A new scheme for approximating the weakly efficient solution set of vector rational optimization problems

Feng Guo · Liguo Jiao

Received: 25 May 2022 / Accepted: 17 April 2023 / Published online: 2 May 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

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.

Keywords  Vector optimization · Polynomial optimization · Achievement function, Lasserre’s hierarchy · Weakly efficient solution set

Mathematics Subject Classification 90C29 · 90C32 · 90C23 · 90C22

1 Introduction
Vector optimization forms an important field of research in optimization theory; see, e.g., [9, 14, 15, 39, 49], and many practical applications in various areas, such as engineering [15], humanitarian aid [20], medical health [10] and so on. In this paper, we will be concerned with the approximation of the weakly efficient solution set of vector rational optimization problems 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.

This paper is dedicated to Professor Do Sang Kim on the occasion of his 70th birthday with respect.
with the following constrained vector rational optimization problem of the form

\[
\operatorname{Min}_{\mathbb{R}^m_+} \left\{ f(x) := \left( \frac{p_1(x)}{q_1(x)}, \ldots, \frac{p_m(x)}{q_m(x)} \right) : x \in \Omega \right\},
\]

(VROP)

where “\( \operatorname{Min}_{\mathbb{R}^m_+} \)” 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 \( x = (x_1, \ldots, x_n) \) for each \( i = 1, \ldots, m \), and the feasible set \( \Omega \) is given by

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

where for each \( j = 1, \ldots, r, \) \( g_j \) is a real polynomial in the variable \( x \). By letting \( q_i = 1 \) for all \( i = 1, \ldots, m \), our model then covers vector polynomial optimization problems [5, 25, 36, 39, 45], 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 [24] 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 vector optimization. Now, let us recall the concepts of optimal solutions to vector optimization problems. A point \( x \in \Omega \) is said to be an efficient solution (or Edgeworth-Pareto (EP) optimal point) to the problem (VROP) if it holds that

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

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

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

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

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

Denote by \( S \) (resp., \( S_w, S^\epsilon_w \)) the set of all efficient (resp., weakly efficient, weakly \( \epsilon \)-efficient) solutions to the problem (VROP), respectively. Clearly, \( S \subset S_w \subset S^\epsilon_w \), but not conversely. We call the image \( f(S_w) \) the Pareto frontier (the Pareto curve if \( m = 2 \)) of (VROP); see [40].

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

(A1) The feasible set \( \Omega \) is nonempty and compact;

(A2) The denominators \( q_i(x) > 0 \) over \( \Omega \) 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 \( \Omega \) is also compact, which ensures the existence of (weakly) efficient solutions to the problem (VROP); see, e.g., [6, Theorem 1], [14, Theorem 2.1] and [49, 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 \( \frac{p_i}{q_i^\epsilon} \) for \( \frac{p_i}{q_i} \), (A2) can be weakened as \( q_i(x) \neq 0 \) over \( \Omega \) for all \( i = 1, \ldots, m \).

The problem (VROP) is an important subclass of the multi-objective fractional programming (MOFP) problems. The (MOFP) problem has various applications in real life problems such as management of finance, production, inventory, banking etc. It can be used to model problem of real life with many objectives like as actual cost/standard cost, output/total employees, profit/cost, inventory/sales, nurses/patients, profit/investment ratios, etc; see the recent survey [3] and the references therein. Many (MOFP) problems can be formulated as (VROP) problems, among which are the widely studied multi-objective linear fractional
programming (MLFP) problems (c.f. [7, 28, 41, 42]). To name a few, the first example is the transportation problem which is a multi-objective decision-making problem [56]. It aims to determine the ideal transportation setup that matches the decision maker’s preferences while taking into account competing objectives/criteria, like total actual cost/standard cost, actual time/standard time, actual deterioration/standard deterioration, and risk assets/capital, etc. The generalized mathematical model of such problems can be formulated as (MLFP). Another example comes from the water resources sustainability problem [1]. A comprehensive cropping pattern planning takes into account the high level of interrelation of the environmental, economic and social aspects of farming systems. To assess the sustainability of water resources and determine an optimal pattern of cropping, we can simultaneously optimize the ratios of net return/water consumption and labor employment/water consumption as the sustainability indicators, which is modelled as an (MLFP) problem.

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 [8, 11, 16, 35, 36, 45, 53–55] 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., [8, 11, 35, 36, 45]), 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., [16] for Newton’s methods, [53–55, 59, 60] 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 optimization in the real world. In fact, there may exist more than one (weekly) efficient solution which are equally acceptable. Which of the (weekly) efficient solutions is preferable to the decision makers depends on the situations, like their financial position, time limit etc. So it is better to find or approximate the whole (weekly) efficient solution set of the (VROP) problem, so that decision makers can choose a better option from alternatives according to their level of satisfaction of objectives. Therefore, the aim and novelty of this paper is to provide a new scheme for approximating the whole set of weakly (ε-)efficient solutions of (VROP).

More precisely, we provide a procedure to obtain a sequence of explicit approximations of $S^\epsilon_w$ (and hence $S_w$ by letting $\epsilon \to 0$). Each approximation is the intersection of the sublevel set of a single polynomial and the feasible set $\Omega$. 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. [14, 46, 58]) associated with the problem (VROP) which is defined as

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

It can be shown that the sets $S_w$ and $S^\epsilon_w$ can be written as the intersection of sublevel sets of $\psi(x)$ and the feasible set $\Omega$ (see Sect. 3). As the function $\psi(x)$ can be fairly complicated, the problem is reduced to construct polynomial approximations of $\psi(x)$. By rewriting the definition of $\psi(x)$ as a parametric polynomial optimization problem, we can construct a sequence of polynomial approximations $\{\psi_k(x)\}_{k \in \mathbb{N}}$ of $\psi(x)$ over the feasible set $\Omega$ from above by invoking the “joint+marginal” approach developed by Lasserre in [31, 33]. Remarkably, the construction of $\{\psi_k(x)\}_{k \in \mathbb{N}}$ can be converted to semidefinite programming (SDP) problems. For $\epsilon \in \mathbb{R}_+^m$ of the form $\epsilon = (\delta, \ldots, \delta)$ with $\delta > 0$, the intersection, denoted by $A(\delta, k)$,
of the sublevel set $\psi_k(x) \leq \delta$ and the feasible set $\Omega$ are inner approximations of $S^e_w$. Under some conditions, we prove that $\text{vol}\left(S^e_w \setminus A(\delta, k)\right) \to 0$ as $k \to \infty$, where “$\text{vol}(\cdot)$” denotes the Lebesgue volume (see Theorem 4.2). Since it holds for $\epsilon = (\delta, \ldots, \delta)$ that $S^e_w \to S_w$ as $\delta \to 0$ (see Proposition 3.2), we may take $A(\delta, k)$ as an approximation of $S_w$ with sufficiently small $\delta > 0$ and 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 Sect. 3, we study the characterization of the weakly efficient solution set of the problem (VROP) by the associated achievement function $\psi(x)$. In Sect. 4, we show how to approximate the weakly ($\epsilon$-)efficient solution set of the problem (VROP), and present some nontrivial illustrating examples. Conclusions are given in Sect. 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 $D \subset \mathbb{R}^n$, we use $\text{cl}(D)$ and $\text{int}(D)$ to denote the closure and interior of $D$, respectively. Denote by $B$ the closed unit ball in $\mathbb{R}^n$ centered at the origin. For a point $u \in \mathbb{R}^n$, $\text{dist}(u, D)$ denotes the Euclidean distance between $u$ and $D$. For $u \in \mathbb{R}^n$, $\|u\|$ denotes the standard Euclidean norm of $u$. For $\alpha := (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $|\alpha| = \alpha_1 + \cdots + \alpha_n$. For $k \in \mathbb{N}$, denote by $\mathbb{N}^n_k = \{\alpha \in \mathbb{N}^n : |\alpha| \leq k\}$ and $|\mathbb{N}^n_k|$ its cardinality. Denote by $\mathbb{R}[x]$ the ring of polynomials in $x := (x_1, \ldots, x_n)$ with real coefficients and by $\mathbb{R}[x]_k$ the set of polynomials in $\mathbb{R}[x]$ 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 $x^n$ stands for the monomial $x_1^{a_1} \cdots x_n^{a_n}$.

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 $h \in \mathbb{R}[x]$ 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 $x$ is denoted by $\Sigma^2[x]$. Let $\{h_1, \ldots, h_s\} \subset \mathbb{R}[x]$ be a finite set of polynomials and

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

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

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

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

**Theorem 2.1** (Putinar’s Positivstellensatz [47]) Suppose that Assumption 2.1 holds. If $h(x) \in \mathbb{R}[x]$ is positive on $S$, then $h(x)$ can be written in the form

$$h(x) = \sigma_0(x) + \sum_{j=1}^s \sigma_jh_j(x),$$

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

Note that if we fix the degrees of $\sigma_j$’s in (1), then checking the above representation of $h(x)$ reduces to an SDP feasibility problem (c.f. [34]). 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. [29, 32]).

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 \( x_I := \{x_i: i \in I\} \) and \( \mathbb{R}[x_I] \) as the polynomial ring in the variables \( x_I \).

**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) \( h_j \in \mathbb{R}[x_{I_i}] \) 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

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

for some sum of squares polynomials \( \sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[x_{I_i}], j \in J_i. \)

The following result enables us to construct sparse SDP relaxations of polynomial optimization problems, which can significantly reduce the computational cost (c.f. [30, 57]).

