Computation and analysis of Pareto critical sets in smooth and nonsmooth multiobjective optimization

Der Fakultät für Elektrotechnik, Informatik und Mathematik der Universität Paderborn zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.) vorgelegte Dissertation von

Bennet Gebken

Paderborn 2021

Abstract

Multiobjective optimization is concerned with the simultaneous optimization of multiple scalar-valued functions. In this case, a point is optimal if there is no other point that is at least as good in all objectives and better in at least one objective. A necessary condition for optimality can be derived based on first-order information of the objectives. The set of points that satisfy this necessary condition is called the Pareto critical set. This thesis presents new results about Pareto critical sets for smooth and nonsmooth multiobjective optimization problems, both in terms of their efficient computation and structural properties.

We begin by deriving new continuation methods for smooth objective functions based on coverings with hypercubes. For the case where the (exact) gradients of the objectives are available, we propose a method that computes a covering of the Pareto critical set based on its smoothness properties and then extend it to the case of constrained problems. For the case where only inexact gradients are available together with upper bounds for the error, we derive a tight superset of the Pareto critical set and then show how the superset can be approximated numerically. Afterwards, we consider the solution of multiobjective optimization problems where the objectives are merely locally Lipschitz continuous, i.e., potentially nonsmooth. Here, we propose a new descent method and show its convergence to Pareto critical points. A comparison to the proximal bundle method, which is currently regarded as the most efficient solution method for nonsmooth problems, suggests that the performance of our method is competitive.

After discussing the computation of Pareto critical sets, we shift our view to their analytical structure. For a certain class of well-behaved objective functions, it is well-known that the Pareto critical set is diffeomorphic to the standard simplex. In this case, the boundary of the Pareto critical set consists of points which are also Pareto critical for subproblems in which only a subset of the set of objectives is considered. We show that this result about the boundary can be generalized to a more general class of objective functions. Furthermore, we present first results about the extension to the case of constrained problems and nonsmooth problems. Finally, we use the structural results about Pareto critical sets to solve inverse multiobjective optimization problems, where a data set is given and a set of objective functions is sought for which the data points are Pareto critical.

Zusammenfassung

Mehrzieloptimierung behandelt Probleme, bei denen mehrere skalare Zielfunktionen simultan optimiert werden sollen. Ein Punkt ist in diesem Fall optimal, wenn es keinen anderen Punkt gibt, der mindestens genauso gut ist in allen Zielfunktionen und besser in mindestens einer Zielfunktion. Ein notwendiges Optimalitätskriterium lässt sich über Ableitungsinformationen erster Ordnung der Zielfunktionen herleiten. Die Menge der Punkte, die dieses notwendige Kriterium erfüllen, wird als Pareto-kritische Menge bezeichnet. Diese Arbeit enthält neue Resultate über Pareto-kritische Mengen für glatte und nicht-glatte Mehrzieloptimierungsprobleme, sowohl was deren Berechnung betrifft als auch deren Struktur.

Es werden zunächst neue Fortsetzungsverfahren hergeleitet für glatte Zielfunktionen, die auf Überdeckungen durch Hyperwürfel basieren. Für den Fall, dass (exakte) Gradienten verfügbar sind, wird eine Methode vorgestellt, die eine Überdeckung der Pareto-kritischen Menge basierend auf deren Glattheitseigenschaften berechnet. Danach wird diese Methode auf beschränkte Probleme erweitert. Für den Fall, dass nur inexakte Gradienten verfügbar sind, zusammen mit oberen Schranken für den Fehler, wird zunächst eine Obermenge der Pareto-kritischen Menge hergeleitet. Anschließend wird gezeigt, wie sich diese Menge numerisch berechnen lässt. Nach dem glatten Fall werden Mehrzieloptimierungsprobleme betrachtet, bei denen die Zielfunktionen lediglich lokale Lipschitz-stetig sind, also potenziell nicht-glatt. Für deren Lösung wird ein neues Abstiegsverfahren vorgestellt, für das Konvergenz gegen Pareto-kritische Punkte gezeigt wird. Ein Vergleich mit der Proximal Bundle Method, die momentan als effizienteste Lösungsmethode für nicht-glatte Probleme gilt, zeigt, dass unsere Methode konkurrenzfähig ist bezüglich ihrer Effizienz.

Nach der Berechnung Pareto-kritischer Mengen wird deren analytische Struktur behandelt. Für eine bestimmte Klasse gutartiger Zielfunktionen ist bekannt, dass die Pareto-kritische Menge diffeomorph zum Standardsimplex ist. In diesem Fall besteht der Rand der Pareto-kritischen Menge aus Punkten, die auch Pareto-kritisch sind für Subprobleme, bei denen nur ein Teil der Zielfunktionen betrachtet wird. In dieser Arbeit wird gezeigt, dass sich dieses Resultat über die Struktur des Randes auf eine allgemeinere Klasse von Zielfunktionen erweitern lässt. Darüber hinaus werden erste Resultate zur Erweiterung auf den beschränkten Fall und den nicht-glatten Fall präsentiert. Schlussendlich werden die Resultate zur Struktur von Pareto-kritischen Mengen dazu genutzt, inverse Mehrzieloptimierungsprobleme zu lösen. Hier geht es darum Zielfunktionen zu finden, für die alle Punkte einer gegebenen Datenmenge Pareto-kritisch sind.

Acknowledgments

First of all, I thank Prof. Dr. Michael Dellnitz for supervising my thesis, for supporting me during the past four years and for giving me a long leash with respect to my research. Furthermore, I thank Prof. Dr. Stefan Volkwein for the ongoing collaboration and especially for several opportunities to visit him and his group in Konstanz, which I always thoroughly enjoyed. Moreover, I thank the Deutsche Forschungsgemeinschaft for the financial support via the Priority Programme 1962.

I thank my past and present colleagues Manuel Berkemeier, Katharina Bieker, Raphael Gerlach, Emina Hadzialic, Marianne Kalle, Lukas Lanza, Karin Mora, Feliks Nüske, Sebastian Peitz, Veronika Schulze and Adrian Ziessler, for the fruitful discussions and for always turning the daily work into an enjoyable time.

Finally, I thank my parents for supporting me in all my endeavors.

Contents

1 Introduction

Humans have an inherent drive for doing things in the optimal way. In a race, the person with the lowest time from start to finish wins and on the market, the company that can produce at the lowest cost can offer the lowest price. If the quantity to be optimized, i.e., the time or the price in the previous examples, can be modeled via a scalar function $f : X \to \mathbb{R}$ on some set X, then an optimal solution is a point x^* in X such that there is no other point in X with a smaller function value than $f(x^*)$. While many real-world problems fit into this framework, there are also problems where a single scalar function is not sufficient to model all relevant quantities. For instance, in the economic example from above, the price might not be the only factor that determines whether a product sells or not. An additional factor could be the quality of the product. Since the cheapest version of the product will (in most cases) not be of the highest quality, these two objectives contradict each other and cannot be modeled with a single scalar objective.

Instead, multiple objectives $f_i : X \to \mathbb{R}$, $i \in \{1, \dots, k\}$, have to optimized at the same time or, in other words, an objective vector $f = (f_i)_{i \in \{1,\dots,k\}} : X \to \mathbb{R}^k$ has to be optimized. As in the scalar case, a point x^* in X is said to be optimal if there is no other point with a smaller function value than $f(x^*)$. In this case, this means that there is no point in X that is at least as good as x^* in all objective functions, but strictly better than x^* in at least one objective function. One of the first to introduce this concept of optimality was Vilfredo Pareto in [Par06] (1906), which is why in the presence of multiple objectives, optimal points are called *Pareto optimal*. Unlike optimality in the scalar case, Pareto optimality of x^* does not imply that $f(x^*)$ is less or equal to all other function values of f, as there can be points that are superior in one objective function but inferior in another. Thus, there are generally multiple Pareto optimal points and the set of all these points is called the *Pareto set*. Its image under f is the so-called *Pareto front*, which corresponds to the optimal value in the scalar case. The task of finding the Pareto set is called a *multiobjective optimization problem* (MOP), simply denoted by

$$\min_{x \in X} f(x).$$

There is a wide range of different real-world applications for multiobjective optimization across many different areas. In [Oba+00], it is used for the optimization of the aerodynamic design of a wing for supersonic transport. Here, the objectives are the drag during supersonic cruise, the drag during transonic cruise and the bending moment at the wing root during supersonic cruise, which should all be minimized. In [LSB03], it is used for the planning of radiotherapy treatment for cancer patients. The goals here are to apply a sufficient amount of radiation to kill as many cancerous cells as possible, while minimizing the impact non-cancerous cells. In [Naj+14],

it is used for optimizing the design parameters of a gas turbine, where the goals are the maximization of the efficiency and the minimization of the costs. Finally, in [Del+09; Sch+09], it is used for finding optimal trajectories for orbital and interplanetary transfers. The objectives are to minimize both fuel consumption and flight time.

When it comes to the practical solution of multiobjective optimization problems, there are different philosophies that lead to different solution approaches:

- **Scalarization methods:** The solution of scalar optimization problems is well understood and there are many different solvers for different classes of objective functions. Thus, if the multiobjective problem is transformed into a scalar problem, then existing methods can be used for its solution. The most popular method in this class is the *weighting method* [Mie98], where each objective function is assigned a non-negative weight and the sum of the weighted objectives is minimized. By varying the weights, different Pareto optimal points can be computed.

- **Generalizations of methods from scalar optimization:** Since scalar optimization is a special case of multiobjective optimization, some methods from the scalar case can be generalized to the multiobjective case. For example, descent methods can be generalized by assuring descent for all objectives at the same time, as was done in [FS00] for the steepest descent method and in [FDS09] for Newton's method.

- **Set-based methods:** Since the solution of an MOP is a set, set-based methods were introduced that aim at computing the entire Pareto set instead of just single optimal points. Examples here are the *subdivision method* [DSH05] and the *branch-and-bound method* [NE19], which compute a covering of the Pareto set via boxes, and *continuation methods* [Hil01; SDD05; MS17], which produce an even pointwise discretization of the Pareto set.

- **Evolutionary methods:** As evolutionary computation is generally based on maintaining a set (or "population") of candidate solutions, it can be applied to multiobjective problems by interpreting the population as an approximation of the Pareto set. The approximation is iteratively improved by using stochastic operators that modify and then select points from the population. A popular method from this class is *NSGA-II* [Deb+02].

- **Interactive methods:** Although the solution of a multiobjective problem is a set, there are practical applications where ultimately only a single Pareto optimal point is needed. While it would be possible to first compute the whole Pareto set and then let a human decision maker select one point from it, it can be more efficient to only compute Pareto points that are desirable for the decision maker. To this end, methods were proposed that can be steered interactively during the solution process based on preference. For example, in [Sch+19], continuation-like procedures were proposed that can explore the Pareto set based on a direction specified by the user. For a general survey of interactive methods, see [Bra+08].

Due to these different approaches for solving multiobjective optimization problems, there is a vast amount of different solution methods in addition to the ones mentioned above. (Furthermore, the classes above are not mutually exclusive.)

