NORM MINIMIZATION-BASED CONVEX VECTOR OPTIMIZATION ALGORITHMS

A DISSERTATION SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
INDUSTRIAL ENGINEERING

By
Muhammad Umer
August 2022

Norm minimization-based convex vector optimization algorithms

By Muhammad Umer

August 2022

We certify that we have read this dissertation and that in our opinion it is fully adequate, in scope and in quality, as a dissertation for the degree of Doctor of Philosophy.

Firdevs Ulus(Advisor)

Çağın Ararat(Co-Advisor)

Oya Ekin Karaşan

Özlem Karsu

Refail Kasimbeyli

Mehmet Rüştü Taner

Approved for the Graduate School of Engineering and Science:

Orhan Arıkan
Director of the Graduate School

ABSTRACT

NORM MINIMIZATION-BASED CONVEX VECTOR OPTIMIZATION ALGORITHMS

Muhammad Umer
Ph.D. in Industrial Engineering
Advisor: Firdevs Ulus
Co-Advisor: Çağın Ararat
August 2022

This thesis is concerned with convex vector optimization problems (CVOP). We propose an outer approximation algorithm (Algorithm 1) for solving CVOPs. In each iteration, the algorithm solves a norm-minimizing scalarization for a reference point in the objective space. The idea is inspired by some Benson-type algorithms in the literature that are based on Pascoletti-Serafini scalarization. Since this scalarization needs a direction parameter, the efficiency of these algorithms depend on the selection of the direction parameter. In contrast, our algorithm is free of direction biasedness since it solves a scalarization that is based on minimizing a norm. However, the structure of such algorithms, including ours, has some built-in limitation which makes it difficult to perform convergence analysis. To overcome this, we modify the algorithm by introducing a suitable compact subset of the upper image. After the modification, we have Algorithm 2 in which norm-minimizing scalarizations are solved for points in the compact set. To the best of our knowledge, Algorithm 2 is the first algorithm for CVOPs, which is proven to be finite. Finally, we propose a third algorithm for the purposes of convergence analysis (Algorithm 3), where a modified norm-minimizing scalarization is solved in each iteration. This scalarization includes an additional constraint which ensures that the algorithm deals with only a compact subset of the upper image from the beginning. Besides having the finiteness result, Algorithm 3 is the first CVOP algorithm with an estimate of a convergence rate. The experimental results, obtained using some benchmark test problems, show comparable performance of our algorithms with respect to an existing CVOP algorithm based on Pascoletti-Serafini scalarization.

Keywords: convex vector optimization, multiobjective optimization, approximation algorithm, scalarization, norm minimization.

ÖZET

NORM ENKÜÇÜKLEMEYE DAYALI DIŞBÜKEY VEKTÖR ENİYİLEME ALGORİTMALARI

Muhammad Umer
Endüstri Mühendisliği, Doktora
Tez Danışmanı: Firdevs Ulus
İkinci Tez Danışmanı: Çağın Ararat
Ağustos 2022

Bu tezde dışbükey vektör eniyileme problemleri (DVEP) çalışılmıştır. İlk olarak DVEP'leri çözmek için bir dış yaklaşıklama algoritması (Algoritma 1) önerilmiştir. Algoritmanın her yinelemesinde, amaç uzayındaki bir referans noktası için normu enküçükleyen bir skalerizasyon problemi çözülmektedir. Algoritmanın temel yapısı, bilimsel yazında Pascoletti-Serafini skalerizasyonuna dayanan bazı Benson tipi algoritmalara dayanmaktadır. Pascoletti-Serafini skalerizasyonu bir yön parametresine ihtiyaç duyduğundan, bu algoritmaların verimliliği yön parametresinin seçimine bağlıdır. Buna karşılık, bizim önerdiğimiz Algoritma 1, bir normu enküçüklemeye dayalı bir skalerizasyon problemi çözdüğü için yön yanlılığından arındırılmıştır. Bununla birlikte, bizimki de dahil olmak üzere bu tür algoritmalar, yapılarından dolayı, yakınsama analizi yapmayı zorlaştıran bazı kısıtlamalara sahiptir. Bunun üstesinden gelmek için, üst görüntünün uygun bir tıkız altkümesi oluşturularak Algoritma 1 değiştirilmiştir. Değişiklikten sonra, tıkız kümedeki noktalar için norm enküçükleyen skalerizasyon problemlerinin çözüldüğü Algoritma 2 elde edilmiştir. Bildiğimiz kadarıyla, Algoritma 2, DVEP'ler için sonlu olduğu kanıtlanmış ilk algoritmadır. Son olarak yakınsama analizi yapmanın mümkün olduğu üçüncü bir algoritma (Algoritma 3) geliştirilmiştir. Bu algoritmada da norm enküçükleyen bir skalerizasyon problemi çözülmektedir; ancak problem, önceki algoritmalardaki skalerizasyon problemleriyle aynı değildir, onlardan farklı olarak algoritmanın baştan itibaren üst görüntünün yalnızca tıkız bir altkümesiyle ilgilenmesini sağlayan ek bir kısıt içerir. Algoritma 3, sonluluk sonucuna sahip olmanın yanı sıra, yakınsama oranı tahmini olan ilk DVEP algoritmasıdır. Bazı test problemleri kullanılarak elde edilen deneysel sonuçlar, Pascoletti-Serafini skalerizasyonuna dayalı mevcut bir DVEP algoritmasına göre algoritmalarımızın karşılaştırılabilir performansa sahip olduğunu göstermektedir.

Anahtar sözcükler: dışbükey vektör eniyileme, çokamaçlı eniyileme, yaklaşıklama algoritmaları, skalerizasyon, norm enküçükleme.

Acknowledgement

I would like to express my deepest gratitude to my advisors Asst. Prof. Firdevs Ulus and Asst. Prof. Cağın Ararat for their continuous guidance and never-ending patience and support. I feel extremely lucky that they took me under their supervision. I am extremely grateful that they trusted me and walked me through the subject from the very basics till the completion of this dissertation. The endeavor would not have been possible without their endless guidance and support. I am proud of working under their invaluable supervision.

I am deeply indebted to department chairs Prof. Selim Aktürk and Prof. Bahar Yetiş Kara for their kind and continuous support which always remained available through their prompt and favorable responses. I am highly grateful to my thesis committee members Prof. Refail Kasımbeyli, Prof. Oya Ekin Karaşan, Assoc. Prof. Özlem Karsu and Assoc. Prof. Mehmet Rüştü Taner for their valuable suggestions and insights which help a lot to improve and refine my work. I am also grateful to the faculty members of the department, especially the ones I have taken courses from.

My words cannot express my gratitude to my best friend and mentor Milad Malekipirbazari who has taught me the subject from the scratch covering all the aspects and knowledge needed for my research. I am extremely grateful for his time and countless evenings he spend explaining the topics, in the office and cafes, of which I had no prior knowledge whatsoever. Besides his endless technical support, he also supported me in pressing times I have faced throughout this journey. I have found him always sitting by my side whenever I needed a shoulder to cry on. I would like to extend my sincere thanks to Parinaz Toufani for her valuable advices and her kindness which she generously extended to me and my family in our dire times. My special thanks to my friend Can Telkenaroğlu who acted as a tour guide throughout my stay in Turkey. Thanks to him for pinpointing good restaurants where my family and I enjoyed a lot of Turkish cuisines.

I would like to acknowledge my friends who were always willing to support and encourage me by taking their time out of their busy routines. These include Erman Gözü, Bashir Abdullahi Bashir, Ece Çiğdem Karakoyun, Halil İbrahim

Bayrak, Kamyar Kargar, Özlem Yılmaz, Simay Tekgül and Onur Kılınç in Bilkent and Ammar Saleem, Safdar ud Din, Saad Javaid and Muhammad Amjad back in my home country. My special thanks to my mentors sir Messam Abbas Naqvi, sir Farooq bin Akram and sir Faisal Siddiqui whose support and guidance have given me confidence and courage to complete this challenging task.

Finally, I would extend my gratitude to my parents and my family. I am what I am today because of my parents. They have invested their lives in me and I would never be able to pay them back for what they have sacrificed for me. I am also grateful to my brother Muhammad Fahad for his continuous support. My wife and son are the ones who suffered the most in this journey. The patience they have shown and the compromises they have made are beyond anybody's imagination. This work would not have been possible without their continuous patience, support and encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, we discuss the literature review regarding multiobjective optimization problems (MOP), vector optimization problems (VOP), algorithms available for such problems and the scalarization methods employed by the available algorithms. Then, we discuss the contributions of this thesis compared to the existing literature.

1.1 Motivation

In multiobjective optimization, the decision-maker is supposed to consider multiple objective functions simultaneously. In general, these functions conflict in the sense that improving one objective leads to deteriorating some of the others. Consequently, there does not exist a feasible solution which can generate optimal values of all the objectives. Rather, there exists a subset of feasible solutions, called *efficient solutions*, which map to the so called *nondominated points* in the objective space. The image of a feasible solution is said to be nondominated if none of the objective functions can be improved in value without degrading some of the other objective values. The concept of efficient solutions is first introduced by Koopmans [2], although the idea can be found in the earlier works by Edgeworth in [3] and Pareto in [4], see [5, Section 2.7] and [6, Section 2.4].

In vector optimization, the objective function takes values again in a vector space, namely, the objective space. However, rather than comparing the objective function values componentwise as in the multiobjective case, a more general order relation, which is induced by an ordering cone, is used for this purpose. Clearly, multiobjective optimization can be seen as a special case where the ordering cone is the positive orthant. Assuming that the VOP is a minimization problem with respect to an ordering cone C, the concept of nondominated point for a MOP is generalized to *minimal point with respect to* C in the vector optimization case.

A special class of vector optimization problems is the linear VOPs, where the objective function is linear and the feasible region is a polyhedron. There is rich literature available discussing various methods and algorithms for solving linear VOPs. They deal with the problem by generating either the efficient solutions in the decision space [7, 8] or the nondominated points in the objective space [9, 5, 10]. The reader is referred to the books by Ehrgott [11] and by Jahn [12] for the details of these approaches.

In 1998, Benson proposed an outer approximation algorithm for linear MOPs which generates the set of all nondominated points in the objective space rather than the set of all efficient points in the decision space [9]. The idea of this approach can be seen as similar to a method for polyhedral approximation of convex sets, proposed by Kamenev, [13], in 1992. Later, the algorithm in [9] is extended to solve linear VOPs, see [5]. The main principle of the algorithm is that if one adds the ordering cone to the image of the feasible set, then the resulting set, called the *upper image*, contains all nondominated points in its boundary. The algorithm starts with a set containing the upper image and iterates by updating the outer approximating set until it is equal to the upper image.

For nonlinear MOPs/VOPs, there is a further subdivision, namely the convex and nonconvex problems. Note that the methods described for linear MOPs/VOPs may not be directly applicable to these classes as, in general, it is not possible to generate the set of all nondominated/minimal points in the objective space. Therefore, approximation algorithms, which approximate the set of all minimal points in the objective space, are widely explored in the literature, refer for example to the survey paper by Ruzika and Wiecek [14] for the multiobjective case. Nonconvex VOPs are out of the scope of this study, therefore we will not go into it further. However, the details of such methods dealing with nonconvex VOPs can be found in [15, 16, 17, 18, 19].

For bounded convex vector optimization problems (CVOPs), see Section 2.2 for precise definitions, there are several outer approximation algorithms in the literature that work on the objective space. In [20], the algorithm in [9] is extended for the case of convex MOPs. Another extension of Benson's algorithm for the vector optimization case is proposed in [1], which is a simplification and generalization of the algorithm in [20]. This algorithm has already been used for solving mixed-integer convex multiobjective optimization problems [21], as well as problems in stochastic optimization [22] and finance [23, 24]. Recently, in [25], a modification of the algorithm in [1] is proposed. The main idea of these algorithms is to generate a sequence of better approximating polyhedral supersets of the upper image until the approximation is sufficiently fine. This is done by sequentially solving some scalarization models in which the original CVOP is converted into an optimization problem with a single objective. There are many scalarization methods available in the literature for MOPs/VOPs, for instance, ε-constraint, weighted Chebyshev and Benson scalarizations for LVOPs and weighted sum, Pascoletti-Serafini and conic scalarizations for general VOPs. The reader is referred to the book by Eichfelder [26] as well as the recent papers [27, 28, 29] for further details of scalarization methods.

In each iteration of the CVOP algorithms proposed in [25, 20, 1], a *Pascoletti-Serafini scalarization* [30], which requires a reference point v and a direction vector d in the objective space as its parameters, is solved. For the algorithms in [20] and [1], the reference point v is selected to be an arbitrary vertex of the current outer approximation of the upper image. Moreover, in [20], the direction parameter d is computed depending on the reference point v together with a fixed point in the objective space, whereas it is fixed throughout the algorithm proposed in [1]. In [25], a procedure to select a vertex v as well as a direction parameter d, which depends on v and the current approximation, is proposed.

1.2 Norm-Minimizing Scalarization based CVOP Algorithm

In this study, we propose an outer approximation algorithm (Algorithm 1) for CVOPs, which solves a norm-minimizing scalarization in each iteration. Different from Pascoletti-Serafini scalarization, it does not require a direction parameter; hence, one does not need to fix a direction parameter as in [1], or a point in the objective space in order to compute the direction parameter as in [20]. Moreover, when terminates, the algorithm provides the Hausdorff distance between the upper image and its outer approximation, directly.

The scalarization methods based on a norm have been frequently used in the context of MOPs, see for instance [26]. These methods generally depend on the *ideal point* at which all objectives of the MOP attain its optimal value, simultaneously. Since the ideal point is not feasible in general, the idea is to find the minimum distance from the ideal point to the image of the feasible region. One of the well-known methods is the *weighted compromise programming* problem, which utilizes the ℓ_p norm with $p \geq 1$, see for instance [31, 32]. The most commonly used special case is also known as the weighted Chebyshev scalarization, where the underlying norm is taken as the ℓ_∞ norm, see for instance [33, 34, 14]. The weight vector in these scalarization problems is taken such that each component is positive. If the weight vector is taken as the vector of ones, then they are simply called compromise programming ($p \geq 1$) and Chebyshev scalarization ($p = +\infty$), respectively.

Another approximation method for MOPs based on norm was introduced by Schandl et al. in [35]. This method uses the oblique norms, introduced in [36], which is based on polyhedral gauge function. Note that, unlike other norm based scalarization methods, it does not use directly the norm in the scalarization model. Rather, using the properties of oblique norms a linear objective function is obtained [35]. Therefore, it reduces to a classical scalarization method. Klamroth and Tind argue that this is the generalization of weighted sum scalarization and augmented weighted Chebyshev scalarization [17].

The scalarization method that is proposed in this thesis works with any norm defined on the objective space. It simply computes the distance, with respect to a fixed norm, from a given reference point in the objective space to the upper image. This is similar to compromise programming, however it has further advantages compared to it:

The reference point used in the norm-minimizing scalarization is not necessarily the ideal point, which is not well-defined for a VOP. Indeed, within the proposed algorithm, we solve it for the vertices of the outer approximation of the upper image. In weighted compromise programming, finding various nondominated points is done by varying the (nonnegative) weight parameters. It is not straightforward to generalize weighted compromise programming for a vector optimization setting, whereas this can be done directly with the proposed norm-minimizing scalarization.

We discuss some properties of the proposed scalarization under mild assumptions. In particular, we prove that if the feasible region of the VOP is solid and compact, then there exist an optimal solution to it as well as an optimal solution to its Lagrange dual. Moreover, strong duality holds. We further prove that using a dual optimal solution, one can generate a supporting halfspace to the upper image. Note that for these results, the ordering cone is assumed to be a closed convex cone that is solid, pointed and nontrivial. However, different from the similar results regarding Pascoletti-Serafini scalarization, see for instance [1], the ordering cone is not necessarily polyhedral.

The main idea of Algorithm 1 is similar to the Benson-type outer approximation algorithms; iteratively, it finds better outer approximations to the upper image and stops when the approximation is sufficiently fine. As already mentioned, it solves the proposed norm-minimizing scalarization model instead of Pascoletti-Serafini scalarization. Hence, it is free of direction-biasedness. Using the properties of the norm-minimizing scalarization, we prove that the algorithm works correctly, that is, given an approximation error $\epsilon > 0$, when terminates, the algorithm returns an outer approximation to the upper image such that the Hausdorff distance between the two is less than ϵ.

Note that, owing to some limitation due to the structure of CVOP algorithms available in the literature, including Algorithm 1, it becomes difficult to perform convergence analysis. To overcome this, we come up with two different variants of the algorithm by introducing a suitable compact subset of the upper image.

1.3 An Algorithm With Finiteness Result

We propose a modification of Algorithm 1, namely, Algorithm 2. In addition to its correctness, we prove that if the feasible region is compact, then for a given approximation error $\epsilon > 0$, Algorithm 2 stops after finitely many iterations. Note that the finiteness of outer approximation algorithms for linear VOPs are known, see for instance [5]. Also, under compact feasible region assumption, the finiteness of an outer approximation for nonlinear (even for nonconvex) MOPs, proposed in [19], is known. However, to the best of our knowledge, Algorithm 2 is the first CVOP algorithm with a guarantee for finiteness.

Compared to the cases of linear VOPs and nonconvex MOPs, proving the finiteness of Algorithm 2 has the following new challenges which we address by our technical analysis:

1. Since the upper image is polyhedral for a linear VOP, the algorithms find exact solutions, and finiteness follows by the polyhedrality of the upper image. On the other hand, for a CVOP, we look for approximate solutions of a convex and generally nonpolyhedral upper image. Hence, the proof of

finiteness requires completely different arguments.

2. The algorithm for nonconvex MOPs in [19] constructs an outer approximation for the upper image by discarding sets of the form $\{v\} - \operatorname{int} \mathbb{R}^q_+$, where v is a point on the upper image (see Section 2.1 for precise definitions). In this case, the proof of finiteness relies on a hypervolume argument for certain small hypercubes generated by the outer approximation. In the current work, we deal with CVOPs with general ordering cones and our algorithms construct an outer approximation by intersecting certain supporting halfspaces of the upper image (instead of discarding "point minus cone" type sets). To prove the finiteness of Algorithm 2, we propose a novel hypervolume argument which exploits the relationship between these halfspaces and certain subsets of small norm balls (see Lemma 5.1.1). Another important challenge in using supporting halfspaces is to guarantee that the vertices of the outer approximations, which are the reference points for the scalarization models, as well as the minimal points of the upper image found by solving these scalarizations throughout the algorithm are within a compact set. Note that this is naturally the case in [19] by the structure of their outer approximations. For our proposed algorithm, we construct sufficiently large compact sets S and S_2 such that the vertices and the corresponding minimal points of the upper image are within S and S_2, respectively (see Lemmas 4.2.3, 5.1.2 and Remark 4.2.4).

1.4 An Algorithm With Convergence Rate Analysis

We propose a variant of Algorithm 2, namely Algorithm 3, which is based on a modified norm-minimizing scalarization. In particular, the proposed scalarazition is modified using an additional constraint representing a halfspace. The constraint ensures that the algorithm deals with only a compact subset of the upper image throughout the iterations. Subsequently, this helps not only to prove the finiteness of Algorithm 3, but also to estimate the convergence rate of the algorithm. In particular, by the modification of the scalarization, we attain certain conditions (see Section 5.3) which are required in order to refer to the convergence analysis result available in the literature ([13], [37, Chapter 8]). The result originally holds for polyhedral approximations of convex compact bodies, which remain an interesting problem in the literature. For a detailed review on this topic, the reader is referred to, for instance, the reviews by Gruber [38] and by Lotov et al. [37], which mainly discuss the results of [13, 39, 40, 41, 42, 43, 44].

Note that the convergence rate for some MOP algorithms (including nonlinear and nonconvex MOPs) has been determined in the literature, see for instance [17, 45, 46]. To the best of our knowledge, among CVOP algorithms, Algorithm 3 is the first one with an estimate of convergence rate.

Next, in order to find even a better estimate, we analyze the results of Kamenev in [43] in case of our algorithm under the assumption of Euclidean norm. The results in [43] originally hold for a specific class of algorithms (see [37, Chapter 8] for details). Although Algorithm 3 does not belong to that class, we determine a better estimate of convergence rate by proving similar results to [43], for our algorithm.

1.5 Thesis Outline

The rest of the thesis is organized as follows. In Section 2.1, we introduce the notations used in the thesis and recall some well-known concepts and results in convex analysis. In Section 2.2, we present the setting for CVOP, discuss an approximate solution concept from the literature. This is followed by a detailed treatment of norm-minimizing scalarizations in Section 3.1, including some duality results as well as geometric properties of optimal solutions. Similar treatment of modified norm-minimizing scalarizations and its dual is shown in Section 3.2. Sections 4.1, 4.2 and 4.3 are devoted to Algorithms 1, 2 and 3, respectively. We also prove that they works correctly. The analysis of Algorithm 2 and 3 continues in Chapter 5. This includes finiteness result of Algorithm 2 and 3 in Section 5.1

and 5.2, respectively, and an estimate of convergence rate of Algorithm 3 in Section 5.3. In Section 5.4, we provide an improved estimate of the convergence rate of Algorithm 3 using Euclidean norm in its scalrization. We conclude the thesis by providing several examples and discussing the computational performance of the proposed algorithms in Chapter 6 and a future research direction in Chapter 7.

Chapter 2

Problem Definition and Formulation

In this chapter, after introducing some notations and definitions, we describe the convex vector optimization problem (CVOP) and provide solution concepts for it. Then, few assumptions and propositions are given, which are necessary in order to prove the main results of this thesis.

2.1 Preliminaries

In this section, we describe the notations and definitions which will be used throughout the thesis. Let $q \in \mathbb{N} := \{1, 2, \ldots\}$. We denote by $\mathbb{R}^q$ the q-dimensional Euclidean space. When $q = 1$, we have the real line $\mathbb{R} := \mathbb{R}^1$, and the extended real line $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$. On $\mathbb{R}^q$, we fix an arbitrary norm $\|\cdot\|$, and we denote its dual norm by $\|\cdot\|_*$. We will sometimes assume that $\|\cdot\| = \|\cdot\|_p$ is the ℓ^p-norm on $\mathbb{R}^q$, where $p \in [1, +\infty]$. For $y \in \mathbb{R}^q$, the ℓ^p-norm of y is defined by $\|y\|_p := (\sum_{i=1}^q |y_i|^p)^{\frac{1}{p}}$ when $p \in [1, +\infty)$, and by $\|y\|_p := \max_{i \in \{1,\ldots,q\}} |y_i|$ when $p = +\infty$. In this case, the dual norm is $\|\cdot\|_* = \|\cdot\|_{p'}$, where $p' \in [1, +\infty]$ is the conjugate exponent of p via the relation $\frac{1}{p} + \frac{1}{p'} = 1$. For $\epsilon > 0$, we define the closed ball $\mathbb{B}_\epsilon := \{z \in \mathbb{R}^q \mid \|z\| \leq \epsilon\}$ centered at the origin.

Let $f\colon \mathbb{R}^q \to \overline{\mathbb{R}}$ be a convex function and $y_0 \in \mathbb{R}^q$ with $f(y_0) \in \mathbb{R}$. The set $\partial f(y_0) := \{z \in \mathbb{R}^q \mid \forall y \in \mathbb{R}^q\colon f(y) \geq f(y_0) + z^{\mathsf{T}}(y - y_0)\}$ is called the *subdifferential* of f at y_0.

For a set $A \subseteq \mathbb{R}^q$, we denote by $\operatorname{int} A$, $\operatorname{cl} A$, $\operatorname{bd} A$, $\operatorname{conv} A$, $\operatorname{cone} A$, the interior, closure, boundary, convex hull, conic hull of A, respectively. A recession direction of A is a vector $k \in \mathbb{R}^q \setminus \{0\}$ satisfying $A + \{\lambda k \in \mathbb{R}^q \mid \lambda > 0\} \subseteq A$. The set of all recession directions of A, $\operatorname{recc} A = \{k \in \mathbb{R}^q \mid \forall a \in A, \forall \lambda \geq 0:\ a + \lambda k \in A\}$, is the *recession cone* of A. If $A, B \subseteq \mathbb{R}^q$ are nonempty sets and $\lambda \in \mathbb{R}$, then we define the Minkowski operations $A + B := \{y_1 + y_2 \mid y_1 \in A, y_2 \in B\}$, $\lambda A := \{\lambda y \mid y \in A\}$, $A - B := A + (-1)B$.

Let $C \subseteq \mathbb{R}^q$ be a convex cone. The set $C^+ := \{z \in \mathbb{R}^q \mid \forall y \in C : z^{\mathsf{T}} y \geq 0\}$ is a closed convex cone, and it is called the *dual cone* of C. The cone C is said to be *solid* if $\operatorname{int} C \neq \emptyset$, *pointed* if it does not contain any lines, and *nontrivial* if $\{0\} \subsetneq C \subsetneq \mathbb{R}^q$. If C is a solid pointed nontrivial cone, then the relation $\leq_C$ on $\mathbb{R}^q$ defined by $y_1 \leq_C y_2 \iff y_2 - y_1 \in C$ for every $y_1, y_2 \in \mathbb{R}^q$ is a partial order. Let $X \subseteq \mathbb{R}^n$ be a convex set, where $n \in \mathbb{N}$. A function $\Gamma\colon X \to \mathbb{R}^q$ is said to be *C-convex* if $\Gamma(\lambda x_1 + (1-\lambda)x_2) \leq_C \lambda\Gamma(x_1) + (1-\lambda)\Gamma(x_2)$ for every $x_1, x_2 \in X, \lambda \in [0,1]$. In this case, the function $x \mapsto w^{\mathsf{T}}\Gamma(x)$ on X is convex for every $w \in C^+$. Let $\mathcal{X} \subseteq X$. Then, the set $\Gamma(\mathcal{X}) := \{\Gamma(x) \mid x \in \mathcal{X}\}$ is the image of $\mathcal{X}$ under Γ. The function $I_{\mathcal{X}} : \mathbb{R}^q \to [0, +\infty]$ defined by $I_{\mathcal{X}}(x) = 0$ whenever $x \in \mathcal{X}$ and by $I_{\mathcal{X}}(x) = +\infty$ whenever $x \in \mathbb{R}^q \setminus \mathcal{X}$ is called the *indicator function* of $\mathcal{X}$.

Let $A \subseteq \mathbb{R}^q$ be a nonempty set. A point $y \in A$ is called a *C-minimal element* of A if $(\{y\} - C \setminus \{0\}) \cap A = \emptyset$. If the cone C is solid, then y is called a *weakly C-minimal element* of A if $(\{y\} - \operatorname{int} C) \cap A = \emptyset$. We denote by $\operatorname{Min}_C(A)$ the set of all C-minimal elements of A, and by $\operatorname{wMin}_C(A)$ the set of all weakly C-minimal elements of A whenever C is solid.

For each $z \in \mathbb{R}^q$, we define $d(z, A) := \inf_{y \in A}\|z - y\|$. Let $B \subseteq \mathbb{R}^q$ be a nonempty set. We denote by $\delta^H(A,B)$ the Hausdorff distance between A, B. It is well-known that [47, Proposition 3.2]

$$\begin{aligned}\delta^H(A,B) &= \max\left\{\sup_{y\in A} d(y,B), \sup_{z\in B} d(z,A)\right\}\\ &= \inf\left\{\epsilon>0 \mid A\subseteq B+\mathbb{B}_\epsilon,\ B\subseteq A+\mathbb{B}_\epsilon\right\}.\end{aligned} \tag{2.1.1}$$

Suppose that A is a convex set and let $y \in A$, $w \in \mathbb{R}^q \setminus \{0\}$. If $w^\mathsf{T} y = \inf_{z\in A} w^\mathsf{T} z$, then the set $\{z \in \mathbb{R}^q \mid w^\mathsf{T} z = w^\mathsf{T} y\}$ is called a *supporting hyperplane* of A at y and the set $\{z \in \mathbb{R}^q \mid w^\mathsf{T} z \geq w^\mathsf{T} y\} \supseteq A$ is called a *supporting halfspace* of A at y.

Suppose that A is a polyhedral closed convex set. The representation of A as the intersection of finitely many halfspaces, that is, as $A = \bigcap_{i=1}^r \{y \in \mathbb{R}^q \mid (w^i)^\mathsf{T} y \geq a^i\}$ for some $r \in \mathbb{N}$, $w^i \in \mathbb{R}^q \setminus \{0\}$ and $a^i \in \mathbb{R}$, $i \in \{1,\dots,r\}$, is called an *H-representation* of A. Alternatively, A is uniquely determined by a finite set $\{y^1,\dots,y^s\} \subseteq \mathbb{R}^q$ of vertices and a finite set $\{d^1,\dots,d^t\} \subseteq \mathbb{R}^q$ of directions via $A = \operatorname{conv}\{y^1,\dots,y^s\} + \operatorname{conv}\operatorname{cone}\{d^1,\dots,d^t\}$, which is called a *$V$-representation* of A. A point $v \in A$ is an *extreme point* of A if there does not exist $a_1, a_2 \in A$ and $0 < \lambda < 1$ such that $a_1 \neq a_2$ and $v = \lambda a_1 + (1-\lambda)a_2$. We denote the set of all extreme points of A by $\mathcal{V}_A$. If A is polyhedral, then $\mathcal{V}_A$ coincides with the set of vertices of A. The following lemma is a simple observation that will be useful throughout.

Lemma 2.1.1. *If $A \subseteq \mathbb{R}^q$ is a convex compact set and $C \subseteq \mathbb{R}^q$ is a convex cone, then $\mathcal{V}_{A+C} \subseteq \mathcal{V}_A$.*

Proof. The inclusion is trivial if $\mathcal{V}_{A+C} = \emptyset$. Otherwise, let $v \in \mathcal{V}_{A+C} \setminus \mathcal{V}_A$ for a contradiction. Since $v \in A + C$, $v = a + c$ for some $a \in A$ and $c \in C$. Note that if $c = 0$, then $v = a \notin \mathcal{V}_A$ implies that v can be written as a non-trivial convex combination of points from $A \subseteq A + C$. As this contradicts $v \in \mathcal{V}_{A+C}$, we conclude that $c \neq 0$. Next, as A is compact, we have $\operatorname{recc}(A + C) = C$. Hence, for every $\lambda \geq 0$, $a + \lambda c \in A + C$. 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 $v \in \mathcal{V}_{A+C}$. $\square$

2.2 Convex Vector Optimization

We consider a convex vector optimization problem (CVOP) of the form

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

where $C \subseteq \mathbb{R}^q$ is the *ordering cone* of the problem, $\Gamma\colon X \to \mathbb{R}^q$ is the vector-valued *objective function* defined on a convex set $X \subseteq \mathbb{R}^n$, and $\mathcal{X} \subseteq X$ is the *feasible region*. The conditions we impose on $C, \Gamma, \mathcal{X}$ are stated in the next assumption.

Assumption 2.2.1. *The following statements hold.*

(a) C is a closed convex cone that is also solid, pointed, and nontrivial.

(b) Γ is a C-convex and continuous function.

(c) $\mathcal{X}$ is a compact convex set with $\operatorname{int} \mathcal{X} \neq \emptyset$.

The set

$$\mathcal{P} := \operatorname{cl}(\Gamma(\mathcal{X}) + C) \tag{2.2.1}$$

is called the *upper image* of (P). Clearly, $\mathcal{P}$ is a closed convex set with $\mathcal{P} = \mathcal{P} + C$.

Remark 2.2.2. *Note that, under Assumption 2.2.1, $\Gamma(\mathcal{X})$ is a compact set as the image of a compact set under a continuous function. Then, $\Gamma(\mathcal{X}) + C$ is a closed set as the algebraic sum of a compact set and a closed set [48, Lemma 5.2]. Hence, we may drop the closure in (2.2.1), that is, we have $\mathcal{P} = \Gamma(\mathcal{X}) + C$. Figure 2.1 shows an example of upper image, for some $\Gamma(\mathcal{X}) \subseteq \mathbb{R}^2$ and an ordering cone C.*

We recall the notion of boundedness for CVOP next.

