaness is met for the constraints in (1.1), then the condition (i) must hold.

It is possible that  $f_w(x)$  is unbounded below on  $K$  for some weight  $w$ . For instance,  $f_w(x)$  is unbounded below for  $w = (0, 1)$  in Example 4.3. We refer to the appendix for how to detect unboundedness. Moreover, we remark that not every Pareto point is the minimizer of a LSP, as shown in the following.

**Table 1.** Some Pareto points for Example 4.3.

Weight $w$	Pareto point
(0.5, 0.5)	(-0.3371, 0.4659, -0.7504, -0.2807, -0.1655)
(0.25, 0.75)	(-0.0986, 0.3316, -0.6802, -0.5493, -0.3405)
(0.75, 0.25)	(-0.7711, 0.9015, -1.1818, -0.5752, -0.5114)

**Example 4.4.** Consider the MOP with

$$f_1 = -x_1^3 - x_2^3 + (x_3 - x_4)^2, \quad f_2 = x_1^2 - x_2^2 + (x_3 + x_4)^2,$$

and the constraints  $0 \leq x_1, x_2 \leq 1$ . The LSP is

$$\begin{cases} \min & w_1 f_1(x) + w_2 f_2(x) \\ \text{s.t.} & 0 \leq x_1, x_2 \leq 1. \end{cases}$$

For  $w_1 \geq w_2$ , the minimizer is  $(1, 1, 0, 0)$ . For  $w_1 < w_2$ , the minimizer is  $(0, 1, 0, 0)$ . So, the LSP can only give two Pareto points, by exploring all possibilities of weights. However, each  $(x_1, 1, 0, 0)$ , with  $0 \leq x_1 \leq 1$ , is a Pareto point.

#### 4.2. Existence and Choices of Proper Weights

When  $K$  is compact, the LSP (4.1) is bounded below for all weights. When  $K$  is unbounded, (4.1) may be unbounded below for some  $w$  and has no minimizers. To find a (weakly) Pareto point, we look for a nonzero weight  $w \geq 0$  such that (4.1) is bounded below—that is,  $w$  is a proper weight. The set of all proper weights is denoted as

$$\mathcal{W} := \{0 \neq w \in \mathbb{R}_+^m : f_w(x) \text{ is bounded below on } K\}. \quad (4.8)$$

Clearly, the proper weight set  $\mathcal{W}$  is a convex cone.

Note that a nonzero weight  $w \in \mathcal{W}$  if and only if there exists a scalar  $\gamma \in \mathbb{R}$  such that  $f_w(x) - \gamma \in \mathcal{P}_d(K)$ . So,

$$\mathcal{W} = \{0 \neq w \in \mathbb{R}_+^m : f_w(x) \in \mathcal{P}_d(K) + \mathbb{R}\}. \quad (4.9)$$

The cone  $\mathcal{P}_d(K)$  can be approximated by the sum of the ideal  $\text{Ideal}[c_{eq}]$  and the quadratic module  $\text{Qmod}[c_{in}]$ . Thus, we have the following.

**Proposition 4.5.** *It holds that*

$$\{0 \neq w \in \mathbb{R}_+^m : f_w(x) \in \text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}] + \mathbb{R}\} \subseteq \mathcal{W}. \quad (4.10)$$

When  $\text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  is archimedean ( $K$  is bounded for this case), the containment in (4.10) is an equality. This is because if  $f_w(x)$  is bounded below on  $K$ , then  $f_w(x) - \gamma \in \text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  for  $\gamma$  small enough. When  $K$  is unbounded, the sum  $\text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  cannot be archimedean, and the containment in (4.10) is typically not an equality. For instance, for  $K = \mathbb{R}^3$ ,  $f_1 = x_1^2 x_2^2 (x_1^2 + x_2^2)$ ,  $f_2 = x_3^6 - 3x_1^2 x_2^2 x_3^2$ , we have  $(1, 1) \in \mathcal{W}$ , but  $f_{(1,1)} \notin \Sigma[x] + \mathbb{R}$ . For this case,  $\text{Ideal}[c_{eq}] = \{0\}$ ,  $\text{Qmod}[c_{in}] = \Sigma[x]$ , and  $f_{(1,1)}$  is the Motzkin polynomial that is nonnegative, but not SOS.

Among all proper weights  $w \geq 0$  normalized as  $e^T w = 1$ , the smallest possibility of the minimum value of (4.1) is equal to the smallest one of  $f_1^*, \dots, f_m^*$ , where  $f_i^*$  is the minimum value of  $f_i(x)$  on  $K$ . Some of  $f_i^*$  may be  $-\infty$ . For the choice  $w = e_i$ , the minimum value of (4.1) is  $f_i^*$ . Beyond them, people are also interested in  $w$  such that the minimum value of (4.1) is maximum. We discuss how to find such  $w$  in the following.

Assume  $d$  is the maximum degree of  $f_1, \dots, f_m$ . For the minimum value of  $f_w(x)$  on  $K$  to be maximum, we consider the optimization

$$\begin{cases} \max & \gamma \\ \text{s.t.} & 1 - e^T w = 0, w_1 \geq 0, \dots, w_m \geq 0, \\ & \sum_{i=1}^m w_i f_i - \gamma \in \mathcal{P}_d(K). \end{cases} \quad (4.11)$$

The dual cone of  $\mathcal{P}_d(K)$  is  $cl(\mathcal{R}_d(K))$ . (When  $K$  is compact, the moment cone  $\mathcal{R}_d(K)$  is closed.)

The dual optimization of (4.11) is

$$\begin{cases} \min & \mu \\ \text{s.t.} & \mu - \langle f_i, y \rangle \geq 0 \ (i = 1, \dots, m), \\ & y_0 = 1, y \in cl(\mathcal{R}_d(K)). \end{cases} \quad (4.12)$$

The  $k$ th-order SOS relaxation for (4.11) is

$$\begin{cases} \max & \gamma \\ \text{s.t.} & w_1 + \dots + w_m = 1, w_1 \geq 0, \dots, w_m \geq 0, \\ & \sum_{i=1}^m w_i f_i - \gamma \in \text{Ideal}[c_{eq}]_{2k} + \text{Qmod}[c_{in}]_{2k}. \end{cases} \quad (4.13)$$

The dual optimization of (4.13) is the  $k$ th-order moment relaxation for (4.12):

$$\left\{ \begin{array}{ll} \min & \mu \\ \text{s.t.} & \mu - \langle f_i, y \rangle \geq 0 \ (i = 1, \dots, m), \\ & L_{c_i}^{(k)}[y] = 0 \ (i \in \mathcal{E}), \\ & L_{c_j}^{(k)}[y] \succeq 0 \ (j \in \mathcal{I}), \\ & M_k[y] \succeq 0, \\ & y_0 = 1, \ y \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{array} \right. \quad (4.14)$$

As  $k$  increases, the above gives a hierarchy of Moment-SOS relaxations for solving (4.11). When the sum  $\text{Ideal}[c_{eq}] + Q\text{mod}[c_{in}]$  is archimedean, the convergence of the hierarchy was shown in Lasserre [30] and Nie [50].

**Example 4.6.** Consider the objectives

$$\begin{aligned} f_1 &= (x_1^2 + x_2 + x_3)^2 + (x_2^2 + x_3 + x_4)^2 - 3x_1x_2x_3x_4, \\ f_2 &= \sum_{i=1}^4 x_i^4 - (x_1 - x_2)(x_2 - x_3)(x_3 - x_4)(x_4 - x_1), \\ f_3 &= 3 \sum_{i=1}^4 x_i^3 + x_1^2(x_2^2 - x_3^2) + x_2^2(x_3^2 - x_4^2) + x_3^2(x_4^2 - x_1^2), \end{aligned}$$

and the constraints  $x_1x_2 \geq 1, x_2x_3 \geq 1, x_3x_4 \geq 1, x_1 \geq 0$ . Each  $f_i$  is unbounded below on the feasible set  $K$ . The optimization (4.11) can be solved by the Moment-SOS hierarchy of (4.13)–(4.14). The computed optimal weight  $w^*$  and Pareto point  $x^*$  are, respectively,

$$w^* = (0.5769, 0.2229, 0.2003), \quad x^* = (1.0105, 0.9897, 1.0105, 0.9897).$$

The maximum of the minimum value of  $f_w(x)$  on  $K$  is  $\gamma^* = 11.9435$ .

### 4.3. Nonexistence of Proper Weights

When the feasible set  $K$  is unbounded, there may not exist a weight  $w \geq 0$  such that  $f_w(x)$  is bounded below on  $K$ . We discuss how to detect nonexistence of proper weights.

Recall that  $d$  is the maximum degree of  $f_i$  and  $\tilde{f}_w(\tilde{x}) := x_0^d f_w(x/x_0)$ . When  $K$  is closed at  $\infty$ , the Optimization (4.11) is equivalent to

$$\left\{ \begin{array}{ll} \max & \gamma \\ \text{s.t.} & w_1 + \dots + w_m = 1, (w_1, \dots, w_m) \geq 0, \\ & \tilde{f}_w - \gamma x_0^d \in \mathcal{P}_d(\tilde{K}). \end{array} \right. \quad (4.15)$$

The dual optimization of (4.15) is

$$\left\{ \begin{array}{ll} \min & \mu \\ \text{s.t.} & \mu - \langle x_0^d f_i(x/x_0), \tilde{y} \rangle \geq 0 \ (i = 1, \dots, m), \\ & \langle x_0^d, \tilde{y} \rangle = 1, \ \tilde{y} \in \mathcal{R}_d(\tilde{K}). \end{array} \right. \quad (4.16)$$

When (4.16) is unbounded below, the Problem (4.15) must be infeasible, and, hence, there is no proper weight. This is the case if (4.16) has a decreasing ray  $\Delta\tilde{y}$ :

$$\left\{ \begin{array}{ll} -1 \geq \langle x_0^d f_i(x/x_0), \Delta\tilde{y} \rangle \ (i = 1, \dots, m), \\ \langle x_0^d, \Delta\tilde{y} \rangle = 0, \ \Delta\tilde{y} \in \mathcal{R}_d(\tilde{K}). \end{array} \right. \quad (4.17)$$

Let  $f_i^{(d)}$  denote the homogeneous part of degree  $d$  for  $f_i$ —that is,

$$f_i^{(d)} = x_0^d f_i(x/x_0)|_{x_0=0}.$$

The equality  $\langle x_0^d, \Delta\tilde{y} \rangle = 0$  implies that every representing measure for  $\Delta\tilde{y}$  must be supported in the hyperplane  $x_0 = 0$ . Therefore, (4.17) can be reduced to

$$-1 \geq \langle f_i^{(d)}, \Delta y \rangle \quad (i = 1, \dots, m), \quad \Delta y \in \mathcal{R}_d(K^\circ), \quad (4.18)$$

where  $K^\circ$  is the set, as in (A.3). We remark that if  $\deg(f_i) < d$ , then  $f_i^{(d)} = 0$  and, hence,  $\langle f_i^{(d)}, \Delta y \rangle = 0$ , which implies that (4.17) is infeasible. Therefore, the decreasing ray  $\Delta\tilde{y}$  as in (4.17) exists only if all  $f_i$  have the same degree. The following is the nonexistence theorem of proper weights. Like before, the closeness of  $K$  at infinity can be weakened.