In practice, the definition of Pareto optimality cannot directly be used to verify whether a given point is Pareto optimal, as it relies on global information of the objective functions. Fortunately, it is possible to derive a necessary optimality condition that is easier to work with. In the scalar case, a necessary condition for a point $x \in X$ to be optimal is the nonexistence of a direction in which f descends. In the multiobjective case, this condition naturally generalizes to the nonexistence of a direction in which all objectives f_i descend simultaneously. If $X = \mathbb{R}^n$ and all objectives are differentiable, then a basic result from convex analysis shows that this condition is equivalent to the existence of a convex combination of the gradients $\nabla f_i(x)$, $i \in \{1, \dots, k\}$, which is zero. (In the scalar case, this reduces to the classical condition of $\nabla f(x)$ being zero.) Points that satisfy this condition are called *Pareto critical* and the set of all those points is the *Pareto critical set*. By construction, the Pareto critical set is a superset of the actual Pareto set that only relies on first-order information of the objective functions. If all objective functions are convex, then it coincides with the Pareto set. This optimality condition was originally published by Kuhn and Tucker in [KT51] and is also referred to as the *Karush-Kuhn-Tucker (KKT) condition*. (See [Kuh82] for a discussion of the history of necessary optimality conditions.) The Pareto critical set was first defined by Smale in [Sma73].

Since the solution of a multiobjective optimization problem is a set, theoretical results about topological and geometrical properties of the Pareto set are crucial for its efficient approximation. The Pareto set itself is difficult to analyze, as its definition is purely based on the order relation on the image space. In contrast, the Pareto critical set only relies on local information and can thus be analyzed easier. The first result about its structure was already given by Smale in [Sma73], where he proposed (without proof) that under certain assumptions, the Pareto critical set is a *stratification*. Roughly speaking, this means that it is a manifold with boundaries and corners. Furthermore, he proposed that under stronger assumptions, the Pareto critical set is homeomorphic to a closed $(k-1)$-simplex. In this case, each facet of the simplex corresponds to the Pareto critical set of a subset of the objective functions. Smales work relies on transversality and stratification theory. An arguably simpler result was given by Hillermeier in [Hil01], where he showed that under a regularity assumption concerning the Hessians of the objectives, part of the Pareto critical set is the projection of a $(k-1)$-dimensional manifold from a higher dimensional space onto the variable space $\mathbb{R}^n$. So although the Pareto critical set is not a smooth manifold, it still has certain smoothness properties.

The characterization of the Pareto critical set given above via the existence of vanishing convex combinations of gradients requires differentiability of the objective functions. However, there are various practical applications where nondifferentiable (or *nonsmooth*) objective functions occur. For example, in the area of image processing, the *total variation* is a nonsmooth expression that is used as a regularization term for image denoising [Cha04], resulting in a nonsmooth objective function. From a mathematical point of view, many of the nonsmooth objective functions that occur in practice fall into the class of locally Lipschitz continuous functions. For example, all piecewise differentiable functions are locally Lipschitz [Sch12]. Since gradients cannot be used to describe the local behavior of these functions, solution methods

for smooth problems generally fail to work in the nonsmooth case [Lem89].

For multiobjective optimization problems with locally Lipschitz functions, Pareto criticality can be generalized via the *Clarke subdifferentials* $\partial f_i(x)$, $i \in \{1, \dots, k\}$, of the objective functions [Cla90]. These are sets that act as generalizations of gradients to the nonsmooth case. A necessary condition for $x \in \mathbb{R}^n$ to be Pareto optimal is that zero is contained in the convex hull of the union of the Clarke subdifferentials of the objective functions in x [MEK14]. Since the Clarke subdifferentials reduce to the classical gradients if the objectives are continuously differentiable, this condition is a generalization of Pareto criticality to the locally Lipschitz case. In general, Clarke subdifferentials are more difficult to compute than gradients, so computing Pareto critical sets for nonsmooth problems is generally more complicated than for smooth problems. On top of that, the nonsmoothness of the objective functions causes the Pareto critical set to have a nonsmooth structure, such that "kinks" may occur. In particular, the structural results for the smooth case discussed earlier do not apply.

As the above considerations demonstrate, working with Pareto critical sets is a diverse task, as it combines optimization, differential geometry, convex analysis and, in the nonsmooth case, nonsmooth analysis. To contribute to this aspect of multiobjective optimization, the goal of this thesis is the study of the Pareto critical set in the smooth and nonsmooth case, both in terms of its efficient computation and the analysis of its structure. We will begin by proposing novel methods for computing Pareto critical sets of smooth and nonsmooth problems. Afterwards, we will analyze the structure of the Pareto critical set, where we focus on the relationship between the Pareto critical set of the original problem and the Pareto critical sets of subproblems where only subsets of the objectives are considered. Finally, these structural results will allow us to consider the *inverse problem of multiobjective optimization*, where a data set is given and the goal is to find an objective vector for which the data points are Pareto critical. The structure of this thesis is as follows.

Chapter 2 introduces the basics of multiobjective optimization. We first discuss the concept of optimality in the presence of multiple objectives before deriving the optimality conditions for both smooth and nonsmooth problems. In the nonsmooth case, this also requires basic definitions from the area of nonsmooth analysis. Afterwards, we discuss existing solution methods for multiobjective optimization problems, with an emphasis on methods that are relevant for the results in this thesis.

In Chapter 3, we propose continuation methods for smooth problems that compute a covering of the Pareto critical set via hypercubes (or *boxes*). We first consider the case where exact gradients are available in Section 3.1, where we derive a method that is related to the continuation method from [SDD05]. Based on a partition of the variable space $\mathbb{R}^n$ into (small) boxes, the goal is to find all boxes that have a nonempty intersection with the Pareto critical set. Due to the smoothness properties of the Pareto critical set discussed earlier, it is possible to compute its tangent vectors. Given a box B that contains Pareto critical points, this allows us to find all neighboring boxes of B that (potentially) contain Pareto critical points by checking if their intersection with the span of the tangent vectors is empty. We first consider the unconstrained case and then propose an extension that is able to handle equality and inequality constrained multiobjective optimization problems. In Section 3.2, we assume that the objective functions are smooth but we only have inexact approximations of their gradients. This setting occurs in practice when the objectives are

computationally expensive to evaluate and are therefore replaced by cheaper surrogate models. Based on upper bounds for the inexactness, we derive a tight superset of the Pareto critical set that only requires evaluation of the inexact gradients. Since this superset is not a null set, we propose a continuation method for its boundary to efficiently compute it.

Chapter 4 is concerned with the solution of multiobjective optimization problems with locally Lipschitz objective functions. For smooth objectives, Fliege and Svaiter proposed a descent method in [FS00] that produces a sequence with Pareto critical accumulation points. The goal of this chapter is to generalize this method to the nonsmooth case. To this end, we require a way to compute descent directions. In [AGG15], it was shown that a descent direction can be obtained as the element with the smallest norm in the negative convex hull of the Clarke subdifferentials of the objective functions. Unfortunately, there are two reasons why this result cannot directly be used in practice. The first reason is that the descent direction relies on knowing the full Clarke subdifferentials, which are generally not available in practice. The second reason is the fact that the Clarke subdifferentials only contain information about the nonsmoothness of the objective functions if we are exactly in a nonsmooth point. Due to *Rademacher's theorem* [EG15], the set of nonsmooth points is a null set, so in practice we cannot assure that we actually encounter points from it. To solve these issues, we replace the Clarke subdifferential by the so-called *(Goldstein) ε-subdifferential* [Gol77], which makes it easier to detect and handle nonsmoothness. Afterwards, we discuss how ε-subdifferentials can be approximated efficiently by the convex hull of a finite number of subgradients. Combining the resulting descent direction with an Armijo-like step length results in a descent method that produces sequences with Pareto critical accumulation points. To evaluate the performance of the descent method, we compare it to the *multiobjective proximal bundle method* [MKW14] using a set of test problems. The results suggest that our method is superior in terms of subgradient evaluations, but inferior in terms of function evaluations. Finally, we combine the descent method with the subdivision algorithm (as in [DSH05]) to compute entire Pareto (critical) sets of locally Lipschitz problems.

For Chapter 5, we shift our view from the computation to the analysis of the structure of Pareto critical sets. In [Sma73; LP14; Pei17], it was shown that if $k \leq n$, the objectives $f_1, \ldots, f_k$ are convex and a certain rank assumption holds, then the Pareto critical set is diffeomorphic to a $(k-1)$-simplex. In this case, each facet of the Pareto critical set corresponds to the Pareto critical set of a subset of the objectives of size $k-1$. The goal of this chapter is to generalize this result about the structure of the boundary of the Pareto critical set to nonconvex problems with an arbitrary number of objectives. In Section 5.1, we analyze the topological and geometrical properties of the Pareto critical set. We begin by classifying Pareto critical points with respect to zero entries in their *KKT vectors*, which are the coefficient vectors of the vanishing convex combinations of the gradients in the KKT condition. Since the natural topology on $\mathbb{R}^n$ cannot be used to describe the boundary, we then consider a definition of the boundary of the Pareto critical set based on tangent cones. The main result in this section is that under a regularity assumption concerning the weighted Hessians of the objective functions, the boundary consists of points that are also Pareto critical for a (proper) subset of the objectives. In Section 5.2, we derive the smallest size of subproblems required to still obtain a covering of the

boundary via Pareto critical sets of subproblems, which turns out to be related to the rank of the Jacobian of the objective vector. After applying our results to some examples in Section 5.3, we consider further generalizations. In Section 5.4, we consider equality and inequality constrained problems and argue how our results can be generalized by interpreting the constraints as additional objective functions. Finally, Section 5.5 is concerned with the structure of Pareto critical sets in the nonsmooth (i.e., locally Lipschitz) case. Although many of the tools from the smooth case cannot be applied here, it is still possible to show some basic results using tools from convex analysis.

The results from Chapter 5 and the existing results about the structure of Pareto critical sets create an intuition for whether a given subset of $\mathbb{R}^n$ can be expressed as the Pareto critical set of some (smooth) objective vector. This leads to the problem of inverse multiobjective optimization in Chapter 6, where the goal is to find an objective vector for which a given data set is Pareto critical. We begin this chapter by discussing which data is required to obtain an inverse problem with useful solutions. It turns out that we have to assume that not only the Pareto critical points in $\mathbb{R}^n$, but also the corresponding KKT vectors are given. In Section 6.1, we show that by using the span of a finite number of basis functions as the search space, the inverse problem can be formulated as a system of equations that is linear in the coefficients of the basis functions. This allows us to efficiently solve the inverse problem with a simple algorithm that is based on a singular value decomposition. In particular, the smallest singular value of the matrix of the system is a measure for how well the data set can be expressed as a Pareto critical set of objective functions in the span of the basis functions. In Section 6.2, we present three applications of our results. These are the generation of test problems, the estimation of objective vectors in stochastic multiobjective optimization and the generation of surrogate models for (potentially expensive) objective vectors. Finally, we discuss open problems of this inverse approach in Section 6.3.

The main contributions of this thesis can be summarized as follows:

- New results about the hierarchical structure of Pareto critical sets are obtained for the general nonconvex case, which generalize some of the results from [Sma73; LP14; Pei17]. In particular, the constrained case and the nonsmooth case are considered, which, to the best of the authors' knowledge, have not been addressed before.

