

Geometric duality results and approximation algorithms for convex vector optimization problems

Çağın Ararat* Simay Tekgül† Firdevs Ulus‡

September 27, 2022

Abstract

We study geometric duality for convex vector optimization problems. For a primal problem with a q -dimensional objective space, we formulate a dual problem with a $(q+1)$ -dimensional objective space. Consequently, different from an existing approach, the geometric dual problem does not depend on a fixed direction parameter and the resulting dual image is a convex cone. We prove a one-to-one correspondence between certain faces of the primal and dual images. In addition, we show that a polyhedral approximation for one image gives rise to a polyhedral approximation for the other. Based on this, we propose a geometric dual algorithm which solves the primal and dual problems simultaneously and is free of direction-biasedness. We also modify an existing direction-free primal algorithm in a way that it solves the dual problem as well. We test the performance of the algorithms for randomly generated problem instances by using the so-called primal error and hypervolume indicator as performance measures.

Keywords: Convex vector optimization, multiobjective optimization, approximation algorithm, scalarization, geometric duality, hypervolume indicator.

Mathematics Subject Classification (2020): 90B50, 90C25, 90C29.

1 Introduction

Vector optimization is a generalization of multiobjective optimization (MO) where the order relation over the objective vectors is determined by a general ordering cone. It has been applied in many fields including risk-averse dynamic programming [34], financial mathematics [2, 18, 27, 36], economics [48, 50], game theory [19, 28].

In vector optimization, generating efficient solutions which correspond to minimal elements in the objective space can be done either by solving single-objective optimization problems formed by the structure of the original problem, called scalarizations (see [17] and

*Bilkent University, Department of Industrial Engineering, Ankara, 06800 Turkey, cararat@bilkent.edu.tr

†The University of Edinburgh, School of Mathematics, Edinburgh, EH9 3FD United Kingdom, s.tekgul@ed.ac.uk

‡Corresponding author, Bilkent University, Department of Industrial Engineering, Ankara, 06800 Turkey, firdevs@bilkent.edu.tr

references therein), or by iterative algorithms which work directly in the decision space (see e.g. [8, 14, 20, 21, 22]). The overall aim for many applications is however to generate either (an approximation of) the whole set of minimal elements in the objective space [5, 49] or the set of all efficient solutions in the decision space [3]. Since the dimension of the decision space is in general much higher than that of the objective space and consequently many efficient solutions might map into a single objective value, the former is usually a more affordable goal.

By this motivation, Benson [5] introduced an outer approximation algorithm in 1998 that aims to generate the Pareto frontier of linear multiobjective optimization problems. This has led to vast literature in vector optimization. In the linear case, many variations of Benson’s algorithm have been proposed [9, 15, 29, 35, 53, 54]. We discuss the objective space-based algorithms in the literature for convex vector optimization problems (CVOPs) in more detail below. For non-convex vector optimization problems with special structure, we refer the reader to the recent works [10, 41, 44].

1.1 Literature review on objective space-based CVOP algorithms

For CVOPs, there are various approximation techniques in the literature, see e.g. [51]. Here, we focus on the approaches which generate approximations to the entire minimal set in the objective space. In 2011, Ehrgott, Shao and Schöbel [16] proposed an extension of Benson’s algorithm for convex MO problems. Later Löhne, Rudloff and Ulus [37] generalized this algorithm for CVOPs and recently, different variants have been proposed in [13, 33]. These algorithms solve a Pascoletti-Serafini [45] scalarization and a vertex enumeration problem in each iteration. This scalarization problem depends on two parameters: a direction vector and a reference point. Specifically, given a point v in the objective space, it determines the closest point (along a direction c) to v in the objective space as well as a corresponding feasible solution in the decision space. In [37], the reference point v is selected arbitrarily among the vertices of the current approximation and c is fixed throughout the algorithms, whereas in [13], the reference point v and a corresponding direction vector c are selected based on an additional procedure. In [33], several additional rules for selecting the two parameters of the Pascoletti-Serafini scalarization for the same algorithmic setup are proposed and compared.

Recently, for CVOPs, Ararat, Ulus and Umer [1] considered a norm-minimizing scalarization which requires only a reference point but no direction parameter, and proposed an outer approximation algorithm for CVOPs. As the proposed algorithm does not require a direction parameter, direction-biasedness is not a concern. However, the norm-minimizing scalarization has a nonlinear convex objective function, which is not the case for the Pascoletti-Serafini scalarization that is used in the primal algorithm in [37].

Apart from these *primal* algorithms, in [37], the authors also proposed a *geometric dual* variant of their primal algorithm, which is based on a geometric duality theory. The general theory of convex polytopes indicates that two polytopes are dual to each other if there exists an inclusion-reversing one-to-one mapping between their faces [26]. Similar to the duality relation between polytopes, a duality relation between the polyhedral image of the primal problem and the polyhedral image of a dual problem was introduced in [31] for multiobjective linear programming problems. Later, Luc [40] introduced *parametric duality* for these problems and studied the relationship between the parametric and geometric duality

notions. More recently, the equivalence between the two duality notions has been shown in [12].

In [30], a general duality theory is developed for the epigraph of a closed convex function and that of its conjugate function. As a special case, this theory is applied for CVOPs to obtain a duality relation between the primal and dual images; hence, the geometric duality theory in [31] is generalized to the non-polyhedral case. The geometric dual algorithm that is proposed in [37] is based on the duality relation in [30]. Accordingly, the fixed direction c that is used within the primal solution concept and algorithm is also used in the design of the geometric dual problem and the algorithm. Different from the primal algorithms, in the geometric dual algorithm of [37], instead of Pascoletti-Serafini or norm-minimizing scalarizations, only the well-known weighted sum scalarizations, which may be easier to deal with because of their simple structure, are solved.

1.2 The proposed approach and contributions

In this paper, we establish a *direction-free* geometric duality theory and a dual algorithm to solve CVOPs. More precisely, after recalling the primal solution concept in [1], we propose a direction-free geometric dual problem and a solution concept for it. Then, we prove a geometric duality relation between the images of the primal and dual problems, which is based on an inclusion-reversing one-to-one duality mapping. The proof of this relation is based on the general duality theory for epigraphs in [30]. We also prove that a polyhedral approximation for the primal image yields a polyhedral approximation for the dual image, and vice versa. Different from [30], the proposed geometric dual image does not depend on a fixed direction; but in order to handle this issue, the dimension of the objective space for the dual problem is increased by one. In this sense, it can also be seen as a generalization of the parametric duality for linear MO from [40] to CVOPs. Accordingly, the image of the proposed dual problem is not only a convex set but a convex cone. Due to this conic structure of the dual image, the dimension increase compared to [30] is not an additional source of computational burden, as also confirmed by the numerical results.

Based on the geometric duality theory, we propose a dual algorithm which solves the primal and dual problems simultaneously by solving only weighted sum scalarization problems. More precisely, the algorithm gives a finite ϵ -solution to the dual problem; moreover, it gives a finite weak $\tilde{\epsilon}$ -solution to the primal problem, where $\tilde{\epsilon}$ is determined by ϵ and the structure of the underlying ordering cone. We also modify the primal algorithm in [1] in a way that it returns a finite ϵ -solution to the dual problem as well.

We compare the proposed geometric duality and the dual algorithm with the ones in [30] and [37], respectively. In particular, we show that the proposed dual image and the dual image in [30] can be recovered from each other. Moreover, by fixing a suitable norm, we show that the dual algorithm in [37] can be seen as a special case of the proposed dual algorithm.

Finally, we test the performance of the proposed algorithm in comparison to the existing ones using two performance measures: primal error (PE) and hypervolume indicator (HV). While PE simply measures the Hausdorff distance between the primal image and its returned polyhedral approximation, novel to this work, we define HV as a hypervolume-based performance metric for convex vector optimization. Our definition is similar to hypervolume-based metrics for MO, see e.g. [4, 56, 57, 58]. The computational results suggest that the proposed

dual algorithm has promising performance.

The rest of the paper is structured as follows. Section 2 presents the basic concepts and notation. Section 3 introduces the primal problem, dual problem, and solution concepts for these problems. The geometric duality between the primal and dual problems is studied in Section 4. Section 5 provides the primal and dual algorithms. In Section 6, we compare the proposed approach with the existing ones in the literature. In Section 7, we provide some test results for performance comparisons.

2 Preliminaries

In this section, we provide the definitions and notations used throughout the paper. Let $q \in \mathbb{N} := \{1, 2, \dots\}$ be a positive integer. We denote the q -dimensional Euclidean space by \mathbb{R}^q . Let $\|\cdot\|$ be an arbitrary norm on \mathbb{R}^q . The associated dual norm is denoted by $\|\cdot\|_*$, that is, $\|z\|_* = \sup\{z^\top x \mid \|x\| \leq 1\}$ for each $z \in \mathbb{R}^q$. Throughout, $B(0, \epsilon) := \{z \in \mathbb{R}^q \mid \|z\| \leq \epsilon\}$ denotes the norm ball centered at $0 \in \mathbb{R}^q$ with radius $\epsilon > 0$.

For a set $A \subseteq \mathbb{R}^q$, we denote the convex hull, conic hull, interior, relative interior and closure of A by $\text{conv } A$, $\text{cone } A$, $\text{int } A$, $\text{ri } A$ and $\text{cl } A$, respectively. Recall that $\text{cone } A := \{\lambda a \mid a \in A, \lambda \geq 0\}$. The closed convex cone defined by $A^+ = \{y \in \mathbb{R}^q \mid \forall a \in A: y^\top a \geq 0\}$ is called the *dual cone* of A . It is well-known that the dual cone of A^+ is given by $A^{++} := (A^+)^+ = \text{cl cone conv } A$ whenever A is nonempty. The *recession cone* of A is defined by $\text{recc } A = \{c \in \mathbb{R}^q \mid \forall \lambda \geq 0, a \in A: a + \lambda c \in A\}$. An element $d \in \text{recc } A$ is a (*recession*) *direction* of A . Let $A \subseteq \mathbb{R}^q$ be convex and $F \subseteq A$ be a convex subset. If $\lambda y^1 + (1 - \lambda)y^2 \in F$ for some $0 < \lambda < 1$ holds only if both y^1 and y^2 are elements of F , then F is called a *face* of A . A zero-dimensional face is called an *extreme point* and a one-dimensional face is called an *edge* of A . If A is q -dimensional, then a $(q - 1)$ -dimensional face is called a *facet* of A . A face of A that is not the empty set and not A itself is called a *proper face* of A . We call F an *exposed face* of A if it can be written as the intersection of A and a supporting hyperplane of A . If A is q -dimensional, then an exposed face of A is also a proper face; the converse does not hold in general [30, Section 2.2].

Given nonempty sets $A, B \subseteq \mathbb{R}^q$, their Minkowski sum is defined by $A + B := \{a + b \mid a \in A, b \in B\}$. For $\lambda \in \mathbb{R}$, we also define $\lambda A := \{\lambda a \mid a \in A\}$. In particular, we have $A - B = A + (-1)B$.

Let $C \subseteq \mathbb{R}^q$ be a cone. It is called *pointed* if $C \cap -C = \{0\}$, *solid* if it has nonempty interior, and *nontrivial* if $C \neq \emptyset$ and $C \neq \mathbb{R}^q$. If $C \subseteq \mathbb{R}^q$ is a convex pointed cone, then the relation $\leq_C := \{(x, y) \in \mathbb{R}^q \times \mathbb{R}^q \mid y - x \in C\}$ is an antisymmetric partial order on \mathbb{R}^q [32, Theorem 1.18]; we write $x \leq_C y$ whenever $(x, y) \in \leq_C$.

Let $m \in \mathbb{N}$ and $\mathcal{X} \subseteq \mathbb{R}^m$ be a nonempty convex set. A function $f: \mathcal{X} \rightarrow \mathbb{R}^q$ is said to be *C -convex on \mathcal{X}* if $f(\lambda x^1 + (1 - \lambda)x^2) \leq_C \lambda f(x^1) + (1 - \lambda)f(x^2)$ for every $x^1, x^2 \in \mathcal{X}$ and $\lambda \in [0, 1]$ [32, Definition 2.4]. Given a function $g: \mathbb{R}^q \rightarrow \mathbb{R}$, the function $g^*: \mathbb{R}^q \rightarrow [-\infty, +\infty]$ defined by $g^*(w) := \sup_{z \in \mathbb{R}^q} (w^\top z - g(z))$, $w \in \mathbb{R}^q$, is called the *conjugate function* of g . For a set $A \subseteq \mathbb{R}^q$, the function I_A defined by $I_A(z) := 0$ for $z \in A$, and by $I_A(z) := +\infty$ for $z \notin A$ is called the *indicator function* of A . In this case, taking $g = I_A$ gives $g^*(w) = \sup_{z \in A} w^\top z$ for each $w \in \mathbb{R}^q$; g^* is called the *support function* of A ; we also define the *polar* of A as the set $A^\circ := \{w \in \mathbb{R}^q \mid g^*(w) \leq 1\}$.

Let $A \subseteq \mathbb{R}^q$ and let $C \subseteq \mathbb{R}^q$ be a closed convex pointed cone. The sets $\text{Min}_C A := \{y \in A \mid (\{y\} - C \setminus \{0\}) \cap A = \emptyset\}$, $\text{wMin}_C A := \{y \in A \mid (\{y\} - \text{int } C) \cap A = \emptyset\}$, $\text{Max}_C A := \{y \in A \mid (\{y\} + C \setminus \{0\}) \cap A = \emptyset\}$ are called the sets of C -minimal, weakly C -minimal, C -maximal elements of A , respectively [35, Definition 1.41]. An exposed face of A that only consists of (weakly) C -minimal elements is called a (weakly) C -minimal exposed face. An exposed face of A that consists of only C -maximal elements is called a C -maximal exposed face [35, Section 4.5].

Remark 2.1. Let $A \subseteq \mathbb{R}^q$ be a nonempty convex set and $C \subseteq \mathbb{R}^q$ be a convex cone. If $A = A + C$ and $w \in \mathbb{R}^q$ such that $\inf_{a \in A} w^\top a > -\infty$, then $w \in C^+$. Note that $A = A + \text{recc } A$ holds [55, Theorem 5.6]. If $A \subseteq H := \{z \in \mathbb{R}^q \mid w^\top z \geq r\}$ for some $w \in \mathbb{R}^q$ and $r \in \mathbb{R}$, then we have $w \in (\text{recc } A)^+$.

3 Primal and dual problems

In this paper, we consider a convex vector optimization problem and its geometric dual. The primal problem is defined as

$$\text{minimize } f(x) \text{ with respect to } \leq_C \text{ subject to } x \in \mathcal{X}, \quad (\text{P})$$

where the ordering cone $C \subseteq \mathbb{R}^q$ is nontrivial, pointed, solid, closed and convex; the vector-valued objective function $f: \mathcal{X} \rightarrow \mathbb{R}^q$ is C -convex and continuous; and the feasible set $\emptyset \neq \mathcal{X} \subseteq \mathbb{R}^m$ is compact and convex. The *upper image* of (P) is defined as

$$\mathcal{P} := \text{cl}(f(\mathcal{X}) + C),$$

where $f(\mathcal{X}) := \{f(x) \mid x \in \mathcal{X}\}$ is the image of \mathcal{X} under f . The following proposition collects some basic facts about the problem structure. Its proof is straightforward from the fact that \mathcal{X} is compact, hence we omit it.

Proposition 3.1. *The upper image \mathcal{P} is a closed convex set, the image $f(\mathcal{X})$ is a compact set, and it holds $\mathcal{P} = f(\mathcal{X}) + C$. Moreover, the primal problem (P) is bounded in the sense that $\{y\} + C \subseteq \mathcal{P}$ for some $y \in \mathbb{R}^q$.*

For a parameter vector $w \in C^+$, the convex program

$$\text{minimize } w^\top f(x) \text{ subject to } x \in \mathcal{X} \quad (\text{WS}(w))$$

is called the *weighted sum scalarization* of (P). Let p^w be the optimal value of (WS(w)), that is, $p^w := \inf_{x \in \mathcal{X}} w^\top f(x)$. Since \mathcal{X} is a nonempty compact set and f is a continuous function, it follows that $p^w \in \mathbb{R}$. The next proposition is a well-known result that will be used in the design of the geometric dual problem.

Proposition 3.2. [32, Corollary 5.29] *Let $w \in C^+ \setminus \{0\}$. Then, an optimal solution x^w of (WS(w)) is a weak minimizer of (P). The converse also holds: for each weak minimizer $x \in \mathcal{X}$ of (P), there exists $w \in C^+ \setminus \{0\}$ such that x is an optimal solution of (WS(w)).*

Now, let us define the geometric dual problem of (P) as

$$\text{maximize } \xi(w) \text{ with respect to } \leq_K \text{ subject to } w \in \mathcal{W}. \quad (\text{D})$$

In this problem, the objective function $\xi: \mathbb{R}^q \rightarrow \mathbb{R}^{q+1}$ is defined by

$$\xi(w) := (w_1, \dots, w_q, p^w)^\top, \quad w \in \mathcal{W}; \quad (1)$$

the ordering cone K is defined by $K := \text{cone}\{e^{q+1}\} = \{\lambda e^{q+1} \mid \lambda \geq 0\}$, where $e^{q+1} = (0, \dots, 0, 1)^\top \in \mathbb{R}^{q+1}$; and the feasible set is $\mathcal{W} := C^+$. The *lower image* of (D) is defined as

$$\mathcal{D} := \xi(\mathcal{W}) - K = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid w \in \mathcal{W}, \alpha \leq p^w\}.$$

Remark 3.3. *Note that the decision space of the dual problem has dimension q , which in general is much less than n , the number of variables of the primal problem. However, the dual objective function involves solving another optimization problem. Indeed, if the feasible region \mathcal{X} of (P) is given by explicit constraints, then it is also possible to define a geometric dual problem which includes additional dual variables corresponding to these explicit constraints. In particular, the last component of the dual objective function can be defined using the Lagrangian of (WS(w)) instead of its value directly. This construction would lead to the same lower image, see also [37, Remark 3.6].*

The following proposition follows from the definition of \mathcal{D} , we omit its proof.

Proposition 3.4. *The lower image \mathcal{D} is a closed convex cone.*

We define exact and approximate solution concepts for the primal problem (P).

Definition 3.5. [35, Definition 2.20, Proposition 4.7] *A point $\bar{x} \in \mathcal{X}$ is said to be a (weak) minimizer for (P) if $f(\bar{x})$ is a (weakly) C -minimal element of $f(\mathcal{X})$. A nonempty set $\bar{\mathcal{X}} \subseteq \mathcal{X}$ is called an infimizer of (P) if $\text{cl conv}(f(\bar{\mathcal{X}}) + C) = \mathcal{P}$. An infimizer $\bar{\mathcal{X}}$ of (P) is called (weak) solution to (P) if it consists of only (weak) minimizers.*

Since the upper image of a convex vector optimization problem is not a polyhedral set in general, a finite set $\bar{\mathcal{X}}$ may not satisfy the exact solution concept in Definition 3.5. Hence, we give an approximate solution concept for a fixed $\epsilon > 0$ below.

Definition 3.6. [1, Definition 3.5] *A nonempty finite set $\bar{\mathcal{X}} \subseteq \mathcal{X}$ is called a finite ϵ -infimizer of (P) if $\text{conv } f(\bar{\mathcal{X}}) + C + B(0, \epsilon) \supseteq \mathcal{P}$. A finite ϵ -infimizer $\bar{\mathcal{X}}$ of (P) is called a finite (weak) ϵ -solution to (P) if it consists of only (weak) minimizers.*

Note that if $\bar{\mathcal{X}}$ is a finite (weak) ϵ -solution, then we have the following inner and outer approximations of the upper image:

$$\text{conv } f(\bar{\mathcal{X}}) + C + B(0, \epsilon) \supseteq \mathcal{P} \supseteq \text{conv } f(\bar{\mathcal{X}}) + C.$$

Now, similar to Definition 3.5, we provide an exact solution concept for the dual problem (D).

Definition 3.7. [35, Definition 2.53, Corollary 2.54] A point $\bar{w} \in \mathcal{W}$ is called a maximizer for (D) if $\xi(\bar{w})$ is a K -maximal element of $\xi(\mathcal{W})$. A nonempty set $\bar{\mathcal{W}} \subseteq \mathcal{W}$ is called a supremizer of (D) if $\text{cone conv } \xi(\bar{\mathcal{W}}) - K = \mathcal{D}$. A supremizer $\bar{\mathcal{W}}$ of (D) is called a solution to (D) if it consists of only maximizers.

As for the upper image, in general, the lower image cannot be represented by a finite set $\bar{\mathcal{W}}$ using the exact solution concept in Definition 3.7. In the next definition, we propose a novel approximate solution that is tailor-made for the lower image \mathcal{D} , which is a convex cone; see Remark 3.9 below for the technical motivation.

Definition 3.8. A nonempty finite set $\bar{\mathcal{W}} \subseteq \mathcal{W} \cap \mathbb{S}^{q-1}$ is called a finite ϵ -supremizer of (D) if $\text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D}$, where $\mathbb{S}^{q-1} := \{z \in \mathbb{R}^q \mid \|z\|_* = 1\}$ denotes the unit sphere in \mathbb{R}^q with respect to the dual norm. A finite ϵ -supremizer $\bar{\mathcal{W}}$ of (D) is called a finite ϵ -solution to (D) if it consists of only maximizers.

If $\bar{\mathcal{W}}$ is a finite ϵ -solution of (D), then one obtains the following inner and outer approximations of the lower image:

$$\text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D} \supseteq \text{cone conv } \xi(\bar{\mathcal{W}}) - K.$$

Remark 3.9. Let us comment on the particular structure of Definition 3.8. Since the lower image \mathcal{D} is a convex cone by Proposition 3.4, we evaluate the conic hull of the Minkowski sum $\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}$ so that the error values are scaled properly. With this operation, we ensure that the resulting conic hull is comparable with \mathcal{D} (up to the subtraction of the ordering cone K).

The next proposition will be used later to prove some geometric duality results.

Proposition 3.10. Let $w \in \mathcal{W} \setminus \{0\}$. Then, $\xi(w)$ is a K -maximal element of \mathcal{D} .

