SIAM J. OPTIM.
Vol. 34, No. 3, pp. 2700–2728

CONVERGENCE ANALYSIS OF A NORM MINIMIZATION-BASED CONVEX VECTOR OPTIMIZATION ALGORITHM*

ÇAĞIN ARARAT†, FIRDEVS ULUS†, AND MUHAMMAD UMER‡

Abstract. In this work, we propose an outer approximation algorithm for solving bounded convex vector optimization problems (CVOPs). The scalarization model solved iteratively within the algorithm is a modification of the norm-minimizing scalarization proposed in [Ç. Ararat, F. Ulus, and M. Umer, *J. Optim. Theory Appl.*, 194 (2022), pp. 681–712]. For a predetermined tolerance $\epsilon > 0$, we prove that the algorithm terminates after finitely many iterations, and it returns a polyhedral outer approximation to the upper image of the CVOP such that the Hausdorff distance between the two is less than ϵ. We show that for an arbitrary norm used in the scalarization models, the approximation error after k iterations decreases by the order of $\mathcal{O}(k^{1/(1-q)})$, where q is the dimension of the objective space. An improved convergence rate of $\mathcal{O}(k^{2/(1-q)})$ is proved for the special case of using the Euclidean norm.

Key words. convex vector optimization, multiobjective optimization, approximation algorithm, convergence rate, convex compact set, Hausdorff distance

MSC codes. 52A20, 52A27, 90B50, 90C25, 90C29

DOI. 10.1137/23M1574580

1. Introduction.

Vector optimization refers to minimizing a vector-valued objective function with respect to a given order relation over a feasible region. The special case where this order relation is induced by the positive orthant yields multiobjective optimization. We refer the reader to [6, 16, 20, 50, 51] for multiobjective optimization algorithms working in the decision space and to [8, 12, 14, 43, 45] for algorithms working in the objective space. Multiobjective optimization problems appear naturally in the calculation of set-valued risk measures in financial mathematics [2, 18]; they are also found to be linked to other challenging mathematical problems, such as convex projections [37], dynamic control and optimization problems [36, 53], and regularization [21, 22].

In this paper, we consider vector optimization problems in which the order relation is induced by a polyhedral ordering cone. Beyond the aforementioned multiobjective case, such problems have recently appeared naturally in various applications. In financial mathematics, the so-called *solvency cones* for markets with proportional transaction costs show up as ordering cones of certain vector optimization problems [39]. In particular, some unbounded multiobjective optimization problems related to utility indifference pricing become bounded convex vector optimization problems when formulated using solvency cones [49]. In game theory, the set of all Nash equilibria of a noncooperative game can be characterized in terms of a vector optimization problem [19]. In general, different ordering cones can be used to incorporate more/less

*Received by the editors May 23, 2023; accepted for publication (in revised form) April 30, 2024; published electronically July 25, 2024.
https://doi.org/10.1137/23M1574580
Funding: This work was supported by the 3501 program of TÜBİTAK (Scientific & Technological Research Council of Türkiye), Project No: 118M479.
†Bilkent University, Department of Industrial Engineering, Ankara, Türkiye (cararat@bilkent.edu.tr, firdevs@bilkent.edu.tr).
‡National University of Sciences and Technology (NUST), College of Aeronautical Engineering (CAE), Department of Industrial Engineering, Pakistan (mhmdumer@cae.nust.edu.pk).

conservative preferences of the decision maker, as recently done in [4] in a stochastic bandit framework.

Specifically, we focus on convex vector optimization problems (CVOPs) in which the objective function is cone-convex and the feasible region is a convex set. There are different solution concepts and algorithms for CVOPs in the literature. Many solution approaches are objective space-based, and they generate polyhedral approximations to the set of all (weakly) minimal elements in the objective space. To that end, they consider the *upper image,* which is the image of the feasible region plus the ordering cone, and iteratively generate polyhedral outer approximations to it by solving scalarization problems. A CVOP whose upper image can be covered by a shift of the ordering cone is called *bounded* (see the formal definition after Problem 2.3). We refer the reader to [27, 38] for algorithms to solve linear vector optimization problems and to [3, 5, 13, 15, 32, 34, 40] for algorithms to solve bounded CVOPs. Recently, in [54], algorithms and solution concepts are proposed for unbounded CVOPs.

In this paper, we focus on bounded CVOPs, propose an objective space-based outer approximation algorithm, and study its convergence rate. Note that the literature on the convergence analysis of such algorithms is very limited. For some special cases, the finiteness has been studied without determining the convergence rate. This is the case for some linear vector optimization algorithms; see [38, Chapter 5, p. 161]. Moreover, an outer approximation algorithm for nonlinear multiobjective optimization problems (MOPs), in which the ordering cone is the positive orthant, is proposed and proved to be finite in [46]. Recently, Ararat, Ulus, and Umer [5] proposed an outer approximation algorithm for general CVOPs and showed its finiteness. Their result is mainly based on constructing polyhedral outer approximations of the upper image such that the vertices are guaranteed to be within a compact set in each iteration.

On the other hand, in 2003, Klamroth, Tind, and Wiecek [35] proposed inner and outer approximation algorithms for convex and nonconvex MOPs, and the convergence rates of the algorithms are provided for the biobjective case. In particular, they proved that, for biobjective problems, the approximation error after k iterations of the algorithms in [35] decreases by the order of $\mathcal{O}(k^{-2})$. Later, in 2007, Klamroth and Tind [34] provided the convergence rate of similar algorithms for the multiobjective setting under the assumption that the image of the feasible region is bounded and its Minkowski sum with the positive orthant is closed. Accordingly, if there are $q \geq 3$ objective functions, then the approximation error after k iterations decreases by the order of $\mathcal{O}(k^{2/(1-q)})$. It is also noted that this convergence rate is in general the best possible rate; see also [26]. Note that the algorithms in [34, 35] cannot be directly applied to CVOPs, as they require the ordering cone to be the positive orthant, which is not the case for general CVOPs. When the problem has a polyhedral ordering cone, it is possible to reformulate it as a multiobjective optimization problem. However, the dimension of the objective space of the equivalent multiobjective problem may increase depending on the structure of the ordering cone, which may result in a worsened convergence rate overall. This is also discussed in Remark 4.2.

The convergence rate provided in [34] is based on existing results on the convergence rate of algorithms for finding polyhedral approximations of convex compact sets. Indeed, there is rich literature on this subject; see [29, 30, 31, 42], for instance. In general, the results for convex compact sets cannot be directly used to establish the convergence rate of objective space-based outer approximation algorithms for CVOPs since the upper image is an unbounded set by its structure. However, the design of the MOP algorithms in [34], which depends heavily on the symmetry of the orthants, allowed them to use the results in [29, 42] directly to establish the convergence rate.

More specifically, in [34, 35], using the symmetry of the orthants, a convex compact set is constructed such that its boundary contains all (weakly) minimal elements of the image of the feasible region. The algorithms solve scalarization problems where polyhedral gauge functions are used to construct a special and iteration-dependent norm to measure the distance between this convex compact set and its current polyhedral outer approximation.

In this paper, we propose an outer approximation algorithm for bounded CVOPs and study its convergence rate. To the best of our knowledge, this is the first paper providing a convergence analysis of a general CVOP algorithm in the literature. The proposed algorithm is based on a norm-minimizing scalarization, which is similar to the one proposed in [5]. We modify this scalarization by adding a single constraint which enables us to work with a compact subset of the upper image, which contains all weakly minimal elements of the image of the feasible set on its boundary.

Any norm on the objective space can be used to form the scalarization models solved within the proposed algorithm. Even though the norm used in the scalarization problems is arbitrary, it is fixed throughout the iterations. In that sense, the proposed algorithm and each of the algorithms in [34, 35] are not special cases of each other, as the latter uses a special and iteration-dependent norm.

We prove that, for any norm, the proposed algorithm is finite, and the approximation error after k iterations decreases by the order of $\mathcal{O}(k^{1/(1-q)})$, where q is the dimension of the objective space. In particular, we prove that the sequence of outer approximations formed by the algorithm is an *H-sequence of cutting* (see Definition 6.2) and then apply the results in [29, 42] directly.

Moreover, we consider the special case of using the Euclidean norm within the scalarizations separately and prove that the approximation error decreases by the order of $\mathcal{O}(k^{2/(1-q)})$. Different from the previous case, the results regarding the improved convergence rate in [29, 42] are not directly applicable to establish that of the proposed algorithm. Hence, the convergence analysis for the outer approximating polytopes produced by the proposed algorithm for the special case is proved from scratch.

The paper is organized as follows: section 2 provides basic definitions, notations, and the problem description, and it recalls some well-known results used throughout the paper. In section 3, the norm-minimizing scalarization is introduced and corresponding results are presented. The algorithm is explained in section 4, and its finiteness is proved in section 5. The convergence analysis of the algorithm is presented for any norm and for the special case of Euclidean norm in sections 6 and 7, respectively. Finally, we provide computational results in section 8 and conclude the paper in section 9.

2. Preliminaries.

In this section, we provide our notation and the formulation of the problem of interest.

2.1. Notation and conventions.

Throughout the paper, let $\mathbb{R}^q$ denote the q-dimensional Euclidean space, where $q \in \mathbb{N} := \{1, 2, \ldots\}$, and let $\|\cdot\|$ be an arbitrary fixed norm on $\mathbb{R}^q$ with dual norm $\|\cdot\|_*$. For $y \in \mathbb{R}^q$ and $\epsilon > 0$, we denote by $\mathbb{B}_\epsilon(y) := \{z \in \mathbb{R}^q \mid \|y - z\| \leq \epsilon\}$ the closed ball centered at y with radius ϵ.

Let $A \subseteq \mathbb{R}^q$ be a set. We write $\operatorname{int} A$, $\operatorname{cl} A$, $\operatorname{bd} A$, $\operatorname{conv} A$, $\operatorname{cone} A$ for the interior, closure, boundary, convex hull, conic hull of A, respectively. In particular, we denote by $\mathbb{S}^{q-1} := \operatorname{bd} \mathbb{B}_1(0)$ the unit sphere in $\mathbb{R}^q$. The *indicator function* I_A of A is defined by $I_A(z) = 0$ for $z \in A$ and by $I_A(z) = +\infty$ for $z \in \mathbb{R}^q \setminus A$. When $A, B \subseteq \mathbb{R}^q$ are nonempty sets and $\lambda \in \mathbb{R}$, we define $A + B := \{y + z \mid y \in A, z \in B\}$, $\lambda A := \{\lambda y \mid y \in A\}$, and $A - B := A + (-1)B$. When A is a nonempty convex set, the set of all $k \in \mathbb{R}^q$ such that $A + \operatorname{cone}\{k\} \subseteq A$ is the *recession cone* of A, denoted by $\operatorname{recc} A$; a nonzero element of $\operatorname{recc} A$ is a *recession direction* of A. When A is a nonempty convex cone, the *dual cone* of A, given by $A^+ := \{w \in \mathbb{R}^q \,|\, \forall y \in A : w^\mathsf{T} y \geq 0\}$, is a closed convex cone.

A cone $C \subseteq \mathbb{R}^q$ is said to be *solid* if $\operatorname{int} C \neq \emptyset$ and *pointed* if $C \cap -C = \{0\}$. Let C be a pointed convex cone. The relation $\leq_C$ on $\mathbb{R}^q$ defined by $y \leq_C z$ if and only if $z - y \in C$, $y, z \in \mathbb{R}^q$, is a partial order.

Let $A \subseteq \mathbb{R}^q$ be a set, and let $y \in A$. We say that y is a C*-minimal element* of A if there is no $z \in A \setminus \{y\}$ with $z \leq_C y$. When C is solid, we say that y is a *weakly* C*-minimal element* of A if $(\{y\} - \operatorname{int} C) \cap A = \emptyset$. Let $\operatorname{Min}_C A$ denote the set of all C-minimal elements of A, and let $\operatorname{wMin}_C A$ denote the set of all weakly C-minimal elements of A whenever C is solid.

Let $A, B \subseteq \mathbb{R}^q$ be nonempty sets. The Hausdorff distance between A and B is defined as

$$\delta^H(A,B) := \max\left\{\sup_{y \in A} d(y,B),\ \sup_{z \in B} d(z,A)\right\},$$

where $d(z,A) := \inf_{y \in A} \|z - y\|$. By [11, Proposition 3.2], we also have

$$\delta^H(A,B) = \inf\{\epsilon > 0 \,|\, A \subseteq B + \mathbb{B}_\epsilon(0),\ B \subseteq A + \mathbb{B}_\epsilon(0)\}.$$

Let $A \subseteq \mathbb{R}^q$ be a convex set, and let $w \in \mathbb{R}^q \setminus \{0\}$. We introduce the halfspace

$$\mathcal{H}(w,A) := \left\{z \in \mathbb{R}^q \,|\, w^\mathsf{T} z \geq \inf_{z' \in A} w^\mathsf{T} z'\right\}. \tag{2.1}$$

We also write $\mathcal{H}(w,y) := \mathcal{H}(w,\{y\})$ for every $y \in \mathbb{R}^q$. If $y \in A$ is such that $w^\mathsf{T} y = \inf_{z' \in A} w^\mathsf{T} z'$, then we have $\mathcal{H}(w,A) = \mathcal{H}(w,y)$. In this case, $\operatorname{bd} \mathcal{H}(w,A) = \{z \in \mathbb{R}^q \,|\, w^\mathsf{T} z = w^\mathsf{T} y\}$ is called a *supporting hyperplane* of A at y and $\mathcal{H}(w,A) \supseteq A$ is called a *supporting halfspace* of A at y. We say that a point $v \in A$ is an *extreme point* of A if there do not exist $y, z \in A$ and $0 < \lambda < 1$ such that $y \neq z$ and $v = \lambda y + (1-\lambda) z$. We denote the set of all extreme points of A by $\operatorname{ext} A$. If A is polyhedral, then $\operatorname{ext} A$ coincides with the set of all vertices of A. In this case, A can be represented as an intersection of its finitely many supporting halfspaces, which is called an H-representation of A [41, p. 808].

The following lemmata are simple observations that will be used throughout.

LEMMA 2.1. *Let* $w \in \mathbb{R}^q$ *with* $\|w\|_* \leq 1$ *and* $y \in \mathbb{R}^q$. *Then* $\mathcal{H}(w,y) + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \mathcal{H}(w,y,\epsilon)$, *where*

$$\mathcal{H}(w,y,\epsilon) := \left\{z \in \mathbb{R}^q \,|\, w^\mathsf{T} z \geq w^\mathsf{T} y - \frac{\epsilon}{2}\right\}. \tag{2.2}$$

Proof. Let $y' \in \mathcal{H}(w,y)$ and $y'' \in \mathbb{B}_{\frac{\epsilon}{2}}(0)$. Then, using $y' \in \mathcal{H}(w,y)$, Hölder's inequality, and the fact that $\|w\|_* \leq 1$ and $\|y''\| \leq \frac{\epsilon}{2}$, we have

$$w^\mathsf{T}(y' + y'') \geq w^\mathsf{T} y + w^\mathsf{T} y'' \geq w^\mathsf{T} y - \|w\|_* \|y''\| \geq w^\mathsf{T} y - \frac{\epsilon}{2},$$

which completes the proof. ∎

LEMMA 2.2. *Let* $A \subseteq \mathbb{R}^q$ *be a nonempty convex compact set and* $C \subseteq \mathbb{R}^q$ *be a nonempty convex cone. Then* $\operatorname{ext}(A + C) \subseteq \operatorname{ext} A$.

Proof. The inclusion is trivial if $\text{ext}(A+C)=\emptyset$. Otherwise, let $v \in \text{ext}(A+C)\setminus\text{ext}(A)$ for a contradiction. Since $v\in A+C$, we have $v=a+c$ for some $a\in A$ and $c\in C$. Note that if $c=0$, then $v=a\notin \text{ext}\,A$ implies that v can be written as a nontrivial convex combination of points from $A\subseteq A+C$. As this contradicts $v\in\text{ext}(A+C)$, we conclude that $c\neq 0$. Next, for every $\lambda\geq 0$, we have $a+\lambda c\in A+C$. For $\lambda>1$, we can write v as a strict convex combination of $a, a+\lambda c\in A+C$ as $v=a+c=\frac{\lambda-1}{\lambda}a+\frac{1}{\lambda}(a+\lambda c)$. This is a contradiction to $v\in\text{ext}(A+C)$. ∎

2.2. Problem definition and solution concept. The mathematical problem that is the subject of this study is provided next.