- A new descent method is proposed for nonsmooth multiobjective optimization problems. To the best of the authors' knowledge, this is the first (deterministic) method that generalizes the descent method from [FS00] to the nonsmooth case. Our test results suggest that its performance is competitive to the bundle method from [MKW14], which currently seems to be the most efficient solution method for nonsmooth problems. (In addition, our method is arguably easier to implement in practice.)

- A new framework is proposed for the solution of inverse multiobjective optimization problems. The only similar work so far is [DZ18], where only the convex case is considered with a different, heuristic solution methodology.

- A new continuation method for smooth problems is developed which has certain advantages over the methods in [Hil01; SDD05]. In particular, it can be

generalized to problems with equality and inequality constraints.

- For the case of inexact gradients, a superset of the Pareto critical set is derived which is tighter than the superset suggested in [PD17]. Furthermore, the computation of its boundary via a continuation method is a new approach which is tailored to the theoretical properties of the superset.

Finally, in the following, the content of this thesis is discussed in relation to previous publications of its author.

- In Chapter 3, Section 3.2 is based on Section 2 in [Ban+19], to which the author was the main contributor. While the algorithm in Section 3.1 was also briefly described in [Ban+19], this thesis contains a more detailed discussion. Furthermore, the extension to the constrained case in Section 3.1.4 has not been published before.

- Chapter 4 is based on [GP21a], to which the author was the main contributor. Compared to [GP21a], this thesis additionally contains a modified descent method (Algorithm 4.4) for which stronger convergence results can be shown (Lemma 4.4.4 and Corollary 4.4.5).

- In Chapter 5, the Sections 5.1, 5.2 and 5.3 are based on [GPD19], to which the author was the main contributor. The extensions to the constrained and nonsmooth case in Section 5.4 and Section 5.5, respectively, have not been published previously.

- Chapter 6 is based on [GP21b], to which the author was the main contributor.

2 Basics of multiobjective optimization

In this thesis the task of minimizing a vector-valued function is considered. Formally, let $n, k \in \mathbb{N}$, $\emptyset \neq X \subseteq \mathbb{R}^n$ and $f = (f_1, \ldots, f_k)^\top : X \to \mathbb{R}^k$. The problem of minimizing f is denoted by

$$\min_{x \in X} f(x). \tag{MOP}$$

This problem is called a *multiobjective optimization problem* (MOP) with *objective vector* f, *objective functions* f_i, $i \in \{1, \ldots, k\}$, *variable space* $\mathbb{R}^n$, *image space* $\mathbb{R}^k$ and *feasible set* X. Throughout most of this thesis we will consider the case where $X = \mathbb{R}^n$, i.e., where (MOP) is *unconstrained*.

This chapter introduces the basics of multiobjective optimization which are used in the remainder of this thesis. More detailed introductions can be found in [Mie98; Ehr05]. In Section 2.1, the solution concept of (MOP) is introduced and a simple example is given. In Section 2.2, necessary optimality conditions are derived. This part is split up into the smooth case, where f is differentiable, and the nonsmooth case, where f is merely locally Lipschitz continuous. Additionally, in the smooth case, a first result about the structure of the set of points satisfying the optimality condition is established. Finally, Section 2.3 gives an overview over existing solution methods for MOPs.

2.1 Pareto optimality

For $k = 1$, i.e., for the case where (MOP) is a *single-objective* or *scalar* problem, the concept of optimality is clear: $x \in X$ is called *optimal* if $f(x) \leq f(y)$ for all $y \in X$. In other words, $x \in X$ is optimal if $f(x)$ is a minimal element in the totally ordered set $(\mathrm{im}(f), \leq)$, where $\mathrm{im}(f)$ denotes the image $f(X)$ of f. (For an introduction to ordered sets, see, e.g., [Sch16].) To generalize this concept for $k > 1$, we have to define an order relation on $\mathrm{im}(f) \subseteq \mathbb{R}^k$ that reflects our goal of minimizing all objective functions simultaneously. To this end, for $v, w \in \mathbb{R}^k$, we define

$$v \leq w \;:\Leftrightarrow\; v_i \leq w_i \quad \forall i \in \{1, \ldots, k\}. \tag{2.1}$$

It is easy to see that $(\mathbb{R}^k, \leq)$ is a partially ordered set. But in contrast to $(\mathbb{R}, \leq)$ it is not totally ordered, since there can be $v, w \in \mathbb{R}^k$ such that neither $v \leq w$ nor $w \leq v$. This is visualized in the following example.

Example 2.1.1. *For $k = 2$ let $v^1 = (2.5, 3)^\top$, $v^2 = (0.5, 2)^\top$, $v^3 = (2, 0.5)^\top$ and $v^4 = (3, 0.5)^\top$, as shown in Figure 2.1. Then $v^2 \leq v^1$, $v^3 \leq v^1$ and $v^3 \leq v^4$, but*

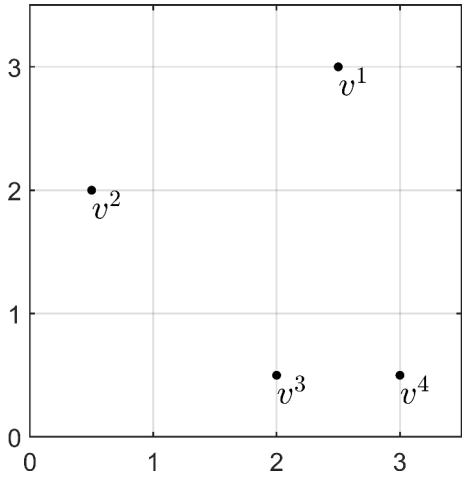


Figure 2.1: Example for the order relation on $\mathbb{R}^k$ for $k = 2$ from Example 2.1.1.

neither $v^2 \leq v^3$ *nor* $v^3 \leq v^2$*. In particular,* v^2 *and* v^3 *are the minimal elements of* $(\{v^1, v^2, v^3, v^4\}, \leq)$*.*

A result of $(\mathbb{R}^k, \leq)$ not being totally ordered for $k > 1$ is the fact that usually, there will be multiple minimal elements in $(\mathrm{im}(f), \leq)$ that have different objective function values. This is something that cannot happen in single-objective optimization and is one of the the fundamental differences between single-objective and multiobjective optimization.

Using the order relation from (2.1) we can now formally define what optimality means in the multiobjective context.

Definition 2.1.2. *A point* $x \in X$ *is called* Pareto optimal *if* $f(x)$ *is a minimal element in* $(\mathrm{im}(f), \leq)$*, i.e., if there is no* $y \in X$ *with* $f(y) \leq f(x)$ *and* $f(y) \neq f(x)$*. In other words,* x *is Pareto optimal if there is no* $y \in X$ *with*

$$\begin{aligned} &f_i(y) \leq f_i(x) \quad \forall i \in \{1, \ldots, k\}, \\ &f_j(y) < f_j(x) \quad \text{for some } j \in \{1, \ldots, k\}. \end{aligned}$$

The set P *of all Pareto optimal points is called the* Pareto set*, its image under* f *is called the* Pareto front*.*

Remark 2.1.3. *a) In the literature, Pareto optimal points are also sometimes referred to as* Edgeworth-Pareto optimal points*,* (Pareto) efficient points *or* nondominated points*. See Table 2.4 in [Ehr05] for a more detailed overview of the common terminology.*

b) Let $C = (\mathbb{R}^{\geq 0})^k = \{v \in \mathbb{R}^k : v_i \geq 0 \ \forall i \in \{1, \ldots, k\}\}$*. Then* C *is a convex, pointed cone (also called the* natural ordering cone*) and* $v^1 \leq v^2$ *for* $v^1, v^2 \in \mathbb{R}^k$ *can be rewritten as* $v^2 \in v^1 + C$*. Due to this, the order relation* $\leq$ *is said to be* induced by C*. In particular,* $x \in X$ *being Pareto optimal can be rewritten as* $(f(x) - C) \cap \mathrm{im}(f) = f(x)$*. For different cones* C *these reformulations lead to a more general solution concept for optimizing vector-valued functions, which is discussed in the area of* vector optimization *[Jah11].*

In addition to the notion of Pareto optimality, we will also make use of the following weaker concepts.

Definition 2.1.4. *Let $x \in X$ (and let X be equipped with the subspace topology of the natural topology of $\mathbb{R}^n$).*

a) The point x is called locally Pareto optimal*, if there is some open set $U \subseteq X$ with $x \in U$ such that $f(x)$ is a minimal element of $(f(U), \leq)$.*

b) The point x is called weakly Pareto optimal*, if there is no $y \in X$ with*

$$f_i(y) < f_i(x) \quad \forall i \in \{1, \dots, k\}. \tag{2.2}$$

c) The point x is called locally weakly Pareto optimal*, if there is some open set $U \subseteq X$ with $x \in U$ such that there is no $y \in U$ that satisfies (2.2).*

By definition, Pareto optimal points are also locally, weakly and locally weakly Pareto optimal, and local Pareto optimal points are also locally weakly Pareto optimal. Furthermore, all (local) minimal points of the individual objective functions are (locally) weakly Pareto optimal, and all unique minimal points are Pareto optimal. The following example visualizes the differences of the different notions of optimality introduced so far.

Example 2.1.5. *For $k = 2$ let $\mathrm{im}(f)$ be given as in Figure 2.2. The red, blue, green and magenta lines show the image of the set of Pareto optimal, weakly Pareto optimal, local Pareto optimal and locally weakly Pareto optimal points, respectively. The colored dots indicate whether the end points of the lines are included or excluded.*

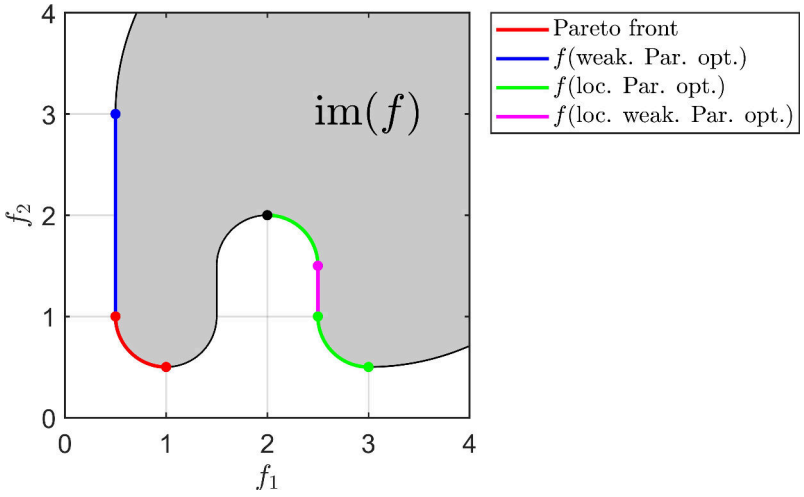


Figure 2.2: Visualization of the different notions of optimality in Example 2.1.5. (Note that the preimage of the point $(3, 0.5)^\top$ is also weakly Pareto optimal.)

Note that since the definitions of local Pareto optimality and local weak Pareto optimality depend on the local behavior of f, the set $\mathrm{im}(f)$ alone cannot be used to identify all those points. Thus, depending on f, there may be more green and magenta lines in the interior of $\mathrm{im}(f)$ in Figure 2.2.