**Theorem 4.7.** *Assume (4.18) has a feasible point  $\Delta y = \lambda_1[z_1]_d + \dots + \lambda_r[z_r]_d$ , with  $\lambda_1, \dots, \lambda_r > 0$  and  $z_1, \dots, z_r \in K^\circ$ . If each  $(0, z_i)$  lies on  $\text{cl}(\tilde{K} \cap \{x_0 > 0\})$ , then the LSP (4.1) is unbounded below for all nonzero  $w \geq 0$  and, hence,  $\mathcal{W} = \emptyset$ .*

**Proof.** For each  $w \geq 0$  with  $e^T w = 1$ , it holds that

$$-1 \geq \left\langle \sum_{i=1}^m w_i f_i^{(d)}, \Delta y \right\rangle = \langle \tilde{f}_w, \Delta y \rangle, \quad \Delta y \in \mathcal{R}_d(K^\circ).$$

Because  $\Delta y = \lambda_1[z_1]_d + \dots + \lambda_r[z_r]_d$ , there exists at least one  $i$  such that

$$-1/r \geq \langle \tilde{f}_w, \lambda_i[z_i]_d \rangle.$$

By Theorem A.1(ii),  $f_w(x)$  is unbounded below on  $K$  because  $(0, z_i)$  lies in the closure of  $\tilde{K} \cap \{x_0 > 0\}$  and  $\lambda_i > 0$ . A nonzero weight  $w \geq 0$  is proper if and only if  $w/(e^T w)$  is proper. Hence, no proper weights exist, and  $\mathcal{W} = \emptyset$ .  $\square$

The Moment System (4.18) is in the form (A.15). Algorithm A.4 can be applied to get a feasible point for (4.18). This can be done by solving a hierarchy of moment relaxations like (A.17). The convergence is shown in Theorem A.5.

**Example 4.8.** Consider the objectives

$$\begin{aligned} f_1 &= -\left(\sum_{i=1}^5 x_i^3\right) - x_2^4 + x_4^4 - x_1 x_2 x_3 - x_3 x_4 x_5, \\ f_2 &= \left(\sum_{i=1}^5 x_i\right)^3 - \sum_{i=1}^4 x_i^4 + x_1 x_2 x_3 x_4 + x_2 x_3 x_4 x_5, \\ f_3 &= x_1^4 - x_2^4 + x_3^4 + x_4^4 - x_1 x_2 x_3 - x_3 x_4 x_5, \\ f_4 &= -(x_1 x_2)^2 + (x_2 x_3)^2 + (x_3 x_4)^2 + (x_4 x_5)^2, \end{aligned}$$

and the constraints  $x_1^2 \geq 1, \dots, x_5^2 \geq 1$ . By Algorithm A.4, we get that  $\Delta y = \lambda[u]_4$  is feasible for (4.18) with

$$u = (-0.7014, -0.7049, 0.0533, -0.0428, 0.0803), \quad \lambda = 4.1146.$$

The set  $\mathcal{C}$  as in (A.12) is empty. By Lemma A.2, the point  $(0, u)$  lies on the closure of  $\tilde{K} \cap \{x_0 > 0\}$ . Therefore, the LSP (4.1) is unbounded below for all nonzero weights  $w \geq 0$  by Theorem 4.7.

We remark that when no proper weights exist, the System (4.18) is still possibly infeasible. For instance, this is the case for

$$K = \mathbb{R}^1, \quad f_1 = x_1^3 + x_1, \quad f_2 = -x_1^3.$$

There is no nonzero  $(w_1, w_2) \geq 0$  such that  $f_w(x)$  is bounded below on  $\mathbb{R}^1$ . However, there is no  $\Delta y$  such that

$$-1 \geq \langle x_1^3, \Delta y \rangle, \quad -1 \geq \langle -x_1^3, \Delta y \rangle, \quad \Delta y \in \mathcal{R}_3(\{x_1^2 = 1\}).$$

Moreover, when no proper weights exist, Pareto points may still exist. For instance, this is the case for

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

For every  $t$ ,  $(t^3, -t)$  is a Pareto point, but there is no nonzero  $w = (w_1, w_2) \geq 0$  such that  $w_1 x_1 + w_2 x_2$  is bounded below on  $x_1 + x_2^3 \geq 0$ .

## 5. The Chebyshev Scalarization

The Chebyshev scalarization problem is

$$\min_{x \in K} \max_{1 \leq i \leq m} w_i(f_i(x) - f_i^*), \quad (5.1)$$

for a nonzero weight  $w := (w_1, \dots, w_m) \geq 0$ . In the above, each  $f_i^*$  is the minimum value of  $f_i$  on  $K$ . In this section, we assume all  $f_i^* > -\infty$ . If one of them is  $-\infty$ , we refer to Subsection 6.2 for how to get PPs and WPPs.

Each minimizer of (5.1) is a weakly Pareto point. Conversely, every weakly Pareto point is a minimizer of the CSP (5.1) for some weight, provided each  $f_i^* > -\infty$ . This is because if  $x^*$  is a weakly Pareto point, then there exist weights  $w_i \geq 0$  such that all  $w_i(f_i(x^*) - f_i^*)$  are equal, because  $f_i(x^*) - f_i^* \geq 0$  for each  $i$ . Then,  $x^*$  is the minimizer for that CSP. Observe that  $f_i^*$  equals the minimum value of the LSP (4.1) for the weight  $w = e_i$ . Algorithm 4.1 can be applied to compute  $f_i^*$ .

After all  $f_i^*$  are obtained, one can solve the CSP (5.1) for a weakly Pareto point. With the new variable  $x_{n+1}$ , the CSP (5.1) is equivalent to

$$\begin{cases} \min & x_{n+1} \\ \text{s.t.} & x_{n+1} - w_i(f_i(x) - f_i^*) \geq 0 \ (i = 1, \dots, m), \\ & c_i(x) = 0 \ (i \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{cases} \quad (5.2)$$

To get convergent Moment-SOS relaxations, we typically need archimedeaness for constraining polynomials. The feasible set of (5.2) is unbounded. To fix this issue, one can select a feasible point  $\xi \in K$  and let

$$B_0 := \max_{1 \leq i \leq m} (w_i(f_i(\xi) - f_i^*)).$$

Then, (5.2) is equivalent to

$$\begin{cases} \min & x_{n+1} \\ \text{s.t.} & x_{n+1} - w_i(f_i(x) - f_i^*) \geq 0 \ (i = 1, \dots, m), \\ & B_0 - x_{n+1} \geq 0, \ x_{n+1} \geq 0, \\ & c_i(x) = 0 \ (i \in \mathcal{E}), \\ & c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{cases} \quad (5.3)$$

For convenience, denote the set

$$\mathcal{G} := \{c_j\}_{j \in \mathcal{I}} \cup \{x_{n+1}, B_0 - x_{n+1}\} \cup \{x_{n+1} - w_i(f_i - f_i^*)\}_{i=1}^m. \quad (5.4)$$

The  $k$ th-order moment relaxation for (5.3) is

$$\begin{cases} \min & \langle x_{n+1}, y \rangle \\ \text{s.t.} & L_{c_i}^{(k)}[y] = 0 \ (i \in \mathcal{E}), \\ & L_p^{(k)}[y] \succeq 0 \ (p \in \mathcal{G}), \\ & M_k[y] \succeq 0, \\ & y_0 = 1, \ y \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{cases} \quad (5.5)$$

Let  $d_0$  be the degree

$$d_0 := \max\{\lceil d/2 \rceil, \lceil \deg(c_i)/2 \rceil \ (i \in \mathcal{E} \cup \mathcal{I})\}. \quad (5.6)$$

Suppose  $y^*$  is a minimizer of (5.5). If there exists  $t \in [d_0, k]$  such that

$$\text{rank } M_t[y^*] = \text{rank } M_{t-d_0}[y^*], \quad (5.7)$$

then we can get  $\text{rank } M_t[y^*]$  minimizers for (5.1) (see Henrion and Lasserre [18] and Nie [48]). The following is about the convergence of the hierarchy of (5.5).

**Theorem 5.1.** *Assume  $\text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  is archimedean. Suppose  $y^{(k)}$  is a minimizer of the Moment Relaxation (5.5) for the order  $k$ . If the CSP (5.1) has finitely many minimizers, then for  $t$  big enough, every accumulation point of  $\{y^{(k)}|_{2t}\}_{k=d_0}^\infty$  must satisfy (5.7).*

**Table 2.** Some Pareto points for Example 5.2.

Weight $w$	Pareto point
(1, 1, 1)	(0.000, 0.000, 0.000, 0.4503)
(1, 2, 2)	(−0.0024, −0.0979, −0.0635, −0.5248)
(1, 2, 3)	(−0.0029, −0.1228, −0.0700, −0.5648)

**Proof.** Because  $\text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  is archimedean, there exists a scalar  $N$  such that  $N - x^T x \in \text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$ . Note that

$$B_0^2 - x_{n+1}^2 = (B_0 - x_{n+1})^2 + 2x_{n+1} \cdot \frac{(B_0 - x_{n+1})^2}{B_0} + 2(B_0 - x_{n+1}) \frac{x_{n+1}^2}{B_0}.$$

Therefore, we get that

$$N - x^T x + B_0^2 - x_{n+1}^2 \in \text{Ideal}[c_{eq}] + \text{Qmod}[\mathcal{G}].$$

This means that  $\text{Ideal}[c_{eq}] + \text{Qmod}[\mathcal{G}]$  is archimedean. When the CSP (5.1) has finitely many minimizers, the conclusion is implied by theorem 3.3 of Nie [48].  $\square$

When  $\text{Ideal}[c_{eq}] + \text{Qmod}[c_{in}]$  is not archimedean (this is the case if  $K$  is unbounded), the homogenization method in Subsection 4.2 can be similarly applied. Moreover, the method in Mai et al. [42] can also be applied to solve (5.1).