Proof. Let $\epsilon > 0$. We prove that $\xi(w) + \epsilon e^{q+1} \notin \mathcal{D} = \xi(\mathcal{W}) - K$. To get a contradiction, assume the existence of $\bar{w} \in \mathcal{W}$ and $\bar{\epsilon} \geq 0$ with $\xi(w) + \epsilon e^{q+1} = \xi(\bar{w}) - \bar{\epsilon} e^{q+1}$. By (1), we have $(w_1, \dots, w_q, p^w)^\top + \epsilon e^{q+1} = (\bar{w}_1, \dots, \bar{w}_q, p^{\bar{w}})^\top - \bar{\epsilon} e^{q+1}$, that is, $(w_1, \dots, w_q, p^w + \epsilon)^\top = (\bar{w}_1, \dots, \bar{w}_q, p^{\bar{w}} - \bar{\epsilon})^\top$. Hence, $\inf_{x \in \mathcal{X}} w^\top f(x) = \inf_{x \in \mathcal{X}} \bar{w}^\top f(x)$ and $\epsilon = -\bar{\epsilon}$. This contradicts $\epsilon > 0$ and $\bar{\epsilon} \geq 0$. Hence, $\xi(w) + \epsilon e^{q+1} \notin \mathcal{D}$. Since $\epsilon > 0$ is arbitrary, $\xi(w)$ is a K -maximal element of \mathcal{D} . \square

4 Geometric duality

In this section, we will investigate the duality relation between the primal and dual problems, (P) and (D). First, we will provide the main duality theorem which relates the weakly C -minimal exposed faces of \mathcal{P} and the K -maximal exposed faces of \mathcal{D} . Then, we will establish some duality properties regarding \mathcal{P} and \mathcal{D} , as well as their polyhedral approximations.

4.1 Geometric duality between \mathcal{P} and \mathcal{D}

We will prove that there is a one-to-one correspondence between the set of all weakly C -minimal exposed faces of the upper image \mathcal{P} and the set of all K -maximal exposed faces of the lower image \mathcal{D} . We will also show that the upper and lower images can be recovered from each other.

Let us start by defining

$$\varphi: \mathbb{R}^q \times \mathbb{R}^{q+1} \rightarrow \mathbb{R}, \quad \varphi(y, w, \alpha) := w^\top y - \alpha, \quad (2)$$

and the following set-valued maps:

$$\begin{aligned} \mathcal{H}, H: \mathbb{R}^{q+1} \rightrightarrows \mathbb{R}^q, \mathcal{H}(w, \alpha) &:= \{y \in \mathbb{R}^q \mid \varphi(y, w, \alpha) \geq 0\}, H(w, \alpha) := \text{bd } \mathcal{H}(w, \alpha), \\ \mathcal{H}^*, H^*: \mathbb{R}^q \rightrightarrows \mathbb{R}^{q+1}, \mathcal{H}^*(y) &:= \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \varphi(y, w, \alpha) \geq 0\}, H^*(y) := \text{bd } \mathcal{H}^*(y). \end{aligned}$$

These functions are essential to define a duality map between \mathcal{P} and \mathcal{D} . Moreover, for arbitrary $y \in \mathbb{R}^q$ and $(w^\top, \alpha)^\top \in \mathbb{R}^{q+1}$, we have the following statement:

$$(w^\top, \alpha)^\top \in H^*(y) \iff y \in H(w, \alpha). \quad (3)$$

With the next proposition we show that, by solving a weighted sum scalarization, one finds supporting hyperplanes to the upper and lower images.

Proposition 4.1. *Let $w \in \mathcal{W} \setminus \{0\}$ and x^w be an optimal solution to $(\text{WS}(w))$. Then, $H(\xi(w))$ is a supporting hyperplane of \mathcal{P} at $f(x^w)$ such that $\mathcal{P} \subseteq \mathcal{H}(\xi(w))$, and $H^*(f(x^w))$ is a supporting hyperplane of \mathcal{D} at $\xi(w)$ such that $\mathcal{D} \subseteq \mathcal{H}^*(f(x^w))$.*

Proof. To prove the first statement, we show that $f(x^w) \in H(\xi(w))$ and $\mathcal{P} \subseteq \mathcal{H}(\xi(w))$. As $\xi(w) = (w^\top, w^\top f(x^w))^\top$, we have $\varphi(f(x^w), \xi(w)) = 0$, which implies that $f(x^w) \in H(\xi(w))$. To prove $\mathcal{P} \subseteq \mathcal{H}(\xi(w))$, let $x \in \mathcal{X}$, $c \in C$ be arbitrary. Then, $\varphi(f(x) + c, \xi(w)) = w^\top f(x) + w^\top c - w^\top f(x^w) \geq 0$ as $w \in C^+$ and x^w is an optimal solution of $(\text{WS}(w))$. Therefore, $\mathcal{H}(\xi(w)) \supseteq f(\mathcal{X}) + C = \mathcal{P}$, where the equality follows by Proposition 3.1. On the other hand, having $\varphi(f(x^w), \xi(w)) = w^\top f(x^w) - w^\top f(x^w) = 0$ also implies that $\xi(w) \in H^*(f(x^w))$. To complete the proof, it is enough to show that $\mathcal{D} \subseteq \mathcal{H}^*(f(x^w))$. Let $(w^\top, \alpha)^\top \in \mathcal{D}$ be arbitrary. It follows from the definition of \mathcal{D} that $\alpha \leq \inf_{x \in \mathcal{X}} w^\top f(x) = w^\top f(x^w)$. Hence, $\varphi(f(x^w), w, \alpha) = w^\top f(x^w) - \alpha \geq 0$, that is, $(w^\top, \alpha)^\top \in \mathcal{H}^*(f(x^w))$. \square

The following propositions show the relationship between the weakly C -minimal elements of \mathcal{P} and the K -maximal exposed faces of \mathcal{D} , as well as that between the K -maximal elements of \mathcal{D} and the weakly C -minimal exposed faces of \mathcal{P} . Their proofs are based on elementary arguments, hence omitted for brevity.

Proposition 4.2. (a) *Let $y \in \mathbb{R}^q$. Then, y is a weakly C -minimal element of \mathcal{P} if and only if $H^*(y) \cap \mathcal{D}$ is a K -maximal exposed face of \mathcal{D} . (b) For every K -maximal exposed face F^* of \mathcal{D} , there exists some $y \in \mathcal{P}$ such that $F^* = H^*(y) \cap \mathcal{D}$.*

Proposition 4.3. (a) *Let $(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \setminus \{0\}$. Then, $(w^\top, \alpha)^\top$ is a K -maximal element of \mathcal{D} if and only if $H(w, \alpha) \cap \mathcal{P}$ is a weakly C -minimal exposed face of \mathcal{P} satisfying $\mathcal{H}(w, \alpha) \supseteq \mathcal{P}$. (b) For every C -minimal exposed face F of \mathcal{P} , there exists some $(w^\top, \alpha)^\top \in \mathcal{D}$ such that $F = H(w, \alpha) \cap \mathcal{P}$.*

We now proceed with the main geometric duality result. Let $\mathcal{F}_{\mathcal{P}}$ be the set of all weakly C -minimal exposed faces of \mathcal{P} and $\mathcal{F}_{\mathcal{D}}^*$ be the set of all K -maximal exposed faces of \mathcal{D} . Consider the set-valued function

$$\Psi: \mathcal{F}_{\mathcal{D}}^* \rightrightarrows \mathbb{R}^q, \quad \Psi(F^*) := \bigcap_{(w^\top, \alpha)^\top \in F^*} H(w, \alpha) \cap \mathcal{P}. \quad (4)$$

Theorem 4.4. Ψ is an inclusion-reversing one-to-one correspondence between $\mathcal{F}_{\mathcal{D}}^* \setminus \{(0^\top, 0)^\top\}$ and $\mathcal{F}_{\mathcal{P}}$. The inverse map is given by $\Psi^*(F) := \bigcap_{y \in F} H^*(y) \cap \mathcal{D}$.

Proof. We use the geometric duality theory for the epigraphs of closed convex functions developed in [30]. To be able to use this theory, we express \mathcal{P} and \mathcal{D} as the epigraphs of closed convex functions, up to transformations. Observe that

$$\mathcal{D} = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid p^w \geq \alpha, w \in C^+\} = -\{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \tilde{p}(w) \leq \alpha\} = -\text{epi } \tilde{p},$$

where $\tilde{p}: \mathbb{R}^q \rightarrow \mathbb{R} \cup \{+\infty\}$ is the support function of \mathcal{P} , i.e., $\tilde{p}(w) = \sup_{y \in \mathcal{P}} w^\top y$ for every $w \in \mathbb{R}^q$. Then, the well-known duality between support and indicator functions ([46, Theorem 13.2]) yields that $\tilde{p} = g^*$, where $g := I_{\mathcal{P}}$ is the indicator function of \mathcal{P} . We also have $\text{epi } g = \mathcal{P} \times \mathbb{R}_+$ and $\mathcal{P} = P(\text{epi } g)$, where $P: 2^{\mathbb{R}^{q+1}} \rightrightarrows \mathbb{R}^q$ is defined by $P(F^*) = \{y \in \mathbb{R}^q \mid (y^\top, 0)^\top \in F^*\}$ for $F^* \subseteq \mathbb{R}^{q+1}$. Let us define a set-valued function $\tilde{\Psi}: 2^{\mathbb{R}^{q+1}} \rightrightarrows \mathbb{R}^{q+1}$ by

$$\tilde{\Psi}(F^*) := \bigcap_{(w^\top, g^*(w))^\top \in F^*} \{(y^\top, g(y))^\top \in \mathbb{R}^q \times \mathbb{R} \mid w \in \partial g(y)\}.$$

Then, by [30, Theorem 3.3], $\tilde{\Psi}$ is an inclusion-reversing one-to-one correspondence between the set of all K -minimal exposed faces of $\text{epi } g^* = -\mathcal{D}$ and the set of all K -minimal exposed faces of $\text{epi } g = \mathcal{P} \times \mathbb{R}_+$. Clearly, a K -minimal exposed face of $-\mathcal{D}$ is of the form $-F^*$, where $F^* \in \mathcal{F}_{\mathcal{D}}^*$; the converse holds as well. The one-to-one correspondence here is inclusion-preserving.

Next, we show that $P(\tilde{F}) \in \mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$ whenever \tilde{F} is a K -minimal exposed face of $\text{epi } g$. Note that a point $(y^\top, r)^\top \in \mathbb{R}^{q+1}$ is a K -minimal element of $\text{epi } g = \mathcal{P} \times \mathbb{R}_+$ if and only if $y \in \mathcal{P}$ and $r = 0$. Let \tilde{F} be a K -minimal exposed face of $\text{epi } g$. The previous observation implies that $\tilde{F} \subseteq \mathcal{P} \times \{0\}$. By definition, we may write $\tilde{F} = (\mathcal{P} \times \mathbb{R}_+) \cap \tilde{H}$ for some supporting hyperplane $\tilde{H} \subseteq \mathbb{R}^{q+1}$ of $\mathcal{P} \times \mathbb{R}_+$. Then, $P(\tilde{F}) = P(\mathcal{P} \times \mathbb{R}_+) \cap P(\tilde{H}) = \mathcal{P} \cap P(\tilde{H})$. By [30, Lemma 3.1(ii)], there exist $w \in \mathbb{R}^q$ and $\alpha \in \mathbb{R}$ such that $\tilde{H} = \{(y^\top, r) \in \mathbb{R}^{q+1} \mid w^\top y - r = \alpha\}$. If $w = 0$, then $\tilde{F} \subseteq \mathcal{P} \times \{0\}$ implies that $\alpha = 0$ and we have $\tilde{H} = \mathbb{R}^q \times \{0\}$ and $P(\tilde{F}) = \mathcal{P}$. Suppose that $w \neq 0$. Then, $P(\tilde{H}) \neq \mathbb{R}^q$ is a supporting hyperplane of $P(\tilde{F})$ since \tilde{H} is a supporting hyperplane of \tilde{F} . It follows that $P(\tilde{F}) = \mathcal{P} \cap P(\tilde{H}) \in \mathcal{F}_{\mathcal{P}}$.

Conversely, we define a function G on $2^{\mathbb{R}^q}$ which maps $\mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$ into the set of all K -minimal exposed faces of $\text{epi } g$. Let $G(F) := F \times \{0\}$ for every $F \subseteq \mathbb{R}^q$. By the definitions of P and G , we obtain $P(G(F)) = F$ for every $F \subseteq \mathbb{R}^q$ and $G(P(\tilde{F})) = \tilde{F}$ for every $\tilde{F} \subseteq \mathbb{R}^q \times \{0\}$. In particular, $P(G(F)) = F$ for every $F \in \mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$ and $G(P(\tilde{F})) = \tilde{F}$ for every K -minimal exposed face \tilde{F} of $\text{epi } g$. It remains to show that $G(F)$ is a K -minimal exposed face of $\text{epi } g$ for all $F \in \mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$. First, note that $G(\mathcal{P}) = \mathcal{P} \times \{0\} = \text{epi } g \cap (\mathbb{R}^q \times \{0\})$ is a K -minimal exposed face of $\text{epi } g$. Next, suppose that $F \in \mathcal{F}_{\mathcal{P}}$. Then, by Proposition 4.3(b), there

exists at least one $(w^\top, \alpha)^\top \in \mathcal{D}$ such that $F = H(w, \alpha) \cap \mathcal{P}$. Moreover, since $H(w, \alpha)$ is a supporting hyperplane of \mathcal{P} , it follows that $\tilde{H}(w, \alpha) := \{(y^\top, r)^\top \in \mathbb{R}^{q+1} \mid w^\top y - r = \alpha\}$ is a supporting hyperplane of $\text{epi } g$ and $\tilde{F} := \tilde{H}(w, \alpha) \cap \text{epi } g$ is a K -minimal exposed face of $\text{epi } g$ by [30, Lemma 3.1(ii)]. Note that $P(\tilde{F}) = P(\tilde{H}(w, \alpha)) \cap P(\text{epi } g) = H(w, \alpha) \cap \mathcal{P} = F$. Hence, $G(F) = G(P(\tilde{F})) = \tilde{F}$ is a K -minimal exposed face of $\text{epi } g$.

Let $F^* \in \mathcal{F}_{\mathcal{D}}^*$. Then, $-F^*$ is a K -minimal exposed face of $\text{epi } g^*$, $\tilde{\Psi}(-F^*)$ is a K -minimal exposed face of $\text{epi } g$, and $P(\tilde{\Psi}(-F^*)) \in \mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$. Moreover, we have

$$\begin{aligned} P(\tilde{\Psi}(-F^*)) &= \bigcap_{(w^\top, g^*(w))^\top \in -F^*} P(\{(y^\top, g(y))^\top \in \mathbb{R}^{q+1} \mid w \in \partial g(y)\}) \\ &= \bigcap_{(-w^\top, -\tilde{p}(w))^\top \in F^*} P(\{(y^\top, 0)^\top \in \mathbb{R}^{q+1} \mid y \in \mathcal{P}, w \in \mathcal{N}_{\mathcal{P}}(y)\}) \\ &= \bigcap_{(w^\top, p^w)^\top \in F^*} \{y \in \mathcal{P} \mid w \in -\mathcal{N}_{\mathcal{P}}(y)\} \\ &= \bigcap_{(w^\top, \alpha)^\top \in F^*} H(w, \alpha) \cap \mathcal{P} = \Psi(F^*). \end{aligned}$$

By [30, Theorem 3.3], the inverse of $\tilde{\Psi}$ is given by

$$\tilde{\Psi}^{-1}(\tilde{F}) = \bigcap_{(y^\top, g(y))^\top \in \tilde{F}} \{(w^\top, g^*(w))^\top \in \mathbb{R}^q \times \mathbb{R} \mid w \in \partial g(y)\}.$$

Let $F \in \mathcal{F}_{\mathcal{P}} \cup \{\mathcal{P}\}$. Then, $-\tilde{\Psi}^{-1}(G(F)) \in \mathcal{F}_{\mathcal{D}}^*$ by the above constructions. Moreover, we have

$$\begin{aligned} -\tilde{\Psi}^{-1}(G(F)) &= \bigcap_{(y^\top, g(y))^\top \in G(F)} \{(-w^\top, -g^*(w))^\top \in \mathbb{R}^q \times \mathbb{R} \mid w \in \partial g(y)\} \\ &= \bigcap_{(y^\top, 0)^\top \in F \times \{0\}} \{(-w^\top, -\tilde{p}(w))^\top \in \mathbb{R}^q \times \mathbb{R} \mid w \in \mathcal{N}_{\mathcal{P}}(y)\} \\ &= \bigcap_{y \in F} \{(w^\top, p^w)^\top \in \mathbb{R}^q \times \mathbb{R} \mid w \in -\mathcal{N}_{\mathcal{P}}(y)\} \\ &= \bigcap_{y \in F} \{(w^\top, p^w)^\top \in \mathbb{R}^q \times \mathbb{R} \mid y \in H(w, p^w)\} \\ &= \bigcap_{y \in F} \{(w^\top, p^w)^\top \in \mathbb{R}^q \times \mathbb{R} \mid (w^\top, p^w)^\top \in H^*(y)\} \\ &= \bigcap_{y \in F} H^*(y) \cap \mathcal{D} = \Psi^*(F). \end{aligned} \tag{5}$$

Finally, note that $\Psi(\{(0^\top, 0)^\top\}) = \mathcal{P}$. Then, by combining the three one-to-one correspondences established above and excluding the pair formed by $\{(0^\top, 0)^\top\}$ and \mathcal{P} , we conclude that Ψ is an inclusion-reversing one-to-one correspondence between $\mathcal{F}_{\mathcal{D}}^* \setminus \{(0^\top, 0)^\top\}$ and $\mathcal{F}_{\mathcal{P}}$, and Ψ^* is its inverse mapping. \square

Remark 4.5. Using the notions of second-order subdifferential and indicatrix for convex functions, one can obtain the following polarity relationship between the curvatures of \mathcal{P} and \mathcal{D} , which is in the spirit of [30, Theorem 5.7]. Suppose that the function $g = I_{\mathcal{P}}$ (see the proof of Theorem 4.4) is twice epi-differentiable ([47, Definition 2.2]). Then, for every K -maximal exposed face F^* of \mathcal{D} , $(w^\top, p^w)^\top \in F^*$, $y \in \Psi(F^*)$, we have

$$\text{Ind}_{\mathcal{D}}((w^\top, p^w) \mid (y^\top, 1)) := \text{Ind } g^*(-w \mid y) = (\text{Ind } g(y \mid -w))^\circ =: (\text{Ind}_{\mathcal{P}}(y \mid -w))^\circ.$$

Here, $\text{Ind } g, \text{Ind } g^*$ are defined as the polars of the corresponding second-order subdifferentials of g, g^* , respectively; see [30, Section 4.1], [43, Proposition 4.1], [52, Section 4]. Moreover, we define $\text{Ind}_{\mathcal{P}}, \text{Ind}_{\mathcal{D}}$ as indicatrices of \mathcal{P}, \mathcal{D} (with suitable dimensions) by using $\text{Ind } g, \text{Ind } g^*$, similar to the construction in [30, Section 5]. The result is a direct consequence of [52, Lemma 4.6(b)], where it is needed to work with (w, y) such that $-w \in \partial g(y)$. In our case, this condition is verified thanks to the structure of $\Psi(F^*)$ in (??) whenever $(w^\top, p^w)^\top \in F^*$, $y \in \Psi(F^*)$. We verify the twice epi-differentiability of g for the problem that will be considered in Example 7.1. To that end, let us take $f(x) = A^\top x$ and $\mathcal{X} = \{x \in \mathbb{R}^n \mid x^\top P x - 1 \leq 0\}$, where $P \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix, and $A \in \mathbb{R}_+^{n \times q}$. By [52, Lemma 4.6(a)], g is twice epi-differentiable at y relative to $-w$ if and only if g^* is so at $-w$ relative to y . To check the latter, note that, for $w \in C^+$, we have

$$\begin{aligned} g^*(-w) &= \sup_{z \in \mathcal{P}} -w^\top z = \sup_{x \in \mathcal{X}} -w^\top f(x) = \sup\{-w^\top A^\top x \mid x^\top P x \leq 1\} \\ &= \sqrt{(Aw)^\top P^{-1} Aw} = \sqrt{w^\top (A^\top P^{-1} A) w}, \end{aligned}$$

which follows from standard calculations for finding the support function of an ellipsoid. It follows that $(g^*)^2$ on \mathbb{R}^q is a piecewise linear-quadratic function in the sense of [47, Definition 1.1], which is twice epi-differentiable by [47, Theorem 3.1]. Hence, g^* is also twice epi-differentiable.

Now, we will see that the upper and dual images \mathcal{P} and \mathcal{D} can be recovered from each other using the function φ introduced in (2). The following definition will be used to simplify the notation in later steps.

Definition 4.6. For closed and convex sets $\bar{\mathcal{P}}$ and $\bar{\mathcal{D}}$, we define

$$\begin{aligned} \mathcal{D}_{\bar{\mathcal{P}}} &:= \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in \bar{\mathcal{P}}: \varphi(y, w, \alpha) \geq 0\}, \\ \mathcal{P}_{\bar{\mathcal{D}}} &:= \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \bar{\mathcal{D}}: \varphi(y, w, \alpha) \geq 0\}. \end{aligned}$$

Remark 4.7. In view of (2), the sets in Definition 4.6 can be rewritten as

$$\begin{aligned} \mathcal{D}_{\bar{\mathcal{P}}} &= \{z \in \mathbb{R}^{q+1} \mid \forall y \in \bar{\mathcal{P}}: (y^\top, -1)z \geq 0\} = (\bar{\mathcal{P}} \times \{-1\})^+, \\ \mathcal{P}_{\bar{\mathcal{D}}} &= \{y \in \mathbb{R}^q \mid \forall z \in \bar{\mathcal{D}}: (y^\top, -1)z \geq 0\} = \{y \in \mathbb{R}^q \mid (y^\top, -1)^\top \in \bar{\mathcal{D}}^+\}. \end{aligned}$$

Proposition 4.8. It holds (a) $\mathcal{D}_{\mathcal{P}} = \mathcal{D}$, (b) $\mathcal{P}_{\mathcal{D}} = \mathcal{P}$.