This section will conclude with a simple example where the concept of Pareto optimality produces a solution that matches the "intuitive" solution one would expect. To this end, let $\|\cdot\|$ be the Euclidean norm on $\mathbb{R}^n$.

Example 2.1.6. *For $k \in \mathbb{N}$ let $c^1, \dots, c^k \in \mathbb{R}^n$ and consider the problem*

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{with} \quad f(x) = \begin{pmatrix} \|x - c^1\|^2 \\ \vdots \\ \|x - c^k\|^2 \end{pmatrix}. \tag{2.3}$$

This problem can be interpreted as follows: We are given k points $c^1, \dots, c^k \in \mathbb{R}^n$ and we want to find points $x \in \mathbb{R}^n$ that minimize the (squared) distance to all $c^1, \dots, c^k$ at the same time. It can be regarded as the multiobjective version of a simple facility location problem *[Meg83; BW02].*

We consider the case where $k = 2$, $n = 2$, $c^1 = (0,0)^\top$ and $c^2 = (1,1)^\top$. Let $x \in \mathbb{R}^2$. Clearly, for all $i \in \{1, \dots, k\}$, all points that have a smaller squared distance to c^i than x form an open ball B^i with radius $\|x - c^i\|$ centered at c^i, and all points with an equal distance are given by its boundary ∂B^i. Thus, for x to be weakly Pareto optimal, $B^1 \cap B^2$ has to be empty. It is easy to see that this is precisely the case when x is on the line connecting c^1 and c^2, i.e., when $x \in \{c^1 + \lambda(c^2 - c^1) : \lambda \in (0,1)\} =: P'$. In particular, for $x \in P'$ it also holds $B^1 \cap \partial B^2 = \emptyset$ and $\partial B^1 \cap B^2 = \emptyset$, so all points in P' are actually Pareto optimal. On top of that, c^i is the unique global minimizer of $\min_{x \in \mathbb{R}^2} f_i(x)$, so c^1 and c^2 are Pareto optimal as well. To summarize, the Pareto set of (2.3) is given by

$$P = \{c^1 + \lambda(c^2 - c^1) : \lambda \in [0,1]\},$$

as shown in Figure 2.3.

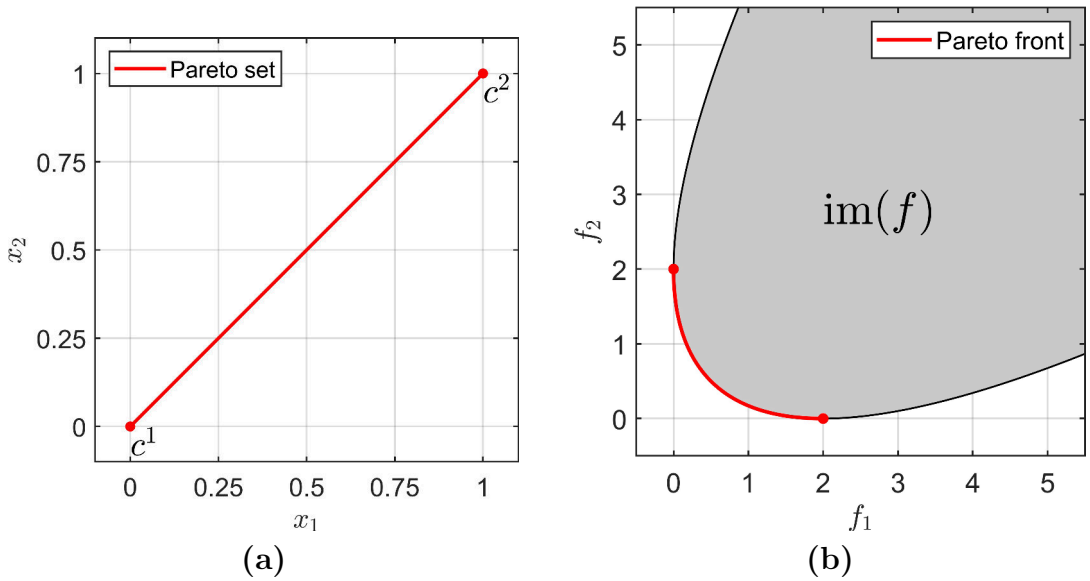


Figure 2.3: Pareto set (a) and Pareto front (b) for the multiobjective location problem in Example 2.1.6.

Analogously to the above discussion, it is easy to see that for the general problem (2.3)*, the Pareto set is given by the convex hull of $\{c^1, \dots, c^k\}$, i.e.,*

$$P = \operatorname{conv}(\{c^1, \dots, c^k\}) := \left\{ \sum_{i=1}^k \lambda_i c^i : \lambda \in (\mathbb{R}^{\geq 0})^k, \sum_{i=1}^k \lambda_i = 1 \right\}. \tag{2.4}$$

Remark 2.1.7. *There are multiple equivalent ways to define the convex hull of $A \subseteq \mathbb{R}^n$. In this thesis we mainly use the following two (cf. [Roc70], Section 2):*

- $\operatorname{conv}(A)$ *is the set of all convex combinations of elements of A, i.e.,*

$$\operatorname{conv}(A) := \left\{ \sum_{i=1}^{m} \lambda_i c^i : m \in \mathbb{N}, \lambda \in (\mathbb{R}^{\geq 0})^m, \sum_{i=1}^{m} \lambda_i = 1, c^i \in A, i \in \{1, \dots, m\} \right\}.$$

- $\operatorname{conv}(A)$ *is the smallest convex set containing A.*

Clearly, the MOP in Example 2.1.6 was only solvable by hand due to its simplicity. For example, merely changing the norms in (2.3) would already make the solution much more involved. Thus, a more sophisticated methodology is needed, which is the goal of the following section.

2.2 Necessary optimality conditions and the Pareto critical set

In Example 2.1.6, the Pareto set was derived by considering the intersection of the *sublevel sets*

$$L_i^<(x) := \{y \in \mathbb{R}^n : f_i(y) < f_i(x)\}$$

over all $i \in \{1, \dots, k\}$. Clearly, x is locally weakly Pareto optimal if and only if

$$U \cap \left(\bigcap_{i \in \{1,\dots,k\}} L_i^<(x) \right) = \emptyset \tag{2.5}$$

for some open set $U \subseteq \mathbb{R}^n$ with $x \in U$. Since U can be chosen arbitrarily small, (2.5) only depends on the local behavior of f. If f is differentiable, then the gradients $\nabla f_i(x)$ are orthogonal to the boundary $\partial L_i^<(x)$ (i.e., to the *level set of* f_i). If f is merely locally Lipschitz, a similar connection can be made using nonsmooth analysis. The goal of this section is to show that in both cases, (2.5) can be reformulated using first-order information of f, resulting in first-order necessary optimality conditions.

The theoretical foundation of this section is the following basic result from convex analysis [CG59]. For a set $W \subseteq \mathbb{R}^n$, let $-W := \{-\xi : \xi \in W\}$.

Lemma 2.2.1. *Let $W \subseteq \mathbb{R}^n$ be convex and compact and*

$$\bar{v} := \operatorname*{arg\,min}_{\xi \in -W} \|\xi\|^2. \tag{2.6}$$

Then either $\bar{v} \neq 0$ and

$$\langle \bar{v}, \xi \rangle \leq -\|\bar{v}\|^2 < 0 \quad \forall \xi \in W, \tag{2.7}$$

or $\bar{v} = 0$ and there is no $v \in \mathbb{R}^n$ with $\langle v, \xi \rangle < 0$ for all $\xi \in W$.

Proof. $\bar{v}$ is well-defined since $\|\cdot\|^2$ is strictly convex and W is convex.
We first assume $\bar{v} \neq 0$. Let $\xi \in W$ and $\lambda \in (0,1)$. Since $\lambda\bar{v} + (1-\lambda)(-\xi) \in -W$ we have

$$\begin{aligned}\|\bar{v}\|^2 &\leq \|\lambda\bar{v} + (1-\lambda)(-\xi)\|^2 = \|\lambda\bar{v} - (1-\lambda)\xi\|^2 \\ &= \lambda^2\|\bar{v}\|^2 - 2\lambda(1-\lambda)\langle\bar{v},\xi\rangle + (1-\lambda)^2\|\xi\|^2.\end{aligned}$$

Subtracting $\|\bar{v}\|^2$ on both sides and dividing by $1-\lambda > 0$ yields

$$0 \leq -(\lambda+1)\|\bar{v}\|^2 - 2\lambda\langle\bar{v},\xi\rangle + (1-\lambda)\|\xi\|^2.$$

Due to continuity, this inequality must also hold for $\lambda = 1$, so

$$\begin{aligned}&0 \leq -2\|\bar{v}\|^2 - 2\langle\bar{v},\xi\rangle \\ \Leftrightarrow\quad &\langle\bar{v},\xi\rangle \leq -\|\bar{v}\|^2 < 0.\end{aligned}$$

Finally, if $\bar{v} = 0$ then $0 \in W$, so in this case there can clearly be no $v \in \mathbb{R}^n$ with $\langle v,\xi\rangle < 0$ for all $\xi \in W$. □

2.2.1 Smooth case

We will begin with the case where f_i is differentiable for all $i \in \{1,\dots,k\}$. In this case, the level sets of f_i can locally be described as the orthogonal complement of $\nabla f_i(x)$. In particular, $-\nabla f_i(x)$ points into the half-space of directions in which f_i descends. Thus, (2.5) can be reformulated to

$$\nexists v \in \mathbb{R}^n : \langle\nabla f_i(x), v\rangle < 0 \ \forall i \in \{1,\dots,k\}.$$

More formally, the following lemma (from [FS00]) holds.

Lemma 2.2.2. *Let $c \in (0,1)$, $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ such that $\langle\nabla f_i(x), v\rangle < 0$ for all $i \in \{1,\dots,k\}$. Then there is some $T > 0$ such that*

$$f_i(x+tv) < f_i(x) + ct\langle\nabla f_i(x), v\rangle \quad \forall t \in (0,T],\ i \in \{1,\dots,k\}.$$

In particular, x is not (locally weakly) Pareto optimal.