**Example 5.2.** Consider the objectives

$$\begin{aligned} f_1 &= \sum_{i=1}^4 x_i^2 - (x_1 x_2 + x_3 x_4)(x_1 x_3 + x_2 x_4), \\ f_2 &= \sum_{i=1}^4 x_i^4 + x_1 x_2 x_3 + x_2 x_3 x_4 + x_1 x_2 x_3 x_4, \\ f_3 &= \sum_{i=1}^4 x_i^6 + (x_1^2 - x_2^2 + 1)(x_2^2 - x_3^2 + 1)(x_3^2 - x_4^2 + 1), \end{aligned}$$

and the constraint  $x_1 x_2 \leq 1, x_2 x_3 \leq 1, x_3 x_4 \leq 1, x_1 x_4 \leq 1$ . The minimum values  $f_1^*, f_2^*, f_3^*$  are 0.0000, −0.0710, 0.6029, respectively. A list of some weights and corresponding weakly Pareto points are in Table 2. Indeed, they are all Pareto points, confirmed by solving the Optimization (6.1).

## 6. Existence and Nonexistence of PPs and WPPs

This section discusses how to check whether a given point is a (weakly) Pareto point and how to detect existence or nonexistence of (weakly) Pareto points.

### 6.1. Detection of PPs and WPPs

For a given point  $x^* \in K$ , how can we detect whether it is a Pareto point or not? To this end, consider the optimization

$$\begin{cases} \min & f_e(x) := f_1(x) + \dots + f_m(x) \\ \text{s.t.} & f_i(x^*) - f_i(x) \geq 0 \ (i = 1, \dots, m), \\ & x \in K. \end{cases} \quad (6.1)$$

This is a kind of lexicographic method (see Marler and Arora [43]). Let  $z^*$  be a minimizer of (6.1), if it exists. Then,  $x^*$  is a Pareto point if and only if the minimum value of (6.1) is equal to  $f_e(x^*)$ . Moreover, if  $x^*$  is not a Pareto point, the minimizer  $z^*$  must be a Pareto point because all the weights are positive. A Pareto point may be obtained by solving (6.1) for given  $x^* \in K$ , provided (6.1) has a minimizer.

Let  $F$  be the feasible set of (6.1) and

$$\mathcal{F} := \{c_j\}_{j \in \mathcal{I}} \cup \{f_i(x^*) - f_i(x)\}_{i=1}^m. \quad (6.2)$$

For a degree  $k \geq d/2$ , the  $k$ th-order moment relaxation for (6.1) is

$$\left\{ \begin{array}{l} \min \quad \langle f_e, y \rangle \\ \text{s.t.} \quad L_{c_i}^{(k)}[y] = 0 \ (i \in \mathcal{E}), \\ \quad L_q^{(k)}[y] \succeq 0 \ (q \in \mathcal{F}), \\ \quad M_k[y] \succeq 0, \\ \quad y_0 = 1, y \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{array} \right. \quad (6.3)$$

Recall that  $d_0$  is the degree as in (5.6). Suppose  $y^*$  is a minimizer of (6.3). If there exists  $t \in [d_0, k]$  such that

$$\text{rank } M_t[y^*] = \text{rank } M_{t-d_0}[y^*], \quad (6.4)$$

then we can get  $r := \text{rank } M_t[y^*]$  minimizers for (6.1). Recall that  $c_{eq}$  is the tuple of equality constraining polynomials. The following result follows from theorem 3.3 of Nie [48].

**Theorem 6.1.** *Assume  $\text{Ideal}[c_{eq}] + \text{Qmod}[\mathcal{F}]$  is archimedean. Suppose  $y^{(k)}$  is a minimizer of the Relaxation (6.3) for the order  $k$ . If (6.1) has only finitely many minimizers, then for  $t$  big enough, every accumulation point of  $\{y^{(k)}|_{2t}\}_{k=1}^{\infty}$  must satisfy (6.4).*

When  $\text{Ideal}[c_{eq}] + \text{Qmod}[\mathcal{F}]$  is not archimedean, the hierarchy of relaxations (6.3) may not converge. For such a case, we refer to the homogenization method in Subsection 4.2 or the method in Mai et al. [42].

**Example 6.2.** (i) Consider the objectives

$$\begin{aligned} f_1 &= x_1^2(x_1 - 2)^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^2, \\ f_2 &= -x_1^2 - x_2^2 - x_3^2 - x_4^2 + x_1x_2 + x_2x_3 + x_3x_4, \end{aligned}$$

and the constraint  $x \geq 0$ . We first solve the CSP (5.1) with  $w_1 = w_2 = 1$  and get the weakly Pareto point  $x^* = (0, 0, 0, 0)$ . It is not a Pareto point. By solving (6.1), we get the Pareto point  $(2.000, 2.001, 2.001, 2.001)$ .

(ii) Consider the objectives

$$f_1 = x_1^3 - x_1^2x_2 - x_2, \quad f_2 = x_2^3 - x_1x_2^2 - x_1,$$

and the constraint  $x_1x_2 \leq 1$ . The LSP (4.1) is unbounded below for all weights  $w_i$ , which is confirmed by a feasible point for (4.18). But we are still able to find a Pareto point by solving (6.1) for some given  $x^*$ . For instance, for  $x^* = (-1, -0.5)$ , solving (6.1) gives the Pareto point  $(1.0000, 1.0000)$ .

We can similarly detect whether a given point  $x^* \in K$  is a weakly Pareto point or not. Consider the optimization

$$\left\{ \begin{array}{l} \min \quad \max_{1 \leq i \leq m} (f_i(x) - f_i(x^*)) \\ \text{s.t.} \quad f_i(x^*) - f_i(x) \geq 0 \ (i = 1, \dots, m), \\ \quad c_i(x) = 0 \ (i \in \mathcal{E}), \\ \quad c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{array} \right. \quad (6.5)$$

Let  $z^*$  be a minimizer of (6.5), if it exists. Then,  $x^*$  is a weakly Pareto point if and only if the optimal value of (6.5) is equal to zero. Moreover, if  $x^*$  is not a weakly Pareto point, then one can show that  $z^*$  is a weakly Pareto point. By introducing the new variable  $x_{n+1}$ , the Optimization (6.5) is equivalent to

$$\left\{ \begin{array}{l} \min \quad x_{n+1} \\ \text{s.t.} \quad x_{n+1} - f_i(x) + f_i(x^*) \geq 0 \ (i = 1, \dots, m), \\ \quad f_i(x^*) - f_i(x) \geq 0 \ (i = 1, \dots, m), \\ \quad c_i(x) = 0 \ (i \in \mathcal{E}), \\ \quad c_j(x) \geq 0 \ (j \in \mathcal{I}). \end{array} \right. \quad (6.6)$$

The optimal value of (6.6) is always less than or equal to zero. A similar hierarchy of moment relaxations like (5.5) can be applied to solve (6.6), and a similar convergence result like Theorem 5.1 holds. When the feasible set

of (6.5) is unbounded, the Moment-SOS hierarchy may not converge. For such a case, we refer to the homogenization method in Subsection 4.2 or the method in Mai et al. [42].

## 6.2. Existence of PPs and WPPs

When  $K$  is unbounded, we discuss how to detect existence of PPs and WPPs. Consider the min-max optimization

$$\min_{x \in K} \max_{1 \leq i \leq m} f_i(x). \quad (6.7)$$

The following is the existence result. See Subsection 3.1 for  $\pi$ -minimal points.

**Theorem 6.3.** *The Min-Max Optimization (6.7) has the following properties:*

- i. If (6.7) is unbounded below, then there is no weakly Pareto point, and, hence, there is no Pareto point. If (6.7) is bounded below, then every minimizer of (6.7) (if it exists) is a weakly Pareto point.
- ii. Let  $S$  be the set of minimizers of (6.7). For each  $x^* \in S$ , if  $f(x^*)$  is a  $\pi$ -minimal point of the image  $f(S)$  for a permutation  $\pi$  of  $(1, \dots, m)$ , then  $x^*$  is a Pareto point. In particular, if  $S$  is compact, then there exists a Pareto point.

**Proof.** (i) If (6.7) is unbounded below, then for every  $x \in K$ , there exists  $z \in K$  such that

$$\max_{1 \leq i \leq m} f_i(z) < \min_{1 \leq i \leq m} f_i(x).$$

This implies  $f(z) < f(x)$ ; hence, there is no weakly Pareto point.

Suppose (6.7) is bounded below and it has a minimizer, say,  $x^*$ . Then,  $x^*$  must be a weakly Pareto point. If, otherwise, there is  $z \in K$  such that  $f(z) < f(x^*)$ , then

$$\max_i f_i(z) < \max_i f_i(x^*),$$

which contradicts that  $x^*$  is a minimizer.

(ii) Suppose  $f(x^*)$  is a  $\pi$ -minimal point of  $f(S)$ . Let  $z \in K$  be a point such that  $f(z) \leq f(x^*)$ . Because  $x^*$  is a minimizer of (6.7), one can see that

$$\max_{1 \leq i \leq m} f_i(x^*) \leq \max_{1 \leq i \leq m} f_i(z) \leq \max_{1 \leq i \leq m} f_i(x^*).$$

This implies that  $z$  is also a minimizer of (6.7), so  $z \in S$ . Because  $f(x^*)$  is  $\pi$ -minimal among  $f(S)$ ,  $f(x^*) \leq f(z)$ , so  $f(x^*) = f(z)$ , and, hence,  $x^*$  is a Pareto point. When  $S$  is compact, the set  $S$  must have a  $\pi$ -minimal point, for every permutation  $\pi$  of  $(1, \dots, m)$ , and, hence, (1.1) has a Pareto point, by Proposition 3.1.  $\square$

Each optimizer  $x^*$  of (6.7) is a weakly Pareto point. One can solve (6.1) to check whether  $x^*$  is a Pareto point or not. If it is not, each minimizer of (6.1) is a Pareto point. We remark that (6.7) can be reformulated as polynomial optimization. By introducing the new variable  $x_{n+1}$ , the Optimization (6.7) is equivalent to

$$\begin{cases} \min & x_{n+1} \\ \text{s.t.} & x_{n+1} \geq f_i(x) \ (i \in [m]), \\ & x \in K. \end{cases} \quad (6.8)$$

The Moment-SOS hierarchy can be applied to solve it. When the set  $K$  is unbounded, the feasible set of (6.8) is also unbounded. The Moment-SOS hierarchy may not converge. For such a case, we refer to the homogenization method in Subsection 4.2 or the method in Mai et al. [42].