Proof. (a) follows directly from the definitions, Remark 2.1 and Proposition 3.1. To prove (b), let $y \in \mathcal{P}$ be arbitrary. By Proposition 3.1, $y = f(x) + c \in \mathcal{P}$ for some $x \in \mathcal{X}, c \in C$. For every $(w^\top, \alpha)^\top \in \mathcal{D}$, we have $\varphi(f(x) + c, w, \alpha) \geq 0$, which follows from $w \in C^+$ and $(w^\top, \alpha)^\top \in \mathcal{D}$. This implies $f(x) + c \in \mathcal{P}_{\mathcal{D}}$, hence, $\mathcal{P}_{\mathcal{D}} \supseteq \mathcal{P}$. For the reverse inclusion, let $y \in \mathcal{P}_{\mathcal{D}}$. By Remark 4.7 and the definition of \mathcal{D} , we have $(y^\top, -1)^\top \in \mathcal{D}^+ = (\xi(\mathcal{W}) - K)^+$. Hence, for every $w \in \mathcal{W}$ and $\alpha \leq p^w$, we have $(y^\top, -1)(w^\top, \alpha)^\top = w^\top y - \alpha \geq 0$. In particular, by choosing $w \in C^+ \setminus \{0\}$ and setting $\alpha = p^w$, we get $w^\top y \geq p^w$ for every $w \in C^+ \setminus \{0\}$. Note that $p^w = \inf_{x \in \mathcal{X}} w^\top f(x) = \inf_{y \in \mathcal{P}} w^\top y$ for every $w \in C^+ \setminus \{0\}$. Since \mathcal{P} is a nonempty closed convex set such that $\mathcal{P} = \mathcal{P} + C$, having $w^\top y \geq p^w$ for every $w \in C^+ \setminus \{0\}$ is equivalent to $y \in \mathcal{P}$. Hence, $\mathcal{P}_{\mathcal{D}} \subseteq \mathcal{P}$. \square

4.2 Geometric duality between the approximations of \mathcal{P} and \mathcal{D}

In Theorem 4.4, we have seen that the upper and lower images can be recovered from each other. In this section, we show that similar relations hold for polyhedral approximations of these sets. Throughout, we call a set $A \subseteq \mathbb{R}^q$ ($A \subseteq \mathbb{R}^{q+1}$) an *upper set* (a *lower set*) if $A = A + C$ ($A = A - K$). We start by showing that a closed convex upper set can be recovered using the transformations introduced in Definition 4.6.

Proposition 4.9. *Let $\emptyset \neq \bar{\mathcal{P}} \subsetneq \mathbb{R}^q$ be a closed convex set. Then, $\bar{\mathcal{P}} = \mathcal{P}_{\mathcal{D}_{\bar{\mathcal{P}}}}$.*

Proof. Let $y \in \mathbb{R}^q$. Remark 4.7 implies that $y \in \mathcal{P}_{\mathcal{D}_{\bar{\mathcal{P}}}}$ is equivalent to $(y^\top, -1)^\top \in (\mathcal{D}_{\bar{\mathcal{P}}})^+ = (\bar{\mathcal{P}} \times \{-1\})^{++}$. By the convexity of $\bar{\mathcal{P}}$ and [46, Theorem 8.2], we have

$$(\bar{\mathcal{P}} \times \{-1\})^{++} = \text{cl cone}(\bar{\mathcal{P}} \times \{-1\}) = \text{cone}(\bar{\mathcal{P}} \times \{-1\}) \cup (\text{recc } \bar{\mathcal{P}} \times \{0\}).$$

Hence, $(y^\top, -1)^\top \in (\bar{\mathcal{P}} \times \{-1\})^{++}$ is equivalent to $(y^\top, -1)^\top \in \text{cone}(\bar{\mathcal{P}} \times \{-1\})$, which is equivalent to $y \in \bar{\mathcal{P}}$. Therefore, $\mathcal{P}_{\mathcal{D}_{\bar{\mathcal{P}}}} = \bar{\mathcal{P}}$. \square

Next, for a closed convex lower set $\bar{\mathcal{D}}$, we want to investigate the relationship between $\bar{\mathcal{D}}$ and $\mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}}$. While the equality of these sets may not hold in general, the next proposition shows that it holds if $\bar{\mathcal{D}}$ is a cone.

Proposition 4.10. *Let $\emptyset \neq \bar{\mathcal{D}} \subseteq \mathbb{R}^{q+1}$ be a closed convex lower set. Suppose further that $\bar{\mathcal{D}}$ is a cone and $\mathcal{P}_{\bar{\mathcal{D}}} \neq \emptyset$. Then, $\mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}} = \bar{\mathcal{D}}$.*

Proof. By Remark 4.7, we have

$$\begin{aligned} \mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}} &= (\mathcal{P}_{\bar{\mathcal{D}}} \times \{-1\})^+ = (\{y \in \mathbb{R}^q \mid (y^\top, -1)^\top \in \bar{\mathcal{D}}^+\} \times \{-1\})^+ \\ &= (\bar{\mathcal{D}}^+ \cap (\mathbb{R}^q \times \{-1\}))^+. \end{aligned}$$

From this observation, it follows that $\mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}} \supseteq \bar{\mathcal{D}}^{++} = \bar{\mathcal{D}}$ since $\bar{\mathcal{D}}$ is a nonempty closed convex cone. To prove the reverse inclusion, let $(w^\top, \alpha)^\top \in \mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}}$. Then, by the above observation, we have

$$(y^\top, -1)(w^\top, \alpha)^\top = w^\top y - \alpha \geq 0 \tag{6}$$

for every $y \in \mathbb{R}^q$ such that $(y^\top, -1)^\top \in \bar{\mathcal{D}}^+$. We show that $(w^\top, \alpha)^\top \in \bar{\mathcal{D}} = \bar{\mathcal{D}}^{++}$. To that end, let $(y^\top, \beta)^\top \in \bar{\mathcal{D}}^+$. We claim that $(y^\top, \beta)(w^\top, \alpha)^\top = w^\top y + \beta\alpha \geq 0$. First, note that

$\beta \leq 0$ since $(y^\top, \beta)^\top \in \bar{\mathcal{D}}^+$ and $\bar{\mathcal{D}}$ is a lower set. We consider the following cases.

Case 1: Suppose that $\beta < 0$. Then, we have $(-\frac{1}{\beta}y^\top, -1)^\top \in \bar{\mathcal{D}}^+$ since $\bar{\mathcal{D}}^+$ is a cone. In particular, applying (6) for this point yields $-\frac{1}{\beta}w^\top y - \alpha \geq 0$, that is, $w^\top y + \beta\alpha \geq 0$. Hence, the claim follows.

Case 2: Suppose that $\beta = 0$. Since $\bar{\mathcal{D}}^+$ is a convex cone, $(y^\top, 0)^\top$ is a recession direction of $\bar{\mathcal{D}}^+$. In particular, for every $\bar{y} \in \mathcal{P}_{\bar{\mathcal{D}}} \neq \emptyset$, we have $(\bar{y}^\top, -1)^\top \in \bar{\mathcal{D}}^+$ by Remark 4.7, which implies that $(\bar{y}^\top, -1)^\top + t(y^\top, 0)^\top = ((\bar{y} + ty)^\top, -1)^\top \in \bar{\mathcal{D}}^+$ for every $t \in \mathbb{R}_+$. Hence, by Remark 4.7 again, $\bar{y} + ty \in \mathcal{P}_{\bar{\mathcal{D}}}$ for every $\bar{y} \in \mathcal{P}_{\bar{\mathcal{D}}}$ and $t \in \mathbb{R}_+$, that is, $y \in \text{recc } \mathcal{P}_{\bar{\mathcal{D}}}$. Furthermore, the proof of [46, Theorem 8.2] states that $\text{recc } \mathcal{P}_{\bar{\mathcal{D}}} \times \{0\}$ consists of the limits of all sequences of the form $(\lambda_n((y^n)^\top, -1)^\top)_{n \in \mathbb{N}}$, where $(\lambda_n)_{n \in \mathbb{N}}$ is a decreasing sequence in \mathbb{R}_+ whose limit is 0, and $((y^n)^\top)_{n \in \mathbb{N}}$ is a sequence in $\mathcal{P}_{\bar{\mathcal{D}}}$. Since $(y^\top, 0)^\top \in \text{recc } \mathcal{P}_{\bar{\mathcal{D}}} \times \{0\}$, such $(\lambda_n)_{n \in \mathbb{N}}$ and $((y^n)^\top)_{n \in \mathbb{N}}$ exist with the additional property that $\lim_{n \rightarrow \infty} \lambda_n y^n = y$. For each $n \in \mathbb{N}$, having $y^n \in \mathcal{P}_{\bar{\mathcal{D}}}$ implies that $((y^n)^\top, -1)^\top \in \bar{\mathcal{D}}^+$ by Remark 4.7, and hence $w^\top y^n - \alpha \geq 0$ by (6). Then, $w^\top \lambda_n y^n - \lambda_n \alpha \geq 0$ for each $n \in \mathbb{N}$. Letting $n \rightarrow \infty$ yields that $w^\top y = w^\top y + \beta\alpha \geq 0$. Hence, the claim follows.

Therefore, $(w^\top, \alpha)^\top \in \bar{\mathcal{D}}^{++} = \bar{\mathcal{D}}$, which completes the proof of $\mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}} \subseteq \bar{\mathcal{D}}$. \square

The next lemma shows that, when computing $\mathcal{P}_{\bar{\mathcal{D}}}$, considering the extreme directions of $\bar{\mathcal{D}}$ is sufficient under a special structure for $\bar{\mathcal{D}}$. Its proof is quite routine, we omit it.

Lemma 4.11. *Let $\bar{\mathcal{D}} = \text{cone conv } \xi(\bar{\mathcal{W}}) - K$ for some $\bar{\mathcal{W}} \subseteq \mathcal{W}$. Then, $\mathcal{P}_{\bar{\mathcal{D}}} = \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \xi(\bar{\mathcal{W}}): \varphi(y, w, \alpha) \geq 0\}$.*

The next corollary provides two special cases of Proposition 4.10 which will be used later.

Corollary 4.12. *Let $\bar{\mathcal{W}} \subseteq \mathcal{W}$ be a nonempty finite set. Then, $\bar{\mathcal{D}} = \mathcal{D}_{\mathcal{P}_{\bar{\mathcal{D}}}}$ in each of the following cases.*

(a) $\bar{\mathcal{D}} = \text{cone conv } \xi(\bar{\mathcal{W}}) - K$. (b) $0 \notin \bar{\mathcal{W}}$ and $\bar{\mathcal{D}} = \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$.

Proof. Since $\bar{\mathcal{W}}$ is finite, $\bar{\mathcal{D}}$ is a closed convex lower set that is also a cone in each case. Next, we show that $\mathcal{P}_{\bar{\mathcal{D}}} \neq \emptyset$.

(a) Let $\bar{x} \in \mathcal{X}$. We have $w^\top f(\bar{x}) \geq \inf_{x \in \mathcal{X}} w^\top f(x) = p^w$ for every $w \in \mathcal{W}$, in particular, for every $w \in \bar{\mathcal{W}}$. Hence, $f(\bar{x}) \in \mathcal{P}_{\bar{\mathcal{D}}}$ by Lemma 4.11.

(b) By the definition of $\mathcal{P}_{\bar{\mathcal{D}}}$ and simple algebraic manipulations, we have

$$\begin{aligned} \mathcal{P}_{\bar{\mathcal{D}}} &= \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \bar{\mathcal{D}}: w^\top y \geq \alpha\} \\ &= \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}: w^\top y \geq \alpha\} \\ &= \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \text{conv } \xi(\bar{\mathcal{W}}): w^\top y \geq \alpha + \epsilon\} \\ &= \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \xi(\bar{\mathcal{W}}): w^\top y \geq \alpha + \epsilon\}. \end{aligned}$$

Since $0 \notin \bar{\mathcal{W}}$ and $\bar{\mathcal{W}}$ is finite, there exists $\bar{c} \in \text{int } C$ such that $w^\top \bar{c} \geq \epsilon$ for every $w \in \bar{\mathcal{W}}$. Let $\bar{x} \in \mathcal{X}$. Then, $w^\top (f(\bar{x}) + \bar{c}) \geq p^w + \epsilon$ for every $w \in \bar{\mathcal{W}}$. Hence, $f(\bar{x}) + \bar{c} \in \mathcal{P}_{\bar{\mathcal{D}}}$. \square

5 Algorithms

In this section, we will present two approximation algorithms, namely the primal and dual algorithms, for solving the primal and dual problems, (P) and (D), simultaneously. First, we will explain the primal algorithm, which is proposed in [1] for solving (P) only, and show that by simple modifications, this algorithm also yields a solution to the dual problem (D). Next, we will describe the dual algorithm which uses the geometric duality results from Section 4.

Recall that C is assumed to be a nontrivial, pointed, solid, closed and convex cone; the vector-valued objective function $f: \mathcal{X} \rightarrow \mathbb{R}^q$ is C -convex and continuous; and the feasible set $\mathcal{X} \subseteq \mathbb{R}^m$ is compact and convex. We further assume the following from now on.

Assumption 5.1. (a) *The feasible region of (P) has nonempty interior, that is, $\text{int } \mathcal{X} \neq \emptyset$; and (b) the ordering cone C is polyhedral.*

Under Assumption 5.1 (b), it is known that C^+ is polyhedral. We denote the generating vectors of C^+ by w^1, \dots, w^J and assume without loss of generality that $\|w^j\|_* = 1$ for each $j \in \{1, \dots, J\}$.

5.1 Primal algorithm

The primal algorithm in [1] is an outer approximation algorithm, that is, it works with polyhedral outer approximations of the upper image \mathcal{P} . In particular, it starts by finding a polyhedral outer approximation \mathcal{P}_0 of the upper image and iterates by updating the outer approximation with the help of supporting halfspaces of \mathcal{P} until the approximation is sufficiently fine.

In each iteration of the primal algorithm (Algorithm 1), the following norm-minimizing scalarization problem is solved:

$$\text{minimize } \|z\| \text{ subject to } f(x) - z - v \in -C, \quad z \in \mathbb{R}^q, \quad x \in \mathcal{X}, \quad (\text{P}(v))$$

where $v \in \mathbb{R}^q$ is a parameter to be set by the algorithm. The Lagrangian dual of (P(v)) is given by

$$\text{maximize } \inf_{x \in \mathcal{X}} w^\top f(x) - w^\top v \text{ subject to } \|w\|_* \leq 1, \quad w \in C^+. \quad (\text{D}(v))$$

Before explaining the details of the algorithm, we present some results regarding (P(v)) and (D(v)); see [1] for details and further results.

Proposition 5.2. *Let $v \in \mathbb{R}^q$. The following statements hold under Assumption 5.1: (a) [1, Proposition 4.2] There exist optimal solutions (x^v, z^v) and w^v to (P(v)) and (D(v)), respectively, and the optimal values coincide. (b) [1, Proposition 4.6] If $v \notin \text{int } \mathcal{P}$, then x^v is a weak minimizer for (P). (c) [1, Remark 4.4] x^v is an optimal solution to (WS(w^v)), i.e., $\inf_{x \in \mathcal{X}} (w^v)^\top f(x) = (w^v)^\top f(x^v)$.*

The primal algorithm is initialized by solving the weighted sum scalarization problem (WS(w)) for each generating vector of the dual ordering cone C^+ . Let $x^j \in \mathcal{X}$ denote the optimal solution of (WS(w^j)) for each $j \in \{1, \dots, J\}$. By Proposition 3.2, each x^j is a weak minimizer for (P). Moreover, from Proposition 3.10 and Definition 3.8, it is known

that each w^j is a maximizer for (D). Hence, we initialize the set to be returned as a weak ϵ -solution for (P) as $\bar{\mathcal{X}} = \{x^1, \dots, x^J\}$; and the set to be returned as an ϵ -solution to (D) as $\bar{\mathcal{W}} = \{w^1, \dots, w^J\}$. Note that by Proposition 4.1, for each $j \in \{1, \dots, J\}$, the set $\mathcal{H}(\xi(w^j))$ is a supporting halfspace of \mathcal{P} such that $\mathcal{H}(\xi(w^j)) \supseteq \mathcal{P}$. Then, the initial outer approximation of \mathcal{P} is defined as $\mathcal{P}_0 := \bigcap_{j=1}^J \mathcal{H}(\xi(w^j))$. As part of the initialization, we introduce a set $\mathcal{V}_{\text{known}}$, which stores the set of vertices that have already been considered by the algorithm and initialize it as $\mathcal{V}_{\text{known}} = \emptyset$, see lines 1-3 of Algorithm 1. We later introduce a set $\mathcal{V}_{\text{unknown}}$, which stores the set of vertices of the current approximation that are not yet considered, see line 7.

In each iteration k , the first step is to compute the vertices \mathcal{V}_k of the current outer approximation \mathcal{P}_k by solving a vertex enumeration problem (line 6). Then, for each vertex in \mathcal{V}_k which have not been considered before, optimal solutions for (P(v)) and (D(v)) are found; the respective solutions are added to sets $\bar{\mathcal{X}}$ and $\bar{\mathcal{W}}$ (see Proposition 5.2(b) and Proposition 3.10); and $\mathcal{V}_{\text{known}}$ is updated (lines 7-10). Note that by Proposition 5.2(c) and Proposition 4.1, $\mathcal{H}(\xi(w^v))$ is a supporting halfspace of \mathcal{P} . If the distance of a vertex v to the upper image, namely $\|z^v\|$, is not sufficiently small, then $\mathcal{H}(\xi(w^v)) = \{y \in \mathbb{R}^q \mid \varphi(y, w^v, (w^v)^\top f(x^v)) \geq 0\}$ is stored in order to be used in updating the current outer approximation. After each vertex in \mathcal{V}_k is considered, then the current approximation is updated by intersecting it with those halfspaces (lines 11-16). The algorithm terminates when all the vertices of \mathcal{P}_k are in ϵ distance to the upper image (lines 5 and 18).

Algorithm 1 Primal algorithm

- 1: Compute an optimal solution x^j of (WS(w^j)) for each $j \in \{1, \dots, J\}$;
 - 2: Let $\mathcal{P}_0 = \bigcap_{j=1}^J \mathcal{H}(\xi(w^j))$;
 - 3: $k \leftarrow 0, \bar{\mathcal{X}} \leftarrow \{x^1, \dots, x^J\}, \bar{\mathcal{W}} \leftarrow \{w^1, \dots, w^J\}, \mathcal{V}_{\text{known}} = \emptyset$;
 - 4: **repeat**
 - 5: $M \leftarrow \mathbb{R}^q$;
 - 6: Compute the set \mathcal{V}_k of vertices of \mathcal{P}_k ;
 - 7: $\mathcal{V}_{\text{unknown}} \leftarrow \mathcal{V}_k \setminus \mathcal{V}_{\text{known}}$;
 - 8: **for** $v \in \mathcal{V}_{\text{unknown}}$ **do**
 - 9: Compute optimal solutions (x^v, z^v) and w^v to (P(v)) and (D(v));
 - 10: $\bar{\mathcal{X}} \leftarrow \bar{\mathcal{X}} \cup \{x^v\}, \bar{\mathcal{W}} \leftarrow \bar{\mathcal{W}} \cup \left\{ \frac{w^v}{\|w^v\|_*} \right\}, \mathcal{V}_{\text{known}} \leftarrow \mathcal{V}_{\text{known}} \cup \{v\}$;
 - 11: **if** $\|z^v\| > \epsilon$ **then**
 - 12: $M \leftarrow M \cap \mathcal{H}(\xi(w^v))$;
 - 13: **end if**
 - 14: **end for**
 - 15: **if** $M \neq \mathbb{R}^q$ **then**
 - 16: $\mathcal{P}_{k+1} = \mathcal{P}_k \cap M, k \leftarrow k + 1$;
 - 17: **end if**
 - 18: **until** $M = \mathbb{R}^q$
 - 19: **return** $\begin{cases} \bar{\mathcal{X}} & : \text{A finite weak } \epsilon\text{-solution to (P);} \\ \bar{\mathcal{W}} & : \text{A finite } \epsilon\text{-solution to (D);} \end{cases}$
-

Remark 5.3. A ‘break’ command can be placed between lines 12 and 13 in the algorithm. In the current version, the algorithm goes over all the vertices of the current outer approximation without updating it. With the ‘break’ command, the algorithm updates the outer approximation as soon as it detects a vertex v with $\|z^v\| > \epsilon$.

The next proposition states that Algorithm 1 gives a finite weak ϵ -solution to (P).

Proposition 5.4. [1, Theorem 5.4] *If the primal algorithm terminates, then it returns a finite weak ϵ -solution $\bar{\mathcal{X}}$ to (P).*

Next, we show that the primal algorithm yields also a finite ϵ -solution, \bar{W} , for the dual problem (D). To that end, we provide the following lemma which shows that inner and outer approximations of the lower image \mathcal{D} can be obtained by using a finite ϵ -solution $\bar{\mathcal{X}}$ of (P). Then, this lemma will be used in order to prove the main result of this section.

Lemma 5.5. *For $\epsilon > 0$, let $\bar{\mathcal{X}}$ be a finite weak ϵ -solution of (P), and $\mathcal{P}_\epsilon := \text{conv } f(\bar{\mathcal{X}}) + C + B(0, \epsilon)$. Then, $\mathcal{D}_\epsilon := \mathcal{D}_{\mathcal{P}_\epsilon}$ is an inner approximation of \mathcal{D} and*

$$\text{cone}((\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D} \supseteq \mathcal{D}_\epsilon. \quad (7)$$

Proof. Since $\bar{\mathcal{X}}$ is a finite weak ϵ -solution of (P), \mathcal{P}_ϵ is an outer approximation of the upper image \mathcal{P} by Definition 3.6, that is, $\mathcal{P}_\epsilon \supseteq \mathcal{P}$. By Proposition 4.8(a), $\mathcal{D} \supseteq \mathcal{D}_\epsilon$.

In order to show the first inclusion in (7), we first note that $\mathcal{P}' := \text{conv } f(\bar{\mathcal{X}}) + C \subseteq \mathcal{P}$ and, by Proposition 4.8(a), we have

$$\mathcal{D}' := \mathcal{D}_{\mathcal{P}'} = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in \mathcal{P}': \varphi(y, w, \alpha) \geq 0\} \supseteq \mathcal{D}. \quad (8)$$

We claim that

$$(\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\} \supseteq \mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R}) \quad (9)$$

holds. Indeed, observe that

$$\begin{aligned} & (\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\} \\ &= \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \|w\|_* = 1, \forall y \in \mathcal{P}_\epsilon: \varphi(y, w, \alpha - \epsilon) \geq 0\} \\ &= \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \|w\|_* = 1, \forall y \in \mathcal{P}', \forall \gamma \in B(0, 1): \varphi(y + \gamma\epsilon, w, \alpha - \epsilon) \geq 0\}. \end{aligned}$$

On the other hand, we have

$$\mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R}) = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \|w\|_* = 1, \forall y \in \mathcal{P}': \varphi(y, w, \alpha) \geq 0\}.$$

Let $(w^\top, \alpha)^\top \in \mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R})$ be arbitrary. Note that $\varphi(y, w, \alpha) = w^\top y - \alpha \geq 0$ for every $y \in \mathcal{P}'$ and $\|w\|_* = 1$. Moreover, for every $\gamma \in B(0, 1)$ we have $|w^\top \gamma| \leq \|\gamma\| \|w\|_* \leq 1$, hence $w^\top \gamma \geq -1$. Then, $\varphi(y + \gamma\epsilon, w, \alpha - \epsilon) = w^\top y + \epsilon w^\top \gamma - \alpha + \epsilon \geq 0$ holds, that is, $(w^\top, \alpha)^\top \in (\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}$, which implies (9).

