A NEW SCHEME FOR APPROXIMATING THE WEAKLY EFFICIENT SOLUTION SET OF VECTOR RATIONAL OPTIMIZATION PROBLEMS

FENG GUO AND LIGUO JIAO*

Abstract. In this paper, we provide a new scheme for approximating the weakly efficient solution set for a class of vector optimization problems with rational objectives over a feasible set defined by finitely many polynomial inequalities. More precisely, we present a procedure to obtain a sequence of explicit approximations of the weakly efficient solution set of the problem in question. Each approximation is the intersection of the sublevel set of a single polynomial and the feasible set. To this end, we make use of the achievement function associated with the considered problem and construct polynomial approximations of it over the feasible set from above. Remarkably, the construction can be converted to semidefinite programming problems. Several nontrivial examples are designed to illustrate the proposed new scheme.

1. Introduction

Vector optimization forms an important field of research in optimization theory; see, e.g., [4, 8, 9, 28, 35], and many practical applications in various areas, such as engineering [9], humanitarian aid [13], medical health [5] and so on. In this paper, we will be concerned with the following constrained vector rational optimization problem of the form

$$\mathrm{Min}_{\mathbb{R}^m_+} \left\{ f(\boldsymbol{x}) := \left(\frac{p_1(\boldsymbol{x})}{q_1(\boldsymbol{x})}, \ldots, \frac{p_m(\boldsymbol{x})}{q_m(\boldsymbol{x})} \right) : \boldsymbol{x} \in \Omega \right\}, \tag{VROP}$$

where "$\mathrm{Min}_{\mathbb{R}^m_+}$" is understood with respect to the ordering non-negative orthant $\mathbb{R}^m_+$, $f : \mathbb{R}^n \to \mathbb{R}^m$ is a rational mapping with $f_i = \frac{p_i}{q_i}$, in which p_i and q_i are real polynomials in the variable $\boldsymbol{x} = (x_1, \ldots, x_n)$ for each $i = 1, \ldots, m$, and the feasible set Ω is given by

$$\Omega := \{\boldsymbol{x} \in \mathbb{R}^n \colon g_j(\boldsymbol{x}) \geq 0,\ j = 1, \ldots, r\},$$

where for each $j = 1, \ldots, r$, g_j is a real polynomial in the variable $\boldsymbol{x}$. By letting $q_i = 1$ for all $i = 1, \ldots, m$, our model then covers vector polynomial optimization problems [1, 17, 25, 28, 31], and by letting p_i, q_i be linear functions for all $i = 1, \ldots, m$, our model also covers linear fractional vector optimization problems [16] as well.

For vector optimization, it is almost impossible to find a single point simultaneously minimizing all the objective functions. Therefore, we usually look for some "best preferred" solutions in

Date: May 26, 2022.

Key words and phrases. Vector optimization, polynomial optimization, achievement function, Lasserre's hierarchy, weakly efficient solution set.

2020 Mathematics Subject Classification. 90C29, 90C32, 90C23, 90C22.

*Corresponding Author.

vector optimization. Now, let us recall the concepts of optimal solutions to vector optimization problems. A point $\boldsymbol{x} \in \Omega$ is said to be an *efficient solution* (or Edgeworth–Pareto (EP) optimal point) to the problem (VROP) if it holds that

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \notin -\mathbb{R}^m_+ \setminus \{0\} \quad \text{for all} \quad \boldsymbol{y} \in \Omega;$$

and a *weakly efficient solution* (or weakly EP optimal point) to the problem (VROP) if it holds that

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) \notin -\mathbb{R}^m_{++} \quad \text{for all} \quad \boldsymbol{y} \in \Omega,$$

where $\mathbb{R}^m_{++}$ denotes the positive orthant of $\mathbb{R}^m$. Let $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$ be given, a point $\boldsymbol{x} \in \Omega$ is said to be a *weakly $\boldsymbol{\epsilon}$-efficient solution* to the problem (VROP) if it holds that

$$f(\boldsymbol{y}) - f(\boldsymbol{x}) + \boldsymbol{\epsilon} \notin -\mathbb{R}^m_{++} \quad \text{for all} \quad \boldsymbol{y} \in \Omega.$$

Denote by $\mathcal{S}$ (resp., $\mathcal{S}_w$, $\mathcal{S}_w^{\boldsymbol{\epsilon}}$) the set of all efficient (resp, weakly efficient, weakly $\boldsymbol{\epsilon}$-effcient) solutions to the problem (VROP), respectively. Clearly, $\mathcal{S} \subset \mathcal{S}_w \subset \mathcal{S}_w^{\boldsymbol{\epsilon}}$, but not conversely. We call the image $f(\mathcal{S}_w)$ the Pareto frontier (the Pareto curve if $m = 2$) of (VROP); see [29].

Throughout this paper, we make the following blanket assumptions on (VROP):

(**A1**) The feasible set Ω is nonempty and compact;

(**A2**) The denominators $q_i(\boldsymbol{x}) > 0$ over Ω for all $i = 1, \ldots, m$.

As each f_i is continuous, (**A1**) implies that the image $f(\Omega)$ of the rational mapping f over Ω is also compact, which ensures the existence of (weakly) efficient solutions to the problem (VROP); see, e.g., [2, Theorem 1], [8, Theorem 2.1] and [35, Corollary 3.2.1]. The problem (VROP) is well defined under (**A2**), which is commonly adopted in the literature when studying fractional programming. Moreover, by substituting $\frac{p_i q_i}{q_i^2}$ for $\frac{p_i}{q_i}$, (**A2**) can be weakened as $q_i(\boldsymbol{x}) \neq 0$ over Ω for all $i = 1, \ldots, m$.

Motivated by its extensive applications, a great deal of attention has been attracted to the development of algorithms for computing (weakly) efficient solutions to vector optimization; see [3, 6, 10, 24, 25, 31, 37–39] and references therein. Among them, there are mainly two different approaches for solving vector optimization, by which we mean finding its (weakly) efficient solutions. One is based on the scalarization methods (e.g., [3, 6, 24, 25, 31]), which computes (weakly) efficient solutions by choosing some parameters in advance and reformulating them as one or several single objective optimization problems. The other is based on descent methods; see e.g., [10] for Newton's methods, [37–39, 42, 43] for (projected) gradient methods.

We would like to emphasize that the aforementioned methods can only find one or some particular (weakly) efficient solutions, rather than giving information about the whole set of (weakly) efficient solutions, which is apparently important for applications of vector optimziation in the real world. Instead, the aim and novelty of this paper is to provide a new scheme for approximating the whole set of weakly ($\boldsymbol{\epsilon}$-)efficient solutions of (VROP). More precisely, we provide a procedure to obtain a sequence of explicit approximations of $\mathcal{S}_w^{\boldsymbol{\epsilon}}$ (and hence $\mathcal{S}_w$ by letting $\boldsymbol{\epsilon} \to 0$). Each approximation is the intersection of the sublevel set of a single polynomial and the feasible set Ω. As far as we know, there are few methods of this type for solving vector optimization problems in the literature.

To this end, we make use of the *achievement function* (c.f. [8, 32, 41]) associated with the problem (VROP) which is defined as

$$\psi(\boldsymbol{x}) := \sup_{\boldsymbol{y}\in\Omega} \min_{i=1,\ldots,m} [f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})].$$

It can be shown that the sets $\mathcal{S}_w$ and $\mathcal{S}_w^{\boldsymbol{\epsilon}}$ can be written as the intersection of sublevel sets of $\psi(\boldsymbol{x})$ and the feasible set Ω (see Section 3). As the function $\psi(\boldsymbol{x})$ can be fairly complicated, the problem is reduced to construct polynomial approximations of $\psi(\boldsymbol{x})$. By rewriting the definition of $\psi(\boldsymbol{x})$ as a parametric polynomial optimization problem, we can contruct a sequence of polynomial approximations $\{\psi_k(\boldsymbol{x})\}_{k\in\mathbb{N}}$ of $\psi(\boldsymbol{x})$ over the feasible set Ω from above by invoking the "joint+marginal" approach developed by Lasserre in [21, 22]. Remarkably, the construction of $\{\psi_k(\boldsymbol{x})\}_{k\in\mathbb{N}}$ can be converted to semidefinite programming (SDP) problems. For $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$ of the form $\boldsymbol{\epsilon} = (\delta, \ldots, \delta)$ with $\delta > 0$, the intersection, denoted by $\mathcal{A}(\delta, k)$, of the sublevel set $\psi_k(\boldsymbol{x}) \leq \delta$ and the feasible set Ω are inner approximations of $\mathcal{S}_w^{\boldsymbol{\epsilon}}$. Under some conditions, we prove that $\mathrm{vol}\,(\mathcal{S}_w^{\boldsymbol{\epsilon}} \setminus \mathcal{A}(\delta, k)) \to 0$ as $k \to \infty$, where "$\mathrm{vol}(\cdot)$" denotes the Lebesgue volume (see Theorem 4.2). Since it holds for $\boldsymbol{\epsilon} = (\delta, \ldots, \delta)$ that $\mathcal{S}_w^{\boldsymbol{\epsilon}} \to \mathcal{S}_w$ as $\delta \to 0$ (see Proposition 3.2), we may take $\mathcal{A}(\delta, k)$ as an approximation of $\mathcal{S}_w$ with sufficiently small $\delta > 0$ and sufficiently large $k \in \mathbb{N}$ (see Corollary 4.1 and Remark 4.1).

The rest of this paper is organized as follows. Section 2 contains some preliminaries on polynomial optimization. In Section 3, we study the characterization of the weakly efficient solution set of the problem (VROP) by the associated achievement function $\psi(\boldsymbol{x})$. In Section 4, we show how to approximate the weakly ($\boldsymbol{\epsilon}$-)efficient solution set of the problem (VROP), and present some nontrivial illustrating examples. Concusions are given in Section 5.

2. Preliminaries

In this section, we collect some notation and preliminary results which will be used in this paper. The symbol $\mathbb{N}$ (resp., $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$) denotes the set of nonnegative integers (resp., real numbers, nonnegative real numbers, positive real numbers). For a set D in $\mathbb{R}^n$, we use $\mathrm{cl}(D)$ and $\mathrm{int}(D)$ to denote the closure and interior of D, respectively. Denote by $\mathbf{B}$ the closed unit ball in $\mathbb{R}^n$ centered at the origin. For a point $\boldsymbol{u} \in \mathbb{R}^n$, $\mathrm{dist}(\boldsymbol{u}, D)$ denotes the Euclidean distance between $\boldsymbol{u}$ and D. For $\boldsymbol{u} \in \mathbb{R}^n$, $\|\boldsymbol{u}\|$ denotes the standard Euclidean norm of $\boldsymbol{u}$. For $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$. For $k \in \mathbb{N}$, denote by $\mathbb{N}^n_k = \{\alpha \in \mathbb{N}^n \colon |\alpha| \leq k\}$ and $|\mathbb{N}^n_k|$ its cardinality. Denote by $\mathbb{R}[\boldsymbol{x}]$ the ring of polynomials in $\boldsymbol{x} := (x_1, \ldots, x_n)$ with real coefficients and by $\mathbb{R}[\boldsymbol{x}]_k$ the set of polynomials in $\mathbb{R}[\boldsymbol{x}]$ of degree up to k. For a polynomial f, we use $\deg(f)$ to denote the total degree of f. For $\alpha \in \mathbb{N}^n$, the notation $\boldsymbol{x}^\alpha$ stands for the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$.