Once a minimizer  $x^*$  for (6.8) is obtained, we can solve (6.1) to detect whether it is a Pareto point or not. If it is not, we may get a Pareto point by solving (6.1).

**Example 6.4.** Consider the MOP with objectives

$$f_1 = x_1^3 + x_2^3 - x_3^3 + x_3^2 x_4^2, f_2 = x_2^3 + x_3^3 - x_4^3 + x_4^2 x_1^2,$$

$$f_3 = x_3^3 + x_4^3 - x_1^3 + x_1^2 x_2^2, f_4 = x_4^3 + x_1^3 - x_2^3 + x_2^2 x_3^2,$$

and with the exterior constraint  $x_1^3 + x_2^3 + x_3^3 + x_4^3 \geq 1$ . All  $f_1, f_2, f_3, f_4$  are unbounded below on  $K$ . The CSP (5.1) does not exist because each  $f_i^* = -\infty$ . However, solving (6.8) gives the Pareto point  $(0.6300, 0.6300, 0.6300, 0.6300)$ .

### 6.3. Nonexistence of WPPs

We discuss how to detect nonexistence of weakly Pareto points, when  $K$  is unbounded. Recall that  $d_i := \deg(f_i)$ . Observe that (6.7) is unbounded below if and only if the following optimization is unbounded below:

$$\begin{cases} \min & x_{n+1} \\ \text{s.t.} & -(-x_{n+1})^{d_i} - f_i(x) \geq 0 \ (i \in [m]), \ x \in K. \end{cases} \quad (6.9)$$

Let  $K_1$  be the feasible set of (6.9), and let its homogenization be (note  $\tilde{x} := (x_0, x)$ ):

$$\tilde{K}_1 := \left\{ (x_0, x, x_{n+1}) \left| \begin{array}{l} -(-x_{n+1})^{d_i} - \tilde{f}_i(\tilde{x}) \geq 0 \ (i \in [m]), \\ \tilde{c}_i(\tilde{x}) = 0 \ (i \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 \ (j \in \mathcal{I}), \\ \|\tilde{x}\|^2 + \|x_{n+1}\|^2 = 1, \ x_0 \geq 0 \end{array} \right. \right\}. \quad (6.10)$$

When  $K_1$  is closed at  $\infty$ ,  $x_{n+1} \geq \gamma$  on  $K_1$  if and only if  $x_{n+1} - \gamma x_0 \geq 0$  on  $\tilde{K}_1$ —that is,  $x_{n+1} - \gamma x_0 \in \mathcal{P}_1(\tilde{K}_1)$ . So, we consider the linear conic optimization

$$\max \gamma \quad \text{s.t.} \quad x_{n+1} - \gamma x_0 \in \mathcal{P}_1(\tilde{K}_1). \quad (6.11)$$

The Optimization (6.7) is unbounded below if and only if (6.11) is infeasible, when  $K_1$  is closed at  $\infty$ . The dual optimization of (6.11) is

$$\min \langle x_{n+1}, \tilde{y} \rangle \quad \text{s.t.} \quad \langle x_0, \tilde{y} \rangle = 1, \ \tilde{y} \in \mathcal{R}_1(\tilde{K}_1). \quad (6.12)$$

Note that (6.12) is feasible if  $K$  is nonempty. So, it is unbounded below if there is a decreasing ray  $\Delta \tilde{y}$ :

$$\langle x_{n+1}, \Delta \tilde{y} \rangle = -1, \quad \langle x_0, \Delta \tilde{y} \rangle = 0, \quad \Delta \tilde{y} \in \mathcal{R}_1(\tilde{K}_1). \quad (6.13)$$

Because  $x_0 \geq 0$  on  $\tilde{K}_1$ , the equality  $\langle x_0, \Delta \tilde{y} \rangle = 0$  implies that every representing measure for  $\Delta \tilde{y}$  is supported in  $x_0 = 0$ . Therefore, (6.13) is equivalent to

$$\langle x_{n+1}, \Delta \tilde{y} \rangle = -1, \quad \Delta \tilde{y} \in \mathcal{R}_1(K_1^\circ), \quad (6.14)$$

where  $K_1^\circ$  is the linear section  $x_0 = 0$  of  $\tilde{K}_1$ :

$$K_1^\circ := \left\{ (x, x_{n+1}) \left| \begin{array}{l} -(-x_{n+1})^{d_i} - f_i^{hom}(x) \geq 0 \ (i \in [m]), \\ c_i^{hom}(x) = 0 \ (i \in \mathcal{E}), \\ c_j^{hom}(x) \geq 0 \ (j \in \mathcal{I}), \\ \|x\|^2 + x_{n+1}^2 = 1 \end{array} \right. \right\}. \quad (6.15)$$

The following is the theorem for nonexistence of WPPs.

**Theorem 6.5.** Suppose  $\Delta \tilde{y} = \lambda v$ , with  $\lambda > 0$  and  $v \in K_1^\circ$ , is a feasible point for (6.14). If the point  $(0, v) \in cl(\tilde{K}_1 \cap \{x_0 > 0\})$ , then (6.9) and (6.7) must be unbounded below, and, hence, there are no weakly Pareto points.

**Proof.** The unboundedness of (6.9) is implied by the item (ii) of Theorem A.1, for the case that  $g^{hom} := x_{n+1}$  and  $K^\circ$  is replaced by  $K_1^\circ$ . Note that (6.7) is unbounded below if and only if (6.9) is unbounded below. So, (6.7) is also unbounded below. By Theorem 6.3, there are no weakly Pareto points.  $\square$

The tms  $\Delta \tilde{y} = \lambda v$  satisfying (6.14) can be obtained by Algorithm A.4 with a minor variation. The only difference is to choose a generic  $R \in int(\Sigma[x, x_{n+1}]_{2d_1})$  and then solve the hierarchy of moment relaxations:

$$\begin{cases} \min & \langle R, z \rangle \\ \text{s.t.} & \langle x_{n+1}, z \rangle = -1, \\ & L_{\|x, x_{n+1}\|^2 - 1}^{(k)}[z] = 0, \\ & L_{c_i^{hom}}^{(k)}[z] = 0 \ (i \in \mathcal{E}), \\ & L_{c_j^{hom}}^{(k)}[z] \succeq 0 \ (j \in \mathcal{I}), \\ & L_{h_i}^{(k)}[z] \succeq 0 \ (i \in [m]), \\ & M_k[z] \succeq 0, \ z \in \mathbb{R}^{\mathbb{N}_{2k}^{n+1}}. \end{cases} \quad (6.16)$$

In the above, each  $h_i := -(-x_0)^{d_i} - f_i^{hom}(x)$ . The convergence property for the hierarchy of (6.16) is similar to that for Theorem A.5.

**Example 6.6.** Consider the MOP with objectives

$$\begin{aligned} f_1 &= (x_1 x_2 + x_3 x_4)(x_1 x_4 + x_2 x_3) + x_1^2 + x_2^2 + x_3^2 + x_4^2, \\ f_2 &= x_1^3 x_2^2 + x_2^3 x_3^2 + x_3^3 x_4^2 + x_4^3 x_1^2, \\ f_3 &= x_1^4 - x_2^4 + x_3^4 - x_4^4 + x_1 x_2 x_4 + x_1 x_3 x_4, \\ f_4 &= (x_1 - x_2)(x_3 - x_4)^2 + (x_1 - x_3)(x_2 - x_4)^2 \\ &\quad + (x_1 - x_4)(x_2 - x_3)^2 + x_1 x_2 + x_2 x_3 + x_3 x_4, \end{aligned}$$

and with the constraints  $x_1 x_2 x_3 \geq 1$ ,  $x_2 x_3 x_4 \geq 1$ . Solving the Moment Relaxation (6.16) gives the feasible point  $\Delta \tilde{y} = 3.3597[v]_5$  with

$$v = (v_1, v_2, v_3, v_4, v_5) = (-0.2761, 0.8737, 0.0000, -0.2680, -0.2976).$$

The set  $K_1$  is not closed at infinity, but  $(0, v)$  still belongs to  $cl(\tilde{K}_1 \cap \{x_0 > 0\})$ . This is implied by Lemma A.2 because  $\Delta x = (0, 0, -1, 0, 0)^T$  satisfies the Condition (A.13). By Theorem 6.5, there is no weakly Pareto point.

#### 6.4. Nonexistence of PPs

When there are no weakly Pareto points, there must exist no Pareto points. So, Theorem 6.5 is also applicable to detect nonexistence of Pareto points. However, a Pareto point may not exist while weakly Pareto points exist. This section discusses how to detect nonexistence of Pareto points for this case.