The next step is to show that

$$\text{cone}(\mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R})) - K = \mathcal{D}'. \quad (10)$$

The inclusion \subseteq is straightforward. For the reverse inclusion, let $(w^\top, \alpha)^\top \in \mathcal{D}'$. First, suppose that $w \neq 0$. Let $\lambda := \frac{1}{\|w\|_*} > 0$. Clearly, $\|\lambda w\|_* = 1$ and $\varphi(y, \lambda w, \lambda \alpha) = \lambda w^\top y - \lambda \alpha \geq 0$ holds for each $y \in \mathcal{P}'$. Hence, $(\lambda w^\top, \lambda \alpha)^\top \in \mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R})$ and $(w^\top, \alpha)^\top \in \text{cone}(\mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R})) - K$. Now, suppose that $w = 0$. By the definition of \mathcal{D}' , we have $\alpha \leq 0$. Note that $0 \in \text{cone}(\mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R}))$ and $(w, \alpha) = (0, \alpha) \in \text{cone}(\mathcal{D}' \cap (\mathbb{S}^{q-1} \times \mathbb{R})) - K$. Therefore, (10) holds.

Finally, (8), (9) and (10) imply $\text{cone}((\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D}$. \square

Proposition 5.6. *If the primal algorithm terminates, then it returns a finite ϵ -solution $\bar{\mathcal{W}}$ to (D).*

Proof. By the structure of the algorithm, $\bar{\mathcal{W}} \neq \emptyset$ and it consists of maximizers by Proposition 3.10. Also the inclusion $\bar{\mathcal{W}} \subseteq \mathcal{W} \cap \mathbb{S}^{q-1}$ holds, see line 10 of Algorithm 1. To prove the statement, it is sufficient to show that $\mathcal{D} \subseteq \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$.

Let $\bar{\mathcal{D}} := \text{cone } \text{conv } \xi(\bar{\mathcal{W}}) - K$, $\bar{\mathcal{P}} := \mathcal{P}_{\bar{\mathcal{D}}}$. By Lemma 4.11, $\bar{\mathcal{P}} = \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \xi(\bar{\mathcal{W}}): \varphi(y, w, \alpha) \geq 0\}$. Though not part of the original algorithm, we introduce an alternative for $\bar{\mathcal{W}}$ that is updated only when the current vertex v is sufficiently far from the upper image, i.e., when $\|z^v\| > \epsilon$. More precisely, let us introduce a set $\bar{\bar{\mathcal{W}}}$ that is initialized as $\bar{\bar{\mathcal{W}}} = \{w^1, \dots, w^J\}$ in line 3 of Algorithm 1 and updated as $\bar{\bar{\mathcal{W}}} \leftarrow \bar{\bar{\mathcal{W}}} \cup \{\frac{w^v}{\|w^v\|_*}\}$ in line 12 throughout the algorithm. Observe that $\bar{\bar{\mathcal{W}}} \subseteq \bar{\mathcal{W}}$.

Suppose that the algorithm terminates at the \bar{k} th iteration and let $\mathcal{P}_{\bar{k}}$ denote the resulting outer approximation of the upper image. By the construction of $\bar{\mathcal{W}}$, $\mathcal{P}_{\bar{k}} = \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \xi(\bar{\mathcal{W}}): \varphi(y, w, \alpha) \geq 0\}$. Since $\bar{\bar{\mathcal{W}}} \subseteq \bar{\mathcal{W}}$, we get $\bar{\mathcal{P}} \subseteq \mathcal{P}_{\bar{k}}$.

Let us define $\mathcal{P}_\epsilon := \text{conv } f(\bar{\mathcal{X}}) + C + B(0, \epsilon)$. By the structure and the termination criterion of the algorithm, for every vertex $v \in \mathcal{V}_{\bar{k}}$ of $\mathcal{P}_{\bar{k}}$, the scalarization problem (P(v)) is solved; x^v is added to $\bar{\mathcal{X}}$; and we have $\|z^v\| \leq \epsilon$. Moreover, $v + z^v \in \{f(x^v)\} + C \subseteq \text{conv } f(\bar{\mathcal{X}}) + C$ holds for every $v \in \mathcal{V}_{\bar{k}}$. Thus, $\mathcal{V}_{\bar{k}} \subseteq \mathcal{P}_\epsilon$. From [1, Lemma 5.2], the recession cone of $\mathcal{P}_{\bar{k}}$ is the ordering cone; hence, $\mathcal{P}_{\bar{k}} = \text{conv } \mathcal{V}_{\bar{k}} + C$. Moreover, \mathcal{P}_ϵ is a convex upper set. As a result, we have $\mathcal{P}_{\bar{k}} \subseteq \mathcal{P}_\epsilon$. Together with $\bar{\mathcal{P}} \subseteq \mathcal{P}_{\bar{k}}$, the last inclusion implies that $\bar{\mathcal{P}} \subseteq \mathcal{P}_\epsilon$.

Define $\mathcal{D}_\epsilon := \mathcal{D}_{\mathcal{P}_\epsilon} = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in \mathcal{P}_\epsilon: \varphi(y, w, \alpha) \geq 0\}$. By Corollary 4.12(a) and since $\bar{\mathcal{P}} \subseteq \mathcal{P}_\epsilon$, we have $\bar{\mathcal{D}} \supseteq \mathcal{D}_\epsilon$. Moreover, using Proposition 5.4 and Lemma 5.5, $\text{cone}((\mathcal{D}_\epsilon \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D}$. With $\bar{\mathcal{D}} \supseteq \mathcal{D}_\epsilon$, this implies

$$\text{cone}((\bar{\mathcal{D}} \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D}. \quad (11)$$

It is straightforward to check that $(\bar{\mathcal{D}} \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\} \subseteq \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$, which implies that

$$\text{cone}((\bar{\mathcal{D}} \cap (\mathbb{S}^{q-1} \times \mathbb{R})) + \epsilon\{e^{q+1}\}) - K \subseteq \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K.$$

By (11), $\text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K \supseteq \mathcal{D}$; hence, $\bar{\mathcal{W}}$ is a finite ϵ -solution to (D). \square

5.2 Dual algorithm

In this section, we describe a geometric dual algorithm for solving problems (P) and (D). The main idea is to construct outer approximations of the lower image \mathcal{D} , iteratively. Recall

that, by Proposition 3.4, \mathcal{D} is a closed convex cone; similarly, we will see that the outer approximations found through the iterations of the dual algorithm are polyhedral convex cones.

The dual algorithm (Algorithm 2) is initialized by solving a weighted sum scalarization for some weight vector from $\text{int } C^+$. In particular, we solve $(\text{WS}(w^0))$ by taking $w^0 := \sum_{j=1}^J w^j / \|\sum_{j=1}^J w^j\|_*$. By Propositions 3.2 and 3.10, an optimal solution x^0 of $(\text{WS}(w^0))$ is a weak minimizer for (P) and w^0 is a maximizer for (D). Hence, we set $\bar{\mathcal{X}} = \{x^0\}$, $\bar{\mathcal{W}} = \{w^0\}$. Moreover, using Proposition 4.1 and the definition of the lower image \mathcal{D} , we define the initial outer approximation of \mathcal{D} as $\mathcal{D}_0 := \mathcal{H}^*(f(x^0)) \cap (C^+ \times \mathbb{R}) \supseteq \mathcal{D}$; see lines 1-3 of Algorithm 2. Note that \mathcal{D}_0 satisfies $\mathcal{D}_0 = \mathcal{D}_0 - K$ and, under Assumption 5.1, it is a polyhedral convex cone.

Throughout the algorithm, weighted sum scalarizations will be solved for some weight vectors from \mathcal{W} . In order to keep track of the already used $w \in \mathcal{W}$, we keep a list $\mathcal{W}_{\text{known}}$ and initialize it as the empty set.

In each iteration k , first, the set $\mathcal{D}_k^{\text{dir}}$ of extreme directions of the current outer approximation \mathcal{D}_k is computed (line 6). The extreme directions in $\mathcal{D}_k^{\text{dir}} \setminus (\{0\} \times \mathbb{R})$ are normalized such that $\|w\|_* = 1$, and $\mathcal{D}_k^{\text{dir}}$ is updated such that it is a subset of $\mathbb{S}^{q-1} \times \mathbb{R}$ (line 7). It will be seen that \mathcal{D}_k is constructed by intersecting \mathcal{D}_0 with a set of halfspaces of the form $\mathcal{H}^*(f(x))$, where x is a weak minimizer. Then, by the definitions of \mathcal{D}_0 and $\mathcal{H}^*(\cdot)$, it is clear that $\mathcal{D}_k = \mathcal{D}_k - K$; moreover, e^{q+1} is not a recession direction for \mathcal{D}_k . Hence, we have

$$\mathcal{D}_k = \text{cone conv}(\mathcal{D}_k^{\text{dir}} \cup \{-e^{q+1}\}) = \text{cone conv } \mathcal{D}_k^{\text{dir}} - K. \quad (12)$$

For each unknown extreme direction $(w^\top, \alpha)^\top$, an optimal solution x^w and the optimal value p^w of $(\text{WS}(w))$ are computed (lines 8-10). Recall that x^w is a weak minimizer by Proposition 3.2 and $\xi(w) = (w, p^w)$ is a K -maximal element of \mathcal{D} by Proposition 3.10; we update $\bar{\mathcal{X}}$ and $\bar{\mathcal{W}}$ accordingly, and add $(w^\top, \alpha)^\top$ to the set of known extreme directions (line 11). Since $(w^\top, \alpha)^\top \in \mathcal{D}_k$ and $\mathcal{D}_k \supseteq \mathcal{D}$, we have $\alpha \geq p^w$. If the difference between α and p^w is greater than the allowed error ϵ , then the supporting halfspace $\mathcal{H}^*(f(x^w))$ of \mathcal{D} is computed and stored. Once all unknown extreme directions are explored, the current outer approximation is updated by using the stored supporting halfspaces (lines 12-14 and 17). The algorithm terminates when every unknown extreme direction $(w^\top, \alpha)^\top$ is sufficiently close to the lower image in terms of the ‘‘vertical distance’’ $\alpha - p^w$ (line 19).

Remark 5.7. A ‘break’ command can be placed between lines 13 and 14 in the algorithm. With the current version, the algorithm goes through all the extreme directions of the current outer approximation without updating it. With the ‘break’ command, the algorithm would update the outer approximation as soon as it detects an extreme direction w with $\alpha - p^w > \epsilon$.

The following lemma will be useful in proving Propositions 5.9 and 5.13.

Lemma 5.8. For $\epsilon > 0$, suppose that the algorithm terminates at the \bar{k}^{th} iteration, returns sets $\bar{\mathcal{X}}, \bar{\mathcal{W}}$. Let $\bar{\mathcal{P}} := \text{conv } f(\bar{\mathcal{X}}) + C$, $\bar{\mathcal{D}} := \mathcal{D}_{\bar{\mathcal{P}}}$ and $\mathcal{D}_\epsilon := \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$. Then, $\mathcal{D}_\epsilon \supseteq \mathcal{D}_{\bar{k}} \supseteq \bar{\mathcal{D}} \supseteq \mathcal{D}$, where $\mathcal{D}_{\bar{k}}$ denotes the resulting outer approximation of \mathcal{D} at termination.

Algorithm 2 Dual algorithm

- 1: Compute an optimal solution x^0 to $(\text{WS}(w^0))$ for $w^0 = \frac{\sum_{j=1}^J w^j}{\|\sum_{j=1}^J w^j\|_*}$;
 - 2: Let $\mathcal{D}_0 = \mathcal{H}^*(f(x^0)) \cap (C^+ \times \mathbb{R})$;
 - 3: $k \leftarrow 0, \bar{\mathcal{X}} \leftarrow \{x^0\}, \bar{\mathcal{W}} \leftarrow \{w^0\}, \mathcal{D}_{\text{known}} = \emptyset$;
 - 4: **repeat**
 - 5: $M \leftarrow \mathbb{R}^{q+1}$;
 - 6: Compute the set $\mathcal{D}_k^{\text{dir}}$ of extreme directions of \mathcal{D}_k ;
 - 7: $\mathcal{D}_k^{\text{dir}} \leftarrow \left\{ \frac{(w^\top, \alpha)^\top}{\|w\|_*} \in \mathbb{R}^{q+1} \mid (w^\top, \alpha)^\top \in \mathcal{D}_k^{\text{dir}} \setminus (\{0\} \times \mathbb{R}) \right\}$;
 - 8: $\mathcal{D}_{\text{unknown}} \leftarrow \mathcal{D}_k^{\text{dir}} \setminus \mathcal{D}_{\text{known}}$;
 - 9: **for** $(w^\top, \alpha)^\top \in \mathcal{D}_{\text{unknown}}$ **do**
 - 10: Compute an optimal solution x^w to $(\text{WS}(w))$ and let $p^w := w^\top f(x^w)$;
 - 11: $\bar{\mathcal{X}} \leftarrow \bar{\mathcal{X}} \cup \{x^w\}, \bar{\mathcal{W}} \leftarrow \bar{\mathcal{W}} \cup \{w\}, \mathcal{D}_{\text{known}} \leftarrow \mathcal{D}_{\text{known}} \cup \{(w^\top, \alpha)^\top\}$;
 - 12: **if** $\alpha - p^w > \epsilon$ **then**
 - 13: $M \leftarrow M \cap \mathcal{H}^*(f(x^w))$;
 - 14: **end if**
 - 15: **end for**
 - 16: **if** $M \neq \mathbb{R}^{q+1}$ **then**
 - 17: $\mathcal{D}_{k+1} = \mathcal{D}_k \cap M, k \leftarrow k + 1$;
 - 18: **end if**
 - 19: **until** $M = \mathbb{R}^{q+1}$
 - 20: **return** $\begin{cases} \bar{\mathcal{X}} & : \text{A finite weak } \tilde{\epsilon}\text{-solution to (P);} \\ \bar{\mathcal{W}} & : \text{A finite } \epsilon\text{-solution to (D);} \end{cases}$
-

Proof. First observe that $\bar{\mathcal{P}} \subseteq \mathcal{P}$ implies $\bar{\mathcal{D}} \supseteq \mathcal{D}$ from Proposition 4.8. Moreover, $\bar{\mathcal{D}} \subseteq \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in f(\bar{\mathcal{X}}): \varphi(y, w, \alpha) \geq 0\}$ holds since $\bar{\mathcal{P}} \supseteq f(\bar{\mathcal{X}})$. Consider an artificial set $\bar{\bar{\mathcal{X}}}$ that is initialized in line 3 of Algorithm 2 as the empty set, and updated in line 13 as $\bar{\bar{\mathcal{X}}} \leftarrow \bar{\bar{\mathcal{X}}} \cup \{x^w\}$. By the definition of $\bar{\bar{\mathcal{X}}}$, we have $\mathcal{D}_{\bar{k}} = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in f(\bar{\bar{\mathcal{X}}}): \varphi(y, w, \alpha) \geq 0\}$ and $\bar{\bar{\mathcal{X}}} \subseteq \bar{\mathcal{X}}$. Therefore, $\mathcal{D}_{\bar{k}} \supseteq \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in f(\bar{\mathcal{X}}): \varphi(y, w, \alpha) \geq 0\}$. Note that for every $(w^\top, \alpha)^\top \in \mathcal{D}_{\bar{k}}^{\text{dir}} \subseteq \mathbb{S}^{q-1} \times \mathbb{R}$, the scalarization problem $(\text{WS}(w))$ is solved; w is added to $\bar{\mathcal{W}}$ and we have $\alpha - p^w \leq \epsilon$. Then, we have $\mathcal{D}_{\bar{k}}^{\text{dir}} \subseteq \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\} - K$. On the other hand, $\mathcal{D}_{\bar{k}} = \text{cone conv } \mathcal{D}_{\bar{k}}^{\text{dir}} - K$, see (12). These imply $\mathcal{D}_{\bar{k}} \subseteq \text{cone conv}(\xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K = \mathcal{D}_\epsilon$. Therefore, $\bar{\mathcal{D}} \subseteq \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \forall y \in f(\bar{\mathcal{X}}): \varphi(y, w, \alpha) \geq 0\} \subseteq \mathcal{D}_{\bar{k}} \subseteq \mathcal{D}_\epsilon$. \square

Next, we prove that the dual algorithm returns a finite ϵ -solution to problem (D).

Proposition 5.9. *If Algorithm 2 terminates, then it returns a finite ϵ -solution $\bar{\mathcal{W}}$ to (D).*

Proof. By the structure of the algorithm and Proposition 3.10, $\bar{\mathcal{W}}$ is nonempty and consists of maximizers. Also the inclusion $\bar{\mathcal{W}} \subseteq \mathcal{W} \cap \mathbb{S}^{q-1}$ holds. From Lemma 5.8, we have $\mathcal{D} \subseteq \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$, which implies that $\bar{\mathcal{W}}$ is a finite ϵ -solution to (D). \square

The next step is to prove that the algorithm returns also a solution to the primal problem (P). In order to prove this result, the following lemma and propositions will be useful. To

that end, for each $n \in \{2, 3, \dots\}$, let us define

$$\Delta^{n-1} := \left\{ \lambda \in \mathbb{R}_+^n \mid \sum_{j=1}^n \lambda_j = 1 \right\}, \quad \Delta_+^{n-1} := \left\{ \lambda \in \mathbb{R}_+^n \mid \sum_{j=1}^n \lambda_j \geq 1 \right\}.$$

Lemma 5.10. *Let d^1, \dots, d^n be the generating vectors of a pointed convex polyhedral cone $D \subseteq \mathbb{R}^q$, where $n \geq 2$. (a) For every $\lambda \in \Delta^{n-1}$, $\sum_{j=1}^n \lambda_j d^j \neq 0$ holds. (b) It holds $\min_{\lambda \in \Delta^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_* > 0$. (c) It holds $\min_{\lambda \in \Delta^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_* = \min_{\lambda \in \Delta_+^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_*$.*

Proof. To prove (a), let $\lambda \in \Delta^{n-1}$. Assume to the contrary that $\sum_{j=1}^n \lambda_j d^j = 0$. Since $\sum_{j=1}^n \lambda_j = 1$, we have $\lambda_j > 0$ for at least one $j \in \{1, \dots, n\}$. Without loss of generality, we may assume that $j = 1$. Hence, $\frac{1}{\lambda_1} \sum_{j=2}^n \lambda_j d^j = -d^1$. Since D is a convex cone, $\frac{1}{\lambda_1} \sum_{j=2}^n \lambda_j d^j \in D$. Therefore, $d^1, -d^1 \in D$, contradicting the pointedness of D [6, Section 2.4.1]. Therefore, $\sum_{j=1}^n \lambda_j d^j \neq 0$.

By (a), $\|\sum_{j=1}^n \lambda_j d^j\|_* > 0$ for every $\lambda \in \Delta^{n-1}$. Since the feasible set Δ^{n-1} is compact, the minimum of the continuous function $\lambda \mapsto \|\sum_{j=1}^n \lambda_j d^j\|_*$ is attained at some $\bar{\lambda} \in \Delta^{n-1}$. Hence, the minimum $\|\sum_{j=1}^n \bar{\lambda}_j d^j\|_*$ is also strictly positive. This proves (b).

Finally, we prove (c). Since $\Delta^{n-1} \subseteq \Delta_+^{n-1}$, we have $\min_{\lambda \in \Delta^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_* \geq \inf_{\lambda \in \Delta_+^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_*$. To prove the reverse inequality, assume to the contrary that there exists $\bar{\lambda} \in \Delta_+^{n-1}$ such that $\|\sum_{j=1}^n \bar{\lambda}_j d^j\|_* < \min_{\lambda \in \Delta^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_*$. Then, using $\bar{\lambda} \in \Delta_+^{n-1}$, we have

$$\left\| \sum_{j=1}^n \bar{\lambda}_j d^j \right\|_* \geq \frac{1}{\sum_{i=1}^n \bar{\lambda}_i} \left\| \sum_{j=1}^n \bar{\lambda}_j d^j \right\|_* = \left\| \sum_{j=1}^n \frac{\bar{\lambda}_j}{\sum_{i=1}^n \bar{\lambda}_i} d^j \right\|_* \geq \min_{\lambda \in \Delta^{n-1}} \left\| \sum_{j=1}^n \lambda_j d^j \right\|_*,$$

a contradiction. Therefore, $\min_{\lambda \in \Delta^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_* = \inf_{\lambda \in \Delta_+^{n-1}} \|\sum_{j=1}^n \lambda_j d^j\|_*$. Since the minimum is attained on the left and $\Delta^{n-1} \subseteq \Delta_+^{n-1}$, the infimum on the right is also a minimum. \square

Next, we show that an inner $\tilde{\epsilon}$ -approximation of the upper image \mathcal{P} can be obtained by using a finite ϵ -solution \bar{W} of (D).

Proposition 5.11. *For $\epsilon > 0$, let \bar{W} be a finite ϵ -solution of (D), and define $\mathcal{D}_\epsilon := \text{cone}(\text{conv } \xi(\bar{W}) + \epsilon\{e^{q+1}\}) - K$. Then, $\mathcal{P}_\epsilon := \mathcal{P}_{\mathcal{D}_\epsilon}$ is an inner approximation of \mathcal{P} and $\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}) \supseteq \mathcal{P} \supseteq \mathcal{P}_\epsilon$, where $\tilde{\epsilon} = \epsilon / \min_{\lambda \in \Delta^{J-1}} \|\sum_{j=1}^J \lambda_j w^j\|_*$.*

Proof. Since \bar{W} is a finite ϵ -solution of (D), \mathcal{D}_ϵ is an outer approximation of the lower image \mathcal{D} by Definition 3.8, that is, $\mathcal{D}_\epsilon \supseteq \mathcal{D}$. Therefore, by Proposition 4.8(b), the inclusion $\mathcal{P} \supseteq \mathcal{P}_\epsilon$ holds.