Now we recall some background about the *sum of squares* representations of nonnegative (positive) polynomials over a set defined by finitely many polynomial inequalities. We say that a polynomial $h \in \mathbb{R}[\boldsymbol{x}]$ is sum of squares of polynomials if there exist polynomials h_j, $j = 1, \ldots, s$, such that $h = \sum_{j=1}^{s} h_j^2$. The set consisting of all sum of squares polynomial in $\boldsymbol{x}$ is denoted by $\Sigma^2[\boldsymbol{x}]$. Let $\{h_1, \ldots, h_s\} \subset \mathbb{R}[\boldsymbol{x}]$ be a finite set of polynomials and

$$S := \{\boldsymbol{x} \in \mathbb{R}^n \colon h_j(\boldsymbol{x}) \geq 0,\ j = 1, \ldots, s\}.$$

Assumption 2.1. There exists some $N \in \mathbb{R}$ such that

$$N - \sum_{i=1}^{n} x_i^2 = \sigma_0(\boldsymbol{x}) + \sum_{j=1}^{s} \sigma_j(\boldsymbol{x}) h_j(\boldsymbol{x}),$$

for some sum of squares polynomials $\sigma_j \in \Sigma^2[\boldsymbol{x}], j = 0, 1, \ldots, s$.

Theorem 2.1 (Putinar's Positivstellensatz [33])**.** *Suppose that Assumption 2.1 holds. If $h(\boldsymbol{x}) \in \mathbb{R}[\boldsymbol{x}]$ is positive on S, then $h(\boldsymbol{x})$ can be written in the form*

$$h(\boldsymbol{x}) = \sigma_0(\boldsymbol{x}) + \sum_{j=1}^{s} \sigma_j h_j(\boldsymbol{x}), \tag{1}$$

for some sum of squares polynomials $\sigma_j \in \Sigma^2[\boldsymbol{x}], j = 0, 1, \ldots, s$.

Note that if we fix the degrees of σ_j's in (1), then checking the above representation of $h(\boldsymbol{x})$ reduces to an SDP feasibility problem (c.f. [23]). The well-known Lasserre's hierarchy of SDP relaxations for polynomial optimization problems is based on Putinar's Positivstellensatz and the dual moment theory (c.f. [18, 21]).

A sparse version of the representation (1) is available if some sparsity pattern is satisfied by h and h_j's. For a subset $I \subseteq \{1, \ldots, n\}$, denote the subset of variables $\boldsymbol{x}_I := \{x_i \colon i \in I\}$ and $\mathbb{R}[\boldsymbol{x}_I]$ as the polynomial ring in the variables $\boldsymbol{x}_I$.

Assumption 2.2. There are partitions $\{1, \ldots, n\} = I_1 \cup \cdots \cup I_l$ and $\{1, \ldots, s\} = J_1 \cup \cdots \cup J_l$ where J_i, $i = 1, \ldots, l$ are disjoint. The collections $\{I_i\}_{i=1}^{l}$ and $\{J_i\}_{i=1}^{l}$ satisfy the following:

(i) $\forall i \in \{1, \ldots, l-1\}$, $\exists k \in \{1, \ldots, i\}$ s.t. $I_{i+1} \cap (I_1 \cup \cdots \cup I_i) \subseteq I_k$;

(ii) $h_j \in \mathbb{R}[\boldsymbol{x}_{I_i}]$ for each $j \in J_i$, $1 \leq i \leq l$.

(iii) For each $i = 1, \ldots, l$, there exists some $N_i \in \mathbb{R}$ such that

$$N_i - \sum_{j \in I_i} x_j^2 = \sigma_{i,0} + \sum_{j \in J_i} \sigma_{i,j} h_j,$$

for some sum of squares polynomials $\sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[\boldsymbol{x}_{I_i}]$, $j \in J_i$.

The following result enables us to construct sparse SDP relaxations of polynomial optimization problems, which can significantly reduce the computational cost (c.f. [19, 40]).

Theorem 2.2 (Sparse version of Putinar's Positivstellensatz [12, 19, 40])**.** *Suppose that Assumption 2.2 holds. If $h(\boldsymbol{x}) \in \sum_{i=1}^{l} \mathbb{R}[\boldsymbol{x}_{I_i}]$ and is positive on S, then $h(\boldsymbol{x})$ can be written as*

$$h(\boldsymbol{x}) = \sum_{i=1}^{l} \left(\sigma_{i,0} + \sum_{j \in J_i} \sigma_{i,j} h_j \right),$$

for some sum of squares polynomials $\sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[\boldsymbol{x}_{I_i}]$, $j \in J_i$, $i = 1, \dots, l$.

3. Charactering the weakly efficient solution set

In this section, we study the achievement function associated with (VROP), which can be used to characterize the weakly ($\boldsymbol{\epsilon}$-)efficient solution set of (VROP).

By defintion of $\mathcal{S}_w$, we have

$$\begin{aligned}
\mathcal{S}_w = &\ \left\{\boldsymbol{x} \in \Omega \colon \forall \boldsymbol{y} \in \Omega,\ f(\boldsymbol{y}) - f(\boldsymbol{x}) \notin -\mathbb{R}^m_{++}\right\} \\
= &\ \{\boldsymbol{x} \in \Omega \colon \forall \boldsymbol{y} \in \Omega,\ \exists i \in \{1, \dots, m\} \text{ such that } f_i(\boldsymbol{x}) - f_i(\boldsymbol{y}) \leq 0\} \\
= &\ \left\{\boldsymbol{x} \in \Omega \colon \forall \boldsymbol{y} \in \Omega,\ \min_{i=1,\dots,m} [f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})] \leq 0\right\} \\
= &\ \left\{\boldsymbol{x} \in \Omega \colon \sup_{\boldsymbol{y}\in\Omega} \min_{i=1,\dots,m} [f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})] \leq 0\right\}.
\end{aligned}$$

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be the function given by

$$\psi(\boldsymbol{x}) := \sup_{\boldsymbol{y}\in\Omega} \min_{i=1,\dots,m} [f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})].$$

The function $\psi(\boldsymbol{x})$ is known as the achievement function in the area of vector optimization in the literature; see [8, Section 4.6] and [32, 41]. Therefore,

$$\mathcal{S}_w = \{\boldsymbol{x} \in \mathbb{R}^n \colon \psi(\boldsymbol{x}) \leq 0\} \cap \Omega.$$

Moreover, we have the following results, which imply that the function $\psi(\boldsymbol{x})$ is indeed a *merit function* (see [7, 26, 38, 39]).

Proposition 3.1. [32, Lemmas 3.1 and 3.2] *The achievement function $\psi(\boldsymbol{x})$ satisfies*

(i) $\psi(\boldsymbol{x}) \geq 0$ *for all* $\boldsymbol{x} \in \Omega$ *and hence* $\mathcal{S}_w = \{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) = 0\}$.

(ii) $\psi(\boldsymbol{x})$ *is locally Lipschitz on* Ω.

Proof. (i) is clear. If the objective in (VROP) is a vector of polynomials, (ii) was proved in [32, Lemma 3.2] which is based on the locally Lipschitz property of polynomial functions. Note that the rational function f_i is locally Lipschitz on Ω under (**A1-2**). Hence, the proof of [32, Lemma 3.2] is still valid for the case studied in this paper. □

So far, we know the weakly efficient solution set $\mathcal{S}_w$ can be *completely* characterized with the help of the achievement function $\psi(\boldsymbol{x})$. Note that, $\psi(\boldsymbol{x})$ can be fairly complicated and computing $\psi(\boldsymbol{x})$ by some descent methods directly might be difficult. However, as shown below in Proposition 3.3, the sublevels of $\psi(\boldsymbol{x})$ have rather close relation with the set of all weakly $\boldsymbol{\epsilon}$-efficient solutions, which in turn yields the information of the set of all weakly efficient solutions.

Recall the definition of the set $\mathcal{S}_w^{\boldsymbol{\epsilon}}$ of all weakly $\boldsymbol{\epsilon}$-efficient solutions to (VROP), and clearly by definition, $\mathcal{S}_w \subset \mathcal{S}_w^{\boldsymbol{\epsilon}}$ for any $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$. Conversely, denote a set-valued mapping $\mathcal{F}(\cdot) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and let $\mathcal{F}(\boldsymbol{\epsilon}) := \mathcal{S}_w^{\boldsymbol{\epsilon}}$ for $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$. The following proposition shows that $\mathcal{F}(\cdot)$ is *continuous at* $\bar{\boldsymbol{\epsilon}} = 0$

relative to $\mathbb{R}^m_+$ in the sense of Painlevé–Kuratowski (see [34, Definition 5.4]), i.e., $\mathcal{F}(\boldsymbol{\epsilon}) \to \mathcal{F}(0)$ as $\boldsymbol{\epsilon} \to 0$. For convenience, we recall the definitions of continuity (outer semicontinuity, inner semicontinuity) for set-valued mapppings; see [34, Chapters 4 & 5] for more information. Given a set-valued mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, we denote by

$$\limsup_{\boldsymbol{y}\to\bar{\boldsymbol{y}}} F(\boldsymbol{y}) := \{\boldsymbol{x} \in \mathbb{R}^n \colon \exists \boldsymbol{y}_k \to \bar{\boldsymbol{y}},\ \exists \boldsymbol{x}_k \to \boldsymbol{x} \text{ with } \boldsymbol{x}_k \in F(\boldsymbol{y}_k)\},$$

$$\liminf_{\boldsymbol{y}\to\bar{\boldsymbol{y}}} F(\boldsymbol{y}) := \{\boldsymbol{x} \in \mathbb{R}^n \colon \forall \boldsymbol{y}_k \to \bar{\boldsymbol{y}},\ \exists \boldsymbol{x}_k \to \boldsymbol{x} \text{ with } \boldsymbol{x}_k \in F(\boldsymbol{y}_k)\},$$

the outer and inner limit of F at $\bar{\boldsymbol{y}}$ in the sense of Painlevé–Kuratowski, respectively.

Definition 3.1. A set-valued mapping $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to be *outer semicontinuous* (osc) at $\bar{\boldsymbol{y}}$ if $\limsup_{\boldsymbol{y}\to\bar{\boldsymbol{y}}} F(\boldsymbol{y}) \subset F(\bar{\boldsymbol{y}})$, and *inner semicontinuous* (isc) at $\bar{\boldsymbol{y}}$ if $F(\bar{\boldsymbol{y}}) \subset \liminf_{\boldsymbol{y}\to\bar{\boldsymbol{y}}} F(\boldsymbol{y})$. It is called *continuous* at $\bar{\boldsymbol{y}}$ if F is simultaneously osc and isc at $\bar{\boldsymbol{y}}$, i.e., $F(\boldsymbol{y}) \to F(\bar{\boldsymbol{y}})$ as $\boldsymbol{y} \to \bar{\boldsymbol{y}}$. These terms are invoked *relative to* X, a subset of $\mathbb{R}^m$ containing $\bar{\boldsymbol{y}}$, if the inclusions hold in restriction to convergence $\boldsymbol{y} \to \bar{\boldsymbol{y}}$ with $\boldsymbol{y} \in X$.