Definition 2.2.3. *[1, Definition 3.1] (P) is said to be bounded if $\mathcal{P} \subseteq \{y\} + C$ for some $y \in \mathbb{R}^q$.*

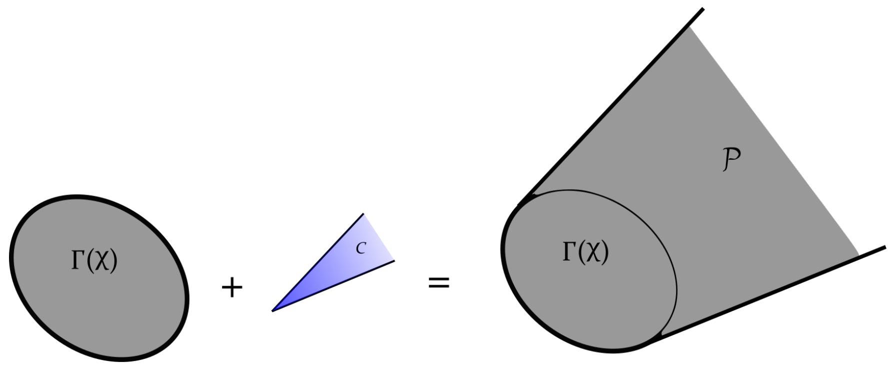


Figure 2.1: Upper image $\mathcal{P} = \Gamma(\mathcal{X}) + C$, for some $\Gamma(\mathcal{X}) \subseteq \mathbb{R}^2$ and an ordering cone C.

In view of Remark 2.2.2, it follows that (P) is bounded under Assumption 2.2.1. The next definition recalls the relevant solution concepts for CVOP.

Definition 2.2.4. *[49, Definition 7.1] A point $\bar{x} \in \mathcal{X}$ is said to be a* (weak) minimizer *for* (P) *if $\Gamma(\bar{x})$ is a (weakly) C-minimal element of $\Gamma(\mathcal{X})$. A nonempty set $\bar{\mathcal{X}} \subseteq \mathcal{X}$ is called an* infimizer *of* (P) *if $\operatorname{cl}\operatorname{conv}(\Gamma(\bar{\mathcal{X}})+C) = \mathcal{P}$. An infimizer $\bar{\mathcal{X}}$ of* (P) *is called a* (weak) solution *of* (P) *if it consists of only (weak) minimizers.*

In CVOP, it may be difficult or impossible to compute a solution in the sense of Definition 2.2.4, in general. Hence, we consider the following notion of approximate solution.

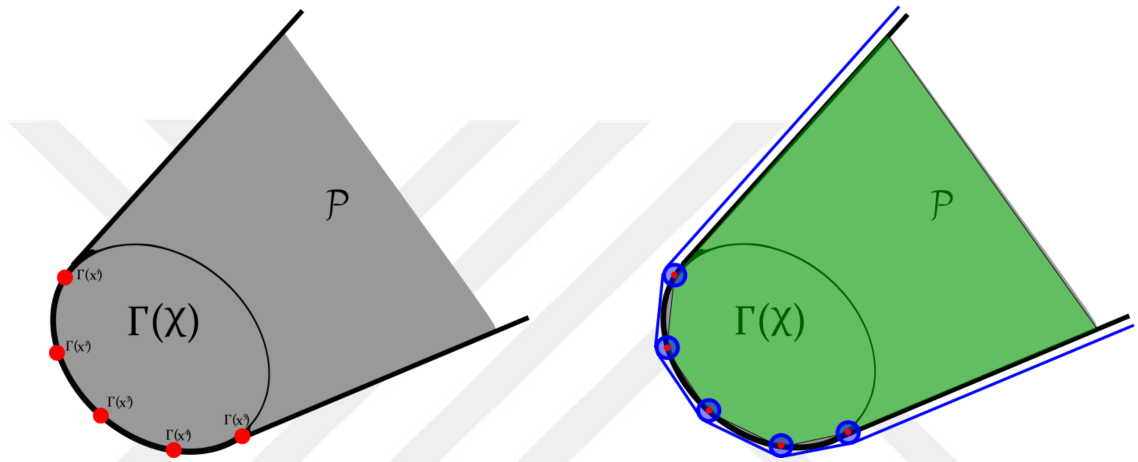
Definition 2.2.5. *[25, Definition 3.3] Suppose that* (P) *is bounded and let $\epsilon > 0$. Let $\bar{\mathcal{X}} \subseteq \mathcal{X}$ be a nonempty finite set and define $\bar{\mathcal{P}} := \operatorname{conv}\Gamma(\bar{\mathcal{X}}) + C$. The set $\bar{\mathcal{X}}$ is called a* finite ϵ-infimizer *of* (P) *if $\bar{\mathcal{P}}_\epsilon = \bar{\mathcal{P}} + B(0,\epsilon) \supseteq \mathcal{P}$. The set $\bar{\mathcal{X}}$ is called a* finite (weak) ϵ-solution *of* (P) *if it is an ϵ-infimizer that consists of only (weak) minimizers.*

For a finite (weak) ϵ-solution $\bar{\mathcal{X}}$, it is immediate from Definition 2.2.5 that

$$\bar{\mathcal{P}}_\epsilon \supseteq \mathcal{P} \supseteq \bar{\mathcal{P}}. \tag{2.2.2}$$

Hence, $\bar{\mathcal{X}}$ provides an inner and an outer approximation for the upper image $\mathcal{P}$. Figure 2.2 gives an intuition of finite (weak) ϵ-solution of (P) using $\mathcal{P} \subseteq \mathbb{R}^2$. Let $\bar{\mathcal{X}}$

consists of five (weak) minimizers, that is, $\bar{\mathcal{X}} = \{x^1, \ldots, x^5\}$, a corresponding set of (weakly) C-minimal elements be $\Gamma(\bar{\mathcal{X}}) = \{\Gamma(x^1), \ldots, \Gamma(x^5)\}$ and $\epsilon > 0$. Then, the inclusion required to satisfy Definition 2.2.5, $\operatorname{conv}\Gamma(\bar{\mathcal{X}}) + C + B(0,\epsilon) \supseteq \mathcal{P}$, is shown in Figure 2.2b.



(a) Upper image $\mathcal{P}$ along with a set of (weakly) C-minimal elements $\Gamma(\bar{\mathcal{X}}) = \{\Gamma(x^1), \ldots, \Gamma(x^5)\}$.

(b) $\operatorname{conv}\Gamma(\bar{\mathcal{X}}) + C + B(0,\epsilon) \supseteq \mathcal{P}$.

Figure 2.2: Intuition of finite (weak) ϵ-solution, $\bar{\mathcal{X}} = \{x^1, \ldots, x^5\}$, of (P) given by Definition 2.2.5.

Remark 2.2.6. *In [25, Definition 3.3] the statement of the definition is slightly different. Instead of $\bar{\mathcal{P}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}$, the requirement is given as $\delta^H(\mathcal{P}, \bar{\mathcal{P}}) \leq \epsilon$. However, both yield equivalent definitions. Indeed, by (2.2.2) we have $\mathcal{P} \subseteq \bar{\mathcal{P}} + \mathbb{B}_\epsilon$ as well as $\bar{\mathcal{P}} \subseteq \mathcal{P} \subseteq \mathcal{P} + \mathbb{B}_\epsilon$. Then, $\delta^H(\mathcal{P}, \bar{\mathcal{P}}) \leq \epsilon$ follows by (2.1.1). The converse holds similarly by (2.1.1).*

Given $w \in C^+ \setminus \{0\}$, the following convex program is the well-known *weighted sum scalarization* of (P):

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

The following proposition is a standard result in vector optimization, it formulates the connection between weighted sum scalarizations and weak minimizers.

Proposition 2.2.7. *[50, Corollary 2.3] Let $w \in C^+ \setminus \{0\}$. Then, every optimal solution of* (WS(w)) *is a weak minimizer of* (P).

For the new notion of approximate solution in Definition 2.2.5, we prove an existence result next.

Proposition 2.2.8. *Suppose that Assumption 2.2.1 holds. Then, there exists a solution of* (P)*. Moreover, for every $\epsilon > 0$, there exists a finite ϵ-solution of* (P).

Proof. The existence of a solution $\bar{\mathcal{X}}$ of (P) follows by [1, Proposition 4.2]. To prove the second statement, we fix $\epsilon > 0$. For each $x \in \bar{\mathcal{X}}$, note that the set $\{\Gamma(x)\} + \operatorname{int} C + B(0, \epsilon)$ is open as the algebraic sum of an open set and an arbitrary set [48, Lemma 5.2]. Then,

$$\left\{\{\Gamma(x)\} + \operatorname{int} C + B(0,\epsilon) \mid x \in \bar{\mathcal{X}}\right\}$$

is an open cover of $\Gamma(\mathcal{X})$. By Remark 2.2.2, the set $\Gamma(\mathcal{X})$ is compact. Therefore, there exists a finite sub-cover of $\Gamma(\mathcal{X})$, that is, there exists a finite set $\hat{\mathcal{X}} \subseteq \bar{\mathcal{X}}$ such that

$$\Gamma(\hat{\mathcal{X}}) + \operatorname{int} C + B(0,\epsilon) = \bigcup_{x \in \hat{\mathcal{X}}} \left(\{\Gamma(x)\} + \operatorname{int} C + B(0,\epsilon)\right) \supseteq \Gamma(\mathcal{X}). \tag{2.2.3}$$

Adding the cone C to both sides of (2.2.3) gives

$$\Gamma(\hat{\mathcal{X}}) + \operatorname{int} C + B(0,\epsilon) = \Gamma(\hat{\mathcal{X}}) + \operatorname{int} C + B(0,\epsilon) + C \supseteq \Gamma(\mathcal{X}) + C$$

since $\operatorname{int} C + C = \operatorname{int} C$. Since $\operatorname{conv} \Gamma(\hat{\mathcal{X}}) \supseteq \Gamma(\hat{\mathcal{X}})$ and $C \supseteq \operatorname{int} C$, it follows that

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

where the last equality follows from Remark 2.2.2. Finally, since $\bar{\mathcal{X}}$ is a solution of (P), we have $\hat{\mathcal{X}} \subseteq \bar{\mathcal{X}} \subseteq \operatorname{Min}_C(\Gamma(\mathcal{X}))$. Hence, by Definition 2.2.5, $\hat{\mathcal{X}}$ is a finite ϵ-solution of (P). □

Chapter 3

Norm-Minimizing Scalarizations

In this chapter, we introduce norm-minimizing scalarization and provide its detailed analysis based on the results obtained in Chapter 2. This is followed by the modified scalarization along with the similar results. The scalarizations would be used in the proposed algorithms in Chapter 4.

3.1 Norm-Minimizing Scalarization

In the algorithm that will be described in the next chapter, we employ a norm-minimizing scalarization to be solved in each iteration. In this section, we describe the scalarization model and provide some results to be used in the design of the algorithm.

Let us fix an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^q$ and a point $v \in \mathbb{R}^q$. We consider the *norm-minimizing scalarization* of (P) given by

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

Note that this is a convex program as both the objective function and the feasible region are convex.

Remark 3.1.1. *The optimal value of* (P(v)) *is equal to* $d(v, \mathcal{P})$*, the distance of* v *to the upper image* $\mathcal{P}$*. Indeed, by Remark 2.2.2, we have*

$$\begin{aligned} d(v, \mathcal{P}) &= \inf_{y \in \mathcal{P}} \|v - y\| = \inf\{\|z\| \mid v + z \in \mathcal{P},\ z \in \mathbb{R}^q\} \\ &= \inf\{\|z\| \mid v + z \in \{\Gamma(x)\} + C,\ x \in \mathcal{X},\ z \in \mathbb{R}^q\} \\ &= \inf\{\|z\| \mid \Gamma(x) - v - z \leq_C 0,\ x \in \mathcal{X},\ z \in \mathbb{R}^q\}. \end{aligned}$$

In order to derive the Lagrangian dual of (P(v)), we first pass to an equivalent formulation of (P(v)). To that end, let us define a scalar function $f\colon X \times \mathbb{R}^q \to \overline{\mathbb{R}}$ and a set-valued function $G : X \times \mathbb{R}^q \rightrightarrows \mathbb{R}^q$ by

$$f(x, z) := \|z\| + I_{\mathcal{X}}(x), \quad G(x, z) := \{\Gamma(x) - z - v\}, \quad x \in X, z \in \mathbb{R}^q. \tag{3.1.1}$$

Clearly, in our case of vector optimization problem, $G(x, z)$ is a singleton. Let's write the constraint of (P(v)), $z + v - \Gamma(x) \in C$, which is equivalent to $G(x, z) \cap -C \neq \emptyset$ since $G(x, z)$ is a singleton. Now, (P(v)) can be rewritten as

$$\inf_{(x,z) \in S} f(x, z), \quad \text{where } S := \big\{(x, z) \in (X, \mathbb{R}^q) \mid G(x, z) \cap -C \neq \emptyset\big\}. \tag{P$'(v)$}$$

To use the results from [5, Section 3.3.1] and [51, Section 8.3.2] for convex programming with set-valued constraints, we define the Lagrangian $L\colon X \times \mathbb{R}^q \times \mathbb{R}^q \to \overline{\mathbb{R}}$ for (P$'(v)$) by

$$L(x, z, w) := f(x, z) + \inf_{u \in G(x,z)+C} w^{\mathsf{T}} u, \quad (x, z, w) \in X \times \mathbb{R}^q \times \mathbb{R}^q. \tag{3.1.2}$$

Then, the dual objective function $\phi : \mathbb{R}^q \to \overline{\mathbb{R}}$ is

$$\phi(w) := \inf_{x \in X, z \in \mathbb{R}^q} L(x, z, w), \quad w \in \mathbb{R}^q.$$

and the dual problem of (P$'(v)$) is formulated as

$$\sup_{w \in \mathbb{R}^q} \phi(w) = \sup_{w \in \mathbb{R}^q} \left\{ \inf_{x \in X, z \in \mathbb{R}^q} L_v(x, z, w) \right\}. \tag{D(v)}$$

Note that

$$\begin{aligned}\inf_{u\in G(x,z)+C} w^T u &= \inf_{u\in\{\Gamma(x)-z-v\}+C} w^T u\\ &= \inf_{c\in C} w^T(\Gamma(x)-z-v+c)\\ &= w^T(\Gamma(x)-z-v)+\inf_{c\in C} w^T c\end{aligned}$$

where

$$\inf_{c\in C} w^T c = \begin{cases} 0 & \text{if } w\in C^+,\\ -\infty & \text{otherwise.}\end{cases}$$

Hence,

$$\inf_{x\in X, z\in\mathbb{R}^q} L_v(x,z,w) = \begin{cases} \inf_{x\in X, z\in\mathbb{R}^q}\left\{f(x,z)+w^T(\Gamma(x)-z-v)\right\} & \text{if } w\in C^+,\\ -\infty & \text{otherwise.}\end{cases}$$

Also, note that for $w\in C^+$

$$\begin{aligned}\inf_{x\in X, z\in\mathbb{R}^q} L_v(x,z,w) &= \inf_{x\in X, z\in\mathbb{R}^q}\left\{f(x,z)+w^T(\Gamma(x)-z-v)\right\}\\ &= \inf_{x\in X, z\in\mathbb{R}^q}\left\{\|z\|+I_{\mathcal{X}}(x)+w^T(\Gamma(x)-z-v)\right\}\\ &= \inf_{x\in \mathcal{X}, z\in\mathbb{R}^q}\left\{\|z\|+w^T(\Gamma(x)-z-v)\right\}\end{aligned}$$

Now, from (D(v)), we get

$$\begin{aligned}\sup_{w\in C^+}\phi(w) &= \sup_{w\in C^+}\left\{\inf_{x\in \mathcal{X}, z\in\mathbb{R}^q}\left\{\|z\|+w^T(\Gamma(x)-z-v)\right\}\right\}.\\ &= \sup_{w\in C^+}\left\{\inf_{x\in\mathcal{X}} w^T\Gamma(x)+\inf_{z\in\mathbb{R}^q}\{\|z\|-w^Tz\}-w^Tv\right\}.\\ &= \sup_{w\in C^+}\left\{\inf_{x\in X} w^T\Gamma(x)-\sup_{z\in\mathbb{R}^q}\{w^Tz-\|z\|\}-w^Tv\right\},\end{aligned}$$

where second term, on right hand side, is the conjugate of a norm. From [52, Example 3.26], we have

$$\sup_{z\in\mathbb{R}^q}\{w^Tz-\|z\|\} = \begin{cases} 0 & \text{if } \|w\|_*\le 1,\\ +\infty & \text{otherwise,}\end{cases}$$

which shows that the conjugate function of $\|\cdot\|$ is the indicator function of the unit ball of the dual norm $\|\cdot\|_*$. Therefore, the optimal value of (D(v)) is given by

$$\sup\left\{\inf_{x\in\mathcal{X}} w^T\Gamma(x) - w^T v \;\middle|\; \|w\|_* \leq 1,\;\; w\in C^+\right\}. \tag{3.1.3}$$

The next proposition shows the strong duality between (P(v)) and (D(v)).

Proposition 3.1.2. *Suppose that Assumption 2.2.1 holds. For every $v\in\mathbb{R}^q$, there exist optimal solutions (x^v, z^v) and w^v to problems (P(v)) and (D(v)), respectively, and the optimal values coincide.*

Proof. Let us fix some $\tilde{x}\in\mathcal{X}$ and define $\tilde{z} := \Gamma(\tilde{x}) - v$. Clearly, $(\tilde{x},\tilde{z})$ is feasible for (P(v)). We consider the following problem with compact feasible region in $\mathbb{R}^{n+q}$:

$$\text{minimize } \|z\| \text{ subject to } \Gamma(x) - z - v \leq_C 0,\;\; \|z\|\leq\|\tilde{z}\|,\;\; x\in\mathcal{X},\;\; z\in\mathbb{R}^q. \tag{3.1.4}$$

An optimal solution (x^*, z^*) for the problem in (3.1.4) exists by Weierstrass Theorem and (x^*, z^*) is also optimal for (P(v)). To show the existence of an optimal solution of (D(v)), we show that the following constraint qualification in [51, 5] holds for (P(v)):

$$G(\operatorname{dom} f)\cap -\operatorname{int} C \neq \emptyset, \tag{3.1.5}$$

where $\operatorname{dom} f := \{(x,z)\in X\times\mathbb{R}^q \mid f(x,z) < +\infty\}$. Since $\operatorname{int}\mathcal{X}\neq\emptyset$ and $\operatorname{int} C\neq\emptyset$ by Assumption 2.2.1, we may fix $x^0\in\operatorname{int}\mathcal{X}$, $y^0\in\Gamma(x^0)+\operatorname{int} C$ and define $z^0 := y^0 - v$. We have $v + z^0 - \Gamma(x^0)\in\operatorname{int} C$, equivalently, $G(x^0, z^0)\subseteq -\operatorname{int} C$. As $(x^0, z^0)\in\operatorname{dom} f = \mathcal{X}\times\mathbb{R}^q$, it follows that (3.1.5) holds. Moreover, the set-valued map $G : X\times\mathbb{R}^q \rightrightarrows \mathbb{R}^q$ is C-convex [51, Section 8.3.2], that is,

$$\lambda G(x^1, z) + (1-\lambda)G(x^2, z) \subseteq G(\lambda(x^1, z) + (1-\lambda)(x^2, z)) + C \tag{3.1.6}$$

for every $x^1, x^2\in X$, $z\in\mathbb{R}^q$, $\lambda\in[0,1]$. Indeed, by the C-convexity of $\Gamma\colon X\to\mathbb{R}^q$, we have

$$\lambda(\Gamma(x^1)-z-v)+(1-\lambda)(\Gamma(x^2)-z-v)\in\{\Gamma(\lambda x^1+(1-\lambda)x^2)-z-v\}+C$$

for every $x^1,x^2\in X$, $z\in\mathbb{R}^q$, and $\lambda\in[0,1]$, from which (3.1.6) follows. Finally, since $f\colon X\times\mathbb{R}^q\to\overline{\mathbb{R}}$ is also convex, by [5, Theorem 3.19], we have strong duality and dual attainment. $\square$

Notation 3.1.3. *From now on, we fix an arbitrary optimal solution (x^v,z^v) of (P(v)) and an arbitrary optimal solution w^v of (D(v)). Their existence is guaranteed by Proposition 3.1.2.*

Remark 3.1.4. *Note that (x^v,z^v,w^v) is a saddle point of the Lagrangian for (P(v)) given by (3.1.2); see [52, Section 5.4.2]. Hence, we have*

$$\sup_{w\in\mathbb{R}^q}L(x^v,z^v,w)\leq L(x^v,z^v,w^v)\leq\inf_{x\in\mathcal{X},z\in\mathbb{R}^q}L(x,z,w^v).$$

From strong duality, we have

$$\sup_{w\in\mathbb{R}^q}L(x^v,z^v,w)=L(x^v,z^v,w^v)=\inf_{x\in\mathcal{X},z\in\mathbb{R}^q}L(x,z,w^v),$$

where second equality gives

$$(w^v)^T\Gamma(x^v)=\inf_{x\in\mathcal{X}}(w^v)^T\Gamma(x).$$

Hence, x^v is an optimal solution of (WS(w)) for $w=w^v$.

In the next lemma, we characterize the cases where $z^v=0$.

Lemma 3.1.5. *Suppose that Assumption 2.2.1 holds. The following statements hold.*

(a) If $v\notin\mathcal{P}$, then $z^v\neq 0$ and $w^v\neq 0$.

(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 $z^v = 0$.

Proof. To prove (a), suppose that $v \notin \mathcal{P}$. To get a contradiction, we assume that $z^v = 0$. Since (x^v, z^v) is feasible for (P(v)), we have $v = v + z^v \in \{\Gamma(x^v)\} + C \subseteq \mathcal{P}$, contradicting the supposition. Hence, $z^v \neq 0$. Moreover, if we had $w^v = 0$, then the optimal value of (D(v)) would be zero and strong duality would imply that $\|z^v\| = 0$, that is, $z^v = 0$. Therefore, we must have $w^v \neq 0$.

To prove (b) and (c), suppose that $v \in \mathcal{P}$. By Remark 2.2.2, there exists $x \in \mathcal{X}$ such that $\Gamma(x) \leq_C v$. Then, $(x, 0)$ is feasible for (P(v)). Hence, the optimal value of (P(v)) is zero so that $\|z^v\| = 0$, that is, $z^v = 0$. Suppose that we further have $v \in \operatorname{int} \mathcal{P}$. Let $\delta \in \operatorname{int} C$ be such that $v - \delta \in \mathcal{P}$. By Remark 2.2.2, there exists $x^\delta \in \mathcal{X}$ such that $\Gamma(x^\delta) \leq_C v - \delta$, which implies $(w^v)^\mathsf{T}\Gamma(x^\delta) \leq (w^v)^\mathsf{T} v - (w^v)^\mathsf{T}\delta$. Moreover, by strong duality, $\inf_{x\in\mathcal{X}} (w^v)^\mathsf{T}(\Gamma(x) - v) = 0$ holds. Combining these gives

$$0 = \inf_{x\in\mathcal{X}} (w^v)^\mathsf{T}(\Gamma(x) - v) \leq (w^v)^\mathsf{T}(\Gamma(x^\delta) - v) \leq -(w^v)^\mathsf{T}\delta \leq 0$$

so that $(w^v)^\mathsf{T}\delta = 0$. As $\delta \in \operatorname{int} C$ and $w^v \in C^+$, we must have $w^v = 0$. □

The next proposition shows that solving (P(v)) when $v \notin \operatorname{int} \mathcal{P}$ yields a weak minimizer for problem (P).

Proposition 3.1.6. *Suppose that Assumption 2.2.1 holds. If $v \notin \operatorname{int} \mathcal{P}$, then x^v is a weak minimizer of (P), and $y^v := v + z^v \in \operatorname{wMin}_C(\mathcal{P})$.*

Proof. As $\mathcal{X}$ is nonempty and compact, we have $\mathcal{P} \neq \emptyset$ and $\mathcal{P} \neq \mathbb{R}^q$. By [5, Definition 1.45 and Corollary 1.48 (iv)], we have $\operatorname{wMin}_C(\mathcal{P}) = \operatorname{bd} \mathcal{P}$. First, suppose that $v \in \operatorname{bd} \mathcal{P}$. Then, $z^v = 0$ by Lemma 3.1.5. Together with primal feasibility, this implies $\Gamma(x^v) \leq_C v$. As $v \in \operatorname{wMin}_C(\mathcal{P})$, by the definition of weakly C-minimal element, we have $\Gamma(x^v) \in \operatorname{wMin}_C(\mathcal{P})$. Hence, x^v is a weak minimizer of (P) in this case. Next, suppose that $v \notin \mathcal{P}$. Then, $w^v \neq 0$ by Lemma 3.1.5. By Remark 3.1.4, x^v is an optimal solution of (WS(w)) for $w = w^v \in C^+ \setminus \{0\}$. Hence, by Proposition 2.2.7, x^v is a weak minimizer of (P).

Since (x^v, z^v) is feasible for (P(v)), $y^v \in \mathcal{P}$ holds. To get a contradiction, assume that $y^v \notin \mathrm{wMin}_C(\mathcal{P})$; hence, $y^v = v + z^v \in \operatorname{int}\mathcal{P}$. Then, there exists $\epsilon > 0$ 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 $\bar{z} := (\|z^v\| - \epsilon)\frac{z^v}{\|z^v\|}$. Then, $(\bar{x}, \bar{z})$ is feasible for (P(v)). This is a contradiction as $\|\bar{z}\| < \|z^v\|$. □

The following result shows that a supporting hyperplane of $\mathcal{P}$ at $y^v = v + z^v$ can be found by using a dual optimal solution w^v.

Proposition 3.1.7. *Suppose that Assumption 2.2.1 holds and $w^v \neq 0$. Then, the halfspace*

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

contains the upper image $\mathcal{P}$. Moreover, $\operatorname{bd}\mathcal{H}$ is a supporting hyperplane of $\mathcal{P}$ both at $\Gamma(x^v)$ and $y^v = v + z^v$. In particular, $(w^v)^\mathsf{T}\Gamma(x^v) = (w^v)^\mathsf{T} y^v$.

Proof. We clearly have $\Gamma(x^v) \in \operatorname{bd}\mathcal{P} \cap \mathcal{H}$ and $y^v \in \mathcal{P} \cap \mathcal{H}$. Let $y \in \mathcal{P}$ be arbitrary and $\bar{x} \in \mathcal{X}$ be such that $\Gamma(\bar{x}) \leq_C y$. Consider the problems (P(y)) and (D(y)). Clearly, $(\bar{x}, 0)$ is feasible for (P(y)). Moreover, the optimal solution w^v of (D(v)) is feasible for (D(y)). Using weak duality for (P(y)) and (D(y)), we obtain $0 \geq \inf_{x\in\mathcal{X}}(w^v)^\mathsf{T}\Gamma(x) - (w^v)^\mathsf{T} y$. Moreover, from strong duality for (P(v)) and (D(v)), we have $\|z^v\| = \inf_{x\in\mathcal{X}}(w^v)^\mathsf{T}\Gamma(x) - (w^v)^\mathsf{T} v$. Hence,

$$(w^v)^\mathsf{T} y \geq \inf_{x\in\mathcal{X}}(w^v)^\mathsf{T}\Gamma(x) = \|z^v\| + (w^v)^\mathsf{T} v.$$

Note that $\|z^v\| \geq (w^v)^\mathsf{T} z^v$ holds as $\|w^v\|_* \leq 1$ by dual feasibility. Then, we obtain $(w^v)^\mathsf{T} y \geq (w^v)^\mathsf{T} y^v$. In particular, we have $(w^v)^\mathsf{T}\Gamma(x^v) \geq (w^v)^\mathsf{T} y^v$ as $\Gamma(x^v) \in \mathcal{P}$. On the other hand, since $\Gamma(x^v) \leq_C y^v$ and $w^v \in C^+$, we also have $(w^v)^\mathsf{T}\Gamma(x^v) \leq (w^v)^\mathsf{T} y^v$. The equality $(w^v)^\mathsf{T} y^v = (w^v)^\mathsf{T}\Gamma(x^v)$ completes the proof as it implies $y \in \mathcal{H}$ (hence $\mathcal{P} \subseteq \mathcal{H}$) as well as $y^v \in \operatorname{bd}\mathcal{H}$. □

Proposition 3.1.7 provides a method to generate a supporting halfspace of $\mathcal{P}$ at $\Gamma(x^v)$ in which one uses an arbitrary dual optimal solution w^v. The next result shows that if the norm in (P(v)) is taken as the ℓ_p-norm for some $p \in [1, +\infty)$, for instance, the Euclidean norm, then it is possible to generate a supporting halfspace to $\mathcal{P}$ at $\Gamma(x^v)$ using z^v instead of w^v.

Corollary 3.1.8. *Suppose that Assumption 2.2.1 holds and $\|\cdot\| = \|\cdot\|_p$ for some $p \in [1, +\infty)$. Assume that $v \notin \mathcal{P}$. Then, the halfspace*

$$\mathcal{H} = \Big\{ y \in \mathbb{R}^q \mid \sum_{i=1}^{q} \operatorname{sgn}(z_i^v) |z_i^v|^{p-1} y_i \geq \sum_{i=1}^{q} \operatorname{sgn}(z_i^v) |z_i^v|^{p-1} \Gamma_i(x^v) \Big\}$$

contains the upper image $\mathcal{P}$, where sgn *is the usual sign function. Moreover,* bd $\mathcal{H}$ *is a supporting hyperplane of $\mathcal{P}$ both at $\Gamma(x^v)$ and $y^v = v + z^v$.*

Proof. Consider (P(v)) and its Lagrange dual (D(v)). Let us define $g(z) := \|z\|_p - (w^v)^\mathsf{T} z$, $z \in \mathbb{R}^q$. The arbitrarily fixed dual optimal solution w^v satisfies the first order condition with respect to z, that is,

$$0 \in \partial g(z^v).$$

By the chain rule for subdifferentials, this is equivalent to

$$w^v \in \Big(\sum_{i=1}^{q} |z_i^v|^p \Big)^{\frac{1-p}{p}} \big(|z_1^v|^{p-1} S_1 \times \cdots \times |z_q^v|^{p-1} S_q \big), \tag{3.1.7}$$

where, for each $i \in \{1, \ldots, q\}$, S_i denotes the subdifferential of the absolute value function at z_i^v. Let $i \in \{1, \ldots, q\}$. Note that if $z_i^v \neq 0$, then we have $S_i = \{\operatorname{sgn}(z_i^v)\}$. On the other hand, if $z_i^v = 0$, then for each $s_i \in S_i$, we have $|z_i^v|^{p-1} s_i = 0$. Hence, by (3.1.7),

$$w_i^v = \Big(\sum_{i=1}^{q} |z_i^v|^p \Big)^{\frac{1-p}{p}} |z_i^v|^{p-1} \operatorname{sgn}(z_i^v).$$

The assertion follows from Proposition 3.1.7. □

3.2 A modified norm-minimizing scalarization

In this section, we propose a modification to the norm-minimizing scalarization model, which is introduced in Section 3.1, and we provide some analytical results for the modified scalarization.

Let $\bar{w} \in \operatorname{int} C^+$ and $\tilde{\alpha} \in \mathbb{R}$ be fixed. The construction of $\bar{w}$ and $\tilde{\alpha}$ will be discussed, in detail, in section 4.2. For a fix arbitrary norm $\|\cdot\|$ on $\mathbb{R}^q$, we consider the *modified norm-minimizing scalarization* of (P) given by

$$\text{minimize } \|z\| \text{ subject to } \Gamma(x) - z - v \leq_C 0,\ \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha} \leq_{\mathbb{R}_+} 0,\ x \in \mathcal{X},\ z \in \mathbb{R}^q. \tag{$\mathrm{P}_m(v)$}$$

Note that this is a convex program as both the objective function and the feasible region are convex.

Remark 3.2.1. *The optimal value of* ($\mathrm{P}_m(v)$) *is equal to* $d(v, \mathcal{P} \cap S)$, *the distance of* v *to the set* $\mathcal{P} \cap S$ *which is defined by the intersection of upper image* $\mathcal{P}$ *with the halfspace* S.