We consider the Optimization (6.1) with  $x^* \in K$ . A Pareto point exists if and only if (6.1) is bounded below and has a minimizer for some  $x^* \in K$ . The “if” implication is clear. When  $x^*$  itself is a Pareto point, then  $x^*$  must be a minimizer for (6.1). This explains the “only if” implication. Let  $K(x^*)$  be the feasible set of (6.1) determined by  $x^*$ , and let  $\tilde{K}(x^*)$  be the homogenization of  $K(x^*)$  similarly as in (A.2). Suppose  $K(x^*)$  is closed at  $\infty$ . Then, (6.1) is bounded below if and only if  $\tilde{f}_e(\tilde{x}) - \gamma x_0^d \in \mathcal{P}_d(\tilde{K}(x^*))$  for some  $\gamma$ . We consider the linear conic optimization

$$\max \gamma \quad s.t. \quad \tilde{f}_e(\tilde{x}) - \gamma x_0^d \in \mathcal{P}_d(\tilde{K}(x^*)). \quad (6.17)$$

Pareto points do not exist if (6.1) is unbounded below for all  $x^* \in K$ . This is equivalent to that (6.17) is infeasible for all  $x^* \in K$ . The dual optimization of (6.17) is

$$\min \langle \tilde{f}_e, \tilde{y} \rangle \quad s.t. \quad \langle x_0^d, \tilde{y} \rangle = 1, \tilde{y} \in \mathcal{R}_d(\tilde{K}(x^*)). \quad (6.18)$$

By weak duality, (6.17) is infeasible if (6.18) is unbounded below. The Problem (6.18) is feasible for all  $x^* \in K$ . Therefore, (6.18) is unbounded below if there is a decreasing ray  $\Delta \tilde{y}$ :

$$\langle \tilde{f}_e, \Delta \tilde{y} \rangle = -1, \langle x_0^d, \Delta \tilde{y} \rangle = 0, \quad \Delta \tilde{y} \in \mathcal{R}_d(\tilde{K}(x^*)). \quad (6.19)$$

Because  $x_0 \geq 0$  on  $\tilde{K}(x^*)$ ,  $\langle x_0^d, \Delta \tilde{y} \rangle = 0$  if and only if every representing measure for  $\Delta \tilde{y}$  is supported in the hyperplane  $x_0 = 0$ . Hence, the existence of  $\Delta \tilde{y}$  satisfying (6.19) is equivalent to the existence of  $\Delta \tilde{y}$  satisfying

$$\langle f_e^{hom}, \Delta \tilde{y} \rangle = -1, \quad \Delta \tilde{y} \in \mathcal{R}_d(\tilde{K}_0^*), \quad (6.20)$$

where  $f_e^{hom}(x) := \tilde{f}_e(0, x)$  and  $K_0^*$  is the section  $x_0 = 0$  of  $K(x^*)$ :

$$K_0^* := \left\{ x \left| \begin{array}{l} c_j^{hom}(x) = 0 \ (j \in \mathcal{E}), \\ c_j^{hom}(x) \geq 0 \ (j \in \mathcal{I}), \\ -f_i^{hom}(x) \geq 0 \ (i \in [m]), \\ x^T x = 1. \end{array} \right. \right\}. \quad (6.21)$$

It is important to observe that  $K_0^*$  and (6.20) do not depend on  $x^*$ . If there exists  $\Delta \tilde{y}$  satisfying (6.20), then (6.1) is unbounded below for all  $x^* \in K$ , and, hence, there are no Pareto points. This implies the following theorem.

**Theorem 6.7.** Suppose  $K(x^*)$  is closed at infinity for all  $x^* \in K$ . If there is  $\Delta\tilde{y}$  satisfying (6.20), then (6.1) is unbounded below for all  $x^* \in K$ , and, hence, Pareto points do not exist.

Theorem 6.7 only shows nonexistence of Pareto points, but it does not imply nonexistence of weakly Pareto points. For instance, consider the MOP

$$\begin{cases} \min & (x_1, x_2) \\ \text{s.t.} & x_1 \geq 0. \end{cases}$$

The tms  $\Delta\tilde{y} := [(0, -1)]_1$  satisfies (6.20), so there are no Pareto points. But each  $(0, x_2)$  is a weakly Pareto point. The existence of  $\Delta\tilde{y}$  satisfying (6.20) can be checked by applying Algorithm A.4 similarly, with the polynomial  $g_1 := f_e^{hom}$  and the set  $K_0^*$ . The properties are summarized in Theorems A.1 and A.5.

**Example 6.8.** Consider the objectives

$$\begin{aligned} f_1 &= x_1^4 + x_3^4 + (x_1 x_2)^2 + (x_2 x_3)^2 + (x_3 x_4)^2 + x_1 x_2 x_3 x_4, \\ f_2 &= x_1^4 + x_2^4 + x_3^4 + x_4^4 - 2x_2^4 - x_1^3 x_2 - x_3^3 x_4, \end{aligned}$$

and the constraint  $x_1 x_2 x_3 x_4 \geq 0$ . Because  $f_1(0, t, 0, 0) = 0$  is the minimum value, the point  $(0, t, 0, 0)$  is a weakly Pareto point for all  $t \in \mathbb{R}$ . Because all the polynomials are homogeneous,  $K(x^*)$  is closed at infinity for all  $x^* \in K$ . By Algorithm A.4, we get  $\Delta\tilde{y} = 1.0023[u]_4$  satisfying (6.20), for  $u = (0.0000, -0.9994, 0.0000, 0.0339)$ . Hence, there is no Pareto point.

## 7. Conclusions and Discussions

This paper studies multi-objective optimization given by polynomials. We characterize the convex geometry for (weakly) Pareto values and give convex representations for them. For LSPs, we show how to use tight relaxations to solve them, how to find proper weights, and how to detect nonexistence of proper weights. For CSPs, we show how to solve them by moment relaxations. Furthermore, we show how to check whether a given point is a (weakly) Pareto point and how to detect existence or nonexistence of (weakly) Pareto points. To detect nonexistence of proper weights and (weakly) Pareto points, we also show how to detect unboundedness of polynomial optimization.

There are some open questions for studying these topics. To detect nonexistence of (weakly) Pareto points, or to detect nonexistence of proper weights, we need to check unboundedness of polynomial optimization. This is discussed in the appendix. A feasible point for the System (A.8) is only a sufficient condition for unboundedness of the Optimization (A.1), but it may not be necessary.

**Question 7.1.** When (A.8) is infeasible, what is a computationally convenient certificate for unboundedness of (A.1)?

Another important question is to detect nonexistence of proper weights. This is discussed in Subsection 4.3. We have seen that (4.18) is sufficient for the proper weight set  $\mathcal{W} = \emptyset$ , but it may not be necessary.

**Question 7.2.** When (4.18) does not have a feasible point, how can we detect nonexistence of proper weights?

In Subsections 6.3 and 6.4, we discussed how to detect nonexistence of (weakly) Pareto points. Under certain conditions, we have shown that (6.14) implies nonexistence of weakly Pareto points and (6.20) implies nonexistence of Pareto points. However, they may not be necessary for nonexistence.

**Question 7.3.** Beyond (6.14) and (6.20), what are computationally convenient certificates for nonexistence of (weakly) Pareto points?

The above questions are mostly open, to the best of the authors' knowledge. They are interesting future work.

## Appendix. Unboundedness in Polynomial Optimization

This section discusses how to detect unboundedness of a polynomial optimization problem. This question is very important for detecting nonexistence of proper weights and (weakly) Pareto points in Sections 4 and 6.

For a polynomial  $g(x)$  of degree  $d$ , consider the optimization

$$\inf g(x) \quad \text{s.t.} \quad x \in K. \quad (\text{A.1})$$

The feasible set  $K$  is the same as for (1.1). When  $K$  is unbounded, (A.1) may be unbounded below—that is, there exists a sequence  $\{u_k\} \subseteq K$  such that  $g(u_k) \rightarrow -\infty$ . We discuss how to detect unboundedness of (A.1). Equivalently, the Problem (A.1) is unbounded

below if and only if

$$\inf\{g(x) | x \in K\} = -\infty.$$

The homogenization of the set  $K$  is  $(\tilde{x} := (x_0, x)$  is the homogenizing variable)

$$\tilde{K} := \left\{ \tilde{x} \left| \begin{array}{l} \tilde{c}_i(\tilde{x}) = 0 (i \in \mathcal{E}), \\ \tilde{c}_j(\tilde{x}) \geq 0 (j \in \mathcal{I}), \\ \tilde{x}^T \tilde{x} = 1, x_0 \geq 0 \end{array} \right. \right\}, \quad (\text{A.2})$$

where  $\tilde{c}_i(\tilde{x}) = x_0^{\deg(c_i)} c_i(x/x_0)$  is the homogenization of  $c_i(x)$ . The ball constraint  $\tilde{x}^T \tilde{x} = 1$  is added to make the set  $\tilde{K}$  compact. The constraint  $x_0 \geq 0$  ensures that  $\tilde{g}(\tilde{x}) - \gamma x_0^{\deg(g)} \geq 0$  on  $\tilde{K}$  implies that  $g(x) - \gamma \geq 0$  on  $K$ . The set  $K$  is said to be *closed at  $\infty$*  (see Nie [47]) if

$$\tilde{K} = \text{cl}(\{\tilde{x} \in \tilde{K} : x_0 > 0\}).$$

The closeness of  $K$  at  $\infty$  is a genericity condition, as shown in Guo et al. [16]. When  $K$  is closed at  $\infty$ , the polynomial  $g(x) - \gamma$  is nonnegative on  $K$  if and only if its homogenization  $\tilde{g}(\tilde{x}) - \gamma x_0^{\deg(g)}$  is nonnegative on  $\tilde{K}$ .

The intersection of  $\tilde{K}$  and  $x_0 = 0$  is

$$K^\circ := \left\{ x \in \mathbb{R}^n \left| \begin{array}{l} c_i^{\text{hom}}(x) = 0 (i \in \mathcal{E}), \\ c_j^{\text{hom}}(x) \geq 0 (j \in \mathcal{I}), \\ x^T x = 1, \end{array} \right. \right\}, \quad (\text{A.3})$$

where each  $c_i^{\text{hom}}(x) = \tilde{c}_i(0, x)$ .

### A.1. A Certificate for Unboundedness

The Optimization (A.1) is bounded below if and only if  $g$  has a lower bound  $\gamma$  on  $K$ —that is,  $g - \gamma \in \mathcal{P}_d(K)$ . So, we consider the optimization

$$\begin{cases} \max & \gamma \\ \text{s.t.} & g - \gamma \in \mathcal{P}_d(K). \end{cases} \quad (\text{A.4})$$

To check infeasibility of (A.4), we use the homogenization trick in Nie [47]. When  $K$  is closed at  $\infty$ , a polynomial  $p \geq 0$  on  $K$  if and only if its homogenization  $\tilde{p} \geq 0$  on  $\tilde{K}$  (see Huang et al. [20] and Nie [47]). So, the membership  $g - \gamma \in \mathcal{P}_d(K)$  is equivalent to  $\tilde{g} - \gamma x_0^d \in \mathcal{P}_d(\tilde{K})$ , and, hence, (A.4) is the same as

$$\begin{cases} \max & \gamma \\ \text{s.t.} & \tilde{g} - \gamma x_0^d \in \mathcal{P}_d(\tilde{K}). \end{cases} \quad (\text{A.5})$$

The dual optimization of (A.5) is

$$\begin{cases} \min & \langle \tilde{g}, y \rangle \\ \text{s.t.} & \langle x_0^d, y \rangle = 1, y \in \mathcal{R}_d(\tilde{K}). \end{cases} \quad (\text{A.6})$$

If (A.6) is unbounded below, then (A.5) must be infeasible, which implies that (A.4) is infeasible and (A.1) is unbounded below, when  $K$  is closed at  $\infty$ .