It follows from Definition 3.1 that $\mathcal{F}(\cdot)$ is continuous at $\bar{\boldsymbol{\epsilon}} = 0$ relative to $\mathbb{R}^m_+$. Similar to [34, Proposition 5.12 and Exercise 5.13], we have the following result. For any $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$, denote $\boldsymbol{\epsilon}_{\max} := \max_{i=1,\dots,m}\{\epsilon_i\}$ and $\boldsymbol{\epsilon}_{\min} := \min_{i=1,\dots,m}\{\epsilon_i\}$.

Proposition 3.2. *For any $d > 0$, there exists a number $\delta(d) > 0$ depending on d such that $\operatorname{dist}(\boldsymbol{u}, \mathcal{S}_w) < d$ for any $\boldsymbol{u} \in \mathcal{S}_w^{\boldsymbol{\epsilon}}$, i.e., $\mathcal{S}_w^{\boldsymbol{\epsilon}} \subset \mathcal{S}_w + d\mathbf{B}$, whenever $\boldsymbol{\epsilon}_{\max} < \delta(d)$.*

Proof. Suppose that the conclusion does not hold for some $d > 0$. Then, for any $k \in \mathbb{N}$, there exist $\boldsymbol{\epsilon}^{(k)}$ with $\boldsymbol{\epsilon}^{(k)}_{\max} < \frac{1}{k}$ and a point $\boldsymbol{u}^{(k)} \in \mathcal{S}_w^{\boldsymbol{\epsilon}^{(k)}}$ such that $\operatorname{dist}(\boldsymbol{u}^{(k)}, \mathcal{S}_w) \geq d$. As Ω is compact, without loss of generality, we can assume that there is a point $\boldsymbol{u}' \in \Omega$ such that $\lim_{k\to\infty} \boldsymbol{u}^{(k)} = \boldsymbol{u}'$. Now we show that $\boldsymbol{u}' \in \mathcal{S}_w$. To the contrary, suppose that there exists $\boldsymbol{y}' \in \Omega$ such that $f(\boldsymbol{y}') - f(\boldsymbol{u}') \in -\mathbb{R}^m_{++}$, i.e., $\max_{i=1,\dots,m}[f_i(\boldsymbol{y}') - f_i(\boldsymbol{u}')] < 0$. Due to the continuity of f_i, there exists $k' \in \mathbb{N}$ such that for each $i = 1, \dots, m$,

$$\max_{i=1,\dots,m}[f_i(\boldsymbol{y}') - f_i(\boldsymbol{u}')] + \frac{1}{k} + f_i(\boldsymbol{u}') - f_i(\boldsymbol{u}^{(k)}) < 0$$

holds for any $k \geq k'$. Then for each $i = 1, \dots, m$,

$$\begin{aligned} f_i(\boldsymbol{y}') - f_i(\boldsymbol{u}^{(k)}) + \boldsymbol{\epsilon}_i^{(k)} &= f_i(\boldsymbol{y}') - f_i(\boldsymbol{u}') + f_i(\boldsymbol{u}') - f_i(\boldsymbol{u}^{(k)}) + \boldsymbol{\epsilon}_i^{(k)} \\ &\leq \max_{i=1,\dots,m}[f_i(\boldsymbol{y}') - f_i(\boldsymbol{u}')] + f_i(\boldsymbol{u}') - f_i(\boldsymbol{u}^{(k)}) + \frac{1}{k} \qquad (\text{by } \boldsymbol{\epsilon}^{(k)}_{\max} < \tfrac{1}{k}) \\ &< 0, \end{aligned}$$

which means that $f(\boldsymbol{y}') - f(\boldsymbol{u}^{(k)}) + \boldsymbol{\epsilon}^{(k)} \in -\mathbb{R}^m_{++}$, i.e., $\boldsymbol{u}^{(k)} \notin \mathcal{S}_w^{\boldsymbol{\epsilon}^{(k)}}$, a contradiction. Hence, $\boldsymbol{u}' \in \mathcal{S}_w$ and $\operatorname{dist}(\boldsymbol{u}', \mathcal{S}_w) = 0$. However, due to the continuity of distance function, one has

$$\operatorname{dist}(\boldsymbol{u}', \mathcal{S}_w) = \lim_{k\to\infty} \operatorname{dist}(\boldsymbol{u}^{(k)}, \mathcal{S}_w) \geq d > 0,$$

a contradiction. $\square$

Furthermore, the following proposition allows us to study the set $\mathcal{S}_w^{\boldsymbol{\epsilon}}$ of all weakly $\boldsymbol{\epsilon}$-efficient solutions by means of sublevels of $\psi(\boldsymbol{x})$.

Proposition 3.3. *For any $\boldsymbol{\epsilon} \in \mathbb{R}^m_+$, we have*

$$\{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) \le \boldsymbol{\epsilon}_{\min}\} \subset \mathcal{S}_w^{\boldsymbol{\epsilon}} \subset \{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) \le \boldsymbol{\epsilon}_{\max}\}. \tag{2}$$

Particularly, if $\boldsymbol{\epsilon}_{\max} = \boldsymbol{\epsilon}_{\min}$, then

$$\{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) \le \boldsymbol{\epsilon}_{\min} = \boldsymbol{\epsilon}_{\max}\} = \mathcal{S}_w^{\boldsymbol{\epsilon}}.$$

Proof. To show the first relation in (2), suppose to the contrary that there exists $\boldsymbol{u} \in \Omega$ such that $\psi(\boldsymbol{u}) \le \boldsymbol{\epsilon}_{\min}$ but $\boldsymbol{u} \notin \mathcal{S}_w^{\boldsymbol{\epsilon}}$. Then, there exists $\boldsymbol{y}' \in \Omega$ such that $f(\boldsymbol{y}') - f(\boldsymbol{u}) + \boldsymbol{\epsilon} \in -\mathbb{R}^m_{++}$, i.e., $f_i(\boldsymbol{u}) - f_i(\boldsymbol{y}') - \boldsymbol{\epsilon}_i > 0$ for each $i = 1, \ldots, m$. Thus, $\min_{i=1,\ldots,m}[f_i(\boldsymbol{u}) - f_i(\boldsymbol{y}')] > \boldsymbol{\epsilon}_{\min}$ which implies that $\psi(\boldsymbol{u}) > \boldsymbol{\epsilon}_{\min}$, a contradiction.

Now, fix a point $\boldsymbol{u} \in \mathcal{S}_w^{\boldsymbol{\epsilon}}$. For any $\boldsymbol{y} \in \Omega$, by definition, there exists $k_{\boldsymbol{y}} \in \{1, \ldots, m\}$ depending on $\boldsymbol{y}$ such that $f_{k_{\boldsymbol{y}}}(\boldsymbol{y}) - f_{k_{\boldsymbol{y}}}(\boldsymbol{u}) + \boldsymbol{\epsilon}_{k_{\boldsymbol{y}}} \ge 0$. Then, $\min_{i=1,\ldots,m}[f_i(\boldsymbol{u}) - f_i(\boldsymbol{y})] \le \boldsymbol{\epsilon}_{k_{\boldsymbol{y}}}$ for all $\boldsymbol{y} \in \Omega$, and hence

$$\psi(\boldsymbol{u}) = \max_{\boldsymbol{y} \in \Omega} \min_{i=1,\ldots,m} [f_i(\boldsymbol{u}) - f_i(\boldsymbol{y})] \le \boldsymbol{\epsilon}_{\max},$$

thus, the second relation in (2) holds. Consequently, the conclusion follows. □

4. Approximations of weakly ($\boldsymbol{\epsilon}$-)efficient solution set

In this section, we will construct polynomial approximations of the achievement function $\psi(\boldsymbol{x})$ from above and use their sublevel sets to approximate the set of all weakly ($\boldsymbol{\epsilon}$-)efficient solutions to (VROP). The construction of these polynomial approximations of $\psi(\boldsymbol{x})$ is inspired by [22] and can be reduced to SDP problems. As Ω is compact, after a possible re-scaling of the g_j's, we may and will assume that $\Delta := [-1, 1]^n \supseteq \Omega$ in the rest of this paper.

4.1. **Approximations of achievement function.** To construct polynomial approximations of $\psi(\boldsymbol{x})$, we need first compute upper and lower bounds of $f_i(\boldsymbol{x})$, $i = 1, \ldots, m$, over Ω. To this end, for each $i = 1, \ldots, m$, we compute a number $f_i^{\text{lower}} \in \mathbb{R}$ satisfying

$$\begin{aligned}
&p_i(\boldsymbol{x}) - f_i^{\text{lower}} q_i(\boldsymbol{x}) = \sigma_{i,0}(\boldsymbol{x}) + \sum_{j=1}^{r} \sigma_{i,j}(\boldsymbol{x}) g_j(\boldsymbol{x}) + \sum_{j=1}^{n} \sigma_{i,r+j}(\boldsymbol{x})(1 - x_j^2),\\
&\sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[\boldsymbol{x}],\ j = 1, \ldots, r+n,\ \deg(\sigma_{i,0}) \le 2k_i,\ k_i \in \mathbb{N},\\
&\deg(\sigma_{i,j} g_j) \le 2k_i,\ j = 1, \ldots, r,\ \deg(\sigma_{i,r+j}(1 - x_j^2)) \le 2k_i,\ j = 1, \ldots, n,
\end{aligned} \tag{3}$$

which is equivalent to an SDP feasibility problem (c.f. [23]). Under (**A1-2**), each $\frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})}$ is bounded from below on Ω and $p_i(\boldsymbol{x}) - f_i^{\text{lower}} q_i(\boldsymbol{x}) > 0$ on Ω for any $f_i^{\text{lower}} < \min_{\boldsymbol{x} \in \Omega} \frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})}$. Hence, by Putinar's Positivstellensatz, a number f_i^{lower} satisfying (3) always exists for k_i large enough (note that Assumption 2.1 holds due to the redundant polynomials $1 - x_j^2$, $j = 1, \dots, n$, added in (3)). Clearly, it holds that

$$f^{\text{lower}} := \min_{i=1,\dots,m} f_i^{\text{lower}} \le \min_{i=1,\dots,m,\ \boldsymbol{x}\in\Omega} \frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})}.$$

Similarly, replace $p_i(\boldsymbol{x}) - f_i^{\text{lower}} q_i(\boldsymbol{x})$ in (3) by $f_i^{\text{upper}} q_i(\boldsymbol{x}) - p_i(\boldsymbol{x})$, where f_i^{upper} denotes another real number. Then similarly, such a number f_i^{upper} exists for k_i large enough and can be computed by solving another SDP feasibility problem. Then, we have

$$f^{\text{upper}} := \max_{i=1,\dots,m} f_i^{\text{upper}} \ge \max_{i=1,\dots,m,\ \boldsymbol{x}\in\Omega} \frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})}.$$

Now, we deal with the achievement function $\psi(\boldsymbol{x})$ over Δ from the viewpoint of polynomial optimization problems. For each $\boldsymbol{x} \in \mathbb{R}^n$, it holds that

$$\begin{aligned}
\psi(\boldsymbol{x}) &:= \sup_{\boldsymbol{y}\in\Omega} \min_{i=1,\dots,m} [f_i(\boldsymbol{x}) - f_i(\boldsymbol{y})] \\
&= \sup_{\boldsymbol{y}\in\Omega} \min_{i=1,\dots,m} \left[\frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})} - \frac{p_i(\boldsymbol{y})}{q_i(\boldsymbol{y})}\right] \\
&= \sup_{\boldsymbol{y}\in\Omega, z\in\mathbb{R}} \left\{ z \colon \frac{p_i(\boldsymbol{x})}{q_i(\boldsymbol{x})} - \frac{p_i(\boldsymbol{y})}{q_i(\boldsymbol{y})} \ge z,\ i = 1, \dots, m \right\}.
\end{aligned}$$

For any $\boldsymbol{x} \in \Delta$, let

$$\begin{cases}
\tilde{\psi}(\boldsymbol{x}) := \max\limits_{\boldsymbol{y}\in\mathbb{R}^n, z\in\mathbb{R}} & z \\
\qquad\qquad \text{s.t.} & p_i(\boldsymbol{x})q_i(\boldsymbol{y}) - p_i(\boldsymbol{y})q_i(\boldsymbol{x}) - z q_i(\boldsymbol{x}) q_i(\boldsymbol{y}) \ge 0,\ i = 1, \dots, m, \\
& \boldsymbol{y} \in \Omega,\ z \in [f^{\text{lower}} - f^{\text{upper}}, f^{\text{upper}} - f^{\text{lower}}].
\end{cases} \tag{4}$$

In other words, $\tilde{\psi}(\boldsymbol{x})$ over Δ can be seen as the optimal value function of the *parameter* polynomial optimization problem (4). Under (**A1-2**), we have

Proposition 4.1. *$\tilde{\psi}(\boldsymbol{x}) = \psi(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega$. Hence, Propositions 3.1 and 3.3 also hold for $\tilde{\psi}$.*

Next, we construct polynomial approximations of $\tilde{\psi}$ over Δ from above by means of the SDP method proposed in [22], and use their sublevel sets to approximate the set of all weakly ($\boldsymbol{\epsilon}$-)efficient solutions to (VROP).