Problem 2.3. *We consider a* convex vector optimization problem (CVOP) *given by*

$$\text{(P)} \qquad \textit{minimize } \Gamma(x) \textit{ with respect to } \leq_C \textit{ subject to } x\in\mathcal{X},$$

where $C\subseteq\mathbb{R}^q$ is a closed convex solid and pointed cone, $X\subseteq\mathbb{R}^n$ is a convex set, $\Gamma\colon X\to\mathbb{R}^q$ is a C-convex and continuous vector-valued function, and $\mathcal{X}\subseteq X$ is a compact convex set with $\text{int}\,\mathcal{X}\neq\emptyset$.

Recall that Γ is called *C-convex* if $\Gamma(\lambda x_1+(1-\lambda)x_2)\leq_C \lambda\Gamma(x_1)+(1-\lambda)\Gamma(x_2)$ for every $x_1,x_2\in X, \lambda\in[0,1]$; equivalently, for every $w\in C^+$, the function $x\mapsto w^\mathsf{T}\Gamma(x)$ on X is convex.

The set $\mathcal{P}:=\text{cl}(\Gamma(\mathcal{X})+C)$ is called the *upper image* of (P). Under the assumptions of Problem 2.3, $\Gamma(\mathcal{X})$ is a compact set since it is the image of a compact set under a continuous function. Then $\Gamma(\mathcal{X})+C$ is a closed set by [1, Lemma 5.2, p. 166] and we have $\mathcal{P}=\Gamma(\mathcal{X})+C=\mathcal{P}+C$.

Moreover, (P) is a *bounded* CVOP in the sense that there exists a point $y\in\mathbb{R}^q$ such that $\mathcal{P}\subseteq\{y\}+C$; see [40, Definition 3.1].

The next definition recalls the relevant solution concepts for the CVOP given in Problem 2.3.

Definition 2.4. *Consider Problem* 2.3.

1. *A point $\bar{x}\in\mathcal{X}$ is called a* (weak) minimizer *for* (P) *if* $\Gamma(\bar{x})\in\text{Min}_C\,\Gamma(\mathcal{X})$ $(\Gamma(\bar{x})\in\text{wMin}_C\,\Gamma(\mathcal{X}))$.
2. [13, *Definition* 3.3] *Let $\bar{\mathcal{X}}\subseteq\mathcal{X}$ be a nonempty finite set of (weak) minimizers, and let $\bar{\mathcal{P}}:=\text{conv}\,\Gamma(\bar{\mathcal{X}})+C$. Then $\bar{\mathcal{X}}$ is called a* finite (weak) ϵ-solution *of* (P) *if $\bar{\mathcal{P}}+\mathbb{B}_\epsilon(0)\supseteq\mathcal{P}$.*

Remark 2.5. $\bar{\mathcal{P}}+\mathbb{B}_\epsilon(0)\supseteq\mathcal{P}$ is equivalent to $\delta^H(\mathcal{P},\bar{\mathcal{P}})\leq\epsilon$; see [5, Remark 3.6].

Below, we list some relevant results from the literature regarding Problem 2.3.

Proposition 2.6 (see [5, Proposition 3.8]). *For every $\epsilon>0$, there exists a finite ϵ-solution of* (P).

For each $w\in C^+\setminus\{0\}$, let us consider the corresponding *weighted sum scalarization* of (P):

$$(\text{WS}(w)) \qquad \text{minimize } w^\mathsf{T}\Gamma(x) \text{ subject to } x\in\mathcal{X}.$$

Proposition 2.7 (see [28, Corollary 2.3]). *Let $w\in C^+\setminus\{0\}$. Then every optimal solution of* (WS(w)) *is a weak minimizer of* (P).

3. A norm-minimizing scalarization. In this section, we discuss a modification of the following norm-minimizing scalarization model, which is introduced in [5] for CVOPs:

$$\text{(NM}(v)) \qquad \text{minimize } \|z\| \text{ subject to } \Gamma(x) - z - v \leq_C 0,\ x \in \mathcal{X},\ z \in \mathbb{R}^q,$$

where $v \in \mathbb{R}^q$. The Lagrange dual of (NM(v)) can be formulated as (see [5, section 4] for details)

$$\text{(dNM}(v)) \qquad \text{maximize } \inf_{x \in \mathcal{X}} w^\mathsf{T}\Gamma(x) - w^\mathsf{T}v \text{ subject to } \|w\|_* \leq 1, w \in C^+.$$

It is not difficult to observe that the norm-minimizing scalarization (NM(v)) computes the distance $d(v, \mathcal{P})$ from the point v to the upper image $\mathcal{P}$; see also [5, Remark 4.1]. Compared to similar existing scalarization models, e.g., the Pascoletti–Serafini scalarization [47], it depends only on a single parameter $v \in \mathbb{R}^q$ but not on an additional direction parameter. Choosing different direction parameters in Pascoletti–Serafini scalarization within an algorithmic framework affects the efficiency of the corresponding algorithm heavily [32]. However, an algorithm based on solving norm-minimizing scalarizations is free of direction-biasedness [5].

For our purposes, we modify (NM(v)) by adding a halfspace constraint as follows:

$$\text{(P}(v)) \quad \text{minimize } \|z\| \text{ subject to } \Gamma(x) - z - v \leq_C 0,\ \bar{w}^\mathsf{T}(v + z) \leq \gamma,\ x \in \mathcal{X},\ z \in \mathbb{R}^q,$$

where $\bar{w} \in \operatorname{int} C^+$ is assumed to be fixed and $\gamma \in \mathbb{R}$ is a parameter. Note that this is a convex program, as both the objective function and the feasible region are convex. The next remark discusses the key difference between the two scalarizations, which will later allow us to study the convergence rate of the proposed algorithm instead of the algorithms in [5]; see also Remark 6.11.

Remark 3.1. As mentioned earlier, (NM(v)) computes the distance $d(v, \mathcal{P})$ from the point v to the upper image $\mathcal{P}$. On the other hand, the optimal value of (P(v)) may not be equal to $d(v, \mathcal{P})$. Indeed, it finds the distance $d(v, \mathcal{P} \cap S(\gamma))$, where

$$S(\gamma) := \{y \in \mathbb{R}^q \mid \bar{w}^\mathsf{T} y \leq \gamma\}. \tag{3.1}$$

Clearly, if γ is not sufficiently large, then $\mathcal{P} \cap S(\gamma)$ may be the empty set, and hence (P(v)) may become infeasible. The construction of a suitable choice of γ is discussed in [5, section 6], and it will be recalled later in section 4. Figure 1 illustrates the sets $\Gamma(\mathcal{X}), \mathcal{P}, \mathcal{P} \cap S(\gamma)$ for a suitable choice of γ as well as the scalarization problems (NM(v)), (P(v)) for two different choices of v over a simple example.

The next lemma is a simple observation about $\mathcal{P} \cap S(\gamma)$. In particular, it implies that the optimal value of (P(v)) is finite whenever $\mathcal{P} \cap S(\gamma) \neq \emptyset$.

LEMMA 3.2. *$\mathcal{P} \cap S(\gamma)$ is a compact set.*

Proof. If $\mathcal{P} \cap S(\gamma) = \emptyset$, then the result holds trivially. Suppose that $\mathcal{P} \cap S(\gamma) \neq \emptyset$. Note that $\mathcal{P} \cap S(\gamma)$ is closed and convex as the intersection of closed convex sets. Let $c \in \operatorname{recc}(\mathcal{P} \cap S(\gamma))$, and fix $z \in \mathcal{P} \cap S(\gamma)$. Then, for every $\lambda \geq 0$, we have $z + \lambda c \in \mathcal{P} \cap S(\gamma)$. Since (P) is a bounded CVOP, we have $\mathcal{P} \subseteq \{y\} + C$ for some $y \in \mathbb{R}^q$. In particular, for every $\lambda \geq 0$, we have $z + \lambda c \in (\{y\} + C) \cap S(\gamma)$ so that

$$\forall w \in C^+ \colon w^\mathsf{T} y = \inf_{\bar{c} \in C} w^\mathsf{T}(y + \bar{c}) \leq w^\mathsf{T}(z + \lambda c) \quad \text{and} \quad \bar{w}^\mathsf{T}(z + \lambda c) \leq \gamma.$$

Then letting $\lambda \to \infty$ implies that

$$\forall w \in C^+ \colon w^\mathsf{T} c \geq 0 \quad \text{and} \quad \bar{w}^\mathsf{T} c \leq 0.$$

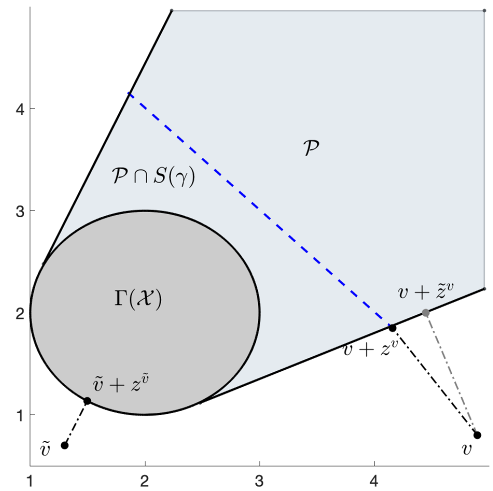


FIG. 1. *The sets* $\Gamma(\mathcal{X}),\mathcal{P},\mathcal{P}\cap S(\gamma)$ *for the problem with* $\Gamma:\mathbb{R}^2\to\mathbb{R}^2,\Gamma(x)=x$, $\mathcal{X}=\mathbb{B}_1(2)$, *and* $C=\operatorname{conv}\operatorname{cone}\{(1,2)^\mathsf{T},(2,1)^\mathsf{T}\}$. *The dashed blue line (color available online only) indicates* $(\operatorname{bd}S(\gamma))\cap\mathcal{P}$. *For* $\tilde{v}=(1.3,0.7)^\mathsf{T}$, $(\mathrm{NM}(v))$ *and* $(\mathrm{P}(v))$ *yield the same optimal solution* $(x^{\tilde{v}},z^{\tilde{v}})$ *with* $\tilde{v}+z^{\tilde{v}}\in\operatorname{bd}\mathcal{P}$. *However, for* $v=(4.9,0.8)^\mathsf{T}$, $(\mathrm{P}(v))$ *yields an optimal solution* (x^v,z^v) *with* $v+z^v\in\operatorname{bd}(\mathcal{P}\cap S(\gamma))$, *whereas* $(\mathrm{NM}(v))$ *yields an optimal solution* $(\tilde{x}^v,\tilde{z}^v)$ *with* $v+\tilde{z}^v\in(\operatorname{bd}\mathcal{P})\setminus S(\gamma)$.

In other words, $c\in C$ and $\bar{w}^\mathsf{T}c\le 0$. Since $\bar{w}\in\operatorname{int}C^+$, we must have $c=0$. Therefore, $\operatorname{recc}(\mathcal{P}\cap S(\gamma))=\{0\}$. Then, by [48, Theorem 8.4, p. 64], $\mathcal{P}\cap S(\gamma)$ is bounded and hence compact. ∎

The Lagrange dual of $(\mathrm{P}(v))$ can be formulated as

$$(\mathrm{D}(v))\qquad \text{maximize}\ \ \phi(w,\lambda)\ \ \text{subject to}\ \ w\in C^+,\ \lambda\ge 0,$$

where the dual objective function $\phi\colon\mathbb{R}^{q+1}\to\overline{\mathbb{R}}$ is given by

$$\phi(w,\lambda):=\inf_{x\in\mathcal{X},z\in\mathbb{R}^q}\left(\|z\|+w^\mathsf{T}(\Gamma(x)-z-v)+\lambda(\bar{w}^\mathsf{T}(v+z)-\gamma)\right).$$

We establish the strong duality between $(\mathrm{P}(v))$ and $(\mathrm{D}(v))$ next. The derivation of $(\mathrm{D}(v))$ and the proofs of the results in this section are provided in Appendix A.

PROPOSITION 3.3. *Let* $\gamma\in\mathbb{R}$ *be such that* $\Gamma(\mathcal{X})\subseteq\operatorname{int}S(\gamma)$, *where* $S(\gamma)$ *is given by* (3.1). *Then, for every* $v\in\mathbb{R}^q$, *there exist optimal solutions* (x^v,z^v) *and* (w^v,λ^v) *of problems* $(\mathrm{P}(v))$ *and* $(\mathrm{D}(v))$, *respectively, and the optimal values coincide.*

NOTATION 3.4. *From now on, by* (x^v,z^v) *and* (w^v,λ^v), *we denote arbitrary optimal solutions of problems* $(\mathrm{P}(v))$ *and* $(\mathrm{D}(v))$, *respectively. Moreover, we define* $\tilde{w}^v:=w^v-\lambda^v\bar{w}$.

The next lemma relates an optimal solution (w^v,λ^v) of $(\mathrm{D}(v))$ with the possible regions for the parameter v. Although the result is analogous to [5, Lemma 4.5], it works under the additional assumption that $v\in S(\gamma)$.

LEMMA 3.5. *Let* $\gamma\in\mathbb{R}$ *be such that* $\Gamma(\mathcal{X})\subseteq\operatorname{int}S(\gamma)$. *Let* $v\in S(\gamma)$. *Then the following statements hold:*
(a) *If* $v\notin\mathcal{P}$, *then* $z^v\neq 0$, $w^v\neq 0$, *and* $\tilde{w}^v\neq 0$.
(b) *If* $v\in\operatorname{bd}\mathcal{P}$, *then* $z^v=0$.
(c) *If* $v\in\operatorname{int}\mathcal{P}$, *then* $z^v=0$ *and* $w^v=0$.
In particular, $v\in\mathcal{P}$ *if and only if* $z^v=0$.

The following two propositions interpret the primal (x^v, z^v) and the dual optimal solutions (w^v, λ^v); that is, a primal solution yields a weak minimizer for problem (P), and a dual solution provides a supporting hyperplane of $\mathcal{P} \cap S(\gamma)$ at $y^v = v + z^v$.

PROPOSITION 3.6. *If $v \notin \operatorname{int} \mathcal{P}$ and $v \in S(\gamma)$, then x^v is a weak minimizer of* (P), *and $y^v := v + z^v \in \operatorname{wMin}_C \mathcal{P}$.*

PROPOSITION 3.7. *Suppose that $\tilde{w}^v \neq 0$. Let $v \in S(\gamma)$ and $\Gamma(\mathcal{X}) \subseteq \operatorname{int} S(\gamma)$. Then*

$$\mathcal{H}(\tilde{w}^v, y^v) = \{y \in \mathbb{R}^q \mid (\tilde{w}^v)^\mathsf{T} y \geq (\tilde{w}^v)^\mathsf{T} y^v\}$$

is a supporting halfspace of $\mathcal{P} \cap S(\gamma)$ at $y^v = v + z^v$.

4. The algorithm. In this section, we propose an algorithm for solving Problem 2.3 whose finiteness will be proved in section 5, followed by its convergence rate in section 6. Algorithm 1 is an outer approximation algorithm which has a structure similar to Algorithm 2 in [5]. It consists of two phases. The first one is an initialization phase which computes a compact initial outer approximation $\bar{\mathcal{P}}_0^{\text{out}}$ of $\mathcal{P} \cap S(\gamma)$ using the original norm-minimizing scalarization (NM(v)) and its dual (dNM(v)). In the second phase, the current outer approximation is refined in each iteration via a supporting halfspace that is calculated by solving the modified norm-minimizing scalarization (P(v)) and its dual (D(v)). The algorithm works under the following assumption.

Assumption 4.1. The ordering cone C is polyhedral.

Assumption 4.1 ensures that the dual cone C^+ is also polyhedral. Letting $w^1, \ldots, w^J$ be the generating vectors of C^+ with $J \in \mathbb{N}$, we have $C^+ = \operatorname{conv} \operatorname{cone}\{w^1, \ldots, w^J\}$. Note that C^+ is solid, as C is assumed to be pointed; see [17, Proposition 2.4.3, p. 174]. In particular, $J \geq q$. Without loss of generality, we assume that $\|w^j\|_* = 1$ for each $j \in \{1, \ldots, J\}$.

Remark 4.2. Under Assumption 4.1, we have $C = \{y \in \mathbb{R}^q \mid Wy \geq 0\}$, where $W := [w^1, \ldots, w^J]^\mathsf{T} \in \mathbb{R}^{J \times q}$. Then problem (P) is equivalent to the multiobjective optimization problem

$$\text{minimize } W\Gamma(x) \text{ with respect to } \leq_{\mathbb{R}^J_+} \text{ subject to } x \in \mathcal{X}$$

in the sense that both problems have the same set of minimizers if the null space of W equals $\{0\}$ [52, Lemma 2.3.4, p. 31]. If one applies the algorithms in [34, 35] to solve the equivalent problem, then the approximation error after k iterations decreases by the order of $\mathcal{O}(k^{2/(1-J)})$, which would yield a better convergence rate than $\mathcal{O}(k^{2/(1-q)})$ only if $J < q$. However, due to the assumption that C is pointed, $J \geq q$ holds. Indeed, for $q \geq 3$, J may be much larger than q in general.