When  $K \neq \emptyset$ , the Linear Conic Optimization (A.6) has a feasible point. It is unbounded below if there is a decreasing ray  $\Delta y$ :

$$\langle \tilde{g}, \Delta y \rangle = -1, \langle x_0^d, \Delta y \rangle = 0, \quad \Delta y \in \mathcal{R}_d(\tilde{K}). \quad (\text{A.7})$$

If  $\nu$  is a representing measure for  $\Delta y$  and is supported in  $\tilde{K}$ , then

$$0 = \langle x_0^d, \Delta y \rangle = \int x_0^d d\nu,$$

implies that  $\text{supp}(\nu) \subseteq \tilde{K} \cap \{x_0 = 0\}$ . Thus, (A.7) is equivalent to

$$\langle g^{\text{hom}}, z \rangle = -1, \quad z \in \mathcal{R}_d(K^\circ), \quad (\text{A.8})$$

where  $K^\circ$  is the set as in (A.3). Let  $d_1$  be the degree

$$d_1 := \lceil d/2 \rceil. \quad (\text{A.9})$$

To check whether (A.8) is feasible or not, we select a generic  $R \in \text{int}(\Sigma[x]_{2d_1})$  and consider the linear moment optimization

$$\begin{cases} \min & \langle R, z \rangle \\ \text{s.t.} & \langle g^{\text{hom}}, z \rangle = -1, z \in \mathcal{R}_{2d_1}(K^\circ). \end{cases} \quad (\text{A.10})$$

The following shows how to detect unboundedness of (A.1).

**Theorem A.1.** Let  $\tilde{K}, K^\circ$  be the sets as in (A.2)–(A.3).

- i. Suppose (A.8) is feasible. If  $R \in \text{int}(\Sigma[x]_{2d_1})$  is generic, then (A.10) has a unique optimizer  $z^*$  and  $z^* = \lambda[u]_{2d_1}$ , with  $u \in K^\circ$  and  $\lambda > 0$ .
- ii. Suppose  $z := \lambda[u]_d$ , with  $u \in K^\circ$  and  $\lambda > 0$ , is a feasible point for (A.8). If the point  $(0, u) \in \text{cl}(\tilde{K} \cap \{x_0 > 0\})$ , then (A.1) is unbounded below.

**Proof.** (i) Because  $R$  is generic in the interior  $\text{int}(\Sigma[x]_{2d_1})$ , there exists  $\epsilon > 0$  such that

$$R - \epsilon \|[x]_{d_0}\|^2 \in \Sigma[x]_{2d_1}.$$

Hence, for all  $z \in \mathcal{R}_{2d_0}(K^\circ)$ , it holds that

$$\langle R, z \rangle \geq \epsilon \|[x]_{d_1}\|^2, z \rangle \geq \epsilon \cdot \text{trace}(M_{d_1}[z]).$$

Because (A.8) is feasible, the Optimization (A.10) is also feasible—say,  $z^{(0)}$  is a feasible point. Then, (A.10) is equivalent to

$$\left\{ \begin{array}{l} \min \quad \langle R, z \rangle \\ \text{s.t.} \quad \text{trace}(M_{d_1}[z]) \leq \frac{1}{\epsilon} \langle R, z^{(0)} \rangle, \\ \quad \langle f^{hom}, z \rangle = -1, \\ \quad z \in \mathcal{R}_{2d_1}(K^\circ). \end{array} \right. \quad (\text{A.11})$$

The feasible set of (A.11) is compact, so it has an optimizer—say,  $z^*$ . When  $R$  is generic, the optimizer  $z^*$  must be unique, and it is an extreme point of the feasible set of (A.10). Because (A.10) has only a single equality constraint, the optimizer  $z^*$  must lie in an extreme ray of the cone  $\mathcal{R}_{2d_1}(K^\circ)$ . This means that  $z^* = \lambda[u]_{2d_1}$  for a point  $u \in K^\circ$  and a scalar  $\lambda > 0$  (note  $z^*$  is nonzero).

(ii) Because  $(0, u) \in \text{cl}(\tilde{K} \cap \{x_0 > 0\})$ , there is a sequence

$$\{(t_k, u_k)\}_{k=1}^\infty \subseteq \tilde{K} \cap \{x_0 > 0\},$$

such that  $\lim_{k \rightarrow \infty} (t_k, u_k) = (0, u)$ . Note that each  $t_k > 0$  and

$$-1 = \langle g^{hom}, \lambda[u]_d \rangle = \lambda g^{hom}(u) = \lim_{k \rightarrow \infty} \lambda \tilde{g}(t_k, u_k).$$

Thus, for  $k$  big enough,  $\lambda \tilde{g}(t_k, u_k) \leq -1/2$  and

$$\lambda \tilde{g}(t_k, u_k) = \lambda \cdot (t_k)^d g(u_k/t_k) \leq -1/2.$$

This implies that  $g(u_k/t_k) \leq -1/(2\lambda(t_k)^d)$  for all  $k$  big enough, so  $g(u_k/t_k) \rightarrow -\infty$  as  $k \rightarrow \infty$ . Because each  $u_k/t_k \in K$ ,  $g$  is unbounded below on  $K$ .  $\square$

In computational practice, the generic polynomial  $R \in \text{int}(\Sigma[x]_{2d_1})$  can be chosen as  $[x]_{d_1}^T A^T A [x]_{d_1}$ , for some randomly generated square matrix  $A$ .

We remark that the closeness of  $K$  at  $\infty$  is a generic condition, as shown in Guo et al. [16]. In Theorem A.1(ii), we use the relaxed condition  $(0, u) \in \text{cl}(\tilde{K} \cap \{x_0 > 0\})$  instead of the closeness at  $\infty$ . For the relaxed condition, we give a sufficient condition in Lemma A.2 to check whether it is satisfied.

**Lemma A.2.** Let  $\tilde{K}, K^\circ$  be the sets as in (A.2) and (A.3) and  $z := \lambda[u]_d$ , with  $u \in K^\circ$  and  $\lambda > 0$ , be a feasible point for (A.8). If there exist  $\Delta x \in \mathbb{R}^n$  and  $\delta_0 > 0$  such that

$$c_i^{hom}(u + t\Delta x) > 0 \quad \forall t \in (0, \delta_0), \quad \forall i \in \mathcal{C} := \{i \in \mathcal{E} \cup \mathcal{I} \mid c_i^{hom}(u) = 0\}, \quad (\text{A.12})$$

then  $(0, u) \in \text{cl}(\tilde{K} \cap \{x_0 > 0\})$ .

**Proof.** The constraint polynomial  $\tilde{c}_i(x_0, x)$  can be rewritten as

$$\tilde{c}_i(x_0, x) = c_i^{hom}(x) + x_0 h_i(x_0, x),$$

for some polynomial  $h_i(x_0, x)$ . When  $i \in \mathcal{C}$ , it satisfies the Condition (A.12). When  $i \notin \mathcal{C}$ , it holds that  $c_i^{hom}(u) > 0$ . Therefore, there are  $M > 0$  and  $0 < \delta < \delta_0$  such that

$$c_i^{hom}(u + t\Delta x) > 0 \quad \text{and} \quad h_i(x_0, x) > -M,$$

for all  $\delta > x_0, t > 0$  and  $i \in \mathcal{E} \cup \mathcal{I}$ . Let  $\{t_k\}_{k=1}^\infty$  be a sequence such that  $\lim_{k \rightarrow \infty} t_k = 0$  and  $\delta > t_k > 0$  for all  $k$ . For each  $k$ , we define

$$s_k := \min \left( \frac{\delta}{2k}, \left\{ \frac{c_i^{hom}(u + t_k \Delta x)}{2M} \right\}_{i \in \mathcal{E} \cup \mathcal{I}} \right) > 0.$$

For all  $i \in \mathcal{E} \cup \mathcal{I}$ , it holds that

$$\begin{aligned}\tilde{c}_i(s_k, u + t_k \Delta x) &= c_i^{hom}(u + t_k \Delta x) + s_k h_i(s_k, u + t_k \Delta x) \\ &\geq c_i^{hom}(u + t_k \Delta x) + \frac{c_i^{hom}(u + t_k \Delta x)}{2M}(-M) \\ &\geq \frac{1}{2} c_i^{hom}(u + t_k \Delta x) \\ &> 0.\end{aligned}$$

For convenience, we denote

$$\tilde{u}_k := (s_k, u + t_k \Delta x) / \| (s_k, u + t_k \Delta x) \|.$$

Each  $\tilde{c}_i$  is homogeneous, so  $\tilde{c}_i(\tilde{u}_k) > 0$  by above inequalities. It implies  $\tilde{u}_k \in \tilde{K}$ . The construction of sequences ensures

$$\lim_{k \rightarrow \infty} s_k = \lim_{k \rightarrow \infty} t_k = 0.$$

Thus,  $\lim_{k \rightarrow \infty} \tilde{u}_k = (0, u)$ . So, it shows that  $(0, u) \in \text{cl}(\tilde{K} \cap \{x_0 > 0\})$ .  $\square$

The Sufficient Condition (A.12) in Lemma A.2 requires that  $\Delta x$  is an increasing direction for  $c_i^{hom}(x)$  at  $x = u$  for  $i \in \mathcal{C}$ . It can be checked numerically by gradients and Hessian matrices. We denote  $\mathcal{C}_0 = \{i \in \mathcal{C} \mid \nabla c_i^{hom}(u) = 0\}$  and  $\mathcal{C}_1 = \{i \in \mathcal{C} \mid \nabla c_i^{hom}(u) \neq 0\}$ . The direction  $\Delta x$  satisfies the condition in (A.12) if it satisfies

$$\begin{cases} \Delta x^T \nabla^2 c_i^{hom}(u) \Delta x > 0 & \forall i \in \mathcal{C}_0, \\ \nabla c_i^{hom}(u)^T \Delta x > 0 & \forall i \in \mathcal{C}_1. \end{cases} \quad (\text{A.13})$$

It can be formulated as the following quadratic optimization problem

$$\begin{cases} \max_{\Delta x, a} & a \\ \text{s.t.} & \Delta x^T \nabla^2 c_i^{hom}(u) \Delta x \geq a \quad \forall i \in \mathcal{C}_0, \\ & \nabla c_i^{hom}(u)^T \Delta x \geq a \quad \forall i \in \mathcal{C}_1, \\ & \|\Delta x\|^2 \leq 1. \end{cases} \quad (\text{A.14})$$

There exists a direction  $\Delta x$  satisfying (A.13) if and only if the Problem (A.14) has the maximum  $a^* > 0$ . The Problem (A.14) can be solved as a polynomial optimization problem.