**Theorem 2.2 (Sparse version of Putinar’s Positivstellensatz [19, 30, 57])** Suppose that Assumption 2.2 holds. If \( h(x) \in \sum_{i=1}^l \mathbb{R}[x_{I_i}] \) and is positive on \( S \), then \( h(x) \) can be written as

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

for some sum of squares polynomials \( \sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[x_{I_i}], j \in J_i, i = 1, \ldots, l. \)

### 3 Characterizing the weakly efficient solution set

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

By definition of \( S_w \), we have

\[
S_w = \left\{ x \in \Omega: \forall y \in \Omega, f(y) - f(x) \notin -\mathbb{R}_+^m \right\} = \left\{ x \in \Omega: \forall y \in \Omega, \exists i \in \{1, \ldots, m\} \text{ such that } f_i(x) - f_i(y) \leq 0 \right\} = \left\{ x \in \Omega: \sup_{y \in \Omega} \min_{i \in \{1, \ldots, m\}} [f_i(x) - f_i(y)] \leq 0 \right\}.
\]

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

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

The function \( \psi(x) \) is known as the achievement function in the area of vector optimization in the literature; see [14, Section 4.6] and [46, 58]. Therefore,

\[
S_w = \{ x \in \mathbb{R}^n: \psi(x) \leq 0 \} \cap \Omega.
\]
Moreover, we have the following results, which imply that the function \( \psi(x) \) is indeed a merit function (see [13, 37, 54, 55]).

**Proposition 3.1** ([46, Lemmas 3.1 and 3.2]) The achievement function \( \psi(x) \) satisfies

(i) \( \psi(x) \geq 0 \) for all \( x \in \Omega \) and hence \( S_w = \{ x \in \Omega : \psi(x) = 0 \} \).

(ii) \( \psi(x) \) is locally Lipschitz on \( \bigcap_{i=1}^{m} \text{dom}(f_i) \), where \( \text{dom}(f_i) = \{ x \in \mathbb{R}^n : q_i(x) \neq 0 \} \).

**Proof** (i) is clear. If the objective in (VROP) is a vector of polynomials, (ii) was proved in [46, 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 \( \bigcap_{i=1}^{m} \text{dom}(f_i) \). Hence, the proof of [46, Lemma 3.2] is still valid for the case studied in this paper.

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

Recall the definition of the set \( S^\epsilon_w \) of all weakly \( \epsilon \)-efficient solutions to (VROP), and clearly by definition, \( S_w \subset S^\epsilon_w \) for any \( \epsilon \in \mathbb{R}^m_+ \). Conversely, denote a set-valued mapping \( F(\cdot) : \mathbb{R}^m \Rightarrow \mathbb{R}^n \) and let \( F(\epsilon) := S^\epsilon_w \) for \( \epsilon \in \mathbb{R}^m_+ \). The following proposition shows that \( F(\cdot) \) is continuous at \( \bar{\epsilon} = 0 \) relative to \( \mathbb{R}^m \) in the sense of Painlevé–Kuratowski (see [48, Definition 5.4]), i.e., \( F(\epsilon) \rightarrow F(0) \) as \( \epsilon \rightarrow 0 \). For convenience, we recall the definitions of continuity (outer semicontinuity, inner semicontinuity) for set-valued mappings; see [48, Chapters 4 & 5] for more information. Given a set-valued mapping \( F : \mathbb{R}^m \Rightarrow \mathbb{R}^n \), we denote by

\[
\begin{align*}
\limsup_{y \rightarrow \bar{y}} F(y) & := \left\{ x \in \mathbb{R}^n : \exists y_k \rightarrow \bar{y}, \exists x_k \rightarrow x \text{ with } x_k \in F(y_k) \right\}, \\
\liminf_{y \rightarrow \bar{y}} F(y) & := \left\{ x \in \mathbb{R}^n : \forall y_k \rightarrow \bar{y}, \exists x_k \rightarrow x \text{ with } x_k \in F(y_k) \right\},
\end{align*}
\]

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

**Definition 3.1** A set-valued mapping \( F : \mathbb{R}^m \Rightarrow \mathbb{R}^n \) is said to be outer semicontinuous (osc) at \( \bar{y} \) if \( \limsup_{y \rightarrow \bar{y}} F(y) \subset F(\bar{y}) \), and inner semicontinuous (isc) at \( \bar{y} \) if \( F(\bar{y}) \subset \liminf_{y \rightarrow \bar{y}} F(y) \).

It is called continuous at \( \bar{y} \) if \( F \) is simultaneously osc and isc at \( \bar{y} \), i.e., \( F(y) \rightarrow F(\bar{y}) \) as \( y \rightarrow \bar{y} \). These terms are invoked relative to \( X \), a subset of \( \mathbb{R}^m \) containing \( \bar{y} \), if the inclinations hold in restriction to convergence \( y \rightarrow \bar{y} \) with \( y \in X \).

It follows from Definition 3.1 that \( F(\cdot) \) is continuous at \( \bar{\epsilon} = 0 \) relative to \( \mathbb{R}^m_+ \). Similar to [48, Proposition 5.12 and Exercise 5.13], we have the following result. For any \( \epsilon \in \mathbb{R}^m_+ \), denote \( \epsilon_{\max} := \max_{i=1, \ldots, m} \{ \epsilon_i \} \) and \( \epsilon_{\min} := \min_{i=1, \ldots, m} \{ \epsilon_i \} \).

**Proposition 3.2** For any \( d > 0 \), there exists a number \( \delta(d) > 0 \) depending on \( d \) such that \( \text{dist}(u, S_w) < d \) for any \( u \in S^\epsilon_w \), i.e., \( S^\epsilon_w \subset S_w + dB \), whenever \( \epsilon_{\max} < \delta(d) \).

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

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

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

$$
f_i(y') - f_i(u^{(k)}) + \epsilon^{(k)}_i = f_i(y') - f_i(u') + f_i(u') - f_i(u^{(k)}) + \epsilon^{(k)}_i
$$

$$
\leq \max_{i=1,\ldots,m} [f_i(y') - f_i(u')] + f_i(u') - f_i(u^{(k)}) + \frac{1}{k'} \left(\text{by } \epsilon^{(k)}_{\max} < \frac{1}{k'}\right)
$$

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

$$
\text{dist}(u', S_w) = \lim_{k \to \infty} \text{dist}(u^{(k)}, S_w) \geq d > 0,
$$

a contradiction.

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

**Proposition 3.3** For any $\epsilon \in \mathbb{R}_+^m$, we have

$$
\{x \in \Omega : \psi(x) \leq \epsilon_{\min}\} \subset S_w^\epsilon \subset \{x \in \Omega : \psi(x) \leq \epsilon_{\max}\}.
$$

**Particularly, if $\epsilon_{\max} = \epsilon_{\min}$, then**

$$
\{x \in \Omega : \psi(x) \leq \epsilon_{\min} = \epsilon_{\max}\} = S_w^\epsilon.
$$

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

Now, fix a point $u \in S_w^\epsilon$. For any $y \in \Omega$, by definition, there exists $k_y \in \{1, \ldots, m\}$ depending on $y$ such that $k_y(y) - f_k(u) + \epsilon_k \geq 0$. Then, $\min_{i=1,\ldots,m}[f_i(u) - f_i(y)] \leq \epsilon_{k_y}$ for all $y \in \Omega$, and hence

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

thus, the second relation in (2) holds. Consequently, the conclusion follows. $\Box$

### 4 Approximations of weakly (ε)-efficient solution set

In this section, we will construct polynomial approximations of the achievement function $\psi(x)$ from above and use their sublevel sets to approximate the set of all weakly (ε)-efficient solutions to (VROP). The construction of these polynomial approximations of $\psi(x)$ is inspired by [33] and can be reduced to SDP problems. As $\Omega$ is compact, after a possible re-scaling of the $g_j$’s, we may and will assume that $\Delta := [-1, 1]^m \supset \Omega$ in the rest of this paper.
4.1 Approximations of achievement function

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

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

which is equivalent to an SDP feasibility problem (c.f. [34]). Under (A1-2), each $\frac{p_i(x)}{q_i(x)}$ is bounded from below on $\Omega$ and $p_i(x) - f_i^{\text{lower}} q_i(x) > 0$ on $\Omega$ for any $f_i^{\text{lower}} < \min_{x \in \Omega} \frac{p_i(x)}{q_i(x)}$. Hence, by Putinar’s Positstellensatz, a number $f_i^{\text{lower}}$ satisfying (3) always exists for $k_i$ large enough (note that Assumption 2.1 holds due to the redundant polynomials $1 - x_j^2, j = 1, \ldots, n$, added in (3)). Clearly, it holds that

\[
f^{\text{lower}} := \min_{i=1,\ldots,m} f_i^{\text{lower}} \leq \min_{i=1,\ldots,m, \ x \in \Omega} \frac{p_i(x)}{q_i(x)}.
\]

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

\[
f^{\text{upper}} := \max_{i=1,\ldots,m} f_i^{\text{upper}} \geq \max_{i=1,\ldots,m, \ x \in \Omega} \frac{p_i(x)}{q_i(x)}.
\]

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

\[
\psi(x) := \sup_{y \in \Omega} \min_{i=1,\ldots,m} \left[ f_i(x) - f_i(y) \right]
= \sup_{y \in \Omega} \min_{i=1,\ldots,m} \left[ \frac{p_i(x)}{q_i(x)} - \frac{p_i(y)}{q_i(y)} \right]
= \sup_{y \in \Omega, \ z \in \mathbb{R}} \left\{ z : \frac{p_i(x)}{q_i(x)} - \frac{p_i(y)}{q_i(y)} \geq z, \ i = 1, \ldots, m \right\}.
\]

For any $x \in \Delta$, let

\[
\tilde{\psi}(x) := \max_{y \in \mathbb{R}^n, z \in \mathbb{R}} \left[ z : \begin{array}{l}
  p_i(x)q_i(y) - p_i(y)q_i(x) - zq_i(x)q_i(y) \geq 0, \ i = 1, \ldots, m, \\
  y \in \Omega, \ z \in [f^{\text{lower}} - f^{\text{upper}}, f^{\text{upper}} - f^{\text{lower}}].
\end{array} \right]
\]

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

\[\text{Proposition 4.1} \quad \tilde{\psi}(x) = \psi(x) \text{ for all } x \in \Omega. \text{ Hence, Propositions 3.1 and 3.3 also hold for } \tilde{\psi}.\]
Next, we construct polynomial approximations of $\tilde{\psi}$ over $\Delta$ from above by means of the SDP method proposed in [33], and use their sublevel sets to approximate the set of all weakly $(\varepsilon)$-efficient solutions to (VROP).