Proof. Since f_i is differentiable for all $i \in \{1,\dots,k\}$, we have

$$0 > c\langle\nabla f_i(x), v\rangle > \langle\nabla f_i(x), v\rangle = \lim_{t\to 0}\frac{f_i(x+tv) - f_i(x)}{t} \quad \forall i \in \{1,\dots,k\}.$$

Thus, for each $i \in \{1,\dots,k\}$, there is some $T_i > 0$ such that

$$\begin{aligned}&c\langle\nabla f_i(x), v\rangle > \frac{f_i(x+tv) - f_i(x)}{t} \quad \forall t \in (0,T_i] \\ \Leftrightarrow\quad &f_i(x+tv) < f_i(x) + ct\langle\nabla f_i(x), v\rangle \quad \forall t \in (0,T_i].\end{aligned}$$

Defining $T := \min_{i\in\{1,\dots,k\}} T_i$ completes the proof. □

Remark 2.2.3. *The inequality in Lemma 2.2.1 is also called the* Armijo condition *[NW06] and will later reappear in the computation of step lengths for descent methods.*

Using Lemma 2.2.1 and Lemma 2.2.2, it is now easy to derive the first-order necessary optimality condition for (locally weak) Pareto optimality, known as the *KKT (Karush-Kuhn-Tucker) condition* [KT51; Mie98]. To this end, we define the $(k-1)$*-standard simplex*

$$\Delta_k := \left\{ \alpha \in \mathbb{R}^k : \alpha_i \geq 0 \ \forall i \in \{1, \dots, k\}, \sum_{i=1}^{k} \alpha_i = 1 \right\}. \tag{2.8}$$

Theorem 2.2.4. *Let x be (locally weakly) Pareto optimal. Then*

$$\exists \alpha \in \Delta_k \text{ such that } \sum_{i=1}^{k} \alpha_i \nabla f_i(x) = 0. \tag{KKT}$$

Proof. Assume that this does not hold, i.e., $0 \notin \operatorname{conv}(\{\nabla f_1(x), \dots, \nabla f_k(x)\}) =: W$. Clearly, W is convex and compact, so by Lemma 2.2.1 we must have some $\bar{v} \in -W$ with $\langle \nabla f_i(x), \bar{v} \rangle < 0$ for all $i \in \{1, \dots, k\}$. By Lemma 2.2.2, this means that x is not locally weakly Pareto optimal, which is a contradiction. □

Since (KKT) is only a necessary condition, we make the following definition.

Definition 2.2.5. *A point $x \in \mathbb{R}^n$ is called* Pareto critical *if it satisfies* (KKT)*. A corresponding $\alpha \in \Delta_k$ is called a* KKT vector of x *and contains* KKT multipliers *α_i, $i \in \{1, \dots, k\}$. The set P_c of all Pareto critical points is the* Pareto critical set*.*

Remark 2.2.6. *a) For $k = 1$,* (KKT) *reduces to the classical optimality condition $\nabla f(x) = 0$ (with KKT vector $\alpha = 1$) from single-objective optimization.*

b) The condition (KKT) *can also be formulated as*

$$\exists \alpha \in \Delta_k \text{ such that } Df(x)^\top \alpha = 0,$$

where Df is the Jacobian of f. This will later be used to derive relationships between the rank and kernel of $Df^\top$ and the Pareto critical set.

c) Pareto critical points are also sometimes referred to as substationary points *[Mie98].*

d) The Pareto critical set was first defined by Smale in [Sma73] via nonexistence of a common descent direction of the objectives. By Lemma 2.2.1, both definitions are equivalent.

e) If f_i is convex for all $i \in \{1, \dots, k\}$, then (KKT) *is also sufficient. In general, sufficient conditions can be derived using second-order derivatives (cf. [Mie98], Theorem 3.2.17).*

f) In Section 3.1.4, a generalization of Theorem 2.2.4 to constrained MOPs will be given (see Theorem 3.1.9).

The KKT vector $\alpha \in \Delta_k$, i.e., the weighting of the gradients of the objectives in the KKT condition, is a quantity that is exclusive to the multiobjective case. It can be geometrically interpreted via the following lemma. Roughly speaking, it implies that if $x \in \mathbb{R}^n$ is Pareto optimal, then corresponding KKT vectors are orthogonal to the Pareto front in $f(x)$ (pointing into the interior of $\operatorname{im}(f)$). (See Section 4.3 in [Hil01] for more details on this.)

Lemma 2.2.7. *Assume that*

$$x : (-1,1) \to \mathbb{R}^n, \quad t \mapsto x(t),$$
$$\alpha : (-1,1) \to (\mathbb{R}^{>0})^k, \quad t \mapsto \alpha(t)$$

are differentiable such that for all $t \in (-1,1)$, $x(t)$ is Pareto critical with KKT vector $\alpha(t)$. Then

$$\alpha(t)^\top D(f \circ x)(t) = 0 \quad \forall t \in (-1,1).$$

In other words, $\alpha(t)$ is orthogonal to the tangent of the curve $f \circ x$ in t.

Proof. By assumption we have $Df(x(t))^\top \alpha(t) = 0$ for all $t \in (-1,1)$. Thus,

$$\begin{aligned}\alpha(t)^\top D(f \circ x)(t) &= \alpha(t)^\top Df(x(t))Dx(t) \\ &= (Df(x(t))^\top \alpha(t))^\top Dx(t) = 0 \quad \forall t \in (-1,1).\end{aligned}$$

□

The following example, which is similar to Example 2.1.6, shows a simple application of the KKT condition.

Example 2.2.8. *For $c \in \mathbb{R}^n$ and a symmetric, positive definite matrix $Q \in \mathbb{R}^{n\times n}$ consider the problem*

$$\min_{x\in\mathbb{R}^n} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} \|x\|_Q^2 \\ \|x-c\|^2 \end{pmatrix}, \tag{2.9}$$

where $\|x\|_Q := \sqrt{x^\top Q x}$ denotes the norm based on the inner product induced by Q. This problem can be seen as the location problem from Example 2.1.6 with a more general norm. We will now derive its Pareto critical set.

First of all, we have

$$\nabla f_1(x) = 2Qx, \quad \nabla f_2(x) = 2(x-c),$$

so (KKT) *is equivalent to*

$$\begin{aligned}0 &= \alpha_1 \nabla f_1(x) + \alpha_2 \nabla f_2(x) = 2\alpha_1 Qx + 2\alpha_2(x-c) \\ &= 2(\alpha_1 Q + \alpha_2 I)x - 2\alpha_2 c\end{aligned} \tag{2.10}$$

for $\alpha \in \Delta_2$, where I is the identity matrix in $\mathbb{R}^{n\times n}$. The matrix $\alpha_1 Q + \alpha_2 I$ is positive definite (and thus invertible) as the sum of two positive definite matrices, and since $\alpha_1 + \alpha_2 = 1$, we have $\alpha_1 = 1 - \alpha_2$. So (2.10) *is equivalent to*

$$x = \alpha_2((1-\alpha_2)Q + \alpha_2 I)^{-1} c$$

for some $\alpha_2 \in [0,1]$, and the Pareto critical set is given by

$$P_c = \{\alpha_2((1-\alpha_2)Q + \alpha_2 I)^{-1} c : \alpha_2 \in [0,1]\}.$$

(Since both objectives are convex, all Pareto critical points are actually Pareto optimal.) Figure 2.4 shows the Pareto critical set for $n = 2$,

$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \textit{and} \quad Q = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

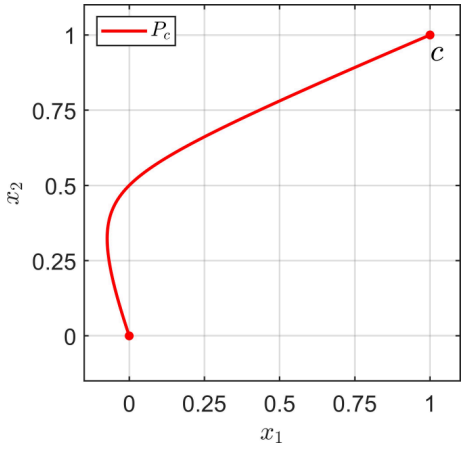


Figure 2.4: Pareto critical set for the generalized location problem in Example 2.2.8.

Since the solution of an MOP is a set, the structure of that set is important to better understand the problem and to be able to efficiently solve it. As the KKT condition is a necessary optimality condition, the Pareto critical set is a superset of the actual Pareto set. Thus, to an extent, results about the structure of the Pareto critical set also apply to the Pareto set. Additionally, those results can motivate new algorithms and strategies to solve MOPs, which will be demonstrated in this thesis.

Note that the Pareto critical set is implicitly defined via the existence of some $\alpha \in \Delta_k$ such that a certain system of equations is satisfied. Thus, by considering Pareto critical points x and their corresponding KKT vectors α simultaneously as points $(x, \alpha) \in \mathbb{R}^n \times \Delta_k$, the Pareto critical set can be expressed as the projection of a zero level set in a higher dimensional space. This opens up the possibility to use tools from differential geometry to study its structure. In the following, this approach will be formalized and a first result will be given. To this end, let

$$F : \mathbb{R}^n \times (\mathbb{R}^{\geq 0})^k \to \mathbb{R}^{n+1}, \quad (x, \alpha) \mapsto \begin{pmatrix} \sum_{i=1}^k \alpha_i \nabla f_i(x) \\ 1 - \sum_{i=1}^k \alpha_i \end{pmatrix}. \tag{2.11}$$

Then

$$x \in P_c \;\Leftrightarrow\; \exists \alpha \in (\mathbb{R}^{\geq 0})^k : F(x, \alpha) = 0,$$

so

$$P_c = \mathrm{pr}_x(F^{-1}(0)), \tag{2.12}$$

where $\mathrm{pr}_x : \mathbb{R}^{n+k} \to \mathbb{R}^n$, $(x, \alpha) \mapsto x$, is the projection onto the variable space. The following result from differential geometry, known as the *level set theorem* [Lee12], can be used to analyze the structure of the zero level set $F^{-1}(0)$ of F.

Theorem 2.2.9 (Level set theorem). *Let $n_1, n_2 \in \mathbb{N}$, $U \subseteq \mathbb{R}^{n_1}$ be open and $h : U \to \mathbb{R}^{n_2}$ be differentiable. If $\mathrm{rk}(Dh(x)) = n_2$ for all $x \in h^{-1}(0)$, then $h^{-1}(0)$ is a $(n_1 - n_2)$-dimensional embedded submanifold of $\mathbb{R}^{n_1}$ with tangent space*

$$T_x(h^{-1}(0)) = \ker(Dh(x)).$$

Proof. Corollary 5.14 and Proposition 5.38 in [Lee12]. □

The level set theorem implies the following result from [Hil01].

Theorem 2.2.10. *Let f be twice continuously differentiable and let*

$$\mathcal{M} := (F|_{\mathbb{R}^n \times (\mathbb{R}^{>0})^k})^{-1}(0).$$

a) If $rk(DF(x,\alpha)) = n+1$ for all $(x,\alpha) \in \mathcal{M}$, then $\mathcal{M}$ is a $(k-1)$-dimensional embedded submanifold of $\mathbb{R}^{n+k}$ with tangent space

$$T_{(x,\alpha)}\mathcal{M} = \ker(DF(x,\alpha)).$$

b) If $(x,\alpha) \in \mathcal{M}$ with $rk(DF(x,\alpha)) = n+1$, then there is an open set $U \subseteq \mathbb{R}^{n+k}$ with $(x,\alpha) \in U$ such that $\mathcal{M} \cap U$ is a manifold as in a).

Proof. a) Application of Theorem 2.2.9 to the differentiable function $F|_{\mathbb{R}^n \times (\mathbb{R}^{>0})^k}$.
b) Since the subset of matrices in $\mathbb{R}^{(n+1)\times(n+k)}$ with full rank is open, and since $DF : \mathbb{R}^n \times (\mathbb{R}^{>0})^k \to \mathbb{R}^{(n+1)\times(n+k)}$ is continuous by assumption, the set

$$U := \{(x,\alpha) \in \mathbb{R}^n \times (\mathbb{R}^{>0})^k : \mathrm{rk}(DF(x,\alpha)) = n+1\}$$

is open. Applying Theorem 2.2.9 to $F|_U$ completes the proof. □

Remark 2.2.11. *The proof of Theorem 2.2.10 given here shows in particular that*

$$\{(x,\alpha) \in \mathcal{M} : rk(DF(x,\alpha)) = n+1\}$$

is open in $\mathcal{M}$ (w.r.t. the subspace topology) and a manifold with the properties from Theorem 2.2.10 a).