Now, to attain the Lagrangian dual of ($\mathrm{P}_m(v)$) we write an equivalent formulation of it. First, note that the two constraints $\Gamma(x) \leq_C v + z$ and $-\tilde{\alpha} \leq_{\mathbb{R}_+} -\bar{w}^\mathsf{T}(v+z)$ can be combined as,

$$\begin{bmatrix} \Gamma(x) - z - v \\ \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}. \end{bmatrix} \leq_{C \times \mathbb{R}_+} 0$$

Then, ($\mathrm{P}_m(v)$) can be expressed as

$$\text{minimize } \|z\| \text{ subject to } (\Gamma(x) - z - v, \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}) \leq_{C \times \mathbb{R}_+} 0,\ x \in \mathcal{X},\ z \in \mathbb{R}^q.$$

Now, define a scalar function $f \colon X \times \mathbb{R}^q \to \overline{\mathbb{R}}$ and a set-valued function $G : X \times \mathbb{R}^q \rightrightarrows \mathbb{R}^{q+1}$ as follows

$$f(x,z) := \|z\| + I_{\mathcal{X}}(x), \quad G(x,z) := \{(\Gamma(x) - z - v, \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha})\}, \tag{3.2.1}$$

for $x \in X, z \in \mathbb{R}^q$. Then, ($\mathrm{P}_m(v)$) can be equivalently written as the following:

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

To use the results from [5, Section 3.3.1] and [51, Section 8.3.2] for convex programming with set-valued constraints, we define the Lagrangian $L \colon X \times \mathbb{R}^q \times \mathbb{R}^{q+1} \to \overline{\mathbb{R}}$ for ($\mathrm{P}'_m(v)$) by

$$L(x,z,w,\lambda) := f(x,z) + \inf_{u \in G(x,z)+(C \times \mathbb{R}_+)} [w^\mathsf{T} \ \lambda] u, \quad (x,z,w,\lambda) \in X \times \mathbb{R}^q \times \mathbb{R}^{q+1}. \tag{3.2.2}$$

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

$$\phi(w,\lambda) := \inf_{x \in X, z \in \mathbb{R}^q} L(x,z,w,\lambda), \quad (w,\lambda) \in \mathbb{R}^{q+1}.$$

Note that

$$\begin{aligned}
\inf_{u \in G(x,z)+(C \times \mathbb{R}_+)} [w^\mathsf{T} \ \lambda] u &= \inf_{u \in \{(\Gamma(x)-z-v, \bar{w}^\mathsf{T}(v+z)-\tilde{\alpha})\}+(C \times \mathbb{R}_+)} [w^\mathsf{T} \ \lambda] u \\
&= \inf_{(c,r) \in (C \times \mathbb{R}_+)} [w^\mathsf{T} \ \lambda] \left((\Gamma(x) - z - v, \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}) + (c,r) \right) \\
&= \inf_{(c,r) \in (C \times \mathbb{R}_+)} \left\{ w^\mathsf{T}(\Gamma(x) - z - v + c) + \lambda(\bar{w}^\mathsf{T}(v+z) - \tilde{\alpha} + r) \right\} \\
&= w^\mathsf{T}(\Gamma(x) - z - v) + \lambda(\bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}) + \inf_{(c,r) \in (C \times \mathbb{R}_+)} (w^\mathsf{T} c + \lambda r) \\
&= w^\mathsf{T}(\Gamma(x) - z - v) + \lambda(\bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}) + \inf_{c \in C} w^\mathsf{T} c + \inf_{r \in \mathbb{R}_+} \lambda r,
\end{aligned}$$

where

$$\inf_{c \in C} w^T c = \begin{cases} 0 & \text{if } w \in C^+, \\ -\infty & \text{otherwise} \end{cases}$$

and

$$\inf_{r \in \mathbb{R}_+} \lambda r = \begin{cases} 0 & \text{if } \lambda \in \mathbb{R}_+, \\ -\infty & \text{otherwise.} \end{cases}$$

Now, we have $\inf_{c \in C} w^{\mathsf{T}} c = -I_{C^+}(w)$ and $\inf_{r \in \mathbb{R}_+} \lambda r = -I_{\mathbb{R}_+}(\lambda)$ for every $(w, \lambda) \in \mathbb{R}^{q+1}$. Then, using the definitions of f and G, we obtain

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

if $(w, \lambda) \in C^+ \times \mathbb{R}_+$ and $\phi(w, \lambda) = -\infty$ otherwise. Finally, the dual problem of ($\mathrm{P}'_m(v)$) is formulated as

$$\text{maximize } \phi(w, \lambda) \text{ subject to} (w, \lambda) \in \mathbb{R}^{q+1}. \qquad (\mathrm{D}_m(v))$$

Then, the optimal value of ($\mathrm{D}_m(v)$) is given by

$$\begin{aligned} &\sup_{(w,\lambda) \in \mathbb{R}^{q+1}} \phi(w, \lambda) \\ &= \sup_{(w,\lambda) \in C^+ \times \mathbb{R}_+} \left\{ \inf_{x \in \mathcal{X}} w^{\mathsf{T}} \Gamma(x) - \sup_{z \in \mathbb{R}^q} \left\{ (w - \lambda \bar{w})^{\mathsf{T}} z - \|z\| \right\} - (w - \lambda \bar{w})^{\mathsf{T}} v - \lambda \tilde{\alpha} \right\} \\ &= \sup \left\{ \inf_{x \in \mathcal{X}} w^{\mathsf{T}} \Gamma(x) - (w - \lambda \bar{w})^{\mathsf{T}} v - \lambda \tilde{\alpha} \mid \|w - \lambda \bar{w}\|_* \leq 1, \ (w, \lambda) \in C^+ \times \mathbb{R}_+ \right\}. \end{aligned} \tag{3.2.3}$$

Here, we use that the conjugate function of $\|\cdot\|$ is the indicator function of the unit ball of the dual norm $\|\cdot\|_*$; see, for instance, [52, Example 3.26].

The problem may become infeasible for very small value of $\tilde{\alpha}$. In Section 4.2, we specify the value of $\tilde{\alpha}$ which ensures the generation of all non-dominated solution using the scalarization model. It can be shown easily that there exists $\tilde{\alpha}$ such that one can generate any non-dominated solution using the model. Define

$$S(\tilde{\alpha}) := \{ y \in \mathbb{R}^q \mid \bar{w}^{\mathsf{T}} y \leq \tilde{\alpha}, \ \bar{w} \in \operatorname{int} C^+, \ \tilde{\alpha} \in \mathbb{R} \}. \tag{3.2.4}$$

The next proposition shows the strong duality between ($\mathrm{P}_m(v)$) and ($\mathrm{D}_m(v)$).

Proposition 3.2.2. *Suppose that Assumption 2.2.1 holds. In the definition of S, given by (3.2.4), let $\tilde{\alpha}$ be large enough such that $\Gamma(\mathcal{X}) \subseteq \operatorname{int} S$. For every $v \in \mathbb{R}^q$, there exist optimal solutions (x^v, z^v) and (w^v, λ^v) of problems ($\mathrm{P}_m(v)$) and ($\mathrm{D}_m(v)$), respectively, and the optimal values coincide.*

Proof. First, we show that there exists some feasible solution $(\tilde{x}, \tilde{z})$ of ($\mathrm{P}_m(v)$), that is, feasible set of ($\mathrm{P}_m(v)$) is nonempty. Fix $\tilde{x} \in \mathcal{X}$, which exists since $\mathcal{X}$ has non-empty interior by our assumption. Define $\tilde{z} := \Gamma(\tilde{x}) - v$. Then, $v + \tilde{z} = \Gamma(\tilde{x}) \in \operatorname{int} S \subseteq S$. We know that, $v + \tilde{z} = \Gamma(\tilde{x}) \in \Gamma(\tilde{x}) + C$, since $0 \in C$ as its a cone. Then, by definition of a partial order $\leq_C$, we get $\Gamma(\tilde{x}) \leq_C v + \tilde{z}$ and hence $(\tilde{x}, \tilde{z})$ is feasible. Next, note that the feasible set of ($\mathrm{P}_m(v)$) is

$$\left\{(x,z) \in \mathbb{R}^{n+q} \;\mid\; x \in \mathcal{X},\; (\Gamma(x) - z - v) \leq_C 0,\; \bar{w}^{\mathsf{T}}(v+z) - \tilde{\alpha} \leq_{\mathbb{R}_+} 0\right\}$$
$$\subseteq \mathcal{X} \times (\mathcal{P} \cap S - \{v\}).$$

Using [53, Lemma 6.3], it is easy to see that $\mathcal{P} \cap S$ is compact. By Assumption 2.2.1, $\mathcal{X}$ is compact, it implies that the product $\mathcal{X} \times (\mathcal{P} \cap S - \{v\})$ is also compact. Then, it is enough to show that the feasible set is closed. Consider a sequence of feasible points $(x_n, z_n)_{n \geq 0}$ such that $\lim_{n\to\infty}(x_n, z_n) = (x, z)$. Then, we have $x \in \mathcal{X}$ as it is compact and $z \in S - v$ as it is a closed halfspace. We have, $v + z_n - \Gamma(x_n) \in C$ for any $n \geq 0$. Since $\Gamma : X \to \mathbb{R}^q$ is continuous and $\lim_{n\to\infty}(x_n, z_n) = (x, z)$,

$$\lim_{n\to\infty} (v + z_n - \Gamma(x_n)) = v + z - \Gamma(x),$$

which is in C as it is closed. Hence, an optimal solution (x^*, z^*) of ($\mathrm{P}_m(v)$) exists by Weierstrass, as the feasible set is compact and $\|\cdot\|$ is continuous.

Next, for the existence of an optimal solution of ($\mathrm{D}_m(v)$), we show that the constraint qualification in [51, 5] holds for ($\mathrm{P}_m(v)$). From Assumption 2.2.1, $\mathcal{X}$ has a non-empty interior, there exists $x^0 \in \operatorname{int} \mathcal{X}$. Then, as $\Gamma(\mathcal{X}) \subseteq \operatorname{int} S$, there exists $\epsilon > 0$ such that $B(\Gamma(x^0), \epsilon) \subseteq \operatorname{int} S$. Let us pick y^0 from $\Gamma(x^0) + \operatorname{int} C$, for instance, we can take

$$y^0 = \Gamma(x^0) + \frac{\epsilon}{2}\frac{c}{\|c\|},$$

where $c \in \operatorname{int} C$. Then, $y^0 \in B(\Gamma(x^0), \epsilon) \cap (\Gamma(x^0) + \operatorname{int} C)$ and $\Gamma(x^0) <_C y^0$. By defining $z^0 := y^0 - v$, we get $\Gamma(x^0) <_C v + z^0$, that is, $v + z^0 - \Gamma(x^0) \in \operatorname{int} C$. Since, $y^0 \in B(\Gamma(x^0), \epsilon) \subseteq \operatorname{int} S$, we have $\bar{w}^{\mathsf{T}} y^0 = \bar{w}^{\mathsf{T}}(v + z^0) < \tilde{\alpha}$. This gives, $-\bar{w}^{\mathsf{T}}(v + z^0) + \tilde{\alpha} \in \operatorname{int} \mathbb{R}_+$. By considering $G : X \times \mathbb{R}^q \rightrightarrows \mathbb{R}^{q+1}$ as in (3.2.1), we get,

$$\begin{aligned} G(x^0, z^0) = \{(\Gamma(x^0) - z^0 - v, \bar{w}^{\mathsf{T}}(v + z^0) - \tilde{\alpha})\} &\subseteq -(\operatorname{int}\ C \times \operatorname{int}\ \mathbb{R}_+) \\ &= -\operatorname{int}(C \times \mathbb{R}_+). \end{aligned} \quad (3.2.5)$$

From (3.2.1), we have $f : (X, \mathbb{R}^q) \to \overline{\mathbb{R}}$ with the domain

$$\operatorname{dom}\ f = \{(x, z) \in X \times \mathbb{R}^q \ \mid \ f(x, z) < +\infty\}.$$

By the definition of $f(x, z)$, we know that

$$f(x, z) = \begin{cases} \|z\| & \text{if } x \in \mathcal{X} \\ +\infty & \text{otherwise.} \end{cases}$$

Hence, $\operatorname{dom} f = \mathcal{X} \times \mathbb{R}^q$. Since $x^0 \in \mathcal{X}$, $G(x^0, z^0) \subseteq G(\operatorname{dom} f)$. Then, from (3.2.5) we have

$$G(\operatorname{dom} f) \cap -\operatorname{int}(C \times \mathbb{R}_+) \neq \emptyset, \quad (3.2.6)$$

Hence, (x^0, z^0) satisfies constraint qualification. Next, we show that our set valued map $G : (X, \mathbb{R}^q) \rightrightarrows \mathbb{R}^{q+1}$ is C-convex. We have $\Gamma : X \to \mathbb{R}^q$ a C-convex function, that is, for $x^1, x^2 \in X$ and $\lambda \in [0, 1]$, $\Gamma(\lambda x^1 + (1-\lambda)x^2) \leq_C \lambda\Gamma(x^1) + (1-\lambda)\Gamma(x^2)$. By partial ordering $\leq_C$,

$$\lambda\Gamma(x^1) + (1 - \lambda)\Gamma(x^2) \in \Gamma(\lambda x^1 + (1 - \lambda)x^2) + C.$$

Introducing $z+v$ above,

$$\lambda(\Gamma(x^1)-z-v)+(1-\lambda)(\Gamma(x^2)-z-v)\in\Gamma(\lambda x^1+(1-\lambda)x^2)-z-v+C.$$

Since, $0\in\mathbb{R}_+$, $\bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}\in\bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}+\mathbb{R}_+$. Then, we have

$$\lambda\begin{bmatrix}\Gamma(x^1)-z-v\\ \bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}\end{bmatrix}+(1-\lambda)\begin{bmatrix}\Gamma(x^2)-z-v\\ \bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}\end{bmatrix}\in$$
$$\begin{bmatrix}\Gamma(\lambda x^1+(1-\lambda)x^2)-z-v\\ \bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}\end{bmatrix}+(C\times\mathbb{R}_+)$$

From the definition of $G(x,z)$, it implies

$$\lambda G(x^1,z)+(1-\lambda)G(x^2,z)\subseteq G(\lambda(x^1,z)+(1-\lambda)(x^2,z))+(C\times\mathbb{R}_+).$$

Hence, $G: X\times\mathbb{R}^q\rightrightarrows\mathbb{R}^{q+1}$ is $C\times\mathbb{R}_+$-convex by definition [51, Subsection 8.3.2].

Since, $f:(X,\mathbb{R}^q)\to\overline{\mathbb{R}}$ is convex, $G: X\times\mathbb{R}^q\rightrightarrows\mathbb{R}^{q+1}$ is $(C\times\mathbb{R}_+)$-convex and (3.2.6) holds, it implies strong duality and dual attainment, see [51, Theorem 8.3.10] and [5, Theorem 3.19]. Therefore, an optimal solution (w^v,λ^v) of ($\mathrm{D}_m(v)$) exists. □

Similar to Lemma 3.1.5, we characterize, through the following lemma, the cases where $z^v=0$ in the modified scalarization.

Lemma 3.2.3. *Suppose that Assumption 2.2.1 holds. Let $\Gamma(\mathcal{X})\subseteq\operatorname{int}S$ and let $v\in\mathbb{R}^q$ such that $v\in S$. The following statements hold.*

(a) If $v\notin\mathcal{P}$, then $z^v\neq0$, $w^v\neq0$ and $\tilde{w}=w^v-\lambda^v\bar{w}\neq0$.

(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 $z^v=0$.

Proof. Clearly, $z^v \neq 0$ and $w^v \neq 0$ in (a), see the proof of Lemma 3.1.5 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 ($\mathrm{D}_m(v)$) and strong duality, we get

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

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

To prove (b) and (c), note that, we have $\bar{w}^\mathsf{T} v \leq \tilde{\alpha}$. Then the assertion follows using the similar arguments as in the proof of Lemma 3.1.5. □

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

Proposition 3.2.4. *Suppose that Assumption 2.2.1 holds. If $v \notin \operatorname{int} \mathcal{P}$ and $v \in S$, then x^v is a weak minimizer of (P), and $y^v := v + z^v \in \mathrm{wMin}_C(\mathcal{P})$.*

Proof. By using similar arguments as in the proof of Proposition 3.1.6, x^v is a weak minimizer of (P).

For the second statement, 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 $\|z^v\| \neq 0$, as $v \notin \operatorname{int} \mathcal{P}$ from 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, $\bar{w}^\mathsf{T}(v + \bar{z}) \leq \tilde{\alpha}$. Assume otherwise,

$$\bar{w}^\mathsf{T}(v + \bar{z}) = \bar{w}^\mathsf{T}\left(v + z^v - \epsilon\frac{z^v}{\|z^v\|}\right) > \tilde{\alpha}. \tag{3.2.7}$$

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

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

This implies, $\bar{w}^\mathsf{T} z^v < 0$. We also know that, $\bar{w}^\mathsf{T} v \leq \tilde{\alpha}$, as $v \in S$. Then from (3.2.7), we get,

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

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

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

Proposition 3.2.5. *Suppose that Assumption 2.2.1 holds and* $\tilde{w}^v = w^v - \lambda^v \bar{w} \neq 0$. *Let,* $v \in S$ *and* $\Gamma(\mathcal{X}) \subseteq \operatorname{int} S$. *Then, the halfspace*

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

contains $\mathcal{P} \cap S$. *Moreover,* $\operatorname{bd} \mathcal{H}$ *is a supporting hyperplane of* $\mathcal{P} \cap S$ *at* $y^v = v + z^v$.

Proof. Let $y \in \mathcal{P} \cap S$ be arbitrary and $\bar{x} \in \mathcal{X}$ be such that $\Gamma(\bar{x}) \leq_C y$. Consider the problems ($\mathrm{P}_m(y)$) and ($\mathrm{D}_M(y)$). Let $(\bar{x},\ \bar{z})$ be feasible for ($\mathrm{P}_m(y)$). Then, $(\bar{x}, 0)$ is feasible, as $y + \bar{z} = y \in \mathcal{P} = \Gamma(\mathcal{X}) + C$ and $\bar{w}^\mathsf{T}(y + \bar{z}) = \bar{w}^\mathsf{T} y \leq \tilde{\alpha}$. Then, for ($\mathrm{P}_m(y)$), we have $\|\bar{z}\| = 0$. Moreover, the optimal solution (λ^v, w^v) of ($\mathrm{D}_m(v)$) is also feasible for ($\mathrm{D}_m(y)$), since they have the same feasible region. Using weak duality for ($\mathrm{P}_m(y)$) and ($\mathrm{D}_m(y)$) we obtain

$$0 = \|\bar{z}\| \geq \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T}(\Gamma(x) - y) + \lambda^v(\bar{w}^\mathsf{T} y - \tilde{\alpha})$$

$$(w^v)^\mathsf{T} y - \lambda^v \bar{w}^\mathsf{T} y \geq \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T} \Gamma(x) - \lambda^v \tilde{\alpha} \tag{3.2.8}$$

Using strong duality for ($\mathrm{P}_m(v)$) and ($\mathrm{D}_m(v)$), we have

$$\|z^v\| = \inf_{x \in \mathcal{X}} (w^v)^\mathsf{T} \Gamma(x) - (w^v - \lambda \bar{w})^\mathsf{T} v - \lambda^v \tilde{\alpha} \tag{3.2.9}$$

From (3.2.8) and (3.2.9) we obtain

$$\begin{aligned} (w^v)^\mathsf{T} y - \lambda^v \bar{w}^\mathsf{T} y &\geq \|z^v\| + (w^v - \lambda \bar{w})^\mathsf{T} v \\ (\tilde{w}^v)^\mathsf{T} y &\geq \|z^v\| + (\tilde{w}^v)^\mathsf{T} v \end{aligned} \tag{3.2.10}$$

Using the definition of dual norm and the dual constraint, $\|\tilde{w}\|_* \leq 1$, we obtain

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

Then from (3.2.10) we get

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

This implies $y \in \mathcal{H}$ and hence $\mathcal{P} \cap S \subseteq \mathcal{H}$.

Next, from Proposition 3.2.4, we have $y^v \in \mathcal{P}$, in particular $y^v \in \mathcal{P} \cap S$ from the feasibility of ($\mathrm{P}_m(v)$). We have $y^v \in \operatorname{bd} \mathcal{H}$ by definition of $\mathcal{H}$. Hence, $y^v \in \mathcal{P} \cap S \cap \operatorname{bd} \mathcal{H}$ which completes the proof. □

Corollary 3.2.6. *Suppose that Assumption 2.2.1 holds and $\|\cdot\| = \|\cdot\|_p$ for some $p \in [1, +\infty)$. Assume that $v \notin \mathcal{P}$. Then, the halfspace*

$$\mathcal{H} = \left\{ y \in \mathbb{R}^q \mid \sum_{i=1}^{q} \operatorname{sgn}(z_i^v) |z_i^v|^{p-1} y_i \geq \sum_{i=1}^{q} \operatorname{sgn}(z_i^v) |z_i^v|^{p-1} y_i^v \right\}$$

contains the upper image $\mathcal{P}$, where sgn *is the usual sign function. Moreover,* $\operatorname{bd} \mathcal{H}$ *is a supporting hyperplane of $\mathcal{P}$ at $y^v = v + z^v$.*

Proof. Consider ($\mathrm{P}_m(v)$) and its Lagrange dual ($\mathrm{D}_m(v)$). Let us define $g(z) := \|z\|_p - (w - \lambda \bar{w})^\mathsf{T} z = \|z\|_p - \tilde{w}^\mathsf{T} z, z \in \mathbb{R}^q$. Then, the assertion follows using the similar arguments as in the proof of Corollary 3.1.8 □

3.3 Linearization of objective function of ($\mathrm{P}_Z(v)$) for ℓ_1 and ℓ_∞ norms

In this section, we linearize the objective function of the modified scalarization, in case of ℓ_1 and ℓ_∞ norms. The aim is to input the linearized version to a solver which might make it easier for the solver to handle the scalarization using ℓ_1 and ℓ_∞ norms. For simplification, we re-define the constraint (regarding the feasiblity of $\mathcal{P}$) of scalarization model in terms of generating vectors of the dual of ordering cone C^+, instead of using a set function as in previous section.

Let C be polyhedral and Z is the matrix whose columns are the generating vectors of C^+. For a fixed arbitrary norm $\|\cdot\|$ on $\mathbb{R}^q$ and a point $v \in \mathbb{R}^q$, the *modified norm-minimizing scalarization* of (P) can be written as

$$\text{minimize } \|z\| \text{ subject to } Z^\mathsf{T}(\Gamma(x)-z-v) \leq 0,\ \bar{w}^\mathsf{T}(v+z)-\tilde{\alpha} \leq 0,\ x \in \mathcal{X},\ z \in \mathbb{R}^q, \tag{$\mathrm{P}_Z(v)$}$$

where $\bar{w} \in \operatorname{int} C^+$ and $\tilde{\alpha} \in \mathbb{R}$ are as in Section 3.2. Before discussing the special cases of ℓ_1 and ℓ_∞ norms, we first find the dual problem of the above formulation. Define the Lagrangian $L\colon X \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R} \to \overline{\mathbb{R}}$ as follows,

$$\begin{aligned} L(x,z,\lambda,u) &:= \|z\| + \lambda^\mathsf{T}(Z^\mathsf{T}(\Gamma(x)-z-v)) + u(\bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}) \\ &= \|z\| + (Z\lambda)^\mathsf{T}(\Gamma(x)-z-v) + u(\bar{w}^\mathsf{T}v + \bar{w}^\mathsf{T}z - \tilde{\alpha}). \end{aligned} \tag{3.3.1}$$

Now, the dual objective function $\phi\colon \mathbb{R}^q \times \mathbb{R} \to \overline{\mathbb{R}}$ is defined by

$$\phi(\lambda,u) := \inf_{x \in X, z \in \mathbb{R}^q} L(x,z,\lambda,u).$$

Note that, $Z\lambda \in C^+$, as it is a conic combination of the generating vectors of C^+.

Define, $w := Z\lambda$. After this change of variables we get,

$$\begin{aligned}\phi(w,u) &= \inf_{x\in\mathcal{X}} w^{\mathsf{T}}\Gamma(x) + \inf_{z\in\mathbb{R}^q}\left(\|z\| - (w-u\bar{w})^{\mathsf{T}}z\right) - (w-u\bar{w})^{\mathsf{T}}v - u\tilde{\alpha} \\ &= \inf_{x\in\mathcal{X}} w^{\mathsf{T}}\Gamma(x) - \sup_{z\in\mathbb{R}^q}\left((w-u\bar{w})^{\mathsf{T}}z - \|z\|\right) - (w-u\bar{w})^{\mathsf{T}}v - u\tilde{\alpha},\end{aligned}$$

for $w \in C^+, u \in \mathbb{R}_+$. Note that, we have

$$\phi(w,u) = \inf_{x\in\mathcal{X}} w^{\mathsf{T}}\Gamma(x) - (w-u\bar{w})^{\mathsf{T}}v - u\tilde{\alpha}$$

if $\|w-u\bar{w}\| \leq 1$ and $\phi(w,u) = -\infty$ otherwise. The dual problem of ($\mathrm{P}_Z(v)$) is formulated as

$$\text{maximize } \phi(w,u) \text{ subject to } (w,u) \in C^+ \times \mathbb{R}_+. \tag{$\mathrm{D}_Z(v)$}$$

Finally, the optimal value of ($\mathrm{D}_Z(v)$) is as follows,

$$\phi(w,u) = \sup\left\{\inf_{x\in\mathcal{X}} w^{\mathsf{T}}\Gamma(x) - (w-u\bar{w})^{\mathsf{T}}v - u\tilde{\alpha} \mid \|w-u\bar{w}\|_* \leq 1,\ (w,u) \in C^+ \times \mathbb{R}_+\right\}. \tag{3.3.2}$$

3.3.1 Linearization of objective function of ($\mathrm{P}_Z(v)$) using ℓ_1 norm

Consider ℓ_1 norm for the scalarization. Then, the objective function can be linearized as

$$\begin{aligned} \text{minimize } & \sum_{i=1}^{q}(t_i + s_i) \qquad\qquad (\mathrm{P}_{\ell_1}(v)) \\ \text{subject to } & Z^{\mathsf{T}}(\Gamma(x) - z - v) \leq 0, \\ & \bar{w}^{\mathsf{T}}(v+z) - \tilde{\alpha} \leq 0, \\ & z = t - s, \\ & x \in \mathcal{X},\ z \in \mathbb{R}^q,\ t, s \in \mathbb{R}^q_+, \end{aligned}$$

where Z is the matrix whose columns are the generating vectors of C^+. We define the Lagrangian $L\colon X \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^q \to \overline{\mathbb{R}}$ as follows,

$$\begin{aligned} L(x,z,t,s,\lambda,u,\bar{\lambda}) &:= e^\mathsf{T}(t+s) + \lambda^\mathsf{T}(Z^\mathsf{T}(\Gamma(x)-z-v)) + u(\bar{w}^\mathsf{T}(v+z)-\tilde{\alpha}) \\ &\qquad + \bar{\lambda}^\mathsf{T}(z-t+s) \\ &= e^\mathsf{T}(t+s) + (Z\lambda)^\mathsf{T}(\Gamma(x)-z-v) + u(\bar{w}^\mathsf{T}v + \bar{w}^\mathsf{T}z - \tilde{\alpha}) \\ &\qquad + \bar{\lambda}^\mathsf{T}(z-t+s), \end{aligned} \tag{3.3.3}$$

where $e = (1,\dots,1)^\mathsf{T} \in \mathbb{R}^q$. Similar to Section 3.3, $w := Z\lambda \in C^+$, and $\phi\colon \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^q \to \overline{\mathbb{R}}$ is defined by

$$\phi(\lambda,u,\bar{\lambda}) := \inf_{x\in X, z\in\mathbb{R}^q, t,s\in\mathbb{R}^q_+} L(x,z,t,s,\lambda,u,\bar{\lambda}).$$

We get

$$\begin{aligned} \phi(w,u,\bar{\lambda}) = \inf_{x\in\mathcal{X}} w^\mathsf{T}\Gamma(x) + \inf_{z\in\mathbb{R}^q}(-w+u\bar{w}+\bar{\lambda})^\mathsf{T}z + \inf_{t\in\mathbb{R}^q_+}(e-\bar{\lambda})^\mathsf{T}t + \inf_{s\in\mathbb{R}^q_+}(e+\bar{\lambda})^\mathsf{T}s \\ -(w-u\bar{w})^\mathsf{T}v - u\tilde{\alpha}, \end{aligned}$$

for $w \in C^+$, $u \in \mathbb{R}_+$ and $\bar{\lambda} \in \mathbb{R}^q$. Above equation gives, $-e \le \bar{\lambda} \le e$. This implies $\|\bar{\lambda}\|_\infty \le 1$. From second term above, we get $w - u\bar{w} = \bar{\lambda}$. This implies

$$\|w-u\bar{w}\|_\infty = \|\bar{\lambda}\|_\infty \le 1.$$

The dual problem of ($\mathrm{P}_{\ell_1}(v)$) is formulated as,

$$\text{maximize } \phi(w,u) \text{ subject to } (w,u) \in C^+ \times \mathbb{R}_+. \tag{$\mathrm{D}_{\ell_1}(v)$}$$

Finally, the optimal value of ($\mathrm{D}_{\ell_1}(v)$) is as follows,

$$\phi(w,u) = \\ \sup\left\{\inf_{x\in\mathcal{X}} w^\mathsf{T}\Gamma(x) - (w-u\bar{w})^\mathsf{T}v - u\tilde{\alpha} \mid \|w-u\bar{w}\|_\infty \le 1,\ (w,u)\in C^+\times\mathbb{R}_+\right\} \tag{3.3.4}$$

3.3.2 Linearization of objective function of $(P_Z(v))$ using ℓ_∞ norm

Consider ℓ_∞ norm for the scalarization. Then, the objective function can be linearized by considering a scalar variable t such that $\|z\|_\infty \le t$. This means that $\max_{i\in\{1,\dots,q\}}|z_i| \le t$, which implies $|z_i| \le t$, for all $i \in \{1,\dots,q\}$. Hence, $-te \le z \le te$, where $e = (1,\dots,1)^\mathsf{T} \in \mathbb{R}^q$. Then, we have

$$\begin{aligned} &\text{minimize } t && (\mathrm{P}_{\ell_\infty}(v))\\ &\text{subject to } Z^\mathsf{T}(\Gamma(x) - z - v) \le 0, \\ &\qquad \bar{w}^\mathsf{T}(v+z) - \tilde{\alpha} \le 0, \\ &\qquad -te \le z \le te, \\ &\qquad x \in \mathcal{X},\ z \in \mathbb{R}^q,\ t \in \mathbb{R}_+, \end{aligned}$$

We define the Lagrangian $L\colon X \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^q \to \overline{\mathbb{R}}$ as follows,

$$\begin{aligned} L(x,z,t,\lambda,u,\lambda',\lambda'') &:= t + \lambda^\mathsf{T}(Z^\mathsf{T}(\Gamma(x) - z - v)) + u(\bar{w}^\mathsf{T}(v+z) - \tilde{\alpha}) \\ &\qquad + (\lambda')^\mathsf{T}(z - te) + (\lambda'')^\mathsf{T}(-z - te) \\ &= t + w^\mathsf{T}(\Gamma(x) - z - v) + u(\bar{w}^\mathsf{T}v + \bar{w}^\mathsf{T}z - \tilde{\alpha}) \\ &\qquad + (\lambda' - \lambda'')^\mathsf{T}z - e^\mathsf{T}(\lambda' + \lambda'')t. \end{aligned} \tag{3.3.5}$$

Now, the dual objective function $\phi\colon \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^q \times \mathbb{R}^q \to \overline{\mathbb{R}}$ is defined by

$$\phi(w,u,\lambda',\lambda'') := \inf_{x\in X, z\in\mathbb{R}^q, t\in\mathbb{R}_+} L(x,z,t,\lambda,u,\lambda',\lambda'').$$

We get

$$\begin{aligned} \phi(w,u,\lambda',\lambda'') = \inf_{x\in\mathcal{X}} w^\mathsf{T}\Gamma(x) + \inf_{z\in\mathbb{R}^q}(-w+u\bar{w}+\lambda'-\lambda'')^\mathsf{T}z + \inf_{t\in\mathbb{R}_+}(1-e^\mathsf{T}\lambda'-e^\mathsf{T}\lambda'')^\mathsf{T}t \\ - (w-u\bar{w})^\mathsf{T}v - u\tilde{\alpha} \end{aligned}$$

From above equation, $e^\mathsf{T}\lambda' + e^\mathsf{T}\lambda'' \le 1$. Note that $\lambda', \lambda'' \ge 0$ as a constraint of the dual problem. The second term above gives, $w - u\bar{w} = \lambda' - \lambda''$. Then, we have, $\|w - u\bar{w}\|_1 = \|\lambda' - \lambda''\|_1 \leq 1$.