Consider the following sets

$$K := \left\{ (x, y, z) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} : \begin{array}{l} p_i(x)q_j(y) - p_i(y)q_j(x) - zq_j(x)q_j(y) \geq 0, \\
i = 1, \ldots, m, \ x \in \Delta, \ y \in \Omega, \\
z \in \left[ f_{\text{lower}} \ - f_{\text{upper}}, f_{\text{upper}} - f_{\text{lower}} \right] \end{array} \right\},$$

and

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

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

As proved in [33, Theorem 1], a sequence of polynomial approximations of $\tilde{\psi}(x)$ on $\Delta$ from above exists mainly due to the Stone–Weierstrass theorem.

**Proposition 4.2** (c.f. [33, Theorem 1]) There exists a sequence of polynomials $\{\psi_k \in \mathbb{R}[x] : k \in \mathbb{N}\}$ such that $\psi_k(x) \geq \tilde{\psi}(x)$ for all $x \in \Delta$, and $\{\psi_k\}_{k \in \mathbb{N}}$ converges to $\tilde{\psi}$ in $L_1(\Delta)$, i.e.,

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

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

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

For any $\delta > 0$, with a slight abuse of notation, we denote $S^\delta_w := S^\epsilon_w$, where $\epsilon = (\delta, \ldots, \delta)$. The following result can be derived by slightly modifying the proof of [33, Theorem 3]. It shows that we can approximate the set $S^\delta_w$ by the sequence $\{A(\delta, k)\}_{k \in \mathbb{N}}$.

**Theorem 4.1** For any $\delta > 0$, we have $A(\delta, k) \subset S^\delta_w$ and

$$\text{vol} \left( \{x \in \Omega : \psi(x) < \delta\} \right) \leq \lim_{k \to \infty} \text{vol} \left( A(\delta, k) \right) \leq \text{vol} \left( \{x \in \Omega : \psi(x) \leq \delta\} \right) = \text{vol} \left( S^\delta_w \right) \quad (5)$$

Consequently, if $\text{vol} \left( \{x \in \Omega : \psi(x) = \delta\} \right) = 0$, then $\lim_{k \to \infty} \text{vol} \left( S^\delta_w \setminus A(\delta, k) \right) = 0.$

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

$$\lim_{k \to \infty} \text{vol} \left( \{x \in \Delta : |\psi_k(x) - \tilde{\psi}(x)| \geq \alpha\} \right) = 0. \quad (6)$$

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

$$\text{vol} \left( \{x \in \Omega : \psi(x) \leq \delta + \frac{1}{\ell}\} \right)$$

$$= \text{vol} \left( \{x \in \Omega : \tilde{\psi}(x) \leq \delta + \frac{1}{\ell}\} \right) \quad \text{(by Proposition 4.1)}$$

$$= \text{vol} \left( \{x \in \Omega : \tilde{\psi}(x) \leq \delta + \frac{1}{\ell}\} \cap \{x \in \Omega : \psi_k(x) > \delta\} \right)$$

$$+ \text{vol} \left( \{x \in \Omega : \tilde{\psi}(x) \leq \delta + \frac{1}{\ell}\} \cap \{x \in \Omega : \psi_k(x) \leq \delta\} \right)$$

$$= \lim_{k \to \infty} \text{vol} \left( \{x \in \Omega : \tilde{\psi}(x) \leq \delta + \frac{1}{\ell}\} \cap \{x \in \Omega : \psi_k(x) \leq \delta\} \right) \quad \text{(by (6))}$$
\[ \lim_{k \to \infty} \text{vol} (\{x \in \Omega : \psi_k(x) \leq \delta\}) \leq \text{vol} \left( \left\{ x \in \Omega : \tilde{\psi}(x) \leq \delta \right\} \right) = \text{vol} (S_w^\delta). \text{ (by Propositions 3.3 and 4.1)} \]

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

\[ A(\delta, k) \subset S_w + d\mathbf{B} \]

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

(ii) For \( d > 0 \) and any \( \delta > 0 \) with \( \text{vol} (\{x \in \Omega : \psi(x) = \delta\}) = 0 \), there exists \( k(d, \delta) \in \mathbb{N} \) depending on \( \delta \) and \( d \) such that

\[ S_w \cap \text{cl} (\text{int} (\Omega \setminus S_w)) \subset A(\delta, k) + d\mathbf{B} \]

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

**Proof** (i) Since \( A(\delta, k) \subset S_w^\delta \) 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 \( u \in S_w \cap \text{cl} (\text{int} (\Omega \setminus S_w)) \neq \emptyset \), then \( \tilde{\psi}(u) = 0 \) by Propositions 3.1 (i) and 4.1, and there exists a sequence \( \{u^{(l)}\}_{l \in \mathbb{N}} \subset \text{int} (\Omega \setminus S_w) \) such that \( \lim_{l \to \infty} u^{(l)} = u \). Fix the numbers \( d, \delta > 0 \). By the continuity of \( \tilde{\psi} \) on \( \Omega \) (Proposition 4.1), there exists \( l_0 \in \mathbb{N} \) depending on \( d \) and \( \delta \) such that \( \tilde{\psi}(u^{(l_0)}) < \delta \) and \( \|u^{(l_0)} - u\| < d \). As \( u^{(l_0)} \in \text{int} (\Omega \setminus S_w) \), by the continuity of \( \tilde{\psi} \) again, there is a neighborhood \( O^{(l_0)} \subset \Omega \) of \( u^{(l_0)} \) such that \( \tilde{\psi}(x) < \delta \) and \( \|x - u\| < d \) for all \( x \in O^{(l_0)} \). Proposition 3.3 implies that \( O^{(l_0)} \subset S_w^\delta \). Then, we show that there exists \( k(d, \delta) \in \mathbb{N} \) such that for any \( k > k(d, \delta) \), it holds that \( A(\delta, k) \cap O^{(l_0)} \neq \emptyset \) which means that \( u \in A(\delta, k) + d\mathbf{B} \) and the conclusion follows. To the contrary, suppose that such \( k(d, \delta) \) does not exist, then there is a subsequence \( \{A(\delta, k_j)\}_{j \in \mathbb{N}} \) with \( k_j \to \infty \) such that \( A(\delta, k_j) \cap O^{(l_0)} = \emptyset \) for all \( k_j \). Then, \( \text{vol}(S_w^\delta \setminus A(\delta, k_j)) \geq \text{vol}(O^{(l_0)}) > 0 \) for all \( k_j \). As \( \text{vol}(\{x \in \Omega : \psi(x) = \delta\}) = 0 \), 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 \( S_w \cap \text{cl} (\text{int} (\Omega \setminus S_w)) \neq \emptyset \), then for any \( \delta > 0 \), \( A(\delta, k) \neq \emptyset \) for \( k \) large enough. In fact, we have \( O^{(l_0)} \subset \{x \in \Omega : \psi(x) < \delta\} \) for the neighborhood \( O^{(l_0)} \) in the proof of Corollary 4.1. Then, (5) implies that \( A(\delta, k) \neq \emptyset \) for \( k \) large enough.

(ii) Suppose there is a sequence \( \{\delta_i\}_{i \in \mathbb{N}} \) with \( \delta_i \downarrow 0 \) such that \( \text{vol}(\{x \in \Omega : \psi(x) = \delta_i\}) = 0 \) holds for all \( i \) and \( S_w = S_w \cap \text{cl} (\text{int} (\Omega \setminus S_w)) \), 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 \( A(\delta, k) \) with sufficiently small \( \delta > 0 \) and sufficiently large \( k \in \mathbb{N} \).

For any \( 0 < \varepsilon < \delta \), denote by \( S_w^{\delta - \varepsilon} \) the weakly \((\delta - \varepsilon, \ldots, \delta - \varepsilon)\)-efficient solution set \( \{x \in \Omega : \psi(x) \leq \delta - \varepsilon\} \) which converges to \( \{x \in \Omega : \psi(x) \leq \delta\} \) as \( \varepsilon \to 0 \).

**Proposition 4.3** For any \( 0 < \varepsilon < \delta \) and \( k \in \mathbb{N} \), we have

\[ \text{vol} \left( S_w^{\delta - \varepsilon} \setminus A(\delta, k) \right) \leq \frac{1}{\varepsilon} \int_{\Delta} (\psi_k - \tilde{\psi})dx. \]
Proof} Note that $\psi_k - \tilde{\psi} \geq 0$ on $\Delta$. For any $k \in \mathbb{N}$, we have

$$\text{vol} \left( S_w^{\delta-k} \setminus \mathcal{A}(\delta, k) \right) = \text{vol} \left( \{ x \in \Omega : \tilde{\psi}(x) \leq \delta - \varepsilon, \psi_k(x) > \delta \} \right)$$

$$\leq \text{vol} \left( \{ x \in \Omega : \psi_k(x) - \tilde{\psi}(x) > \varepsilon \} \right)$$

$$\leq \frac{1}{\varepsilon} \int_{\Omega} (\psi_k - \tilde{\psi}) \, dx$$

$$\leq \frac{1}{\varepsilon} \int_{\Delta} (\psi_k - \tilde{\psi}) \, dx.$$  

\[ \square \]

For a fixed $\delta > 0$, Proposition 4.3 indicates that the smaller the value $\int_\Delta \psi_k \, dx$ is, the better the set $S_w^\delta$ is approximated by $\mathcal{A}(\delta, k)$ from inside. Hence, we can compare the values $\int_\Delta \psi_k \, dx$ to measure quality of the sets $\mathcal{A}(\delta, k)$ with different $k \in \mathbb{N}$ in approximating $S_w^\delta$.