Consider the following sets

$$\mathbf{K} := \left\{ (\boldsymbol{x}, \boldsymbol{y}, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \ :\ \begin{cases} p_i(\boldsymbol{x})q_i(\boldsymbol{y}) - p_i(\boldsymbol{y})q_i(\boldsymbol{x}) - z q_i(\boldsymbol{x}) q_i(\boldsymbol{y}) \ge 0, \\ i = 1, \dots, m,\ \boldsymbol{x} \in \Delta,\ \boldsymbol{y} \in \Omega, \\ z \in [f^{\text{lower}} - f^{\text{upper}}, f^{\text{upper}} - f^{\text{lower}}] \end{cases} \right\},$$

and

$$\mathbf{K}_{\boldsymbol{x}} := \{(\boldsymbol{y}, z) \in \mathbb{R}^n \times \mathbb{R} \ :\ (\boldsymbol{x}, \boldsymbol{y}, z) \in \mathbf{K}\}, \quad \text{for } \boldsymbol{x} \in \Delta.$$

Then it is clear that $\mathbf{K}$ is compact and for any $\boldsymbol{x} \in \Delta$, $\tilde{\psi}(\boldsymbol{x}) = \max_{(\boldsymbol{y},z)\in\mathbf{K}_{\boldsymbol{x}}} z$.

As proved in [22, Theorem 1], a sequence of polynomial approximations of $\tilde{\psi}(\boldsymbol{x})$ on Δ from above exists mainly due to the Stone–Weierstrass theorem.

Proposition 4.2. (c.f. [22, Theorem 1]) *There exists a sequence of polynomials* $\{\psi_k \in \mathbb{R}[\boldsymbol{x}]\colon k \in \mathbb{N}\}$ *such that* $\psi_k(\boldsymbol{x}) \geq \tilde{\psi}(\boldsymbol{x})$ *for all* $\boldsymbol{x} \in \Delta$, *and* $\{\psi_k\}_{k\in\mathbb{N}}$ *converges to* $\tilde{\psi}$ *in* $L_1(\Delta)$, *i.e.,*

$$\lim_{k\to\infty} \int_\Delta |\psi_k(\boldsymbol{x}) - \tilde{\psi}(\boldsymbol{x})| d\boldsymbol{x} = 0.$$

Let $\{\psi_k \in \mathbb{R}[\boldsymbol{x}]\colon k \in \mathbb{N}\}$ be as in Proposition 4.2. For any $\delta > 0$ and $k \in \mathbb{N}$, denote

$$\mathcal{A}(\delta, k) := \{\boldsymbol{x} \in \Omega\colon \psi_k(\boldsymbol{x}) \leq \delta\}.$$

For any $\delta > 0$, with a slight abuse of notation, we denote $\mathcal{S}_w^\delta := \mathcal{S}_w^{\boldsymbol{\epsilon}}$, where $\boldsymbol{\epsilon} = (\delta, \dots, \delta)$. The following result can be derived by slightly modifying the proof of [22, Theorem 3]. It shows that we can approximate the set $\mathcal{S}_w^\delta$ by the sequence $\{\mathcal{A}(\delta, k)\}_{k\in\mathbb{N}}$.

Theorem 4.1. *For any* $\delta > 0$, *we have* $\mathcal{A}(\delta, k) \subset \mathcal{S}_w^\delta$ *and*

$$\mathrm{vol}\left(\{\boldsymbol{x} \in \Omega\colon \psi(\boldsymbol{x}) < \delta\}\right) \leq \lim_{k\to\infty} \mathrm{vol}\left(\mathcal{A}(\delta, k)\right) \leq \mathrm{vol}\left(\{\boldsymbol{x} \in \Omega\colon \psi(\boldsymbol{x}) \leq \delta\}\right) = \mathrm{vol}\left(\mathcal{S}_w^\delta\right). \tag{5}$$

Consequently, if $\mathrm{vol}\left(\{\boldsymbol{x} \in \Omega\colon \psi(\boldsymbol{x}) = \delta\}\right) = 0$, *then* $\lim_{k\to\infty} \mathrm{vol}\left(\mathcal{S}_w^\delta \setminus \mathcal{A}(\delta, k)\right) = 0$.

Proof. By Proposition 3.3, it is clear that $\mathcal{A}(\delta, k) \subset \mathcal{S}_w^\delta$. By Proposition 4.2, ψ_k converges to $\tilde{\psi}$ in measure, that is, for every $\alpha > 0$,

$$\lim_{k\to\infty} \mathrm{vol}\left(\{\boldsymbol{x} \in \Delta\colon |\psi_k(\boldsymbol{x}) - \tilde{\psi}(\boldsymbol{x})| \geq \alpha\}\right) = 0. \tag{6}$$

Consequently, for every $\ell \geq 1$, it holds that

$$\begin{aligned}
&\mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \psi(\boldsymbol{x}) \leq \delta + \tfrac{-1}{\ell}\right\}\right) \\
= {}&\mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \tilde{\psi}(\boldsymbol{x}) \leq \delta + \tfrac{-1}{\ell}\right\}\right) && \text{(by Proposition 4.1)} \\
= {}&\mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \tilde{\psi}(\boldsymbol{x}) \leq \delta + \tfrac{-1}{\ell}\right\} \cap \{\boldsymbol{x} \in \Omega\colon \psi_k(\boldsymbol{x}) > \delta\}\right) \\
&+ \mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \tilde{\psi}(\boldsymbol{x}) \leq \delta + \tfrac{-1}{\ell}\right\} \cap \{\boldsymbol{x} \in \Omega\colon \psi_k(\boldsymbol{x}) \leq \delta\}\right) \\
= {}&\lim_{k\to\infty} \mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \tilde{\psi}(\boldsymbol{x}) \leq \delta + \tfrac{-1}{\ell}\right\} \cap \{\boldsymbol{x} \in \Omega\colon \psi_k(\boldsymbol{x}) \leq \delta\}\right) && \text{(by (6))} \\
\leq {}&\lim_{k\to\infty} \mathrm{vol}\left(\{\boldsymbol{x} \in \Omega\colon \psi_k(\boldsymbol{x}) \leq \delta\}\right) \\
\leq {}&\mathrm{vol}\left(\left\{\boldsymbol{x} \in \Omega\colon \tilde{\psi}(\boldsymbol{x}) \leq \delta\right\}\right) \\
= {}&\mathrm{vol}\left(\mathcal{S}_w^\delta\right). && \text{(by Propositions 3.3 and 4.1)}
\end{aligned}$$

Now, taking $\ell \to \infty$ yields (5) and the conclusion. $\square$

Corollary 4.1. *The following assertions are true.*

(i) *For any* $d > 0$, *there exists* $\delta(d) > 0$ *depending on* d *such that*

$$\mathcal{A}(\delta, k) \subset \mathcal{S}_w + d\mathbf{B}$$

holds for any $\delta < \delta(d)$ *and any* $k \in \mathbb{N}$.

(ii) *For $d > 0$ and any $\delta > 0$ with* $\mathrm{vol}(\{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) = \delta\}) = 0$, *there exists* $k(d, \delta) \in \mathbb{N}$ *depending on δ and d such that*

$$\mathcal{S}_w \cap \mathrm{cl}\,(\mathrm{int}\,(\Omega \setminus \mathcal{S}_w)) \subset \mathcal{A}(\delta, k) + d\mathbf{B}$$

holds for any $k > k(d, \delta)$.

Proof. (i) Since $\mathcal{A}(\delta, k) \subset \mathcal{S}_w^\delta$ for any $k \in \mathbb{N}$ by Theorem 4.1, the existence of $\delta(d)$ is a direct consequence of Proposition 3.2.

(ii) Let $\boldsymbol{u} \in \mathcal{S}_w \cap \mathrm{cl}\,(\mathrm{int}\,(\Omega \setminus \mathcal{S}_w)) \neq \emptyset$, then $\tilde{\psi}(\boldsymbol{u}) = 0$ by Propositions 3.1 (i) and 4.1, and there exists a sequence $\{\boldsymbol{u}^{(l)}\}_{l \in \mathbb{N}} \subset \mathrm{int}\,(\Omega \setminus \mathcal{S}_w)$ such that $\lim_{l \to \infty} \boldsymbol{u}^{(l)} = \boldsymbol{u}$. Fix the numbers $d, \delta > 0$. By the continuity of $\tilde{\psi}$ on Ω (Proposition 4.1), there exists $l_0 \in \mathbb{N}$ depending on d and δ such that $\tilde{\psi}(\boldsymbol{u}^{(l_0)}) < \delta$ and $\|\boldsymbol{u}^{(l_0)} - \boldsymbol{u}\| < d$. As $\boldsymbol{u}^{(l_0)} \in \mathrm{int}\,(\Omega \setminus \mathcal{S}_w)$, by the continuity of $\tilde{\psi}$ again, there is a neighborhood $\mathcal{O}^{(l_0)} \subset \Omega$ of $\boldsymbol{u}^{(l_0)}$ such that $\tilde{\psi}(\boldsymbol{x}) < \delta$ and $\|\boldsymbol{x} - \boldsymbol{u}\| < d$ for all $\boldsymbol{x} \in \mathcal{O}^{(l_0)}$. Proposition 3.3 implies that $\mathcal{O}^{(l_0)} \subset \mathcal{S}_w^\delta$. Then, we show that there exists $k(d, \delta) \in \mathbb{N}$ such that for any $k > k(d, \delta)$, it holds that $\mathcal{A}(\delta, k) \cap \mathcal{O}^{(l_0)} \neq \emptyset$ which means that $\boldsymbol{u} \in \mathcal{A}(d, \delta) + d\mathbf{B}$ and the conclusion follows. To the contrary, suppose that such $k(d, \delta)$ does not exist, then there is subsequence $\{\mathcal{A}(\delta, k_j)\}_{j \in \mathbb{N}}$ with $k_j \to \infty$ such that $\mathcal{A}(\delta, k_j) \cap \mathcal{O}^{(l_0)} = \emptyset$ for all k_j. Then, $\mathrm{vol}(\mathcal{S}_w^\delta \setminus \mathcal{A}(\delta, k_j)) \geq \mathrm{vol}(\mathcal{O}^{(l_0)}) > 0$ for all k_j. As $\mathrm{vol}(\{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) = \delta\}) = 0$, it contradicts the conclusion in Theorem 4.2. $\square$

Remark 4.1. From Corollary 4.1 and its proof, we can see that

(i) If $\mathcal{S}_w \cap \mathrm{cl}\,(\mathrm{int}\,(\Omega \setminus \mathcal{S}_w)) \neq \emptyset$, then for any $\delta > 0$, $\mathcal{A}(\delta, k) \neq \emptyset$ for k large enough. In fact, we have $\mathcal{O}^{(l_0)} \subset \{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) < \delta\}$ for the neighborhood $\mathcal{O}^{(l_0)}$ in the proof of Corollary 4.1. Then, (5) implies that $\mathcal{A}(\delta, k) \neq \emptyset$ for k large enough.