The dual problem of ($\mathrm{P}_{\ell_\infty}(v)$) is formulated as,

$$\text{maximize } \phi(w,u) \text{ subject to } (w,u) \in C^+ \times \mathbb{R}_+. \tag{$\mathrm{D}_{\ell_\infty}(v)$}$$

Finally, the optimal value of ($\mathrm{D}_{\ell_\infty}(v)$) is as follows,

$$\phi(w,u) = \sup \left\{ \inf_{x \in \mathcal{X}} w^{\mathsf{T}} \Gamma(x) - (w - u\bar{w})^{\mathsf{T}} v - u\tilde{\alpha} \mid \|w - u\bar{w}\|_1 \leq 1,\ (w,u) \in C^+ \times \mathbb{R}_+ \right\}. \tag{3.3.6}$$

Chapter 4

Approximation Algorithms for Solving CVOP

In this chapter, we propose an algorithm which finds an outer approximation of the upper image in order to solve CVOPs. Moreover, we also introduce two variants of the algorithm. All three algorithms are based on norm-minimizng scalarizations introduced in Chapter 2. Here, we provide certain propositions that lead to prove that the algorithms work correctly provided they terminate after some iterations. The finiteness results in case of later two algorithms are given in Chapter 5.

4.1 The Main Algorithm: Algorithm 1

We propose an outer approximation algorithm for finding a finite weak ϵ-solution to CVOP as in Definition 2.2.5. The algorithm is based on solving norm minimization scalarizations iteratively. The design of the algorithm is similar to the "Benson-type algorithms" in the literature; see, for instance, [9, 20, 1]. It starts by finding a polyhedral outer approximation $\mathcal{P}_0^{\text{out}}$ of $\mathcal{P}$ and iterates in order to form a sequence $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}_1^{\text{out}} \supseteq \ldots \supseteq \mathcal{P}$ of finer approximating sets.

Before providing the details regarding the algorithm, we impose a further assumption on C.

Assumption 4.1.1. *The ordering cone C is polyhedral.*

Assumption 4.1.1 implies that the dual cone C^+ is also polyhedral. We denote the set of generating vectors of C^+ by $\{w^1, \dots, w^J\}$, where $J \in \mathbb{N}$. Hence, $C^+ = \operatorname{conv\,cone}\{w^1, \dots, w^J\}$. Moreover, under Assumption 2.2.1, C^+ is solid since C is pointed [54, Propostion 2.4.3]. Hence, we have $J \geq q$.

The algorithm starts by solving the weighted sum scalarizations corresponding to each generating vector of C^+ (WS(w^1)), $\dots$, (WS(w^J)). For each $j \in \{1, \dots, J\}$, the existence of an optimal solution $x^j \in \mathcal{X}$ of (WS(w^j)) is guaranteed by Assumption 2.2.1 (b, c). The initial set of weak minimizers is set as $\mathcal{X}_0 := \{x^1, \dots, x^J\}$, see Proposition 2.2.7.[1] The set $\mathcal{V}^{\text{known}}$, which keeps the set of all points $v \in \mathbb{R}^q$ for which (P(v)) and (D(v)) are solved throughout the algorithm, is initialized as the empty set. Moreover, similar to the primal algorithm in [1], the initial outer approximation is set as

$$\mathcal{P}_0^{\text{out}} := \bigcap_{j=1}^{J} \{y \in \mathbb{R}^q \mid (w^j)^\mathsf{T} y \geq (w^j)^\mathsf{T} \Gamma(x^j)\} \tag{4.1.1}$$

(see lines 1-3 of Algorithm 1). It is not difficult to see that $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$. Indeed, for each $\bar{y} \in \mathcal{P}$, there exists $\bar{x} \in \mathcal{X}$ such that $\Gamma(\bar{x}) \leq_C \bar{y}$. Then, for each $j \in \{1, \dots, J\}$, we have $(w^j)^\mathsf{T}\Gamma(\bar{x}) \leq (w^j)^\mathsf{T}\bar{y}$ which implies $(w^j)^\mathsf{T}\bar{y} \geq \inf_{x \in \mathcal{X}} (w^j)^\mathsf{T}\Gamma(x) = (w^j)^\mathsf{T}\Gamma(x^j)$ so that $\bar{y} \in \mathcal{P}_0^{\text{out}}$. Moreover, as C is pointed and (P) is bounded, $\mathcal{P}_0^{\text{out}}$ has at least one vertex, see [55, Corollary 18.5.3] (as well as [1, Section 4.1]).

At an arbitrary iteration $k \geq 0$ of the algorithm, the set $\mathcal{V}_k$ of vertices of the current outer approximation $\mathcal{P}_k^{\text{out}}$ is computed first (line 6).[2] Then, for each

[1]Alternatively, one may start with $\mathcal{X}_0 = \emptyset$ in line 2 of Algorithm 1. This would decrease $|\mathcal{X}|$ by J, the number of generating vectors C.

[2]This is done by solving a vertex enumeration problem for $\mathcal{P}_k^{\text{out}}$, that is, from the H-representation of $\mathcal{P}_k^{\text{out}}$, its V-representation is computed. For the computational tests of Section 6.1, we use *bensolve tools* for this purpose [10].

$v \in \mathcal{V}_k$, if not done before, the norm-minimizing scalarization (P(v)) and its dual (D(v)) are solved in order to find optimal solutions (x^v,z^v) and w^v, respectively (see Proposition 3.1.2).[3] Moreover, v is added to $\mathcal{V}^{\text{known}}$ (lines 7-10). If the distance from v to $\mathcal{P}$, namely $\|z^v\|$, is less than or equal to the predetermined approximation error $\epsilon > 0$, then x^v is added to the set of weak minimizers (see Proposition 3.1.6)[4] and the algorithm continues by considering the remaining vertices of $\mathcal{P}_k$ (line 18). Otherwise, the supporting halfspace

$$\mathcal{H}_k := \{y \in \mathbb{R}^q \mid (w^v)^\mathsf{T} y \geq (w^v)^\mathsf{T} \Gamma(x^v)\} \tag{4.1.2}$$

of $\mathcal{P}$ at $\Gamma(x^v)$ is found (see Proposition 3.1.7); and the current approximation is updated as $\mathcal{P}_{k+1}^{\text{out}} = \mathcal{P}_k^{\text{out}} \cap \mathcal{H}_k$ (lines 12). The algorithm terminates if all the vertices in $\mathcal{V}_k$ are within ϵ distance to the upper image (lines 5, 15, 16, 22).

By the design of the algorithm, for each iteration $k \geq 0$, the set $\mathcal{P}_k^{\text{out}} \supseteq \mathcal{P}$ is an outer approximation of the upper image; similarly, we define an inner approximation of $\mathcal{P}$ by

$$\mathcal{P}_k^{\text{in}} := \operatorname{conv} \Gamma(\mathcal{X}_k) + C \subseteq \mathcal{P}. \tag{4.1.3}$$

In order to visualize an iteration of the algorithm, consider $\mathcal{P} \in \mathbb{R}^2$ with an initial outer approximation $\mathcal{P}_0^{\text{out}}$ and $\mathcal{V}_0 = \{v\}$, see Figure 4.1. The scalarization is solved against the vertex v and optimal solutions (x^v, z^v) and w^v are obtained. Subsequently, the current outer approximation is updated to $\mathcal{P}_1^{\text{out}} = \mathcal{P}_0^{\text{out}} \cap \mathcal{H}_0$ with a set of vertices $\mathcal{V}_1 = \{v^1, v^2\}$.

Now, we present two lemmas regarding these inner and outer approximations. The first one shows that, for each $k \geq 0$, the sets $\mathcal{P}_k^{\text{out}}$ and $\mathcal{P}_k^{\text{in}}$ have the same recession cone, which is the ordering cone C. With the second lemma, we see

[3] Note that many solvers yield both primal and dual optimal solutions when called only for one of the problems.

[4] Note that the solution x^v found in line 9 of Algorithm 1 is a weak minimizer, it is also possible to update the set of weak minimizers right after line 9 (without checking the value of $\|z^v\|$) and subsequently ignore lines 13 and 18. This would yield a finite weak ϵ-solution with an increased cardinality.

Algorithm 1 Outer Approximation Algorithm for (P)

Compute an optimal solution x^j of (WS(w^j)) for each $j \in \{1, \dots, J\}$;
Set $k = 0, \mathcal{X}_0 = \{x^1, \dots, x^J\}, \mathcal{V}^{\text{known}} = \emptyset$;
Store an H-representation of $\mathcal{P}_0^{\text{out}}$ according to (4.1.1);
repeat
  Stop $\leftarrow$ **true**;
  Compute the set $\mathcal{V}_k$ of vertices of $\mathcal{P}_k^{\text{out}}$ from its H-representation;
  for $v \in \mathcal{V}_k$ **do**
    if $v \notin \mathcal{V}^{\text{known}}$ **then**
      Solve (P(v)) and (D(v)) to compute (x^v, z^v) and w^v;
      $\mathcal{V}^{\text{known}} \leftarrow \mathcal{V}^{\text{known}} \cup \{v\}$;
      if $\|z^v\| > \epsilon$ **then**
        $\mathcal{P}_{k+1}^{\text{out}} = \mathcal{P}_k^{\text{out}} \cap \mathcal{H}_k$;
        $\mathcal{X}_{k+1} = \mathcal{X}_k$;
        $k \leftarrow k + 1$;
        Stop $\leftarrow$ **false**;
        break;
      else
        $\mathcal{X}_k \leftarrow \mathcal{X}_k \cup \{x^v\}$;
      end if
    end if
  end for
until Stop
return $\begin{cases} \mathcal{X}_k & : \text{A finite weak } \epsilon\text{-solution to (P)}; \\ \mathcal{P}_k^{\text{out}} & : \text{An outer approximation of } \mathcal{P}. \end{cases}$

that, in order to compute the Hausdorff distance between $\mathcal{P}_k^{\text{out}}$ and $\mathcal{P}$ (or $\mathcal{P}_k^{\text{in}}$), it is sufficient to consider the set $\mathcal{V}_k$ of vertices of $\mathcal{P}_k^{\text{out}}$.

Lemma 4.1.2. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $k \geq 0$. Then,*

$$\operatorname{recc} \mathcal{P}_k^{out} = \operatorname{recc} \mathcal{P}_k^{in} = \operatorname{recc} \mathcal{P} = C.$$

Proof. As $\operatorname{conv} \Gamma(\mathcal{X}_k)$ is a compact set, we have $\operatorname{recc} \mathcal{P}_k^{\text{in}} = C$ directly from (4.1.3). Similarly, since $\Gamma(\mathcal{X})$ is compact by Assumption 2.2.1 and $\mathcal{P} = \Gamma(\mathcal{X}) + C$ by Remark 2.2.2, we have $\operatorname{recc} \mathcal{P} = C$. Since $\mathcal{P}_k^{\text{out}} \supseteq \mathcal{P}$, we have $\operatorname{recc} \mathcal{P}_k^{\text{out}} \supseteq \operatorname{recc} \mathcal{P} = C$; see [56, Proposition 2.5]. In order to conclude that $\operatorname{recc} \mathcal{P}_k^{\text{out}} = C$, it is enough to show that $\operatorname{recc} \mathcal{P}_0^{\text{out}} \subseteq C$. Indeed, we have $\mathcal{P}_k^{\text{out}} \subseteq \mathcal{P}_0^{\text{out}}$, which implies that $\operatorname{recc} \mathcal{P}_k^{\text{out}} \subseteq \operatorname{recc} \mathcal{P}_0^{\text{out}}$. To prove $\operatorname{recc} \mathcal{P}_0^{\text{out}} \subseteq C$, let $\bar{y} \in \operatorname{recc} \mathcal{P}_0^{\text{out}}$. Then, for

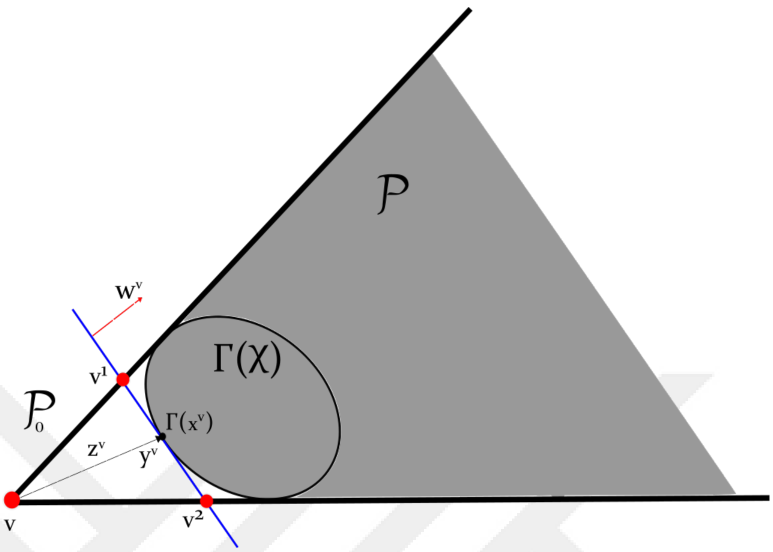


Figure 4.1: An iteration of Algorithm 1.

each $y \in \mathcal{P}_0^{\text{out}}$, we have $y + \bar{y} \in \mathcal{P}_0^{\text{out}}$. By the definition of $\mathcal{P}_0^{\text{out}}$ in (4.1.1), we have $(w^j)^\mathsf{T}(y + \bar{y}) \geq (w^j)^\mathsf{T}\Gamma(x^j)$ for each $j \in \{1, \ldots, J\}$. In particular, as $\Gamma(x^j) \in \mathcal{P} \subseteq \mathcal{P}_0^{\text{out}}$, we have $(w^j)^\mathsf{T}(\Gamma(x^j) + \bar{y}) \geq (w^j)^\mathsf{T}\Gamma(x^j)$, hence $(w^j)^\mathsf{T}\bar{y} \geq 0$ for each $j \in \{1, \ldots, J\}$. By the definition of dual cone and using $(C^+)^+ = C$, we have $\bar{y} \in C$. The assertion holds as $\bar{y} \in \operatorname{recc} \mathcal{P}_0^{\text{out}}$ is arbitrary. $\square$

It is well known that for a convex set $P \subseteq \mathbb{R}^q$, the distance function $d(\cdot, P)$ is convex [57, Proposition 1.77]. Here, we give the result for completeness. Let $x^1, x^2 \in \mathbb{R}^q$. As $d(x^i, P) = \inf_{p \in P} \|x^i - p\|$ and P is convex, for any $\epsilon > 0$, there exists $p^i \in P$ such that

$$\|x^i - p^i\| < d(x^i, P) + \epsilon,$$

for $i = 1, 2$. Then, for $\lambda \in [0, 1]$ we have

$$\begin{aligned} d(\lambda x^1 + (1-\lambda)x^2, P) &\leq d(\lambda x^1 + (1-\lambda)x^2, \lambda p^1 + (1-\lambda)p^2) \\ &= \|\lambda x^1 + (1-\lambda)x^2 - \lambda p^1 - (1-\lambda)p^2\| \\ &= \|\lambda(x^1 - p^1) + (1-\lambda)(x^2 - p^2)\| \\ &\leq \lambda\|x^1 - p^1\| + (1-\lambda)\|x^2 - p^2\| \\ &< \lambda\big(d(x^1, P) + \epsilon\big) + (1-\lambda)\big(d(x^2, P) + \epsilon\big) \\ &= \lambda d(x^1, P) + (1-\lambda)d(x^2, P) + \epsilon \end{aligned}$$

As $\epsilon > 0$ is arbitrary, we have

$$d(\lambda x^1 + (1-\lambda)x^2, P) \leq \lambda d(x^1, P) + (1-\lambda)d(x^2, P).$$

Lemma 4.1.3. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $k \geq 0$. Then,*

$$\delta^H(\mathcal{P}_k^{out}, \mathcal{P}_k^{in}) = \max_{v \in \mathcal{V}_k} d(v, \mathcal{P}_k^{in}) \ \ \textit{and} \ \ \delta^H(\mathcal{P}_k^{out}, \mathcal{P}) = \max_{v \in \mathcal{V}_k} d(v, \mathcal{P}),$$

where $\mathcal{V}_k$ is the set of vertices of $\mathcal{P}_k^{out}$.

Proof. To see the first equality, note that

$$\delta^H(\mathcal{P}_k^{\text{out}}, \mathcal{P}_k^{\text{in}}) = \max\left\{\sup_{y \in \mathcal{P}_k^{\text{out}}} d(y, \mathcal{P}_k^{\text{in}}), \sup_{y \in \mathcal{P}_k^{\text{in}}} d(y, \mathcal{P}_k^{\text{out}})\right\} = \sup_{y \in \mathcal{P}_k^{\text{out}}} d(y, \mathcal{P}_k^{\text{in}})$$

as $\mathcal{P}_k^{\text{in}} \subseteq \mathcal{P}_k^{\text{out}}$. Moreover, since $\operatorname{recc} \mathcal{P}_k^{\text{in}} = \operatorname{recc} \mathcal{P}_k^{\text{out}}$ by Lemma 4.1.2, $\delta^H(\mathcal{P}_k^{\text{out}}, \mathcal{P}_k^{\text{in}}) < \infty$ holds correct; see [54, Lemma 6.3.15]. Since $\mathcal{P}_k^{\text{out}}$ is a polyhedron with at least one vertex, $d(\cdot, \mathcal{P}_k^{\text{in}})$ is a convex function (see, for instance, [57, Proposition 1.77]) and $\delta^H(\mathcal{P}_k^{\text{out}}, \mathcal{P}_k^{\text{in}}) < +\infty$, we have

$$\sup_{y \in \mathcal{P}_k^{\text{out}}} d(y, \mathcal{P}_k^{\text{in}}) = \max_{v \in \mathcal{V}_k} d(v, \mathcal{P}_k^{\text{in}})$$

from [58, Propositions 7-8]. The second equality can be shown similarly by noting that $\delta^H(\mathcal{P}_k^{\text{out}}, \mathcal{P}) \leq \delta^H(\mathcal{P}_k^{\text{out}}, \mathcal{P}_k^{\text{in}}) < \infty$ thanks to $\mathcal{P}_k^{\text{in}} \subseteq \mathcal{P} \subseteq \mathcal{P}_k^{\text{out}}$. $\square$

Theorem 4.1.4. *Under Assumptions 2.2.1 and 4.1.1, Algorithm 1 works correctly: if the algorithm terminates, then it returns a finite weak ϵ-solution to* (P).

Proof. For all $j \in \{1, \dots, J\}$, an optimal solution x^j of (WS(w^j)) exists since $\mathcal{X}$ is compact and $x \mapsto (w^j)^\mathsf{T}\Gamma(x)$ is continuous by the continuity of $\Gamma\colon X \to \mathbb{R}^q$ provided by Assumption 2.2.1. Moreover, x^j is a weak minimizer of (P) by Proposition 2.2.7. Thus, $\mathcal{X}_0$ consists of weak minimizers.

Since (P) is a bounded problem and C is a pointed cone, the set $\mathcal{P}_0^{\text{out}}$ contains no lines. Hence, $\mathcal{P}_0^{\text{out}}$ has at least one vertex [55, Corollary 18.5.3], that is, $\mathcal{V}_0 \neq \emptyset$. Moreover, as detailed after (4.1.1), $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$, hence we have $v \notin \operatorname{int}\mathcal{P}$ for any $v \in \mathcal{V}_0$. As it will be discussed below, for $k \geq 1$, $\mathcal{P}_k^{\text{out}}$ is constructed by intersecting $\mathcal{P}_0^{\text{out}}$ with supporting halfspaces of $\mathcal{P}$. This implies $\mathcal{V}_k \neq \emptyset$ and $\mathcal{V}_k \subseteq \mathbb{R}^q \setminus \operatorname{int}\mathcal{P}$ hold for any $k \geq 0$.

By Proposition 3.1.2, optimal solutions (x^v,z^v) and w^v to (P(v)) and (D(v)), respectively, exist. Moreover, by Proposition 3.1.6, x^v is a weak minimizer of (P). If $\|z^v\| > 0$, then $v \notin \mathcal{P}$, hence $w^v \neq 0$ by Lemma 3.1.5. By Proposition 3.1.7, $\mathcal{H}_k$ given by (4.1.2) is a supporting halfspace of $\mathcal{P}$ at $\Gamma(x^v)$. Then, since $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$ and $\mathcal{P}_{k+1}^{\text{out}} = \mathcal{P}_k^{\text{out}} \cap \mathcal{H}_k$, we have $\mathcal{P}_k^{\text{out}} \supseteq \mathcal{P}$ for all $k \geq 0$.

Assume that the algorithm stops after $\hat{k}$ iterations. Since $\mathcal{X}_{\hat{k}}$ is finite and consists of weak minimizers, to prove $\mathcal{X}_{\hat{k}}$ is a finite weak ϵ-solution of (P) as in Definition 2.2.5, it is sufficient to show that $\mathcal{P}_{\hat{k}}^{\text{in}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}$, where $\mathcal{P}_{\hat{k}}^{\text{in}} = \operatorname{conv}\Gamma(\mathcal{X}_{\hat{k}}) + C$.

Note that by the stopping condition, we have $\|z^v\| \leq \epsilon$, hence $x^v \in \mathcal{X}_{\hat{k}}$, for all $v \in \mathcal{V}_{\hat{k}}$. Moreover, since (x^v, z^v) is feasible for (P(v)), $v + z^v \in \{\Gamma(x^v)\} + C \subseteq \operatorname{conv}\Gamma(\mathcal{X}_{\hat{k}}) + C = \mathcal{P}_{\hat{k}}^{\text{in}}$. Then, by Lemma 4.1.3,

$$\delta^H(\mathcal{P}_{\hat{k}}^{\text{out}}, \mathcal{P}_{\hat{k}}^{\text{in}}) = \max_{v \in \mathcal{V}_{\hat{k}}} d(v, \mathcal{P}_{\hat{k}}^{\text{in}}) = \max_{v \in \mathcal{V}_{\hat{k}}} \inf_{u \in \mathcal{P}_{\hat{k}}^{\text{in}}} \|u - v\| \leq \max_{v \in \mathcal{V}_{\hat{k}}} \|z^v\| \leq \epsilon.$$

Consequently, $\mathcal{P}_{\hat{k}}^{\text{in}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}$ follows since $\mathcal{P}_{\hat{k}}^{\text{in}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}_{\hat{k}}^{\text{out}} \supseteq \mathcal{P}$. $\square$

4.2 A Modified Algorithm: Algorithm 2

In this section, we propose a modification of Algorithm 1, namely Algorithm 2, and prove that the algorithm works correctly. Later, we will also provide the finiteness result of the algorithm separately in Section 5.1.

The main feature of Algorithm 2 is that, in each iteration, it intersects the current outer approximation of $\mathcal{P}$ with a fixed halspace S and considers only the vertices of the intersection. As we describe next, the halfspace S is formed in such a way that $\mathcal{P} \cap S$ is compact and $\Gamma(\mathcal{X}) \subseteq \mathcal{P} \cap S$.

Algorithm 2 Modified Outer Approximation Algorithm for (P)

Compute an optimal solution x^j of (WS(w^j)) for each $j \in \{1, \dots, J\}$;
Set $k = 0, \bar{\mathcal{X}}_0 = \{x^1, \dots, x^J\}, \mathcal{V}^{\text{known}} = \emptyset$;
Store an H-representation of $\mathcal{P}_0^{\text{out}}$ according to (4.1.1);
Compute the set $\mathcal{V}_0$ of vertices of $\mathcal{P}_0^{\text{out}}$ from its H-representation;
$\bar{\mathcal{P}}_0^{\text{out}} = \mathcal{P}_0^{\text{out}}$;
for $v \in \mathcal{V}_0$ **do**
    Solve (P(v)) and (D(v)) to compute (x^v, z^v), w^v, and $d(v, \mathcal{P})$;
    $\mathcal{V}^{\text{known}} \leftarrow \mathcal{V}^{\text{known}} \cup \{v\}$;
    if $\|z^v\| > \epsilon$ **then**
        $\bar{\mathcal{P}}_0^{\text{out}} \leftarrow \bar{\mathcal{P}}_0^{\text{out}} \cap \{y \in \mathbb{R}^q \mid (w^v)^{\mathsf{T}} y \geq (w^v)^{\mathsf{T}} \Gamma(x^v)\}$;
    else
        $\bar{\mathcal{X}}_0 \leftarrow \bar{\mathcal{X}}_0 \cup \{x^v\}$;
    end if
end for
Compute β by Remark 4.2.1 and α by (4.2.3);
Store an H-representation of S according to (4.2.4);
repeat
    Stop $\leftarrow$ **true**;
    Compute the set $\bar{\mathcal{V}}_k$ of vertices of $\bar{\mathcal{P}}_k^{\text{out}} \cap S$ from its H-representation;
    for $v \in \bar{\mathcal{V}}_k$ **do**
        Follow lines 8-20 of Algorithm 1 using $\bar{\mathcal{P}}_k^{\text{out}}$, $\bar{\mathcal{P}}_{k+1}^{\text{out}}$, $\bar{\mathcal{X}}_k$, $\bar{\mathcal{X}}_{k+1}$[5];
    end for
until Stop
return $\begin{cases} \bar{\mathcal{X}}_k & : \text{A finite weak } \epsilon\text{-solution to (P)}; \\ \bar{\mathcal{P}}_k^{\text{out}} & : \text{An outer approximation of } \mathcal{P}. \end{cases}$

[5]More precisely, in Algorithm 1, we apply $\bar{\mathcal{P}}_{k+1}^{\text{out}} = \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$ in line 12, $\bar{\mathcal{X}}_{k+1} = \bar{\mathcal{X}}_k$ in line 13, $\bar{\mathcal{X}}_k \leftarrow \bar{\mathcal{X}} \cup \{x^v\}$ in line 18.

For the construction of S, let us define

$$\bar{w} := \frac{\sum_{j=1}^{J} w^j}{\|\sum_{j=1}^{J} w^j\|_*} \in \operatorname{int} C^+, \tag{4.2.1}$$

and fix $\beta \in \mathbb{R}$ such that

$$\beta \geq \sup_{x \in \mathcal{X}} \bar{w}^\mathsf{T} \Gamma(x). \tag{4.2.2}$$

The existence of such β is guaranteed by Assumption 2.2.1 since $\mathcal{X} \subseteq \mathbb{R}^n$ is a compact set and $x \mapsto \bar{w}^\mathsf{T}\Gamma(x)$ is a continuous function on X under this assumption. Note that β is an upper bound on the optimal value of an optimization problem that may fail to be convex, in general. We address some possible ways of computing β in Remark 4.2.1 below.

Remark 4.2.1. *Note that, $\bar{w}^\mathsf{T}\Gamma(x)$ is convex as $\Gamma : X \to \mathbb{R}^q$ is C-convex and $\bar{w} \in \operatorname{int} C^+$. Hence, $\sup_{x \in \mathcal{X}} \bar{w}^\mathsf{T}\Gamma(x)$ is a concave minimization problem over a compact convex set $\mathcal{X}$. Concave minimization is a well-known problem type in optimization for which numerous algorithms available in the literature to find a global optimal solution, see for instance [59, 60]. In our case, it is enough to run a single iteration of one such algorithm to find an upper bound β.*

In addition to $\bar{w}$ and β, we fix $\alpha \in \mathbb{R}$ such that

$$\alpha > \max_{v \in \mathcal{V}_0} (\bar{w}^\mathsf{T} v - \beta)^+ + \delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P}), \tag{4.2.3}$$

where $\mathcal{P}_0^{\text{out}}$ is the initial outer approximation used in Algorithm 1, $\mathcal{V}_0$ is the set of vertices of $\mathcal{P}_0^{\text{out}}$, and $a^+ := \max\{a, 0\}$ for $a \in \mathbb{R}$. By Lemma 4.1.3, we have $\delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P}) = \max_{v \in \mathcal{V}_0} d(v, \mathcal{P})$. Moreover, for each $v \in \mathcal{V}_0$, we have $d(v, \mathcal{P}) = \|z^v\|$, where (x^v, z^v) is an optimal solution to (P(v)). Hence, α can be computed once (P(v)) is solved for each $v \in \mathcal{V}_0$. Finally, using $\bar{w}, \alpha, \beta$, we define

$$S := \{y \in \mathbb{R}^q \mid \bar{w}^\mathsf{T} y \leq \beta + \alpha\}. \tag{4.2.4}$$

Figure 4.2 helps to understand the construction of such S by considering $\mathcal{P} \subseteq \mathbb{R}^2$.

Algorithm 2 starts with an initialization phase followed by a main loop that

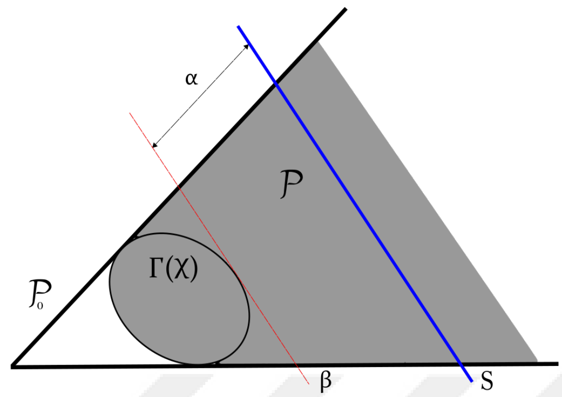


Figure 4.2: A visualization of Halfspace S, in case of $\mathcal{P} \subseteq \mathbb{R}^2$.

is similar to Algorithm 1. The initialization phase starts by constructing $\mathcal{P}_0^{\text{out}}$ according to (4.1.1) (lines 1-4 of Algorithm 2) and computing the set $\mathcal{V}_0$ of its vertices. For each $v \in \mathcal{V}_0$, the problems (P(v)) and (D(v)) are solved (line 7). The common optimal value $\|z^v\| = d(v, \mathcal{P})$ is used in the calculation of $\delta^H(\mathcal{P}_0^{\text{out}}, \mathcal{P})$ as described above. Moreover, these problems yield a supporting halfspace of $\mathcal{P}$ which is used to refine the outer approximation (line 10) if $\|z^v\|$ exceeds the predetermined error $\epsilon > 0$. Otherwise, the solution of (P(v)) is added to the set of weak minimizers (line 12). We denote by $\bar{\mathcal{P}}_0^{\text{out}}$ the refined outer approximation that is obtained at the end of the initialization phase.

The main loop of Algorithm 2 (lines 17-23) follows the same structure as Algorithm 1 except that it computes the set $\bar{\mathcal{V}}_k$ of all vertices of $\bar{\mathcal{P}}_k^{\text{out}} \cap S$ (as opposed to that of $\mathcal{P}_k^{\text{out}}$) at each iteration $k \geq 0$ (line 19). The algorithm terminates if all the vertices in $\bar{\mathcal{V}}_k$ are within ϵ distance to $\mathcal{P}$.

As opposed to Algorithm 1, in Algorithm 2, the norm-minimizing scalarization (P(v)) is not solved for a vertex v of $\bar{\mathcal{P}}_k^{\text{out}}$ if it is not in S. In Theorem 4.2.6, we will prove that Algorithm 2 works correctly even if it ignores such vertices. The next proposition, even though it is not directly used in the proof of Theorem 4.2.6, provides a geometric motivation for this result. In particular, it shows that if (P(v)) is solved for some $v \notin \operatorname{int} S$, then the supporting halfspace obtained as in Proposition 3.1.7 supports the upper image at a weakly C-minimal but not C-minimal element of the upper image.