### 4.2 Computational aspects

Now we follow the scheme proposed in [33, Section 3.3] to construct a sequence of polynomials $(\psi_k)_{k \in \mathbb{N}} \in \mathbb{R}[x]$ as defined in Proposition 4.2.

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

$$h_{1,1}(x, y, z) = p_1(x)q_1(y) - p_1(y)q_1(x) - zq_1(x)q_1(y), \ldots,$$

$$h_{1,m}(x, y, z) = p_m(x)q_m(y) - p_m(y)q_m(x) - zq_m(x)q_m(y),$$

$$h_{2,1}(x, y, z) = g_1(y), \ldots, h_{2,r}(x, y, z) = g_r(y),$$

$$h_{2,r+1}(x, y, z) = 1 - y^2, \ldots, h_{2,r+n}(x, y, z) = 1 - y_n^2,$$

$$h_{3,1}(x, y, z) = 1 - x_1^2, \ldots, h_{3,n}(x, y, z) = 1 - x_n^2,$$

$$h_{4,1}(x, y, z) = (f^{\text{upper}} - f^{\text{lower}})^2 - z^2.$$  \hspace{1cm} (7)

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

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

Let $\lambda$ be the scaled Lebesgue measure on $\Delta$, i.e., $d\lambda(x) = dx/2^n$, and

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

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

For each $k \in \mathbb{N}$, with $k \geq \bar{k}$ where

$$\bar{k} := \max \left\{ \left\lfloor \frac{\deg h_{i, j}}{2} \right\rfloor, i = 1, \ldots, 4, j \in J_i \right\}.$$
consider the following optimization problem,

\[
\rho_k^* := \inf_{\phi, \sigma_0, \sigma_{i,j}} \int_{\Delta} \phi(x) d\lambda(x) \left( = \sum_{a \in \mathbb{N}_2^{|x|^2}} c_a Y_a \right)
\]

\[
\begin{aligned}
\text{s.t. } & \phi(x) = \sum_{a \in \mathbb{N}_2^{|x|^2}} c_a x^a \in \mathbb{R}[x]_{2k}, \ c_a \in \mathbb{R}, \\
& \phi(x) - z = \sigma_0 + \sum_{i=1}^{4} \sum_{j \in J_i} \sigma_{i,j} h_{i,j}, \ \sigma_0, \sigma_{i,j} \in \Sigma^2[x,y,z], \\
& \deg(\sigma_0), \deg(\sigma_{i,j} h_{i,j}) \leq 2k, \ i = 1, \ldots, 4, \ j \in J_i,
\end{aligned}
\]

(P_k)

which can be reduced to an SDP problem (c.f. [29, 31]). Clearly, for any \((\phi, \sigma_0, \sigma_{i,j})\) feasible to (P_k), we have \(\phi(x) \geq \tilde{\psi}(x)\) on \(\Delta\). The following result follows directly from [33, Theorem 5] and we include here a brief proof for the sake of completeness. It shows that we can compute the sequence of polynomials \(\{\psi_k \in \mathbb{R}[x] : k \in \mathbb{N}\}\) in Proposition 4.2 by solving (P_k).

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

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

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

\[
\rho^* := \inf_{\phi} \int_{\Delta} \phi(x) d\lambda(x) \left( = \sum_{a \in \mathbb{N}_2^{|x|^2}} c_a Y_a \right)
\]

\[
\begin{aligned}
\text{s.t. } & \phi(x) = \sum_{a \in \mathbb{N}_2^{|x|^2}} c_a x^a \in \mathbb{R}[x], \ c_a \in \mathbb{R}, \\
& \phi(x) - z \geq 0, \ \forall (x, y, z) \in K.
\end{aligned}
\]

It is clear that \(\Delta, K\) are compact and \(K_x\) is nonempty for every \(x \in \Delta\). Then, by [31, Corollary 2.6], it holds that \(\rho^* = \int_{\Delta} \tilde{\psi}(x) d\lambda(x)\). Let \((\phi_\ell)_{\ell \in \mathbb{N}}\) be a minimizing sequence of the above problem. For any \(\ell \in \mathbb{N}\), let \(\phi_\ell(x) = \phi_\ell(x) + 1/\ell\), then we have \(\phi_\ell(x) - z \geq 1/\ell > 0\) on \(K\). Notice that

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

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

\[
\int_{\Delta} \tilde{\psi}(x) d\lambda(x) = \rho^* \leq \rho_{k_\ell}^* \leq \int_{\Delta} \phi_\ell(x) d\lambda(x) + \frac{1}{\ell} \downarrow \rho^* = \int_{\Delta} \tilde{\psi}(x) d\lambda(x).
\]
As $\rho_k^*$ is monotone, we have $\lim_{k \to \infty} \rho_k^* = \int_\Delta \tilde{y}(x) d\lambda(x)$.

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

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

in (7) and reset $J_1 = \{1, \ldots, m+1\}$. Denote the following subsets of variables $I_1 = \{x, y, z\}$, $I_2 = \{y\}$, $I_3 = \{x\}$ 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) $\sum_{\alpha \in \mathbb{N}^n_{2k}} c_\alpha x^\alpha - z$ 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

$$\tilde{\rho}_k^* := \inf_{\phi, \sigma_{i,0}, \sigma_{i,j}} \int_\Delta \phi(x) d\lambda(x) \;
\left( = \sum_{\alpha \in \mathbb{N}^n_{2k}} c_\alpha x^\alpha \right)$$

s.t. $\phi(x) = \sum_{\alpha \in \mathbb{N}^n_{2k}} c_\alpha x^\alpha \in \mathbb{R}[x]_{2k}, \; c_\alpha \in \mathbb{R}$, 

$$\phi(x) - z = \sum_{i=1}^4 \left( \sigma_{i,0} + \sum_{j \in J_i} \sigma_{i,j} h_{i,j} \right), \; \sigma_{i,0}, \sigma_{i,j} \in \Sigma^2[I_i].$$

$$\deg(\sigma_{i,0}), \deg(\sigma_{i,j} h_{i,j}) \leq 2k, \; i = 1, \ldots, 4, \; j \in J_i.$$

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

**Proof** Let $\phi'_k$ 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 $\tilde{k} \in \mathbb{N}$ and $\sigma^{(\ell)}_{i,0}, \sigma^{(\ell)}_{i,j} \in \Sigma^2[I_i], i = 1, \ldots, 4, \; j \in J_i$ such that $(\phi'_k, \sigma^{(\ell)}_{i,0}, \sigma^{(\ell)}_{i,j})$ is a feasible solution to $(SP_{k'}^\ell)$. Hence, the conclusion follows from the proof of Theorem 4.2.

Notice that an $n$-variate sum of squares of polynomials of degree $2k$ can be represented by a positive semidefinite matrix of size $\binom{n+k}{k}$ in the canonical monomial basis. Comparing $(SP_k)$ with $(P_k)$, although there are some extra sums of squares of polynomials in $(SP_k)$ and it might happen that $\tilde{\rho}_k^* > \rho_k^*$, the positive semidefinite matrices to represent $\sigma_{i,j}$ for $i = 2, 3, 4, \; j \in J_i$ can be significantly reduced especially when the numbers $n$ and $k$ are large. See Examples 4.1 and 4.2 for numerical comparisons.

### 4.3 Approximation estimate of $\mathcal{A}(\delta, k)$

In this subsection, let us denote by $(\psi_k, \sigma_0^{(k)}, \sigma_{i,j}^{(k)})$ an optimal solution to $(P_k)$ for any $k \in \mathbb{N}$ with $k \geq \tilde{k}$, and consider the set $\mathcal{A}(\delta, k) = \{x \in \Omega : \psi_k(x) \leq \delta\}$ for $\delta > 0$. According to
Theorem 4.1, Corollary 4.1 and Remark 4.1, the set $\mathcal{A}(\delta, k)$, which is an inner approximation of $S_w^\delta$, can be taken as an approximation of $S_w$ with sufficiently small $\delta > 0$ and large $k \in \mathbb{N}$ under some conditions. Next, based on some existing work in the literature, we establish some results about the approximation quality of $\mathcal{A}(\delta, k)$ as $\delta \to 0$ and $k \to \infty$, respectively.

**Definition 4.1** We say that the Mangasarian–Fromovitz constraint qualification (MFCQ) holds at a given point $x \in \Omega$ if there exists a vector $d \in \mathbb{R}^n$ such that $\langle \nabla g_j(x), d \rangle > 0$ for all $j \in \{ j : g_j(x) = 0, \ j = 1, \ldots, r \}$.