**Example A.3.** Consider the following optimization problem

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

Note that  $g^{hom} = x_1 x_2 x_3$  and a feasible point of (A.8) is the tms  $3\sqrt{3}[u]_6$ , for  $u = (1, 1, -1)/\sqrt{3}$ . One can check that  $\nabla c^{hom}(u) = 0$  and  $e^T \nabla^2 c^{hom}(u) e > 0$  for  $e = (1, 1, 1)^T$ . It demonstrates  $(0, u)$  lies on the closure  $\text{cl}(\tilde{K} \cap \{x_0 > 0\})$ , so this optimization problem is unbounded below.

When (A.6) is unbounded below, it is not necessary that (A.6) has a decreasing ray—that is, the System (A.8) may be infeasible. That is, (A.8) is sufficient for unboundedness of (A.1), but it may not be necessary. For instance, consider the optimization

$$\begin{cases} \min & g(x) := x_1 x_2 x_3 + x_1^2 x_2^2 (x_1^2 + x_2^2) + x_3^6 - 3x_1^2 x_2^2 x_3^2 \\ \text{s.t.} & x_1^2 + x_2^2 - 2x_3^2 = 0, x_1 x_2 \geq 0. \end{cases}$$

It is unbounded below, because  $g(t, t, -t) = -t^3 \rightarrow -\infty$  as  $t \rightarrow +\infty$  and  $(t, t, -t)$  is feasible for all  $t \geq 0$ . However, the Certificate (A.8) is infeasible. This is because

$$g^{hom} = x_1^2 x_2^2 (x_1^2 + x_2^2) + x_3^6 - 3x_1^2 x_2^2 x_3^2,$$

is the Motzkin polynomial, and  $\langle g^{hom}, z \rangle \geq 0$  for all  $z \in \mathcal{R}_d(K^\circ)$ . When (A.8) fails to be feasible, the question of detecting unboundedness of (A.1) is mostly open.

## A.2. Solving Linear Moment Systems

Semidefinite relaxations can be applied to solve (A.8) and (A.10). For more generality, we consider the moment system

$$a_i \geq \langle g_i, z \rangle \quad (i = 1, \dots, m), \quad z \in \mathcal{R}_d(K^\circ), \quad (\text{A.15})$$

for given polynomials  $g_1, \dots, g_m \in \mathbb{R}[x]_d$  and given scalars  $a_1, \dots, a_m \in \mathbb{R}$ . It is worthy to note that in (A.8) and (A.10), the equality is equivalent to the inequality like the above, due to the conic membership condition.

Select a generic  $R \in \text{int}(\Sigma[x]_{2d_1})$  and consider the moment optimization

$$\begin{cases} \min & \langle R, z \rangle \\ \text{s.t.} & a_i - \langle g_i, z \rangle \geq 0 \ (i \in [m]), \\ & z \in \mathcal{R}_{2d_1}(K^\circ). \end{cases} \quad (\text{A.16})$$

Let  $d_2 := \max\{d_1, d_c\}$ , where  $d_c$  is as in (2.7). For  $k = d_2, d_2 + 1, \dots$ , we solve the hierarchy of semidefinite relaxations

$$\begin{cases} \min & \langle R, z \rangle \\ \text{s.t.} & a_i - \langle g_i, z \rangle \geq 0 \ (i = 1, \dots, m), \\ & L_{c_i^{\text{hom}}}^{(k)}[z] = 0 \ (i \in \mathcal{E}), \\ & L_{c_j^{\text{hom}}}^{(k)}[z] \succeq 0 \ (j \in \mathcal{I}), \\ & L_{x^T x - 1}^{(k)}[z] = 0, \\ & M_k[z] \succeq 0, z \in \mathbb{R}^{\mathbb{N}_{2k}^n}. \end{cases} \quad (\text{A.17})$$

Suppose  $z^{(k)}$  is a minimizer of (A.17) for a relaxation order  $k$ . If there is an integer  $t \in [d_c, k]$  such that the Rank Condition (2.6) holds, then the truncation  $z^{(k)}|_{2t}$  has a  $r$ -atomic representing measure supported in  $K^\circ$ —that is,

$$z^{(k)}|_{2t} = \lambda_1[u_1]_{2t} + \dots + \lambda_r[u_r]_{2t},$$

for scalars  $\lambda_1, \dots, \lambda_r > 0$ , distinct points  $u_1, \dots, u_r \in K^\circ$  and  $r = \text{rank } M_t[z^{(k)}]$ . Then, the truncation  $z^{(k)}|_d$  is a feasible point for (A.15).

**Algorithm A.4** Let  $k := d_2$ . Do the following loop:

Step 1. Solve the Semidefinite Relaxation (A.17) for a minimizer  $z^{(k)}$ .

Step 2. Check if there exists  $t \in [d_c, k]$  such that (2.6) holds. If it does, then the truncation  $z^{(k)}|_d$  is a feasible point for (A.8).

Step 3. If (2.6) fails for all  $t \in [d_c, k]$ , let  $k := k + 1$  and go to Step 1.

Algorithm A.4 can be implemented in the software **GloptiPoly 3** (Henrion et al. [19]). The following is the convergence property for the hierarchy of Relaxations (A.17).

**Theorem A.5.** Assume the System (A.15) is feasible and  $R \in \text{int}(\Sigma[x]_{2d_1})$  is generic. Then, we have:

i. The Optimization (A.16) has a unique minimizer  $z^*$  and

$$z^* = \lambda_1[u_1]_{2d_1} + \dots + \lambda_r[u_r]_{2d_1} \quad (\text{A.18})$$

for scalars  $\lambda_1, \dots, \lambda_r > 0$ , distinct points  $u_1, \dots, u_r \in K^\circ$  and  $r \leq m$ .

ii. For each fixed  $t \geq d_1$ , the sequence  $\{z^{(k)}|_{2t}\}_{k=d_2}^\infty$  is bounded, and every accumulation point  $z^{**}$  of  $\{z^{(k)}|_{2t}\}_{k=d_2}^\infty$  satisfies  $z^* = z^{**}|_{2d_1}$ .

**Proof.** (i) As in the proof for item (i) of Theorem A.1, the trace of  $M_{d_1}[z]$  can be bounded by a constant. Similarly, it implies that (A.16) has a minimizer  $z^*$ . The minimizer  $z^*$  is unique because the objective  $\langle R, z \rangle$  is linear in  $z$  and has generic coefficients. The membership  $z^* \in \mathcal{R}_{2d_1}(K^\circ)$  implies that  $z^*$  has a decomposition like (A.18). We only need to show that  $r \leq m$ . Consider the following linear program in  $(\tau_1, \dots, \tau_r)$ :

$$\begin{cases} \min & \tau_1 R(u_1) + \dots + \tau_r R(u_r) \\ \text{s.t.} & -1 \geq \sum_{j=1}^r \tau_j g_i(u_j), i = 1, \dots, m, \\ & \tau_1 \geq 0, \dots, \tau_r \geq 0. \end{cases} \quad (\text{A.19})$$

Note that (A.17) and (A.19) have the same optimal value. Because it is a linear program, (A.19) has a minimizer  $\tau^* = (\tau_1^*, \dots, \tau_r^*)$  of at most  $m$  nonzero entries (see Bertsimas and Tsitsiklis [4]). This implies that the number  $r$  in (A.18) can be chosen to be at most  $m$ .

(ii) Because  $R$  lies in the interior of  $\Sigma[x]_{2d_1}$ , there is  $\epsilon > 0$  such that  $R - \epsilon \in \Sigma[x]_{2d_1}$ . Then, the constraint  $M_k[z] \succeq 0$  implies that

$$\langle R, z \rangle - \epsilon(z)_0 = \langle R - \epsilon, z \rangle \geq 0.$$

So, we get that  $(z)_0 \leq \langle R, z \rangle / \epsilon$ . The optimal value of (A.17) is always less than or equal to that of (A.10). Therefore, the sequence  $\{(z^{(k)})_0\}_{k=d_2}^\infty$  is bounded. Moreover, the constraint  $L_{x^T x - 1}^{(k)}[z] = 0$  implies that

$$(z)_{2\alpha} = (z)_{2e_1+2\alpha} + \dots + (z)_{2e_n+2\alpha} \geq \max((z)_{2e_1+2\alpha}, \dots, (z)_{2e_n+2\alpha}),$$

for all monomial powers  $\alpha$ . The diagonal entries of the positive semidefinite moment matrix  $M_k[z]$  are precisely the entries  $(z)_{2\beta}$  for powers  $\beta$ . This implies that the sequence  $\{(z^{(k)})_{2\beta}\}_{k=d_2}^\infty$  is bounded for all powers  $\beta$ . Therefore, for each fixed  $t \geq d_1$ , the

sequence of each diagonal entry of  $M_t[z^{(k)}]$  is bounded, and so is the truncated sequence  $\{z^{(k)}|_{2t}\}_{k=d_2}^\infty$ . Let  $\mathcal{H}_k$  be the set of feasible points  $z$  in (A.17) for the relaxation order  $k$ , except the first  $m$  inequalities. Denote the truncation:

$$G_k := \{z|_{2t} : z \in \mathcal{H}_k\}.$$

Then,  $G_{k+1} \subseteq G_k$  for all  $k$ . Because there is a sphere constraint  $x^T x = 1$ , the quadratic module for the set  $K^\circ$  is archimedean, so (see proposition 3.3 of Nie [50])

$$\mathcal{R}_{2t}(K^\circ) = \bigcap_{k=d_2}^\infty G_k.$$

If  $z^{**}$  is an accumulation point of  $\{z^{(k)}|_{2t}\}_{k=d_2}^\infty$ , then  $z^{**} \in G_k$  for all  $k$ , and, hence,  $z^{**} \in \mathcal{R}_{2t}(K^\circ)$ . Note that the truncation  $z^{**}|_{2d_1}$  is also a minimizer of (A.16). Because the minimizer is unique, we must have  $z^* = z^{**}|_{2d_1}$ .  $\square$

The Optimization (A.16) is a linear conic optimization problem with the moment cone. It can also be viewed as a generalized moment problem. When the constraining set is compact, we refer to Lasserre [30] and Nie [50] for how to solve it; when the set is unbounded, we refer to the recent works Huang et al. [20] and Huang et al. [21].