For $x \in \mathbb{R}^n$ and $\alpha \in \Delta_k$ we have

$$DF(x,\alpha) = \begin{pmatrix} \sum_{i=1}^k \alpha_i \nabla^2 f_i(x) & Df(x)^\top \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+k)}. \tag{2.13}$$

We will sometimes refer to the matrix $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ as the *weighted Hessian matrix* of f. The following lemma (which is similar to Theorem 5.3 in [Hil01]) gives sufficient conditions for the rank of DF to be full, i.e., for the requirements of Theorem 2.2.10 to be satisfied. In particular, it shows that if all objective functions are *strongly convex* (cf. [BV04]), then the requirement of Theorem 2.2.10 a) is satisfied.

Lemma 2.2.12. *Let $x \in \mathbb{R}^n$.*

a) If $\alpha \in \Delta_k$ such that $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ is regular, then $\mathrm{rk}(DF(x,\alpha)) = n+1$.

b) If $\nabla^2 f_i(x)$ is positive definite for all $i \in \{1,\dots,k\}$, then $\mathrm{rk}(DF(x,\alpha)) = n+1$ for all $\alpha \in \Delta_k$.

Proof. a) Assume that $n+1 > \mathrm{rk}(DF(x,\alpha)) = \mathrm{rk}(DF(x,\alpha)^\top)$. Then there must be some $0 \neq v = (v^1, v^2)^\top \in \mathbb{R}^n \times \mathbb{R}$ such that

$$0 = DF(x,\alpha)^\top v = \begin{pmatrix} \sum_{i=1}^k \alpha_i \nabla^2 f_i(x) v^1 \\ Df(x)v^1 + (v^2,\dots,v^2)^\top \end{pmatrix}. \tag{2.14}$$

Since $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ is regular, it follows that $v^1 = 0$, which also implies $v^2 = 0$. This is a contradiction.
b) The matrix $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ is positive definite (and in particular regular) for all $\alpha \in \Delta_k$, since all $\nabla^2 f_i(x)$ are positive definite and $\alpha_j > 0$ for at least one $j \in \{1,\dots,k\}$. Thus, the proofs follows from a). □

As an example, Theorem 2.2.10 can be used to show that $\mathcal{M}$ is a manifold for the (strongly convex) MOP in Example 2.2.8. In the following, in order to show what kind of "singularities" may occur in the general case, we will consider a simple problem where $\mathcal{M}$ is not a manifold.

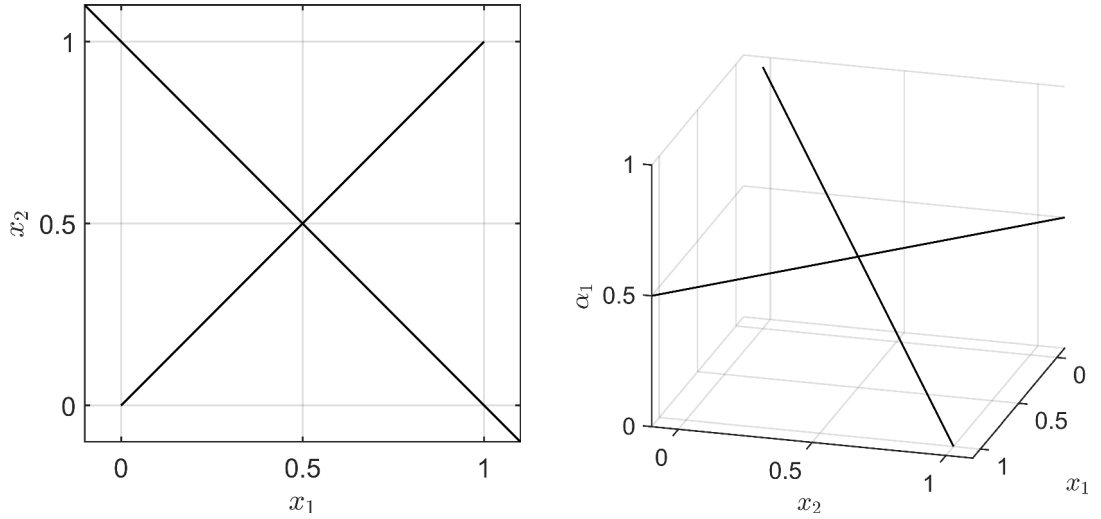
Example 2.2.13. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} -\frac{1}{2}x_1^2 - \frac{1}{2}x_2^2 \\ x_1 + x_2 - x_1 x_2 \end{pmatrix}.$$

It is straight-forward to show that the Pareto critical set for this problem is given by

$$P_c = \{x \in [0,1]^2 : x_1 = x_2\} \cup \{x \in \mathbb{R}^n : x_1 + x_2 = 1\},$$

as shown in Figure 2.5(a). The KKT vector corresponding to $x \in P_c$ is given by



(a) P_c **(b)** $\mathcal{M}$

Figure 2.5: Pareto critical set and $\mathcal{M}$ in Example 2.2.13.

$$\alpha = \begin{cases} (1 - x_1, x_1)^\top, & \textit{if } x_1 = x_2, \\ (\frac{1}{2}, \frac{1}{2})^\top, & \textit{if } x_1 + x_2 = 1. \end{cases}$$

A projection of the resulting set $\mathcal{M}$ onto x_1, x_2 and α_1 is shown in Figure 2.5(b). We see that $\mathcal{M}$ cannot be a manifold with dimension $k - 1 = 1$ in this case, since the intersection at $((\frac{1}{2}, \frac{1}{2})^\top, (\frac{1}{2}, \frac{1}{2})^\top) \in \mathcal{M}$ is not diffeomorphic to a one-dimensional line. The matrix DF (cf. (2.13)) for this problem is given by

$$DF(x, \alpha) = \begin{pmatrix} -\alpha_1 & -\alpha_2 & -x_1 & 1 - x_2 \\ -\alpha_2 & -\alpha_1 & -x_2 & 1 - x_1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

It is easy to see that $rk(DF(x, \alpha)) = 2 \neq 3 = n + 1$ in the intersection point from above.

If Theorem 2.2.10 can be applied, then the Pareto critical set P_c is the projection of a smooth manifold in $\mathbb{R}^{n+k}$ onto $\mathbb{R}^n$, except for the points where at least one KKT multiplier is zero. So although P_c set itself is not a manifold, there is still a smooth structure on $\mathcal{M}$ that we can project onto P_c. This will later be used to derive the box-continuation method (Chapter 3) and to analyze the hierarchical structure of P_c (Chapter 5).

2.2.2 Nonsmooth case

We will now derive an optimality condition for the case where f is not necessarily differentiable. First, it has to be specified how much "nonsmoothness" is allowed. Since we want to work with local information about f, local change of x should only cause local change of $f(x)$. In other words, f should still be continuous. Unfortunately, the class of continuous functions includes pathological functions like the Weierstrass function (cf. [Har16]), whose graph is a fractal curve that does not yield any useful local information. So in addition to continuity, the evaluation of f should be somehow numerically "stable", i.e., the local change of f should be bounded. More precisely, we assume f to be *locally Lipschitz (continuous)*, i.e., for each $i \in \{1, \dots, k\}$ and $x \in \mathbb{R}^n$, there is some $L_i > 0$ and $\varepsilon > 0$ with

$$|f_i(y) - f_i(z)| \leq L_i \|y - z\| \quad \forall y, z \in B_\varepsilon(x), \tag{2.15}$$

where $B_\varepsilon(x) := \{y \in \mathbb{R}^n : \|x - y\| < \varepsilon\}$. As it turns out, the class of locally Lipschitz objective functions is broad enough to cover many relevant problems from nonsmooth optimization (cf. [BKM14], Part II).

An important result about locally Lipschitz functions is the following theorem, known as *Rademacher's theorem* [EG15]. Let $\Omega_i \subseteq \mathbb{R}^n$ be the set of points where f_i is not differentiable.

Theorem 2.2.14 (**Rademacher's theorem**). *The set Ω_i is a (Lebesgue) null set for all $i \in \{1, \dots, k\}$.*

Due to Rademacher's theorem, even if f_i is merely locally Lipschitz, we can still work with the classical gradient almost everywhere. In particular, $\mathbb{R}^n \setminus \Omega_i$ is dense in $\mathbb{R}^n$, so for each $x \in \Omega_i$, there is a sequence of differentiable points that converges to x. This can be used to generalize the concept of differentiability to the locally Lipschitz case [Cla90]:

Definition 2.2.15. *Let $x \in \mathbb{R}^n$ and $i \in \{1, \dots, k\}$. The set*

$$\partial f_i(x) = \operatorname{conv}(\{\xi \in \mathbb{R}^n : \exists (x^j)_j \in \mathbb{R}^n \setminus \Omega_i \text{ with } \lim_{j\to\infty} x^j = x \text{ and } \lim_{j\to\infty} \nabla f_i(x^j) = \xi\})$$

is the (Clarke) subdifferential of f_i in x. *An element $\xi \in \partial f_i(x)$ is a* (Clarke) subgradient.

Clearly, if f_i is continuously differentiable in x, then $\partial f_i(x) = \{\nabla f_i(x)\}$. In this sense, the Clarke subdifferential reduces to the classical gradient in the continuously differentiable case. The following example shows how the Clarke subdifferential can be used to obtain derivatives of the ℓ_1-norm, which is notoriously nonsmooth.

Example 2.2.16. *Let*

$$f : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \|x\|_1 := |x_1| + \ldots + |x_n|.$$

It is easy to see that the set of nondifferentiable points is

$$\Omega = \{x \in \mathbb{R}^n : x_i = 0 \text{ for some } i \in \{1, \dots, k\}\}$$

and that f is continuously differentiable outside of Ω. For $x \notin \Omega$ we have

$$\nabla f(x) = \begin{pmatrix} \operatorname{sgn}(x_1) \\ \vdots \\ \operatorname{sgn}(x_n) \end{pmatrix}.$$

For $x \in \Omega$ with $x_i = 0$, sequences in Definition 2.2.15 can converge from the negative ($x_i < 0$) or the positive ($x_i > 0$) half-space, so the subgradients are either 1 or -1 in the i-th component. Thus, the Clarke subdifferential is given by

$$\partial f(x) = \left\{ \xi \in \mathbb{R}^n : \xi_i \begin{cases} = \operatorname{sgn}(x_i) & \text{if } x_i \neq 0 \\ \in [-1, 1] & \text{if } x_i = 0 \end{cases}, i \in \{1, \dots, n\} \right\}.$$

The following lemma summarizes some of the well-known properties of the Clarke subdifferential [Cla90].