Next, we show that $\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$. Assume that there exists $\bar{y} \in \mathcal{P} \setminus (\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}))$. Hence, there exists $\bar{w} \in \mathbb{R}^q \setminus \{0\}$ such that $\bar{w}^\top \bar{y} < \inf_{y \in \mathcal{P}_\epsilon} \bar{w}^\top y + \inf_{\gamma \in B(0,1)} \tilde{\epsilon} \bar{w}^\top \gamma$. Without loss of generality, we may assume that $\|\bar{w}\|_* = 1$ so that $\inf_{\gamma \in B(0,1)} \tilde{\epsilon} \bar{w}^\top \gamma = -\tilde{\epsilon} \|\bar{w}\|_* = -\tilde{\epsilon}$. Hence, we have

$$\bar{w}^\top \bar{y} + \tilde{\epsilon} < \inf_{y \in \mathcal{P}_\epsilon} \bar{w}^\top y =: \bar{\alpha}. \quad (13)$$

The definition of $\bar{\alpha}$ ensures that $\varphi(y, \bar{w}, \bar{\alpha}) = \bar{w}^\top y - \bar{\alpha} \geq 0$ for each $y \in \mathcal{P}_\epsilon$, that is, $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{D}_{\mathcal{P}_\epsilon}$. By Corollary 4.12(b), we have $\mathcal{D}_{\mathcal{P}_\epsilon} = \mathcal{D}_{\mathcal{P}_{\mathcal{D}_\epsilon}} = \mathcal{D}_\epsilon$. Therefore, $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{D}_\epsilon$. Hence, there exist $\delta \geq 0$, $k \geq 0$, $n \in \mathbb{N}$ and $\mu \in \Delta^{n-1}$, $((\bar{w}^i)^\top, \alpha_i)^\top \in \xi(\bar{\mathcal{W}})$ for each $i \in \{1, \dots, n\}$ such that

$$(\bar{w}^\top, \bar{\alpha})^\top = \delta \left(\sum_{i=1}^n \mu_i ((\bar{w}^i)^\top, \alpha_i)^\top + \epsilon e^{q+1} \right) - k e^{q+1}. \quad (14)$$

Using (14), we have $\bar{w} = \delta \sum_{i=1}^n \mu_i \bar{w}^i$ and $\bar{\alpha} = \delta (\sum_{i=1}^n \mu_i \alpha_i + \epsilon) - k$. In particular, having $\|\bar{w}\|_* = 1$ implies that $\delta = \frac{1}{\|\sum_{i=1}^n \mu_i \bar{w}^i\|_*}$.

Next, we claim that $(\bar{w}^\top, \bar{\alpha} - \tilde{\epsilon}) \in \mathcal{D}$. Since $\mathcal{W} = C^+$ is a convex cone, we have $\bar{w} \in \mathcal{W}$. Therefore, using the definition of \mathcal{D} , proving the inequality

$$\bar{\alpha} - \tilde{\epsilon} \leq \inf_{x \in \mathcal{X}} \bar{w}^\top f(x) \quad (15)$$

is enough to conclude that $(\bar{w}^\top, \bar{\alpha} - \tilde{\epsilon}) \in \mathcal{D}$. Since $((\bar{w}^i)^\top, \alpha_i) \in \xi(\bar{\mathcal{W}})$, $\alpha_i = \inf_{x \in \mathcal{X}} (\bar{w}^i)^\top f(x)$ for each $i \in \{1, \dots, n\}$. Hence, using (14) and $k \geq 0$, we have

$$\bar{\alpha} = \delta \left(\sum_{i=1}^n \mu_i \alpha_i + \epsilon \right) - k \leq \delta \left(\sum_{i=1}^n \mu_i \inf_{x \in \mathcal{X}} (\bar{w}^i)^\top f(x) + \epsilon \right) = \inf_{x \in \mathcal{X}} \bar{w}^\top f(x) + \delta \epsilon.$$

Hence, we have $\bar{\alpha} - \delta \epsilon \leq \inf_{x \in \mathcal{X}} \bar{w}^\top f(x)$.

Let us show that $\delta \epsilon \leq \tilde{\epsilon}$ so that (15) follows. For each $i \in \{1, \dots, n\}$, since $\bar{w}^i \in \bar{\mathcal{W}} \subseteq \mathcal{W} \cap (\mathbb{S}^{q-1} \times \mathbb{R})$, there exists $(\gamma_{i1}, \dots, \gamma_{iJ})^\top \in \Delta^{J-1}$ such that $\bar{w}^i = \sum_{j=1}^J \frac{\gamma_{ij}}{\|\sum_{j'=1}^J \gamma_{ij'} w^{j'}\|_*} w^j$, where w^1, \dots, w^J are the generating vectors of C^+ . It follows that $\sum_{i=1}^n \mu_i \bar{w}^i = \sum_{j=1}^J \lambda_j w^j$, where $\lambda_j := \sum_{i=1}^n \mu_i \frac{\gamma_{ij}}{\|\sum_{j'=1}^J \gamma_{ij'} w^{j'}\|_*} \geq 0$ for $j \in \{1, \dots, J\}$. Note that $\sum_{j=1}^J \lambda_j \geq 1$ since $\|\sum_{j=1}^J \gamma_{ij} w^j\|_* \leq \sum_{j=1}^J \|\gamma_{ij} w^j\|_* = \sum_{j=1}^J \gamma_{ij} \|w^j\|_* = 1$, $\sum_{j=1}^J \gamma_{ij} = 1$ for each $i \in \{1, \dots, n\}$ and $\sum_{i=1}^n \mu_i = 1$. Therefore, $\lambda = (\lambda_1, \dots, \lambda_J)^\top \in \Delta_+^{J-1}$. Since the dual cone C^+ is convex and pointed, we may use Lemma 5.10 to write

$$\delta \epsilon = \frac{\epsilon}{\left\| \sum_{j=1}^J \lambda_j w^j \right\|_*} \leq \frac{\epsilon}{\min_{\lambda' \in \Delta_+^{J-1}} \left\| \sum_{j=1}^J \lambda'_j w^j \right\|_*} = \frac{\epsilon}{\min_{\lambda' \in \Delta^{J-1}} \left\| \sum_{j=1}^J \lambda'_j w^j \right\|_*} = \tilde{\epsilon}.$$

Therefore, (15) follows and we have $(\bar{w}^\top, \bar{\alpha} - \tilde{\epsilon})^\top \in \mathcal{D}$. However, by (13), we have $\varphi(\bar{y}, \bar{w}, \bar{\alpha} - \tilde{\epsilon}) = \bar{w}^\top \bar{y} - \bar{\alpha} + \tilde{\epsilon} < 0$ for $\bar{y} \in \mathcal{P}$ and $(\bar{w}^\top, \bar{\alpha} - \tilde{\epsilon})^\top \in \mathcal{D}$, a contradiction to Proposition 4.8. Hence, $\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$. \square

The next proposition provides a better bound on the realized approximation error compared to Proposition 5.11; however, it requires post-processing the finite ϵ -solution of (D) provided by the algorithm. We omit its proof for brevity.

Proposition 5.12. *For $\epsilon > 0$, let $\bar{\mathcal{W}}$ be a finite ϵ -solution of (D). Define $\mathcal{D}_\epsilon := \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon \{e^{q+1}\}) - K$. Let $\mathcal{F} = \{F_1, \dots, F_T\}$ be the set of K -maximal facets of \mathcal{D}_ϵ . For each $i \in \{1, \dots, T\}$, let $\{((w^{i1})^\top, \alpha_{i1})^\top, \dots, ((w^{iJ_i})^\top, \alpha_{iJ_i})^\top\}$ be the set of extreme directions of F_i and define $f_{\min}^i := \min_{\lambda \in \Delta^{J_i-1}} \left\| \sum_{j=1}^{J_i} \lambda_j w^{ij} \right\|_*$. Then, $\mathcal{P}_\epsilon := \mathcal{P}_{\mathcal{D}_\epsilon}$ is an inner approximation of \mathcal{P} and $\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}) \supseteq \mathcal{P} \supseteq \mathcal{P}_\epsilon$, where $\tilde{\epsilon} = \epsilon / \min\{f_{\min}^1, \dots, f_{\min}^T\}$.*

Proposition 5.13. *If the algorithm terminates, then it returns a finite weak $\tilde{\epsilon}$ -solution $\bar{\mathcal{X}}$ to (P), where $\tilde{\epsilon}$ is either as in Proposition 5.11 or as in Proposition 5.12.*

Proof. Note that every element of $\bar{\mathcal{X}}$ is of the form x^w which is an optimal solution to (WS(w)) for some $w \in C^+ \setminus \{0\}$; by Proposition 3.2, x^w is a weak minimizer of (P). To prove the statement, we need to show that $\text{conv } f(\bar{\mathcal{X}}) + C + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$ holds. Let us define $\bar{\mathcal{P}} := \text{conv } f(\bar{\mathcal{X}}) + C$, $\bar{\mathcal{D}} := \mathcal{D}_{\bar{\mathcal{P}}}$. From Lemma 5.8, we have $\bar{\mathcal{D}} \subseteq \mathcal{D}_\epsilon = \text{cone}(\text{conv } \xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) - K$. Then, Proposition 4.9 implies $\bar{\mathcal{P}} = \mathcal{P}_{\bar{\mathcal{D}}} \supseteq \mathcal{P}_{\mathcal{D}_\epsilon}$. By Proposition 5.9 and Proposition 5.11 (or Proposition 5.12), we have $\mathcal{P}_{\mathcal{D}_\epsilon} + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$. As $\bar{\mathcal{P}} \supseteq \mathcal{P}_{\mathcal{D}_\epsilon}$, we get $\bar{\mathcal{P}} + B(0, \tilde{\epsilon}) = \text{conv } f(\bar{\mathcal{X}}) + C + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$, i.e., $\bar{\mathcal{X}}$ is a finite $\tilde{\epsilon}$ -solution to (P). \square

6 Relationships to similar approaches from literature

In this section, we compare our approach with the approaches in two closely related works.

6.1 Connection to the geometric dual problem by Heyde [30]

In [30], a geometric dual image, which corresponds to the upper image \mathcal{P} of problem (P), is constructed in \mathbb{R}^q as follows. For a fixed $c \in \text{int } C$, a matrix $E \in \mathbb{R}^{q \times (q-1)}$ is taken such that $T = [E \ c] \in \mathbb{R}^{q \times q}$ is orthogonal. Then, the dual image is defined as

$$\mathcal{D}^H := \{(t^\top, s)^\top \in \mathbb{R}^{q-1} \times \mathbb{R} \mid c^*(t) \in C^+, s \leq \inf_{x \in \mathcal{X}} (c^*(t))^\top f(x)\},$$

where $c^* : \mathbb{R}^{q-1} \rightarrow \mathbb{R}^q$ is given by $c^*(t) := T^{-\top}(t^\top, 1)^\top$. Throughout, we denote the identity matrix in $\mathbb{R}^{n \times n}$ by $I_n \in \mathbb{R}^{n \times n}$. Observe from $T^\top T^{-\top} = I_q$ that $c^\top T^{-\top} = e_q$ and $E^\top T^{-\top} = [I_{q-1} \ 0] \in \mathbb{R}^{(q-1) \times q}$ hold true.

We will show that \mathcal{D}^H and the dual image \mathcal{D} constructed in Section 3 are related. For this purpose, we define $\hat{T} \in \mathbb{R}^{(q+1) \times (q+1)}$, $S \in \mathbb{R}^{(q+1) \times (q+1)}$ and $P \in \mathbb{R}^{q \times (q+1)}$ as

$$\hat{T} := \begin{bmatrix} T & 0 \\ 0 & 1 \end{bmatrix}, S := \begin{bmatrix} I_{q-1} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, P := [I_q \ 0], \quad (16)$$

where 0 is the matrix or vector of zeros with respective sizes. The following lemma will be useful for proving the next proposition.

Lemma 6.1. *Let $H_y := \{(t^\top, s)^\top \in \mathbb{R}^{q-1} \times \mathbb{R} \mid c^*(t) \in C^+, s \leq (c^*(t))^\top y\}$, where $y \in \mathbb{R}^q$ is fixed. Then,*

$$\bar{H}_y := \{(w^\top, \alpha)^\top \in C^+ \times \mathbb{R} \mid \alpha \leq w^\top y\} = \hat{T}^{-\top} S [\text{cone}(H_y \times \{1\})].$$

Proof. Note that for any $(t^\top, s)^\top \in H_y$, we have $\hat{T}^{-\top} S (t^\top, s, 1)^\top = (c^*(t)^\top, s)^\top$, $c^*(t) \in C^+$ and $s \leq (c^*(t))^\top y$. Hence, $\hat{T}^{-\top} S [H_y \times \{1\}] \subseteq \bar{H}_y$ holds. As \bar{H}_y is a cone, we obtain

$$\hat{T}^{-\top} S [\text{cone}(H_y \times \{1\})] = \text{cone } \hat{T}^{-\top} S [H_y \times \{1\}] \subseteq \bar{H}_y.$$

For the reverse inclusion, let $(w^\top, \alpha)^\top \in \bar{H}_y$. Consider $t = \frac{E^\top w}{c^\top w}, s = \frac{\alpha}{c^\top w}$. Then,

$$c^*(t) = \frac{1}{k^\top w} T^{-\top} (w^\top E, w^\top c)^\top = \frac{1}{c^\top w} T^{-\top} T^\top w = \frac{w}{c^\top w} \in C^+.$$

Moreover, $s \leq (c^*(t))^\top y$ holds. Hence, $(t^\top, s)^\top \in H_y$. Similar to the previous case, we have

$$\hat{T}^{-\top} S(t^\top, s, 1)^\top = (c^*(t)^\top, s)^\top = \frac{1}{c^\top w} (w^\top, \alpha)^\top \in \hat{T}^{-\top} S[\mathcal{D}^H \times \{1\}],$$

which implies $(w^\top, \alpha)^\top \in \text{cone} \hat{T}^{-\top} S[\mathcal{D}^H \times \{1\}]$. \square

The next proposition shows that \mathcal{D} and \mathcal{D}^H can be recovered from each other.

Proposition 6.2. *The following relations hold true for \mathcal{D} and \mathcal{D}^H :*

$$(a) \mathcal{D}^H = P \left[S\hat{T}^\top[\mathcal{D}] \cap \{z \in \mathbb{R}^{q+1} \mid z_{q+1} = 1\} \right], \quad (b) \mathcal{D} = \hat{T}^{-\top} S \left[\text{cone}(\mathcal{D}^H \times \{1\}) \right].$$

Proof. To see (a), let $(t^\top, s)^\top \in P[S\hat{T}^\top[\mathcal{D}] \cap \{z \in \mathbb{R}^{q+1} \mid z_{q+1} = 1\}]$ be arbitrary. There exists $(w^\top, \alpha)^\top \in \mathcal{D}$ such that $(t^\top, s)^\top = P S\hat{T}^\top (w^\top, \alpha)^\top = (w^\top E, \alpha)^\top$ together with $c^\top w = 1$. In particular, $t = E^\top w, s = \alpha$. We have $w^\top T = (w^\top E, w^\top c) = (t^\top, 1)$ and $c^*(t) = T^{-\top} (t^\top, 1)^\top = T^{-\top} T^\top w = w \in C^+$ as $(w^\top, \alpha)^\top \in \mathcal{D}$. Moreover, $s \leq \inf_{x \in \mathcal{X}} (c^*(t))^\top f(x)$ holds. Hence, $(t^\top, s)^\top \in \mathcal{D}^H$. For the reverse inclusion, let $(t^\top, s)^\top \in \mathcal{D}^H$. Then, $(c^*(t)^\top, s)^\top \in \mathcal{D}$ and $S\hat{T}^\top (c^*(t)^\top, s)^\top = ((c^*(t))^\top E, s, c^\top c^*(t))^\top = (t^\top, s, 1)^\top$. Hence, $(t^\top, s)^\top \in P[S\hat{T}^\top[\mathcal{D}] \cap \{z \in \mathbb{R}^{q+1} \mid z_{q+1} = 1\}]$.

To see (b), let H_y, \bar{H}_y be defined as in Lemma 6.1. Noting that $\mathcal{D}^H = \bigcap_{x \in \mathcal{X}} H_{f(x)}$ and $\mathcal{D} = \bigcap_{x \in \mathcal{X}} \bar{H}_{f(x)}$, Lemma 6.1 implies

$$\begin{aligned} \hat{T}^{-\top} S \left[\text{cone}(\mathcal{D}^H \times \{1\}) \right] &= \hat{T}^{-\top} S \left[\text{cone} \left(\bigcap_{x \in \mathcal{X}} (H_{f(x)} \times \{1\}) \right) \right] \\ &= \bigcap_{x \in \mathcal{X}} \hat{T}^{-\top} S \left[\text{cone} (H_{f(x)} \times \{1\}) \right] = \bigcap_{x \in \mathcal{X}} \bar{H}_{f(x)} = \mathcal{D}. \end{aligned}$$

\square

Remark 6.3. *For \mathcal{D}^H and \mathcal{D} , one can check that the following results hold: (i) A subset $F \subseteq \mathcal{D}^H$ is an exposed face of \mathcal{D}^H if and only if $\hat{T}^{-\top} S[\text{cl cone}(F \times \{1\})]$ is an exposed face of \mathcal{D} . (ii) For every exposed face G of \mathcal{D} , not lying in the hyperplane $\hat{T}^{-\top} S[\mathbb{R}^q \times \{0\}]$, there is an exposed face F of \mathcal{D}^H such that $\hat{T}^{-\top} S[\text{cl cone}(F \times \{1\})] = G$. The proof relies on a generalization of [40, Lemma 3.1] for the exposed faces of (possibly nonpolyhedral) closed convex sets.*

6.2 Connection to the geometric dual algorithm by Löhne, Rudloff and Ulus [37]

The geometric dual algorithm proposed in [37] is based on the geometric duality results from [30]. We will now show that by considering a special norm $\|\cdot\|$ used in Algorithm 2,

we recover the dual algorithm from [37]. In particular, assume that the dual norm satisfies $\|w\|_* = c^\top w$ for all $w \in C^+$, where $c \in \text{int } C$ is fixed.¹

We denote by $\mathcal{D}_k^H, k \geq 0$ the outer approximations of \mathcal{D}^H that is obtained in the k^{th} iteration of [37, Algorithm 2]. Let w^0, x^0 be as defined in Section 5.2. Noting that $\|w^0\|_* = c^\top w^0 = 1$ and $x^0 \in \arg \min_{x \in \mathcal{X}} (w^0)^\top f(x)$, in [37, Algorithm 2], the initial outer approximation for \mathcal{D}^H is

$$\mathcal{D}_0^H = \{(t^\top, s)^\top \in \mathbb{R}^{q-1} \times \mathbb{R} \mid c^*(t) \in C^+, s \leq c^*(t)^\top f(x^0)\} = H_{f(x^0)}.$$

By Lemma 6.1, we obtain $\mathcal{D}_0 = \hat{T}^{-\top} S [\text{cone}(\mathcal{D}_0^H \times \{1\})]$. Now, assume for some $k \geq 1$, $\mathcal{D}_k = \hat{T}^{-\top} S [\text{cone}(\mathcal{D}_k^H \times \{1\})]$ holds. Then, for any vertex $(t^\top, s)^\top$ of \mathcal{D}_k^H , $\hat{T}^{-\top} S(t^\top, s, 1)^\top$ is an extreme direction of \mathcal{D}_k . Moreover, for an extreme direction $(w^\top, \alpha)^\top$ of \mathcal{D}_k , there exists a vertex $v = (t^\top, s)^\top$ of \mathcal{D}_k^H such that $\frac{(w^\top, \alpha)^\top}{\|w\|_*} = \frac{\hat{T}^{-\top} S(t^\top, s, 1)^\top}{\|P\hat{T}^{-\top} S(t^\top, s, 1)^\top\|_*}$.

In Algorithm 2, we consider an extreme direction of \mathcal{D}_k , say $(w^\top, \alpha)^\top$ with $\|w\|_* = 1$ and solve (WS(w)). Assume that $v = (t^\top, s)^\top \in \mathcal{D}_k^H$ is a vertex satisfying $(w^\top, \alpha)^\top = \frac{\hat{T}^{-\top} S(v^\top, 1)^\top}{\|P\hat{T}^{-\top} S(v^\top, 1)^\top\|_*}$. Note that with the specified norm, we have $\|P\hat{T}^{-\top} S(t^\top, s, 1)^\top\|_* = c^\top P\hat{T}^{-\top} S(t^\top, s, 1)^\top = c^\top T^{-\top} (t^\top, 1)^\top = e_q^\top (t^\top, 1)^\top = 1$. Then, we have $\hat{T}^{-\top} S(t^\top, s, 1)^\top = (w^\top, \alpha)^\top$, hence $(t^\top, 1, s)^\top = S(t^\top, s, 1)^\top = \hat{T}^\top (w^\top, \alpha)^\top = (w^\top T, \alpha)^\top$. In particular, $(t^\top, 1)^\top = T^\top w$ holds. This implies $c^*(t) = T^{-\top} (t^\top, 1)^\top = w$ and $s = \alpha$. In particular, both algorithms solve the same weighted sum scalarization problem (WS(w)), hence may find the same minimizer x^w . Moreover, both update the current outer approximation if $\alpha - p^w > \epsilon$ holds. In that case, the updated outer approximations are $\mathcal{D}_{k+1}^H = \mathcal{D}_k^H \cap H_{f(x^w)}$ and $\mathcal{D}_{k+1} = \mathcal{D}_k \cap \bar{H}_{f(x^w)}$, respectively. Then, by Lemma 6.1, we have

$$\begin{aligned} \hat{T}^{-\top} S [\text{cone}(\mathcal{D}_{k+1}^H \times \{1\})] &= \hat{T}^{-\top} S [\text{cone}((\mathcal{D}_k^H \cap H_{f(x^w)}) \times \{1\})] \\ &= \hat{T}^{-\top} S [\text{cone}((\mathcal{D}_k^H \times \{1\}) \cap (H_{f(x^w)} \times \{1\}))] \\ &= \hat{T}^{-\top} S [\text{cone}(\mathcal{D}_k^H \times \{1\}) \cap \text{cone}(H_{f(x^w)} \times \{1\})] \\ &= \mathcal{D}_k \cap \bar{H}_{f(x^w)} = \mathcal{D}_{k+1}. \end{aligned}$$

We have shown by induction that $\mathcal{D}_k = \hat{T}^{-\top} S [\text{cone}(\mathcal{D}_k^H \times \{1\})]$ hold for all $k \geq 0$ as long as the algorithms consider the extreme directions, respectively the corresponding vertices, in the same order. In that case, \mathcal{X} returned by both algorithms would also be the same. We conclude that by considering a norm satisfying $\|w\|_* = c^\top w$, the proposed geometric dual algorithm recovers [37, Algorithm 2].

7 Numerical examples

We implement the primal and dual algorithms given in Section 5 using MATLAB.² For problems (P(v)) and (WS(w)), we employ CVX, a framework for specifying and solving convex programs [24, 25]. For solving vertex enumeration problems within the algorithms, we use *bensolve tools* [7, 38, 39].

¹For instance, $\|w\|_* := \inf\{c^\top(v+u) \mid v, u \in C^+, v-u=w\}$ would satisfy this.

²We run the experiments in Section 7.2 on a computer with i5-8265U CPU and 8 GB RAM, and the ones in Section 7.3 on a computer with i5-10310U CPU and 16 GB RAM.