(ii) Suppose there is a sequence $\{\delta_i\}_{i \in \mathbb{N}}$ with $\delta_i \downarrow 0$ such that $\mathrm{vol}(\{\boldsymbol{x} \in \Omega \colon \psi(\boldsymbol{x}) = \delta_i\}) = 0$ holds for all i and $\mathcal{S}_w = \mathcal{S}_w \cap \mathrm{cl}\,(\mathrm{int}\,(\Omega \setminus \mathcal{S}_w))$, then Corollary 4.1 (i) and (ii) indicate that the whole set of the weakly efficient solutions of (VROP) can be approximated arbitrarily well by $\mathcal{A}(\delta, k)$ with sufficiently small $\delta > 0$ and sufficiently large $k \in \mathbb{N}$.

4.2. **Computational aspects.** Now we follow the scheme proposed in [22, Section 3.3] to construct a sequence of polynomials $(\psi_k)_{k \in \mathbb{N}} \in \mathbb{R}[\boldsymbol{x}]$ as defined in Proposition 4.2.

We denote the following $m + r + 2n + 1$ polynomials in $\mathbb{R}[\boldsymbol{x}, \boldsymbol{y}, z]$

$$\begin{aligned}
&h_{1,1}(\boldsymbol{x}, \boldsymbol{y}, z) = p_1(\boldsymbol{x})q_1(\boldsymbol{y}) - p_1(\boldsymbol{y})q_1(\boldsymbol{x}) - zq_1(\boldsymbol{x})q_1(\boldsymbol{y}), \ldots, \\
&h_{1,m}(\boldsymbol{x}, \boldsymbol{y}, z) = p_m(\boldsymbol{x})q_m(\boldsymbol{y}) - p_m(\boldsymbol{y})q_m(\boldsymbol{x}) - zq_m(\boldsymbol{x})q_m(\boldsymbol{y}), \\
&h_{2,1}(\boldsymbol{x}, \boldsymbol{y}, z) = g_1(\boldsymbol{y}), \ldots, h_{2,r}(\boldsymbol{x}, \boldsymbol{y}, z) = g_r(\boldsymbol{y}), \\
&h_{2,r+1}(\boldsymbol{x}, \boldsymbol{y}, z) = 1 - y_1^2, \ldots, h_{2,r+n}(\boldsymbol{x}, \boldsymbol{y}, z) = 1 - y_n^2, \\
&h_{3,1}(\boldsymbol{x}, \boldsymbol{y}, z) = 1 - x_1^2, \ldots, h_{3,n}(\boldsymbol{x}, \boldsymbol{y}, z) = 1 - x_n^2, \\
&h_{4,1}(\boldsymbol{x}, \boldsymbol{y}, z) = (f^{\mathrm{upper}} - f^{\mathrm{lower}})^2 - z^2.
\end{aligned} \tag{7}$$

Denote by $J_1 = \{1, \ldots, m\}$, $J_2 = \{1, \ldots, r+n\}$, $J_3 = \{1, \ldots, n\}$ and $J_4 = \{1\}$. Then,

$$\mathbf{K} = \{(\boldsymbol{x}, \boldsymbol{y}, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} : \ h_{i,j}(\boldsymbol{x}, \boldsymbol{y}, z) \geq 0, \ i = 1, \ldots, 4, \ j \in J_i\}.$$

Let λ be the scaled Lebesgue measure on Δ, i.e., $d\lambda(\boldsymbol{x}) = d\boldsymbol{x}/2^n$, and

$$\gamma_\alpha := \int_\Delta \boldsymbol{x}^\alpha d\lambda(\boldsymbol{x}) = \begin{cases} 0, & \text{if } \alpha_i \text{ is odd for some } i \\ \displaystyle\prod_{i=1}^n (\alpha_i + 1)^{-1}, & \text{otherwise} \end{cases}$$

be the moment of λ for each $\alpha \in \mathbb{N}^n$.

For each $k \in \mathbb{N}$, with $k \geq \max\left\{\lceil \frac{\deg h_{i,j}}{2} \rceil, \ i = 1, \ldots, 4, \ j \in J_i\right\}$, consider the following optimization problem,

$$\begin{cases} \rho_k^* := \inf\limits_{\phi, \sigma_0, \sigma_{i,j}} \displaystyle\int_\Delta \phi(\boldsymbol{x}) d\lambda(\boldsymbol{x}) \left(= \sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \gamma_\alpha \right) \\ \qquad \text{s.t.} \quad \phi(\boldsymbol{x}) = \displaystyle\sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \boldsymbol{x}^\alpha \in \mathbb{R}[\boldsymbol{x}]_{2k}, \ c_\alpha \in \mathbb{R}, \\ \qquad \qquad \phi(\boldsymbol{x}) - z = \sigma_0 + \displaystyle\sum_{i=1}^4 \sum_{j \in J_i} \sigma_{i,j} h_{i,j}, \ \sigma_0, \sigma_{i,j} \in \Sigma^2[\boldsymbol{x}, \boldsymbol{y}, z], \\ \qquad \qquad \deg(\sigma_0), \deg(\sigma_{i,j} h_{i,j}) \leq 2k, \ i = 1, \ldots, 4, \ j \in J_i, \end{cases} \tag{P$_k$}$$

which can be reduced to an SDP problem (c.f. [18, 20]). Clearly, for any $(\phi, \sigma_0, \sigma_{i,j})$ feasible to (P_k), we have $\phi(\boldsymbol{x}) \geq \tilde{\psi}(\boldsymbol{x})$ on Δ. The following result follows directly from [22, Theorem 5] and we include here a brief proof for the sake of completeness. It shows that we can compute the sequence of polynomials $\{\psi_k \in \mathbb{R}[\boldsymbol{x}] \colon k \in \mathbb{N}\}$ in Proposition 4.2 by solving (P_k).

Theorem 4.2. *We have* $\lim_{k\to\infty} \rho_k^* = \int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x})$. *Consequently, let* $\left(\psi_k, \sigma_0^{(k)}, \sigma_{i,j}^{(k)}\right)$ *be a nearly optimal solution to* (P_k), *e.g.,* $\int_\Delta \psi_k d\lambda(\boldsymbol{x}) \leq \rho_k^* + 1/k$, *then* $\psi_k(\boldsymbol{x}) \geq \tilde{\psi}(\boldsymbol{x})$ *on* Δ *and*

$$\lim_{k\to\infty} \int_\Delta |\psi_k(\boldsymbol{x}) - \tilde{\psi}(\boldsymbol{x})| d\lambda(\boldsymbol{x}) = 0.$$

Proof. We only need to prove that $\lim_{k\to\infty} \rho_k^* = \int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x})$. Consider the following infinite-dimensional linear program

$$\begin{cases} \rho^* := \inf\limits_{\phi} \displaystyle\int_\Delta \phi(\boldsymbol{x}) d\lambda(\boldsymbol{x}) \left(= \sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \gamma_\alpha \right) \\ \qquad \text{s.t. } \phi(\boldsymbol{x}) = \displaystyle\sum_{\alpha \in \mathbb{N}^n} c_\alpha \boldsymbol{x}^\alpha \in \mathbb{R}[\boldsymbol{x}], \ c_\alpha \in \mathbb{R}, \\ \qquad \qquad \phi(\boldsymbol{x}) - z \geq 0, \quad \forall \ (\boldsymbol{x}, \boldsymbol{y}, z) \in \mathbf{K}. \end{cases}$$

It is clear that Δ, $\mathbf{K}$ are compact and $\mathbf{K}_{\boldsymbol{x}}$ is nonempty for every $\boldsymbol{x} \in \Delta$. Then, by [22, Corollary 2.6], it holds that $\rho^* = \int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x})$. Let $(\phi_\ell)_{\ell \in \mathbb{N}}$ be a minimizing sequence of the above problem. For any $\ell \in \mathbb{N}$, let $\phi'_\ell(\boldsymbol{x}) = \phi_\ell(\boldsymbol{x}) + 1/\ell$, then we have $\phi'_\ell(\boldsymbol{x}) - z \geq 1/\ell > 0$ on $\mathbf{K}$. Notice that

$$2n + (f^{\text{upper}} - f^{\text{lower}})^2 - \sum_{i=1}^{n} \left(x_i^2 + y_i^2\right) - z^2 = \sum_{j=r+1}^{r+n} h_{2,j} + \sum_{j=1}^{n} h_{3,j} + h_{4,1},$$

that is, Assumption 2.1 holds for the defining polynomials of $\mathbf{K}$. Therefore, by Putinar's Positivstellensatz (Theorem 2.1), there exists $k_\ell \in \mathbb{N}$ and $\sigma_0^{(\ell)}, \sigma_{i,j}^{(\ell)} \in \Sigma^2[\boldsymbol{x}, \boldsymbol{y}, z]$ such that $(\phi'_\ell, \sigma_0^{(\ell)}, \sigma_{i,j}^{(\ell)})$ is a feasible solution to (P_{k_ℓ}). Note that $\rho^* \leq \rho_k^*$ holds for any $k \in \mathbb{N}$. Then, it implies that

$$\int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x}) = \rho^* \leq \rho_{k_\ell}^* \leq \int_\Delta \phi_\ell(\boldsymbol{x}) d\lambda(\boldsymbol{x}) + \frac{1}{\ell} \downarrow \rho^* = \int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x}).$$

As ρ_k^* is monotone, we have $\lim_{k\to\infty} \rho_k^* = \int_\Delta \tilde{\psi}(\boldsymbol{x}) d\lambda(\boldsymbol{x})$. □

Next, we propose a sparse version of the SDP problem (P_k) by exploiting its sparsity pattern, which reduces the computational costs at the order k. Add a redundant polynomial

$$h_{1,m+1}(\boldsymbol{x}, \boldsymbol{y}, z) = 2n + (f^{\text{upper}} - f^{\text{lower}})^2 - \sum_{i=1}^{n} \left(x_i^2 + y_i^2\right) - z^2$$

in (7) and reset $J_1 = \{1, \ldots, m+1\}$. Denote the following subsets of variables $I_1 = \{\boldsymbol{x}, \boldsymbol{y}, z\}$, $I_2 = \{\boldsymbol{y}\}$, $I_3 = \{\boldsymbol{x}\}$ and $I_4 = \{z\}$. For $i = 1, \ldots, 4$, denote by $\mathbb{R}[I_i]$ the ring of real polynomials in the variables in I_i. Then, the following conditions hold.