The initialization phase starts by solving the scalarizations (WS(w^j)), $j \in \{1, \ldots, J\}$: an optimal solution x^j of (WS(w^j)) exists by the compactness of $\mathcal{X}$ and the continuity of Γ, and x^j is a weak minimizer of (P) by Proposition 2.7. This gives the set $\mathcal{X}_0 := \{x^1, \ldots, x^J\}$ of weak minimizers. We define

$$\mathcal{P}_0^{\text{out}} := \bigcap_{j=1}^{J} \{y \in \mathbb{R}^q \mid (w^j)^\mathsf{T} y \geq (w^j)^\mathsf{T} \Gamma(x^j)\}. \tag{4.1}$$

Note that $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$; see [5, section 5]. Then $\mathcal{P}_0^{\text{out}}$ is further intersected with the halfspace $S(\gamma) = \{y \in \mathbb{R}^q \mid \bar{w}^\mathsf{T} y \leq \gamma\}$ (see (3.1)), where $\bar{w} \in \operatorname{int} C^+$, $\gamma \in \mathbb{R}$ are fixed such that

Algorithm 1 Outer Approximation Algorithm for (P).

Compute an optimal solution x^j of (WS(w^j)) for each $j \in \{1,\ldots,J\}$;
$k=0, \bar{\mathcal{X}}_0 = \{x^1,\ldots,x^J\}, \mathcal{V}^{\text{known}} = \emptyset, \mathcal{V}^{\text{known2}} = \emptyset$;
Store an H-representation of $\mathcal{P}_0^{\text{out}}$ according to (4.1);
Compute $\operatorname{ext} \mathcal{P}_0^{\text{out}}$ from the H-representation of $\mathcal{P}_0^{\text{out}}$;
Compute γ by (4.2) using [5, Remark 6.1];
$\bar{\mathcal{P}}_0^{\text{out}} = \mathcal{P}_0^{\text{out}} \cap S(\gamma)$;
repeat
  Stop $\leftarrow$ **true**;
  Compute $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$ from the H-representation of $\bar{\mathcal{P}}_k^{\text{out}}$;
  for $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}} \setminus \mathcal{V}^{\text{known}}$ **do**
    if $v \notin \mathcal{V}^{\text{known2}}$ **then**
      Solve (P(v)) and (D(v)) to compute (x^v, z^v) and (w^v, λ^v);
    end if
    if $\|z^v\| \le \epsilon$ **then**
      $\bar{\mathcal{X}}_k \leftarrow \bar{\mathcal{X}}_k \cup \{x^v\}$, $\mathcal{V}^{\text{known}} \leftarrow \mathcal{V}^{\text{known}} \cup \{v\}$;
    else
      $\mathcal{V}^{\text{known2}} \leftarrow \mathcal{V}^{\text{known2}} \cup \{v\}$;
    end if
  end for
  Compute $v^k \in \arg\max\{\|z^v\| \mid v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}\}$;
  if $\|z^{v^k}\| > \epsilon$ **then**
    $\bar{\mathcal{P}}_{k+1}^{\text{out}} = \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$, $\bar{\mathcal{X}}_{k+1} = \bar{\mathcal{X}}_k$;
    $k \leftarrow k+1$;
    Stop $\leftarrow$ **false**;
  end if
until Stop
return $\begin{cases} \bar{\mathcal{X}}_k & : \text{A finite weak } \epsilon\text{-solution to (P)}; \\ \bar{\mathcal{P}}_k^{\text{out}} & : \text{An outer approximation of } \mathcal{P} \cap S(\gamma). \end{cases}$

$$\bar{w} := \frac{\sum_{j=1}^J w^j}{\left\|\sum_{j=1}^J w^j\right\|_*}, \qquad \gamma > \sup_{x\in\mathcal{X}} \bar{w}^{\mathsf{T}}\Gamma(x) + \max_{v \in \operatorname{ext} \mathcal{P}_0^{\text{out}}} (\bar{w}^{\mathsf{T}} v - \beta)^+ + \delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P}). \tag{4.2}$$

Here $\beta \ge \sup_{x\in\mathcal{X}} \bar{w}^{\mathsf{T}}\Gamma(x)$ is a constant and $a^+ := \max\{a, 0\}$ for $a \in \mathbb{R}$. The calculation of γ is discussed in [5, section 6] in detail.[1] The choice of γ in (4.2) ensures that $\Gamma(\mathcal{X}) \subseteq \mathcal{P} \cap S(\gamma)$. Finally, we obtain the initial outer approximation of $\mathcal{P} \cap S(\gamma)$ as $\bar{\mathcal{P}}_0^{\text{out}} := \mathcal{P}_0^{\text{out}} \cap S(\gamma)$, which is compact by [5, Lemma 6.3].

Given $k \ge 0$, let $\bar{\mathcal{P}}_k^{\text{out}}$ be the current outer approximation of $\mathcal{P} \cap S(\gamma)$, to be defined recursively. In the second phase of the algorithm, the set $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$ of all vertices of $\bar{\mathcal{P}}_k^{\text{out}}$ is computed, $k \ge 0$. Then, for each $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$, optimal solutions (x^v, z^v) and (w^v, λ^v) are obtained by solving the modified norm-minimizing scalarization (P(v)) and its dual (D(v)), respectively. If $\|z^v\| \le \epsilon$, where $\epsilon > 0$ is a predetermined approximation

[1] Note that computing $\sup_{x\in\mathcal{X}} \bar{w}^{\mathsf{T}}\Gamma(x)$ would require solving a concave minimization problem. However, since γ is an upper bound, it is sufficient to run a few iterations of an algorithm dealing with concave minimization such as the ones in [7, 9].

error, then the set $\bar{\mathcal{X}}_k$ of weak minimizers is updated with x^v; see Proposition 3.6. To find one of the farthest vertices to the set $\mathcal{P} \cap S(\gamma)$, we compute

$$v^k \in \arg\max\{\|z^v\| \mid v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}\}. \tag{4.3}$$

If $\|z^{v^k}\| > \epsilon$, then the supporting halfspace of $\mathcal{P}$ at y^{v^k}, namely

$$\mathcal{H}_k := \mathcal{H}(\tilde{w}^{v^k}, y^{v^k}) = \{y \in \mathbb{R}^q \mid (\tilde{w}^{v^k})^\mathsf{T} y \geq (\tilde{w}^{v^k})^\mathsf{T} y^{v^k}\}, \tag{4.4}$$

is used to update the current outer approximation as $\bar{\mathcal{P}}_{k+1}^{\text{out}} := \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$; see Proposition 3.7. Otherwise, the algorithm terminates.

Remark 4.3. By the definition of $S(\gamma)$, we have $\Gamma(\mathcal{X}) \subseteq S(\gamma)$. Since $\bar{\mathcal{P}}_0^{\text{out}} \supseteq \mathcal{P} \cap S(\gamma)$ and $\bar{\mathcal{P}}_{k+1}^{\text{out}} = \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$, we have $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \mathcal{P} \cap S(\gamma) \supseteq \Gamma(\mathcal{X})$ for every $k \geq 0$. Then, using [5, Remark 3.2], we obtain $\mathcal{P} = \Gamma(\mathcal{X}) + C \subseteq \bar{\mathcal{P}}_k^{\text{out}} + C$. Hence, $\bar{\mathcal{P}}_k^{\text{out}} + C$ gives an outer approximation of $\mathcal{P}$.

Remark 4.4. The upper image $\mathcal{P}$ is Motzkin decomposable (see, for instance, [23]), as it can be written as the sum of a compact convex set and a closed convex cone. Indeed, we have $\mathcal{P} = (\mathcal{P} \cap S(\gamma)) + C$. Note that $\mathcal{P} = \Gamma(\mathcal{X}) + C \subseteq (\mathcal{P} \cap S(\gamma)) + C$ as $\Gamma(\mathcal{X}) \subseteq S(\gamma)$. The other inclusion is by $(\mathcal{P} \cap S(\gamma)) + C \subseteq \mathcal{P} + C = \mathcal{P}$.

LEMMA 4.5. *For every $k \geq 0$ and $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{out}$, we have $v \notin \operatorname{int} \mathcal{P}$.*

Proof. Clearly, $\operatorname{bd} \mathcal{P}_0^{\text{out}} \subseteq \mathbb{R}^q \setminus \operatorname{int} \mathcal{P}$ since $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$. Then, for every $v \in \operatorname{ext} \bar{\mathcal{P}}_0^{\text{out}} \subseteq \operatorname{bd} \mathcal{P}_0^{\text{out}}$, we have $v \notin \operatorname{int} \mathcal{P}$.

Assume that $k \geq 1$ is the first iteration number for which there exists $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$ such that $v \in \operatorname{int} \mathcal{P}$. Note that $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}} \subseteq S(\gamma)$ by the construction of the algorithm. Since v is a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$, it must be true that $v \in \operatorname{bd} \mathcal{H}_{\bar{k}}$ for some $\bar{k} \leq k$. Moreover, $\operatorname{bd} \mathcal{H}_{\bar{k}}$ is a supporting hyperplane of the closed set $\mathcal{P} \cap S(\gamma)$. Hence,

$$v \notin \operatorname{int}(\mathcal{P} \cap S(\gamma)) = \operatorname{int} \mathcal{P} \cap \operatorname{int} S(\gamma).$$

Then we must have $v \in \operatorname{bd} S(\gamma)$ since $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}} \subseteq S(\gamma)$. As $v \in \operatorname{int} \mathcal{P}$, there exists $\delta > 0$ such that $\mathbb{B}_\delta(v) \subseteq \operatorname{int} \mathcal{P}$. Clearly, $\mathbb{B}_\delta(v) \cap \operatorname{bd} S(\gamma) \neq \emptyset$. Let us fix $v' \in \operatorname{bd} S(\gamma)$, where $v' \neq v$, and define

$$s := \delta \frac{v - v'}{\|v - v'\|}.$$

Then $v \pm s$ is an affine combination of $v, v' \in \operatorname{bd} S(\gamma)$. Since $\operatorname{bd} S(\gamma)$ is an affine set, we get $v \pm s \in \operatorname{bd} S(\gamma)$. This implies that

$$v \pm s \in \operatorname{int} \mathcal{P} \cap \operatorname{bd} S(\gamma) \subseteq \mathcal{P} \cap S(\gamma) \subseteq \bar{\mathcal{P}}_k^{\text{out}}.$$

This implies that v can be written as a convex combination of $v \pm s \in \bar{\mathcal{P}}_k^{\text{out}}$ since $v = \frac{1}{2}(v+s) + \frac{1}{2}(v-s)$. This is a contradiction to $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$. Hence, $v \notin \operatorname{int} \mathcal{P}$. ∎

THEOREM 4.6. *Under Assumption* 4.1*, Algorithm* 1 *works correctly: if the algorithm terminates, then it returns a finite weak ϵ-solution to* (P).

Proof. Note that $\bar{\mathcal{X}}_0$ consists of weak minimizers and $\operatorname{ext} \bar{\mathcal{P}}_0^{\text{out}}$ is a nonempty set of vertices; see the proof of [5, Theorem 6.6] for details. Let $k \geq 0$ and $v \in \operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$. Then optimal solutions (x^v, z^v) and (w^v, λ^v) to (P(v)) and (D(v)), respectively, exist by Proposition 3.3. By Lemma 4.5, $v \notin \operatorname{int} \mathcal{P}$. Hence, by Proposition 3.6, x^v is a weak minimizer of (P). If $v^k \in \operatorname{bd}\mathcal{P}$ (see line 20), then we have $z^{v^k} = 0$ by Lemma 3.5. In this case, line 21 ensures that the outer approximation is not updated and the algorithm terminates. If $v^k \notin \mathcal{P}$, which implies $\tilde{w}^{v^k} \neq 0$ by Lemma 3.5, then $\mathcal{H}_k$ given by (4.4) is a supporting halfspace of $\mathcal{P} \cap S(\gamma)$ by Proposition 3.7. We have $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \mathcal{P} \cap S(\gamma)$ as $\bar{\mathcal{P}}_0^{\text{out}} \supseteq \mathcal{P} \cap S(\gamma)$; $\mathcal{H}_1, \dots, \mathcal{H}_k$ are supporting halfspaces of $\mathcal{P} \cap S(\gamma)$; and $\bar{\mathcal{P}}_k^{\text{out}} = \bar{\mathcal{P}}_0^{\text{out}} \cap \mathcal{H}_1 \cap \cdots \cap \mathcal{H}_k$. The latter also implies that $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}} \neq \emptyset$.

Assume that the algorithm stops after $\hat{k}$ iterations. Consequently, $\bar{\mathcal{X}}_{\hat{k}}$ is a finite set of weak minimizers. Then, using arguments similar to those in the proof of [5, Theorem 6.6], one can show that Definition 2.4 holds, that is,

$$\operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}}) + C + \mathbb{B}_\epsilon(0) \supseteq \mathcal{P},$$

which finishes the proof of correctness. □

5. Finiteness of the algorithm. In this section, we prove the finiteness of Algorithm 1 by showing that a subset of fixed volume is cut off of the current outer approximation in each iteration. We show that the subsets are nonoverlapping and also contained in a compact set, which implies a finite upper bound on the number of iterations of the algorithm. The result is based on the following lemma.

LEMMA 5.1. *Suppose that Assumption* 4.1 *holds. Let $v \notin \mathcal{P}$. Then the following statements hold:*

(a) $\|z^v\| = (\tilde{w}^v)^\mathsf{T} z^v$.

(b) $\|\tilde{w}^v\|_* = 1$.

(c) *If $\|z^v\| \geq \epsilon$, then $\mathbb{B}_{\frac{\epsilon}{4}}(v) \cap \mathcal{H}(\tilde{w}^v, y^v, \epsilon) = \emptyset$, where $\mathcal{H}(\tilde{w}^v, y^v, \epsilon) = \{y \in \mathbb{R}^q \,|\, (\tilde{w}^v)^\mathsf{T} y \geq (\tilde{w}^v)^\mathsf{T} y^v - \frac{\epsilon}{2}\}$.*

(d) *$\mathcal{H}(\tilde{w}^v, y^v) + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \mathcal{H}(\tilde{w}^v, y^v, \epsilon)$, where $\mathcal{H}(\tilde{w}^v, y^v) = \{y \in \mathbb{R}^q \,|\, (\tilde{w}^v)^\mathsf{T} y \geq (\tilde{w}^v)^\mathsf{T} y^v\}$.*

Proof. Using arguments similar to those in the proof of [5, Lemma 7.1], (a) and (c) follow. Moreover, (d) follows from Lemma 2.1. To see (b), note that

$$\|z^v\| = (\tilde{w}^v)^\mathsf{T} z^v \leq \|\tilde{w}^v\|_* \|z^v\| \leq \|z^v\|,$$

which follows by (a), Hölder's inequality, and the feasibility of $\tilde{w}^v$ for the formulation in (A.2). Hence, all terms are equal. Since $z^v \neq 0$ by Lemma 3.5(a), we get $\|\tilde{w}^v\|_* = 1$. □

THEOREM 5.2. *Suppose that Assumption* 4.1 *holds. Then Algorithm* 1 *terminates after a finite number of iterations.*

Proof. First, note that $\operatorname{ext} \bar{\mathcal{P}}_k^{\text{out}}$ consists of finitely many vertices for each $k \geq 0$. Next, we show the existence of some $K \geq 0$ such that $\|z^v\| \leq \epsilon$ holds for every $v \in \operatorname{ext} \bar{\mathcal{P}}_K^{\text{out}}$. Assume otherwise; that is, for every $k \geq 0$, we have $\|z^{v^k}\| > \epsilon$, where v^k is as in (4.3).