Proposition 4.2.2. *Let v be a vertex of $\bar{\mathcal{P}}_k^{out}$ for some $k \geq 1$. If $v \notin \operatorname{int} S$, then $y^v = v + z^v \in \mathrm{wMin}_C(\mathcal{P}) \setminus \mathrm{Min}_C(\mathcal{P})$.*

Proof. Suppose that $v \notin \operatorname{int} S$. As v is a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$, we have $v \notin \operatorname{int} \mathcal{P}$. By Proposition 3.1.6, $y^v \in \mathrm{wMin}_C(\mathcal{P})$; in particular, $y^v \in \mathcal{P}$. Using Remark 2.2.2, there exist $\tilde{x} \in \mathcal{X}, \tilde{c} \in C$ such that $y^v = \Gamma(\tilde{x}) + \tilde{c}$. Next, we show that $\tilde{c} \neq 0$, which implies $y^v \notin \mathrm{Min}_C(\mathcal{P})$. From Hölder's inequality and (4.2.1), we have $\bar{w}^{\mathsf{T}}(v - y^v) \leq \|y^v - v\| \|\bar{w}\|_* = \|y^v - v\| = \|z^v\|$. Moreover, using $\bar{\mathcal{P}}_k^{\text{out}} \subseteq \mathcal{P}_0^{\text{out}}$, we obtain

$$\|z^v\| = d(v, \mathcal{P}) \leq \sup_{v' \in \bar{\mathcal{P}}_k^{\text{out}}} d(v', \mathcal{P}) \leq \sup_{v' \in \mathcal{P}_0^{\text{out}}} d(v', \mathcal{P}) = \delta^H(\mathcal{P}, \mathcal{P}_0^{\text{out}}) < \alpha,$$

where the last inequality follows from (4.2.3). Together, these imply $\bar{w}^{\mathsf{T}} y^v > \bar{w}^{\mathsf{T}} v - \alpha$. Using (4.2.4), (4.2.2) and $v \notin \operatorname{int} S$, we also have $\bar{w}^{\mathsf{T}} v - \alpha \geq \sup_{x \in \mathcal{X}} \bar{w}^{\mathsf{T}} \Gamma(x)$. Now, since $\bar{w}^{\mathsf{T}} y^v > \sup_{x \in \mathcal{X}} \bar{w}^{\mathsf{T}} \Gamma(x)$, it must be true that $y^v \notin \Gamma(\mathcal{X})$, which implies $\tilde{c} \neq 0$. □

It may happen that a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$, $k \geq 1$, falls outside S. We will illustrate this case in Remark 6.2.1 of Section 6.1.

With the following lemma and remark, we show that S satisfies the required properties mentioned at the beginning of this section. Note that Lemma 4.2.3 implies the compactness of $\mathcal{P} \cap S$ since $\mathcal{P} \subseteq \mathcal{P}_0^{\text{out}}$.

Lemma 4.2.3. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $\mathcal{P}_0^{out}$ and S be as in (4.1.1) and (4.2.4), respectively. Then, $\mathcal{P}_0^{out} \cap S$ is a compact set.*

Proof. Since $\mathcal{P}_0^{\text{out}}$ and S are closed sets, $\mathcal{P}_0^{\text{out}} \cap S$ is closed. Let $r \in \operatorname{recc}(\mathcal{P}_0^{\text{out}} \cap S)$ and $a \in \mathcal{P}_0^{\text{out}} \cap S$ be arbitrary. For every $\lambda \geq 0$, we have $a + \lambda r \in \mathcal{P}_0^{\text{out}} \cap S$. By the definition of $\mathcal{P}_0^{\text{out}}$, this implies

$$(w^j)^{\mathsf{T}}(a + \lambda r) \geq (w^j)^{\mathsf{T}} \Gamma(x^j) = \inf_{x \in \mathcal{X}} (w^j)^{\mathsf{T}} \Gamma(x), \tag{4.2.5}$$

for each $j \in \{1, \ldots, J\}$. On the other hand, by the definition of S in (4.2.4), we have

$$\bar{w}^\mathsf{T}(a + \lambda r) \leq \beta + \alpha. \tag{4.2.6}$$

Since (4.2.5) and (4.2.6) hold for every $\lambda \geq 0$, we have $(w^j)^\mathsf{T} r \geq 0$ for each $j \in \{1, \ldots, J\}$ and $\bar{w}^\mathsf{T} r = \sum_{j=1}^{J} (w^j)^\mathsf{T} r / \| \sum_{j=1}^{J} w^j \|_* \leq 0$, respectively. Together, these imply that

$$(w^j)^\mathsf{T} r = 0 \tag{4.2.7}$$

for each $j \in \{1, \ldots, J\}$. Recall that $J \geq q$ is implied by Assumptions 2.2.1 and 4.1.1. Consider the $q \times J$ matrix, say W, whose columns are the generating vectors of C^+. Since C^+ is solid, which follows from C being pointed, rank $W = q$, see for instance [61, Theorem 3.1]. Consider a $q \times q$ invertible submatrix $\tilde{W}$ of W. From (4.2.7), we have $\tilde{W}^\mathsf{T} r = 0 \in \mathbb{R}^q$, which implies $r = 0$. As $r \in \operatorname{recc}(\mathcal{P}_0^{\text{out}} \cap S)$ is chosen arbitrarily, $\mathcal{P}_0^{\text{out}} \cap S$ is bounded, hence compact. ☐

Remark 4.2.4. *It is clear by the definition of S that $\Gamma(\mathcal{X}) \subseteq S$. Let $k \geq 0$. Since $\mathcal{P} \subseteq \bar{\mathcal{P}}_k^{out}$, we also have $\Gamma(\mathcal{X}) \subseteq \bar{\mathcal{P}}_k^{out} \cap S$. Then, using Remark 2.2.2, we obtain $\mathcal{P} = \Gamma(\mathcal{X}) + C \subseteq (\bar{\mathcal{P}}_k^{out} \cap S) + C$. Also note that, if the algorithm terminates, then all the vertices in $\bar{\mathcal{V}}_k$ are within ϵ distance to $\mathcal{P}$.*

Remark 4.2.5. *In line 19 of Algorithm 2, if we compute $\mathcal{V}_k \cap S$ instead of $\bar{\mathcal{V}}_k$, that is, if we just ignore the vertices of $\bar{\mathcal{P}}_k^{out}$ which are outside S, then we cannot guarantee returning a finite weak ϵ-solution. This is because there may exist some vertices of $\bar{\mathcal{P}}_k^{out}$ that are out of S with distance to the upper image being larger than ϵ. Moreover, $\operatorname{conv}(\mathcal{V}_k \cap S) + C$ may not contain $\mathcal{P}$. This will be illustrated in Remark 6.2.1 of Section 6.1.*

Theorem 4.2.6. *Under Assumptions 2.2.1 and 4.1.1, Algorithm 2 works correctly: if the algorithm terminates, then it returns a finite weak ϵ-solution to* (P).

Proof. Similar to the proof of Theorem 4.1.4, the set $\bar{\mathcal{X}}_0$ is initialized by weak minimizers of (P), and $\mathcal{V}_0$ is nonempty. Moreover, for each $v \in \mathcal{V}_0$, optimal solutions (x^v, z^v) and w^v exist (Proposition 3.1.2); x^v is a weak minimizer (Proposition 3.1.6) and $\{y \in \mathbb{R}^q \mid (w^v)^\mathsf{T} y \geq (w^v)^\mathsf{T}\Gamma(x^v)\}$ is a supporting halfspace of $\mathcal{P}$ at $\Gamma(x^v)$ (Proposition 3.1.7). Hence, $\bar{\mathcal{P}}_0^{\text{out}} \supseteq \mathcal{P}$ is an outer approximation and $\bar{\mathcal{X}}_0$ consists of weak minimizers. By the definition of S, we have $\mathcal{V}_0 \subseteq S$. Hence, the set $\bar{\mathcal{V}}_0$ is nonempty.

Considering the main loop of the algorithm, we know by Proposition 3.1.2 that optimal solutions $(x^v$,$z^v)$ and w^v to (P(v)) and (D(v)), respectively, exist. Moreover, if $\|z^v\| \geq \epsilon$, then $w^v \neq 0$ by Lemma 3.1.5. Hence, by Proposition 3.1.7, $\mathcal{H}_k$ given by (4.1.2) is a supporting halfspace of $\mathcal{P}$. This implies $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \mathcal{P}$ for all $k \geq 0$. Since $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \mathcal{P} \supseteq \Gamma(\mathcal{X})$ and $S \supseteq \Gamma(\mathcal{X})$, see Remark 4.2.4, the set $\bar{\mathcal{P}}_k^{\text{out}} \cap S$ is nonempty. Moreover, as $\bar{\mathcal{P}}_k^{\text{out}} \subseteq \mathcal{P}_0^{\text{out}}$, it is true that $\bar{\mathcal{P}}_k^{\text{out}} \cap S$ is compact by Lemma 4.2.3. Then, $\bar{\mathcal{V}}_k$ is nonempty for all $k \geq 0$. Note that every vertex $v \in \bar{\mathcal{V}}_k$ satisfies $v \notin \operatorname{int} \mathcal{P}$. Indeed, since v is a vertex of $\bar{\mathcal{P}}_k^{\text{out}} \cap S$, it must be true that $v \in \operatorname{bd} \mathcal{H}_{\bar{k}}$ for some $\bar{k} \leq k$. The assertion follows since $\operatorname{bd} \mathcal{H}_{\bar{k}}$ is a supporting hyperplane of $\mathcal{P}$. Then, by Proposition 3.1.6, x^v is a weak minimizer of (P).

Assume that the algorithm stops after $\hat{k}$ iterations. Clearly, $\bar{\mathcal{X}}_{\hat{k}}$ is finite and consists of weak minimizers. By Definition 2.2.5, it remains to show that $\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}$ holds, where $\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} := \operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}}) + C$. By the stopping condition, for every $v \in \bar{\mathcal{V}}_{\hat{k}}$, we have $\|z^v\| \leq \epsilon$, hence $x^v \in \bar{\mathcal{X}}_{\hat{k}}$. Moreover, since (x^v, z^v) is feasible for (P(v)),

$$v + z^v \in \{\Gamma(x^v)\} + C \subseteq \operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}}) + C = \bar{\mathcal{P}}_{\hat{k}}^{\text{in}}. \tag{4.2.8}$$

By Remark 4.2.4, it is true that $\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} \subseteq \mathcal{P} \subseteq (\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S) + C$. Moreover, as $\operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}})$ and $\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S$ are compact sets, $\operatorname{recc} \bar{\mathcal{P}}_{\hat{k}}^{\text{in}} = \operatorname{recc}((\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S) + C) = C$. By repeating the arguments in the proof of Lemma 4.1.3, it is easy to check that

$$\delta^H(\bar{\mathcal{P}}_{\hat{k}}^{\text{in}}, (\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S) + C) = \max_{v \in \bar{\mathcal{V}}_{\hat{k}}^C} d(v, \bar{\mathcal{P}}_{\hat{k}}^{\text{in}}),$$

where $\bar{\mathcal{V}}_{\hat{k}}^C$ denotes the set of all vertices of $(\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S) + C$. By Lemma 2.1.1, every vertex v of $(\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S) + C$ is also a vertex of $\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} \cap S$, that is, $\bar{\mathcal{V}}_{\hat{k}}^C \subseteq \bar{\mathcal{V}}_{\hat{k}}$. Then, we obtain

$$\delta^H(\bar{\mathcal{P}}^{\text{in}}_{\hat{k}}, (\bar{\mathcal{P}}^{\text{out}}_{\hat{k}} \cap S) + C) \leq \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} d(v, \bar{\mathcal{P}}^{\text{in}}_{\hat{k}}) = \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} \inf_{u \in \bar{\mathcal{P}}^{\text{in}}_{\hat{k}}} \|u - v\| \leq \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} \|z^v\| \leq \epsilon,$$

where the penultimate inequality follows by (4.2.8). Since

$$\bar{\mathcal{P}}^{\text{in}}_{\hat{k}} + \mathbb{B}_\epsilon \supseteq (\bar{\mathcal{P}}^{\text{out}}_{\hat{k}} \cap S) + C \supseteq \mathcal{P}, \tag{4.2.9}$$

$\bar{\mathcal{P}}^{\text{in}}_{\hat{k}} + \mathbb{B}_\epsilon \supseteq \mathcal{P}$ follows. □

4.3 An Algorithm based on modified scalarization: Algorithm 3

In this section, we propose an algorithm (Algorithm 3) whose finiteness and convergence rate will be considered in Section 5.2 and in Section 5.3, respectively.

The Algorithm 3 has a similar design as of Algorithm 2, except that it computes the outer approximation of $\mathcal{P} \cap S$, instead of $\mathcal{P}$. Moreover, Algorithm 3 is based on solving modified norm-minimizing scalarizations ($\text{P}_m(v)$) and its dual ($\text{D}_m(v)$) iteratively. It has two phases. The first phase is same as the initialization phase (line 1-16) of Algorithm 2, which uses the norm-minimizing scalarizations ($\text{P}(v)$) and its dual ($\text{D}(v)$) to find S. Here, we extend the phase with an additional step of finding $\bar{\mathcal{P}}^{\text{out}}_0 := \mathcal{P}^{\text{out}}_0 \cap S$, where S is given by (4.2.4).

In the second phase of the algorithm, the set $\bar{\mathcal{V}}_k$ of all vertices of $\bar{\mathcal{P}}^{\text{out}}_k$ is computed, $k \geq 0$. Then, for all $v \in \bar{\mathcal{V}}_k$, optimal solutions (x^v, z^v) and (w^v, λ^v) are obtained by solving the modified norm-minimizing scalarization ($\text{P}_m(v)$) and its dual ($\text{D}_m(v)$), respectively. If, for $v \in \bar{\mathcal{V}}_k$, $\|z^v\| \leq \epsilon$, where ϵ is a predetermined approximation error, the set $\mathcal{X}_k$ of weak minimizers is updated with x^v, see Proposition 3.2.4. In order to find the farthest vertex to the upper image in $\bar{\mathcal{V}}_k$, we compute

$$v^{k*} \in \arg\max\{\|z^v\| \mid v \in \bar{\mathcal{V}}_k\}. \tag{4.3.1}$$

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

$$\mathcal{H}_k := \{y \in \mathbb{R}^q \mid (\tilde{w}^{v^{k*}})^{\mathsf{T}} y \geq (\tilde{w}^{v^{k*}})^{\mathsf{T}} y^{v^{k*}}\}, \tag{4.3.2}$$

is determined, see Proposition 3.2.5, and the current outer approximation is updated as $\bar{\mathcal{P}}^{\text{out}}_{k+1} = \bar{\mathcal{P}}^{\text{out}}_k \cap \mathcal{H}_k$. Otherwise, the algorithm terminates.

Algorithm 3 Outer Approximation Algorithm based on Modified Scalarization

Compute an optimal solution x^j of (WS(w^j)) for each $j \in \{1, \dots, J\}$;
Set $k = 0, \bar{\mathcal{X}}_0 = \{x^1, \dots, x^J\}, \mathcal{V}^{\text{known}} = \emptyset, \mathcal{V}^{\text{known2}} = \emptyset$;
Follow lines 3-16 of Algorithm 2;
$\bar{\mathcal{P}}^{\text{out}}_0 = \mathcal{P}^{\text{out}}_0 \cap S$;
repeat
  Stop $\leftarrow$ **true**;
  Compute the set $\bar{\mathcal{V}}_k$ of vertices of $\bar{\mathcal{P}}^{\text{out}}_k$ from its H-representation;
  for $v \in \bar{\mathcal{V}}_k \setminus \mathcal{V}^{\text{known}}$ **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 $\|z^v\| \leq \epsilon$ **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 $v^{k*} \in \arg\max\{\|z^v\| \mid v \in \bar{\mathcal{V}}_k\}$
  if $\|z^{v^{k*}}\| > \epsilon$ **then**
    $\bar{\mathcal{P}}^{\text{out}}_{k+1} = \bar{\mathcal{P}}^{\text{out}}_k \cap \mathcal{H}_k$;
    $\bar{\mathcal{X}}_{k+1} = \bar{\mathcal{X}}_k$;
    $k \leftarrow k + 1$;
    Stop $\leftarrow$ **false**;
  end if
until Stop
return $\begin{cases} \bar{\mathcal{X}}_k & : \text{A finite weak } \epsilon\text{-solution to (P);} \\ \bar{\mathcal{P}}^{\text{out}}_k & : \text{An outer approximation of } \mathcal{P} \cap S. \end{cases}$

Remark 4.3.1. *By definition of S, $\Gamma(\mathcal{X}) \subseteq S$. Since $\bar{\mathcal{P}}^{out}_0 \supseteq \mathcal{P} \cap S$ and $\bar{\mathcal{P}}^{out}_{k+1} = \bar{\mathcal{P}}^{out}_k \cap \mathcal{H}_k$, we have $\bar{\mathcal{P}}^{out}_k \supseteq \mathcal{P} \cap S \supseteq \Gamma(\mathcal{X})$ for all $k \geq 0$. Then, using Remark 2.2.2, we obtain $\mathcal{P} = \Gamma(\mathcal{X}) + C \subseteq \bar{\mathcal{P}}^{out}_k + C$. Hence, $\bar{\mathcal{P}}^{out}_k + C$ gives an outer approximation of $\mathcal{P}$.*

Remark 4.3.2. *We have, $\mathcal{P} = (\mathcal{P} \cap S) + C$. Indeed, $\mathcal{P} = \Gamma(\mathcal{X}) + C \subseteq (\mathcal{P} \cap S) + C$, as $\Gamma(\mathcal{X}) \subseteq S$. For the other inclusion, $(\mathcal{P} \cap S) + C \subseteq \mathcal{P} + C = \mathcal{P}$.*

Lemma 4.3.3. *Let $k \geq 1$. For any vertex $v \in \bar{\mathcal{V}}_k$, $v \notin \operatorname{int} \mathcal{P}$.*

Proof. Clearly, $\mathcal{V}_0 \subseteq \mathbb{R}^q \setminus \operatorname{int} \mathcal{P}$, as $\mathcal{P}_0^{\text{out}} \supseteq \mathcal{P}$. Then, for any $v \in \bar{\mathcal{V}}_0$, we have $v \notin \operatorname{int} \mathcal{P}$, where $\bar{\mathcal{V}}_0$ is the set of vertices of $\bar{\mathcal{P}}_0^{\text{out}} = \mathcal{P}_0^{\text{out}} \cap S \supseteq \mathcal{P} \cap S$.

Assume contrary for a contradiction. Let k be first iteration number in which there exists $v \in \bar{\mathcal{V}}_k$ such that $v \in \operatorname{int} \mathcal{P}$. Note that, $\bar{\mathcal{V}}_k \subseteq S$ by the construction of the algorithm (as $\bar{\mathcal{P}}_0^{\text{out}} = \mathcal{P}_0^{\text{out}} \cap S$ and $\bar{\mathcal{P}}_{k+1}^{\text{out}} = \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$, $k \geq 0$). 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 $\bar{k} \leq k$. We know that $\operatorname{bd} \mathcal{H}_{\bar{k}}$ is a supporting hyperplane of $\mathcal{P} \cap S$ which is closed. Hence, $v \notin \operatorname{int}(\mathcal{P} \cap S) = \operatorname{int} \mathcal{P} \cap \operatorname{int} S$. Then, v is on the boundary of S, as $\bar{\mathcal{V}}_k \subseteq S$. On the other hand, since $v \in \operatorname{int} \mathcal{P}$, there exists $\delta > 0$ such that $B(v, \delta) \subseteq \operatorname{int} \mathcal{P}$. Clearly, $B(v, \delta) \cap \operatorname{bd} S \neq \emptyset$. Fix $v' \in \operatorname{bd} S$, where $v' \neq v$. Define,

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

Then, $v \pm s$ is an affine combination of v and v', where v and v' both are in $\operatorname{bd} S$. Since $\operatorname{bd} S$ is an affine set, we get $v \pm s \in \operatorname{bd} S$. This implies,

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

On the other hand, clearly v can be written as a convex combination of $v \pm s \subseteq \bar{\mathcal{P}}_k^{\text{out}}$, that is, $v = \frac{1}{2}(v + s) + \frac{1}{2}(v - s)$. This is a contradiction, as v is a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$. Hence, $v \notin \operatorname{int} \mathcal{P}$. □

Now, similar to Theorem 4.2.6, we show that Algorithm 3 returns a finite weak ϵ-solution to (P) if it terminates.

Theorem 4.3.4. *Under Assumptions 2.2.1 and 4.1.1, Algorithm 3 works correctly: if the algorithm terminates, then it returns a finite weak ϵ-solution to (P).*

Proof. Similar to the proof of Theorem 4.2.6, the set $\bar{\mathcal{X}}_0$ in line 11 consists of weak minimizers and the set $\bar{\mathcal{V}}_0$ is nonempty.

Considering the main loop of the algorithm, we know by Proposition 3.2.2 that optimal solutions (x^v, z^v) and (w^v, λ^v) to ($\mathrm{P}_m(v)$) and ($\mathrm{D}_m(v)$), respectively, exist. By Proposition 3.2.4, x^v is a weak minimizer of (P). Moreover, if $\|z^v\| \geq \epsilon$, then $\tilde{w} \neq 0$ by Lemma 3.2.3. Hence, by Proposition 3.2.5, $\mathcal{H}_k$ given by (4.3.2) is a supporting halfspace of $\mathcal{P} \cap S$. Then, since $\bar{\mathcal{P}}_0^{\text{out}} \supseteq \mathcal{P} \cap S$ and $\bar{\mathcal{P}}_k^{\text{out}}$ $(k \geq 1)$ is constructed by intersecting $\bar{\mathcal{P}}_0^{\text{out}}$ with supporting halfspaces $\mathcal{H}_k$ of $\mathcal{P} \cap S$, we have $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \mathcal{P} \cap S$ for all $k \geq 0$. Also note that, $\bar{\mathcal{P}}_{k+1}^{\text{out}} = \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k$ implies $\bar{\mathcal{V}}_k \neq \emptyset$ hold for any $k \geq 0$.

By Lemma 4.3.3, $v \notin \operatorname{int} \mathcal{P}$. Hence, by Proposition 3.2.4, x^v is a weak minimizer of (P).

If $v \in \operatorname{bd} \mathcal{P}$, there may exist optimal solution of ($\mathrm{D}_m(v)$) such that $\tilde{w} = 0$. However, in this case we have $\|z^v\| = 0$ from Lemma 3.2.3. Then, line 32 ensures that the corresponding halfspace will not be used to update an outer approximation.

Assume that the algorithm stops after $\hat{k}$ iterations. Clearly, $\bar{\mathcal{X}}_{\hat{k}}$ is finite and consists of weak minimizers. By Definition 2.2.5, it remains to show that

$$\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} + B(0, \epsilon) \supseteq \mathcal{P} \tag{4.3.3}$$

holds, where $\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} := \operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}}) + C$.

By the stopping condition, for every $v \in \bar{\mathcal{V}}_{\hat{k}}$, we have $\|z^v\| \leq \epsilon$, hence $x^v \in \bar{\mathcal{X}}_{\hat{k}}$. Note that for every vertex v of $\bar{\mathcal{P}}_{\hat{k}}^{\text{out}}$, $y^v = v + z^v \in \{\Gamma(x^v)\} + C$, since (x^v, z^v) is feasible for ($\mathrm{P}_m(v)$). Hence,

$$y^v \in \{\Gamma(x^v)\} + C \subseteq \operatorname{conv}(\Gamma(\bar{\mathcal{X}}_{\hat{k}})) + C = \bar{\mathcal{P}}_{\hat{k}}^{\text{in}}. \tag{4.3.4}$$

By Remark 4.3.1, it is true that $\bar{\mathcal{P}}_{\hat{k}}^{\text{in}} \subseteq \mathcal{P} \subseteq \bar{\mathcal{P}}_{\hat{k}}^{\text{out}} + C$. Moreover, as $\operatorname{conv} \Gamma(\bar{\mathcal{X}}_{\hat{k}})$ and $\bar{\mathcal{P}}_{\hat{k}}^{\text{out}}$ are compact sets, $\operatorname{recc} \bar{\mathcal{P}}_{\hat{k}}^{\text{in}} = \operatorname{recc}(\bar{\mathcal{P}}_{\hat{k}}^{\text{out}} + C) = C$. By using the similar arguments as in the proof of Lemma 4.1.3, we obtain

$$\delta^H(\bar{\mathcal{P}}^{\text{in}}_{\hat{k}}, \bar{\mathcal{P}}^{\text{out}}_{\hat{k}} + C) = \max_{v \in \bar{\mathcal{V}}^C_{\hat{k}}} d(v, \bar{\mathcal{P}}^{\text{in}}_{\hat{k}}),$$

where $\bar{\mathcal{V}}^C_{\hat{k}}$ denotes the set of all vertices of $\bar{\mathcal{P}}^{\text{out}}_{\hat{k}} + C$. By Lemma 2.1.1, $\bar{\mathcal{V}}^C_{\hat{k}} \subseteq \bar{\mathcal{V}}_{\hat{k}}$, that is, every vertex v of $\bar{\mathcal{P}}^{\text{out}}_{\hat{k}} + C$ is also a vertex of $\bar{\mathcal{P}}^{\text{out}}_{\hat{k}}$. Then, we obtain

$$\delta^H(\bar{\mathcal{P}}^{\text{in}}_{\hat{k}}, \bar{\mathcal{P}}^{\text{out}}_{\hat{k}} + C) \leq \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} d(v, \bar{\mathcal{P}}^{\text{in}}_{\hat{k}}) = \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} \inf_{u \in \bar{\mathcal{P}}^{\text{in}}_{\hat{k}}} \|u - v\| \leq \max_{v \in \bar{\mathcal{V}}_{\hat{k}}} \|z^v\| \leq \epsilon,$$

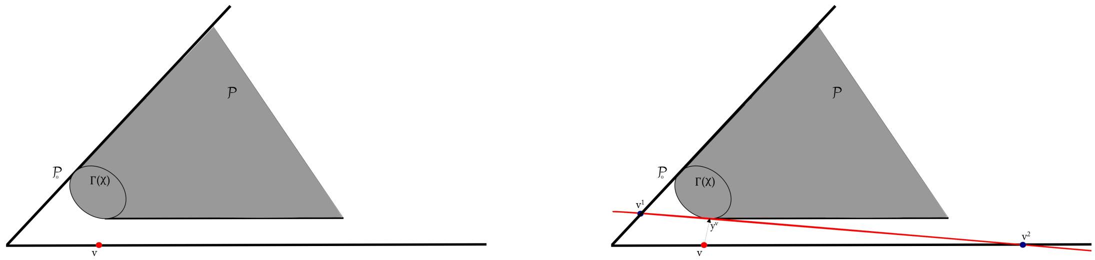
where the penultimate inequality follows by Equation (4.3.4). Consequently, (4.3.3) follows since $\bar{\mathcal{P}}^{\text{in}}_{\hat{k}} + B(0, \epsilon) \supseteq \bar{\mathcal{P}}^{\text{out}}_{\hat{k}} + C \supseteq \mathcal{P}$. □

Chapter 5

Convergence Analysis

In this chapter, we examine the Algorithms 2 and 3 further. We show that for any norm used in the scalarizations the algorithms are finite. Then, the rate of convergence is determined for Algorithm 3 using some results from the literature [13, 37, 43].

Let us first explain the limitation of Algorithm 1 which makes the convergence analysis difficult. An intuition is given, using Figure 5.1, by conidering $\mathcal{P} \subseteq \mathbb{R}^2$ and an outer approximation $\mathcal{P}_k^{\text{out}}$, for some $k \geq 0$, for illustration purposes. Note that, such scenario may not be encountered in two dimensional objective space, but cannot be ruled out in three or higher dimensional. Consider solving the scalarization for a point $v \in \mathbb{R}^2$ (Figure 5.1a). The resulting supporting halfspace may generate a vertex too far away from $\Gamma(\mathcal{X})$ (Figure 5.1b). This shows that the region, containing the vertices obtained throughout the iterations of the algorithm, is not compact which contributes to problems in proving the finiteness. The reader is referred to the finiteness result of Algorithm 2 (proof of Theorem 5.1.3) for further insight. Note that, the same limitation and its repercussion hold for the similar CVOP algorithms available in the literature, for instance the Primal algorithm in [1].



(a) Picked a point v at some iteration of Algorithm 1.

(b) Supporting halfspace, corresponding to point v, may generate a vertex too far away from $\Gamma(\mathcal{X})$.

Figure 5.1: Illustration of the problem, while proving the finiteness of Algorithm 1, that the region which may contain the vertices, obtained throughout the iterations of the algorithm, is not compact.

5.1 Finiteness of Algorithm 2

In this section, we show that Algorithm 2 terminates after finitely many iterations. The idea is to use compactness of a set. In particular, we show that through the iterations of the algorithm, a sequence of non-overlapping subsets of fixed volume, contained in a compact set, can be constructed. We provide two technical results before proceeding to the main theorem.

Lemma 5.1.1. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $v \notin \mathcal{P}$ and $\mathcal{H}$ be the halfspace defined by Proposition 3.1.7. If $\|z^v\| \geq \epsilon$, then $B \cap \mathcal{H} = \emptyset$, where*

$$B := \left\{ y \in \{v\} + C \mid \|y - v\| \leq \frac{\epsilon}{2} \right\}.$$

Proof. Consider (P(v)) and its Lagrange dual (D(v)). The arbitrarily fixed dual optimal solution w^v satisfies the first order condition with respect to z, which can be expressed as $w^v \in \partial\|z^v\|$. Note that the subdifferential of $\|\cdot\|$ at z^v has the variational characterization

$$\partial\|z^v\| = \left\{ w \in \mathbb{R}^q \mid \sup\{(w')^\mathsf{T} z^v \mid \|w'\|_* \leq 1\} = w^\mathsf{T} z^v,\ \|w\|_* \leq 1 \right\},$$

which follows by applying [55, Theorem 23.5]. Since the dual norm of $\|\cdot\|_*$ is $\|\cdot\|$,

$$w^v \in \partial\|z^v\| = \{w \in \mathbb{R}^q \mid \|z^v\| = w^\mathsf{T} z^v,\ \|w\|_* \leq 1\}. \tag{5.1.1}$$

Let $\bar{y} \in \mathcal{H}$ be arbitrary. From the definition of $\mathcal{H}$ and (5.1.1), we have $(w^v)^\mathsf{T}\bar{y} \geq (w^v)^\mathsf{T}(v + z^v) = (w^v)^\mathsf{T}v + \|z^v\|$. Equivalently, $(w^v)^\mathsf{T}(\bar{y} - v) \geq \|z^v\|$. On the other hand, from Hölder's inequality and (5.1.1), we have $|(w^v)^\mathsf{T}(\bar{y} - v)| \leq \|w^v\|_*\|\bar{y} - v\| \leq \|\bar{y}-v\|$. If $\|z^v\| \geq \epsilon$, then from the last two inequalities, we obtain $\|\bar{y}-v\| \geq |(w^v)^\mathsf{T}(\bar{y} - v)| \geq \|z^v\| \geq \epsilon$. Therefore, $\bar{y} \notin B$, which implies $B \cap \mathcal{H} = \emptyset$. □

We know that the set of vertices $\bar{\mathcal{V}}_k$ of $\bar{\mathcal{P}}_k^{\text{out}} \cap S$, $k \geq 0$, is contained in the compact set $\mathcal{P}_0 \cap S$. In the following lemma, we enlarge the compact set such that it would also contain every C-minimal element y^v of $\mathcal{P}$ corresponding to a vertex $v \in \bar{\mathcal{V}}_k, k \geq 0$. This will help to prove the finiteness of the algorithm.

Lemma 5.1.2. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $k \geq 0$, v be a vertex of $\bar{\mathcal{P}}_k^{out}$, S be as in (4.2.4); and define*

$$S_2 := \{y \in \mathbb{R}^q \mid \bar{w}^\mathsf{T} y \leq \beta + 2\alpha\},$$

where $\bar{w}, \beta$ and α are defined by (4.2.1), (4.2.2) and (4.2.3), respectively. If $v \in S$, then $v + z^v \in \operatorname{int} S_2$.