(i) For each $i = 1, 2, 3$, there exists some $s \leq i$ such that $I_{i+1} \cap \bigcup_{j=1}^{i} I_j \subseteq I_s$;

(ii) For each $i = 1, \ldots, 4$, and each $j \in J_i$, $h_{i,j} \in \mathbb{R}[I_i]$;

(iii) $\sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \boldsymbol{x}^\alpha - z$ in (P_k) is the difference of two polynomials in $\mathbb{R}[I_3]$ and $\mathbb{R}[I_4]$, respectively.

Then, by the sparse version of Putinar's Positivstellensatz (Theorem 2.2), we can construct a sparse version of (P_k) as

$$\begin{cases} \tilde{\rho}_k^* := \inf\limits_{\phi, \sigma_{i,0}, \sigma_{i,j}} \displaystyle\int_\Delta \phi(\boldsymbol{x}) d\lambda(\boldsymbol{x}) \left(= \sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \gamma_\alpha \right) \\ \qquad\qquad \text{s.t.} \quad \phi(\boldsymbol{x}) = \displaystyle\sum_{\alpha \in \mathbb{N}_{2k}^n} c_\alpha \boldsymbol{x}^\alpha \in \mathbb{R}[\boldsymbol{x}]_{2k}, \ c_\alpha \in \mathbb{R}, \\ \qquad\qquad\qquad \phi(\boldsymbol{x}) - z = \displaystyle\sum_{i=1}^{4} \left(\sigma_{i,0} + \sum_{j \in J_i} \sigma_{i,j} h_{i,j} \right), \ \sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[I_i], \\ \qquad\qquad\qquad \deg(\sigma_{i,0}), \deg(\sigma_{i,j} h_{i,j}) \leq 2k, \ i = 1, \ldots, 4, \ j \in J_i. \end{cases} \tag{SP$_k$}$$

Theorem 4.3. *The statements for* (P_k) *in Theorem 4.2 also hold for* (SP_k).

Proof. Let ϕ'_ℓ be the polynomial in the proof of Theorem 4.2. Note that Assumption 2.2 holds by adding the redundant polynomial $h_{1,m+1}$. Then, by Theorem 2.2, there exists $\tilde{k}_\ell \in \mathbb{N}$ and $\sigma_{i,0}^{(\ell)}, \sigma_{i,j}^{(\ell)} \in \Sigma^2[I_i]$, $i = 1, \ldots, 4$, $j \in J_i$ such that $\left(\phi'_\ell, \sigma_{i,0}^{(\ell)}, \sigma_{i,j}^{(\ell)}\right)$ is a feasible solution to $(\mathrm{SP}_{\tilde{k}_\ell})$. Hence, the conclusion follows from the proof of Theorem 4.2. □

4.3. **Comparisons with existing SDP relaxation methods.** Now, we compare our method with the recent existing work in [31] and [29]. All the three methods can deal with vector (nonlinear) polynomial optimization problems by SDP relaxations, without convexity assumptions on the involved functions. For convenience, we assume that all objectives f_i's in (VROP) are polynomials, i.e., $q_i(\boldsymbol{x}) = 1$, $i = 1, \ldots, m$.

To get weakly efficient solutions to (VROP), Nie and Yang [31] used the linear scalarization and the Chebyshev scalarization techniques to scalarize (VROP) to a single objective polynomial optimization problem and solve it by the SDP relaxation method proposed in [30]. Precisely, for a given nonzero weighting parameter $w := (w_1, \ldots, w_m) \in \mathbb{R}^m$, the linear scalarization scalarizes the problem (VROP) to

$$\min\ w_1 f_1(\boldsymbol{x}) + \cdots + w_m f_m(\boldsymbol{x}) \quad \text{s.t.} \quad \boldsymbol{x} \in \Omega, \tag{8}$$

and the Chebyshev scalarization scalarizes the problem (VROP) to

$$\min_{\boldsymbol{x}\in\Omega} \max_{1\leq i\leq m} w_i(f_i(\boldsymbol{x}) - f_i^*), \tag{9}$$

where each f_i^* is the goal which decision maker wants to achieve for the objective f_i. In general, by the scalarizations (8) and (9), we can only find one or some particular (weakly) efficient solutions for a given weight w. Moreover, a serious drawback of linear scalarization is that it can not provide a solution among sunken parts of Pareto frontier due to "duality gap" of nonconvex cases (see Example 4.3). Instead, the sets $\{\mathcal{A}(\delta, k)\}$ computed by our method can approximate the whole set of weakly efficient solutions in some sense under certain conditions. The representation of $\mathcal{A}(\delta, k)$ as the intersection of the sublevel set of a single polynomial and the feasible set is more desirable in some applications. For example, it can be used in optimization problems with Pareto constraints (c.f. [14]). A Pareto constraint can be replaced by the polynomial inequality $\psi_k(\boldsymbol{x}) \leq \delta$ with small $\delta > 0$ and large $k \in \mathbb{N}$ (see Example 4.3).

On the other hand, Magron et al. [29] studied the problem (VROP) with $m = 2$. Rather than computing the weakly efficient solutions, they presented a method to approximate as closely as desired the Pareto curve which is the image of the objective functions over the set of weakly efficient solutions. To this end, they also considered the scalarizations (8) and (9), as well as the parametric sublevel set approximation method which is inspired by [11] and amounts to solving the following parametric problem

$$\min_{\boldsymbol{x}\in\Omega}\ f_2(\boldsymbol{x}) \quad \text{s.t.} \quad f_1(\boldsymbol{x}) \leq w, \tag{10}$$

with a parameter $w \in [\min_{\boldsymbol{x}\in\Omega} f_1(\boldsymbol{x}), \max_{\boldsymbol{x}\in\Omega} f_1(\boldsymbol{x})]$. By treating w in (8), (9) and (10) as a parameter and employing the "joint+marginal" approach proposed in [20], they associated each scalarization problem a hierarchy of SDP relaxations and obtained an approximation of the Pareto curve by solving an inverse problem (for (8) and (9)) or by building a polynomial underestimator (for (10)). Again, comparing with the approximate Pareto curve obtained in [29], it is more convenient to apply our explicit approximation $\mathcal{A}(\delta, k)$ of the weakly efficient solution set to optimization problems with Pareto constraint. Moreover, when using the scalarization problems (8) and (9), the approach in [29] requires that for almost all the values of the parameter w, these parametric problems (8) and (9) have a unique global minimizer. Namely, there should be a one-to-one correspondence between the points on the computed Pareto curves and the associated weakly efficient solutions in the feasible set. Note that our method does not have such restriction when approximating the set of weakly efficient solutions (see Example 4.4).

4.4. **Numerical experiments.** Here we present some numerical examples to illustrate the behavior of the sets $\mathcal{A}(\delta, k)$ in approximating $\mathcal{S}_w$ as $\delta \to 0$ and $k \to \infty$. We use the software Yalmip [27] to implement the problems (SP_k) and call the SDP solver SeDuMi [36] to solve the resulting SDP problems. For the examples with $m = 2$, to show how close the sets $\mathcal{A}(\delta, k)$ in approximating $\mathcal{S}_w$, we illustrate the corresponding images of $f(\Omega)$ and $f(\mathcal{A}(\delta, k))$. To this end, we choose a square containing Ω. For each point u on a uniform discrete grid inside the square, we check if $u \in \Omega$ (resp., $u \in \mathcal{A}(\delta, k)$). If so, we have $(f_1(u), f_2(u)) \in f(\Omega)$ (resp., $(f_1(u), f_2(u)) \in f(\mathcal{A}(\delta, k)))$ and we plot the point $(f_1(u), f_2(u))$ in grey (resp., in red) in the image plane.

Example 4.1. Consider the problem

$$\begin{cases} \mathrm{Min}_{\mathbb{R}^3_+} & \left(x_1,\ x_2,\ x_1^2 + x_2^2\right) \\ \text{s.t.} & \boldsymbol{x} \in \Omega_1 := \{\boldsymbol{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 \le 1\}. \end{cases}$$

Clearly, the set of all weakly efficient solution to this problem is

$$\mathcal{S}_w = \left\{\boldsymbol{x} \in \mathbb{R}^2 \colon x_1 \le 0, x_2 \le 0, x_1^2 + x_2^2 \le 1\right\}.$$

For any $\delta > 0$, by considering the four quadrants of $\mathbb{R}^2$ one by one, it is easy to check by definition that the set $\mathcal{S}_w^\delta$ consists of the following four sets

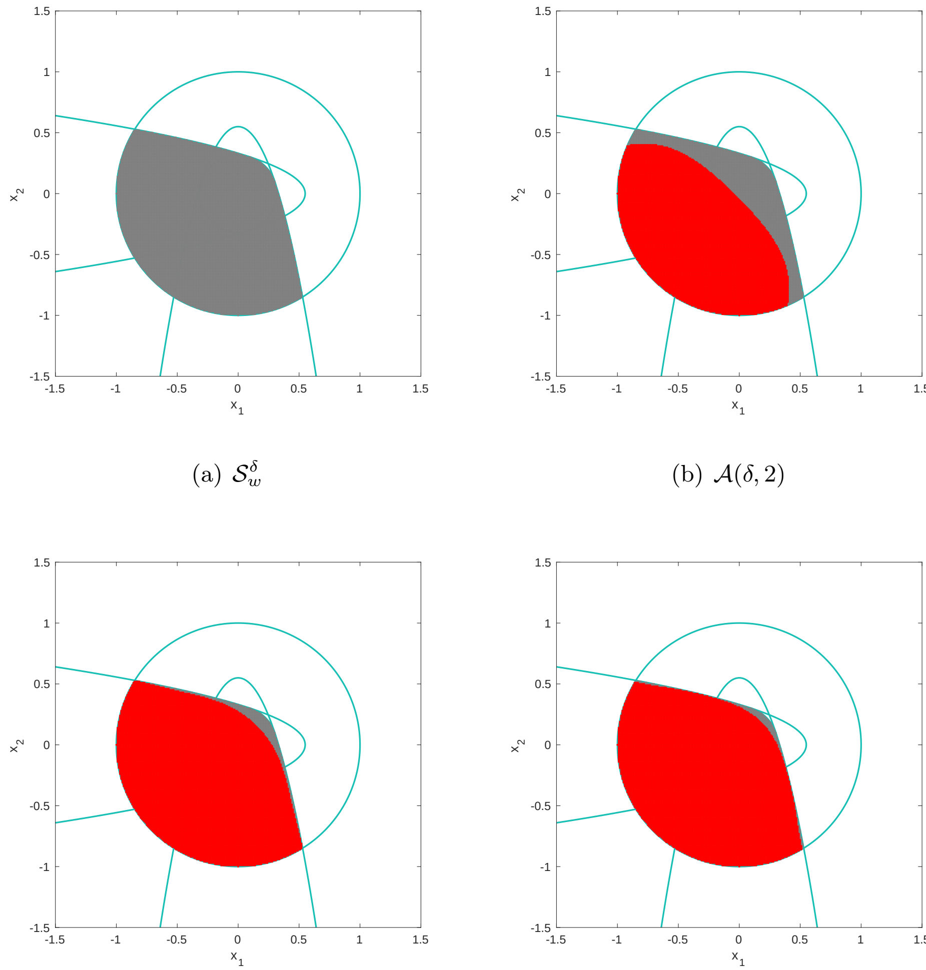
$$\begin{aligned} &\{\boldsymbol{x} \in \mathbb{R}^2 \colon x_1 \ge \delta,\ x_2 \ge \delta,\ x_1^2 + x_2^2 \le \delta\}, \\ &\{\boldsymbol{x} \in \mathbb{R}^2 \colon x_1 \le \delta,\ x_2 \ge \delta,\ x_2^2 + 2\delta x_1 - \delta - \delta^2 \le 0,\ x_1^2 + x_2^2 \le 1\}, \\ &\{\boldsymbol{x} \in \mathbb{R}^2 \colon x_1 \le \delta,\ x_2 \le \delta,\ x_1^2 + x_2^2 \le 1\}, \\ &\{\boldsymbol{x} \in \mathbb{R}^2 \colon x_1 \ge \delta,\ x_2 \le \delta,\ x_1^2 + 2\delta x_2 - \delta - \delta^2 \le 0,\ x_1^2 + x_2^2 \le 1\}. \end{aligned}$$