7.1 Proximity measures

We measure the performance of the primal and dual algorithms using two indicators. The distance between \mathcal{P} and the furthest point from the outer approximation \mathcal{P}_o returned by the algorithm to \mathcal{P} is referred to as the *primal error indicator*. If the recession cone of \mathcal{P}_o is C , then the primal error indicator is nothing but the Hausdorff distance between \mathcal{P}_o and \mathcal{P} , see [1, Lemma 5.3]. The primal error indicator of an algorithm is calculated by solving $(P(v))$ for the set \mathcal{V}_o of the vertices of \mathcal{P}_o . For each $v \in \mathcal{V}_o$, we find an optimal solution z^v and the corresponding optimal value $\|z^v\|$ by solving $(P(v))$. Then, the primal error indicator is calculated by $\text{PE} := \max\{\|z^v\| \mid v \in \mathcal{V}_o\}$.

Note that the norm used in the definition of PE is the same (primal) norm that is used in the algorithm. Hence, this indicator depends on the choice of the norm. Motivated by this, we define a *hypervolume indicator for CVOPs* which is free of norm-biasedness. Recall that, for a multiobjective optimization problem, the hypervolume of a set $S \subseteq \mathbb{R}^q$ of points with respect to a reference point $r \in \mathbb{R}^q$ is computed as $\Lambda(\{y \in \mathbb{R}^q \mid s \leq_{\mathbb{R}_+^q} y \leq_{\mathbb{R}_+^q} r, s \in S\})$, where Λ is the Lebesgue measure on \mathbb{R}^q [58]. Different from this hypervolume measure, which can be used for convex as well as nonconvex multiobjective optimization problems, we use the fact that (P) is a convex problem and the underlying order relation is induced by cone C . To that end, we define the hypervolume of a set $S \subseteq \mathbb{R}^q$ with respect to a bounding polytope $\mathcal{Q} \subseteq \mathbb{R}^q$ as $\text{HV}(S, \mathcal{Q}) := \Lambda((\text{conv } S + C) \cap \mathcal{Q})$. Let $\mathcal{P}_o, \mathcal{P}_i$ be, respectively, the outer and inner approximations of \mathcal{P} returned by an algorithm and $\mathcal{V}_o, \mathcal{V}_i$ be the set of vertices of them. Then, we compute the *hypervolume indicator* by $\text{HV} := \left(\frac{\text{HV}(\mathcal{V}_o, \mathcal{Q}) - \text{HV}(\mathcal{V}_i, \mathcal{Q})}{\text{HV}(\mathcal{V}_o, \mathcal{Q})}\right) \times 100$, where $\mathcal{Q} \subseteq \mathbb{R}^q$ is a polytope satisfying $\mathcal{V}_o \cup \mathcal{V}_i \subseteq \mathcal{Q}$. Suppose that the problem is solved by finitely many algorithms and let $A \subseteq \mathbb{R}^q$ be the set of all vertices of the outer and inner approximations returned by *all* algorithms. For our computational tests, in order to have a fair comparison, we fix the polytope \mathcal{Q} such that $A \subseteq \mathcal{Q}$. For this, we set $\mathcal{Q} := \bigcap_{j=1}^J \{y \in \mathbb{R}^q \mid (w^j)^\top y \leq \max_{a \in A} (w^j)^\top a\}$, where w^1, \dots, w^J are the generating vectors of C^+ . Note that a smaller hypervolume indicator is more desirable in terms of an algorithm's performance since it means less of a difference between the inner and outer sets.

7.2 Computational results on Algorithms 1 and 2

We assess the performance of the primal and dual algorithms by solving randomly generated problem instances. A problem structure that is simple and versatile is required for scaling purposes, in terms of both the decision space and the image space. To this end, we work with a linear objective function and a quadratic constraint as described in Example 7.1.

Example 7.1. *Consider the problem*

$$\text{minimize } f(x) = A^\top x \text{ with respect to } \leq_{\mathbb{R}_+^q} \text{ subject to } x^\top P x - 1 \leq 0,$$

where $P \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix, and $A \in \mathbb{R}_+^{n \times q}$. In the computational experiments, we generate A and P as instances of random matrices.³

³More precisely, we generate A as the instance of a random matrix with independent entries having the uniform distribution over $[0, 50]$. To construct P , we first create a matrix $U \in \mathbb{R}^{n \times n}$ following the same

7.2.1 Results for multiobjective problem instances

This subsection provides the results obtained by solving randomly generated instances of Example 7.1. The computational results are presented in Table 1, which shows the stopping criteria (Stop), number of optimization problems solved (Opt), time taken to solve optimization problems (T_{opt}), number of vertex enumeration problems solved (En), time taken to solve vertex enumeration problems (T_{en}), and total runtime of the algorithm (T) in terms of seconds. As a measure of efficiency for the algorithms, we also provide the average runtime spent per optimization problem ($T_{\text{opt}}/\text{Opt}$) and total runtime per weak minimizer in $\bar{\mathcal{X}}$ ($T/|\bar{\mathcal{X}}|$).

The performance indicators for the tests are primal error (PE) and hypervolume (HV). For the computations of these, we take the following: For Algorithm 1, $\mathcal{P}_o = \mathcal{P}_{\bar{k}}$, where \bar{k} is the iteration number at which the algorithm terminates; for Algorithm 2, $\mathcal{P}_o = \bigcap_{w \in \bar{W}} \mathcal{H}(w, p^w)$ where \bar{W} is the solution to (D) that Algorithm 2 returns. For both algorithms, $\mathcal{P}_i = \text{conv } f(\bar{\mathcal{X}}) + C$, where $\bar{\mathcal{X}}$ is the solution to (P) that the corresponding algorithm returns.

We take $q = 3$. The results for various choices of the dimension of the decision space (n) can be found in Table 1. For each value of n , 20 random problem instances are generated using the structure of Example 7.1 and solved by both Algorithms 1 and 2 twice with different stopping criteria. The averages of the results and performance indicators of the 20 instances are presented in the table in corresponding cells.

n	Alg	Stop	Opt	T_{opt}	$T_{\text{opt}}/\text{Opt}$	En	T_{en}	T	$T/ \bar{\mathcal{X}} $	PE	HV
10	1	$\epsilon_1 = 0.5000$	52.35	19.92	0.3801	5.05	0.09	46.41	0.91	0.4209	3.5805
	2	$\epsilon_2 = 0.2887$	86.25	29.03	0.3363	4.40	0.20	71.36	70.84	0.1106	1.1973
	1	$\epsilon_3 = 0.1106$	232.80	85.32	0.3655	6.90	0.16	200.13	0.92	0.1052	1.2846
	2	$T = 46.41$	60.70	20.32	0.3336	4.05	0.24	48.43	0.80	0.2244	1.8735
15	1	$\epsilon_1 = 0.5000$	70.85	26.98	0.3777	5.60	0.09	63.54	0.93	0.4694	2.6458
	2	$\epsilon_2 = 0.2887$	105.15	34.87	0.3311	4.70	0.26	84.58	0.80	0.1033	0.7146
	1	$\epsilon_3 = 0.1033$	295.30	108.74	0.3677	7.45	0.21	251.56	0.91	0.0991	0.5690
	2	$T = 63.54$	81.75	27.01	0.3307	4.40	0.36	65.51	0.81	0.2022	1.0703
20	1	$\epsilon_1 = 0.5000$	65.45	24.62	0.3797	5.50	0.09	57.12	0.92	0.4569	3.4845
	2	$\epsilon_2 = 0.2887$	101.95	33.66	0.3302	4.65	0.24	81.31	0.80	0.1052	1.0783
	1	$\epsilon_3 = 0.1052$	284.25	104.18	0.3672	7.35	0.19	238.73	0.91	0.1020	0.7881
	2	$T = 57.12$	74.40	24.63	0.3324	4.30	0.29	59.06	0.80	0.1996	1.5307
25	1	$\epsilon_1 = 0.5000$	85.95	32.63	0.3786	5.95	0.12	82.51	0.98	0.4650	2.3063
	2	$\epsilon_2 = 0.2887$	139.25	46.13	0.3306	4.95	0.34	112.65	0.81	0.1071	0.6154
	1	$\epsilon_3 = 0.1071$	368.60	137.91	0.3736	7.70	0.26	340.28	1.01	0.1034	0.5063
	2	$T = 82.51$	106.10	35.02	0.3297	4.80	0.56	84.55	0.80	0.1878	0.8167
30	1	$\epsilon_1 = 0.5000$	95.70	36.59	0.3818	6.00	0.12	91.27	0.99	0.4690	2.2715
	2	$\epsilon_2 = 0.2887$	150.70	51.13	0.3386	5.15	0.43	131.15	0.87	0.1066	0.6110
	1	$\epsilon_3 = 0.1066$	452.70	170.22	0.3751	8.05	0.32	419.02	1.02	0.1046	0.5420
	2	$T = 91.27$	109.00	36.89	0.3381	4.60	0.48	93.41	0.86	0.2272	0.8947

Table 1: Results of randomly generated problems.

In Table 1, the first two rows for each value of n show the results of the primal and dual algorithms when the given ϵ_i value is fed to the algorithms as stopping criterion. It can be seen that the ϵ_i values that are used in the algorithms are different. We run Algorithm 1 and obtain a weak ϵ_1 -solution to problem (P). When working with Algorithm 2, to obtain

procedure as for A . We assume that U has at least one nonzero entry, which occurs with probability one. Then, $S := (U + U^T)/2$ is a symmetric matrix and we calculate its diagonalization as $S = Q\bar{D}Q^T$, where $\bar{D} \in \mathbb{R}^{n \times n}$ is the diagonal matrix of the eigenvalues of S , and Q is the orthogonal matrix of the corresponding eigenvectors. Denoting by $D \in \mathbb{R}^{n \times n}$ the diagonal matrix whose entries are the absolute values of the entries of \bar{D} , we let $P := QDQ^T$, which is guaranteed to be symmetric and positive definite.

also a weak ϵ_1 -solution to problem (P), we take $\epsilon_2 = \epsilon_1 \min_{\lambda \in \Delta^{J-1}} \|\sum_{j=1}^J \lambda_j w^j\|_*$ based on Proposition 5.11. As a result, we take $\epsilon_1 = 0.5$ for Algorithm 1 and $\epsilon_2 = 0.2887$ for Algorithm 2 as stopping criteria. We observe from the first two rows for each n in Table 1 that Algorithm 1 stops in shorter runtime; however, Algorithm 2 returns smaller primal error and hypervolume indicators.

For further comparison of the algorithms, we also run each algorithm with different stopping criteria. In the third row, we aim to observe the runtime it takes for Algorithm 1 to reach similar primal error that Algorithm 2 returns. Therefore, the PE value in the second row is fed to Algorithm 1 as stopping criterion. Finally, in the fourth row, we aim to observe the primal error and hypervolume indicators of Algorithm 2 when it is terminated after a similar runtime as of Algorithm 1 from the first row. Therefore, T from the first row is fed to Algorithm 2 as stopping criterion.⁴

From Table 1, one can observe that Algorithms 1 and 2 in the first two rows are not comparable. Indeed, while Algorithm 1 in the first row has shorter runtime for each n value, it also gives larger hypervolume results, which indicates worse performance in comparison with Algorithm 2 in the second row. The main reason is that, when Algorithm 2 is run with a stopping criterion that guarantees obtaining a weak ϵ_1 -solution to (P), the solution that it returns has much higher proximity, for instance, compare $\epsilon_2 = 0.2887$ and PE= 0.1106 in the second row of Table 1.⁵ On the other hand, when we set ϵ_3 in a way that Algorithms 1 and 2 return similar PE values, Algorithm 1 may return slightly better HV results but it requires higher runtime.

In order to have further insights, we analyze the results of Algorithm 1 from the first row and Algorithm 2 from the fourth row in more detail, since they have similar runtimes. Similarly, as they yield similar PE values, we analyze the results of Algorithm 1 from the third row and Algorithm 2 from the second row. In Figure 1, the plots of the PE and HV values for rows one and four of Table 1 are shown (first two figures). One can see that Algorithm 2 has consistently better performance for each value of n in terms of both primal error and hypervolume indicator. Moreover, the plots of total runtimes and HV values corresponding to rows two and three of Table 1 can be seen together with the PE values on the right vertical axis (last two figures). We observe that the difference between the primal error indicators of both algorithms are very similar to each other, although, Algorithm 1 has smaller primal error as expected. With very similar primal error indicators, we observe that Algorithm 2 has around half of the runtime of Algorithm 1. In line with the primal error results, Algorithm 1 has a better HV value than Algorithm 2.

7.2.2 Results for different ordering cones and norms

In order to test the performance of the algorithms on problems with different ordering cones and with different norms employed in (P(v)) and (D(v)), we design some additional experi-

⁴Note that the actual termination time is slightly higher than the predetermined time limit as we check the time only at the beginning of the loop in the implementation and it takes a couple of more seconds to terminate the algorithm afterwards.

⁵After running Algorithm 2 with the predetermined ϵ_2 value (to get a weak ϵ_1 -solution for (P)), in order to compute the realized PE for (P), we solve (P(v)) problems as explained in Section 7.1. Instead, in order to compute an upper bound for the PE value (which would be a better bound than ϵ_1), one can also use Proposition 5.12.

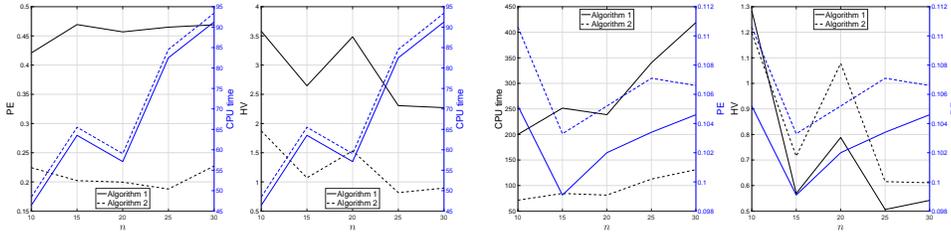


Figure 1: Example 7.1: Average PE (first) and HV (second) values of 20 random instances under nearly equal runtime (rows one and four of Table 1); average CPU time (third) and HV (fourth) values of these instances when the dual algorithm uses the PE of the primal algorithm as approximation error (rows two and three of Table 1).

ments. Although the original stopping criteria in Algorithms 1 and 2 are based on ϵ -closeness, we implement an alternative stopping criterion for the rest of our analysis. In particular, motivated by our analysis in Section 7.2.1, we set a predefined time limit for the algorithms. Note that since ϵ -closeness depends on the choice of the norm used in the scalarization models, using runtime as the stopping criterion results in a fair comparison of the algorithms when considering different norms.

We denote the nonnegative cone by $C_1 = \mathbb{R}_+^3$. The non-standard ordering cones that will be used throughout are⁶:

$$C_2 = \text{conv cone}\{(4, 2, 2)^\top, (2, 4, 2)^\top, (4, 0, 2)^\top, (1, 0, 2)^\top, (0, 1, 2)^\top, (0, 4, 2)^\top\},$$

$$C_3 = \text{conv cone}\{(-1, -1, 3)^\top, (2, 2, -1)^\top, (1, 0, 0)^\top, (0, -1, 2)^\top, (-1, 0, 2)^\top, (0, 1, 0)^\top\}.$$

We consider the 20 instances of Example 7.1 that were generated randomly for our analysis in Section 7.2.1. We solve all instances under the runtime limit of 50 seconds. We consider ordering cones C_1, C_2, C_3 and ℓ_p norms for $p \in \{1, 2, \infty\}$. While solving some of the instances with different ordering cones or different norms, we encountered some errors in the solvers that we have employed in the algorithms. Hence, we consider a subset of instances which can be solved in all settings and we indicate the size of this subset in a separate column in Table 2 (Size).

First, we fix the ℓ_2 norm and solve the problem instances with respect to the ordering cones C_1, C_2 and C_3 ; second, we fix the ordering cone as C_1 and solve the problem instances where we take the norms in $(P(v))$ as ℓ_1, ℓ_2 and ℓ_∞ , see the left and the right columns of Table 2, respectively. In order to summarize the results, we plot the average PE and HV values obtained by the algorithms, see Figures 2 and 3. From these figures, we observe that for all considered ordering cones and norms, Algorithm 2 has better performance in terms of both PE and HV under time limit.

7.3 Comparison with algorithms from the literature

We compare the performance of Algorithms 1 and 2 with similar ones from the literature; in particular with the following algorithms which guarantee returning polyhedral inner and outer approximations to the upper image: the primal (LRU-P) and dual (LRU-D) algorithms

⁶These cones are taken from [1].

n	Size	Cone	Alg	Opt	T_{opt}	T_{opt}/Opt	En	T_{en}	T/\mathcal{X}	PE	HV	Size	p	Alg	Opt	T_{opt}	T_{opt}/Opt	En	T_{en}	T/\mathcal{X}	PE	HV
10	15	C_1	1	57.47	21.60	0.3763	5.00	0.09	0.93	0.4024	1.9112	17	1	53.71	21.09	0.3931	5.00	0.08	1.00	0.5665	2.1873	
			2	60.73	20.42	0.3363	4.00	0.16	0.86	0.2527	1.0413		2	57.35	19.79	0.3450	4.00	0.17	0.91	0.3473	1.2154	
		C_2	1	60.60	22.65	0.3743	3.80	0.07	0.92	0.8190	1.5472		1	56.00	21.14	0.3777	5.00	0.07	0.95	0.3973	2.0040	
			2	60.73	20.49	0.3376	3.80	0.35	0.86	0.4035	0.8266		2	57.88	19.98	0.3453	4.00	0.16	0.90	0.2574	1.1113	
		C_3	1	61.20	22.73	0.3719	3.93	0.07	0.91	0.1337	2.5455		1	53.41	21.23	0.3976	5.06	0.07	0.99	0.2826	1.7489	
			2	60.20	20.40	0.3392	3.93	0.39	0.87	0.0802	1.7527		2	58.18	20.07	0.3450	4.00	0.17	0.90	0.1634	1.0354	
15	12	C_1	1	54.92	21.38	0.3908	5.00	0.10	0.97	0.5544	2.3338	17	1	57.65	22.04	0.3831	5.00	0.08	0.92	0.7394	2.1807	
			2	60.17	20.44	0.3401	4.00	0.17	0.87	0.2843	1.3300		2	63.24	20.72	0.3278	4.00	0.16	0.82	0.3808	0.8593	
		C_2	1	58.58	22.44	0.3839	3.67	0.07	0.95	1.1556	2.2875		1	61.82	22.47	0.3637	5.00	0.07	0.86	0.6205	2.2184	
			2	60.25	20.49	0.3405	3.92	0.41	0.87	0.4582	0.9734		2	63.88	20.83	0.3261	4.00	0.16	0.82	0.2966	1.0616	
		C_3	1	60.33	22.81	0.3785	4.00	0.07	0.92	0.1754	2.2426		1	60.94	22.54	0.3700	5.18	0.07	0.87	0.3284	2.0562	
			2	61.33	20.72	0.3380	4.00	0.42	0.85	0.0958	1.6479		2	64.06	20.90	0.3264	4.00	0.16	0.81	0.1807	1.2320	
20	15	C_1	1	58.40	22.36	0.3837	5.00	0.09	0.91	0.4571	2.8328	18	1	61.17	22.51	0.3686	5.00	0.08	0.87	0.7294	3.0080	
			2	62.67	21.24	0.3390	4.00	0.16	0.83	0.2631	1.5824		2	63.83	20.90	0.3274	4.00	0.15	0.82	0.3499	1.2593	
		C_2	1	60.73	23.10	0.3809	3.80	0.07	0.91	0.9700	2.0552		1	60.11	22.12	0.3684	5.00	0.07	0.88	0.4588	2.6591	
			2	61.80	20.88	0.3382	4.00	0.41	0.85	0.4777	1.3532		2	64.00	20.98	0.3279	4.00	0.16	0.81	0.2757	1.4626	
		C_3	1	62.33	23.39	0.3756	4.00	0.07	0.89	0.1390	1.9080		1	59.28	22.21	0.3753	5.00	0.07	0.90	0.2694	2.6331	
			2	62.27	21.06	0.3384	4.00	0.41	0.84	0.0953	1.3192		2	64.00	21.10	0.3300	4.00	0.15	0.81	0.1729	1.5567	
25	12	C_1	1	54.58	21.19	0.3895	5.00	0.08	0.98	0.5988	2.2838	14	1	58.79	21.74	0.3703	5.00	0.07	0.90	0.9830	2.3330	
			2	61.08	20.54	0.3366	4.00	0.16	0.85	0.3339	1.3099		2	64.64	21.23	0.3286	4.00	0.16	0.81	0.4813	0.8697	
		C_2	1	56.67	21.77	0.3845	3.58	0.06	0.98	1.7656	2.6061		1	57.93	21.49	0.3714	5.00	0.07	0.91	0.6971	2.3461	
			2	61.75	20.79	0.3369	4.00	0.41	0.85	0.5857	1.4343		2	64.79	21.25	0.3282	4.00	0.15	0.81	0.3578	1.2700	
		C_3	1	57.67	21.89	0.3799	3.83	0.07	0.96	0.2698	2.4895		1	57.36	21.60	0.3767	5.00	0.06	0.92	0.4166	2.6032	
			2	62.17	20.88	0.3360	4.00	0.41	0.84	0.1346	1.5884		2	64.71	21.26	0.3286	4.00	0.16	0.81	0.2241	1.4588	
30	11	C_1	1	55.45	21.11	0.3829	5.00	0.08	0.98	0.6844	2.8334	16	1	57.44	21.51	0.3751	5.00	0.08	0.93	1.0316	2.1772	
			2	59.55	20.08	0.3375	4.00	0.16	0.88	0.4049	1.5810		2	59.75	20.08	0.3363	4.00	0.16	0.88	0.5999	0.8015	
		C_2	1	58.91	22.26	0.3783	3.73	0.07	0.95	1.9284	2.1519		1	55.63	20.98	0.3777	5.00	0.07	0.96	0.7395	2.5195	
			2	59.36	20.03	0.3378	3.91	0.43	0.88	0.7094	1.1776		2	60.00	20.11	0.3355	4.00	0.16	0.87	0.4591	1.3388	
		C_3	1	60.09	22.50	0.3747	3.82	0.07	0.93	0.2666	2.6265		1	56.00	21.35	0.3815	5.00	0.07	0.95	0.4387	2.4787	
			2	59.73	20.18	0.3382	3.91	0.39	0.87	0.1357	1.6780		2	60.69	20.29	0.3345	4.00	0.16	0.86	0.2922	1.4639	

Table 2: Results for randomly generated instances of Example 7.1 with different ordering cones (left) and with different norms used in $(P(v))$ (right), when the algorithms are run for $T=50$ seconds.