By [5, Lemma 6.3], the set $\bar{\mathcal{P}}_0^{\text{out}}$ is compact. The ball $\mathbb{B}_{\frac{\epsilon}{2}}(0)$ is also compact. Hence, $\bar{\mathcal{P}}_0^{\text{out}} + \mathbb{B}_{\frac{\epsilon}{2}}(0)$ is compact by [1, Lemma 5.2, p. 166]. For each $k \geq 0$, since $v^k \in \bar{\mathcal{P}}_k^{\text{out}}$, we have

$$\mathbb{B}_{\frac{\epsilon}{4}}(v^k) \subseteq \{v^k\} + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \bar{\mathcal{P}}_k^{\text{out}} + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \bar{\mathcal{P}}_0^{\text{out}} + \mathbb{B}_{\frac{\epsilon}{2}}(0). \tag{5.1}$$

Next, we show that $\mathbb{B}_{\frac{\epsilon}{4}}(v^i) \cap \mathbb{B}_{\frac{\epsilon}{4}}(v^j) = \emptyset$ for every $j > i \geq 0$. Clearly, $\bar{\mathcal{P}}_j^{\text{out}} \subseteq \bar{\mathcal{P}}_{i+1}^{\text{out}}$. By Lemma 5.1(c), we have $\mathbb{B}_{\frac{\epsilon}{4}}(v^i) \cap \mathcal{H}(\tilde{w}^{v^i}, y^{v^i}, \epsilon) = \emptyset$. Moreover, we have

$$\bar{\mathcal{P}}_j^{\text{out}} + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \bar{\mathcal{P}}_{i+1}^{\text{out}} + \mathbb{B}_{\frac{\epsilon}{2}}(0) = (\bar{\mathcal{P}}_i^{\text{out}} \cap \mathcal{H}_i) + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \mathcal{H}_i + \mathbb{B}_{\frac{\epsilon}{2}}(0),$$

where $\mathcal{H}_i$ is the supporting halfspace of $\mathcal{P} \cap S(\gamma)$ at y^i as obtained by Proposition 3.7. Using Lemma 5.1(d) with the above inclusion, we get

$$\bar{\mathcal{P}}_j^{\rm out} + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \mathcal{H}_i + \mathbb{B}_{\frac{\epsilon}{2}}(0) \subseteq \mathcal{H}(\tilde{w}^{v^i}, y^{v^i}, \epsilon).$$

This implies $\mathbb{B}_{\frac{\epsilon}{4}}(v^i) \cap (\bar{\mathcal{P}}_j^{\rm out} + \mathbb{B}_{\frac{\epsilon}{2}}(0)) = \emptyset$. From (5.1), $\mathbb{B}_{\frac{\epsilon}{4}}(v^j) \subseteq \bar{\mathcal{P}}_j^{\rm out} + \mathbb{B}_{\frac{\epsilon}{2}}(0)$. Hence, $\mathbb{B}_{\frac{\epsilon}{4}}(v^i) \cap \mathbb{B}_{\frac{\epsilon}{4}}(v^j) = \emptyset$. Note that the sets $\mathbb{B}_{\frac{\epsilon}{4}}(v^k)$, $k \geq 0$, have the same volume, which is strictly positive. With (5.1), these imply that the compact set $\bar{\mathcal{P}}_0^{\rm out} + \mathbb{B}_{\frac{\epsilon}{2}}(0)$ contains an infinite number of disjoint subsets of identical and nonzero volume, a contradiction. ∎

6. Convergence rate of the algorithm. In order to study the convergence rate of Algorithm 1, we aim to use the results of Kamenev [29] and Lotov, Bushenkov, and Kamenev [42, Chapter 8, p. 237], which originally hold for convex compact bodies. Even though the upper image $\mathcal{P}$ is not compact, it is Motzkin decomposable by Remark 4.4. Indeed, Algorithm 1 is basically a cutting method as defined in [29] (see Definition 6.2 below) applied to the compact component $\mathcal{P} \cap S(\gamma)$ of the upper image. Hence, the results in [29, 42] can still be partially used for the convergence analysis of Algorithm 1.

Note that the algorithmic schemes described in [42, Chapter 8, p. 237] are different from Algorithm 1 in terms of computing the points on the boundary of the convex compact set to be approximated and the corresponding supporting halfspaces of it. The convergence analysis in [42] is specifically designed for these schemes. We will see that we can still use some of their results by proving that Algorithm 1 satisfies their assumptions; see Theorem 6.9. However, Algorithm 1 does not necessarily satisfy the assumptions of their main result [42, Theorem 8.14, p. 251] regarding an improved convergence rate. Hence, in section 7, we will prove the improved convergence rate from scratch, while we can still use some lemmata directly from [42].

For the theoretical analysis in this section and section 7, we ignore the stopping criterion in Algorithm 1 and assume that it runs indefinitely while updating the current outer approximation in each iteration. This can be done by ignoring lines 21 and 25 of the algorithm.

Remark 6.1. Under Assumption 4.1, for the nonterminating version of Algorithm 1, it holds that

$$\lim_{k\to\infty} \delta^H(\bar{\mathcal{P}}_k^{\rm in}, \mathcal{P}) = \lim_{k\to\infty} \delta^H(\bar{\mathcal{P}}_k^{\rm out} + C, \mathcal{P}) = 0,$$

where, for each $k \geq 0$, $\bar{\mathcal{P}}_k^{\rm in} := \operatorname{conv}\Gamma(\bar{\mathcal{X}}_k) + C$ and the sets $\bar{\mathcal{X}}_k, \bar{\mathcal{P}}_k^{\rm out}$ are as described in Algorithm 1. The proof follows along the same lines as [5, Corollary 7.4] combined with Theorem 5.2.

In the following, we recall a definition from [42, Chapter 8, p. 237]. Note that the notations used in the original definition are replaced with the ones in this paper to remain consistent. To use the convergence result, [42, Theorem 8.6, p. 249], our algorithm has to produce the type of sequence of outer approximating polytopes provided in Definition 6.2 below.

Definition 6.2 (see [42, Definition 8.3, p. 246]). *Consider a nonempty convex compact set $A \subseteq \mathbb{R}^q$ and a sequence $(A_k)_{k\geq 0}$ of polytopes in $\mathbb{R}^q$. Assume that $A_0 = \bigcap_{i=1}^I \mathcal{H}(\omega^i, A)$, where $I \in \mathbb{N}$ and $\omega^i \in \mathbb{R}^q \setminus \{0\}$ for $i \in \{1, \dots, I\}$. We say that $(A_k)_{k\geq 0}$ is* generated by a cutting method *if, for every $k \geq 0$, it holds that $A_k \supseteq A$ and there exists a supporting halfspace $H_k \subseteq \mathbb{R}^q$ of A such that $A_{k+1} = A_k \cap H_k$. In this case,*

$(A_k)_{k\geq 0}$ *is called an* $H(r,A)$-sequence of cutting *for a given constant* $r>0$ *if, for every* $k\geq 0$*, it holds that*

$$\delta^H(A_k, A_{k+1}) \geq r\delta^H(A_k, A).$$

Note that, in each iteration $k\geq 0$ of the algorithm, we choose a farthest vertex, denoted as v^k, from the upper image of the current outer approximation and generate a halfspace $\mathcal{H}_k$, given by (4.4), to update the current outer approximation. Using this update structure, we show in Theorem 6.4 that the Hausdorff distance between any two consecutive outer approximations, $\bar{\mathcal{P}}_k^{\text{out}}$ and $\bar{\mathcal{P}}_{k+1}^{\text{out}}$, is equal to the Hausdorff distance between the former outer approximation $\bar{\mathcal{P}}_k^{\text{out}}$ and $\mathcal{P}\cap S(\gamma)$.

The next lemma is a simple observation that will be used in Theorem 6.4.

LEMMA 6.3. *Suppose that Assumption* 4.1 *holds. Let* $v\in S(\gamma)\setminus\mathcal{P}$*. Then* $\operatorname{int}\mathbb{B}_{\|z^v\|}(v)\cap\mathcal{H}(\tilde{w}^v,y^v)=\emptyset$*, where the halfspace* $\mathcal{H}(\tilde{w}^v,y^v)$ *is defined by Proposition* 3.7*.*

Proof. First, note that $v\in S(\gamma)\setminus\mathcal{P}$ implies $\|z^v\|>0$ and $\tilde{w}^v\neq 0$ by Lemma 3.5(a). Hence, $\operatorname{int}\mathbb{B}_{\|z^v\|}(v)\neq\emptyset$ and $\mathcal{H}(\tilde{w}^v,y^v)$ is well-defined. Using arguments similar to those in the proof of [5, Lemma 7.1], the assertion of the lemma follows. ∎

THEOREM 6.4. *Suppose that Assumption* 4.1 *holds, and consider the sequence of outer approximating polytopes* $(\bar{\mathcal{P}}_k^{out})_{k\geq 0}$ *produced by Algorithm* 1*. Then, for every* $k\geq 0$*, we have*

$$\delta^H(\bar{\mathcal{P}}_k^{out},\bar{\mathcal{P}}_{k+1}^{out})=\delta^H(\bar{\mathcal{P}}_k^{out},\mathcal{P}\cap S(\gamma)).$$

Proof. Let $k\geq 0$. By construction, we have $\bar{\mathcal{P}}_{k+1}^{\text{out}}\subseteq\bar{\mathcal{P}}_k^{\text{out}}$. By [5, Lemma 5.3], we get

$$\delta^H(\bar{\mathcal{P}}_k^{\text{out}},\bar{\mathcal{P}}_{k+1}^{\text{out}})=\max_{v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}} d(v,\bar{\mathcal{P}}_{k+1}^{\text{out}}).$$

First, we consider the case $\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}\subseteq\mathcal{P}$. Since $\bar{\mathcal{P}}_k^{\text{out}}$ is a polytope, we have $\bar{\mathcal{P}}_k^{\text{out}}\subseteq\mathcal{P}$ by the convexity of $\mathcal{P}$. Then, from $\bar{\mathcal{P}}_k^{\text{out}}\subseteq\bar{\mathcal{P}}_0^{\text{out}}=\mathcal{P}_0^{\text{out}}\cap S(\gamma)\subseteq S(\gamma)$, we get $\bar{\mathcal{P}}_k^{\text{out}}\subseteq\mathcal{P}\cap S(\gamma)$. The reverse inclusion is trivial by construction so that $\bar{\mathcal{P}}_k^{\text{out}}=\mathcal{P}\cap S(\gamma)$. Then $\bar{\mathcal{P}}_{k+1}^{\text{out}}=\bar{\mathcal{P}}_k^{\text{out}}$ and $\delta^H(\bar{\mathcal{P}}_k^{\text{out}},\bar{\mathcal{P}}_{k+1}^{\text{out}})=0$. Hence, the equality is trivial.

Next, let us assume that there exists some $v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}$ such that $v\notin\mathcal{P}$. Without loss of generality, we may assume that $v=v^k$; recall that v^k is a farthest vertex to $\mathcal{P}\cap S(\gamma)$ (see (4.3)). Then $\operatorname{int}\mathbb{B}_{\|z^{v^k}\|}(v^k)\cap\mathcal{H}_k=\emptyset$ by Lemma 6.3. Hence, $v^k\notin\bar{\mathcal{P}}_k^{\text{out}}\cap\mathcal{H}_k=\bar{\mathcal{P}}_{k+1}^{\text{out}}$. This implies $\delta^H(\bar{\mathcal{P}}_k^{\text{out}},\bar{\mathcal{P}}_{k+1}^{\text{out}})\geq\|z^{v^k}\|>0$, where the second inequality is due to $v^k\notin\mathcal{P}$; see Lemma 3.5. By [5, Lemma 5.3], $\delta^H(\bar{\mathcal{P}}_k^{\text{out}},\mathcal{P}\cap S(\gamma))$ is attained at a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$ so that

$$\delta^H(\bar{\mathcal{P}}_k^{\text{out}},\mathcal{P}\cap S(\gamma))=\max_{v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}} d(v,\mathcal{P}\cap S(\gamma))=\max_{v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}}\|z^v\|=\|z^{v^k}\|.$$

Let $\bar{y}\in\mathcal{H}_k$ be arbitrary. By the definition of $\mathcal{H}_k$ in (4.4) and Lemma 5.1(a), we have $(\tilde{w}^{v^k})^{\mathsf{T}}\bar{y}\geq(\tilde{w}^{v^k})^{\mathsf{T}}(v^k+z^{v^k})=(\tilde{w}^{v^k})^{\mathsf{T}}v^k+\|z^{v^k}\|$, that is,

$$(\tilde{w}^{v^k})^{\mathsf{T}}(\bar{y}-v^k)\geq\|z^{v^k}\|. \tag{6.1}$$

On the other hand, by Hölder's inequality and (A.2), we have

$$|(\tilde{w}^{v^k})^{\mathsf{T}}(\bar{y}-v^k)|\leq\|(\tilde{w}^{v^k})\|_*\|\bar{y}-v^k\|\leq\|\bar{y}-v^k\|. \tag{6.2}$$

Then, using (6.1) and (6.2), we obtain $\|\bar{y}-v^k\| \geq |(\tilde{w}^{v^k})^\mathsf{T}(\bar{y}-v^k)| \geq \|z^{v^k}\|$. We get

$$\begin{aligned}\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P}\cap S(\gamma)) = \|z^{v^k}\| &\leq \min_{y\in\mathcal{H}_k} \|y-v^k\| \\ &\leq \min_{y\in\bar{\mathcal{P}}_k^{\text{out}}\cap\mathcal{H}_k} \|y-v^k\| = d(v^k, \bar{\mathcal{P}}_k^{\text{out}}\cap\mathcal{H}_k) \\ &\leq \max_{v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}} d(v, \bar{\mathcal{P}}_k^{\text{out}}\cap\mathcal{H}_k) \\ &= \max_{v\in\operatorname{ext}\bar{\mathcal{P}}_k^{\text{out}}} d(v, \bar{\mathcal{P}}_{k+1}^{\text{out}}) = \delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}),\end{aligned}$$

where the last equality follows from the fact that $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}})$ is attained at a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$; see [5, Lemma 5.3]. Hence, $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) \geq \delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P}\cap S(\gamma))$. Next, since $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \bar{\mathcal{P}}_{k+1}^{\text{out}} \supseteq \mathcal{P}\cap S(\gamma)$, we have $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) \leq \delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P}\cap S(\gamma))$. Hence, the equality in the theorem follows. ∎

From now on, we use the following notation:

$$\mathcal{A} := \mathcal{P}\cap S(\gamma); \qquad \mathcal{A}_k := \bar{\mathcal{P}}_k^{\text{out}}, \quad k\geq 0. \tag{6.3}$$

The next corollary verifies that these sets form an instance of Definition 6.2.

Corollary 6.5. *Suppose that Assumption* 4.1 *holds. Then the sequence of outer approximating polytopes $(\mathcal{A}_k)_{k\geq 0}$ is an $H(1,\mathcal{A})$-sequence of cutting, that is, $\delta^H(\mathcal{A}_k,\mathcal{A}_{k+1}) \geq \delta^H(\mathcal{A}_k,\mathcal{A})$ holds for every $k\geq 0$.*

Proof. Let $k \geq 0$. Clearly, $\mathcal{A}$ and $\mathcal{A}_k$ satisfy the required conditions for the sets in Definition 6.2 since $\mathcal{A}\subseteq\mathcal{A}_k$ is compact by [5, Lemma 6.3]; $\mathcal{A}_0$ is defined as the intersection of supporting halfspaces of $\mathcal{A}$ (see lines 3 and 6 of Algorithm 1); and $\mathcal{A}_k$ has the required property (see line 22 of Algorithm 1 and Proposition 3.7). Moreover, by the definitions of $\mathcal{A}$ and $\mathcal{A}_k$ and Theorem 6.4, we have $\delta^H(\mathcal{A}_k,\mathcal{A}_{k+1}) = \delta^H(\mathcal{A}_k,\mathcal{A})$. ∎

We restate an important theorem from [42] that will be useful in proving the convergence of $(\mathcal{A}_k)_{k\geq 0}$.

Theorem 6.6 (see [42, Theorem 8.5, p. 247]). *Let $A\subseteq\mathbb{R}^q$ be a nonempty convex compact set, and let $r>0$. For an $H(r,A)$-sequence $(A_k)_{k\geq 0}$, we have*

$$\lim_{k\to\infty} \delta^H(A_k, A) = 0.$$

Next, we relate the Hausdorff distance between each approximation produced by Algorithm 1 and the upper image to the one between their compact versions in (6.3).

Lemma 6.7. *Suppose that Assumption* 4.1 *holds. Then, for every $k\geq 0$, we have*

$$\delta^H(\bar{\mathcal{P}}_k^{out}+C, \mathcal{P}) \leq \delta^H(\mathcal{A}_k,\mathcal{A}).$$

Proof. Let $k\geq 0$. By [5, Lemma 5.3], $\delta^H(\mathcal{A}_k,\mathcal{A})$ is attained at a vertex of $\mathcal{A}_k$. Hence,

$$\delta^H(\mathcal{A}_k,\mathcal{A}) = \max_{v\in\operatorname{ext}\mathcal{A}_k} d(v,\mathcal{A}).$$

By Lemma 2.2, we have $\operatorname{ext}(\mathcal{A}_k+C)\subseteq\operatorname{ext}\mathcal{A}_k$. Hence,

$$\begin{aligned}\delta^H(\mathcal{A}_k,\mathcal{A}) = \max_{v\in\operatorname{ext}\mathcal{A}_k} d(v,\mathcal{A}) &\geq \max_{v\in\operatorname{ext}\mathcal{A}_k} d(v,\mathcal{A}+C) \\ &\geq \max_{v\in\operatorname{ext}(\mathcal{A}_k+C)} d(v,\mathcal{A}+C) = \delta^H(\mathcal{A}_k+C,\mathcal{A}+C) \\ &= \delta^H(\bar{\mathcal{P}}_k^{\text{out}}+C, (\mathcal{P}\cap S(\gamma))+C),\end{aligned}$$

where the penultimate equality is a consequence of Remark 4.3, Remark 4.4, and [5, Lemma 5.3]. Moreover, by Remark 4.4, we have $\mathcal{P} = (\mathcal{P} \cap S(\gamma)) + C$. Hence, $\delta^H(\mathcal{A}_k, \mathcal{A}) \geq \delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P})$. ∎

Remark 6.8. By Corollary 6.5, $(\mathcal{A}_k)_{k\geq 0}$ is an $H(1,\mathcal{A})$-sequence of polytopes. Then Theorem 6.6 and Lemma 6.7 together imply the following result, which is already discussed in Remark 6.1:

$$\lim_{k\to\infty} \delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P}) = \lim_{k\to\infty} \delta^H(\mathcal{A}_k, \mathcal{A}) = 0.$$

We restate an important theorem from [42] on convergence rates. Its proof can be found in [29].