For $\delta = 0.1$, we show the set $\mathcal{S}_w^\delta$ and its approximations $\mathcal{A}(\delta, k)$, $k = 2, 3, 4$, in Figure 1. □

Example 4.2. To illustrate how the set $\mathcal{A}(\delta, k)$ behaves in approximating the set of weakly efficient solutions $\mathcal{S}_w$ as $\delta \to 0$ and $k \to \infty$, we consider the problem

$$\begin{cases} \mathrm{Min}_{\mathbb{R}^2_+} & \left(x_1,\ \dfrac{x_2^2 - 2x_1x_2 + 1}{x_2^2 + 1}\right) \\ \text{s.t.} & \boldsymbol{x} \in \Omega_2 := \left\{\boldsymbol{x} \in \mathbb{R}^2 \colon 1 - x_1^2 - x_2^2 \ge 0\right\}. \end{cases}$$

We plot the images $f(\mathcal{A}(\delta, k))$ with $\delta = 0.1, 0.05, 0.02$ and $k = 3, 4, 5$, as well as $f(\Omega_2)$, in Figure 2. □



(a) $\mathcal{S}_w^\delta$ (b) $\mathcal{A}(\delta, 2)$

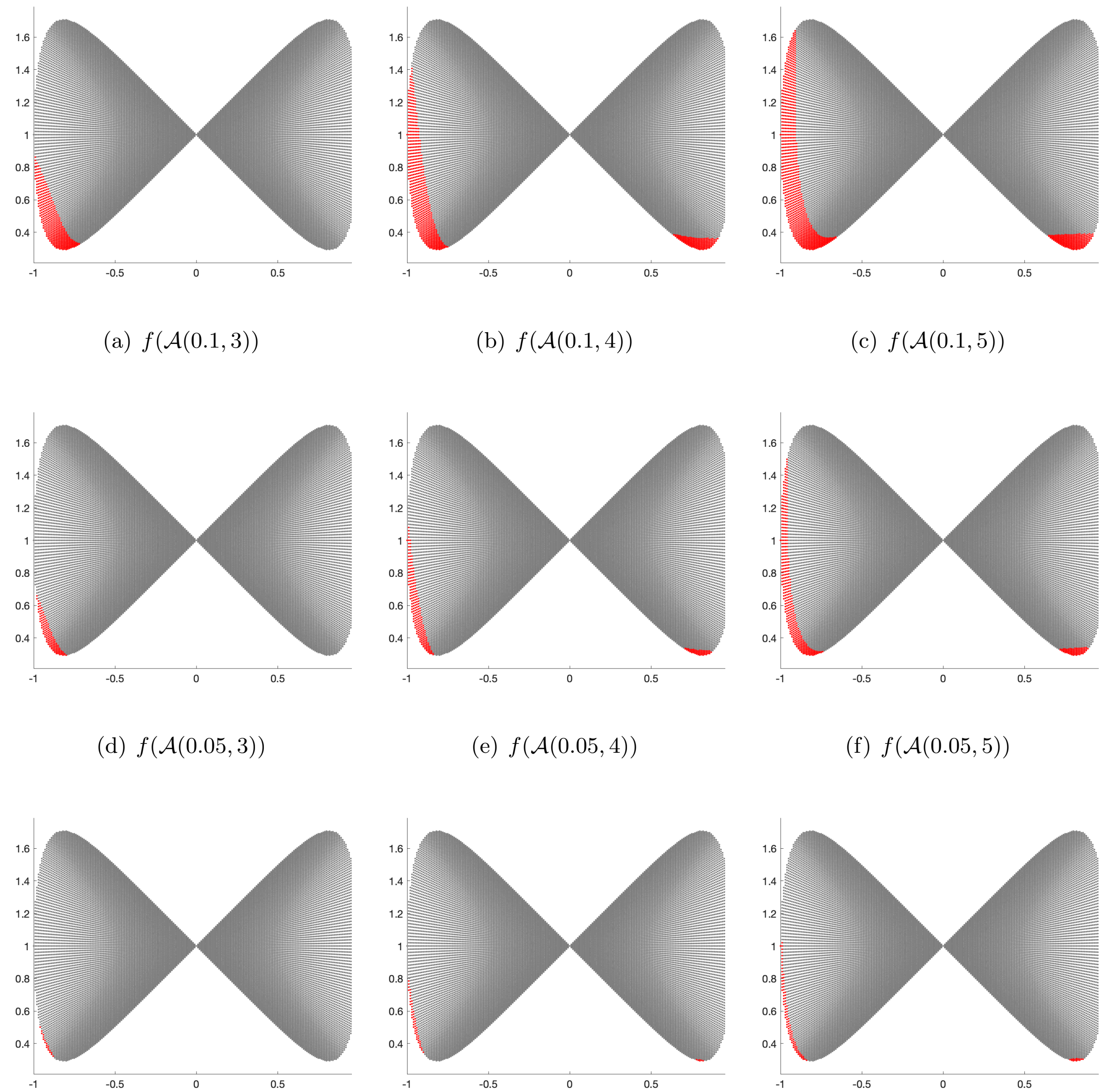
(c) $\mathcal{A}(\delta, 3)$ (d) $\mathcal{A}(\delta, 4)$

FIGURE 1. The set $\mathcal{S}_w^\delta$ and its approximations $\mathcal{A}(\delta, k)$ with $\delta = 0.1$, $k = 2, 3, 4$, in Example 4.1.

Example 4.3. Consider the problem

$$\begin{cases} \mathrm{Min}_{\mathbb{R}^2_+} & \left(\dfrac{\sqrt{2}}{2}(-x_1 + x_2),\ \dfrac{\sqrt{2}}{2}(x_1 + x_2)\right) \\ \text{s.t.} & \boldsymbol{x} \in \Omega_3 := \left\{\boldsymbol{x} \in \mathbb{R}^2 \colon g(\boldsymbol{x}) := x_2^2(1 - x_1^2) - (x_1^2 + 2x_2 - 1)^2 \geq 0\right\}, \end{cases}$$

In fact, the equality $g(\boldsymbol{x}) = 0$ defines the so-called *bicorn curve* as show in Figure 3 (a). Hence, the feasible set Ω_3 of this problem is the region enclosed by the bicorn curve and the image $f(\Omega_3)$ is obtained by rotating Ω_3 clockwise by 45° (Figure 3 (b)). It is clear that the weakly efficient solution set $\mathcal{S}_w$ consists of the points on the shorter path connecting the two singular points of the bicorn curve. As discussed in subsection 4.3, the linear scalarization (8) can only enable us to compute two points in $\mathcal{S}_w$, namely, the two singular points of the bicorn curve.



(a) $f(\mathcal{A}(0.1, 3))$ (b) $f(\mathcal{A}(0.1, 4))$ (c) $f(\mathcal{A}(0.1, 5))$

(d) $f(\mathcal{A}(0.05, 3))$ (e) $f(\mathcal{A}(0.05, 4))$ (f) $f(\mathcal{A}(0.05, 5))$

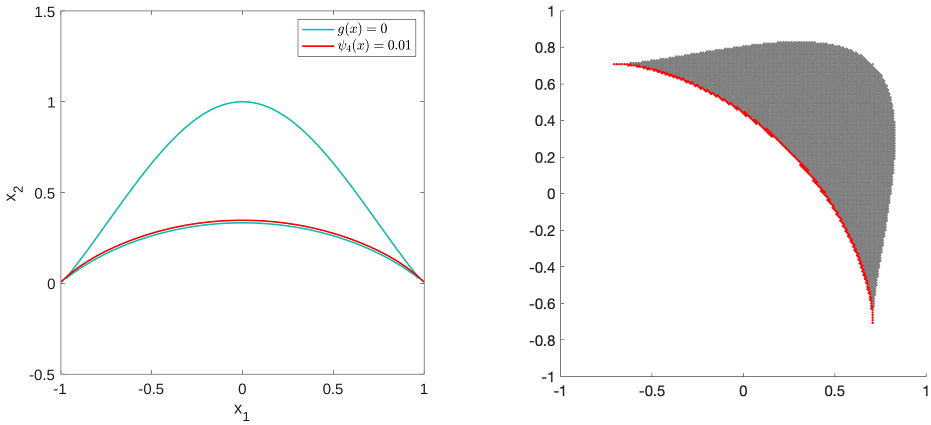
(g) $f(\mathcal{A}(0.02, 3))$ (h) $f(\mathcal{A}(0.02, 4))$ (i) $f(\mathcal{A}(0.02, 5))$

FIGURE 2. The images $f(\mathcal{A}(\delta, k))$ (in red) and $f(\Omega_2)$ (in gray) in Example 4.2

By our method, we compute the approximation $\mathcal{A}(0.01, 4)$ and show it in Figure 3 (a), which is the intersection of Ω_3 and the area under the red curve defined by $\psi_4(\boldsymbol{x}) = 0.01$. The image $f(\mathcal{A}(0.01, 4))$ is illustrated in Figure 3 (c), which shows that we can obtain good approximations of $\mathcal{S}_w$ including the ones corresponding to the sunken part of Pareto curve.