For any positive integers $a$ and $b$, let

$$\mathcal{R}(a, b) := \begin{cases} b(3b - 3)^{a-1} & \text{if } b \geq 2, \\ 1 & \text{if } b = 1. \end{cases}$$

**Theorem 4.4** (c.f. [21, Corollary 1.2], [46, Theorem 1.1]) There exist constants $c > 0$ and $\alpha > 0$ such that

$$c \mathrm{dist}(x, S_w) \leq \lceil \psi(x) \rceil^\alpha \text{ for all } x \in \Omega.$$  

Moreover, if $q_i = 1$ for all $i = 1, \ldots, m$, and the (MFCQ) holds on $\Omega$, then we can choose

$$\alpha = \frac{1}{\mathcal{R}(n + m + r + 2)(n + 1) + r - 1, d + 4},$$

where $d := \max\{\deg(p_i), i = 1, \ldots, m, \ \deg(g_j), j = 1, \ldots, r\}$.

**Proof** Since $\Omega$ is a semi-algebraic compact set and $\psi$ is a semialgebraic function due to the Tarski–Seidenberg Theorem (c.f. [21, Theorem 1.5]), the existence of $c$ and $\alpha$ follows from the classic Łojasiewicz inequality (c.f. [21, Corollary 1.2]). When $q_i = 1$ for all $i = 1, \ldots, m$, and the (MFCQ) holds on $\Omega$, the explicit exponent $\alpha$ is given in [46, Theorem 1.1 and Remark 1.1 (i)].

**Corollary 4.2** Suppose that $q_i = 1$ for all $i = 1, \ldots, m$, and the (MFCQ) holds on $\Omega$, then there exists a constant $c > 0$ such that for any $d > 0$, $\mathcal{A}(\delta, k) \subset S_w^\delta \subset S_w + dB$ holds for any $k \in \mathbb{N}$ whenever $\delta < \frac{1}{(cd)^{\frac{1}{\alpha}}}$ where $\alpha$ in defined in (8).

**Proof** Let $c$ be the constant in Theorem 4.4 and fix two numbers $d > 0$, $\delta > 0$ with $\delta < \frac{1}{(cd)^{\frac{1}{\alpha}}}$. For any point $x \in S_w^\delta$, by Theorem 4.4, it holds that

$$\mathrm{dist}(x, S_w) \leq \frac{1}{c} \lceil \psi(x) \rceil^\alpha \leq \frac{1}{c} \delta^\alpha \leq d,$$

which implies that $S_w^\delta \subset S_w + dB$. Since $\mathcal{A}(\delta, k) \subset S_w^\delta$ any $k \in \mathbb{N}$, the conclusion follows.

Now let us fix a $\delta > 0$. By Theorem 4.1, the set $\mathcal{A}(\delta, k)$ approximates $S_w^\delta$ from inside as $k \to \infty$ in the following sense

$$\mathrm{vol}(\{ x \in \Omega : \psi(x) < \delta \}) \leq \lim_{k \to \infty} \mathrm{vol}(\mathcal{A}(\delta, k)) \leq \mathrm{vol}(S_w^\delta).$$

By Proposition 4.3 and Theorem 4.2, for any $0 < \varepsilon_1 < \delta$, $\varepsilon_2 > 0$, there exists $k(\varepsilon_1, \varepsilon_2) \in \mathbb{N}$ such that $\mathrm{vol}(S_w^{\delta-\varepsilon} \setminus \mathcal{A}(\delta, k)) < \varepsilon_2$ whenever $k > k(\varepsilon_1, \varepsilon_2)$. Next, inspired by [26], we give an estimation of the number $k(\varepsilon_1, \varepsilon_2)$ for given $\varepsilon_1, \varepsilon_2 > 0.$
Assumption 4.1 (i) The origin belongs to the interior of $K$; (ii) The denominators $q_i(x) > 0$ over $\Delta$ for all $i = 1, \ldots, m$.

Remark 4.2 (i) If $\Omega$ has a non-empty interior, then $K$ has a non-empty interior and hence we can translate $K$ such that the origin belongs to its interior. In fact, suppose that there exists an open set $\mathcal{O} \subset \Omega$ and fix an $\varepsilon > 0$ such that $\varepsilon < f^\text{upper} - f^\text{lower}$. Due to the continuity of $f_i(x)$’s on $\Omega$, by shrinking $\mathcal{O}$ if necessary, we can assume that $|f_i(x) - f_i(y)| < \varepsilon$ for all $x, y \in \mathcal{O}$ and $i = 1, \ldots, m$. Then, under (A2), it holds that $p_i(x)q_i(y) - p_i(y)q_i(x) > -\varepsilon q_i(x)q_i(y)$ for all $x, y \in \mathcal{O}$ and $i = 1, \ldots, m$. Therefore, the open set $\mathcal{O} \times \mathcal{O} \times (f^\text{lower} - f^\text{upper}, -\varepsilon) \subset K$.

(ii) If Assumption 4.1 (ii) holds, then $\hat{\psi}(x) = \psi(x)$ on $\Delta$. Consequently, $\hat{\psi}(x)$ is Lipschitz on $\Delta$ by the compactness of $\Delta$ and Proposition 3.1 (ii).

(iii) By substituting $\frac{p_iq_i}{q_i}$ for $\frac{p_i}{q_i}$, Assumption 4.1 (ii) can be weakened as $\Delta \subseteq \bigcap_{i=1}^m \text{dom}(f_i)$.

Recall that Assumption 2.1 holds for the defining polynomials of $K$. Applying [26, Theorem 3], which is a consequence of the fundamental result [44, Theorem 6] and [27, Corollary 1], to the set $K$ with defining polynomials $h_{i,j}$’s in (7), we obtain

Proposition 4.4 Let $\phi(x, y, z) \in \mathbb{R}[x, y, z]$ be strictly positive on $K$ and let Assumption 4.1 (i) hold. Then, $\phi$ has the representation

$$\phi = \sigma_0 + \sum_{i=1}^4 \sum_{j \in J_i} \sigma_{i,j}h_{i,j},$$

where $\sigma_0, \sigma_{i,j} \in \Sigma^2[x, y, z]$, $\deg(\sigma_0), \deg(\sigma_{i,j}h_{i,j}) \leq 2k$ for $i = 1, \ldots, 4$, $j \in J_i$, whenever

$$k \geq c_K \exp \left[ \left( 3^{\deg(\phi) + 1}k^{\deg(\phi)}(\deg(\phi))^2(2n + 1)^{\deg(\phi)}\frac{\max_{(x, y, z) \in K} \phi(x, y, z)}{\min_{(x, y, z) \in K} \phi(x, y, z)} \right)^{\frac{cK}{\max(\phi)}} \right],$$

for some constant $c_K$ depending only on $h_{i,j}$’s and

$$c := \sup\{t > 0 : [-t, t]^{2n+1} \subset K\}. \tag{9}$$

Consider the following assumption about the finite Gibbs phenomenon happening when approximating the function $\hat{\psi}$ on $\Delta$ with a series of polynomials.

Assumption 4.2 (c.f. [26, Assumption 2]) There is a sequence $\{\hat{\psi}_k\}_{k \in \mathbb{N}}$ with

$$\hat{\psi}_k := \arg\min_{\phi \in \mathbb{R}[x]} \int_{\Delta} (\phi - \hat{\psi})dx \quad \text{s.t.} \quad \phi \geq \hat{\psi} \text{ on } \Delta, \tag{10}$$

and a finite constant $c_G$ such that $\max_{x \in \Delta} \hat{\psi}_k(x) \leq c_G$ holds for all $k \in \mathbb{N}$.

Proposition 4.5 Suppose Assumptions 4.1 and 4.2 hold. Let $\hat{\psi}_k \in \mathbb{R}[x]$ be the polynomial in Assumption 4.2, then for any $k \in \mathbb{N}$ it holds that

$$\int_{\Delta} (\hat{\psi}_k - \hat{\psi})dx \leq \frac{c}{k}$$

for some constant $c$ depending on $\hat{\psi}$. 

\(\square\) Springer
Proposition 4.5

Let \( \hat{\psi}(x) \) with \( \hat{\psi}(x) \) as the modulus of continuity of \( \psi \). According to Assumptions 4.2, \( \hat{\psi}(x) \) is Lipschitz on \( \Delta \). That is, there exists \( L > 0 \) such that for every \( x, y \in \Delta \) with \( \| y - x \| \leq t \), it holds \( |\hat{\psi}(y) - \hat{\psi}(x)| \leq L t \). It implies that \( \omega^\Delta_{\psi}(t) \leq L t \) for any \( x \in \Delta \). Then, we have \( \hat{\omega}_{\psi}(1/k) \leq \omega^\Delta_{\psi}(1/k) \) for some constant \( c \) depending on \( \hat{\psi} \). By Remark 4.2 (ii), \( \hat{\psi}(x) \) is Lipschitz on \( \Delta \). That is, there exists \( L > 0 \) such that for every \( x, y \in \Delta \) with \( \| y - x \| \leq t \), it holds \( |\hat{\psi}(y) - \hat{\psi}(x)| \leq L t \). It implies that \( \omega^\Delta_{\psi}(t) \leq L t \) for any \( x \in \Delta \). Then, we have \( \hat{\omega}_{\psi}(t) \leq \omega^\Delta_{\psi}(t) \). The conclusion follows by letting \( \hat{c} := 2^n L c \).

Theorem 4.5

Suppose Assumptions 4.1 and 4.2 hold. Then for any \( 0 < \varepsilon_1 < \delta, \varepsilon_2 > 0 \), it holds that \( \Omega \left( S^\Delta_{\psi} \setminus A(\delta, k) \right) < \varepsilon_2 \) whenever

\[
k \geq c K \exp \left[ \left( \frac{3 \left( \frac{2 \hat{c}}{\delta} \right)^2}{6 k n + 3 k} \varepsilon + 2^{n+1} \left( c G + f^{\text{upper}} - f^{\text{lower}} \right) \right) \right]c K
\]

\[
= O \left( \exp \left[ \frac{1}{\varepsilon c K} \left( 6 k n + 3 k \right)^{2 \varepsilon K} \right] \right)
\]

where \( \varepsilon = \varepsilon_1 \varepsilon_2 \), \( c K, c G \) and \( \hat{c} \) are the constants in Proposition 4.4, Assumption 4.2 and Proposition 4.5, respectively, and \( k \) is defined in (9).