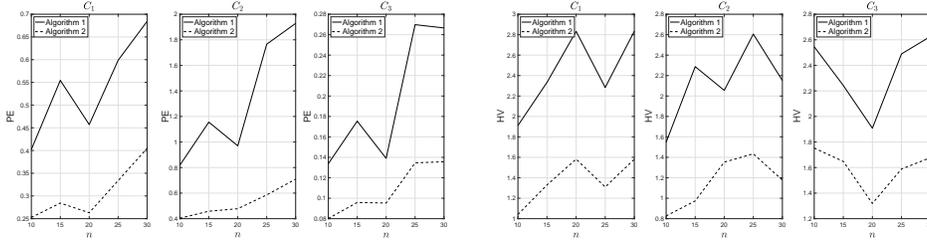


Figure 2: Average primal error (first group) and HV (second group) values of random instances of Example 7.1 for ordering cones C_1 (left), C_2 (middle) and C_3 (right) when the algorithms are run under time limit of 50 seconds.

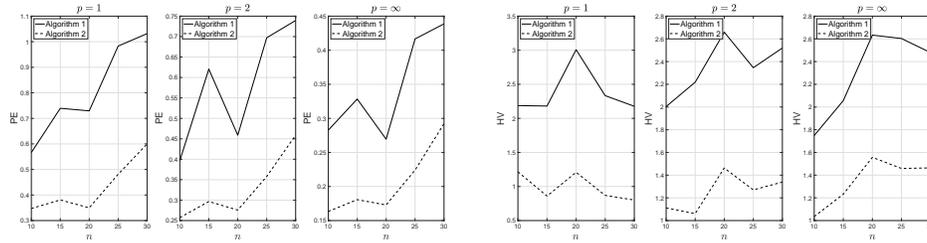


Figure 3: Average primal error (first group) and HV (second group) values of random instances of Example 7.1 with ℓ_p norms for $p = 1$ (left), $p = 2$ (middle) and $p = \infty$ (right), when the algorithms are run for 50 seconds.

from [37] and the (primal) algorithm (DLSW) from [13].⁷

For the first set of experiments we use the hundred randomly generated problem instances from Section 7.2.1. We solve each instance under runtime limits of 50 seconds and 100 seconds and compare the performances of Algorithms 1, 2, LRU-P, LRU-D and DLSW via empirical cumulative distribution functions of the proximity measures, see Figure 4. In [11], when comparing different solvers with respect to CPU times, they scale the data points by the minimum over different solvers and plot the corresponding empirical cumulative distributions, which are called *performance profiles*.

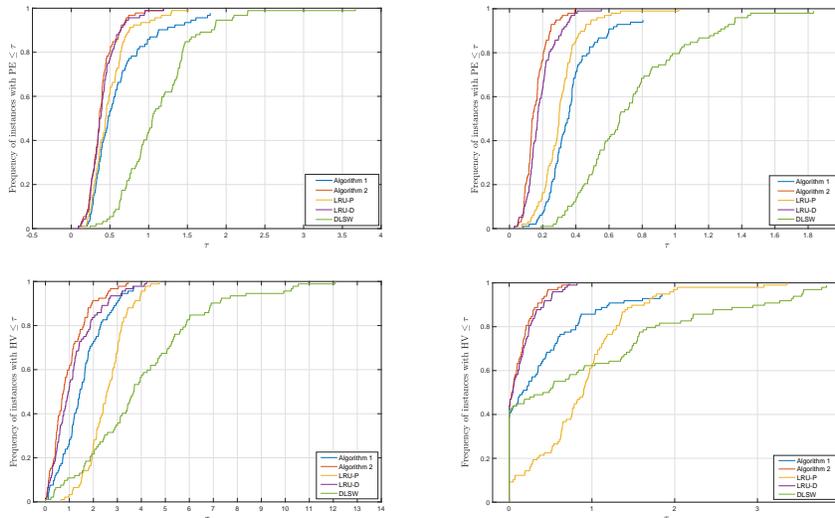


Figure 4: Empirical cumulative distribution functions for PE (first row) and HV (second row) values of random instances from Section 7.2.1 when the algorithms are run under time limit of 50 seconds (left) and 100 seconds (right).

From Figure 4, we see that the dual algorithms perform better than the primal ones in HV and PE under fixed run time. Moreover, Algorithm 2 performs slightly better than LRU-D.

Next, we compare the algorithms over different examples from the literature. Example 7.2 is a special case of Example 7.1, which can be also seen in [16, 37]. In Example 7.3, the objective functions are nonlinear while the constraints are linear, see [16, Examples 5.8], [42]. We solve these examples for different norms and ordering cones as in Section 7.2.2 under fixed runtime of 50 seconds. Since the objective function of Example 7.3 is quadratic, it is not C -convex for $C = C_2$, hence we solve it only with ordering cones C_1 and C_3 .

Example 7.2. We consider the following problem

$$\begin{aligned} & \text{minimize } f(x) = x \text{ with respect to } \leq_C \\ & \text{subject to } \|x - e\|_2 \leq 1, x \in \mathbb{R}^3. \end{aligned}$$

⁷We use MATLAB implementations of these algorithms that were also used in [37] and [33], respectively.

Example 7.3. Let $a^1 = (1, 1)^\top$, $a^2 = (2, 3)^\top$, $a^3 = (4, 2)^\top$. Consider

$$\begin{aligned} & \text{minimize } f(x) = (\|x - a^1\|_2^2, \|x - a^2\|_2^2, \|x - a^3\|_2^2) \text{ with respect to } \leq_C \\ & \text{subject to } x_1 + 2x_2 \leq 10, 0 \leq x_1 \leq 10, 0 \leq x_2 \leq 4, x \in \mathbb{R}^2. \end{aligned}$$

Table 3 shows the Opt, En, $|\bar{\mathcal{X}}|$, PE and HV values that the algorithms return when run for 50 seconds. For both examples the minimum HV values are attained by Algorithm 2 in each setting. The same holds true also for PE values for Example 7.3. However, the PE values returned by the algorithms are comparable for Example 7.2.

Cone	p	Alg	Example 7.2					Example 7.3				
			Opt	En	$ \bar{\mathcal{X}} $	PE	HV	Opt	En	$ \bar{\mathcal{X}} $	PE	HV
C_1	1	1	38	5	38	0,0339	1,1788	31	4	31	0,4289	0,00507
		2	38	4	38	0,0354	0,9296	36	4	36	0,2046	0,00130
		LRU-P	41	5	41	0,0319	2,5017	33	4	33	0,4288	0,35351 ⁸
		LRU-D	40	5	34	0,0354	1,1046	34	4	32	0,2790	0,00174
		DLSW	46	11	13	0,0898	3,7211	-	-	-	-	-
	2	1	39	5	39	0,0267	3,9998	31	4	31	0,3990	0,00451
		2	41	4	41	0,0259	3,9608	36	4	36	0,1762	0,00129
		LRU-P	44	5	44	0,0266	4,0791	34	4	34	0,3989	0,35351
		LRU-D	40	5	34	0,0259	4,7979	33	4	31	0,2672	0,00174
		DLSW	44	13	15	0,0824	10,3820	44	8	10	0,8495	0,01205
	∞	1	34	5	34	0,0183	3,9642	27	4	27	0,3126	0,00003
		2	37	4	37	0,0189	3,6921	36	4	36	0,1206	0,00001
		LRU-P	37	5	37	0,0183	4,1461	34	4	34	0,3124	0,04624
		LRU-D	37	5	31	0,0189	4,3730	37	5	34	0,1206	0,00001
		DLSW	42	9	11	0,0635	12,5425	45	10	12	0,3155	0,00006
C_3	2	1	36	3	36	0,0166	2,3381	31	3	31	0,1964	2,08492
		2	41	3	41	0,0148	2,1469	36	3	36	0,1346	1,66544
		LRU-P	40	3	40	0,0541	4,1121	30	3	30	0,1964	4,38963
		LRU-D	35	4	33	0,0274	3,0475	30	4	24	0,3065	3,34918
		DLSW	45	9	14	0,0541	4,1913	42	9	14	0,1964	3,58336
C_2	2	1	36	3	36	0,1926	6,0955					
		2	41	3	41	0,1150	4,1466					
		LRU-P	44	3	44	0,0498	4,5592					
		LRU-D	36	4	33	0,1393	8,7210					
		DLSW	43	9	14	0,1064	9,0697					

Table 3: Results for Examples 7.2 and 7.3 with different ordering cones and with different norms used in $(P(v))$, when the algorithms are run for T=50 seconds.

Acknowledgments

This work is supported by the 3501 program of TÜBİTAK (Scientific & Technological Research Council of Turkey), Project No. 118M479. The authors thank three anonymous reviewers and an associate editor, as well as Andreas Löhne and Özlem Karsu, for their valuable feedback and suggestions for improvement on earlier versions of the paper. They also thank İrem Nur Keskin for sharing MATLAB implementation of DLSW algorithm which was also used in [33].

⁸In Example 7.2 with cone C_1 , the outer approximations returned by LRU-P contain outlier vertices in the sense that their distances to $f(\mathcal{X})$ are quite large while distances to \mathcal{P} are small. This explains high HV values compared to small PE values.

A (Alternative) Proofs of some results

Proof of Proposition 3.4. Let $((w^i)^\top, \alpha_i)^\top \in \mathcal{D}$ and $\lambda_i \geq 0$ for each $i \in \{1, \dots, n\}$, where $n \in \mathbb{N}$. Since $\mathcal{W} = C^+$ is a convex cone and we have $w^i \in C^+$ for each $i \in \{1, \dots, n\}$, we have $\sum_{i=1}^n \lambda_i w^i \in \mathcal{W}$. Moreover, we have $\alpha_i \leq \inf_{x \in \mathcal{X}} (w^i)^\top f(x)$ for each $i \in \{1, \dots, n\}$, which implies that

$$\sum_{i=1}^n \lambda_i \alpha_i \leq \sum_{i=1}^n \lambda_i \inf_{x \in \mathcal{X}} (w^i)^\top f(x) \leq \inf_{x \in \mathcal{X}} \left(\sum_{i=1}^n \lambda_i w^i \right)^\top f(x).$$

Hence, $\sum_{i=1}^n \lambda_i ((w^i)^\top, \alpha_i)^\top \in \mathcal{D}$. It follows that \mathcal{D} is a convex cone.

Note that the function $w \mapsto p^w$ is continuous as a concave function on \mathcal{W} with finite values. Moreover, $\mathcal{D} \subseteq \mathcal{W} \times \mathbb{R}$ coincides with the hypograph of this function. Hence, \mathcal{D} is a closed set. \square

Proof of Proposition 4.2. (a) Suppose that y is a weakly C -minimal element of \mathcal{P} . By Proposition 3.1, $y = f(\bar{x}) + \bar{c}$ for some $\bar{x} \in \mathcal{X}$ and $\bar{c} \in C$. Moreover, as y is a weakly C -minimal element of \mathcal{P} , we have $y \in \text{bd } \mathcal{P}$ [35, Corollary 1.48]. Then, there exist $\bar{w} \in \mathbb{R}^q$ and $\bar{\alpha} \in \mathbb{R}$ such that $H := \{\tilde{y} \in \mathbb{R}^q \mid \bar{w}^\top \tilde{y} = \bar{\alpha}\}$ supports \mathcal{P} at y . In particular, we have

$$\inf_{\tilde{y} \in \mathcal{P}} \bar{w}^\top \tilde{y} = \bar{w}^\top y = \bar{\alpha}. \quad (17)$$

By Remark 2.1, (17) implies that $\bar{w} \in C^+$. Moreover, using Proposition 3.1 and the fact that $y = f(\bar{x}) + \bar{c}$, (17) can be rewritten as

$$\inf_{x \in \mathcal{X}} \bar{w}^\top f(x) + \inf_{c \in C} \bar{w}^\top c = \bar{w}^\top f(\bar{x}) + \bar{w}^\top \bar{c}.$$

Noting that we have $\inf_{c \in C} \bar{w}^\top c = 0$, $\inf_{x \in \mathcal{X}} \bar{w}^\top f(x) \leq \bar{w}^\top f(\bar{x})$ and $\bar{w}^\top \bar{c} \geq 0$, we obtain $\bar{w}^\top \bar{c} = 0$ and $\bar{w}^\top f(\bar{x}) = \inf_{x \in \mathcal{X}} \bar{w}^\top f(x)$. It follows that $\mathcal{H}^*(f(\bar{x})) = \mathcal{H}^*(y)$ since $y = f(\bar{x}) + \bar{c}$ and $\bar{w}^\top \bar{c} = 0$. Moreover, \bar{x} is an optimal solution to the problem $(\text{WS}(\bar{w}))$. Hence, $H^*(f(\bar{x})) = H^*(y)$ is a supporting hyperplane of \mathcal{D} at $\xi(\bar{w})$ such that $\mathcal{D} \subseteq \mathcal{H}^*(f(\bar{x}))$ by Proposition 4.1. Therefore, $H^*(f(\bar{x})) \cap \mathcal{D}$ is a proper face of \mathcal{D} [46, Section 18, page 162].

To show that $H^*(f(\bar{x})) \cap \mathcal{D}$ is a K -maximal proper face of \mathcal{D} , let $(\tilde{w}^\top, \tilde{\alpha})^\top \in H^*(f(\bar{x})) \cap \mathcal{D}$ be arbitrary. Note that $\varphi(f(\bar{x}), \tilde{w}, \tilde{\alpha}) = \tilde{w}^\top f(\bar{x}) - \tilde{\alpha} = 0$. On the other hand, the fact $\mathcal{D} \subseteq \mathcal{H}^*(f(\bar{x}))$ implies that $\varphi(f(\bar{x}), w, \alpha) = w^\top f(\bar{x}) - \alpha \geq 0$ for each $(w^\top, \alpha)^\top \in \mathcal{D}$. Together, these imply that $(\tilde{w}^\top, \hat{\alpha})^\top \notin \mathcal{D}$ for every $\hat{\alpha} > \tilde{\alpha}$. Hence, $(\tilde{w}^\top, \tilde{\alpha})^\top$ is a K -maximal element of \mathcal{D} .

Conversely, suppose that $H^*(y) \cap \mathcal{D}$ is a K -maximal proper face of \mathcal{D} . Hence, $H^*(y)$ is a supporting hyperplane of \mathcal{D} [46, Section 18, page 162], and we have either $\mathcal{D} \subseteq \mathcal{H}^*(y)$ or $\mathcal{D} \subseteq \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \varphi(y, w, \alpha) \leq 0\}$. We claim that the former relation holds. Indeed, letting $(\tilde{w}^\top, \tilde{\alpha})^\top \in H^*(y) \cap \mathcal{D}$, we have $\tilde{\alpha} \leq \inf_{x \in \mathcal{X}} \tilde{w}^\top f(x)$ and $\varphi(y, \tilde{w}, \tilde{\alpha}) = \tilde{w}^\top y - \tilde{\alpha} = 0$. Then, $(\tilde{w}^\top, \tilde{\alpha} - 1)^\top \in \mathcal{D}$ and $\varphi(y, \tilde{w}, \tilde{\alpha} - 1) = \tilde{w}^\top y - \tilde{\alpha} + 1 > 0$. Hence, $(\tilde{w}^\top, \tilde{\alpha} - 1)^\top \notin \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid \varphi(y, w, \alpha) \leq 0\}$. Therefore, the claim holds and we have $\mathcal{D} \subseteq \mathcal{H}^*(y)$.

Next, we show that $y \in \mathcal{P}$. To get a contradiction, suppose that $y \notin \mathcal{P}$. By separation theorem, there exists $\bar{w} \in \mathbb{R}^q \setminus \{0\}$ such that

$$\bar{w}^\top y < \inf_{\bar{y} \in \mathcal{P}} \bar{w}^\top \bar{y} =: \bar{\alpha}.$$

Note that $\bar{w} \in C^+$ by Remark 2.1. Therefore, we have $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{D} \subseteq \mathcal{H}^*(y)$, which contradicts $\bar{w}^\top y - \bar{\alpha} < 0$. Hence, we have $y \in \mathcal{P}$.

Finally, we show that y is a weakly C -minimal element of \mathcal{P} . To get a contradiction, suppose that there exists $c \in \text{int } C$ with $y - c \in \mathcal{P}$. Without loss of generality, assume that $y - c \in \text{wMin}_C \mathcal{P}$. From the proof of the previous implication, $\mathcal{D} \subseteq \mathcal{H}^*(y - c)$. Let $(w^\top, \alpha)^\top \in H^*(y) \cap \mathcal{D} \subseteq \mathcal{H}^*(y - c)$. Then, we have $w^\top(y - c) - \alpha = w^\top y - w^\top c - \alpha \geq 0$. Note that $w^\top y - \alpha = 0$ since $(w^\top, \alpha)^\top \in H^*(y) \cap \mathcal{D}$. Hence, we have $w^\top y - w^\top c - \alpha = -w^\top c \geq 0$, which contradicts $w \in \mathcal{W}$ and $c \in \text{int } C$. Therefore, y is a weakly C -minimal element of \mathcal{P} .

- (b) Let F^* be a K -maximal proper face of \mathcal{D} . Then, there exists a supporting hyperplane $H^* \subseteq \mathbb{R}^{q+1}$ of \mathcal{D} such that $F^* = H^* \cap \mathcal{D}$ [46, Section 18, page 162]. Then, we may write

$$H^* = \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid a^\top(w^\top, \alpha) = b\}$$

for some $a \in \mathbb{R}^{q+1}, b \in \mathbb{R}$. Without loss of generality, we may assume that

$$\mathcal{H}^* := \{(w^\top, \alpha)^\top \in \mathbb{R}^{q+1} \mid a^\top(w^\top, \alpha)^\top \geq b\} \supseteq \mathcal{D}.$$

Since \mathcal{D} is a convex cone, F^* is also a convex cone [23, Lemma 10.2]. Moreover, F^* is closed as \mathcal{D} is closed by Proposition 3.4. Therefore, $0 \in F^* \subseteq H^*$. Hence, $b = 0$.

Next, we show that $a_{q+1} < 0$. For every $(\bar{w}^\top, \bar{\alpha})^\top \in F^*$, the point $(\bar{w}^\top, \bar{\alpha})^\top$ is a K -maximal element of \mathcal{D} and it holds $a^\top(\bar{w}^\top, \bar{\alpha}) = 0$. Moreover, $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{H}^*$ implies $a^\top(w^\top, \alpha) \geq 0$. For every $\gamma > 0$, since $(\bar{w}^\top, \bar{\alpha} - \gamma)^\top \in \mathcal{D} \subseteq \mathcal{H}^*$, we have $a^\top(\bar{w}^\top, \bar{\alpha} - \gamma) = a^\top(w^\top, \alpha) - \gamma a_{q+1} \geq 0$. Therefore, $a_{q+1} \leq 0$ holds. If $a_{q+1} = 0$, then $a^\top(\bar{w}^\top, \bar{\alpha} - 1)^\top = 0$ implies that $(\bar{w}^\top, \bar{\alpha} - 1)^\top \in F^*$ contradicting the K -maximality of F^* . Therefore, $a_{q+1} < 0$.

By setting

$$y := \left(\frac{-a_1}{a_{q+1}}, \frac{-a_2}{a_{q+1}}, \dots, \frac{-a_q}{a_{q+1}} \right)^\top,$$

we obtain $H^* = H^*(y)$ and $\mathcal{D} \subseteq \mathcal{H}^*(y)$.

Finally, we show that $y \in \mathcal{P}$. Assuming otherwise, there exists $\tilde{w} \in C^+$ such that $\tilde{w}^\top y < \inf_{x \in \mathcal{X}} \tilde{w}^\top f(x) =: \tilde{\alpha}$ by separation arguments. Then, we obtain $(\tilde{w}^\top, \tilde{\alpha})^\top \in \mathcal{D} \setminus \mathcal{H}^*(y)$, which contradicts with $\mathcal{D} \subseteq \mathcal{H}^*(y)$. □

Proof of Proposition 4.3. (a) Suppose that $(w^\top, \alpha)^\top$ is a K -maximal point of \mathcal{D} . Clearly, $w \in C^+$, $\alpha = \inf_{x \in \mathcal{X}} w^\top f(x)$ and $(w^\top, \alpha)^\top = \xi(w)$. Since \mathcal{X} is a compact set, there exists an optimal solution $x^w \in \mathcal{X}$ to $(\text{WS}(w))$. By Proposition 4.1, $H(\xi(w)) = H(w, \alpha)$ is a supporting hyperplane of \mathcal{P} at $f(x^w)$ satisfying $\mathcal{H}(w, \alpha) \supseteq \mathcal{P}$. Then, $H(w, \alpha) \cap \mathcal{P}$

is a proper face of \mathcal{P} [46, Section 18, page 162]. To show that $H(w, \alpha) \cap \mathcal{P}$ is weakly C -minimal, let $\bar{y} \in H(w, \alpha) \cap \mathcal{P}$ be arbitrary. Since $\bar{y} \in H(w, \alpha)$ and $w \in C^+$, we have $\varphi(\bar{y} - c, w, \alpha) = w^\top \bar{y} - w^\top c - \alpha < 0$ for every $c \in \text{int } C$. Note that each $y \in \mathcal{P} \subseteq \mathcal{H}(w, \alpha)$ satisfies $\varphi(y, w, \alpha) = w^\top y - \alpha \geq 0$. Then, $(\{\bar{y}\} - \text{int } C) \cap \mathcal{P} = \emptyset$, hence \bar{y} is weakly C -minimal.

Conversely, suppose that $H(w, \alpha) \cap \mathcal{P}$ is a weakly C -minimal proper face of \mathcal{P} such that $\mathcal{H}(w, \alpha) \supseteq \mathcal{P}$. Then, $H(w, \alpha)$ is a supporting hyperplane of \mathcal{P} [46, Section 18, page 162]. By Remark 2.1, we have $w \in (\text{recc } \mathcal{P})^+ \subseteq C^+$. Moreover, for each $y \in \mathcal{P}$, $\varphi(y, w, \alpha) = w^\top y - \alpha \geq 0$. This implies $\alpha \leq \inf_{y \in \mathcal{P}} w^\top y$, hence $(w^\top, \alpha)^\top \in \mathcal{D}$. On the other hand, let $y = f(x) + c \in H(w, \alpha) \cap \mathcal{P}$ for some $x \in \mathcal{X}$ and $c \in C$. Since $\varphi(y, w, \alpha) = w^\top f(x) + w^\top c - \alpha = 0$, we have $\varphi(y, w, \alpha + \epsilon) = w^\top f(x) + w^\top c - \alpha - \epsilon < 0$ for every $\epsilon > 0$. This implies

$$\alpha + \epsilon > \inf_{x \in \mathcal{X}} w^\top f(x) + \inf_{c \in C} w^\top c = \inf_{x \in \mathcal{X}} w^\top f(x).$$

Since $\epsilon > 0$ is arbitrary, we have $\alpha \geq \inf_{x \in \mathcal{X}} w^\top f(x)$. Together, we obtain $\alpha = \inf_{x \in \mathcal{X}} w^\top f(x)$, which implies that $(w^\top, \alpha)^\top$ is K -maximal.