Lemma 2.2.17. *Let $x \in \mathbb{R}^n$ and $i \in \{1, \dots, k\}$.*

a) $\partial f_i(x)$ is nonempty, convex and compact in $\mathbb{R}^n$.

b) Let $L_i > 0$ be a Lipschitz constant of f_i at x. Then

$$\partial f_i(x) \subseteq \overline{B_{L_i}(0)},$$

i.e., $\|\xi\| \leq L_i$ for all $\xi \in \partial f_i(x)$.

c) As a set-valued map, ∂f_i is upper semicontinuous*: For all open sets $V \subseteq \mathbb{R}^n$ with $\partial f_i(x) \subseteq V$ there is some open set $U \subseteq \mathbb{R}^n$ with $x \in U$ such that*

$$\partial f_i(y) \subseteq V \quad \forall y \in U.$$

Furthermore, the mean value theorem can be extended to the nonsmooth case [BKM14; Cla90].

Theorem 2.2.18 (**Mean value theorem**). *Let $x, y \in \mathbb{R}^n$ with $x \neq y$ and let $h : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz. Then there is some $u \in \operatorname{conv}(\{x, y\}) \setminus \{x, y\}$ and $\xi \in \partial h(u)$ such that*

$$h(y) - h(x) = \langle \xi, y - x \rangle.$$

Remark 2.2.19. *In [Cla90], the mean value theorem is stated for Banach spaces, in which case "global" Lipschitz continuity of h on an open set containing $\operatorname{conv}(\{x, y\})$ is required. In our finite-dimensional case, this immediately follows from the local Lipschitz continuity of h, as every locally Lipschitz function is "globally" Lipschitz continuous on compact sets. For a proof, see, e.g., Proposition 3.3.2 in [SC16].*

Using the generalized concept of derivatives, Lemma 2.2.2 can now be extended from the smooth to the nonsmooth case.

Lemma 2.2.20. *Let $c \in (0, 1)$, $x \in \mathbb{R}^n$ and $v \in \mathbb{R}^n$ such that $\langle \xi, v \rangle < 0$ for all $\xi \in \partial f_i(x)$ and all $i \in \{1, \dots, k\}$. Then there is some $T > 0$ such that*

$$f_i(x + tv) < f_i(x) + tc \max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle \quad \forall t \in (0, T],\ i \in \{1, \dots, k\}.$$

In particular, x is not (locally weakly) Pareto optimal.

Proof. Let $i \in \{1, \ldots, k\}$. Since $\partial f_i(x)$ is compact we have

$$\max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle < 0.$$

Thus, using continuity of $\langle \cdot, v \rangle$, there has to be an open set $V_i \subseteq \mathbb{R}^n$ with $\partial f_i(x) \subseteq V_i$ such that

$$\langle \xi', v \rangle < c \max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle < 0 \quad \forall \xi' \in V_i.$$

Due to upper semicontinuity of ∂f_i, there is some open set $U_i \subseteq \mathbb{R}^n$ with $x \in U_i$ such that

$$\partial f_i(y) \subseteq V_i \quad \forall y \in U_i.$$

Let $U := U_1 \cap \ldots \cap U_k$. Then for each $i \in \{1, \ldots, k\}$ we have

$$\langle \xi', v \rangle < c \max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle < 0 \quad \forall \xi' \in \partial f_i(y),\ y \in U.$$

Since U is open and $x \in U$, there is some $T > 0$ such that $x + tv \in U$ for all $t \in (0, T]$.
Now let $t \in (0, T]$ and $i \in \{1, \ldots, k\}$. Applying the mean value theorem to x and $x + tv$ yields the existence of some $u \in \operatorname{conv}(\{x, x + tv\})$ and $\xi \in \partial f_i(u)$ such that

$$f_i(x + tv) - f_i(x) = t\langle \xi, v \rangle.$$

By construction, $u \in U$, so $\langle \xi, v \rangle < c \max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle$ and thus

$$f_i(x + tv) < f_i(x) + tc \max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle.$$

□

Remark 2.2.21. *In general,* $\max_{\xi \in \partial f_i(x)} \langle \xi, v \rangle =: \partial^\circ f_i(x, v)$ *is known as the* generalized directional derivative of f_i at x in the direction v *[Cla90].*

By combining Lemma 2.2.20 with Lemma 2.2.1, we obtain a first-order optimality condition for the nonsmooth case [MEK14]. To this end, let

$$\partial^\cup f(x) := \operatorname{conv}\left(\bigcup_{i=1}^{k} \partial f_i(x) \right). \tag{2.16}$$

Theorem 2.2.22. *Let x be (locally weakly) Pareto optimal. Then*

$$0 \in \partial^\cup f(x). \tag{2.17}$$

Proof. Analogously to the proof of Theorem 2.2.4 with $W = \partial^\cup f(x)$. (W is compact as the convex hull of a compact set. See, e.g., Exercise 2.4.11 in [BV10].) □

It is easy to see that (2.17) reduces to the smooth KKT condition if f is continuously differentiable. Thus, the following definition extends Pareto criticality to the nonsmooth case.

Definition 2.2.23. *A point $x \in \mathbb{R}^n$ is called* Pareto critical *if $0 \in \partial^{\cup} f(x)$. The set P_c of all Pareto critical points is the* Pareto critical set.

Remark 2.2.24. *a) In the smooth case, the KKT vectors are defined via the coefficients of the vanishing convex combinations of the gradients. In the nonsmooth case, these coefficients also depend on which elements from the individual $\partial f_i(x)$ are chosen. A possible generalization of KKT vectors in the nonsmooth case will be considered in Section 5.5 (see Definition 5.5.3).*

b) If f_i is convex for all $i \in \{1, \ldots, k\}$, then (2.17) is sufficient for weak Pareto optimality (cf. [MEK14], Theorem 14).

In the following, a simple example for a Pareto critical set in the nonsmooth case is considered. (It is inspired by [BGP21].)

Example 2.2.25. *For a continuously differentiable $g : \mathbb{R}^n \to \mathbb{R}$, consider the problem*

$$\min_{x \in \mathbb{R}^n} f(x) \quad \text{with} \quad f(x) = \begin{pmatrix} g(x) \\ \|x\|_1 \end{pmatrix}.$$

Since the ℓ_1-norm enforces sparsity (cf. [Tib96], Section 2.3), this problem can be interpreted as finding minimal points of g that are as sparse as possible.

Let $n = 2$ and $g(x) := \|x - (2,1)^\top\|^2$, i.e., $\nabla g(x) = 2(x - (2,1)^\top)$. Let Ω_2 be the set of nondifferentiable points of $\|\cdot\|_1$. By Example 2.2.16, $x \notin \Omega_2$ is Pareto critical if and only if

$$\begin{aligned} 0 &= 2\alpha_1 \left(x - \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right) + \alpha_2 \begin{pmatrix} \operatorname{sgn}(x_1) \\ \operatorname{sgn}(x_2) \end{pmatrix} \\ \Leftrightarrow x &= \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{\alpha_2}{2\alpha_1} \begin{pmatrix} \operatorname{sgn}(x_1) \\ \operatorname{sgn}(x_2) \end{pmatrix} \end{aligned}$$

for some $\alpha \in \Delta_2$ with $\alpha_1 > 0$. This condition holds precisely when x lies on the line connecting $(2,1)^\top$ and $(1,0)^\top$ (excluding $(1,0)^\top$ itself). For $x \in \mathbb{R}^2 \setminus \{0\}$ with $x_2 = 0$, x is Pareto critical if and only if

$$\begin{aligned} 0 &\in \operatorname{conv}\left(\left\{ 2\left(\begin{pmatrix} x_1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right) \right\} \cup \left(\{\operatorname{sgn}(x_1)\} \times [-1, 1] \right) \right) \\ &= \operatorname{conv}\left(\left\{ \begin{pmatrix} 2x_1 - 4 \\ -2 \end{pmatrix}, \begin{pmatrix} \operatorname{sgn}(x_1) \\ -1 \end{pmatrix}, \begin{pmatrix} \operatorname{sgn}(x_1) \\ 1 \end{pmatrix} \right\} \right). \end{aligned}$$

It is possible to show that this is equivalent to x lying on the line connecting $(1,0)^\top$ and $(0,0)^\top$ (excluding $(0,0)^\top$ itself). For $x \in \mathbb{R}^2 \setminus \{0\}$ with $x_1 = 0$, no Pareto critical points can be found. Finally, $(0,0)^\top$ is Pareto critical as the unique minimizer of $\|x\|_1$.

The complete Pareto critical set is shown in Figure 2.6. Since all the objectives are convex, it coincides with the Pareto set.

Note that since the optimality condition (2.17) is based on sets, it is not directly possible to write the Pareto critical set as the projection of the level set of some smooth function. In other words, (2.12) cannot be generalized to the nonsmooth

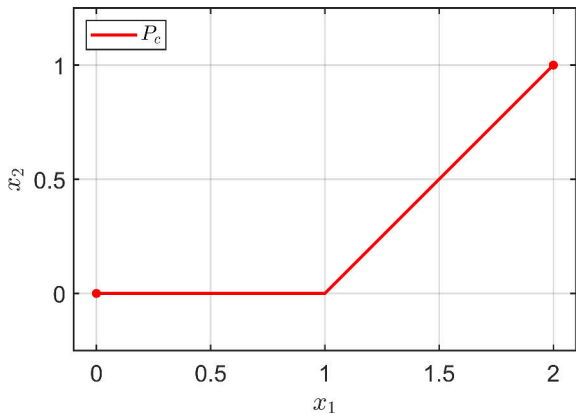


Figure 2.6: Pareto (critical) set for the nonsmooth problem from Example 2.2.25.

case and we lose the "projected smoothness" of P_c. For instance, in Example 2.2.25, there is a kink in P_c when going from the smooth to the nonsmooth part of the variable space.

Another important disparity to the smooth case is the fact that subdifferentials, in contrast to gradients, cannot be easily computed in a practical setting. For example, while finite differences can be used in the smooth case to obtain a good approximation of the gradient, they fail to produce a good approximation of the full subdifferential in the nonsmooth case (cf. [Lem89], Section 3.3). Thus, any method for nonsmooth MOPs that explicitly relies on subdifferentials has to contain a procedure to efficiently and properly approximate them. Examples for this will be given in Section 2.3.2 and Chapter 4.

Finally, we will briefly discuss an approach to obtain a stronger first-order necessary optimality condition than (2.17) in the following remark.