Proof

By Proposition 4.3, for any \( k \in \mathbb{N} \) with \( k \geq \tilde{k} \), we have

\[
\Omega \left( S^\Delta_{\psi} \setminus A(\delta, k) \right) \leq \frac{1}{\varepsilon_1} \int_{\Delta} (\psi_k - \hat{\psi})dx.
\]

Let \( \tilde{k} = \lceil \frac{2 \hat{c}}{\varepsilon_1 \varepsilon_2} \rceil \) where \( \hat{c} > 0 \) is the constant in Proposition 4.5 and let \( \hat{\psi}_k \) be the minimizer in (10) with \( k = \tilde{k} \). By Proposition 4.5, it holds

\[
\int_{\Delta} (\psi_k - \hat{\psi})dx \leq \frac{\hat{c}}{\tilde{k}} \leq \frac{\varepsilon_1 \varepsilon_2}{2}.
\]

Let \( \hat{\phi} := \frac{\varepsilon_1 \varepsilon_2}{2n+1} + \hat{\psi}_k \). As \( \hat{\psi}_k \geq \hat{\psi} \) on \( \Delta \), it is easy to check that \( \hat{\phi}(x) - z \geq \frac{\varepsilon_1 \varepsilon_2}{2n+1} \) for all \( (x, y, z) \in K \). According to Assumption 4.2, \( \hat{\phi}(x) - z \leq \frac{\varepsilon_1 \varepsilon_2}{2n+1} + c G + f^{\text{upper}} - f^{\text{lower}} \) for all \( (x, y, z) \in K \). By Proposition 4.4, \( \hat{\phi} - z \) has the representation

\[
\hat{\phi} - z = \hat{\sigma}_0 + \sum_{i=1}^{4} \sum_{j \in J_i} \hat{\sigma}_{i,j} h_{i,j}.
\]
where $\hat{\sigma}_0, \hat{\sigma}_{i,j} \in \Sigma^2[x, y, z]$, $\deg(\hat{\sigma}_0), \deg(\hat{\sigma}_{i,j}h_{i,j}) \leq 2k$ for $i = 1, \ldots, 4$, $j \in J_i$, whenever

$$k \geq c_K \exp \left[ \left( \frac{3}{\varepsilon_1} \right)^{2k} \left( 2n + 1 \right)^{2k} \left[ \frac{2c}{\varepsilon_1} \right]^{2k} \left[ \frac{2c}{\varepsilon_1} \right]^{\frac{\varepsilon_1 \varepsilon_2}{2n+1}} \frac{cG + f^{\text{upper}} - f^{\text{lower}}}{\varepsilon_1 \varepsilon_2} \right]^{c_K},$$

$$= c_K \exp \left[ \left( 3 \right)^{2k-2} \left( 6\kappa n + 3\kappa \right)^{2k} \left[ \frac{2c}{\varepsilon} \right]^{2k} \left[ \frac{2c}{\varepsilon} \right]^{\frac{\varepsilon_1 \varepsilon_2}{2n+1}} \frac{cG + f^{\text{upper}} - f^{\text{lower}}}{\varepsilon_1 \varepsilon_2} \right]^{c_K} =: k(\varepsilon_1, \varepsilon_2)$$

with the constant $c_K$ depending on $h_{i,j}$’s and $k$ defined in (9). In other words, $(\hat{\phi}, \hat{\sigma}_0, \hat{\sigma}_{i,j})$ is feasible to (P) whenever $k \geq k(\varepsilon_1, \varepsilon_2)$. Then, for any $k \geq k(\varepsilon_1, \varepsilon_2)$, due to (11), it holds that

$$\text{vol} \left( S_{w^\delta - \varepsilon_1} \setminus A(\delta, k) \right) \leq \frac{1}{\varepsilon_1} \int_{\Delta} (\psi_1 - \tilde{\psi}) \, dx$$

$$\leq \frac{1}{\varepsilon_1} \int_{\Delta} (\hat{\phi} - \tilde{\psi}) \, dx$$

$$= \frac{1}{\varepsilon_1} \int_{\Delta} (\hat{\phi} - \tilde{\psi}_k) \, dx + \frac{1}{\varepsilon_1} \int_{\Delta} (\tilde{\psi}_k - \tilde{\psi}) \, dx$$

$$\leq \frac{1}{\varepsilon_1} \frac{\varepsilon_1 \varepsilon_2}{2^{n+1}} \text{vol}(\Delta) + \frac{1}{\varepsilon_1} = \varepsilon_2$$

The conclusion follows.

4.4 Comparisons with existing SDP relaxation methods

Now, we compare our method with the recent existing work in [45] and [40]. 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., $q_i(x) = 1, i = 1, \ldots, m$.

To get weakly efficient solutions to (VROP), Nie and Yang [45] 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 [43]. 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

$$\min \ w_1 f_1(x) + \cdots + w_m f_m(x) \quad \text{s.t.} \quad x \in \Omega, \quad (12)$$

and the Chebyshev scalarization scalarizes the problem (VROP) to

$$\min_{x \in \Omega} \max_{1 \leq i \leq m} w_i (f_i(x) - f_i^*), \quad (13)$$

where $(f_1^*, \ldots, f_m^*)$ is the ideal point of the problem (VROP) with each $f_i^*$ being the goal which decision maker wants to achieve for the objective $f_i$. In general, by the scalarizations (12) and (13), 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 $\{A(\delta, k)\}$ computed by our method can approximate the whole set of weakly efficient solutions in some sense under certain conditions.

On the other hand, Magron et al. [40] 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 (12) and (13), as well as the parametric sublevel set approximation method which is inspired by [18] and amounts to solving the following parametric problem

\[
\min_{x \in \Omega} f_2(x) \quad \text{s.t.} \quad f_1(x) \leq w,
\]  

(14)

with a parameter \( w \in [\min_{x \in \Omega} f_1(x), \max_{x \in \Omega} f_1(x)] \). By treating \( w \) in (12), (13) and (14) as a parameter and employing the “joint+marginal” approach proposed in [31], they associated each scalarization problem a hierarchy of SDP relaxations and obtained an approximation of the Pareto curve by solving an inverse problem (for (12) and (13)) or by building a polynomial underestimator (for (14)). However, when using the scalarization problems (12) and (13), the approach in [40] requires that for almost all the values of the parameter \( w \), these parametric problems (12) and (13) 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).

 Compared with the methods proposed in [45] and [40], our approximation \( \mathcal{A}(\delta, k) \) of \( S_w \) represented as the intersection of the sublevel set of a single polynomial and the feasible set is more desirable in some applications. For instance, it can be used in optimization problems with (weakly) efficient constraints. An (weakly) efficient constraint can be replaced by the polynomial inequality \( \psi_k(x) \leq \delta \) with small \( \delta > 0 \) and large \( k \in \mathbb{N} \) (see Example 4.3). To give a motivation example, we briefly introduce a problem from the efficiency measurement approach of data envelopment analysis (see [17, 22] for more details). Given input vectors \( u^{(i)} \in \mathbb{R}^{n_1} \) and output vectors \( v^{(i)} \in \mathbb{R}^{n_2}, i = 1, \ldots, N \), collected from \( N \) firms, to assess the (relative) efficiency of the \( k \)-th firm, its manager need solve a linear fractional programming problem

\[
\max_{x_1, x_2 \geq 0} \frac{x_1^T v^{(k)}}{x_1^T u^{(k)}} \quad \text{s.t.} \quad \frac{x_1^T v^{(i)}}{x_1^T u^{(i)}} \leq 1, \quad i = 1, \ldots, N.
\]  

(LFP)

In other words, the manager need select weights \( x = (x_1, x_2) \) to maximize the \( k \)-th firm productivity ratio, with the constraints that the productivity ratios are bounded by 1 for all firms. In practice, most production systems produce output through a sequence of stages. Assuming that the \( k \)-th firm is a two-stage firm with subunit (I) and subunit (II). Subunit (I) uses exogenous inputs to produce the intermediate outputs, which are used as input by subunit (II) together with its usage of the exogenous inputs to produce the final outputs of the firm. To assess the efficiency of the two subunits, each submanager need solve an (LFP) problem with respective inputs and outputs. However, the weights chosen by the two submanagers should be consistent and usually cannot simultaneously maximize productivity ratios of both subunits. Hence, the submanagers should select weights from all (weakly) efficient solutions to a multiobjective (LFP) problem, which can be formulated as a (VROP) problem studied in this paper. A further question is that what weights the two submanagers should select among all (weakly) efficient solutions. A natural approach is to select the weights that make the overall firm as efficient as possible. Therefore, one should solve an (LFP) problem with an extra (weakly) efficient constraint corresponding to the multiobjective (LFP) subproblem, which is apparently not easy. To get an approximation solution to such a problem, the extra constraint can be replaced by \( \psi_k(x) \leq \delta \) from our approximation \( \mathcal{A}(\delta, k) \) of the weakly efficient solution set of the multiobjective (LFP) subproblem, with sufficiently small \( \delta > 0 \).
and large $k \in \mathbb{N}$. Then, the problem becomes a standard fractional programming problem which is well studied in the literature (c.f. [2, 12, 50, 51]).