Proof. Let $\tilde{\mathcal{V}}_k$ denote the set of all vertices of $\bar{\mathcal{P}}_k^{\text{out}}$. It is given that $v \in \tilde{\mathcal{V}}_k$. Using Remark 3.1.1 and the arguments in the proof of Lemma 4.1.3, we obtain $\delta^H(\mathcal{P}, \bar{\mathcal{P}}_k^{\text{out}}) = \max_{\tilde{v}\in\tilde{\mathcal{V}}_k} d(\tilde{v}, \mathcal{P}) \geq d(v, \mathcal{P}) = \|z^v\|$. From (4.2.3) and the inclusion $\mathcal{P}_0^{\text{out}} \supseteq \bar{\mathcal{P}}_k^{\text{out}}$, we have

$$\delta^H(\mathcal{P}, \bar{\mathcal{P}}_k^{\text{out}}) \leq \delta^H(\mathcal{P}, \mathcal{P}_0^{\text{out}}) < \alpha,$$

which implies $\|z^v\| < \alpha$. Then, using Hölder's inequality together with $\|\bar{w}\|_* = 1$, we obtain $\bar{w}^\mathsf{T} z^v \leq \|z^v\|\|\bar{w}\|_* = \|z^v\| < \alpha$. On the other hand, $v \in S$ implies that $\bar{w}^\mathsf{T} v \leq \beta + \alpha$. Then, $v + z^v \in \operatorname{int} S_2$ follows since $\bar{w}^\mathsf{T}(v + z^v) < \beta + 2\alpha$. □

Theorem 5.1.3. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Algorithm 2 terminates after a finite number of iterations.*

Proof. By the construction of the algorithm, the number of vertices of $\bar{\mathcal{P}}_k^{\text{out}}$ is finite for every $k \geq 0$. It is sufficient to prove that there exists $K \geq 0$ such that for every vertex $v \in \bar{\mathcal{V}}_K$ of $\bar{\mathcal{P}}_K^{\text{out}} \cap S$, we have $\|z^v\| \leq \epsilon$. To get a contradiction, assume that for every $k \geq 0$, there exists a vertex $v^k \in \bar{\mathcal{V}}_k$ such that $\|z^{v^k}\| > \epsilon$. For convenience, an optimal solution (x^{v^k}, z^{v^k}) of $(\mathrm{P}(v^k))$ is denoted by (x^k, z^k) throughout the rest of the proof. Let S_2 be as in Lemma 5.1.2. Then, following similar arguments as presented in the proof of Lemma 4.2.3, one can show that $\mathcal{P}_0^{\text{out}} \cap S_2$ is compact.

Let $k \geq 0$ be arbitrary. We define $B^k := \left\{y \in \{v^k\} + C \mid \|y - v^k\| \leq \frac{\epsilon}{2}\right\}$. Note that B^k is a compact set in $\mathbb{R}^q$ and, as C is solid by Assumption 2.2.1, B^k has nonzero volume, which is free of the choice of k. Next, we show that $B^k \subseteq \mathcal{P}_0^{\text{out}} \cap S_2$. Repeating the arguments in the proof of Lemma 4.1.2, it can be shown that $\operatorname{recc} \bar{\mathcal{P}}_k^{\text{out}} = C$. Then, since $v^k \in \bar{\mathcal{P}}_k^{\text{out}}$, it holds

$$\{v^k\} + C \subseteq \bar{\mathcal{P}}_k^{\text{out}} \subseteq \mathcal{P}_0^{\text{out}}. \tag{5.1.2}$$

Hence, $B^k \subseteq \mathcal{P}_0^{\text{out}}$. To see $B^k \subseteq S_2$, let $y \in B^k$. From Hölder's inequality and (4.2.1),

$$\bar{w}^{\mathsf{T}}(y - v^k) \leq \|y - v^k\| \|\bar{w}\|_* = \|y - v^k\| \leq \frac{\epsilon}{2}. \tag{5.1.3}$$

As there exists $v^0 \in \bar{\mathcal{V}}_0$ with $\|z^0\| > \epsilon$, it holds true that $\delta^H(\mathcal{P}, \mathcal{P}_0^{\text{out}}) > \epsilon$. From (4.2.3), it follows that $\alpha > \epsilon$. Then, from (5.1.3), we obtain $\bar{w}^{\mathsf{T}}(y - v^k) \leq \frac{\epsilon}{2} < \alpha$. Since $v^k \in S$, this implies $\bar{w}^{\mathsf{T}} y < \bar{w}^{\mathsf{T}} v^k + \alpha \leq \beta + 2\alpha$, hence $y \in S_2$.

Next, we prove that $B^i \cap B^j = \emptyset$ for every $i, j \geq 0$ with $i \neq j$. Assume without loss of generality that $i < j$. Thus, $\bar{\mathcal{P}}_j^{\text{out}} \subseteq \bar{\mathcal{P}}_{i+1}^{\text{out}}$. From Lemma 5.1.1, we have $B^i \cap \mathcal{H}_i = \emptyset$, where $\mathcal{H}_i$ is the supporting halfspace at $\Gamma(x^i)$ as obtained in Proposition 3.1.7. This implies $B^i \cap \bar{\mathcal{P}}_j^{\text{out}} = \emptyset$ as $\bar{\mathcal{P}}_j^{\text{out}} \subseteq \bar{\mathcal{P}}_{i+1}^{\text{out}} = \bar{\mathcal{P}}_i^{\text{out}} \cap \mathcal{H}_i \subseteq \mathcal{H}_i$. On the other hand, we have $B^j \subseteq \bar{\mathcal{P}}_j^{\text{out}}$ from (5.1.2). Thus, $B^i \cap B^j = \emptyset$. These imply that there is an infinite number of non-overlapping sets, having strictly positive fixed volume, contained in a compact set $\mathcal{P}_0^{\text{out}} \cap S_2$, a contradiction. $\square$

We conclude this section with a convergence result regarding the Hausdorff distance between the upper image and its polyhedral approximations.

Corollary 5.1.4. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Let $\epsilon = 0$ and Algorithm 2 be modified by introducing a cutting order based on selecting a farthest away vertex (instead of an arbitrary vertex) in line 20. Then,*

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

where $\bar{\mathcal{P}}_k^{in} := \operatorname{conv}\Gamma(\bar{\mathcal{X}}_k) + C$, the sets $\bar{\mathcal{X}}_k, \bar{\mathcal{P}}_k^{out}$ are as described in Algorithm 2, and S is given by (4.2.4).

Proof. Note that Algorithm 2 is finite by Theorem 5.1.3 for an arbitrary vertex selection rule, hence, also when a farthest away vertex is selected in line 20. Therefore, given $\varepsilon > 0$, there exists $K(\varepsilon) \in \mathbb{N}$ such that the set $\bar{\mathcal{X}}_{K(\varepsilon)}$ is a finite weak ε-solution as in Definition 2.2.5 and $\delta^H(\bar{\mathcal{P}}^{\text{in}}_{K(\varepsilon)}, \mathcal{P}) \leq \varepsilon$ by Remark 2.2.6. Let us consider the modified algorithm (with $\epsilon = 0$). If $\delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}) = 0$ for some $k \in \mathbb{N}$, then it is clear that $\lim_{k\to\infty} \delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}) = 0$. Suppose that $\delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}) > 0$ for every $k \in \mathbb{N}$. With the farthest away vertex selection rule, when we run the algorithm with $\epsilon = 0$ and $\epsilon = \varepsilon > 0$, the two work in the same way until the one with $\epsilon = \varepsilon$ stops. Hence, they find the same inner approximation $\bar{\mathcal{P}}^{\text{in}}_{K(\varepsilon)}$ at step $K(\varepsilon)$. Let $k_0 := K(\varepsilon)$. By an induction argument, for every $n \in \mathbb{N}$, the inequality $\delta^H(\bar{\mathcal{P}}^{\text{in}}_{k_n}, \mathcal{P}) \leq \frac{\varepsilon}{n}$ will be satisfied by the algorithm with $\epsilon = 0$ for some $k_n > k_{n-1}$. Hence, $\lim_{n\to\infty} \delta^H(\bar{\mathcal{P}}^{\text{in}}_{k_n}, \mathcal{P}) = 0$, which implies that $\lim_{k\to\infty} \delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}) = 0$ by the monotonicity of $(\delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}))_{k\in\mathbb{N}}$.

Moreover, similar to the discussion in the proof of Theorem 4.2.6, see (4.2.9), it can be shown that $\delta^H((\bar{\mathcal{P}}_k^{\text{out}} \cap S) + C, \mathcal{P}) \leq \delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P})$ holds for each $k \in \mathbb{N}$. Hence, $\lim_{k\to\infty} \delta^H((\bar{\mathcal{P}}_k^{\text{out}} \cap S) + C, \mathcal{P}) = 0$. $\square$

Remark 5.1.5. *In Corollary 5.1.4, choosing the farthest away vertex in each iteration is critical. Indeed, without this rule, the algorithm run with $\epsilon = 0$ and*

$\epsilon = \varepsilon > 0$ may not work in the same way due to line 11 in Algorithm 2. In such a case, we may not use Theorem 5.1.3 to argue that the algorithm with $\epsilon = 0$ satisfies $\delta^H(\bar{\mathcal{P}}_k^{in}, \mathcal{P}) \leq \varepsilon$ for some $k \in \mathbb{N}$. Indeed, it might happen in case of non-polyhedral $\mathcal{P}$ that the algorithm keeps updating the outer approximation by focusing only on one part of $\mathcal{P}$. Thus, the Hausdorff distance at the limit may not be zero.

5.2 Finiteness of Algorithm 3

In this section, we show the finiteness of Algorithm 3. The idea is similar to Section 5.1 where we have constructed a sequence of non-overlapping subsets of fixed volume inside a compact set. The finiteness result is based on the following two lemmas.

Lemma 5.2.1. *Suppose that Assumptions 2.2.1 holds. Let $v \notin \mathcal{P}$ and $\mathcal{H}'$ be the corresponding halfspace to v defined by*

$$\mathcal{H}' := \left\{ y \in \mathbb{R}^q \mid (\tilde{w}^v)^\mathsf{T} y \geq (\tilde{w}^v)^\mathsf{T} y^v - \frac{\epsilon}{2} \right\}. \tag{5.2.1}$$

If $\|z^v\| \geq \epsilon$, then $B(v, \frac{\epsilon}{4}) \cap \mathcal{H}' = \emptyset$.

Proof. Consider ($\mathrm{P}_m(v)$) and its Lagrange dual ($\mathrm{D}_m(v)$). Then, by using the similar arguments as in the proof of Lemma 5.1.1, we get

$$\tilde{w}^v = w^v - \lambda^v \bar{w} \in \partial\|z^v\| = \{w \in \mathbb{R}^q \mid \|z^v\| = w^\mathsf{T} z^v,\ \|w\|_* \leq 1\}, \tag{5.2.2}$$

Let $\bar{y} \in \mathcal{H}'$ be arbitrary. From the definition of $\mathcal{H}'$ and (5.2.2), we have $(\tilde{w}^v)^\mathsf{T}\bar{y} \geq (\tilde{w}^v)^\mathsf{T}(v + z^v) - \frac{\epsilon}{2} = (\tilde{w}^v)^\mathsf{T} v + \|z^v\| - \frac{\epsilon}{2}$. Equivalently,

$$(\tilde{w}^v)^\mathsf{T}(\bar{y} - v) \geq \|z^v\| - \frac{\epsilon}{2}. \tag{5.2.3}$$

On the other hand, from Hölder's inequality and (5.1.1), we have

$$|(\tilde{w}^v)^\mathsf{T}(\bar{y} - v)| \leq \|(\tilde{w}^v)\|_* \|\bar{y} - v\| \leq \|\bar{y} - v\|. \tag{5.2.4}$$

Then using (5.2.3) and (5.2.4), we obtain

$$\|\bar{y} - v\| \geq |(\tilde{w}^v)^\mathsf{T}(\bar{y} - v)| \geq \|z^v\| - \frac{\epsilon}{2} \geq \epsilon - \frac{\epsilon}{2} = \frac{\epsilon}{2}.$$

Therefore, $\bar{y} \notin B(v, \frac{\epsilon}{4})$, which implies $B(v, \frac{\epsilon}{4}) \cap \mathcal{H}' = \emptyset$. □

Lemma 5.2.2. *Suppose that Assumptions 2.2.1 holds. Let $v \notin \mathcal{P}$ and $\mathcal{H}$ and $\mathcal{H}'$ be the corresponding halfspaces to v defined by Proposition 3.2.5 and (5.2.1), respectively. Then, $\mathcal{H} + B(0, \frac{\epsilon}{2}) \subseteq \mathcal{H}'$.*

Proof. Let $y + y' \in \mathcal{H} + B(0, \frac{\epsilon}{2})$. Assume to the contrary that $y + y' \notin \mathcal{H}'$, then $(\tilde{w}^v)^\mathsf{T}(y + y') < (\tilde{w}^v)^\mathsf{T} y^v - \frac{\epsilon}{2}$. Since, $y \in \mathcal{H}$,

$$(\tilde{w}^v)^\mathsf{T} y^v + (\tilde{w}^v)^\mathsf{T} y' \leq (\tilde{w}^v)^\mathsf{T}(y + y') < (\tilde{w}^v)^\mathsf{T} y^v - \frac{\epsilon}{2}.$$

This gives, $(\tilde{w}^v)^\mathsf{T} y' < -\frac{\epsilon}{2}$. Then, $|(\tilde{w}^v)^\mathsf{T} y'| > \frac{\epsilon}{2}$. Now, by using Hölder's inequality, we get,

$$\frac{\epsilon}{2} < |(\tilde{w}^v)^\mathsf{T} y'| \leq \|\tilde{w}^v\|_* \|y'\| \leq \|\tilde{w}^v\|_* \frac{\epsilon}{2}.$$

This implies, $\|\tilde{w}^v\|_* > 1$, which is a contradiction to (5.2.2). □

Theorem 5.2.3. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. Algorithm 3 terminates after a finite number of iterations.*

Proof. We use the similar arguments as in the proof of Theorem 5.1.3 in order to get a contradiction. Assume that for every $k \geq 0$, there exists a vertex $v^k \in \bar{\mathcal{V}}_k$ such that $\|z^{v^k}\| > \epsilon$. From Lemma 4.2.3, $\bar{\mathcal{P}}_0^{\text{out}}$ is compact. Consider, a ball $B(0, \frac{\epsilon}{2})$ which is also compact. Hence, $\bar{\mathcal{P}}_0^{\text{out}} + B(0, \frac{\epsilon}{2})$ is compact by [48, Lemma 5.3]. Define, for $k \geq 0$, $B^k := B(v^k, \frac{\epsilon}{4})$. We show that $B^k \subseteq \bar{\mathcal{P}}_0^{\text{out}} + B(0, \frac{\epsilon}{2})$. Since

$v^k \in \bar{\mathcal{P}}_k^{\text{out}}$, it holds true that

$$B^k \subseteq \{v^k\} + B(0, \frac{\epsilon}{2}) \subseteq \bar{\mathcal{P}}_k^{\text{out}} + B(0, \frac{\epsilon}{2}) \subseteq \bar{\mathcal{P}}_0^{\text{out}} + B(0, \frac{\epsilon}{2}). \tag{5.2.5}$$

Hence, $B^k \subseteq \bar{\mathcal{P}}_0^{\text{out}} + B(0, \frac{\epsilon}{2})$.

Now to prove $B^i \cap B^j = \emptyset$ for every $i, j \geq 0$ with $i \neq j$, without loss of generality assume that $i < j$. Thus, $\bar{\mathcal{P}}_j^{\text{out}} \subseteq \bar{\mathcal{P}}_{i+1}^{\text{out}}$. From Lemma 5.2.1, we have $B^i \cap \mathcal{H}_i' = \emptyset$, where $\mathcal{H}_i' := \left\{y \in \mathbb{R}^q \mid (\tilde{w}^i)^\mathsf{T} y \geq (\tilde{w}^i)^\mathsf{T} y^i - \frac{\epsilon}{2}\right\}$. Moreover, we have

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

where $\mathcal{H}_i$ is the supporting halfspace at y^i as obtained in Proposition 3.2.5. Using Lemma 5.2.2 with above inclusion, we get,

$$\bar{\mathcal{P}}_j^{\text{out}} + B(0, \frac{\epsilon}{2}) \subseteq \mathcal{H}_i + B(0, \frac{\epsilon}{2}) \subseteq \mathcal{H}_i'.$$

This implies $B^i \cap (\bar{\mathcal{P}}_j^{\text{out}} + B(0, \frac{\epsilon}{2})) = \emptyset$. On the other hand, $B^j \subseteq \bar{\mathcal{P}}_j^{\text{out}} + B(0, \frac{\epsilon}{2})$ from (5.2.5). Thus, $B^i \cap B^j = \emptyset$. These imply that there is an infinite number of non-overlapping sets, having strictly positive and fixed volume, contained in the compact set $\bar{\mathcal{P}}_0^{\text{out}} + B(0, \frac{\epsilon}{2})$, which is a contradiction. $\square$

Corollary 5.2.4. *Suppose that Assumptions 2.2.1 and 4.1.1 hold. It holds true that*

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

where $\bar{\mathcal{P}}_k^{in} = \operatorname{conv} \Gamma(\bar{\mathcal{X}}_k) + C$ *and the sets* $\bar{\mathcal{X}}_k, \bar{\mathcal{P}}_k^{out}$ *are as described in Algorithm 3.*

Proof. Algorithm 3 is finite by Theorem 5.2.3. Hence, given $\epsilon > 0$, there exists $K(\epsilon) \in \mathbb{N}$ (the iteration number for which the algorithm stops) such that, for each $k > K(\epsilon)$, the set $\bar{\mathcal{X}}_k$ is a finite weak ϵ-solution as in Definition 2.2.5. Then, for each $k > K(\epsilon)$, we have $\delta^H(\bar{\mathcal{P}}_k^{\text{in}}, \mathcal{P}) \leq \epsilon$ by Remark 2.2.6. Hence, by the definition of limit,

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

thanks to the vertex selection rule of choosing the farthest away vertex in each iteration, see Remark 5.1.5. The second statement follows similar to the corresponding result in Corollary 5.1.4. □

5.3 Convergence Rate of Algorithm 3

In order to study the convergence rate of Algorithm 3, we aim to use the results of Kamenev [13] and Lotov et al. [37, Chapter 8], which originally hold for convex compact bodies. In the following, we recall a definition from [37, Chapter 8]. Note that, the notations used in the original definition are replaced with the ones in this thesis to remain consistent. In order to use the convergence result, [37, Theorem 8.6], our algorithm has to satisfy the type of sequence of outer approximating polytopes in Definition 5.3.1 provided below. Throughout the section, we assume, without loss of generality, that $0 \in \operatorname{int}(\mathcal{P} \cap S)$.

For the theoretical analysis in this section, we ignore the stopping criteria and assume that the algorithm runs indefinitely while updating the current outer approximation in each iteration. This can be done by ignoring the lines 32 and 36 of Algorithm 3.

Definition 5.3.1. *[37, Definition 8.3]* A sequence of outer approximating polytopes $\{A_k\}_{k=1,2,\ldots}$ generated for a convex compact set A, such that $A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots \supseteq A$, by a cutting method, that is, $A_{k+1} = A_k \cap H_k$ where H_k is a half-space, is called an $H(\gamma, A)$-*sequence of cutting* if there exists a constant $\gamma > 0$ such that for any $k \geq 0$ it holds that

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

The next lemma is a simple observation that will be used in Theorem 5.3.3 regarding H-sequence of cutting.

Lemma 5.3.2. *Suppose that Assumptions 2.2.1 holds. Let $v \notin \mathcal{P}$, $\mathcal{H}$ be the corresponding halfspace to v defined by Proposition 3.2.5 and $B(v, \|z^v\|) = \{y \mid \|y - v\| \leq \|z^v\|\}$. Then $\operatorname{int} B(v, \|z^v\|) \cap \mathcal{H} = \emptyset$.*

Proof. First note that $\operatorname{int} B(v, \|z^v\|)$ is well defined, as $v \notin \mathcal{P}$ implies $\|z^v\| > 0$ by Lemma 3.2.3. Using the similar arguments as in the proof of Lemma 5.2.1, the assertion of the lemma follows. □

Note that, in each iteration of the algorithm, we choose the farthest vertex v^{k*} to the upper image of the current outer approximation and generate a halfspace $\mathcal{H}_k$, given by (4.3.2), to update the current outer approximation. The distance from v^{k*} to $\mathcal{H}_k$ is equal to the distance between v^{k*} and $\mathcal{P} \cap S$, that is,

$$d(v^{k*}, \mathcal{H}_k) = d(v^{k*}, \mathcal{P} \cap S).$$

This can be seen in the following theorem where we show that the choice of the farthest vertex v^{k*} ensures 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 $\mathcal{P} \cap S$.

Theorem 5.3.3. *For the sequence of outer approximating polytopes $\{\bar{\mathcal{P}}_k^{out}\}_{k \geq 0}$, we have*

$$\delta^H(\bar{\mathcal{P}}_k^{out}, \bar{\mathcal{P}}_{k+1}^{out}) = \delta^H(\bar{\mathcal{P}}_k^{out}, \mathcal{P} \cap S). \tag{5.3.1}$$

Proof. By definition $\bar{\mathcal{P}}_{k+1}^{\text{out}} \subseteq \bar{\mathcal{P}}_k^{\text{out}}$. By Lemma 4.1.3, we get,

$$\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) = \max\left\{ \sup_{y \in \bar{\mathcal{P}}_k^{\text{out}}} d(y, \bar{\mathcal{P}}_{k+1}^{\text{out}}), \sup_{y \in \bar{\mathcal{P}}_{k+1}^{\text{out}}} d(y, \bar{\mathcal{P}}_k^{\text{out}}) \right\} = \sup_{v \in \bar{\mathcal{V}}_k} d(y, \bar{\mathcal{P}}_{k+1}^{\text{out}}),$$

where $\bar{\mathcal{V}}_k$ is the set of vertices of $\bar{\mathcal{P}}_k^{\text{out}}$. First, consider the case that $\bar{\mathcal{V}}_k \subseteq \mathcal{P}$. Since $\bar{\mathcal{P}}_k^{\text{out}}$ is a polytope, by convexity of $\mathcal{P}$, $\bar{\mathcal{P}}_k^{\text{out}} \subseteq \mathcal{P}$. Then, from $\bar{\mathcal{P}}_k^{\text{out}} \subseteq \bar{\mathcal{P}}_0^{\text{out}} =$

$\mathcal{P}_0^{\text{out}} \cap S \subseteq S$, we get $\bar{\mathcal{P}}_k^{\text{out}} \subseteq \mathcal{P} \cap S$. The other inclusion is trivial by construction. This means that $\bar{\mathcal{P}}_k^{\text{out}} = \mathcal{P} \cap S$. 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.

Now, assume that there exists some $v^k \in \bar{\mathcal{P}}_k^{\text{out}}$ such that $v^k \notin \mathcal{P}$. Assume without loss of generality that v^k is a farthest away vertex from $\mathcal{P} \cap S$. Define, $B_k^o := \operatorname{int} B(v^k, \|z^k\|)$. Then, $B_k^o \cap \mathcal{H}_k = \emptyset$ from Lemma 5.3.2. Hence, $v^k \notin \bar{\mathcal{P}}_k^{\text{out}} \cap \mathcal{H}_k = \bar{\mathcal{P}}_{k+1}^{\text{out}}$. This implies, $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) \geq \|z^k\| > 0$, where $\|z^k\| > 0$ as $v^k \notin \mathcal{P}$, see Lemma 3.2.3. By Lemma 4.1.3, $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P} \cap S)$ is attained at a vertex of $\bar{\mathcal{P}}_k^{\text{out}}$. Hence,

$$
\begin{aligned}
\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P} \cap S) &= \max \left\{ \sup_{y \in \bar{\mathcal{P}}_k^{\text{out}}} d(y, \mathcal{P} \cap S), \sup_{y \in \mathcal{P} \cap S} d(y, \bar{\mathcal{P}}_k^{\text{out}}) \right\} \\
&= \sup_{y \in \bar{\mathcal{P}}_k^{\text{out}}} d(y, \mathcal{P} \cap S) \\
&= \max_{v \in \bar{\mathcal{V}}_k} d(v, \mathcal{P} \cap S) \\
&= \max_{v \in \bar{\mathcal{V}}_k} \|z^v\| \\
&= \|z^{v^{k*}}\|,
\end{aligned}
$$

where v^{k*} is one of the farthest away vertices, see (4.3.1). Note that, $\bar{\mathcal{P}}_k^{\text{out}}$ is updated using $\mathcal{H}_k$ corresponding to v^{k*}. Let $\bar{y} \in \mathcal{H}_k$ be arbitrary. From the definition of $\mathcal{H}_k$, given by 4.3.2, and (5.2.2), we have $(\tilde{w}^{v^{k*}})^{\mathsf{T}} \bar{y} \geq (\tilde{w}^{v^{k*}})^{\mathsf{T}} (v^{k*} + z^{v^{k*}}) = (\tilde{w}^{v^{k*}})^{\mathsf{T}} v^{k*} + \|z^{v^{k*}}\|$. Equivalently,

$$
(\tilde{w}^{v^{k*}})^{\mathsf{T}} (\bar{y} - v^{k*}) \geq \|z^{v^{k*}}\|. \tag{5.3.2}
$$

On the other hand, from Hölder's inequality and (5.2.2), we have

$$
|(\tilde{w}^{v^{k*}})^{\mathsf{T}} (\bar{y} - v^{k*})| \leq \|(\tilde{w}^{v^{k*}})\|_* \|\bar{y} - v^{k*}\| \leq \|\bar{y} - v^{k*}\|. \tag{5.3.3}
$$

Then using (5.3.2) and (5.3.3), we obtain $\|\bar{y} - v^{k*}\| \geq |(\tilde{w}^{v^{k*}})^{\mathsf{T}} (\bar{y} - v^{k*})| \geq \|z^{v^{k*}}\|$.

We get,

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

The last equality is coming 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 Lemma 4.1.3. Hence, $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) \geq \delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P} \cap S)$. Next, since $\bar{\mathcal{P}}_k^{\text{out}} \supseteq \bar{\mathcal{P}}_{k+1}^{\text{out}} \supseteq \mathcal{P} \cap S$, we have $\delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \bar{\mathcal{P}}_{k+1}^{\text{out}}) \leq \delta^H(\bar{\mathcal{P}}_k^{\text{out}}, \mathcal{P} \cap S)$. Hence, the equality follows. □

From now on, we use the following notations to represent the sets for simplicity

$$\mathcal{A} := (\mathcal{P} \cap S), \qquad \mathcal{A}_k := \bar{\mathcal{P}}_k^{\text{out}}, \qquad k \geq 0. \tag{5.3.4}$$

Corollary 5.3.4. *The sequence of outer approximating polytopes $\{\mathcal{A}_k\}_{k\geq 0}$ is an $H(\gamma, \mathcal{A})$-sequence of cutting with $\gamma = 1$, that is,*

$$\delta^H(\mathcal{A}_k, \mathcal{A}_{k+1}) \geq \delta^H(\mathcal{A}_k, \mathcal{A}).$$

Proof. From the definition of $\mathcal{A}$ and $\mathcal{A}_k$ and Theorem 5.3.3, we have

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

Clearly, $\mathcal{A}$ and $\mathcal{A}_k$ satisfy the required conditions for the sets in Definition 5.3.1, since $0 \in \operatorname{int} \mathcal{A} \subseteq \operatorname{int} \mathcal{A}_k$ by our assumption, $\mathcal{A} \subseteq \mathcal{A}_k$ is compact by Lemma 4.2.3 and $\mathcal{A}_k$ is a polytope. Hence, the above equality implies that the sequence of outer approximating polytopes $\{\mathcal{A}_k\}_{k\geq 0}$ is an $H(1, \mathcal{A})$-sequence of cutting. □

By the following theorem, which holds true for any H-sequence, given by Definition 5.3.1, we have an important convergent argument for $\{\mathcal{A}_k\}_{k\geq 0}$.

Theorem 5.3.5. *[37, Theorem 8.5] Let $\{A_k\}_{k\geq 0}$ be an H-sequence for a convex compact set A. It holds true that*

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

Now, before using the convergence rate result, Theorem 5.3.8, we show that the Hausdorff distance between an outer approximation $\bar{\mathcal{P}}_k^{\text{out}} + C, k \geq 0$, and the upper image is bounded above by the Hausdorff distance between the compact versions of the sets defined by (5.3.4).

Lemma 5.3.6. *We have,*

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

Proof. By Lemma 4.1.3, $\delta^H(\mathcal{A}_k, \mathcal{A})$ is attained at a vertex of $\mathcal{A}_k$. Hence,

$$\delta^H(\mathcal{A}_k, \mathcal{A}) = \max\left\{\sup_{y\in\mathcal{A}_k} d(y,\mathcal{A}), \sup_{y\in\mathcal{A}} d(y,\mathcal{A}_k)\right\} = \sup_{y\in\mathcal{A}_k} d(y,\mathcal{A}) = \max_{v\in\mathcal{V}_k^{\mathcal{A}}} d(v,\mathcal{A}),$$

where $\mathcal{V}_k^{\mathcal{A}}$ is the set of vertices of $\mathcal{A}_k$. Let $\mathcal{V}_k^C$ be the set of vertices of $\mathcal{A}_k + C$. By Lemma 2.1.1, we have $\mathcal{V}_k^C \subseteq \mathcal{V}_k^{\mathcal{A}}$. We get,

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

From Remark 4.3.2, $\mathcal{P} = (\mathcal{P}\cap S) + C$. Hence, $\delta^H(\mathcal{A}_k, \mathcal{A}) \geq \delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P})$. $\square$

Remark 5.3.7. *By Corollary 5.3.4, $\{\mathcal{A}_k\}_{k\geq 0}$ is an $H(\gamma,\mathcal{A})$-sequence of polytopes* *with $\gamma = 1$. Then, Theorem 5.3.5 and Lemma 5.3.6 together deduce the following result which is already shown by Corollary 5.2.4,*

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

In the following, we state the results from [37] concerning the convergence rate of H-sequences. The proofs can be found in [13].

Theorem 5.3.8. *[37, Theorem 8.6] Let $\{A_k\}_{k=1,2,\dots}$ be an $H(\gamma, A)$-sequence of cutting for a convex compact set A. Then for any ϵ, $0 < \epsilon < 1$, there exists N such that for $k \ge N$ it holds that*

$$\delta^H(A_k, A) \le (1+\epsilon)\lambda(\gamma)k^{1/1-q},$$

where $\lambda(\gamma)$ depends on the topological properties of A and can be found in [13],

$$\lambda(\gamma) = \left[\frac{q}{(q-1)\pi_{q-1}}\frac{\sigma(A)}{\gamma^q}\right]^{\frac{1}{q-1}} \omega(A),$$

$\omega(A) = \frac{R(A)}{r(A)}$ is the asphericity of A which is the minimum ratio of the concentric balls circumscribed and inscribed around A, $R(A)$ is the radius of the smallest ball circumscribed around A, $r(A)$ is the radius of the largest ball inscribed in A, π_{q-1} is the volume of unit ball in $\mathbb{R}^{q-1}$ and $\sigma(A)$ is the surface area of A.

Theorem 5.3.9. *The approximation error obtained through the iterations of Algorithm 3 decreases by the order of $\mathcal{O}(k^{1/1-q})$.*

Proof. By Corollary 5.3.4, the sequence of outer approximating polytopes $\{\mathcal{A}_k\}_{k\ge0}$ is an $H(\gamma, \mathcal{A})$-sequence of cutting. Then, by Lemma 5.3.6 and Theorem 5.3.8, we get

$$\delta^H(\bar{\mathcal{P}}_k^{\text{out}} + C, \mathcal{P}) \le \delta^H(\mathcal{A}_k, \mathcal{A}) \le (\text{const})\ k^{1/1-q}.$$

Hence the result follows. □

5.4 Improved Estimate of the Convergence Rate of Algorithm 3 with Euclidean Norm

In this section, we find an improved estimate of convergence rate of Algorithm 3, when ℓ_2 norm is used in the scalarizations. We first state the theorem which is similar to the one in the literature [37, 43], except that instead of the type of sequences of outer approximating polytopes in the literature, see [37, Definition 8.4 and Theorem 8.14], it is stated for Algorithm 3. Similar to Section 5.3, for the purpose of analysis, we assume that $0 \in \operatorname{int}(\mathcal{P} \cap S)$.