THEOREM 6.9 (see [42, Theorem 8.6, p. 249]). *Let $A \subseteq \mathbb{R}^q$ be a nonempty convex compact set, and let $r > 0$. Let $(A_k)_{k\geq 0}$ be an $H(r, A)$-sequence of cutting. Then, for every $\epsilon \in (0,1)$, there exists $N \geq 0$ such that*

$$\delta^H(A_k, A) \leq (1+\epsilon)\lambda(r,A)k^{\frac{1}{1-q}}$$

holds for every $k \geq N$. Here $\lambda(r,A)$ is a constant depending on r and the topological properties of A and can be found in [29, *Theorem* 2].

We conclude this section with an application of the previous theorem in our setting.

COROLLARY 6.10. *Suppose that Assumption* 4.1 *holds. Then the approximation error obtained through the iterations of Algorithm* 1 *decreases by the order of* $\mathcal{O}(k^{1/(1-q)})$.

Proof. By Corollary 6.5, the sequence of outer approximating polytopes $(\mathcal{A}_k)_{k\geq 0}$ is an $H(1,\mathcal{A})$-sequence of cutting. Then, by Lemma 6.7 and Theorem 6.9, we get

$$\delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P}) \leq \delta^H(\mathcal{A}_k, \mathcal{A}) \leq (1+\epsilon)\lambda(r,\mathcal{A})k^{\frac{1}{1-q}}$$

for all sufficiently large k. Hence, the result follows. ∎

Remark 6.11. [5, Algorithm 2] also computes outer approximations to $\mathcal{P} \cap S(\gamma)$ in an iterative manner. Different from the proposed algorithm in the present paper, it solves the norm-minimizing scalarization (NM(v)). By [5, Proposition 4.7], using a dual optimal solution of (NM(v)), one can generate a halfspace which supports $\mathcal{P}$ at $\Gamma(x^v)$. Indeed, for γ satisfying (4.2), we have $\Gamma(x^v) \in S(\gamma)$, which shows that this halfspace supports $\mathcal{P} \cap S(\gamma)$. Hence, the sequence of outer approximations of $\mathcal{P} \cap S(\gamma)$ produced by [5, Algorithm 2] is also generated by a cutting method as defined in Definition 6.2. However, Theorem 6.4 does not apply to [5, Algorithm 2], as the proof requires computing the distance from each vertex to the set $\mathcal{P} \cap S(\gamma)$ but not to the upper image itself; see Remark 3.1 and Figure 1. In particular, it is crucial that the algorithm selects a vertex that is farthest away from $\mathcal{P} \cap S(\gamma)$ in line 20 of Algorithm 1, which is not the case in [5, Algorithm 2].

7. Improved convergence rate under Euclidean norm. In this section, we find an improved estimate of convergence rate for Algorithm 1 when the ℓ_2-norm is used in the scalarizations. The statement of Theorem 7.2 below is similar to those of [42, Theorem 8.14, p. 251] and [30, Corollary 1]. However, our statement is valid for the outer approximating polytopes produced by Algorithm 1 instead of the polytopes described in these references. We need the following assumption for the improved convergence rate.

Assumption 7.1. $\|\cdot\|$ is the ℓ_2-norm, that is, $\|z\| = \sqrt{z^\mathsf{T} z}$ for every $z \in \mathbb{R}^q$.

The next theorem is the main result of this section.

THEOREM 7.2. *Suppose that Assumptions* 4.1*and* 7.1 *hold, and consider the sequence of outer approximating polytopes* $(\bar{\mathcal{P}}_k^{out})_{k\geq 0}$ *produced by Algorithm* 1. *Then, for every* $\epsilon \in (0,1)$, *there exists* $N \geq 0$ *such that*

$$\delta^H(\bar{\mathcal{P}}_{k-1}^{out}, \mathcal{P} \cap S(\gamma)) \leq (1+\epsilon)\bar{\lambda}(\mathcal{P} \cap S(\gamma)) k^{\frac{2}{1-q}}$$

holds for every $k \geq N$. *Here* $\bar{\lambda}(\mathcal{P} \cap S(\gamma))$ *is a constant that depends on the topological properties of* $\mathcal{P} \cap S(\gamma)$ *and can be seen in Notation* 7.3.

NOTATION 7.3. *Let* $A \subseteq \mathbb{R}^q$ *be a nonempty convex compact set.*

(a) $R(A)$ *is the radius of the smallest ball circumscribed around* A.

(b) *The hypervolume of the unit ball* $\mathbb{B}_1(0) \subseteq \mathbb{R}^q$ *is denoted by* π_q, *and* $\bar{\lambda}(A) := 16R(A)(\frac{q\pi_q}{\pi_{q-1}})^{\frac{2}{q-1}}$.

We present five lemmata to prepare for the proof of Theorem 7.2. While two of these lemmata are directly paraphrased from [42], we provide proofs for the remaining three.

LEMMA 7.4 (see [42, Lemma 8.15, p. 252]). *Suppose that Assumption* 7.1 *holds. Let* $y, y' \in \mathbb{R}^q$ *and* $w, w' \in \mathbb{S}^{q-1}$ *be such that* $w^\mathsf{T} w' > 0$, $y' \in \mathcal{H}(w,y)$, *and* $y \in \mathcal{H}(w',y')$. *Then, for every* $\eta > 0$, *we have*

$$d(y, \operatorname{bd}\mathcal{H}(w',y')) \leq \frac{\|(y-\eta w) - (y'-\eta w')\|^2}{\eta}.$$

The next lemma is an analogue of [42, Lemma 8.17, p. 252] in our setting.[2]

LEMMA 7.5. *Suppose that Assumptions* 4.1 *and* 7.1 *hold. Let* $v \in S(\gamma) \setminus \mathcal{P}$, *and set* $y := y^v$, $w := \tilde{w}^v$. *Let* $w' \in \mathbb{S}^{q-1} \cap C^+$, $y' \in \mathcal{P} \cap S(\gamma)$ *be such that* $(w')^\mathsf{T} y' = \inf_{z \in \mathcal{P} \cap S(\gamma)} (w')^\mathsf{T} z$. *Assume that* $v \in \mathcal{H}(w', \mathcal{P} \cap S(\gamma))$. *Then, for every* $\eta > 0$, *we have*

$$\|(y-\eta w) - (y'-\eta w')\| \geq \min\{\eta\sqrt{2}, \sqrt{\eta h} - h\},$$

where $h := w^\mathsf{T}(y-v)$.

Proof. Note that $\|w\| = 1$ by Lemma 5.1(b). First, let us suppose that $w^\mathsf{T} w' \leq 0$. By Proposition 3.7, we have $w^\mathsf{T}(y'-y) \geq 0$; by the definition of y', we have $(w')^\mathsf{T}(y-y') \geq 0$. Hence,

$$\begin{aligned}
&\|(y-\eta w) - (y'-\eta w')\|^2 \\
&\quad = \|y-y'\|^2 + \eta^2 \|w'-w\|^2 + 2\eta w^\mathsf{T}(y'-y) + 2\eta(w')^\mathsf{T}(y-y') \\
&\quad \geq \eta^2 \|w'-w\|^2 = \eta^2 \|w'\|^2 + \eta^2 \|w\|^2 - 2\eta^2 w^\mathsf{T} w' \geq \eta^2 \|w'\|^2 + \eta^2 \|w\|^2 = 2\eta^2
\end{aligned}$$

so that $\|(y-\eta w) - (y'-\eta w')\| \geq \eta\sqrt{2}$.

[2] [42, Lemma 8.17, p. 252] assumes that the origin is included in the set to be approximated, which is not necessarily the case for $\mathcal{P} \cap S(\gamma)$.

Next, suppose that $w^{\mathsf{T}} w' > 0$. By the triangle inequality, we have

$$\|(y-\eta w)-(y'-\eta w')\| \geq \|(v-\eta w)-(y'-\eta w')\| - \|(v-\eta w)-(y-\eta w)\|, \tag{7.1}$$

and we control each term on the right separately. For the first term in (7.1), note that $y' \in \mathcal{P} \cap S(\gamma) \subseteq \mathcal{H}(w, \mathcal{P} \cap S(\gamma)) = \mathcal{H}(w,y)$. Moreover, by Lemma 5.1(a), we have $\mathcal{H}(w,y) \subseteq \mathcal{H}(w,v)$. Hence, $y' \in \mathcal{H}(w,v)$. On the other hand, since $v \in S(\gamma) \setminus \mathcal{P}$, Lemma 6.3 ensures that $v \notin \mathcal{H}(w,y)$. By our assumption, we have

$$v \in \mathcal{H}(w', \mathcal{P} \cap S(\gamma)) = \mathcal{H}(w', y').$$

Then, by Lemma 7.4, we get

$$\|(v-\eta w)-(y'-\eta w')\|^2 \geq \eta d(y', \operatorname{bd} \mathcal{H}(w,v)).$$

Since $\operatorname{bd} \mathcal{H}(w,v) = \{z \in \mathbb{R}^q | w^{\mathsf{T}} z = w^{\mathsf{T}} v\}$ and $w \in \mathbb{S}^{q-1}$, by elementary geometry, we have

$$d(y', \operatorname{bd} \mathcal{H}(w,v)) = |w^{\mathsf{T}}(y'-v)| \geq w^{\mathsf{T}}(y'-y) + w^{\mathsf{T}} y - w^{\mathsf{T}} v \geq w^{\mathsf{T}} y - w^{\mathsf{T}} v = h$$

so that $\|(v-\eta w)-(y'-\eta w')\| \geq \sqrt{\eta h}$. For the second term in (7.1), by Lemma 5.1(a), we have

$$\|(v-\eta w)-(y-\eta w)\| = \|v-y\| = w^{\mathsf{T}}(y-v) = h.$$

Hence, $\|(y-\eta w)-(y'-\eta w')\| \geq \sqrt{\eta h} - h$ by (7.1). ∎

Let us fix some $\eta > 0$. Similar to the construction in [42, Chapter 8, p. 253], we define a new sequence $(\mathcal{Z}_k)_{k \geq 0}$ as follows:

(i) We set $\mathcal{U}_0 := \{(w^j, \Gamma(x^j)) | j \in \{1, \ldots, J\}\} \cup \{(-\bar{w}, \bar{y})\}$, where $\{(x^j, w^j) | j \in \{1, \ldots, J\}\}$ is as in (4.1), $\bar{w}$ is defined in (4.2), and $\bar{y} \in \mathcal{P} \cap S(\gamma)$ is such that $\bar{w}^{\mathsf{T}} \bar{y} = \gamma$. The existence of $\bar{y}$ is guaranteed since the supremum is attained (hence finite) by the compactness of $\mathcal{X}$ and the continuity of $x \mapsto \bar{w}^{\mathsf{T}} \Gamma(x)$. Let us define

$$\mathcal{Z}_0 := \{y - \eta w | (w,y) \in \mathcal{U}_0\} = \{\Gamma(x^1) - \eta w^1, \ldots, \Gamma(x^J) - \eta w^J, \bar{y} + \eta \bar{w}\}.$$

(ii) For each $k \geq 0$, if $v^k \notin \mathcal{P}$, then we define

$$\begin{aligned} \mathcal{U}_{k+1} &:= \{(\tilde{w}^{v^k}, y^{v^k})\} \cup \mathcal{U}_k, \\ \mathcal{Z}_{k+1} &:= \{y - \eta w | (w,y) \in \mathcal{U}_{k+1}\} = \{y^{v^k} - \eta \tilde{w}^{v^k}\} \cup \mathcal{Z}_k. \end{aligned} \tag{7.2}$$

(iii) If $v^{\bar{k}} \in \mathcal{P}$ for some $\bar{k} \geq 0$, then by the vertex selection rule in Algorithm 1 (line 20), the current outer approximation is the same as $\mathcal{P} \cap S(\gamma)$. In this case, we set $\mathcal{Z}_k = \mathcal{Z}_{\bar{k}}$ for all $k \geq \bar{k}$. (Note that this case is realized only if $\mathcal{P} \cap S(\gamma)$ is a polyhedral set.)

Remark 7.6. For each $k \geq 0$, we have $\mathcal{Z}_k \subseteq \operatorname{bd}(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0))$. Let $k \geq 0$ and $(w,y) \in \mathcal{U}_k$. Note that we have

$$\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0) \subseteq \mathcal{H}(w,y) + \mathbb{B}_\eta(0).$$

Moreover, as $\|w\| = 1$, $\mathcal{H}(w,y) + \mathbb{B}_\eta(0) \subseteq \mathcal{H}(w,y,2\eta)$ holds by Lemma 2.1. Thus, we obtain $\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0) \subseteq \mathcal{H}(w,y,2\eta)$. Clearly, $y - \eta w \in \mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0)$. Moreover, we have $y - \eta w \in \operatorname{bd} \mathcal{H}(w,y,2\eta)$ since $w^{\mathsf{T}}(y - \eta w) = w^{\mathsf{T}} y - \eta$. These imply that $y - \eta w \in \operatorname{bd}(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0))$.

The next lemma shows that the cardinality of the sequence $(\mathcal{Z}_k)_{k \geq 0}$ grows linearly.

LEMMA 7.7. *Suppose that Assumptions* 4.1 *and* 7.1 *hold. Suppose also that* $\mathcal{P}\cap S(\gamma)$ *is a nonpolyhedral set. Then* $|\mathcal{Z}_k| = J+1+k$ *for each* $k\geq 0$.

Proof. Note that since $\mathcal{P}\cap S(\gamma)$ is a nonpolyhedral set, $\delta^H(\bar{\mathcal{P}}_k^{\rm out}, \mathcal{P}\cap S(\gamma)) > 0$ holds for all $k\geq 0$. This implies that for all $k\geq 0$, (1) $v^k\notin\mathcal{P}$ by the vertex selection rule of Algorithm 1 and (2) $\bar{\mathcal{P}}_{k+1}^{\rm out}\neq\bar{\mathcal{P}}_k^{\rm out}$ since, by Corollary 6.5, we have

$$\delta^H(\bar{\mathcal{P}}_{k+1}^{\rm out}, \bar{\mathcal{P}}_k^{\rm out}) \geq \delta^H(\bar{\mathcal{P}}_k^{\rm out}, \mathcal{P}\cap S(\gamma)) > 0.$$

Now, by construction, $|\mathcal{Z}_0| = J+1$ and $|\mathcal{Z}_k|\leq J+1+k$ for each $k\geq 0$. To get a contradiction, let $\bar{k}\geq 0$ be the smallest index at which we have $|\mathcal{Z}_{\bar{k}+1}| < J+1+\bar{k}+1$. Then there exists $(w,y)\in\mathcal{U}_{\bar{k}}$ such that

$$a := y^{v^{\bar{k}}} - \eta\tilde{w}^{v^{\bar{k}}} = y - \eta w. \tag{7.3}$$

We claim that $y^{v^{\bar{k}}} = y$. To prove this, let us note that, for an arbitrary $b\in\operatorname{bd}\mathcal{H}_{\bar{k}}$, we can calculate the distance from a to the hyperplane $\operatorname{bd}\mathcal{H}_{\bar{k}}$ as

$$d(a,\mathcal{H}_{\bar{k}}) = \frac{|(\tilde{w}^{v^{\bar{k}}})^{\mathsf{T}}(a-b)|}{\|\tilde{w}^{v^{\bar{k}}}\|} = |(\tilde{w}^{v^{\bar{k}}})^{\mathsf{T}}(a-b)|,$$

where we use Lemma 5.1(b). In particular, taking $b = y^{v^{\bar{k}}}\in\operatorname{bd}\mathcal{H}_{\bar{k}}$ gives

$$d(a,\mathcal{H}_{\bar{k}}) = |(\tilde{w}^{v^{\bar{k}}})^{\mathsf{T}}(a-y^{v^{\bar{k}}})| = \eta|(\tilde{w}^{v^{\bar{k}}})^{\mathsf{T}}\tilde{w}^{v^{\bar{k}}}| = \eta\|\tilde{w}^{v^{\bar{k}}}\|^2 = \eta.$$

The above distance is attained at $y^{v^{\bar{k}}}$ since $\|a-y^{v^{\bar{k}}}\| = \eta\|\tilde{w}^{v^{\bar{k}}}\| = \eta$. Moreover, due to the strict convexity of the ℓ_2-norm, we have $\|a-y'\| > \|a-y^{v^{\bar{k}}}\| = \eta$ for every $y'\in\mathcal{H}_{\bar{k}}$ with $y'\neq y^{v^{\bar{k}}}$. In particular, having $y^{v^{\bar{k}}}\neq y$ would yield $y\in\mathcal{P}\cap S(\gamma)\subseteq\mathcal{H}_{\bar{k}}$ so that $\|a-y\|>\eta$. However, we also have $\|a-y\| = \eta\,\|w\| = \eta$. Hence, we must have $y^{v^{\bar{k}}} = y$. Moreover, by (7.3), we also have $\tilde{w}^{v^{\bar{k}}} = w$ and $\mathcal{H}_{\bar{k}} = \mathcal{H}(w,y)$. Note that by the structure of Algorithm 1 we have

$$\bar{\mathcal{P}}_{\bar{k}}^{\rm out} = \bigcap_{(w,y)\in\mathcal{U}_{\bar{k}}}\mathcal{H}(w,y).$$

Then $\bar{\mathcal{P}}_{\bar{k}+1}^{\rm out} = \bar{\mathcal{P}}_{\bar{k}}^{\rm out}\cap\mathcal{H}_{\bar{k}} = \bar{\mathcal{P}}_{\bar{k}}^{\rm out}$, which is a contradiction. Hence, $|\mathcal{Z}_k| = J+1+k$ for each $k\geq 0$. ∎

Before proceeding further, we recall the following definition, which is critical in proving the convergence rate of the algorithm in Theorem 7.2.