Next, we consider the following optimization problem with a Pareto constraint

$$\min \ x_1^2 + (x_2 - 1)^2 \quad \text{s.t. } (x_1, x_2) \in \mathcal{S}_w,$$



(a) (b)

FIGURE 3. (a) The bicorn curve and the curve defined by $\psi_4(\boldsymbol{x}) = 0.01$; (b) The images $f(\mathcal{A}(0.01, 4))$ (in red) and $f(\Omega_3)$ (in gray) in Example 4.3

which is to compute the square of the Euclidean distance between the point $(0, 1)$ and the curve $\mathcal{S}_w$. It is easy to see that the unique minimizer of the above problem is $\left(0, \frac{1}{3}\right)$ and the minimum is $\frac{4}{9} \approx 0.444$. With the approximation $\mathcal{A}(0.01, 4)$ of $\mathcal{S}_w$, we consider the polynomial optimization problem

$$\min\ x_1^2 + (x_2 - 1)^2 \quad \text{s.t. } \boldsymbol{x} \in \Omega_3,\ \psi_4(\boldsymbol{x}) \leq 0.01.$$

We solve this problem by Lasserre's hierarchy of SDP relaxations (c.f. [18, 21]) with the software GloptiPoly [15], and get the certified minimizer $(-0.0000, 0.3473)$ and minimum 0.4260. $\square$

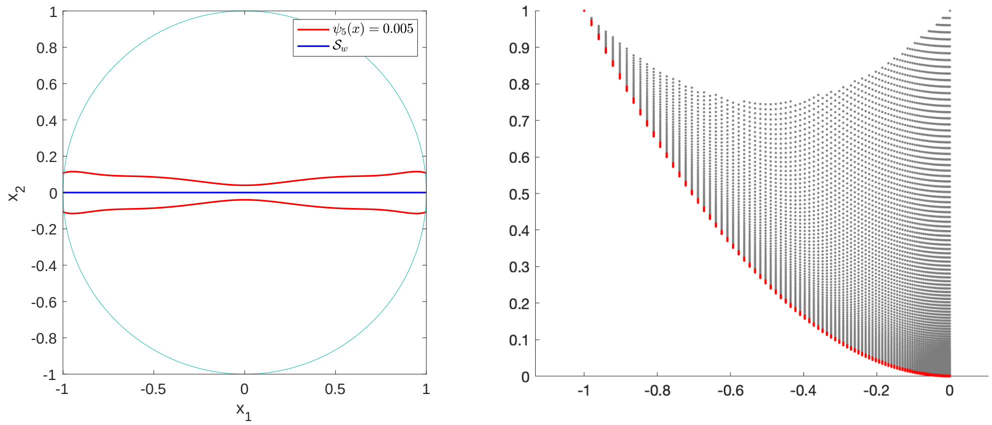
Example 4.4. Consider the problem

$$\begin{cases} \mathrm{Min}_{\mathbb{R}^2_+}\ \left(-x_1^2,\ x_1^4 + x_2^2\right) \\ \text{ s.t. } \quad \boldsymbol{x} \in \Omega_4 := \left\{\boldsymbol{x} \in \mathbb{R}^2 \colon 1 - x_1^2 - x_2^2 \geq 0\right\}. \end{cases}$$

It is easy to see that the set of weakly efficient solutions $\mathcal{S}_w = [-1, 1] \times \{0\}$ and the image $f(\mathcal{S}_w)$ (the Pareto curve) is the curve

$$\left\{(t_1, t_2) \in \mathbb{R}^2 : t_2 = t_1^2,\ t_1 \in [-1, 0]\right\}$$

in the objective plane where $t_1 = -x_1^2$ and $t_2 = x_1^4 + x_2^2$. Clearly, for every point $(t_1, t_2) \in f(\mathcal{S}_w)$, there are two weakly efficient solutions $(-\sqrt{-t_1}, 0)$ and $(\sqrt{-t_1}, 0)$. Therefore, this problem does not satisfy the assumptions of the approach proposed in [29] when using the scalarizations (8) and (9). By our method, we compute the set $\mathcal{A}(0.005, 5)$, which is the intersection of the unit disk and the area enclosed by the red curve defined by $\psi_5(\boldsymbol{x}) = 0.005$ in Figure 4 (a). The images $f(\Omega_4)$ and $f(\mathcal{A}(0.005, 5))$ is shown in Figure 4 (b), which illustrates that we can approximate the set of weakly efficient solutions as closely as possible. $\square$



(a) (b)

FIGURE 4. (a) The set $\mathcal{S}_w$ and the curve defined by $\psi_5(\boldsymbol{x}) = 0.005$; (b) The images $f(\mathcal{A}(0.005, 5))$ (in red) and $f(\Omega_4)$ (in gray) in Example 4.4

.

5. Conclusions

In this paper, we provide a new scheme for approximating the set of all weakly ($\boldsymbol{\epsilon}$-)efficient solutions to the problem (VROP). The procedure mainly relies on the achievement function associated with (VROP) and the "joint+marginal" approach proposed by Lasserre [22]. The obtained results seem new in the area of vector optimization with polynomial structures, in the sense that we approximate the whole set of weakly ($\boldsymbol{\epsilon}$-)efficient solutions to the problem (VROP). Moreover, the obtained results also significantly develop the recent achievements in [6, 24, 25] for vector polynomial optimization problems from convex settings to nonconvex settings.

Acknowledgments. The authors thank Professor Tien-Son Pham for his helpful comments on the early version of this manuscript. Feng Guo was supported by the Chinese National Natural Science Foundation under grant 11571350, and the Fundamental Research Funds for the Central Universities. Liguo Jiao was supported by Natural Science Foundation of Jilin Province (no. 20220101302JC) and the Fundamental Research Funds for the Central Universities.

References

[1] V. Blanco, J. Puerto, and S. E. H. B. Ali. A semidefinite programming approach for solving multiobjective linear programming. *Journal of Global Optimization*, 58(3):465–480, 2014.

[2] J. M. Borwein. On the existence of Pareto efficient points. *Mathematics of Operations Research*, 8(1):64–73, 1983.

[3] R. S. Burachik, C. Y. Kaya, and M. M. Rizvi. A new scalarization technique and new algorithms to generate Pareto fronts. *SIAM Journal on Optimization*, 27(2):1010–1034, 2017.

[4] V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making: Theory and Methodology.* Amsterdam, North-Holland, 1983.

[5] W. Chen, J. Unkelbach, A. Trofimov, T. Madden, H. Kooy, T. Bortfeld, and D. Craft. Including robustness in multi-criteria optimization for intensity modulated proton therapy. *Physics in Medicine and Biology*, 57(3):591–608, 2012.

[6] T. D. Chuong. Second-order cone programming relaxations for a class of multiobjective convex polynomial problems. *Annals of Operations Research*, 311(2):1017–1033, 2022.

[7] J. Dutta, P. Kesarwani, and S. Gupta. Gap functions and error bounds for nonsmooth convex vector optimization problem. *Optimization*, 66(11):1807–1836, 2017.

[8] M. Ehrgott. *Multicriteria Optimization (2nd ed.).* Springer, Berlin, 2005.

[9] H. Eschenauer and J. Koski. *Multicriteria Design Optimization.* Springer, Berlin, 1990.

[10] J. Fliege, L. M. G. Drummond, and B. F. Svaiter. Newton's method for multiobjective optimization. *SIAM Journal on Optimization*, 20(2):602–626, 2009.

[11] B. L. Gorissen and D. d. Hertog. Approximating the Pareto set of multiobjective linear programs via robust optimization. *Operations Research Letters*, 40(5):319–324, 2012.

[12] D. Grimm, T. Netzer, and M. Schweighofer. A note on the representation of positive polynomials with structured sparsity. *Archiv der Mathematik*, 89(5):399–403, 2007.

[13] W. J. Gutjahr and P. C. Nolz. Multicriteria optimization in humanitarian aid. *European Journal of Operational Research*, 252(2):351–366, 2016.

[14] S. T. Hackman and U. Passy. Maximizing a linear fractional function on a Pareto efficient frontier. *Journal of Optimization Theory and Applications*, 113(1):83–103, 2002.

[15] D. Henrion, J. B. Lasserre, and J. Löfberg. GloptiPoly 3: moments, optimization and semidefinite programming. *Optimization Methods and Software*, 24(4-5):761–779, 2009.

[16] N. T. T. Huong, J.-C. Yao, and N. D. Yen. Geoffrion's proper efficiency in linear fractional vector optimization with unbounded constraint sets. *Journal of Global Optimization*, 78(3):545–562, 2020.

[17] D. S. Kim, T.-S. Phạm, and N. V. Tuyen. On the existence of Pareto solutions for polynomial vector optimization problems. *Mathematical Programming Ser. A*, 177(1-2):321–341, 2019.

[18] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[19] J. B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity. *SIAM Journal on Optimization*, 17(3):822–843, 2006.

[20] J. B. Lasserre. A "joint+marginal" approach to parametric polynomial optimization. *SIAM Journal on Optimization*, 20(4):1995–2022, 2010.

[21] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications.* Imperial College Press, London, 2010.

[22] J. B. Lasserre. Tractable approximations of sets defined with quantifiers. *Mathematical Programming Ser. B*, 151(1):507–527, 2015.

[23] M. Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging Applications of Algebraic Geometry*, volume 149 of *The IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, New York, NY, 2009.

[24] J. H. Lee and L. G. Jiao. Solving fractional multicriteria optimization problems with sum of squares convex polynomial data. *Journal of Optimization Theory and Applications*, 176(2):428–455, 2018.

[25] J. H. Lee, N. Sisarat, and L. G. Jiao. Multi-objective convex polynomial optimization and semidefinite programming relaxations. *Journal of Global Optimization*, 80(1):117–138, 2021.

[26] C. G. Liu, K. F. Ng, and W. H. Yang. Merit functions in vector optimization. *Mathematical Programming*, 119:215–237, 2009.

[27] J. Löfberg. YALMIP : a toolbox for modeling and optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508)*, pages 284–289, 2004.

[28] D. T. Luc. *Multiobjective Linear Programming: An Introduction.* Springer International Publishing, Switzerland, 2016.

[29] V. Magron, D. Henrion, and J. B. Lasserre. Approximating Pareto curves using semidefinite relaxations. *Operations Research Letters*, 42(6-7):432–437, 2014.

[30] J. Nie. Tight relaxations for polynomial optimization and Lagrange multiplier expressions. *Mathematical Programming*, 178(1):1–37, 2019.

[31] J. Nie and Z. Yang. The multi-objective polynomial optimization. 2021. arXiv:2108.04336.

[32] T.-S. Phạm, X. D. T. Hà, and J.-C. Yao. The global weak sharp minima with explicit exponents in polynomial vector optimization problems. *Positivity*, 22(1):219–244, 2018.

[33] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.

[34] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis.* Springer Berlin, Heidelberg, 1998.

[35] Y. Sawaragi, H. Nakayama, and T. Tanino. *Theory of Multiobjective Optimization.* Academic Press, Inc., Orlando, FL, 1985.

[36] J. F. Sturm. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11(1-4):625–653, 1999.

[37] H. Tanabe, E. H. Fukuda, and N. Yamashita. Proximal gradient methods for multiobjective optimization and their applications. *Computational Optimization and Applications*, 72(2):339–361, 2019.

[38] H. Tanabe, E. H. Fukuda, and N. Yamashita. New merit functions and error bounds for non-convex multiobjective optimization. 2020. arXiv: 2010.09333.

[39] H. Tanabe, E. H. Fukuda, and N. Yamashita. Convergence rates analysis of multiobjective proximal gradient methods. *Optimization Letters*, 2022. https://doi.org/10.1007/s11590-022-01877-7.

[40] H. Waki, S. Kim, M. Kojima, and M. Muramatsu. Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. *SIAM Journal on Optimization*, 17(1):218–242, 2006.

[41] A. P. Wierzbicki. On the completeness and constructiveness of parametric characterizations to vector optimization problems. *OR Spektrum*, 8(2):73–87, 1986.

[42] X. P. Zhao, M. A. Köis, Y. H. Yao, and J.-C. Yao. A projected subgradient method for non-differentiable quasiconvex multiobjective optimization problems. *Journal of Optimization Theory and Applications*, 190(1):82–107, 2021.

[43] X. P. Zhao and J.-C. Yao. Linear convergence of a nonmonotone projected gradient method for multiobjective optimization. *Journal of Global Optimization*, 82(3):577–594, 2022.

(Feng Guo) School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning Province, China.

Email address: `fguo@dlut.edu.cn`

(Liguo Jiao) Academy for Advanced Interdisciplinary Studies, Northeast Normal University, Changchun 130024, Jilin Province, China.

Email address: `hanchezi@163.com; jiaolg356@nenu.edu.cn`