Theorem 5.4.1. *Let $\|\cdot\|$ be $\|\cdot\|_2$. Let $\{\bar{\mathcal{P}}_k^{out}\}_{k=0,1,\dots}$ be the outer approximating sets of $\mathcal{P} \cap S$ computed by Algorithm 3. Then for any ϵ, $0 < \epsilon < 1$, there exists N such that for $k \geq N$ it holds that*

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

where k indicates number of iteration or number of halfspaces of $\bar{\mathcal{P}}_{k-1}^{out}$,

$$\lambda(\mathcal{P} \cap S) = 16R(\mathcal{P} \cap S) \left(\frac{q\pi_q}{\pi_{q-1}}\right)^{\frac{2}{q-1}},$$

$R(\mathcal{P} \cap S)$ is a radius of the smallest ball circumscribed around $\mathcal{P} \cap S$, $\pi_q = \mu(B(0,1))$ is the volume of unit ball, μ is the volume measure and $B(0,1) \in \mathbb{R}^q$ is the unit ball centered in the origin.

Before presenting a proof of Theorem 5.4.1, we state the following few lemmas along with their proofs wherever necessary.

Lemma 5.4.2. *[37, Lemma 8.15] Let $y^1, y^2 \in \mathbb{R}^q$ and $\eta > 0$. Then for any $u, v \in \mathbb{S}^{q-1}$, where $u^T v > 0, y^2 \in L(u, y^1)$ and $y^1 \in L(v, y^2)$, where $L(w, y)$ is a halfspace at y with unit normal direction w, we have*

$$d(y^1, \operatorname{bd} L(v, y^2)) \leq \frac{\|(y^1 - \eta u) - (y^2 - \eta v)\|^2}{\eta}.$$

Lemma 5.4.3. *Let $\|\cdot\|$ be $\|\cdot\|_2$. Let $0 \in \operatorname{int}(\mathcal{P} \cap S)$, a point $v \notin \mathcal{P} \cap S$, $y =$*

$v+z \in \operatorname{bd}(\mathcal{P}\cap S)$, where $\|z\| = d(v,\mathcal{P}\cap S)$, and $\eta > 0$. Denote by $H(w,y,\mathcal{P}\cap S)$ a supporting halfspace of $\mathcal{P}\cap S$ at y with unit normal direction w and $\mathcal{P}\cap S \subseteq H(w,y,\mathcal{P}\cap S)$. For a unit vector w' and $y' \in \operatorname{bd}\mathcal{P}\cap S$, let $H'(w',y',\mathcal{P}\cap S)$ be such that $v \in H'(w',y',\mathcal{P}\cap S)$. Then, it holds that

$$\|(y-\eta w)-(y'-\eta w')\| \geq \min\left\{\eta\sqrt{2}, \sqrt{\eta h} - \omega_0(\mathcal{P}\cap S)h\right\},$$

where $h = w^\mathsf{T} y - w^\mathsf{T} v$, $\omega_0(\mathcal{P}\cap S) = \frac{R(\mathcal{P}\cap S)}{r(\mathcal{P}\cap S)}$ is an asphericity of $\mathcal{P}\cap S$, $R(\mathcal{P}\cap S)$ is the radius of the smallest ball circumscribed around $\mathcal{P}\cap S$ and centered at origin, $r(\mathcal{P}\cap S)$ is the radius of the largest ball inscribed in $\mathcal{P}\cap S$ and centered at origin.

Proof. If $w^\mathsf{T} w' \leq 0$,

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

We know that $w^\mathsf{T}(y'-y) \geq 0$ and $(w')^\mathsf{T}(y-y') \geq 0$. Hence,

$$\begin{aligned}
&\|(y-y')\|^2 + \eta^2\|(w'-w)\|^2 + 2\eta w^\mathsf{T}(y'-y) + 2\eta(w')^\mathsf{T}(y-y') \geq \eta^2\|(w'-w)\|^2 \\
&= \eta^2\|w'\|^2 + \eta^2\|w\|^2 - 2\eta^2 w^\mathsf{T} w' \\
&= \eta^2\|w'\|^2 + \eta^2\|w\|^2 + 2\eta^2 |w^\mathsf{T} w'| \\
&\geq \eta^2\|w'\|^2 + \eta^2\|w\|^2 \\
&= \eta^2(1+1).
\end{aligned}$$

We get, $\|(y-\eta w)-(y'-\eta w')\| \geq \eta\sqrt{2}$.

Let $w^\mathsf{T} w' > 0$. We have

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

where $H(w,y)$ is a halfspace at y with unit normal direction w and $\mathcal{P}\cap S \subseteq H(w,y)$. From Lemma 5.2.1, It can be seen that, $v \notin \mathcal{P}$ implies $v \notin H(w,y)$.

Hence,

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

From our assumption, we have

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

Then, from Lemma 5.4.2, we get

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

Note that, distance between a point and a hyperplane is the length of the orthogonal projection of the vector (between the point and an element from the hyperplane) on a normal vector to the hyperplane, [62]. This is determined as follows. We know that, $v \in \operatorname{bd} H(w, v)$ and $w/\|w\|$ is the unit normal vector to $\operatorname{bd} H(w, v)$. In order to find the distance between y' and $\operatorname{bd} H(w, v)$, consider the vector $y' - v$. Then, the length of the orthogonal projection is obtained by taking the norm of the projection of $y' - v$ onto $w/\|w\|$, that is, the norm of their inner product. Hence,

$$\begin{aligned}
\frac{\|(v - \eta w) - (y' - \eta w')\|^2}{\eta} &\geq d(y', \operatorname{bd} H(w, v)) \\
&= \frac{|(w)^\mathsf{T}(y' - v)|}{\|w\|} \\
&= w^\mathsf{T} y' - w^\mathsf{T} y + w^\mathsf{T} y - w^\mathsf{T} v \\
&\geq w^\mathsf{T} y - w^\mathsf{T} v = h.
\end{aligned}$$

Secondly, we have,

$$\begin{aligned}
\|(v - \eta w) - (y - \eta w)\| &= d(v, y) = \|z\| \\
&\leq w^\mathsf{T}(y - v) \\
&\leq \omega_0(\mathcal{P} \cap S) w^\mathsf{T}(y - v) \\
&= \omega_0(\mathcal{P} \cap S) h.
\end{aligned}$$

First inequality above is coming from the definition of $\mathcal{H}$ (supporting halfspace with the normal direction w^v) and (5.2.2). Finally, by using triangle inequality,

$$\|(v-\eta w)-(y'-\eta w')\| \leq \|(v-\eta w)-(y-\eta w)\| + \|(y-\eta w)-(y'-\eta w')\|,$$

we get,

$$\begin{aligned}\|(y-\eta w)-(y'-\eta w')\| &\geq \|(v-\eta w)-(y'-\eta w')\| - \|(v-\eta w)-(y-\eta w)\| \\ &\geq \sqrt{\eta h} - \omega_0(\mathcal{P}\cap S)h.\end{aligned}$$

$\square$

Consider $\{\bar{\mathcal{P}}_k^{\text{out}}\}_{k=0,1,\ldots}$ which is the sequence of cutting obtained by Algorithm 3 and $\eta > 0$. Define the sequence $\{Z_k\}_{k=0,1,\ldots}$ as follows:

(i) $Z_0 = \{\Gamma(x^1)-\eta w^1, \ldots, \Gamma(x^J)-\eta w^J, \bar{y}-\eta\bar{w}\}$, where J is the total number of generating vectors of the dual of the ordering cone C, $\bar{w}$ is as defined in (4.2.1) and let $\bar{y}$ be such that $\bar{y} \geq \sup\{\bar{w}^{\mathsf{T}}\Gamma(x) \mid x \in \mathcal{X}\}$ which exists since the supremum is finite as it is attained due to the fact that $\mathcal{X} \subseteq \mathbb{R}^n$ is a compact set and $x \mapsto \bar{w}^{\mathsf{T}}\Gamma(x)$ is a continuous function on X by Assumption 2.2.1.

(ii) For $y^k = v^k + z^k \in \operatorname{bd}(\mathcal{P}\cap S)$ and w^k, the unit normal direction of the supporting halfspace $\mathcal{H}_k$ at y^k, obtained at $(k+1)^{\text{th}}$ iteration of Algorithm 3,

$$Z_{k+1} = \{y^k - \eta w^k\} \cup Z_k, \tag{5.4.1}$$

Remark 5.4.4. *We have, $Z_k \subseteq \operatorname{bd}(\mathcal{P}\cap S - \eta B(0,1))$. Indeed, for any $\tilde{k} < k$,*

$$\mathcal{P}\cap S - \eta B(0,1) \subseteq \mathcal{H}_{\tilde{k}} - \eta B(0,1),$$

where $\mathcal{H}_{\tilde{k}}$ is the supporting halfspace of $\mathcal{P}\cap S$ at $y^{\tilde{k}}$ with unit normal direction $w^{\tilde{k}}$. Define shifted halfspace, $\mathcal{H}'_{\tilde{k}} := \{y \in \mathbb{R}^q \mid w^{\tilde{k}}y \geq (w^{\tilde{k}})^{\mathsf{T}}y^{\tilde{k}} - \eta\}$. Then, from

Lemma 5.2.2, $\mathcal{H}_{\tilde{k}} + \eta B(0,1) \subseteq \mathcal{H}'_{\tilde{k}}$. *It gives,*

$$\begin{aligned}\mathcal{P} \cap S - \eta B(0,1) &\subseteq \mathcal{H}_{\tilde{k}} - \eta B(0,1) \\ &= \mathcal{H}_{\tilde{k}} + \eta B(0,1) \\ &\subseteq \mathcal{H}'_{\tilde{k}}.\end{aligned}$$

Clearly, $y^{\tilde{k}} - \eta w^{\tilde{k}} \in \mathcal{P} \cap S - \eta B(0,1)$. *Also, note that,* $y^{\tilde{k}} - \eta w^{\tilde{k}} \in \operatorname{bd} \mathcal{H}'_{\tilde{k}}$*, since*

$$(w^{\tilde{k}})^\mathsf{T}(y^{\tilde{k}} - \eta w^{\tilde{k}}) = (w^{\tilde{k}})^\mathsf{T} y^{\tilde{k}} - \eta (w^{\tilde{k}})^\mathsf{T} w^{\tilde{k}} = (w^{\tilde{k}})^\mathsf{T} y^{\tilde{k}} - \eta \|w^{\tilde{k}}\|_2^2 = (w^{\tilde{k}})^\mathsf{T} y^{\tilde{k}} - \eta.$$

Hence, the result follows.

Lemma 5.4.5. *Let* $\|\cdot\|$ *be* $\|\cdot\|_2$. *Consider* $\mathcal{P} \cap S$ *be non-polyhedral set. Then,* $\operatorname{card} Z_k = J + 1 + k$*, where* Z_k *is given by (5.4.1),* J *is the total number of generating vectors of the dual of the ordering cone* C *and* $k \geq 0$.

Proof. Assume to the contrary that $\operatorname{card} Z_{k+1} \neq J + 1 + (k+1)$. Then, by definition of Z_{k+1}, see (5.4.1), we get $\operatorname{card} Z_{k+1} < J + 1 + (k+1)$. Assume, without loss of generality, that the strict inequality is observed, for the first time, at $(k+1)^{th}$ iteration. Then, for y^k and some $y^{\tilde{k}}$, $\tilde{k} < k$, both on $\operatorname{bd} \mathcal{P} \cap S$, it holds that

$$a := y^k - \eta w^k = y^{\tilde{k}} - \eta w^{\tilde{k}}.$$

We have following two cases.

1. Let, $y^k \neq y^{\tilde{k}}$. Consider, $\mathcal{H}_k$ be the supporting halfspace of $\mathcal{P} \cap S$ at y^k with unit normal direction w^k. Fix a point, say b, on $\operatorname{bd} \mathcal{H}_k$ to define a vector $a - b$. Then, the distance from a to the hyperplane $\operatorname{bd} \mathcal{H}_k$ is the norm of the orthogonal projection of $a - b$ on w^k [62], that is,

$$d(a, \mathcal{H}_k) = \frac{|(w^k)^\mathsf{T}(a - b)|}{\|w^k\|}.$$

In particular, since y^k is on $\operatorname{bd}\mathcal{H}_k$ and w^k is the unit vector,

$$d(a, \mathcal{H}_k) = |(w^k)^\mathsf{T}(a - y^k)| = |(w^k)^\mathsf{T}(y^k - \eta w^k - y^k)| = \eta|(w^k)^\mathsf{T}w^k| = \eta\|w^k\|_2^2 = \eta.$$

Note that, the distance is attained at y^k, since

$$d(a, y^k) = \|a - y^k\| = \|y^k - \eta w^k - y^k\| = \eta\|w^k\| = \eta.$$

Now, for any y in $\mathcal{H}_k$, $y \neq y^k$, $d(a, y) > d(a, y^k)$. We have, $y^{\tilde{k}} \in \mathcal{P} \cap S \subseteq \mathcal{H}_k$ and $y^{\tilde{k}} \neq y^k$. Then,

$$\eta = d(a, y^{\tilde{k}}) > d(a, y^k) = \eta,$$

a contradiction. Hence, as long as the boundry points, $y^k, k \geq 0$, are different in each iteration, it will generate a unique element of Z_k.

2. Let $y^k = y^{\tilde{k}}$. Then, $w^k \neq w^{\tilde{k}}$ implies that $y^k - \eta w^k \neq y^{\tilde{k}} - \eta w^{\tilde{k}}$, a contradiction.

From the two cases above, we conclude that

$$y^k - \eta w^k = y^{\tilde{k}} - \eta w^{\tilde{k}} \qquad \Longrightarrow \qquad y^k = y^{\tilde{k}} \quad \text{and} \quad w^k = w^{\tilde{k}}.$$

Then, the corresponding halfspaces $\mathcal{H}_k$ and $\mathcal{H}_{\tilde{k}}$ at y^k and $y^{\tilde{k}}$ with unit normal directions w^k and $w^{\tilde{k}}$, respectively, at two different iterations, are equal, that is, $\mathcal{H}_k = \mathcal{H}_{\tilde{k}}$. Since, $\tilde{k} < k$, we get

$$\bar{\mathcal{P}}^{\text{out}}_{k+1} = \bar{\mathcal{P}}^{\text{out}}_{k} \cap \mathcal{H}_k = \bar{\mathcal{P}}^{\text{out}}_{k} \cap \mathcal{H}_{\tilde{k}} = \bar{\mathcal{P}}^{\text{out}}_{k}.$$

In case of nonpolyhedral set, we have $\delta^H(\bar{\mathcal{P}}^{\text{out}}_k, \mathcal{P}) > 0$, for any k. Then, from H-seq, we get

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_k, \bar{\mathcal{P}}^{\text{out}}_k) = \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k+1}, \bar{\mathcal{P}}^{\text{out}}_k) = \delta^H(\bar{\mathcal{P}}^{\text{out}}_k, \mathcal{P}) > 0,$$

a contradiction. Hence, $\operatorname{card} Z_k = J + 1 + k$. □

Before proceeding further, we recall the following definition, which is critical in proving the convergence rate of the algorithm in Theorem 5.4.1.

Definition 5.4.6. *[37, Chapter 8] A set Z is called the base of an ϵ-packing if $d(x, y) \geq 2\epsilon$, for any $x, y \in Z$.*

The following lemma is originally proved for an H_1-sequence of outer approximating polytopes, see [37, Definition 8.4 and Lemma 8.19]. Here, we show that the lemma still holds for the sequence of outer approximating polytopes generated by Algorithm 3 which, in general, may not be an H_1-sequence.

Lemma 5.4.7. *Let $\|\cdot\|$ be $\|\cdot\|_2$. Let $\{\bar{\mathcal{P}}_k^{out}\}_{k=0,1,\ldots}$ be the sequence of sets computed by Algorithm 3, $\{Z_k\}_{k=0,1,\ldots}$ be the sequence as introduced by (5.4.1) and $\eta > 0$. Then for any ϵ, $0 < \epsilon < 1$, there exists N such that for $k \geq N$ the set Z_k is the base of an $\epsilon_k^N(\epsilon)$-packing, where*

$$\epsilon_k^N(\epsilon) = \frac{1}{2}\min\left\{\eta\sqrt{2}, 2\epsilon_N, (1-\epsilon)\sqrt{\eta\delta^H(\bar{\mathcal{P}}_{k-1}^{out}, \mathcal{P} \cap S)}\right\} \tag{5.4.2}$$

and $\epsilon_N = \min\{d(x, y) \mid x, y \in Z_N, x \neq y\}/2$.

Proof. Let us fix N to be determined later. Clearly, Z_k is the base of an ϵ_k-packing, where

$$\epsilon_k = \min\{d(x, y) \mid x, y \in Z_k, x \neq y\}/2.$$

Then, for $k = N, d(x, y) \geq 2\epsilon_N \geq 2\epsilon_N^N(\epsilon)$, for any $x, y \in Z_N$. Hence, the assertion is trivial.

Assume that Z_{k-1} is the base of an $\epsilon_{k-1}^N(\epsilon)$-packing, where $k - 1 \geq N$. By the induction hypothesis,

$$d(x, y) \geq 2\epsilon_{k-1}^N(\epsilon) \qquad \forall\, x, y \in Z_{k-1}.$$

Moreover, since $\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-2}, \mathcal{P} \cap S) \geq \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P} \cap S)$, from (5.4.2) we have

$$\epsilon^N_{k-1}(\epsilon) \geq \epsilon^N_k(\epsilon).$$

Then, Z_{k-1} is the base of an $\epsilon^N_k(\epsilon)$-packing.

Let $b^{k-1} = y^{k-1} - \eta w^{k-1} = v^{k-1} + z^{k-1} - \eta w^{k-1}$. We know that, $Z_k = \{b^{k-1}\} \cup Z_{k-1}$. To show that Z_k is the base of an $\epsilon^N_k(\epsilon)$-packing, it is enough to show that,

$$d(b^{k-1}, b) \geq 2\epsilon^N_k(\epsilon) \qquad \forall\, b \in Z_{k-1}.$$

Let $b = y' - \eta w'$. We know that $v^{k-1} \notin \mathcal{P} \cap S$ and $v^{k-1} \in H'(w', y', \mathcal{P} \cap S)$. Also, $y' \in \operatorname{bd} H'(w', y', \mathcal{P} \cap S) \cap (\mathcal{P} \cap S)$ by definition. Then, by Lemma 5.4.3, we have

$$\|b^{k-1} - b\| = \|(y^{k-1} - \eta w^{k-1}) - (y' - \eta w')\| \geq \min\left\{\eta\sqrt{2},\ \sqrt{\eta h^{k-1}} - \omega_0(\mathcal{P} \cap S)h^{k-1}\right\},$$

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

$$\begin{aligned}
(w^{k-1})^\mathsf{T} y^{k-1} &= (w^{k-1})^\mathsf{T} v^{k-1} + (w^{k-1})^\mathsf{T} z^{k-1} \\
&= (w^{k-1})^\mathsf{T} v^{k-1} + \|z^{k-1}\| \\
&= (w^{k-1})^\mathsf{T} v^{k-1} + \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P} \cap S),
\end{aligned}$$

where second equality is coming from (5.2.2) and third equality is a consequence of choosing farthest away vertex in each iteration. Hence,

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P} \cap S) = h^{k-1}.$$

We choose N so that

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P} \cap S) \leq \delta^H(\bar{\mathcal{P}}^{\text{out}}_{N}, \mathcal{P} \cap S) \leq \frac{\eta\epsilon^2}{\omega_0^2(\mathcal{P} \cap S)},$$

where $k-1 \geq N$.

$$
\begin{aligned}
\sqrt{\eta h^{k-1}} - \omega_0(\mathcal{P}\cap S)h^{k-1} &= \sqrt{\eta \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)} - \omega_0(\mathcal{P}\cap S)\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S) \\
&\geq \sqrt{\eta \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)} - \sqrt{\frac{\eta\epsilon^2}{\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)}}\ \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S) \\
&= \sqrt{\eta \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)} - \epsilon\sqrt{\eta \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)} \\
&= (1-\epsilon)\sqrt{\eta \delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P}\cap S)}.
\end{aligned}
$$

Hence, we get,

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

$\square$

Lemma 5.4.8. *[37, Lemma 8.20] Let A be a convex compact set. Then any base of an ϵ-packing on* bd A*, $0 < \epsilon < R(A)$, contains no more than $N_1(\epsilon, R(A))$ elements, where*

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

Now, using the above Lemmas, we are ready to prove Theorem 5.4.1.

Proof of Theorem 5.4.1. If $\mathcal{P}\cap S$ is a polytope, then there exists $K' \in \mathbb{N}$ such that, for every $k \geq K'$, $\delta^H(\bar{\mathcal{P}}^{\text{out}}_k, \mathcal{P}\cap S) = 0$ and the assertion of the theorem holds trivially. For the rest of the proof, we assume that $\mathcal{P}\cap S$ is a non-polyhedral set.

For this part of the proof, let us fix $\epsilon' \in (0,1)$ and $\eta > 0$ arbitrarily. Let $\{Z_k\}_{k\in\mathbb{Z}_+}$ be the sequence as introduced by (5.4.1). By Lemma 5.4.7, there exists

$N' \in \mathbb{N}$ such that, for every $k \geq N'$, the set Z_k is the base of an $\epsilon_k^{N'}(\epsilon')$-packing, where

$$\epsilon_k^{N'}(\epsilon') = \min\left\{\frac{\eta\sqrt{2}}{2}, \epsilon_{N'}, \tau_k\right\}, \quad \epsilon_{N'} = \frac{1}{2}\min\left\{d(x,y) \mid x, y \in Z_{N'}, x \neq y\right\},$$
$$\tau_k := \frac{(1-\epsilon')\sqrt{\eta\delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P}\cap S)}}{2}.$$

Moreover, by Corollary 5.3.4, $\{\bar{\mathcal{P}}_k^{\text{out}}\}_{k\in\mathbb{Z}_+}$ is an $H(1, \mathcal{P}\cap S)$-sequence. Hence, by Theorem 5.3.5, there exists $N'' \geq N'$ such that

$$\delta^H(\bar{\mathcal{P}}_{N''-1}^{\text{out}}, \mathcal{P}\cap S) \leq \frac{4}{\eta}\left(\frac{\epsilon_{N'}^{N'}(\epsilon')}{1-\epsilon'}\right)^2 \tag{5.4.3}$$

provided that $\epsilon_{N'}^{N'}(\epsilon') > 0$. However, if $\epsilon_{N'}^{N'}(\epsilon') = 0$, then (5.4.2) implies that $\delta^H(\bar{\mathcal{P}}_{N'-1}^{\text{out}}, \mathcal{P}\cap S) = 0$ since $\epsilon_{N'} > 0$ and $\eta > 0$ so that the above inequality still holds. Hence, we get

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

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

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

in particular, Z_k is the base of a τ_k-packing. Similar to (5.4.3), using Theorem 5.3.5, we may find $K_R \in \mathbb{N}$ such that, for every $k \geq K_R$,

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

which is equivalent to

$$\tau_k < R(\mathcal{P}\cap S + \eta B(0,1)).$$

By Remark 5.4.4, $Z_k \subseteq \text{bd}(\mathcal{P}\cap S + \eta B(0,1))$. Then, from Lemma 5.4.8, for every

$k \geq \max\{N'', K_R\}$, we get

$$\operatorname{card} Z_k \leq N_1(\tau_k, R(\mathcal{P} \cap S + \eta B(0,1))),$$

where

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

Since $\mathcal{P} \cap S$ is a non-polyhedral set by assumption, from Lemma 5.4.5, for every $k \in \mathbb{Z}_+$, we have

$$\operatorname{card} Z_k = J + 1 + k$$

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

$$k < \operatorname{card} Z_k \leq N_1(\tau_k, R(\mathcal{P} \cap S) + \eta),$$

which implies that

$$k < \frac{q\pi_q}{\pi_{q-1}} \left\{ 1 + \frac{4(R(\mathcal{P} \cap S) + \eta)^2}{(1-\epsilon')^2 \eta \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S)} \right\}^{\frac{q-1}{2}},$$

or equivalently,

$$\left[\frac{q\pi_q}{\pi_{q-1}} \right]^{\frac{2}{1-q}} k^{\frac{2}{q-1}} < 1 + \frac{4}{(1-\epsilon')^2 \delta^H(\bar{\mathcal{P}}_{k-1}^{\text{out}}, \mathcal{P} \cap S)} \frac{(R(\mathcal{P} \cap S) + \eta)^2}{\eta}.$$

Note that, letting $\eta = 1$ would provide an upper bound, but in order to find a tighter bound, we minimize the right hand side over η [43]. Define,

$$f(\eta) := \frac{(R(\mathcal{P} \cap S) + \eta)^2}{\eta} = \frac{R^2(\mathcal{P} \cap S)}{\eta} + 2R(\mathcal{P} \cap S) + \eta,$$

which is convex over $\eta \in \mathbb{R}_{++}$. Then, from first order condition, we get

$$f'(\eta) = 0$$
$$-\frac{R^2(\mathcal{P} \cap S)}{\eta^2} + 1 = 0,$$

that is, the minimum over η attains at $\eta = R(\mathcal{P} \cap S)$. Continuing with this choice of η (and the corresponding K_R), we get

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

or equivalently,

$$\delta^H(\bar{\mathcal{P}}^{\text{out}}_{k-1}, \mathcal{P} \cap S) \leq \frac{16R(\mathcal{P} \cap S)}{(1-\epsilon')^2 \left[\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{1-q}} - 1\right]} \tag{5.4.4}$$

for every $k \geq \max\{N'', K_R\}$.

In the next part of the proof, we relate the right hand side of (5.4.4) to the assertion of the theorem. Let $\epsilon \in (0,1)$. We would like to choose $\epsilon' \in (0,1)$ in such a way that, for all sufficiently large k, it holds

$$\frac{16R(\mathcal{P} \cap S)}{(1-\epsilon')^2 \left[\left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{1-q}} - 1\right]} \leq (1+\epsilon) 16R(\mathcal{P} \cap S) \left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{q-1}}. \tag{5.4.5}$$

Note that this inequality is equivalent to

$$g(k) \leq (1+\epsilon)(1-\epsilon')^2 - 1, \tag{5.4.6}$$

where

$$g(k) := \frac{1}{1 - \left(\frac{q\pi_q}{k\pi_{q-1}}\right)^{\frac{2}{q-1}}} - 1.$$

Here, we have $(1+\epsilon)(1-\epsilon')^2 - 1 > 0$ if and only if $\epsilon' < 1 - \frac{1}{\sqrt{1+\epsilon}}$. Let us choose

$$\epsilon' := \frac{1}{2}\left(1 - \frac{1}{\sqrt{1+\epsilon}}\right). \tag{5.4.7}$$

Then, (5.4.6) becomes

$$g(k) \leq \frac{(1+\sqrt{1+\epsilon})^2}{4} - 1 = \frac{\epsilon + 2\sqrt{1+\epsilon} - 2}{4}. \tag{5.4.8}$$

Note that g is a decreasing function on $\mathbb{N}$ with $\lim_{k\to\infty} g(k) = 0$. Hence, there exists $N''' \in \mathbb{N}$ such that (5.4.8) holds for every $k \geq N'''$; then, (5.4.5) holds for such k as well with the choice of ϵ' in (5.4.7).

Finally, for given $\epsilon \in (0,1)$, we combine (5.4.4) and (5.4.5) with ϵ' as in (5.4.7). Hence, for every $k \geq \max\{N'', K_R, N'''\}$, we obtain

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

as desired. □

By Theorem 5.4.1, we prove that the approximation error obtained through the iterations of Algorithm 3 when the Euclidean norm is used in the scalarization, decreases by the order of $\mathcal{O}(k^{2/1-q})$.

Chapter 6

Experimental Results

In this chapter, we examine few numerical examples to evaluate the performance of Algorithms 1 and 2 in comparison with the primal algorithm in [1]. We also compare Algorithms 1, 2 and 3 after some modifications in the first two with respect to Algorithm 3 in order to have a fair comparison. The algorithms are implemented using MATLAB R2018a along with CVX, a package to solve convex programs [63, 64], and *bensolve tools* [10] to solve the scalarization and vertex enumeration problems in each iteration, respectively. The tests are performed using a 3.6 GHz Intel Core i7 computer with a 64 GB RAM.

6.1 Test Problems and Computational Results

We consider three examples: Example 6.1.1 is a standard illustrative example with a linear objective function, see for instance [20, 1], in which both the feasible region and its image are the Euclidean unit ball centered at the vector $e = (1, \dots, 1)^\mathsf{T} \in \mathbb{R}^q$. In Example 6.1.2, the objective functions are nonlinear while the constraints are linear (see [20, Example 5.8], [65]); in Example 6.1.3, nonlinear terms appear both in the objective function and in the constraints (see [20, Example 5.10], [65]).

Example 6.1.1. *We consider the following problem for* $q \in \{2,3,4\}$*, where the ordering cone is* $C = \mathbb{R}^q_+$*:*

$$\begin{aligned} &\text{minimize} \;\; \Gamma(x) = x \;\; \text{with respect to} \leq_C \\ &\text{subject to} \;\; \|x - e\|_2 \leq 1, \;\; x \in \mathbb{R}^q. \end{aligned} \tag{6.1.1}$$

Example 6.1.2. *Let* $a^1 = (1,1)^\mathsf{T}, a^2 = (2,3)^\mathsf{T}, a^3 = (4,2)^\mathsf{T}$*. Consider the following problem:*

$$\begin{aligned} &\text{minimize} \quad \Gamma(x) = (\|x-a^1\|_2^2, \|x-a^2\|_2^2, \|x-a^3\|_2^2)^\mathsf{T} \;\; \text{with respect to} \leq_{\mathbb{R}^3_+} \\ &\text{subject to} \quad x_1 + 2x_2 \leq 10, \;\; 0 \leq x_1 \leq 10, \;\; 0 \leq x_2 \leq 4, \;\; x \in \mathbb{R}^2. \end{aligned} \tag{6.1.2}$$

Example 6.1.3. *Let* $\hat{b}^1 = (0,10,120), \hat{b}^2 = (80,-448,80), \hat{b}^3 = (-448,80,80)$ *and* $b^1, b^2, b^3 \in \mathbb{R}^n$*. Consider*

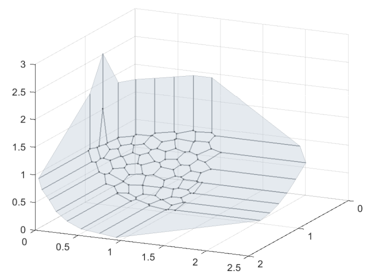
$$\begin{aligned} &\text{minimize} \;\; \Gamma(x) = (\|x\|_2^2 + b^1x, \|x\|_2^2 + b^2x, \|x\|_2^2 + b^3x)^\mathsf{T} \;\; \text{with respect to} \leq_{\mathbb{R}^3_+} \\ &\text{subject to} \;\; \|x\|_2^2 \leq 100, \;\; 0 \leq x_i \leq 10 \text{ for } i \in \{1,\dots,n\}, \;\; x \in \mathbb{R}^n. \end{aligned} \tag{6.1.3}$$

(a) Let $n = 3$ *and* $b^1 = \hat{b}^1, b^2 = \hat{b}^2, b^3 = \hat{b}^3$.

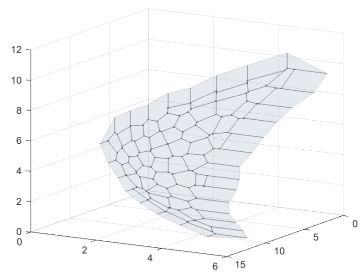
(b) Let $n = 9$ *and* $b^1 = (\hat{b}^1, \hat{b}^1, \hat{b}^1), b^2 = (\hat{b}^2, \hat{b}^2, \hat{b}^2), b^3 = (\hat{b}^3, \hat{b}^3, \hat{b}^3)$.

We solve these examples with Algorithms 1, 2 and the one from the literature [1], where the norm in (P(v)) is taken as the ℓ_p norm for $p \in \{1,2,\infty\}$. An outer approximation of the upper image for each example is shown in Figure 6.1. For the Algorithm from the literature [1], the fixed direction vector for the scalarization model is taken as $\frac{e}{\|e\|_p} \in \mathbb{R}^q$, again for $p \in \{1,2,\infty\}$. This way, it is guaranteed that when the Algorithm returns a finite weak ϵ-solution in the sense of [1, Definition 3.3], this solution is also a finite weak ϵ-solution in the sense of Definition 2.2.5 for the corresponding norm-ball, see [25, Remark 3.4]. We solve Examples 6.1.2 and 6.1.3 for the approximation errors as chosen by Ehrgott et al. in [20] for the same examples.