DEFINITION 7.8 (see [42, Chapter 8, p. 253]). *A set* $Z\subseteq\mathbb{R}^q$ *is called* the base of an ϵ-packing *if* $\|y-z\|\geq 2\epsilon$ *for every* $y,z\in Z$.

The following lemma was originally proved for an H_1-sequence of outer approximating polytopes; see [42, Definition 8.4, p. 247 and Lemma 8.19, p. 254]. Here we show that the lemma still holds for the sequence of outer approximating polytopes generated by Algorithm 1, which, in general, may not be an H_1-sequence.

LEMMA 7.9. *Suppose that Assumptions* 4.1 *and* 7.1 *hold. Suppose that* $0<\epsilon<1$. *Then there exists* $N\geq 0$ *such that, for each* $k\geq N$, *the set* $\mathcal{Z}_k$ *is the base of an* $\varepsilon_k^N(\epsilon)$*-packing, where*

$$\varepsilon_k^N(\epsilon) := \frac{1}{2}\min\left\{\eta\sqrt{2}, 2\epsilon_N, (1-\epsilon)\sqrt{\eta\delta^H(\bar{\mathcal{P}}_{k-1}^{out}, \mathcal{P}\cap S(\gamma))}\right\} \tag{7.4}$$

and $\epsilon_N := \frac{1}{2}\min\{\|y-z\|\,|y,z\in\mathcal{Z}_N, y\neq z\}$.

Proof. Let us fix $N \geq 0$ to be determined later. We argue by induction on $k \geq N$. By the definitions of ϵ_N and $\varepsilon_N^N(\epsilon)$, we have $\|y-z\| \geq 2\epsilon_N \geq 2\varepsilon_N^N(\epsilon)$ for every $y, z \in \mathcal{Z}_N$. Hence, the assertion is trivial for $k = N$.

Assume that $\mathcal{Z}_{k-1}$ is the base of an $\varepsilon_{k-1}^N(\epsilon)$-packing for some $k \geq N+1$, that is,

$$\|y-z\| \geq 2\varepsilon_{k-1}^N(\epsilon)$$

for every $y, z \in \mathcal{Z}_{k-1}$. Moreover, since $\delta^H(\bar{\mathcal{P}}_{k-2}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \geq \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))$, we have $\varepsilon_{k-1}^N(\epsilon) \geq \varepsilon_k^N(\epsilon)$ by (7.4). Hence, $\mathcal{Z}_{k-1}$ is also the base of an $\varepsilon_k^N(\epsilon)$-packing.

For convenience, let us define $b^{k-1} := y^{v^{k-1}} - \eta \tilde{w}^{v^{k-1}} = v^{k-1} + z^{v^{k-1}} - \eta \tilde{w}^{v^{k-1}}$. Since $\mathcal{Z}_k = \{b^{k-1}\} \cup \mathcal{Z}_{k-1}$, to show that $\mathcal{Z}_k$ is the base of an $\varepsilon_k^N(\epsilon)$-packing, it is enough to verify that

$$\left\|b^{k-1} - b\right\| \geq 2\varepsilon_k^N(\epsilon)$$

for every $b \in \mathcal{Z}_{k-1}$. Let us fix $b \in \mathcal{Z}_{k-1}$. We may write $b = y' - \eta w'$, where $(w', y') \in \mathcal{U}_{k-1}$.

We know that $v^{k-1} \in S(\gamma) \setminus \mathcal{P}$ and $v^{k-1} \in \mathcal{H}(w', y') = \mathcal{H}(w', \mathcal{P} \cap S(\gamma))$. Also, $y' \in \operatorname{bd} \mathcal{H}(w', y')$ and $y' \in \mathcal{P} \cap S(\gamma)$ by definition. Then, by Lemma 7.5, we have

$$\|b^{k-1} - b\| = \|(y^{v^{k-1}} - \eta \tilde{w}^{v^{k-1}}) - (y' - \eta w')\| \geq \min\left\{\eta\sqrt{2},\ \sqrt{\eta h^{k-1}} - h^{k-1}\right\}, \tag{7.5}$$

where $h^{k-1} := (\tilde{w}^{v^{k-1}})^\mathsf{T}(y^{v^{k-1}} - v^{k-1})$. Note that

$$h^{k-1} = (\tilde{w}^{v^{k-1}})^\mathsf{T}(y^{v^{k-1}} - v^{k-1}) = (\tilde{w}^{v^{k-1}})^\mathsf{T} z^{v^{k-1}} = \|z^{v^{k-1}}\| = \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)).$$

We choose N such that $\delta^H(\bar{\mathcal{P}}_N^{\text{out}}, \mathcal{P} \cap S(\gamma)) \leq \eta\epsilon^2$. In particular, for each $k \geq N+1$, we have $\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \leq \eta\epsilon^2$, which implies

$$\sqrt{\frac{\eta\epsilon^2}{\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}} \geq 1.$$

Then

$$\begin{aligned}
\sqrt{\eta h^{k-1}} - h^{k-1} &= \sqrt{\eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))} - \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \\
&\geq \sqrt{\eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))} - \sqrt{\frac{\eta\epsilon^2}{\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}}\, \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \\
&= (1-\epsilon)\sqrt{\eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}.
\end{aligned}$$

Hence, from (7.5), we have

$$\begin{aligned}
\left\|b^{k-1} - b\right\| &\geq \min\left\{\eta\sqrt{2},\ \sqrt{\eta h^{k-1}} - h^{k-1}\right\} \\
&\geq \min\left\{\eta\sqrt{2},\ (1-\epsilon)\sqrt{\eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}\right\} \\
&\geq \min\left\{\eta\sqrt{2}, 2\epsilon_N, (1-\epsilon)\sqrt{\eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}\right\} = 2\varepsilon_k^N(\epsilon).
\end{aligned}$$

□

LEMMA 7.10 (see [42, Lemma 8.20, p. 255]). *Let $A \subseteq \mathbb{R}^q$ be a nonempty convex compact set. Suppose that $0 < \epsilon < R(A)$, and let $Z \subseteq \operatorname{bd} A$ be the base of an ϵ-packing. Then*

$$|Z| \leq N_1(\epsilon, R(A)) := \frac{q\pi_q}{\pi_{q-1}} \left(1 + \frac{R(A)^2}{\epsilon^2}\right)^{\frac{q-1}{2}}.$$

Using the above lemmata, we are ready to prove Theorem 7.2.

Proof of Theorem 7.2. If $\mathcal{P} \cap S(\gamma)$ is a polytope, then there exists $K' \in \mathbb{N}$ such that $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P} \cap S(\gamma)) = 0$ for every $k \geq K'$ and the assertion of the theorem holds trivially. For the rest of the proof, we assume that $\mathcal{P} \cap S(\gamma)$ is nonpolyhedral.

Let $\epsilon' := \frac{1}{2}(1 - \frac{1}{\sqrt{1+\epsilon}}) \in (0,1)$ and $\eta := R(\mathcal{P} \cap S(\gamma)) > 0$. By Lemma 7.9, there exists $N' \in \mathbb{N}$ such that, for every $k \geq N'$, the set $\mathcal{Z}_k$ is the base of an $\varepsilon_k^{N'}(\epsilon')$-packing, where

$$\varepsilon_k^{N'}(\epsilon') = \min\left\{\frac{\eta\sqrt{2}}{2}, \epsilon_{N'}, \tau_k\right\},$$

$$\epsilon_{N'} = \frac{1}{2}\min\left\{\|y - z\| \,|y, z \in \mathcal{Z}_{N'}, y \neq z\right\},$$

$$\tau_k := \frac{(1-\epsilon')\sqrt{\eta\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}}{2}.$$

Moreover, by Corollary 6.5, $(\bar{\mathcal{P}}_k^{\text{out}})_{k\geq 0}$ is an $H(1, \mathcal{P}\cap S(\gamma))$-sequence of cutting. Hence, by Theorem 6.6, there exists $N'' \geq N'$ such that

$$\delta^H(\bar{\mathcal{P}}_{N''-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \leq \frac{4}{\eta}\left(\frac{\varepsilon_{N'}^{N'}(\epsilon')}{1-\epsilon'}\right)^2, \tag{7.6}$$

where $\varepsilon_{N'}^{N'}(\epsilon') > 0$ holds, as $\mathcal{P} \cap S(\gamma)$ is not a polyhedral set. Hence, we get

$$\tau_{N''} = \frac{(1-\epsilon')\sqrt{\eta\delta^H(\bar{\mathcal{P}}_{N''-1}^{\text{out}}, \mathcal{P} \cap S(\gamma))}}{2} \leq \varepsilon_{N'}^{N'}(\epsilon') \leq \min\left\{\frac{\eta\sqrt{2}}{2}, \epsilon_{N'}\right\}.$$

Observe that for every $k \geq N''$, we have $\tau_k \leq \tau_{N''}$; hence,

$$\varepsilon_k^{N'}(\epsilon') = \min\left\{\frac{\eta\sqrt{2}}{2}, \epsilon_{N'}, \tau_k\right\} = \tau_k;$$

in particular, $\mathcal{Z}_k$ is the base of a τ_k-packing. Similar to (7.6), using Theorem 6.6, we may find $N''' \in \mathbb{N}$ such that, for every $k \geq N'''$,

$$\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S(\gamma)) \leq \frac{4(R(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0)))^2}{(1-\epsilon')^2\eta},$$

which is equivalent to

$$\tau_k \leq R(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0)).$$

By Remark 7.6, $\mathcal{Z}_k \subseteq \operatorname{bd}(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0))$. Then, from Lemma 7.10, for every $k \geq \max\{N'', N'''\}$, we get

$$|\mathcal{Z}_k| \leq N_1(\tau_k, R(\mathcal{P} \cap S(\gamma) + \mathbb{B}_\eta(0))).$$

Moreover, by the definition of $R(\cdot)$ and by the choice of η, we have

$$\begin{aligned} N_1(\tau_k, R(\mathcal{P}\cap S(\gamma)+\mathbb{B}_\eta(0))) &= N_1(\tau_k, R(\mathcal{P}\cap S(\gamma))+\eta) \\ &= \frac{q\pi_q}{\pi_{q-1}}\left(1+\left(\frac{2R(\mathcal{P}\cap S(\gamma))}{\tau_k}\right)^2\right)^{\frac{q-1}{2}}. \end{aligned}$$

Since $\mathcal{P}\cap S(\gamma)$ is a nonpolyhedral set, by Lemma 7.7, for every $k\geq 0$, we have

$$|\mathcal{Z}_k| = J+1+k.$$

Hence, for every $k\geq \max\{N'', N'''\}$, we have

$$k\leq |\mathcal{Z}_k| \leq N_1(\tau_k, 2R(\mathcal{P}\cap S(\gamma))) = \frac{q\pi_q}{\pi_{q-1}}\left(1+\frac{16R(\mathcal{P}\cap S(\gamma))}{(1-\epsilon')^2\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S(\gamma))}\right)^{\frac{q-1}{2}},$$

which implies that

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S(\gamma)) \leq \frac{16R(\mathcal{P}\cap S(\gamma))}{(1-\epsilon')^2\left(\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{1-q}}-1\right)}. \tag{7.7}$$

Next, we will show that

$$\frac{16R(\mathcal{P}\cap S(\gamma))}{(1-\epsilon')^2\left(\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{1-q}}-1\right)} \leq (1+\epsilon)16R(\mathcal{P}\cap S(\gamma))\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{q-1}} \tag{7.8}$$

holds for sufficiently large k. Note that (7.8) holds if

$$g(k) := \frac{1}{1-\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{q-1}}} - 1 \leq (1+\epsilon)(1-\epsilon')^2 - 1 \tag{7.9}$$

holds. Here g is a decreasing function on $\mathbb{N}$ with $\lim_{k\to\infty} g(k) = 0$. Moreover, by the choice of ϵ', we have $(1+\epsilon)(1-\epsilon')^2-1>0$. Then there exists $N'''' \in \mathbb{N}$ such that (7.9), and hence (7.8), hold for all $k\geq N''''$.

From (7.7) and (7.8), we obtain that

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S(\gamma)) \leq (1+\epsilon)16R(\mathcal{P}\cap S(\gamma))\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{q-1}}$$

holds for $k\geq \max\{N'', N''', N''''\}$, as desired. ◻

By Theorem 7.2, we prove that the approximation error obtained through the iterations of Algorithm 1, when the Euclidean norm is used in the scalarization, decreases by the order of $\mathcal{O}(k^{2/1-q})$.

8. Examples and computational results. In this section, we consider some numerical examples to observe the performance of Algorithm 1, which is implemented using MATLAB R2018a along with CVX, a package for solving convex programs [25, 24], and *Bensolve tools* [41] for solving the scalarization and vertex enumeration problems in each iteration, respectively. The tests are performed using a 3.6 GHz Intel Core i7 computer with 64 GB RAM.