### 4.5 Numerical experiments

Here we present some numerical examples to illustrate the behavior of the sets $A(\delta, k)$ in approximating $S_w$ as $\delta \to 0$ and $k \to \infty$. We use the software Yalmip [38] to implement the problems (SP$^k$) and call the SDP solver SeDuMi [52] to solve the resulting SDP problems. For the examples with $m = 2$, to show how close the sets $A(\delta, k)$ in approximating $S_w$, we illustrate the corresponding images of $f(\Omega)$ and $f(A(\delta, k))$. To this end, we choose a square containing $\Omega$. For each point $u$ on a uniform discrete grid inside the square, we check if $u \in \Omega$ (resp., $u \in A(\delta, k)$). If so, we have $(f_1(u), f_2(u)) \in f(\Omega)$ (resp., $(f_1(u), f_2(u)) \in f(A(\delta, k))$) and we plot the point $(f_1(u), f_2(u))$ in grey (resp., in red) in the image plane. All numerical experiments in the sequel were carried out on a PC with 4-Core Intel i5 2 GHz CPUs and 16 G RAM.

**Example 4.1** Consider the problem

$$
\begin{align*}
\min_{\mathbb{R}^3} & \quad (x_1, x_2, x_1^2 + x_2^2) \\
\text{s.t.} & \quad x \in \Omega_1 := \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}.
\end{align*}
$$

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

$$S_w = \{x \in \mathbb{R}^2 : x_1 \leq 0, x_2 \leq 0, x_1^2 + x_2^2 \leq 1\}.
$$

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 $S^{\delta}_w$ consists of the following four sets

- $\{x \in \mathbb{R}^2 : x_1 \geq \delta, x_2 \geq \delta, x_1^2 + x_2^2 \leq \delta\}$,
- $\{x \in \mathbb{R}^2 : x_1 \leq \delta, x_2 \geq \delta, x_1^2 + 2\delta x_1 - \delta - \delta^2 \leq 0, x_1^2 + x_2^2 \leq 1\}$,
- $\{x \in \mathbb{R}^2 : x_1 \leq \delta, x_2 \leq \delta, x_1^2 + x_2^2 \leq 1\}$,
- $\{x \in \mathbb{R}^2 : x_1 \geq \delta, x_2 \leq \delta, x_1^2 + 2\delta x_2 - \delta - \delta^2 \leq 0, x_1^2 + x_2^2 \leq 1\}$.

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

Now we compare the efficiency of the SDP problems (P$^k$) and (SP$^k$) in computing the approximations $\psi_k$ of $\tilde{\psi}$ over $\Delta$. According to Theorems 4.2 and 4.3, both of the optimal values $\rho^*_k$ and $\tilde{\rho}^*_k$ decreasingly converge to $\int_\Delta \tilde{\psi}(x) d\lambda(x)$ as $k \to \infty$. For fixed $\delta > 0$, by Proposition 4.3, the smaller the values $\rho^*_k$ and $\tilde{\rho}^*_k$ are, the better the set $S^{\delta}_w$ is approximated by $A(\delta, k)$ computed by (P$^k$) and (SP$^k$). Hence, we can measure the efficiency of (P$^k$) and (SP$^k$) by the optimal values and consumed CPU times. Denote by $t_k$ and $\tilde{t}_k$ the CPU times for computing $\psi_k$ using (P$^k$) and (SP$^k$), respectively. The results $(t_k, \rho^*_k)$ and $(\tilde{t}_k, \tilde{\rho}^*_k)$ for $k = 2, 3, 4, 5$, are reported in Table 1. Although $\tilde{\rho}^*_k$ is a little larger than $\rho^*_k$ as expected, the time for computing $\tilde{\rho}^*_k$ is much less than that for computing $\rho^*_k$ as $k$ increases. 

\[\square\]
Fig. 1 The set $S_{w}^{\delta}$ and its approximations $A(\delta, k)$ with $\delta = 0.1$, $k = 2, 3, 4$, in Example 4.1

Table 1 Comparison of numerical results for computing $(P_k)$ and $(SP_k)$ with $k = 2, 3, 4, 5$, in Example 4.1

| $k$ | $(t_k, \rho_k^*)$ | $(\tilde{t}_k, \tilde{\rho}_k^*)$ |
|-----|-------------------|---------------------|
| 2   | (3.8 s, 0.2547)   | (3.5 s, 0.2707)     |
| 3   | (20 s, 0.2256)    | (8 s, 0.2292)       |
| 4   | (603 s, 0.2186)   | (123 s, 0.2209)     |
| 5   | (9860 s, 0.2151)  | (1890 s, 0.2164)    |

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

$$\begin{align*}
\min_{x \in \mathbb{R}^2} & \quad x_2^2 - 2x_1x_2 + 1 \\
\text{s.t.} & \quad x \in \Omega_2 := \{x \in \mathbb{R}^2 : 1 - x_1^2 - x_2^2 \geq 0\}.
\end{align*}$$
Fig. 2  The images $f(A(\delta, k))$ (in red) and $f(\Omega_2)$ (in gray) in Example 4.2

Table 2  Comparison of numerical results for computing $(P_k)$ and $(SP_k)$ with $k = 3, 4, 5$, in Example 4.2

|        | $k = 3$       | $k = 4$       | $k = 5$       |
|--------|----------------|----------------|----------------|
| $(t_k, \rho^*_k)$ | (17 s, 0.8129) | (336 s, 0.6114) | (6108 s, 0.5529) |
| $(\tilde{t}_k, \tilde{\rho}^*_k)$ | (11 s, 0.8338) | (45 s, 0.6214)  | (616 s, 0.5592)  |

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

Like in Example 4.1, we compare the numbers $(t_k, \rho^*_k)$ and $(\tilde{t}_k, \tilde{\rho}^*_k)$ to show the advantage of $(SP_k)$ compared with $(P_k)$. The results $(t_k, \rho^*_k)$ and $(\tilde{t}_k, \tilde{\rho}^*_k)$ for $k = 3, 4, 5$, are reported in Table 2. As we can see, the two numbers $\rho^*_k$ and $\tilde{\rho}^*_k$ get very close as $k$ increases, while the times for computing $\tilde{\rho}^*_k$ are significantly less than those for computing $\rho^*_k$.
Example 4.3 Consider the problem

\[
\begin{aligned}
\text{Min} & \quad R^2 + \left( \frac{\sqrt{2}}{2} (x_1 - x_2), \frac{\sqrt{2}}{2} (x_1 + x_2) \right) \\
\text{s.t.} & \quad x \in \Omega_3 := \{ x \in \mathbb{R}^2 : g(x) := x_2^2 (1 - x_1^2) - (x_1^2 + 2x_2 - 1)^2 \geq 0 \}.
\end{aligned}
\]

In fact, the equality \( g(x) = 0 \) defines the so-called bicorn curve as shown in Fig. 3a. Hence, the feasible set \( \Omega_3 \) of this problem is the region enclosed by the bicorn curve and the image \( f(\Omega_3) \) is obtained by rotating \( \Omega_3 \) clockwise by 45° (Fig. 3b). It is clear that the weakly efficient solution set \( S_w \) consists of the points on the shorter path connecting the two singular points of the bicorn curve. As discussed in subsection 4.4, the linear scalarization (12) can only enable us to compute two points in \( S_w \), namely, the two singular points of the bicorn curve. By our method, we compute the approximation \( A(0.01, 4) \) and show it in Fig. 3a, which is the intersection of \( \Omega_3 \) and the area under the red curve defined by \( \psi_4(x) = 0.01 \). The image \( f(A(0.01, 4)) \) is illustrated in Fig. 3 (c), which shows that we can obtain good approximations of \( S_w \) including the ones corresponding to the sunken part of Pareto curve.

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

\[
\min \ x_1^2 + (x_2 - 1)^2 \quad \text{s.t.} \ (x_1, x_2) \in S_w,
\]

which is to compute the square of the Euclidean distance between the point (0, 1) and the curve \( S_w \). It is easy to see that the unique minimizer of the above problem is \( (0, \frac{1}{2}) \) and the minimum is \( \frac{4}{9} \approx 0.444 \). With the approximation \( A(0.01, 4) \) of \( S_w \), we consider the polynomial optimization problem

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

We solve this problem by Lasserre’s hierarchy of SDP relaxations (c.f. [29, 32]) with the software GloptiPoly [23], and get the certified minimizer \((-0.0000, 0.3473)\) and minimum 0.4260. \( \square \)
Example 4.4 Consider the problem

\[
\begin{align*}
\min_{\mathbb{R}^2_+} & \quad (-x_1^2, x_1^4 + x_2^2) \\
\text{s.t. } & \quad x \in \Omega_4 := \{x \in \mathbb{R}^2 : 1 - x_1^2 - x_2^2 \geq 0\}.
\end{align*}
\]

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

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

in the objective plane where \( t_1 = -x_1^2 \) and \( t_2 = x_1^4 + x_2^2 \). Clearly, for every point \((t_1, t_2) \in f(S_w)\), 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 [40] when using the scalarizations (12) and (13). 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 \( \psi_5(x) = 0.005 \) in Fig. 4a. The images \( f(\Omega_4) \) and \( f(\mathcal{A}(0.005, 5)) \) is shown in Fig. 4b, which illustrates that we can approximate the set of weakly efficient solutions as closely as possible.

5 Conclusions

In this paper, we provide a new scheme for approximating the set of all weakly (\( \epsilon \)-)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 [31, 33]. 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 (\( \epsilon \)-)efficient solutions to the problem (VROP). Moreover, the obtained results also significantly develop the recent achievements in [11, 35, 36] for vector polynomial optimization problems from convex settings to nonconvex settings.