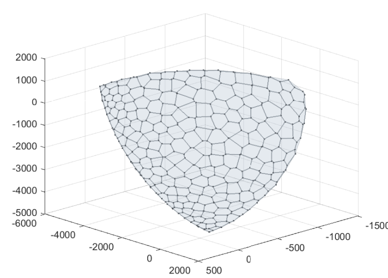
The computational results are presented in Tables 6.1-6.7, which show the approximation error (ϵ), the algorithm (Alg), the cardinality of finite weak ϵ-solution ($|\bar{\mathcal{X}}|$), the number of optimization problems (Opt), and the number of vertex enumeration problems (En) solved through the algorithm, along with the respective times taken to solve these problems ($\mathrm{T_{opt}}$, $\mathrm{T_{en}}$), as well as the total runtime of the algorithm (T), where $\mathrm{T_{opt}}$, $\mathrm{T_{en}}$ and T are in seconds.



(a) Example 6.1.1 with $\epsilon = 0.01$



(b) Example 6.1.2 with $\epsilon = 0.05$



(c) Example 6.1.3 with $\epsilon = 5$

Figure 6.1: Outer approximations obtained from Algorithm 1 using ℓ_2 norm

Note that we could not run the algorithms in a few settings. We cannot solve Example 6.1.1, $q = 4$ for $\epsilon = 0.1$ by Algorithms 1 and 2 when $p = 2$, and by Algorithms 2 and 3 when $p = \infty$. Similarly, we cannot solve Example 6.1.2 by Algorithm 3 for any $p \in \{1, 2, \infty\}$. Moreover, we cannot solve this example by Algorithm 1 when $p = \infty$ for both ϵ values and when $p = 1$ for $\epsilon = 0.01$. Since it is not possible to provide a comparison, we do not report the results for $p = \infty$ in Table 6.2. Finally, Example 6.1.3 cannot be solved by Algorithms 1 and 2 when $p = 1$. Hence, Table 6.3 does not show the results for $p = 1$ for any setting. The main reason that the algorithms cannot solve these instances is the limitations of *bensolve tools* in vertex enumeration.

6.2 Comparison of Algorithms 1 and 2 with the primal algorithm in [1]

Table 6.1: Computational results for Example 6.1.1

			q=3							q=4						
p	Alg	ϵ	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T	ϵ	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T
1	1	0.05	33	52	13.29	20	0.31	13.68	0.5	30	41	11.66	12	0.22	11.95
	2		42	59	15.59	17	0.26	15.99		57	69	19.27	11	0.28	19.71
	lit[1]		56	89	17.32	34	1.02	18.56		33	44	9.71	12	0.20	9.97
2	1		29	45	10.49	17	0.23	10.79		29	34	8.70	6	0.07	8.80
	2		44	61	14.24	17	0.24	14.68		94	99	25.98	5	0.07	26.20
	lit		32	50	9.76	19	0.27	10.10		31	42	9.18	12	0.19	9.43
∞	1		21	34	8.15	14	0.16	8.35		8	9	2.34	2	0.02	2.38
	2		37	51	12.39	13	0.15	12.61		11	15	4.23	1	0.02	4.39
	lit		21	34	6.76	14	0.15	6.96		8	9	2.05	2	0.02	2.09
1	1	0.01	175	262	69.13	88	20.17	92.99	0.1	143	177	52.94	34	2.67	56.25
	2		161	235	62.55	73	10.56	75.28		232	273	78.37	38	5.05	84.42
	lit		256	397	76.54	142	113.29	212.22		412	510	111.49	91	48.25	165.40
2	1		128	196	46.51	69	8.41	56.52					-		
	2		145	209	49.42	64	6.56	57.29					-		
	lit		139	213	41.49	75	10.39	53.88		208	265	57.93	46	5.08	63.67
∞	1		93	145	35.34	53	3.44	39.47		68	82	22.59	12	0.22	22.87
	2		107	154	37.20	47	2.43	40.15					-		
	lit		87	137	26.72	51	3.03	30.36					-		

Table 6.2: Computational results for Example 6.1.2

ϵ	p	Alg	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T
0.05	1	1	188	310	107.26	87	19.30	130.34
		2	157	233	79.79	70	9.88	91.94
	2	1	145	225	70.22	76	11.46	84.07
		2	141	206	61.72	64	7.18	70.47
0.01	1	1				-		
		2	772	1187	401.00	340	3197.85	4276.37
	2	1	869	1421	420.01	311	2302.12	3171.41
		2	655	957	285.04	279	1529.54	2162.02

In line with the theory, Tables 6.1-6.3 illustrate that Opt, En as well as T increase when a smaller approximation error is used, irrespective of the algorithm considered.

According to Table 6.1, for $p = 1$ in terms of all performance measures, Algorithms 1 and 2 perform better than the referenced algorithm, except for $q = 4, \epsilon = 0.5$ for which Algorithm 1 and the referenced algorithm have similar

[1]Primal Algorithm, [1].

Table 6.3: Computational results for Example 6.1.3

			n=3						n=9					
ϵ	p	Alg	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T
10	2	1	502	943	285.43	132	100.12	401.95	1561	2754	1159.60	225	753.05	2046.88
		2	958	3924	1194.30	137	122.28	1339.09	1770	4213	1819.70	218	733.58	2682.37
		lit	305	965	259.85	164	211.23	502.43	2718	4520	1772.08	259	1324.24	3295.96
	inf	1	197	592	179.71	106	41.61	227.61	1231	2106	901.81	152	150.43	1076.14
		2	199	1206	390.99	100	36.67	432.99	3638	9222	4045.45	166	219.60	4301.63
		lit	180	586	157.46	101	33.60	196.24	2628	5057	1994.39	164	202.33	2231.07
5	2	1	1178	3127	932.65	245	1059.10	2175.25	4461	7968	3371.91	390	8008.67	12795.23
		2	1207	5557	1702.79	259	1488.53	3441.28	7052	15662	6546.29	409	9908.03	18268.17
		lit	579	3932	1049.79	309	3046.66	4529.37	7046	11149	4307.21	476	16898.76	23603.25
	inf	1	412	1740	526.33	185	325.59	907.19	3153	4538	1889.06	294	2164.45	4390.16
		2	465	2655	837.24	188	374.95	1268.01	3570	8155	3482.44	305	2589.99	6470.71
		lit	342	1412	380.10	185	352.70	787.48	3146	4712	1845.91	300	2455.39	4663.98

performances. For $p \in \{2, \infty\}$, Algorithm 1 and the referenced algorithm perform similar to each other and better than Algorithm 2 in terms of Opt, T_{opt} and T. When we compare Algorithms 1 and 2, we observe that Algorithm 2 solves a larger number of optimization problems (Opt) compared to Algorithm 1 in all settings except $p = 1, \epsilon = 0.01$. The reason may be that the former algorithm deals with a higher number of vertices, coming from the intersection of $\bar{\mathcal{P}}_k^{out}$ with $\operatorname{bd} S$, $k \geq 0$ (line 19 of Algorithm 2).

Table 6.2 indicates that, in solving Example 6.1.2, Algorithm 2 performs better than Algorithm 1 with respect to all indicators.

Finally, for Example 6.1.3 when we compare the performances under $n = 3$ with $p = 2$, Algorithm 1 works better compared to the others in terms of Opt, En and T. Under $n = 3$ with $p = \infty$, the same holds for the referenced algorithm, see Table 6.3. However, under $n = 9$, Algorithm 1 performs better than the others in all instances.

When we compare the results for Example 6.1.3(a) and (b), we observe that even with the same precision level ϵ, the number of minimizers is at least twice as and generally much higher than the number of minimizers in Example 8.3(a), which also affects the total times. The reason may be that due to the increase in the dimension of the feasible region and the structure of the objective functions, the range of the objective function changes and the difficulty of the problem increases in Example 8.3(b).

6.2.1 Example 6.1.1 with varying ordering cones

We consider Example 6.1.1 for $q \in \{2, 3\}$ with different ordering cones than the positive orthant, see Table 6.4 and Figure 6.2. These cones are given below in terms of their generating vectors:

$$\begin{aligned}
C_1 &= \operatorname{conv}\operatorname{cone}\{(1,2)^\mathsf{T}, (2,1)^\mathsf{T}\}, \\
C_2 &= \operatorname{conv}\operatorname{cone}\{(2,-1)^\mathsf{T}, (-1,2)^\mathsf{T}\}, \\
C_3 &= \operatorname{conv}\operatorname{cone}\{(4,2,2)^\mathsf{T}, (2,4,2)^\mathsf{T}, (4,0,2)^\mathsf{T}, (1,0,2)^\mathsf{T}, (0,1,2)^\mathsf{T}, (0,4,2)^\mathsf{T}\}, \\
C_4 &= \operatorname{conv}\operatorname{cone}\{(-1,-1,3)^\mathsf{T}, (2,2,-1)^\mathsf{T}, (1,0,0)^\mathsf{T}, (0,-1,2)^\mathsf{T}, (-1,0,2)^\mathsf{T}, \\
&\qquad (0,1,0)^\mathsf{T}\}.
\end{aligned}$$

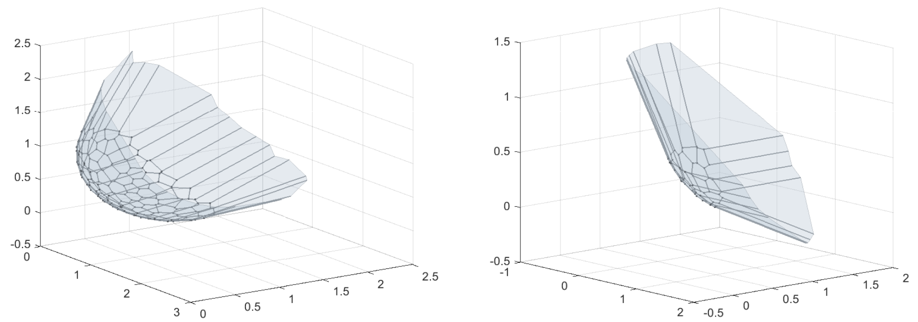
Note that we have $C_1 \subsetneq \mathbb{R}^2_+ \subsetneq C_2 = C_1^+$ and C_3[2] $\subsetneq \mathbb{R}^3_+ \subsetneq C_4 = C_3^+$. We solve these examples with Algorithms 1, 2 and from [1], where the norm in (P(v)) is the ℓ_2 norm. As before, due to the limitations of *bensolve tools*, Table 6.4 does not show the result for Algorithm 2 when the ordering cone is C_3 and $\epsilon = 0.01$. According to Table 6.4, for C_1 and C_2, Algorithms 1 and 2 are comparable in terms of T. For C_3 with $\epsilon = 0.05$, Algorithm 2 gives smaller T. However, for C_4, Algorithm 1 has better runtime. Keeping everything constant (p, q, ϵ and Algorithm), $|\bar{\mathcal{X}}|$ decreases for larger ordering cones, which agrees with [26, Lemma 1.6].

From the results of the test problems above, we deduce a comparable performance of our proposed algorithms with respect to the primal algorithm in [1]. The performance of Algorithm 1 further improves with the increase in the dimension of the objective space.

We conclude this section by a remark that illustrates the necessity of intersecting $\bar{\mathcal{P}}_k^{\text{out}}$ with S in Algorithm 2.

Remark 6.2.1. *As noted before in Section 4.2, it is possible that some vertices of $\bar{\mathcal{P}}_k^{out}$ falls outside S. Consider Example 6.1.1 with $q = 3$. Note that $\Gamma(\mathcal{X})$ is the unit ball centered at $e \in \mathbb{R}^3$, and $\mathcal{P}_0$ is the positive orthant. In the initialization*

[2] The same cone is used as a dual cone in [66, Example 9].



(a) C_3 (b) C_4

Figure 6.2: Outer approximations obtained from Algorithm 1 using ℓ_2 norm for Example 6.1.1 with $q = 3$ under different cones

Table 6.4: Computational results for Example 6.1.1 with different ordering cones and $p = 2$

	$q=2$									$q=3$							
Alg	ϵ	C	$\lvert\bar{\mathcal{X}}\rvert$	Opt	$\mathrm{T_{opt}}$	En	$\mathrm{T_{en}}$	T	ϵ	C	$\lvert\bar{\mathcal{X}}\rvert$	Opt	$\mathrm{T_{opt}}$	En	$\mathrm{T_{en}}$	T
1	0.005	C_1	19	34	6.95	16	0.12	7.12	0.05	C_3	62	89	20.54	28	0.79	21.50
2			21	36	7.51	15	0.11	7.69			57	77	18.27	18	0.35	18.82
1		C_2	6	9	1.87	4	0.02	1.91		C_4	22	29	6.57	8	0.08	6.68
2			8	11	2.30	3	0.02	2.36			28	34	7.90	4	0.04	8.01
1	0.001	C_1	37	69	14.39	32	0.45	14.96	0.01	C_3	229	346	80.66	117	62.00	154.04
2			36	67	14.09	31	0.44	14.69						-		
1		C_2	10	17	3.44	8	0.05	3.52		C_4	71	107	24.99	37	1.45	26.78
2			12	19	4.04	7	0.04	4.13			90	123	28.94	31	1.05	30.31

phase of Algorithm 2, we obtain $S = \{y \in \mathbb{R}^3 \mid \bar{w}^\mathsf{T} y \leq 3.56\}$, *where* $\bar{w} = \frac{1}{\sqrt{3}}e$. *For illustrative purposes, consider the supporting halfspace* $\mathcal{H} = \{y \in \mathbb{R}^q \mid w^\mathsf{T} y \geq 0.68\}$ *of the upper image, where the normal direction is* $w = (1, 1, 0.1)^\mathsf{T}$. *This would support the upper image at the* C*-minimal point* $(0.2947, 0.2947, 0.9295)^\mathsf{T}$. *Note that* $\operatorname{bd}\mathcal{H}$ *intersects with* $\mathcal{P}_0$ *to give three vertices:* $v^1 = (0.68, 0, 0)^\mathsf{T}$, $v^2 = (0, 0.68, 0)^\mathsf{T}$, $v^3 = (0, 0, 6.8)^\mathsf{T}$. *Clearly,* $v^1, v^2 \in S$ *and* $v^3 \notin S$. *Moreover, as stated in Remark 4.2.5 the approximation generated by* v^1, v^2, *namely,* $\operatorname{conv}(\{v^1, v^2\}) + \mathbb{R}^3_+$ *does not contain the upper image; see Figure 6.3.*

We practically encounter vertices which fall outside S, *for instance, while running Algorithm 1 for Example 6.1.1 with* $q = 3$, $p = 2$ *and* $\epsilon = 0.01$. *Figure 6.4 shows the outer approximation, after iteration* $k = 37$, *with one of the vertices outside* $\mathcal{P}_0^{out} \cap S$.

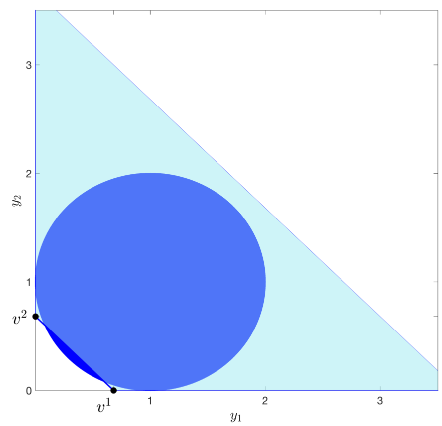


Figure 6.3: Projections of $\Gamma(\mathcal{X})$ *(dark blue),* $\mathrm{conv}(\{v^1, v^2\}) + \mathbb{R}^3_+$ *(light blue) from Remark 6.2.1 on the* $y_3 = 0$ *plane*

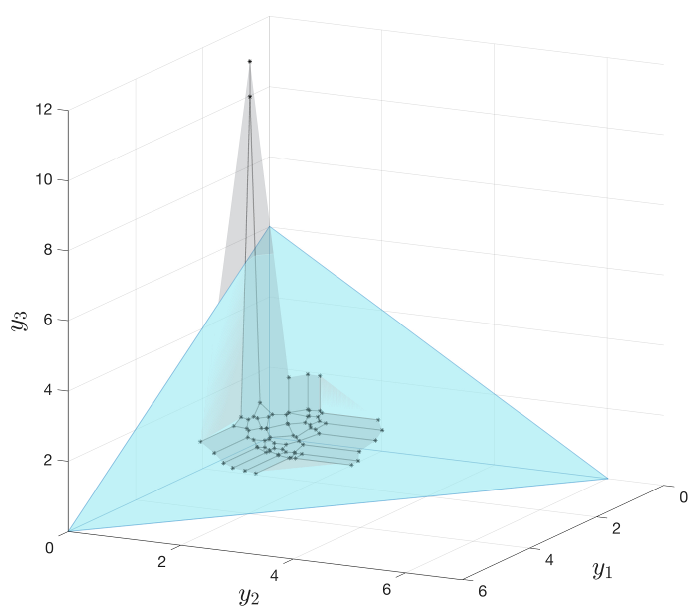


Figure 6.4: Outer approximation $\mathcal{P}_k^{out}$*,* $k = 37$ *(vertices and representative points on unbounded edges indicated by black markers) for Example 6.1.1 obtained using Algorithm 1 (for* $p = 2$ *and* $\epsilon = 0.01$*) has one vertex outside* S *(blue).*

6.3 Comparison of Algorithms 1, 2 and 3

In this section, we compare the three algorithms proposed in Chapter 4. Recall that in Algorithm 1 and 2, we choose arbitrary vertex in each iteration in contrast to choosing farthest away vertex in Algorithm 3. In order to have a fair comparison, we modify Algorithm 1 and 2 by implementing the vertex selection rule as we have in Algorithm 3. The following Tables 6.5-6.7 show the results corresponding to Examples 6.1.1, 6.1.2 and 6.1.3(a), respectively. Note that, in contrast to Algorithm 1, Algorithms 2 and 3 both generate additional vertices due to the intersection of $\operatorname{bd} S$ with the current outer approximation. So, they are expected to solve more optimization problems than Algorithm 1. Next, it is also observed that Algorithm 3 solves even a larger number of optimization problems compared to Algorithm 2. This may be due to the fact that Algorithm 2 is approximating $\mathcal{P}$, whereas Algorithm 3 provides an outer approximation of $\mathcal{P} \cap S$. Clearly, the distance of a vertex of the current approximation, from the set being approximated, is greater in case of Algorithm 3, that is,

$$d(v, \mathcal{P}) \leq d(v, \mathcal{P} \cap S).$$

Hence, Algorithm 3 has to solve larger number of problems to achieve the same approximation error ϵ.

Table 6.5: Computational results for Example 6.1.1

ϵ	p	Alg	$\lvert\bar{\mathcal{X}}\rvert$	Opt	$\mathrm{T_{opt}}$	En	$\mathrm{T_{en}}$	T
0.05	1	1	33	69	23.93	20	0.41	40.70
		2	45	85	28.01	19	0.39	45.80
		3	60	113	41.57	28	0.95	73.38
	2	1	30	60	18.98	17	0.85	35.97
		2	44	93	27.64	16	0.29	46.91
		3	70	122	37.57	25	0.71	60.61
	inf	1	30	48	15.22	14	0.19	23.27
		2	39	65	15.45	13	0.19	25.98
		3	51	85	20.86	22	0.49	35.97

Table 6.6: Computational results Example 6.1.2

ϵ	p	Alg	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T
0.05	1	1	184	522	153.13	86	27.37	238.61
		2	228	660	285.58	85	23.05	447.48
		3	664	1189	527.13	122	96.77	871.16
	2	1	227	561	165.35	79	18.72	244.03
		2	207	475	139.24	78	20.09	196.91
		3	657	1213	277.44	105	59.51	424.63

Table 6.7: Computational results Example 6.1.3

ϵ	p	Alg	$\lvert\bar{\mathcal{X}}\rvert$	Opt	T_{opt}	En	T_{en}	T
100	1	1	80	397	203.43	28	0.85	281.56
		2	117	597	309.88	27	0.92	434.84
		3	-					
	2	1	71	228	86.33	19	0.38	125.19
		2	107	385	188.41	18	0.41	272.42
		3	180	705	356.73	31	1.76	529.34
	inf	1	24	93	35.35	13	0.17	44.67
		2	27	136	73.01	12	0.22	109.39
		3	54	281	160.71	20	0.48	223.03

Chapter 7

Conclusion

In this dissertation, we have proposed an algorithm (Algorithm 1) for CVOPs which is based on a norm-minimizing scalarization. It is different from the similar class of algorithms available in the literature in the sense that it does not need a direction parameter as an input. We have also proposed a modification of the algorithm (Algorithm 2) and proved, for the first time, its finiteness under the assumption of compact feasible region. Moreover, we have proposed another algorithm (Algorithm 3) based on a modified norm-minimizing scalarization. The modification is done by introducing an additional constraint in the norm-minimizing scalarization.

Algorithm 3 provides an H-sequence of outer approximating polytopes to a certain compact set in the image space, which contains the Pareto frontier of the corresponding CVOP. Thus, using a result from the literature, we estimate the convergence rate of Algorithm 3. To the best of our knowledge this is the first CVOP algorithm with a convergence rate guarantee in the literature. We further analyze the algorithm, using the Euclidean norm in the modified scalarization, and found even a better estimate of the convergence rate under this assumption.

As a future research direction, the result of the improved estimate of convergence rate of Algorithm 3 may be generalized for an arbitrary norm used in the scalarization. Note that, Algorithm 3 can also be used for a polyhedral outer approximation of convex compact bodies. Using benchmark test problems, the computational performance of the new algorithms is found to be comparable to a CVOP algorithm in the recent literature which uses the Pascoletti-Serafini scalarization.

Bibliography

[1] A. Löhne, B. Rudloff, and F. Ulus, "Primal and dual approximation algorithms for convex vector optimization problems," *Journal of Global Optimization*, vol. 60, no. 4, pp. 713–736, 2014.

[2] T. Koopmans, "Activity analysis of production and allocation," 1951.

[3] F. Y. Edgeworth, "Mathematical psychics. kegan paul, london," 1881.

[4] V. Pareto, *Cours d'economie politique.* F. Rouge, Lausanne, 1896.

[5] A. Löhne, *Vector Optimization with Infimum and Supremum.* Springer Science & Business Media, 2011.

[6] R. I. Bot, S.-M. Grad, and G. Wanka, *Duality in vector optimization.* Springer Science & Business Media, 2009.

[7] J. P. Evans and R. E. Steuer, "A revised simplex method for linear multiple objective programs," *Mathematical Programming*, vol. 5, no. 1, pp. 54–72, 1973.

[8] B. Rudloff, F. Ulus, and R. Vanderbei, "A parametric simplex algorithm for linear vector optimization problems," *Mathematical Programming*, vol. 163, no. 1-2, pp. 213–242, 2017.

[9] H. P. Benson, "An outer approximation algorithm for generating all efficient extreme points in the outcome set of a multiple objective linear programming problem," *Journal of Global Optimization*, vol. 13, no. 1, pp. 1–24, 1998.

[10] A. Löhne and B. Weißing, "The vector linear program solver bensolve–notes on theoretical background," *European Journal of Operational Research*, vol. 260, no. 3, pp. 807–813, 2017.

[11] M. Ehrgott and M. M. Wiecek, "Mutiobjective programming," in *Multiple Criteria Decision Analysis: State of the Art Surveys* (J. Figueira, S. Greco, and M. Ehrgott, eds.), pp. 667–708, Springer, 2005.

[12] J. Jahn, *Vector Optimization*. Springer, 2nd ed., 2011.

[13] G. K. Kamenev, "A class of adaptive algorithms for approximating convex bodies by polyhedra," *Computational mathematics and mathematical physics*, vol. 32, no. 1, pp. 114–127, 1992.

[14] S. Ruzika and M. M. Wiecek, "Approximation methods in multiobjective programming," *Journal of Optimization Theory and Applications*, vol. 126, no. 3, pp. 473–501, 2005.

[15] A. Chinchuluun and P. M. Pardalos, "A survey of recent developments in multiobjective optimization," *Annals of Operations Research*, vol. 154, no. 1, pp. 29–50, 2007.

[16] R. Kasimbeyli, "A nonlinear cone separation theorem and scalarization in nonconvex vector optimization," *SIAM Journal on Optimization*, vol. 20, no. 3, pp. 1591–1619, 2010.

[17] K. Klamroth and J. Tind, "Constrained optimization using multiple objective programming," *Journal of Global Optimization*, vol. 37, no. 3, pp. 325–355, 2007.

[18] P. M. Pardalos, A. Žilinskas, and J. Žilinskas, *Non-convex multi-objective optimization*. Springer, 2017.

[19] S. Nobakhtian and N. Shafiei, "A Benson type algorithm for nonconvex multiobjective programming problems," *TOP*, vol. 25, no. 2, pp. 271–287, 2017.

[20] M. Ehrgott, L. Shao, and A. Schöbel, “An approximation algorithm for convex multi-objective programming problems,” *Journal of Global Optimization*, vol. 50, no. 3, pp. 397–416, 2011.

[21] M. De Santis, G. Eichfelder, J. Niebling, and S. Rocktäschel, “Solving multiobjective mixed integer convex optimization problems,” *SIAM Journal on Optimization*, vol. 30, no. 4, pp. 3122–3145, 2020.

[22] Ç. Ararat, Ö. Çavuş, and A. İ. Mahmutoğulları, “Multi-objective risk-averse two-stage stochastic programming problems.” arXiv preprint 1711.06403, 2017.

[23] Z. Feinstein and B. Rudloff, “A recursive algorithm for multivariate risk measures and a set-valued bellman’s principle,” *Journal of Global Optimization*, vol. 68, 2017.

[24] B. Rudloff and F. Ulus, “Certainty equivalent and utility indifference pricing for incomplete preferences via convex vector optimization,” *Mathematics and Financial Economics*, vol. 15, pp. 397–430, 2021.

[25] D. Dörfler, A. Löhne, C. Schneider, and W. Benjamin, “A Benson-type algorithm for bounded convex vector optimization problems with vertex selection,” *Optimization Methods and Software*, 2021.

[26] G. Eichfelder, *Adaptive Scalarization Methods in Multiobjective Optimization*. Springer, 2008.

[27] R. Kasimbeyli, Z. K. Ozturk, N. Kasimbeyli, G. D. Yalcin, and B. I. Erdem, “Comparison of some scalarization methods in multiobjective optimization,” *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 42, no. 5, pp. 1875–1905, 2019.

[28] R. S. Burachik, C. Y. Kaya, and M. M. Rizvi, “A new scalarization technique and new algorithms to generate pareto fronts,” *SIAM Journal on Optimization*, vol. 27, no. 2, pp. 1010–1034, 2017.

[29] C. G. Liu, K. F. Ng, and W. H. Yang, “Merit functions in vector optimization,” *Mathematical Programming*, 2009.

[30] A. Pascoletti and P. Serafini, "Scalarizing vector optimization problems," *Journal of Optimization Theory and Applications*, vol. 42, no. 4, pp. 499–524, 1984.

[31] J. G. Lin, "On min-norm and min-max methods of multi-objective optimization," *Mathematical Programming*, vol. 103, no. 1, pp. 1–33, 2005.

[32] M. Zeleny, "Compromise programming," in *Multiple Criteria Decision Making* (M. Zeleny and J. L. Cochrane, eds.), pp. 373–391, University of South Carolina Press, 1973.

[33] M. Ehrgott, *Multicriteria Optimization*, vol. 491. Springer Science & Business Media, 2005.

[34] K. Miettinen, *Nonlinear Multiobjective Optimization*, vol. 12. Springer Science & Business Media, 2012.

[35] B. Schandl, K. Klamroth, and M. M. Wiecek, "Norm-based approximation in multicriteria programming," *Computers & Mathematics with Applications*, vol. 44, no. 7, pp. 925–942, 2002.

[36] B. Schandl, K. Klamroth, and M. M. Wiecek, "Introducing oblique norms into multiple criteria programming," *Journal of Global Optimization*, vol. 23, no. 1, pp. 81–97, 2002.

[37] A. V. Lotov, V. A. Bushenkov, and G. K. Kamenev, *Interactive decision maps: Approximation and visualization of Pareto frontier*, vol. 89. Springer Science & Business Media, 2013.

[38] P. M. Gruber, "Aspects of approximation of convex bodies," in *Handbook of convex geometry*, pp. 319–345, Elsevier, 1993.

[39] E. M. Bronshteyn and L. D. Ivanov, "The approximation of convex sets by polyhedra," *Siberian Mathematical Journal*, vol. 16, no. 5, pp. 852–853, 1975.

[40] R. M. Dudley, "Metric entropy of some classes of sets with differentiable boundaries," *Journal of Approximation Theory*, vol. 10, no. 3, pp. 227–236, 1974.

[41] P. M. Gruber, "Approximation of convex bodies," in *Convexity and its Applications*, pp. 131–162, Springer, 1983.

[42] G. K. Kamenev, "An algorithm for approximating polyhedra," *Computational mathematics and mathematical physics*, vol. 4, no. 36, pp. 533–544, 1996.

[43] G. K. Kamenev, "Conjugate adaptive algorithms for polyhedral approximation of convex bodies," *Computational Mathematics and Mathematical Physics*, vol. 42, no. 9, pp. 1301–1316, 2002.

[44] R. Schneider and J. A. Wieacker, "Approximation of convex bodies by polytopes," *Bulletin of the London Mathematical Society*, vol. 13, no. 2, pp. 149–156, 1981.

[45] G. K. Kamenev, "Study of an adaptive single-phase method for approximating the multidimensional Pareto frontier in nonlinear systems," *Computational Mathematics and Mathematical Physics*, vol. 49, no. 12, pp. 2006–2016, 2009.

[46] G. K. Kamenev, "Study of convergence rate and efficiency of two-phase methods for approximating the edgeworth-pareto hull," *Computational Mathematics and Mathematical Physics*, vol. 53, no. 4, pp. 375–385, 2013.

[47] A. Conci and C. Kubrusly, "Distances between sets - a survey," *Advances in Mathematical Sciences and Applications*, vol. 26, no. 1, 2017.

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

[49] F. Heyde and A. Löhne, "Solution concepts in vector optimization: a fresh look at an old story," *Optimization*, vol. 60, no. 12, pp. 1421–1440, 2011.

[50] J. Jahn, "Scalarization in vector optimization," *Mathematical Programming*, vol. 29, no. 2, pp. 203–218, 1984.

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

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

[53] Ç. Ararat, F. Ulus, and M. Umer, "A norm minimization-based convex vector optimization algorithm," *arXiv preprint arXiv:2104.10282*, 2021.

[54] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems.* Springer Science & Business Media, 2007.

[55] R. T. Rockafellar, *Convex Analysis*, vol. 28. Princeton university press, 1970.

[56] D. T. Luc, *Theory of Vector Optimization.* Springer, 1989.

[57] B. S. Mordukhovich and N. M. Nam, "An easy path to convex analysis and applications," *Synthesis Lectures on Mathematics and Statistics*, vol. 6, no. 2, pp. 1–218, 2013.

[58] A. Löhne and A. Wagner, "Solving DC programs with a polyhedral component utilizing a multiple objective linear programming solver," *Journal of Global Optimization*, vol. 69, no. 2, pp. 369–385, 2017.

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

[60] H. P. Benson and R. Horst, "A branch and bound-outer approximation algorithm for concave minimization over a convex set," *Computers & Mathematics with Applications*, vol. 21, no. 6-7, pp. 67–76, 1991.

[61] F. Burns, M. Fiedler, and E. Haynsworth, "Polyhedral cones and positive operators," *Linear Algebra and its Applications*, vol. 8, no. 6, pp. 547–559, 1974.

[62] H. Anton and C. Rorres, *Elementary linear algebra: applications version.* John Wiley & Sons, 2013.

[63] M. C. Grant and S. P. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.2.," 2020.

[64] M. C. Grant and S. P. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (V. D. Blondel, S. P. Boyd, and H. Kimura, eds.), Lecture Notes in Control and Information Sciences, pp. 95–110, Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.

[65] K. Miettinen, M. M. Mäkelä, and K. Kaario, "Experiments with classification-based scalarizing functions in interactive multiobjective optimization," *European Journal of Operational Research*, vol. 175, no. 2, pp. 931–947, 2006.

[66] A. Löhne and B. Weißing, "Equivalence between polyhedral projection, multiple objective linear programming and vector linear programming," *Mathematical Methods of Operations Research*, vol. 84, no. 2, pp. 411–426, 2016.