We consider three examples:

1. Let $q \in \{2,3,4\}$, $e = (1,\dots,1)^\mathsf{T} \in \mathbb{R}^q$.

$$\text{minimize } x \text{ with respect to } \leq_C \text{ subject to } \|x - e\|_2 \leq 1, \; x \in \mathbb{R}^q.$$

2. Let $a^1 = (1,1)^\mathsf{T}, a^2 = (2,3)^\mathsf{T}, a^3 = (4,2)^\mathsf{T}$.

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

3. Let $b^1 = (0,10,-120), b^2 = (80,-448,80), b^3 = (-448,80,80)$.

$$\begin{aligned}&\text{minimize } \left(\|x\|_2^2 + b^1 x, \|x\|_2^2 + b^2 x, \|x\|_2^2 + b^3 x\right)^\mathsf{T} \text{ with respect to} \leq_C\\ &\text{subject to } \|x\|_2^2 \leq 100, \; x_i \geq 0 \text{ for } i \in \{1,2,3\}, \; x \in \mathbb{R}^3.\end{aligned}$$

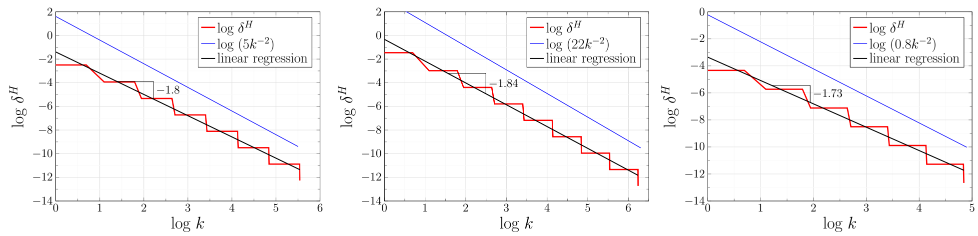
Example 1 is a standard illustrative example with a linear objective function; see [15, 40]. In Example 2, the objective function is nonlinear, while the constraints are linear; in Example 3, nonlinear terms appear both in the objective function and the constraints [15, Examples 5.8 and 5.10], [44]. These examples have also been used recently in [5] to test the performances of similar CVOP algorithms. In addition to the positive orthant, the examples are considered with several other cones (see Figures 2 and 3), given as $C_i = \operatorname{conv}\operatorname{cone} D_i$, $i \in \{1,\dots,6\}$, where

$$\begin{aligned}&D_1 = \{(2,1)^\mathsf{T}, (1,2)^\mathsf{T}\}, \quad D_2 = \{(2,-1)^\mathsf{T}, (-1,2)^\mathsf{T}\},\\ &D_3 = \{(4,2,2)^\mathsf{T}, (2,4,2)^\mathsf{T}, (4,0,2)^\mathsf{T}, (1,0,2)^\mathsf{T}, (0,1,2)^\mathsf{T}, (0,4,2)^\mathsf{T}\},\\ &D_4 = \{(-1,-1,3)^\mathsf{T}, (2,2,-1)^\mathsf{T}, (1,0,0)^\mathsf{T}, (0,-1,2)^\mathsf{T}, (-1,0,2)^\mathsf{T}, (0,1,0)^\mathsf{T}\}.\\ &D_5 = \{(4,2,2,3)^\mathsf{T}, (2,4,2,3)^\mathsf{T}, (4,0,2,3)^\mathsf{T}, (1,0,2,3)^\mathsf{T}, (0,1,2,3)^\mathsf{T}, (0,4,2,3)^\mathsf{T}\},\\ &D_6 = \{(-1,-1,3,-3)^\mathsf{T}, (2,2,-1,3)^\mathsf{T}, (1,0,0,-3)^\mathsf{T}, (0,-1,2,3)^\mathsf{T},\\ &\qquad (-1,0,2,-3)^\mathsf{T}, (0,1,0,3)^\mathsf{T}\}.\end{aligned}$$

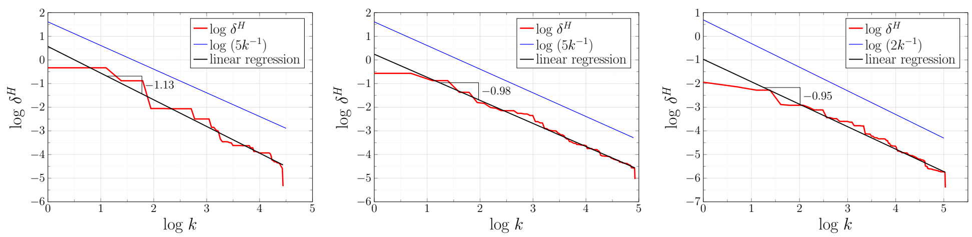
To fix γ satisfying (4.2), we first compute $\sup_{x\in\mathcal{X}} \bar{w}^\mathsf{T}\Gamma(x)$, which is possible by simple geometric observations for each of the examples that we consider. Then we fix $\beta = \sup_{x\in\mathcal{X}} \bar{w}^\mathsf{T}\Gamma(x)$, find an initial outer approximation $\mathcal{P}_0^{\text{out}}$ to $\mathcal{P}$ via (4.1), and compute the Hausdorff distance $\delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P}) = \max\{\|z^v\| \mid v \in \operatorname{ext}\mathcal{P}_0^{\text{out}}\}$, where $\|z^v\|$ is the optimal value of (NM(v)). Finally, we set $\gamma = \beta + \max_{v\in\operatorname{ext}\mathcal{P}_0^{\text{out}}}(\bar{w}^\mathsf{T}v - \beta)^+ + \delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P}) + \epsilon$, where $\epsilon > 0$ is the same as in Algorithm 1.

We run Algorithm 1 for these examples where we use the Euclidean norm within the scalarizations. To observe the convergence behavior, we measure the actual Hausdorff distance between the outer approximation $\bar{\mathcal{P}}_k^{\text{out}}$ and the compact set $\mathcal{P} \cap S(\gamma)$ at each iteration. In Figures 2 and 3, we plot $\log\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P}\cap S(\gamma))$, which is an upper bound on $\log\delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P})$ by Lemma 6.7, versus $\log k$; we compare the linear regression of this graph with the graph of $\log(ck^{\frac{2}{1-q}})$, where $c \in \mathbb{R}_+$ is selected such that the slopes of the lines can be observed easily.[3] Except for Example 1 with $q = 2$, the estimated rate of the decrease in the actual Hausdorff distance is better than the theoretical bound. Note that in Example 1 the image of the feasible region is the unit ball centered at $e \in \mathbb{R}^q$. This implies that the vertices of the outer approximations

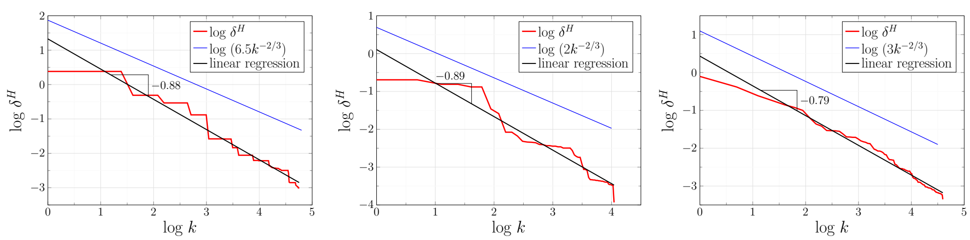
[3] The slopes of the regression lines are indicated in the figures.



(a) $q=2$ with cone $\mathbb{R}^q_+$, $\epsilon=10^{-5}$ (left), narrower cone C_1, $\epsilon=10^{-5}$ (center) and wider cone C_2, $\epsilon=10^{-5}$ (right).



(b) $q=3$ with cone $\mathbb{R}^q_+$, $\epsilon=0.01$ (left), narrower cone C_3, $\epsilon=0.01$ (center) and wider cone C_4, $\epsilon=0.003$ (right).

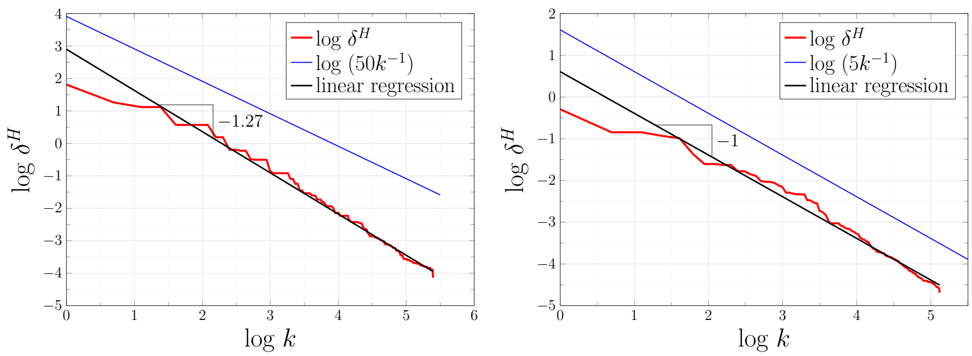


(c) $q=4$ with cone $\mathbb{R}^q_+$, $\epsilon=0.0496$ (left), narrower cone C_5, $\epsilon=0.03$ (center) and wider cone C_6, $\epsilon=0.039$ (right).

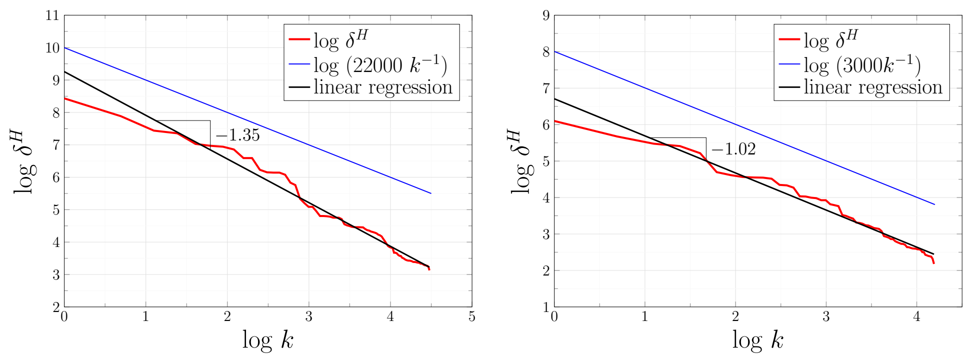
FIG. 2. *Computational results for Example* 1. *We plot* $\log \delta^H(\bar{\mathcal{P}}^{out}_k, \mathcal{P}\cap S(\gamma))$ *versus* $\log k$ *and its linear regression; we compare it with the graph of* $\log(ck^{\frac{2}{1-q}})$ *for some suitable* c. *Except for* $q=2$, *the estimated rate of the decrease in the actual Hausdorff distance is better than the theoretical bound.*

obtained through the iterations of the algorithm are distributed symmetrically with respect to the upper image, and many vertices have the same Euclidean distance to the upper image. Moreover, the number of vertices increase exponentially through the iterations. Hence, the actual Hausdorff distance remains constant for many iterations, which explains the step-function-like behavior observed in Figure 2, especially for $q=2$.

9. Conclusion. To the best of our knowledge, we study the convergence rate of a CVOP algorithm for the first time. The approximation error of the proposed algorithm decreases by the order of $\mathcal{O}(k^{1/(1-q)})$, where q is the dimension of the objective space. Moreover, we consider the special case of using the Euclidean norm within the scalarizations separately and prove that the approximation error decreases by the order of $\mathcal{O}(k^{2/(1-q)})$. Since it is known that in multiobjective optimization ($C=\mathbb{R}^q_+$) this is the best possible convergence rate [34, 26], it cannot be improved in general. It is an open problem to prove the improved convergence rate for the proposed algorithm in the general setting when the underlying norm in the scalarizations is not necessarily the Euclidean norm.



(a) Example 2 with cone $\mathbb{R}^3_+$, $\epsilon = 0.02$ (left) and wider cone C_4, $\epsilon = 0.01$ (right).



(b) Example 3 with cone $\mathbb{R}^3_+$, $\epsilon = 25$ (left) and wider cone C_4, $\epsilon = 10$ (right).

FIG. 3. *Computational results for Examples* 2 *and* 3. *We plot* $\log \delta^H(\bar{\mathcal{P}}^{out}_k, \mathcal{P} \cap S(\gamma))$ *versus* $\log k$ *and its linear regression; we compare it with the graph of* $\log(ck^{-1})$ *for some suitable* c. *The estimated rate of the decrease in the actual Hausdorff distance is better than the theoretical bound.*

Appendix A. Derivations and proofs for section 3. Let $v \in \mathbb{R}^q$. Note that $(\mathrm{P}(v))$ stated in section 3 can be rewritten as

$$(\mathrm{P}'(v)) \qquad \text{minimize } f(x,z) \text{ subject to } G(x,z) \cap -(C \times \mathbb{R}_+) \neq \emptyset,\ (x,z) \in X \times \mathbb{R}^q,$$

where $f\colon X \times \mathbb{R}^q \to \overline{\mathbb{R}}$ is a scalar function and $G\colon X \times \mathbb{R}^q \rightrightarrows \mathbb{R}^{q+1}$ is a set-valued function defined by

$$f(x,z) := \|z\| + I_{\mathcal{X}}(x), \quad G(x,z) := \{(\Gamma(x) - z - v, \bar{w}^{\mathsf{T}}(v+z) - \gamma)\} \tag{A.1}$$

for each $x \in X$, $z \in \mathbb{R}^q$. Using the equivalent problem $(\mathrm{P}'(v))$ and following steps similar to those for constructing the Lagrange dual $(\mathrm{dNM}(v))$ of $(\mathrm{NM}(v))$ (see [5, section 4]), the Lagrange dual of $(\mathrm{P}(v))$ can be formulated as $(\mathrm{D}(v))$ (see section 3). Then the optimal value of $(\mathrm{D}(v))$ is given by

$$\sup_{(w,\lambda) \in C^+ \times \mathbb{R}_+} \phi(w,\lambda) = \sup_{\substack{(w,\lambda) \in C^+ \times \mathbb{R}_+ : \\ \|w - \lambda \bar{w}\|_* \leq 1}} \left(\inf_{x \in \mathcal{X}} w^{\mathsf{T}}(\Gamma(x) - v) + \lambda(\bar{w}^{\mathsf{T}} v - \gamma) \right),$$

where we use the fact that the conjugate function of $\|\cdot\|$ is the indicator function of the unit ball of the dual norm $\|\cdot\|_*$; see [10, Example 3.26, p. 93]. Next, we define $\tilde{w} := w - \lambda \bar{w}$ and rewrite the value of $(\mathrm{D}(v))$ as

$$\sup\left\{\inf_{x\in\mathcal{X}} \tilde{w}^{\mathsf{T}}(\Gamma(x)-v)+\lambda(\bar{w}^{\mathsf{T}}\Gamma(x)-\gamma) \,|\, \|\tilde{w}\|_*\leq 1,\ \tilde{w}+\lambda\bar{w}\in C^+, \lambda\geq 0\right\}. \tag{A.2}$$

We continue with the proofs of the results in section 3.

Proof of Proposition 3.3. First, we show that there exists a feasible solution $(\tilde{x},\tilde{z})$ of $(\mathrm{P}(v))$. Fix $\tilde{x}\in\mathcal{X}$, which exists since $\operatorname{int}\mathcal{X}\neq\emptyset$ is assumed. Let us define $\tilde{z}:=\Gamma(\tilde{x})-v$. Then $v+\tilde{z}=\Gamma(\tilde{x})\in\operatorname{int}S(\gamma)\subseteq S(\gamma)$ and $v+\tilde{z}=\Gamma(\tilde{x})\in\{\Gamma(\tilde{x})\}+C$, that is, $\Gamma(\tilde{x})\leq_C v+\tilde{z}$. Hence, $(\tilde{x},\tilde{z})$ is feasible.

Next, note that the feasible region of $(\mathrm{P}(v))$ is a subset of $\mathcal{X}\times((\mathcal{P}\cap S(\gamma))-\{v\})\subseteq\mathbb{R}^{n+q}$. Moreover, $\mathcal{X}$ is compact by the problem definition and, using [5, Lemma 6.3], it is easy to see that $\mathcal{P}\cap S(\gamma)$ is compact as well. Hence, $\mathcal{X}\times(\mathcal{P}\cap S(\gamma)-\{v\})$ is a compact set, which implies that the feasible region of $(\mathrm{P}(v))$ is bounded. To show that it is also closed, consider a sequence of feasible points $(x_n,z_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty}(x_n,z_n)=(x,z)$. We have $x\in\mathcal{X}$ as $\mathcal{X}$ is compact and $z\in S(\gamma)-v$ as $S(\gamma)$ is a closed halfspace. Since $\Gamma\colon X\to\mathbb{R}^q$ is continuous, C is a closed cone, and $v+z_n-\Gamma(x_n)\in C$ for every $n\in\mathbb{N}$, we obtain $\lim_{n\to\infty}(v+z_n-\Gamma(x_n))=v+z-\Gamma(x)\in C$. Hence, an optimal solution (x^*,z^*) of $(\mathrm{P}(v))$ exists by the Weierstrass theorem.

For strong duality, we show that the following constraint qualification in [33, Theorem 8.3.10, p. 330] and [38, Theorem 3.19, p. 93] holds for $(\mathrm{P}(v))$:

$$G(\operatorname{dom} f)\cap -\operatorname{int}(C\times\mathbb{R}_+)\neq\emptyset, \tag{A.3}$$

where f and G are as in (A.1), and $\operatorname{dom} f:=\{(x,z)\in X\times\mathbb{R}^q | f(x,z)<+\infty\}=\mathcal{X}\times\mathbb{R}^q$. Since $\mathcal{X}$ has nonempty interior, there exists $x^0\in\operatorname{int}\mathcal{X}$. Moreover, since $\Gamma(\mathcal{X})\subseteq\operatorname{int}S(\gamma)$, there exists $\epsilon>0$ such that $\mathbb{B}_\epsilon(\Gamma(x^0))\subseteq\operatorname{int}S(\gamma)$. Let $c\in\operatorname{int}C$ be fixed, and define $y^0:=\Gamma(x^0)+\frac{\epsilon}{2}\frac{c}{\|c\|}$. Clearly, $y^0\in\mathbb{B}_\epsilon(\Gamma(x^0))\cap(\{\Gamma(x^0)\}+\operatorname{int}C)$. Defining $z^0:=y^0-v$, we obtain $v+z^0-\Gamma(x^0)\in\operatorname{int}C$. Moreover, since $y^0\in\mathbb{B}_\epsilon(\Gamma(x^0))\subseteq\operatorname{int}S(\gamma)$, we have $\bar{w}^{\mathsf{T}}(v+z^0)<\gamma$. Then

$$G(x^0,z^0)=\{(\Gamma(x^0)-z^0-v,\bar{w}^{\mathsf{T}}(v+z^0)-\gamma)\}\subseteq -\operatorname{int}(C\times\mathbb{R}_+).$$

On the other hand, since $\operatorname{dom} f=\mathcal{X}\times\mathbb{R}^q$, we have $G(x^0,z^0)\subseteq G(\operatorname{dom} f)$. Hence, (A.3) is satisfied.