Acknowledgements The authors are very grateful for the comments of the anonymous referees which helped to improve the presentation. 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 (YDZJ202201ZYTS302) and the Fundamental Research Funds for the Central Universities.

References

1. Amini Fasakhodi, A., Nouri, S.H., Amini, M.: Water resources sustainability and optimal cropping pattern in farming systems; a multi-objective fractional goal programming approach. Water Resour. Manage. 24(15), 4639–4657 (2010)
2. Bajalinov, E.B.: Linear-Fractional Programming: Theory, Methods, Applications and Software. Springer, New York (2003)
3. Bhati, D., Singh, P., Arya, R.: A taxonomy and review of the multi-objective fractional programming (mofp) problems. Int. J. Appl. Comput. Math. 3(3), 2695–2717 (2017)
4. Bhaya, E.S.: Interpolating operators for multiapproximation. J. Math. Stat. 6, 240–245 (2010)
5. Blanco, V., Puerto, J., Ali, S.E.H.B.: A semidefinite programming approach for solving multiobjective linear programming. J. Global Optim. 58(3), 465–480 (2014)
6. Borwein, J.M.: On the existence of Pareto efficient points. Math. Oper. Res. 8(1), 64–73 (1983)
7. Borza, M., Rambely, A.S.: A new method to solve multi-objective linear fractional problems. Fuzzy Inf. Eng. 13(3), 323–334 (2021)
8. Burachik, R.S., Kaya, C.Y., Rzizi, M.M.: A new scalarization technique and new algorithms to generate Pareto fronts. SIAM J. Optim. 27(2), 1010–1034 (2017)
9. Chankong, V., Haimes, Y.Y.: Multiobjective Decision Making: Theory and Methodology. North-Holland, Amsterdam (1983)
10. Chen, W., Unkelbach, J., Trofimov, A., Kooy, H., Bortfeld, T., Craft, D.: Including robustness in multi-criteria optimization for intensity modulated proton therapy. Phys. Med. Biol. 57(3), 591–608 (2012)
11. Chuong, T.D.: Second-order cone programming relaxations for a class of multiojective convex polynomial programs. Ann. Oper. Res. 311(2), 1017–1033 (2022)
12. Dinkelbach, W.: On nonlinear fractional programming. Manage. Sci. 13(7), 492–498 (1967)
13. Dutta, J., Kesarwani, P., Gupta, S.: Gap functions and error bounds for nonsmooth convex vector optimization problem. Optimization 66(11), 1807–1836 (2017)
14. Ehrmann, M.: Multicriteria Optimization, 2nd edn. Springer, Berlin (2005)
15. Fliege, J., Drummond, L.M.G., Svaiter, B.F.: Newton’s method for multiobjective optimization. SIAM J. Optim. 20(2), 602–626 (2009)
16. Golany, B., Hackman, S.T., Passy, U.: An efficiency measurement framework for multi-stage production systems. Ann. Oper. Res. 145(1), 51–68 (2006)
17. Gorissen, B.L., Hertog, D.D.: Approximating the Pareto set of multiobjective linear programs via robust optimization. Oper. Res. Lett. 40(5), 319–324 (2012)
18. Grimm, D., Netzer, T., Schweighofer, M.: A note on the representation of positive polynomials with structured sparsity. Arch. Math. 89(5), 399–403 (2007)
19. Gutjahr, W.J., Nolz, F.C.: Multicriteria optimization in humanitarian aid. Eur. J. Oper. Res. 252(2), 351–366 (2016)
20. Hà, H.-V., Pham, T.-S.: Genericity in Polynomial Optimization. World Scientific Publishing, Singapore (2017)
21. Hackman, S.T., Passy, U.: Maximizing a linear fractional function on a Pareto efficient frontier. J. Optim. Theory Appl. 113(1), 83–103 (2002)
22. Henrion, D., Lasserre, J.B., Löfberg, J.: GloptiPoly 3: moments, optimization and semidefinite programming. Optim. Methods Softw. 24(4–5), 761–779 (2009)
23. Huong, N.T.T., Yao, J.-C., Yen, N.D.: Geoffrion’s proper efficiency in linear fractional vector optimization with unbounded constraint sets. J. Global Optim. 78(3), 545–562 (2020)
24. Kim, D.S., Pham, T.-S., Tuyen, N.V.: On the existence of Pareto solutions for polynomial vector optimization problems. Math. Program. Ser. A 177(1–2), 321–341 (2019)
25. Korda, M., Henrion, D.: Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets. Optim. Lett. 12(3), 435–442 (2018)
26. Korda, M., Henrion, D., Jones, C.N.: Convergence rates of moment-sum-of-squares hierarchies for optimal control problems. Syst. Control Lett. 100, 1–5 (2017)
28. Kornbluth, J.S.H., Steuer, R.E.: Multiple objective linear fractional programming. Manage. Sci. 27(9), 1024–1039 (1981)
29. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. 11(3), 796–817 (2001)
30. Lasserre, J.B.: Convergent SDP-relaxations in polynomial optimization with sparsity. SIAM J. Optim. 17(3), 822–843 (2006)
31. Lasserre, J.B.: A “joint+ marginal” approach to parametric polynomial optimization. SIAM J. Optim. 20(4), 1995–2022 (2010)
32. Lasserre, J.B.: Moments, Positive Polynomials and their Applications. Imperial College Press, London (2010)
33. Lasserre, J.B.: Tractable approximations of sets defined with quantifiers. Math. Program. Ser. B 151(1), 507–527 (2015)
34. Laurent, M.: 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, pp. 157–270. Springer, New York, NY (2009)
35. Lee, J.H., Jiao, L.G.: Solving fractional multicriteria optimization problems with sum of squares convex polynomial data. J. Optim. Theory Appl. 176(2), 428–455 (2018)
36. Lee, J.H., Sisarat, N., Jiao, L.G.: Multi-objective convex polynomial optimization and semidefinite programming relaxations. J. Global Optim. 80(1), 117–138 (2021)
37. Liu, C.G., Ng, K.F., Yang, W.H.: Merit functions in vector optimization. Math. Program. 119(2), 215–237 (2009)
38. Löfberg, J.: YALMIP : a toolbox for modeling and optimization in MATLAB. In: 2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508), pp. 284–289 (2004)
39. Luc, D.T.: Multiobjective Linear Programming: An Introduction. Springer International Publishing, Switzerland (2016)
40. Magron, V., Henrion, D., Lasserre, J.B.: Approximating Pareto curves using semidefinite relaxations. Oper. Res. Lett. 42(6–7), 432–437 (2014)
41. Magron, V., Henrion, D., Lasserre, J.B.: Approximating Pareto curves using semidefinite relaxations. Oper. Res. Lett. 42(6–7), 432–437 (2014)
42. Metev, B., Gueorguieva, D.: A simple method for obtaining weakly efficient points in multiobjective linear fractional programming problems. Eur. J. Oper. Res. 126(2), 386–390 (2000)
43. Nayak, S., Ojha, A.K.: Multi-objective linear fractional programming problem with fuzzy parameters. In: Bansal, J.C., Das, K.N., Nagar, A., Deep, K., Ojha, A.K. (eds.) Soft Computing for Problem Solving, pp. 79–90. Springer Singapore, Singapore (2019)
44. Nie, J.: Tight relaxations for polynomial optimization and Lagrange multiplier expressions. Math. Program. 178(1), 1–37 (2019)
45. Nie, J., Schweighofer, M.: On the complexity of Putinar’s positivstellensatz. J. Complex. 23(1), 135–150 (2007)
46. Nie, J., Yang, Z.: The multi-objective polynomial optimization (2021). arXiv:2108.04336
47. Putinar, M.: Positive polynomials on compact semi-algebraic sets. Indiana Univ. Math. J. 42(3), 969–984 (1993)
48. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
49. Sawaragi, Y., Nakayama, H., Tanino, T.: Theory of Multiobjective Optimization. Academic Press Inc, Orlando, FL (1985)
50. Schaible, S., Shi, J.M.: Fractional programming: the sum-of-ratios case. Optim. Methods Softw. 18(2), 219–229 (2003)
51. Stancu-Minasian, I.M.: Fractional Programming: Theory, Methods and Applications. Springer, Netherlands (1997)
52. Sturm, J.F.: Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. Optim. Methods Softw. 11(1–4), 625–653 (1999)
53. Tanabe, H., Fukuda, E.H., Yamashita, N.: Proximal gradient methods for multiobjective optimization and their applications. Comput. Optim. Appl. 72(2), 339–361 (2019)
54. Tanabe, H., Fukuda, E.H., Yamashita, N.: New merit functions and error bounds for non-convex multi-objective optimization (2020). arXiv:2010.09333
55. Tanabe, H., Fukuda, E.H., Yamashita, N.: Convergence rates analysis of multiobjective proximal gradient methods. Optim. Lett. 17(2), 333–350 (2023)
56. Veeramani, C., Edalatpanah, S.A., Sharanya, S.: Solving the multiobjective fractional transportation problem through the neutrosophic goal programming approach. Discret. Dyn. Nat. Soc. 2021, 7308042 (2021)
57. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite program relaxations for polynomial optimization problems with structured sparsity. SIAM J. Optim. 17(1), 218–242 (2006)
58. Wierzbicki, A.P.: On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR Spektrum 8, 73–87 (1986)
59. Zhao, X.P., Köis, M.A., Yao, Y.H., Yao, J.-C.: A projected subgradient method for nondifferentiable quasiconvex multiobjective optimization problems. J. Optim. Theory Appl. 190, 82–107 (2021)
60. Zhao, X.P., Yao, J.-C.: Linear convergence of a nonmonotone projected gradient method for multiobjective optimization. J. Global Optim. 82, 577–594 (2022)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.