## Endnote

<sup>1</sup> Throughout the paper, all computational results are displayed with four decimal digits.

## References

- [1] Anagnostopoulos A, Becchetti L, Castillo C, Gionis A, Leonardi S (2012) Online team formation in social networks. *WWW'12 Proc. 21st Internat. Conf. World Wide Web* (Association for Computing Machinery, New York), 839–848.
- [2] Bao T, Mordukhovich B (2007) Variational principles for set-valued mappings with applications to multiobjective optimization. *Control Cybernetics* 36(3):531–562.
- [3] Bao TQ, Mordukhovich BS (2010) Relative Pareto minimizers for multiobjective problems: Existence and optimality conditions. *Math. Program.* 122(2):301–347.
- [4] Bertsimas D, Tsitsiklis JN (1997) *Introduction to Linear Optimization*, Athena Scientific Series in Optimization and Neural Computation, vol. 6 (Athena Scientific, Belmont, MA).
- [5] Blanco V, Puerto J, El Haj Ben Ali S (2014) A semidefinite programming approach for solving multiobjective linear programming. *J. Global Optim.* 58(3):465–480.
- [6] Chen B, Lin Y, Zeng W, Xu H, Zhang D (2017) The mean-variance cardinality constrained portfolio optimization problem using a local search-based multi-objective evolutionary algorithm. *Appl. Intelligence* 47:505–525.
- [7] Cho J, Wang Y, Chen R, Chan KS, Swami A (2017) A survey on modeling and optimizing multi-objective systems. *IEEE Comm. Surveys Tutorials* 19(3):1867–1901.
- [8] Clayton ER, Weber WE, Taylor BW III (1982) A goal programming approach to the optimization of multi response simulation models. *IIE Trans.* 14(4):282–287.
- [9] Curto RE, Fialkow LA (1996) *Solution of the Truncated Complex Moment Problem for Flat Data*, Memoirs of the American Mathematical Society, vol. 568 (American Mathematical Society, Providence, RI).
- [10] Curto RE, Fialkow LA (2005) Truncated  $k$ -moment problems in several variables. *J. Operator Theory* 54(1):189–226.
- [11] Donoso Y, Fabregat R (2016) *Multi-Objective Optimization in Computer Networks Using Metaheuristics* (CRC Press, Boca Raton, FL).
- [12] Emmerich MT, Deutz AH (2018) A tutorial on multiobjective optimization: Fundamentals and evolutionary methods. *Nat. Comput.* 17:585–609.
- [13] Fan J, Nie J, Zhou A (2018) Tensor eigenvalue complementarity problems. *Math. Program.* 170:507–539.
- [14] Fleming PJ (1986) Computer aided design of regulators using multiobjective optimization. Di Pillo G, ed. *Control Applications of Nonlinear Programming and Optimization* (Elsevier, Amsterdam), 47–52.
- [15] Geiger MJ, Sevaux M (2011) The biobjective inventory routing problem–problem solution and decision support. *INOC 2011 Internat. Conf. Network Optim.* (Springer, Berlin, Heidelberg), 365–378.
- [16] Guo F, Wang L, Zhou G (2014) Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities. *J. Global Optim.* 58:261–284.
- [17] Helton JW, Nie J (2010) Semidefinite representation of convex sets. *Math. Program.* 122:21–64.
- [18] Henrion D, Lasserre JB (2005) Detecting global optimality and extracting solutions in GloptiPoly. Henrion D, Garulli A, eds. *Positive Polynomials in Control* (Springer, Berlin, Heidelberg), 293–310.
- [19] Henrion D, Lasserre JB, Löfberg J (2009) GloptiPoly 3: Moments, optimization and semidefinite programming. *Optim. Methods Software* 24(4–5):761–779.
- [20] Huang L, Nie J, Yuan YX (2023a) Homogenization for polynomial optimization with unbounded sets. *Math. Program.* 200(1):105–145.
- [21] Huang L, Nie J, Yuan YX (2023b) Generalized truncated moment problems with unbounded sets. *J. Sci. Comput.* 95(1):15.
- [22] Jahn J (2009) *Vector Optimization: Theory, Applications, and Extensions* (Springer, Berlin, Heidelberg).
- [23] Jiao L, Lee JH (2021) Finding efficient solutions in robust multiple objective optimization with SOS-convex polynomial data. *Ann. Oper. Res.* 296(1–2):803–820.
- [24] Jiao L, Lee JH, Zhou Y (2020b) A hybrid approach for finding efficient solutions in vector optimization with SOS-convex polynomials. *Oper. Res. Lett.* 48(2):188–194.
- [25] Jiao L, Lee JH, Ogata Y, Tanaka T (2020a) Multi-objective optimization problems with SOS-convex polynomials over an LMI constraint. *Taiwanese J. Math.* 24(4):1021–1043.

[26] Jones D, Tamiz M (2010) *Practical Goal Programming, International Series in Operations Research & Management Science*, vol. 141 (Springer, New York).

[27] Kim DS, Pham TS, Tuyen NV (2019) On the existence of Pareto solutions for polynomial vector optimization problems. *Math. Program.* 177(1–2):321–341.

[28] Koski J, Silvennoinen R (1987) Norm methods and partial weighting in multicriterion optimization of structures. *Internat. J. Numer. Methods Engng.* 24(6):1101–1121.

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

[30] Lasserre JB (2008) A semidefinite programming approach to the generalized problem of moments. *Math. Program.* 112:65–92.

[31] Lasserre JB (2009) Convexity in semialgebraic geometry and polynomial optimization. *SIAM J. Optim.* 19(4):1995–2014.

[32] Lasserre JB (2015) *An Introduction to Polynomial and Semi-Algebraic Optimization, Cambridge Texts in Applied Mathematics*, vol. 52 (Cambridge University Press, Cambridge, UK).

[33] Lasserre JB (2018) The moment-SOS hierarchy. Sirakov B, Ney de Souza P, Viana M, eds. *Proc. Internat. Congress Math. ICM 2018* (World Scientific, Singapore), 3773–3794.

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

[35] Laurent M (2009) Sums of squares, moment matrices and optimization over polynomials. Putinar M, Sullivant S, eds. *Emerging Applications of Algebraic Geometry* (Springer, New York), 157–270.

[36] Laurent M (2014) Optimization over polynomials: Selected topics. Jang SY, Kim YR, Lee D-W, Yie I, eds. *Proc. Internat. Congress Mathematicians 2014* (Kyung Moon SA Co. Ltd., Seoul), 843–869.

[37] Lee JH, Jiao L (2018) Solving fractional multicriteria optimization problems with sum of squares convex polynomial data. *J. Optim. Theory Appl.* 176:428–455.

[38] Lee JH, Sisarat N, Jiao L (2021) Multi-objective convex polynomial optimization and semidefinite programming relaxations. *J. Global Optim.* 80(1):117–138.

[39] Liu DY, Hu R, Fang YP (2023) Solvability of a regular polynomial vector optimization problem without convexity. *Optimization* 72(3):821–841.

[40] Magron V, Henrion D, Lasserre JB (2014) Approximating Pareto curves using semidefinite relaxations. *Oper. Res. Lett.* 42(6–7):432–437.

[41] Magron V, Henrion D, Lasserre JB (2015) Semidefinite approximations of projections and polynomial images of semialgebraic sets. *SIAM J. Optim.* 25(4):2143–2164.

[42] Mai N, Lasserre JB, Magron V (2022) Positivity certificates and polynomial optimization on non-compact semialgebraic sets. *Math. Program.* 194:443–485.

[43] Marler RT, Arora JS (2004) Survey of multi-objective optimization methods for engineering. *Structural Multidisciplinary Optim.* 26:369–395.

[44] Matsatsinis NF, Delias P (2003) AgentAllocator: An agent-based multi-criteria decision support system for task allocation. Mařík V, McFarlane D, Valckenaers P, eds. *HoloMAS 2003 Internat. Conf. Indust. Appl. Holonic Multi-Agent Systems* (Springer, Berlin, Heidelberg), 225–235.

[45] Miettinen K (1999) *Nonlinear Multiobjective Optimization, International Series in Operations Research & Management Science*, vol. 12 (Springer Science & Business Media, New York).

[46] Nie J (2011) Polynomial matrix inequality and semidefinite representation. *Math. Oper. Res.* 36(3):398–415.

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

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

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

[50] Nie J (2015) Linear optimization with cones of moments and nonnegative polynomials. *Math. Program.* 153:247–274.

[51] Nie J (2017) Generating polynomials and symmetric tensor decompositions. *Found. Comput. Math.* 17:423–465.

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

[53] Nie J, Zhang X (2018) Real eigenvalues of nonsymmetric tensors. *Comput. Optim. Appl.* 70(1):1–32.

[54] Nie J, Yang Z, Zhou G (2022) The saddle point problem of polynomials. *Found. Comput. Math.* 22:1133–1169.

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

[56] Rosenthal S, Borschbach M (2017) Design perspectives of an evolutionary process for multi-objective molecular optimization. Trautmann H, Rudolph G, Klamroth K, Schütze O, Wiecek M, Jin Y, Grimme C, eds. *EMO 2017 Internat. Conf. Evolutionary Multi-Criterion Optim.* (Springer, Cham, Switzerland), 529–544.

[57] Ruiz-Canales P, Rufián-Lizana A (1995) A characterization of weakly efficient points. *Math. Program.* 68(1–3):205–212.

[58] van der Horst E, Marqués-Gallego P, Mulder-Krieger T, Veldhoven JV, Kruisselbrink J, Aleman A, Emmerich MT et al. (2012) Multi-objective evolutionary design of adenosine receptor ligands. *J. Chem. Inform. Model.* 52(7):1713–1721.

[59] Wang P, Emmerich M, Li R, Tang K, Bäck T, Yao X (2014) Convex hull-based multiobjective genetic programming for maximizing receiver operating characteristic performance. *IEEE Trans. Evol. Comput.* 19(2):188–200.

[60] Zonts S (1989) Multiple criteria mathematical programming: An updated overview and several approaches. Karpak B, Zonts S, eds. *Multiple Criteria Decision Making and Risk Analysis Using Microcomputers* (Springer, Berlin, Heidelberg), 7–60.