Recall that $\mathcal{X}\subseteq\mathbb{R}^n$ is convex and $\Gamma\colon X\to\mathbb{R}^q$ is a C-convex function. Then, by standard arguments, it can be shown that $f: X\times\mathbb{R}^q\to\overline{\mathbb{R}}$ is convex and $G\colon X\times\mathbb{R}^q\rightrightarrows\mathbb{R}^{q+1}$ is $(C\times\mathbb{R}_+)$-convex as a set-valued function. Together with (A.3), these imply strong duality and dual attainment by [38, Theorem 3.19, p. 93]. ∎

Proof of Lemma 3.5. By the feasibility of $(\mathrm{P}(v))$ and strong duality, $z^v\neq 0$ and $w^v\neq 0$ in (a); see the proof of [5, Lemma 4.5] for details. For the other claim of (a), assume to the contrary that $w^v-\lambda^v\bar{w}=0$. From the optimal value of $(\mathrm{D}(v))$ and strong duality, we get

$$\lambda^v\left(\inf_{x\in\mathcal{X}}\bar{w}^{\mathsf{T}}\Gamma(x)-\gamma\right)=\|z^v\|,$$

where $\lambda^v\geq 0$ and $\bar{w}^{\mathsf{T}}\Gamma(x)-\gamma\leq 0$ for any $x\in\mathcal{X}$, as $\Gamma(\mathcal{X})\subseteq S(\gamma)$. Hence, $\|z^v\|\leq 0$. This implies $\|z^v\|=0$, a contradiction.

To prove (b) and (c), note that we have $\bar{w}^{\mathsf{T}}v\leq\gamma$. Then the assertion follows using arguments similar to those in the proof of [5, Lemma 4.5]. ∎

Proof of Proposition 3.6. By using arguments similar to those in the proof of [5, Proposition 4.6], x^v is a weak minimizer of (P). For the second claim, first note that $y^v \in \mathcal{P}$ since (x^v, z^v) is feasible for $(\mathrm{P}(v))$. To get a contradiction, assume that $y^v \notin \mathrm{wMin}_C\,\mathcal{P}$; hence, $y^v = v + z^v \in \operatorname{int}\mathcal{P}$ with $\|z^v\| \neq 0$ since $v \notin \operatorname{int}\mathcal{P}$ by our assumption. Then there exists $0 < \epsilon \leq \|z^v\|$ such that

$$v + z^v - \epsilon \frac{z^v}{\|z^v\|} \in \mathcal{P},$$

which implies the existence of $\bar{x} \in \mathcal{X}$ with

$$v + (\|z^v\| - \epsilon)\frac{z^v}{\|z^v\|} \in \{\Gamma(\bar{x})\} + C.$$

Let $\bar{z} := (\|z^v\| - \epsilon)\frac{z^v}{\|z^v\|}$. We show that $\bar{w}^\mathsf{T}(v + \bar{z}) \leq \gamma$. Assume otherwise that

$$\bar{w}^\mathsf{T}(v + \bar{z}) = \bar{w}^\mathsf{T}\left(v + z^v - \epsilon\frac{z^v}{\|z^v\|}\right) > \gamma. \tag{A.4}$$

Since $\bar{w}^\mathsf{T}(v + z^v) \leq \gamma$ by the feasibility of $(\mathrm{P}(v))$,

$$-\epsilon\frac{\bar{w}^\mathsf{T} z^v}{\|z^v\|} > \gamma - \bar{w}^\mathsf{T}(v + z^v) \geq \gamma - \gamma = 0.$$

This implies $\bar{w}^\mathsf{T} z^v < 0$. We also know that $\bar{w}^\mathsf{T} v \leq \gamma$, as $v \in S(\gamma)$. Then, from (A.4), we get

$$\bar{w}^\mathsf{T} z^v \left(1 - \frac{\epsilon}{\|z^v\|}\right) > \gamma - \bar{w}^\mathsf{T} v \geq \gamma - \gamma = 0.$$

This implies $1 - \frac{\epsilon}{\|z^v\|} < 0$, that is, $\|z^v\| < \epsilon$. This contradicts the upper bound of ϵ.

Now $(\bar{x}, \bar{z})$ is feasible for $(\mathrm{P}(v))$. But this gives $\|\bar{z}\| < \|z^v\|$, which is a contradiction to the optimality of (x^v, z^v). ∎

Proof of Proposition 3.7. Let us fix $y \in \mathcal{P} \cap S(\gamma)$ consider the problems $(\mathrm{P}(v))$ and $(\mathrm{D}(v))$. Then

$$0 \geq \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T}(\Gamma(x) - y) + \lambda^v(\bar{w}^\mathsf{T} y - \gamma);$$

see the proof of [5, Proposition 4.7]. This gives

$$(w^v)^\mathsf{T} y - \lambda^v \bar{w}^\mathsf{T} y \geq \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T}\Gamma(x) - \lambda^v\gamma. \tag{A.5}$$

Using the strong duality between $(\mathrm{P}(v))$ and $(\mathrm{D}(v))$, we get

$$\begin{aligned} \|z^v\| &= \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T}(\Gamma(x) - v) + \lambda^v(\bar{w}^\mathsf{T} v - \gamma) \\ &= \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T}\Gamma(x) - \lambda^v\gamma - (w^v)^\mathsf{T} v + \lambda^v \bar{w}^\mathsf{T} v. \end{aligned} \tag{A.6}$$

From (A.5) and (A.6), we obtain

$$(w^v)^\mathsf{T} y - \lambda^v \bar{w}^\mathsf{T} y \geq \|z^v\| + (w^v)^\mathsf{T} v - \lambda^v \bar{w}^\mathsf{T} v,$$

that is,

$$(\tilde{w}^v)^\mathsf{T} y \geq \|z^v\| + (\tilde{w}^v)^\mathsf{T} v \geq (\tilde{w}^v)^\mathsf{T} z^v + (\tilde{w}^v)^\mathsf{T} v = (\tilde{w}^v)^\mathsf{T} y^v,$$

where the second inequality is a consequence of the definition of dual norm and the dual constraint $\|\tilde{w}^v\|_* \leq 1$. Hence, $y \in \mathcal{H}(\tilde{w}^v, y^v)$. Since y is arbitrary, this implies $\mathcal{P} \cap S(\gamma) \subseteq \mathcal{H}(\tilde{w}^v, y^v)$.

Next, by Proposition 3.6, we have $y^v \in \mathcal{P}$; in particular, $y^v \in \mathcal{P} \cap S(\gamma)$ by the feasibility of (P(v)). We have $y^v \in \operatorname{bd} \mathcal{H}(\tilde{w}^v, y^v)$ by the definition of $\mathcal{H}(\tilde{w}^v, y^v)$. Hence, $y^v \in \mathcal{P} \cap S(\gamma) \cap \operatorname{bd} \mathcal{H}(\tilde{w}^v, y^v)$, which completes the proof. $\square$

REFERENCES

[1] C. D. Aliprantis and K. C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, Springer, 2006.

[2] Ç. Ararat and N. Meimanjan, *Computation of systemic risk measures: A mixed-integer programming approach*, Oper. Res., 71 (2023), pp. 2130–2145.

[3] Ç. Ararat, S. Tekgül, and F. Ulus, *Geometric duality results and approximation algorithms for convex vector optimization problems*, SIAM J. Optim., 33 (2023), pp. 116–146, https://doi.org/10.1137/21M1458788.

[4] Ç. Ararat and C. Tekin, *Vector optimization with stochastic bandit feedback*, in Proceedings of the 26th International Conference on Artificial Intelligence and Statistics, Vol. 206, F. Ruiz, J. Dy, and J.-W. van de Meent, eds., PMLR, 2023, pp. 2165–2190.

[5] Ç. Ararat, F. Ulus, and M. Umer, *A norm minimization-based convex vector optimization algorithm*, J. Optim. Theory Appl., 194 (2022), pp. 681–712.

[6] P. Armand, *Finding all maximal efficient faces in multiobjective linear programming*, Math. Program., 61 (1993), pp. 357–375.

[7] H. P. Benson, *Concave minimization: Theory, applications and algorithms*, in Handbook of Global Optimization, R. Horst and P. M. Pardalos, eds., Springer, 1995, pp. 43–148.

[8] H. P. Benson, *An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem*, J. Global Optim., 13 (1998), pp. 1–24.

[9] H. P. Benson and R. Horst, *A branch and bound-outer approximation algorithm for concave minimization over a convex set*, Comput. Math. Appl., 21 (1991), pp. 67–76.

[10] S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[11] A. Conci and C. Kubrusly, *Distances between sets—a survey*, Adv. Math. Sci. Appl., 26 (2017), pp. 1–18.

[12] M. De Santis, G. Eichfelder, J. Niebling, and S. Rocktäschel, *Solving multiobjective mixed integer convex optimization problems*, SIAM J. Optim., 30 (2020), pp. 3122–3145, https://doi.org/10.1137/19M1264709.

[13] D. Dörfler, A. Löhne, C. Schneider, and B. Weissing, *A Benson-type algorithm for bounded convex vector optimization problems with vertex selection*, Optim. Methods Softw., 37 (2022), pp. 1006–1026.

[14] M. Ehrgott and L. Shao, *Approximately solving multiobjective linear programmes in objective space and an application in radiotherapy treatment planning*, Math. Methods Oper. Res., 68 (2008), pp. 257–276.

[15] M. Ehrgott, L. Shao, and A. Schöbel, *An approximation algorithm for convex multi-objective programming problems*, J. Global Optim., 50 (2011), pp. 397–416.

[16] J. P. Evans and R. E. Steuer, *A revised simplex method for linear multiple objective programs*, Math. Programming, 5 (1973), pp. 54–72.

[17] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, 2007.

[18] Z. Feinstein and B. Rudloff, *A recursive algorithm for multivariate risk measures and a set-valued Bellman's principle*, J. Global Optim., 68 (2017), pp. 47–69.

[19] Z. Feinstein and B. Rudloff, *Technical note—Characterizing and computing the set of Nash equilibria via vector optimization*, Oper. Res., (2023), https://doi.org/10.1287/opre.2023.2457.

[20] J. Fliege, L. M. Graña Drummond, and B. F. Svaiter, *Newton's method for multiobjective optimization*, SIAM J. Optim., 20 (2009), pp. 602–626, https://doi.org/10.1137/08071692X.

[21] J. Giesen, S. Laue, A. Löhne, and C. Schneider, *Using Benson's algorithm for regularization parameter tracking*, in Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33, 2019, pp. 3689–3696.

[22] J. Giesen, F. Nussbaum, and C. Schneider, *Efficient regularization parameter selection for latent variable graphical models via bi-level optimization*, in IJCAI, 2019, pp. 2378–2384.

[23] M. A. Goberna, A. Iusem, J. E. Martinez-Legaz, and M. I. Todorov, *Motzkin decomposition of closed convex sets via truncation*, J. Math. Anal. Appl., 400 (2013), pp. 35–47.

[24] M. C. Grant and S. P. Boyd, *Graph implementations for nonsmooth convex programs*, in Recent Advances in Learning and Control, Lect. Notes Control Inf. Sci. 371, V. D. Blondel, S. P. Boyd, and H. Kimura, eds., Springer, 2008, pp. 95–110, https://web.stanford.edu/~boyd/papers/graph_dcp.html.

[25] M. C. Grant and S. P. Boyd, *CVX: Matlab Software for Disciplined Convex Programming, Version* 2.2, 2020, http://cvxr.com/cvx.

[26] P. M. Gruber, *Aspects of approximation of convex bodies*, in Handbook of Convex Geometry, Vols. A, B, P. M. Gruber and J. M. Wills, eds., Elsevier, 1993, pp. 319–345.

[27] A. H. Hamel, A. Löhne, and B. Rudloff, *Benson type algorithms for linear vector optimization and applications*, J. Global Optim., 59 (2014), pp. 811–836.

[28] J. Jahn, *Scalarization in vector optimization*, Math. Programming, 29 (1984), pp. 203–218.

[29] G. K. Kamenev, *A class of adaptive algorithms for approximating convex bodies by polyhedra*, Comput. Math. Math. Phys., 32 (1992), pp. 114–127.

[30] G. K. Kamenev, *Conjugate adaptive algorithms for polyhedral approximation of convex bodies*, Comput. Math. Math. Phys., 42 (2002), pp. 1301–1316.

[31] G. K. Kamenev and A. I. Pospelov, *Polyhedral approximation of convex compact bodies by filling methods*, Comput. Math. Math. Phys., 52 (2012), pp. 680–690.

[32] I. N. Keskin and F. Ulus, *Outer approximation algorithms for convex vector optimization problems*, Optim. Methods Softw., 38 (2023), pp. 723–755.

[33] A. A. Khan, C. Tammer, and C. Zălinescu, *Set-Valued Optimization: An Introduction with Applications*, Springer, 2016.

[34] K. Klamroth and J. Tind, *Constrained optimization using multiple objective programming*, J. Global Optim., 37 (2007), pp. 325–355.

[35] K. Klamroth, J. Tind, and M. M. Wiecek, *Unbiased approximation in multicriteria optimization*, Math. Methods Oper. Res., 56 (2003), pp. 413–437.

[36] G. Kováčová and B. Rudloff, *Time consistency of the mean-risk problem*, Oper. Res., 69 (2021), pp. 1100–1117.

[37] G. Kováčová and B. Rudloff, *Convex projection and convex multi-objective optimization*, J. Global Optim., 83 (2022), pp. 301–327.

[38] A. Löhne, *Vector Optimization with Infimum and Supremum*, Springer, 2011.

[39] A. Löhne and B. Rudloff, *An algorithm for calculating the set of superhedging portfolios in markets with transaction costs*, Int. J. Theor. Appl. Finance, 17 (2014), 1450012.

[40] A. Löhne, B. Rudloff, and F. Ulus, *Primal and dual approximation algorithms for convex vector optimization problems*, J. Global Optim., 60 (2014), pp. 713–736.

[41] A. Löhne and B. Weissing, *The vector linear program solver Bensolve–notes on theoretical background*, European J. Oper. Res., 260 (2017), pp. 807–813.

[42] A. V. Lotov, V. A. Bushenkov, and G. K. Kamenev, *Interactive Decision Maps: Approximation and Visualization of Pareto Frontier*, Appl. Optim. 89, Springer, 2004.

[43] K. Miettinen and M. M. Mäkelä, *On generalized trade-off directions in nonconvex multiobjective optimization*, Math. Program., 92 (2002), pp. 141–151.

[44] K. Miettinen, M. M. Mäkelä, and K. Kaario, *Experiments with classification-based scalarizing functions in interactive multiobjective optimization*, European J. Oper. Res., 175 (2006), pp. 931–947.

[45] J. Niebling and G. Eichfelder, *A branch-and-bound-based algorithm for nonconvex multiobjective optimization*, SIAM J. Optim., 29 (2019), pp. 794–821, https://doi.org/10.1137/18M1169680.

[46] S. Nobakhtian and N. Shafiei, *A Benson type algorithm for nonconvex multiobjective programming problems*, TOP, 25 (2017), pp. 271–287.

[47] A. Pascoletti and P. Serafini, *Scalarizing vector optimization problems*, J. Optim. Theory Appl., 42 (1984), pp. 499–524.

[48] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.

[49] B. Rudloff and F. Ulus, *Certainty equivalent and utility indifference pricing for incomplete preferences via convex vector optimization*, Math. Financ. Econ., 15 (2021), pp. 397–430.

[50] B. Rudloff, F. Ulus, and R. Vanderbei, *A parametric simplex algorithm for linear vector optimization problems*, Math. Program., 163 (2017), pp. 213–242.

[51] A. Ruszczyński and R. J. Vanderbei, *Frontiers of stochastically nondominated portfolios*, Econometrica, 71 (2003), pp. 1287–1297.

[52] Y. Sawaragi, H. Nakayama, and T. Tanino, *Theory of Multiobjective Optimization*, Elsevier, 1985.

[53] L. Shao, F. Zhao, and G. Hu, *A numerical method for the approximation of reachable sets of linear control systems*, IMA J. Math. Control Inform., 36 (2019), pp. 423–441.

[54] A. Wagner, F. Ulus, B. Rudloff, G. Kováčová, and N. Hey, *Algorithms to solve unbounded convex vector optimization problems*, SIAM J. Optim., 33 (2023), pp. 2598–2624, https://doi.org/10.1137/22M1507693.