- (b) Let F be a C -minimal proper face of \mathcal{P} . Then, there exists a supporting hyperplane H of \mathcal{P} such that $F = H \cap \mathcal{P}$ [46, Section 18, page 162]. We may write $H = \{y \in \mathbb{R}^q \mid w^\top y = \alpha\}$ for some $w \in \mathbb{R}^q$ and $\alpha \in \mathbb{R}$, and assume that $\mathcal{P} \subseteq \mathcal{H} := \{y \in \mathbb{R}^q \mid w^\top y \geq \alpha\}$ without loss of generality. By Remark 2.1, we have $w \in (\text{recc } \mathcal{P})^+ \subseteq C^+$. Moreover, as $\mathcal{P} \subseteq \mathcal{H}$, it holds true that $\alpha \leq \inf_{x \in \mathcal{X}} w^\top f(x)$. Hence, $(w^\top, \alpha)^\top \in \mathcal{D}$. □

Alternative proof of Theorem 4.4. First, for a K -maximal proper face F^* of \mathcal{D} , we show that $\Psi(F^*)$ is a weakly C -minimal proper face of \mathcal{P} . By Proposition 4.3 (a), $H(w, \alpha) \cap \mathcal{P}$ is a weakly C -minimal proper face of \mathcal{P} for each $(w^\top, \alpha)^\top \in F^*$. From the definition given by (4), $\Psi(F^*)$ is a weakly C -minimal proper face of \mathcal{P} if it is nonempty. By Proposition 4.2 (b), there exists some $y \in \mathcal{P}$ such that $F^* = H^*(y) \cap \mathcal{D}$. Therefore, for each $(w^\top, \alpha)^\top \in F^*$, we have $(w^\top, \alpha)^\top \in H^*(y)$, equivalently, $y \in H(w, \alpha)$, see (3). Then, $\Psi(F^*)$ is nonempty as $y \in \Psi(F^*)$.

For a weakly C -minimal proper face F of \mathcal{P} , define $\Psi^*(F) := \bigcap_{y \in F} (H^*(y) \cap \mathcal{D})$. To show that $\Psi^*(F)$ is a K -maximal proper face of \mathcal{D} , let $y \in F$. By Proposition 4.2 (a), $H^*(y) \cap \mathcal{D}$ is a K -maximal proper face of \mathcal{D} . Therefore, $\Psi(F^*)$ is a K -maximal proper face of \mathcal{D} , if it is nonempty. From Proposition 4.3 (b), there exists some $(w^\top, \alpha)^\top \in \mathcal{D}$ such that $F = H(w, \alpha) \cap \mathcal{P}$. Therefore, for each $y \in F$, we have $y \in H(w, \alpha)$, equivalently, $(w^\top, \alpha)^\top \in H^*(y)$, see (3). Then, $\Psi^*(F)$ is nonempty as $(w^\top, \alpha)^\top \in \Psi^*(F)$.

In order to show that Ψ is a bijection and $\Psi^{-1} = \Psi^*$, we will show the following two statements:

- (a) $\Psi^*(\Psi(F^*)) = F^*$ for every K -maximal proper face F^* of \mathcal{D} ,
(b) $\Psi(\Psi^*(F)) = F$ for every weakly C -minimal proper face F of \mathcal{P} .

(a) Let F^* be a K -maximal proper face of \mathcal{D} . Assume for a contradiction that $F^* \not\subseteq \Psi^*(\Psi(F^*))$. Let $(w^\top, \alpha)^\top \in F^* \setminus \Psi^*(\Psi(F^*))$. Since $\Psi(F^*)$ is nonempty, $(w^\top, \alpha)^\top \notin \Psi^*(\Psi(F^*))$ means that there exists $\bar{y} \in \Psi(F^*)$ such that $(w^\top, \alpha)^\top \notin H^*(\bar{y}) \cap \mathcal{D}$. This implies $(w^\top, \alpha)^\top \notin H^*(\bar{y})$ since $(w^\top, \alpha)^\top \in \mathcal{D}$. Using (3), we have $\bar{y} \notin H(w, \alpha)$. Therefore $\bar{y} \notin \Psi(F^*)$, a contradiction. Hence, $F^* \subseteq \Psi^*(\Psi(F^*))$. For the reverse inclusion, first note that, from Proposition 4.2 (b), there exists $\bar{y} \in \mathcal{P}$ such that $F^* = H^*(\bar{y}) \cap \mathcal{D}$. Therefore, for each $(w^\top, \alpha)^\top \in F^*$, we have $(w^\top, \alpha)^\top \in H^*(\bar{y})$, equivalently, $\bar{y} \in H(w, \alpha)$, see (3). Hence,

$$\bar{y} \in \bigcap_{(w^\top, \alpha)^\top \in F^*} (H(w, \alpha) \cap \mathcal{P}) = \Psi(F^*).$$

Therefore,

$$\Psi^*(\Psi(F^*)) = \bigcap_{y \in \Psi(F^*)} (H^*(y) \cap \mathcal{D}) \subseteq H^*(\bar{y}) \cap \mathcal{D} = F^*.$$

Hence, the equality $\Psi^*(\Psi(F^*)) = F^*$ holds.

(b) Let F be a weakly C -minimal proper face of \mathcal{P} . Assume for a contradiction that $F \not\subseteq \Psi(\Psi^*(F))$. Let $y \in F \setminus \Psi(\Psi^*(F))$. Then, there exists $(\bar{w}^\top, \bar{\alpha})^\top \in \Psi^*(F)$ such that $y \notin H(\bar{w}, \bar{\alpha}) \cap \mathcal{P}$. This implies $y \notin H(\bar{w}, \bar{\alpha})$ since $y \in \mathcal{P}$. By (3), $(\bar{w}^\top, \bar{\alpha})^\top \notin H^*(y)$, which implies $(\bar{w}^\top, \bar{\alpha})^\top \notin \Psi^*(F)$, a contradiction. Hence, $\Psi(\Psi^*(F)) \supseteq F$. For the reverse inclusion, first note that, by Proposition 4.3 (b), there exists $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{D}$ such that $F = H(\bar{w}, \bar{\alpha}) \cap \mathcal{P}$. Then, for each $y \in F$, we have $y \in H(\bar{w}, \bar{\alpha})$, equivalently, $(\bar{w}^\top, \bar{\alpha})^\top \in H^*(y)$, see (3). Hence,

$$(\bar{w}^\top, \bar{\alpha})^\top \in \bigcap_{y \in F} (H^*(y) \cap \mathcal{D}) = \Psi^*(F).$$

Therefore,

$$\Psi(\Psi^*(F)) = \bigcap_{(w^\top, \alpha)^\top \in \Psi^*(F)} (H(w, \alpha) \cap \mathcal{P}) \subseteq H(\bar{w}, \bar{\alpha}) \cap \mathcal{P} = F.$$

Hence, the equality $\Psi(\Psi^*(F)) = F$ holds. \square

Proof of Lemma 4.11. Let $\tilde{\mathcal{P}} := \{y \in \mathbb{R}^q \mid \forall (w^\top, \alpha)^\top \in \xi(\bar{\mathcal{W}}): \varphi(y, w, \alpha) \geq 0\}$. Since $\bar{\mathcal{D}} \supseteq \xi(\bar{\mathcal{W}})$, we obtain $\tilde{\mathcal{P}} \supseteq \mathcal{P}_{\bar{\mathcal{D}}}$. To show the reverse inclusion, let us fix $y \in \tilde{\mathcal{P}}$. Let $(w^\top, \alpha)^\top \in \bar{\mathcal{D}}$, that is, there exist $n \in \mathbb{N}$, $\lambda_i \geq 0$, $(\tilde{w}_i^\top, \tilde{\alpha}_i)^\top \in \xi(\bar{\mathcal{W}})$ for $i \in \{1, \dots, n\}$, $\beta \geq 0$ such that $(w^\top, \alpha + \beta)^\top = \sum_{i=1}^n \lambda_i (\tilde{w}_i^\top, \tilde{\alpha}_i)^\top$. In particular, $\tilde{w}_i^\top y - \tilde{\alpha}_i \geq 0$ for each $i \in \{1, \dots, n\}$. Then,

$$\varphi(y, w, \alpha) = (w^\top y - \alpha - \beta) + \beta = \sum_{i=1}^n \lambda_i \tilde{w}_i^\top y - \sum_{i=1}^n \lambda_i \tilde{\alpha}_i + \beta = \sum_{i=1}^n \lambda_i (\tilde{w}_i^\top y - \tilde{\alpha}_i) + \beta \geq 0.$$

Since $\varphi(y, w, \alpha) \geq 0$ for each $(w^\top, \alpha)^\top \in \bar{\mathcal{D}}$, we conclude that $y \in \mathcal{P}_{\bar{\mathcal{D}}}$. \square

Proof of Proposition 5.12. To start with the proof, we have the following observation:

$$\mathcal{D}_\epsilon = \text{cone conv} \left((\xi(\bar{\mathcal{W}}) + \epsilon\{e^{g+1}\}) \cup \{-e^{g+1}\} \right).$$

This implies that the set of extreme directions of \mathcal{D}_ϵ is a subset of $(\xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) \cup \{-e^{q+1}\}$. Let $i \in \{1, \dots, T\}$. By [46, Section 18, page 162], we also have

$$\{((w^{i1})^\top, \alpha_{i1})^\top, \dots, ((w^{iJ_i})^\top, \alpha_{iJ_i})^\top\} \subseteq (\xi(\bar{\mathcal{W}}) + \epsilon\{e^{q+1}\}) \cup \{-e^{q+1}\}.$$

In particular, for $j \in \{1, \dots, J_i\}$, we have $w^{ij} \in \bar{\mathcal{W}} \cup \{0\}$ and $\alpha_{ij} = \inf_{x \in \mathcal{X}} (w^{ij})^\top f(x) + \epsilon = p^{w^{ij}} + \epsilon$ if $w^{ij} \neq 0$.

By Proposition 5.11, we have $\mathcal{P} \supseteq \mathcal{P}_\epsilon$. Following similar steps as in the proof of Proposition 5.11, in order to show that $\mathcal{P}_\epsilon + B(0, \tilde{\epsilon}) \supseteq \mathcal{P}$, we assume the contrary. Then, we obtain $\bar{y} \in \mathcal{P}$, $\bar{w} \in \mathbb{R}^q$ with $\|\bar{w}\|_* = 1$ such that

$$\bar{w}^\top \bar{y} + \tilde{\epsilon} < \inf_{y \in \mathcal{P}_\epsilon} \bar{w}^\top y =: \bar{\alpha},$$

and we can check that $(\bar{w}^\top, \bar{\alpha})^\top \in \mathcal{D}_\epsilon$.

By the construction of \mathcal{D}_ϵ , there exists $k \geq 0$ such that $(\bar{w}^\top, \bar{\alpha})^\top + ke^{q+1} \in \text{bd } \mathcal{D}_\epsilon$ is a K -maximal element of \mathcal{D}_ϵ . Hence, there exists $I \in \{1, \dots, T\}$ such that F_I is a K -maximal facet of \mathcal{D}_ϵ and

$$(\bar{w}^\top, \bar{\alpha})^\top + ke^{q+1} \in F_I = \text{cone conv}\{((w^{I1})^\top, \alpha_{I1})^\top, \dots, ((w^{IJ_I})^\top, \alpha_{IJ_I})^\top\}. \quad (18)$$

Using the K -maximality of F_I , we may assume that $w^{Ij} \neq 0$, hence $\alpha_{Ij} = p^{w^{Ij}} + \epsilon$ for each $j \in \{1, \dots, J_I\}$. Then, we can rewrite (18) as

$$(\bar{w}^\top, \bar{\alpha})^\top + ke^{q+1} \in F_I = \text{cone}(\text{conv}\{((w^{I1})^\top, p^{w^{I1}})^\top, \dots, ((w^{IJ_I})^\top, p^{w^{IJ_I}})^\top\} + \epsilon\{e^{q+1}\}).$$

Hence, there exist $\delta \geq 0$ and $\mu \in \Delta^{J_I-1}$ such that

$$(\bar{w}^\top, \bar{\alpha})^\top = \delta \left(\sum_{j=1}^{J_I} \mu_j ((w^{Ij})^\top, p^{w^{Ij}})^\top + \epsilon e^{q+1} \right) - ke^{q+1}. \quad (19)$$

As exactly in the proof of Proposition 5.11, the aim is to show that $(\bar{w}^\top, \bar{\alpha} - \tilde{\epsilon})^\top \in \mathcal{D}$ in order to get a contradiction. For this purpose, it is sufficient to show that

$$\bar{\alpha} - \tilde{\epsilon} \leq \inf_{x \in \mathcal{X}} \bar{w}^\top f(x). \quad (20)$$

Following the same steps as in the proof of Proposition 5.11, one can use (19) in order to obtain $\bar{\alpha} - \delta\epsilon \leq \inf_{x \in \mathcal{X}} \bar{w}^\top f(x)$ as well as

$$\delta\epsilon \leq \frac{\epsilon}{\min_{\lambda \in \Delta^{J_I-1}} \left\| \sum_{j=1}^{J_I} \lambda_j w_{Ij} \right\|_*} = \frac{\epsilon}{f_{\min}^I} \leq \frac{\epsilon}{\min\{f_{\min}^1, \dots, f_{\min}^T\}} = \tilde{\epsilon}.$$

Then, (20) follows from $\bar{\alpha} - \delta\epsilon \leq \inf_{x \in \mathcal{X}} \bar{w}^\top f(x)$ and $\delta\epsilon \leq \tilde{\epsilon}$. \square

References

- [1] Ç. Ararat, F. Ulus, and M. Umer. A norm minimization-based convex vector optimization algorithm. *Journal of Optimization Theory and Applications*, 194(2):681–712, 2022.
- [2] Ç. Ararat and N. Meimanjan. Computation of systemic risk measures: a mixed-integer programming approach. *arXiv preprint arXiv:1903.08367*, 2019.
- [3] P. Armand. Finding all maximal efficient faces in multiobjective linear programming. *Mathematical Programming*, 61:357–375, 1993.
- [4] A. Auger, J. Bader, D. Brockhoff, and E. Zitzler. Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science*, 425:75–103, 2012.
- [5] H. P. Benson. An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem. *Journal of Global Optimization*, 13:1–24, 1998.
- [6] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [7] D. Ciripoi, A. Löhne, and B. Weißing. A vector linear programming approach for certain global optimization problems. *Journal of Global Optimization*, 72(2):347–372, 2018.
- [8] G. Cocchi and M. Lapucci. An augmented Lagrangian algorithm for multi-objective optimization. *Computational Optimization and Applications*, 77:29–56, 2020.
- [9] L. Csirmaz. Using multiobjective optimization to map the entropy region. *Computational Optimization and Applications*, 63(1):45–67, 2016.
- [10] M. De Santis, G. Eichfelder, J. Niebling, and S. Rocktäschel. Solving multiobjective mixed integer convex optimization problems. *SIAM Journal on Optimization*, 30:3122–3145, 2020.
- [11] E. D. Dolan and J. J. More. Benchmarking optimization software with performance profile. *Mathematical Programming*, 91:201–213, 2002.
- [12] D. Dörfler and A. Löhne. Geometric duality and parametric duality for multiple objective linear programs are equivalent. *Journal of Nonlinear and Convex Analysis*, 19(7):1181–1188, 2018.
- [13] D. Dörfler, A. Löhne, C. Schneider, and B. Weißing. A Benson-type algorithm for bounded convex vector optimization problems with vertex selection. *Optimization Methods and Software*, pages 1–21, 2021.
- [14] L. G. Drummond and A. N. Iusem. A projected gradient method for vector optimization problems. *Computational Optimization and Applications*, 28:5–29, 2004.

- [15] M. Ehrgott, A. Löhne, and L. Shao. A dual variant of Benson’s “outer approximation algorithm” for multiple objective linear programming. *Journal of Global Optimization*, 52(4):757–778, 2012.
- [16] M. Ehrgott, L. Shao, and A. Schöbel. An approximation algorithm for convex multi-objective programming problems. *Journal of Global Optimization*, 50(3):397–416, 2011.
- [17] G. Eichfelder. *Adaptive Scalarization Methods in Multiobjective Optimization*. Springer, 2008.
- [18] Z. Feinstein and B. Rudloff. A recursive algorithm for multivariate risk measures and a set-valued Bellman’s principle. *Journal of Global Optimization*, 68:47–69, 2017.
- [19] Z. Feinstein and B. Rudloff. Characterizing and computing the set of Nash equilibria via vector optimization. *arXiv preprint arXiv:2109.14932*, 2021.
- [20] J. Fliege, L. M. Grana Drummond, and B. F. Svaiter. Newton’s method for multiobjective optimization. *SIAM Journal on Optimization*, 20:602–626, 2009.
- [21] E. H. Fukuda and L. G. Drummond. Inexact projected gradient method for vector optimization. *Computational Optimization and Applications*, 54:473–493, 2013.
- [22] E. H. Fukuda and L. G. Drummond. A barrier-type method for multiobjective optimization. *Computational Optimization and Applications*, 69:2471–2487, 2020.
- [23] H. Glöckner. Positive definite functions on infinite-dimensional convex cones. *Memoirs of the American Mathematical Society*, 166(789), 2003.
- [24] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.
- [25] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. <http://cvxr.com/cvx>, 2014.
- [26] B. Grünbaum. *Convex Polytopes*, volume 221. Springer Science & Business Media, 2013.
- [27] A. Hamel, A. Rudloff, and M. Yankova. Set-valued average value at risk and its computation. *Mathematics and Financial Economics*, 7:229–246, 2013.
- [28] A. H. Hamel and A. Löhne. A set optimization approach to zero-sum matrix games with multi-dimensional payoffs. *Mathematical Methods of Operations Research*, 88:369–397, 2018.
- [29] A. H. Hamel, A. Löhne, and B. Rudloff. Benson type algorithms for linear vector optimization and applications. *Journal of Global Optimization*, 59(4):811–836, 2014.

- [30] F. Heyde. Geometric duality for convex vector optimization problems. *Journal of Convex Analysis*, 20(3):813–832, 2013.
- [31] F. Heyde and A. Löhne. Geometric duality in multiple objective linear programming. *SIAM Journal of Optimization*, 19(2):836–845, 2008.
- [32] J. Jahn. *Vector Optimization: Theory, Applications, and Extensions*. Springer, 2004.
- [33] İ. N. Keskin and F. Ulus. Outer approximation algorithms for convex vector optimization problems. *arXiv preprint arXiv:2109.07119*, 2021.
- [34] G. Kováčová and B. Rudloff. Time consistency of the mean-risk problem. *Operations Research*, 69:1100–1117, 2021.
- [35] A. Löhne. *Vector Optimization with Infimum and Supremum*. Springer, 2011.
- [36] A. Löhne and B. Rudloff. An algorithm for calculating the set of superhedging portfolios in markets with transaction costs. *International Journal of Theoretical and Applied Finance*, 17:1450012 (33 pages), 2014.
- [37] A. Löhne, B. Rudloff, and F. Ulus. Primal and dual approximation algorithms for convex vector optimization problems. *Journal of Global Optimization*, 60:713–736, 2014.
- [38] A. Löhne and B. Weißing. Equivalence between polyhedral projection, multiple objective linear programming and vector linear programming. *Mathematical Methods of Operations Research*, 84(2):411–426, 2016.
- [39] A. Löhne and B. Weißing. The vector linear program solver *Bensolve* – notes on theoretical background. *European Journal of Operational Research*, 260:807–813, 2017.
- [40] D. T. Luc. On duality in multiple objective linear programming. *European Journal of Operational Research*, 210:158–168, 2011.
- [41] K. Miettinen and M. M. Mäkelä. On generalized trade-off directions in nonconvex multiobjective optimization. *Mathematical Programming*, 92:141–151, 2002.
- [42] K. Miettinen, M. M. Mäkelä, and K. Kaario. Experiments with classification-based scalarizing functions in interactive multiobjective optimization. *European Journal of Operational Research*, 175(2):931–947, 2006.
- [43] J. L. Ndoutoume. Calcul différentiel du second ordre. *Publications AVAMAC, University of Perpignan*, 1987.
- [44] J. Niebling and G. Eichfelder. A branch–and–bound–based algorithm for nonconvex multiobjective optimization. *SIAM Journal on Optimization*, 29:794–821, 2019.
- [45] A. Pascoletti and P. Serafini. Scalarizing vector optimization problems. *Journal of Optimization Theory and Applications*, 42(4):499–524, 1984.
- [46] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

- [47] R. T. Rockafellar. First- and second-order epi-differentiability in nonlinear programming. *Transactions of the American Mathematical Society*, 307:75–108, 1988.
- [48] B. Rudloff and F. Ulus. Certainty equivalent and utility indifference pricing for incomplete preferences via convex vector optimization. *Mathematics and Financial Economics*, 15:397–430, 2021.
- [49] B. Rudloff, F. Ulus, and R. Vanderbei. A parametric simplex algorithm for linear vector optimization problems. *Mathematical Programming*, 163:213–242, 2017.
- [50] A. Ruszczyński and R. J. Vanderbei. Frontiers of stochastically nondominated portfolios. *Econometrica*, 71:1287–1297, 2003.
- [51] S. Ruzika and M. M. Wiecek. Approximation methods in multiobjective programming. *Journal of Optimization Theory and Applications*, 126(3):473–501, 2005.
- [52] A. Seeger. Second derivatives of a convex function and of its Legendre-Fenchel transform. *SIAM Journal on Optimization*, 2:405–424, 1992.
- [53] L. Shao and L. Ehrgott. Approximately solving multiobjective linear programmes in objective space and an application in radiotherapy treatment planning. *Mathematical Methods of Operations Research*, 68(2):257–276, 2008.
- [54] L. Shao and L. Ehrgott. Approximating the nondominated set of an MOLP by approximately solving its dual problem. *Mathematical Methods of Operations Research*, 68(3):469–492, 2008.
- [55] V. Soltan. *Lectures on Convex Sets*. World Scientific, 2015.
- [56] E. Zitzler, D. Brockhoff, and L. Thiele. The hypervolume indicator revisited: On the design of Pareto-compliant indicators via weighted integration. In *International Conference on Evolutionary Multi-Criterion Optimization*, pages 862–876. Springer, 2007.
- [57] E. Zitzler and L. Thiele. Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999.
- [58] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Da Fonseca. Performance assessment of multiobjective optimizers: An analysis and review. *IEEE Transactions on evolutionary computation*, 7(2):117–132, 2003.