Remark 2.2.26. *In addition to the generalization of gradients of f_i via the Clarke subdifferential (cf. Definition 2.2.15), it is possible to generalize the Jacobian of f as*

$$\partial f(x) := \operatorname{conv}(\{\xi \in \mathbb{R}^{k\times n} : \exists (x^j)_j \in \mathbb{R}^n \setminus \Omega \text{ with } \lim_{j\to\infty} x^j = x \text{ and } \lim_{j\to\infty} Df(x^j) = \xi\}),$$

where Ω is the set of points in which f is nondifferentiable. The set $\partial f(x) \subseteq \mathbb{R}^{k\times n}$ is called the generalized Jacobian of f at x *[Cla90].*

In the smooth case, the Jacobian of f can be used to obtain the equivalent formulation

$$Df(x)^\top \alpha = 0, \quad \alpha \in \Delta_k,$$

for the optimality condition (KKT)*. In other words, in the smooth case, it does not matter if we express the derivative of f as the set of gradients ∇f_i, $i \in \{1,\dots,k\}$, or the Jacobian Df. In the general nonsmooth case, this equivalence is lost (cf. [Cla90], Remark 6.1.2, and [Gut+16], p. 4): Although the i-th row of all matrices in the generalized Jacobian is a subgradient $\xi_i \in \partial f_i(x)$, not all combinations of subgradients of f_i, $i \in \{1,\dots,k\}$ produce elements in $\partial f(x)$. Only those $\xi_i \in \partial f_i(x)$*

that arise by using the same sequence $(x^j)_j$ in Definition 2.2.15 can be combined into an element $(\xi_1, \dots, \xi_k)^\top \in \mathbb{R}^{k\times n}$ of the generalized Jacobian. More formally,

$$\partial f(x) \subsetneq \{(\xi_1, \dots, \xi_k)^\top \in \mathbb{R}^{k\times n} : \xi_i \in \partial f_i(x),\ i \in \{1, \dots, k\}\}.$$

An optimality condition based on the generalized Jacobian can be found in [Gut+16], Definition 3. As this condition is slightly more complicated than the optimality condition (2.17)*, we will leave its utilization in our context for future research.*

2.3 Existing solution methods

In this section, several popular solution methods for MOPs will be presented. They are roughly separated into four classes: The first class contains scalarization methods, which are based on turning MOPs into parameter-dependent scalar problems and then applying methods from single-objective optimization. The second class contains methods that arise by directly generalizing methods from single-objective optimization (without scalarizing the MOP beforehand). The methods in the first two classes only compute a single Pareto optimal point at a time. In contrast to this, the third class contains methods that compute the entire Pareto set at once instead of just single optimal points. Finally, the fourth class contains methods which are motivated by evolutionary computation.

This section acts as a short overview. For a more detailed discussion, see [Mie98; Ehr05; Deb01].

2.3.1 Scalarization methods

The idea of scalarization methods is to convert MOPs into (parameter-dependent) scalar problems, for which there is a large quantity of solvers available. Since the actual optimization is carried out in the single-objective setting, these methods tend to be easy to implement. On top of that, in many cases, the reformulation of the MOP is intuitive and easy to interpret, requiring almost no understanding of general multiobjective optimization. Both reasons led to scalarization methods being among the most popular methods for solving MOPs. The methods considered here are the *weighting method* and the *ε-constraint method*.

Weighting method

Instinctively, when getting tasked to minimize multiple objectives at the same time, the first idea one might come up with is to minimize the sum of all objectives, as this does embody a sense of simultaneous minimization. Since this approach is clearly highly dependent on the scaling of the objective function values, one might additionally add a weighting coefficient in front of each objective. Formally, for coefficients $\alpha \in \Delta_k$, the resulting scalar problem is

$$\min_{x\in\mathbb{R}^n} \sum_{i=1}^{k} \alpha_i f_i(x), \tag{WS}$$

called the *weighted sum*. Solving this problem is referred to as the *weighting method*.

It is possible to show that for all $\alpha \in \Delta_k$, solutions of (WS) are indeed at least weakly Pareto optimal. (In particular, if f is differentiable, it is easy to see that α is a corresponding KKT vector.) But only if all objective functions are convex, it is possible to show the opposite implication, i.e., that for all Pareto optimal points $x \in P$, there is some weighting $\alpha \in \Delta_k$ such that x is a solution of (WS). The required convexity is one of the main drawbacks of the weighting method.

ε-constraint method

A different intuition for minimizing multiple objectives at the same time is to minimize only one of the objectives, with the constraint that the values of the other objectives are at least below certain levels. Formally, let $j \in \{1, \dots, k\}$, $\varepsilon_i \in \mathbb{R}$ for $i \in \{1, \dots, k\} \setminus \{j\}$ and consider the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f_j(x), \\ s.t. \quad & f_i(x) \leq \epsilon_i \ \forall i \in \{1, \dots, k\} \setminus \{j\}. \end{aligned} \tag{EC}$$

Solving (EC) is referred to as the *ε-constraint method*.

It is again possible to show that solutions of (EC) are at least weakly Pareto optimal. Additionally, $x \in \mathbb{R}^n$ is Pareto optimal if and only if x is a solution of (EC) for all $j \in \{1, \dots, k\}$ and $\varepsilon_i = f_i(x)$, $i \in \{1, \dots, k\} \setminus \{j\}$. So in this case, convexity of f is not required, but (EC) has to be solved k-times to ensure Pareto optimality.

2.3.2 Generalizations of methods from scalar optimization

Since multiobjective optimization is a generalization of single-objective optimization, certain methods for the single-objective case can be generalized. An advantage of this approach is that the original methods are well understood and in many cases, theoretical results only need minor adjustments to be transferred to the multiobjective case. Here, we will only consider the *steepest descent method* [FS00] (for smooth MOPs) and the *multiobjective proximal bundle method* [MKW14] (for nonsmooth MOPs), but other methods can be derived in a similar way.

Steepest descent method

The idea of descent methods is to iteratively generate a sequence $(x^j)_j \in \mathbb{R}^n$, starting in some $x^1 \in \mathbb{R}^n$, where each element is "more optimal" than the previous element. In the single-objective case, this means $f(x^{j+1}) < f(x^j)$ for all $j \geq 1$. If f is continuously differentiable, this can be realized by setting

$$x^{j+1} = x^j + tv,$$

where $v = -\nabla f(x^j)$ and $t > 0$ is some step length assuring descent. Since $-\nabla f(x)$ is the direction that locally promises the steepest (first-order) descent of f in x, this method is called the *steepest descent method* (or *gradient descent method*) [NW06].

In the multiobjective setting, by Lemma 2.2.2, the descent of f in a direction $v \in \mathbb{R}^n$ can be measured by

$$\max_{i \in \{1, \dots, k\}} \langle \nabla f_i(x), v \rangle. \tag{2.18}$$

We say that v is a *descent direction of f in x* if (2.18) is negative. Let $\bar{v} : \mathbb{R}^n \to \mathbb{R}^n$, where $\bar{v}(x)$ is the solution of (2.6) for

$$W = \operatorname{conv}(\{\nabla f_1(x), \ldots, \nabla f_k(x)\}),$$

i.e.,

$$\bar{v}(x) := \operatorname*{arg\,min}_{v \in -\operatorname{conv}(\{\nabla f_1(x), \ldots, \nabla f_k(x)\})} \|v\|^2. \tag{2.19}$$

By Lemma 2.2.1, if x is not Pareto critical, then $\bar{v}(x)$ is a descent direction for f in x. In fact, $\bar{v}(x)$ is the steepest descent direction of f in x in the following sense:

Lemma 2.3.1. *Let $x \in \mathbb{R}^n$. If $\bar{v}(x) \neq 0$ then*

$$\frac{\bar{v}(x)}{\|\bar{v}(x)\|} = \operatorname*{arg\,min}_{v \in \overline{B_1(0)}} \max_{i \in \{1,\ldots,k\}} \langle \nabla f_i(x), v \rangle. \tag{2.20}$$

Proof. By inequality (2.7) we have

$$\max_{i \in \{1,\ldots,k\}} \left\langle \nabla f_i(x), \frac{\bar{v}(x)}{\|\bar{v}(x)\|} \right\rangle = \frac{1}{\|\bar{v}(x)\|} \max_{i \in \{1,\ldots,k\}} \langle \nabla f_i(x), \bar{v}(x) \rangle \leq -\|\bar{v}(x)\|.$$

Assume that (2.20) does not hold, i.e., that there is some $w \in \overline{B_1(0)}$ such that

$$\max_{i \in \{1,\ldots,k\}} \langle \nabla f_i(x), w \rangle < \max_{i \in \{1,\ldots,k\}} \left\langle \nabla f_i(x), \frac{\bar{v}(x)}{\|\bar{v}(x)\|} \right\rangle \leq -\|\bar{v}(x)\|.$$

Since $\bar{v}(x) \in -\operatorname{conv}(\{\nabla f_1(x), \ldots, \nabla f_k(x)\})$, there is some coefficient vector $\alpha \in \Delta_k$ such that $\bar{v}(x) = -\sum_{i=1}^k \alpha_i \nabla f_i(x)$. In particular,

$$\begin{aligned}
& -\langle \bar{v}(x), w \rangle = \sum_{i=1}^k \alpha_i \langle \nabla f_i(x), w \rangle \leq \sum_{i=1}^k \alpha_i \max_{j \in \{1,\ldots,k\}} \langle \nabla f_j(x), w \rangle < -\|\bar{v}(x)\| \\
\Leftrightarrow \quad & -\langle \bar{v}(x), \|\bar{v}(x)\| w \rangle < -\|\bar{v}(x)\|^2 = -\langle \bar{v}(x), \bar{v}(x) \rangle \\
\Leftrightarrow \quad & \langle \bar{v}(x), \|\bar{v}(x)\| w - \bar{v}(x) \rangle > 0.
\end{aligned}$$

As a result,

$$\begin{aligned}
\|\bar{v}(x)\|^2 \geq \|\bar{v}(x)\|^2 \|w\|^2 &= \|\|\bar{v}(x)\| w\|^2 = \|\bar{v}(x) + (\|\bar{v}(x)\| w - \bar{v}(x))\|^2 \\
&= \|\bar{v}(x)\|^2 + 2\langle \bar{v}(x), \|\bar{v}(x)\| w - \bar{v}(x) \rangle + \|\|\bar{v}(x)\| w - \bar{v}(x)\|^2 \\
&> \|\bar{v}(x)\|^2,
\end{aligned}$$

which is a contradiction. □

Remark 2.3.2. *a) In practice, $\bar{v}$ can be computed via $\bar{v}(x) = -\sum_{i=1}^k \bar{\alpha}_i \nabla f_i(x)$, where $\bar{\alpha}$ is the solution of*

$$\begin{aligned}
\min_{\alpha \in \mathbb{R}^k} \quad & \left\| \sum_{i=1}^k \alpha_i \nabla f_i(x) \right\|^2, \\
s.t. \quad & \sum_{i=1}^k \alpha_i = 1, \\
& \alpha_i \geq 0 \quad \forall i \in \{1, \ldots, k\}.
\end{aligned} \tag{2.21}$$

b) Instead of requiring $v \in \overline{B_1(0)}$, the boundedness of the optimization problem on the right-hand side of (2.20) could also be ensured by adding $\frac{1}{2}\|v\|^2$ to the objective, i.e., by considering the problem

$$\min_{v\in\mathbb{R}^n}\left(\max_{i\in\{1,\dots,k\}}\langle\nabla f_i(x),v\rangle+\frac{1}{2}\|v\|^2\right). \tag{2.22}$$

To get rid of the nondifferentiability caused by the max *function, this problem can be equivalently written as*

$$\begin{aligned}\min_{(v,\beta)\in\mathbb{R}^{n+1}}\quad & \beta+\frac{1}{2}\|v\|^2,\\ s.t.\quad & \nabla f_i(x)^\top v\le\beta\quad\forall i\in\{1,\dots,k\}.\end{aligned}$$

This is the dual problem of (2.21) (cf. [FS00]) and an alternative way of computing $\bar{v}$.

c) Problem (2.22) can also be written as

$$\begin{aligned}\min_{(v,\beta)\in\mathbb{R}^{n+1}}\quad & \beta,\\ s.t.\quad & \nabla f_i(x)^\top v+\frac{1}{2}\|v\|^2\le\beta\quad\forall i\in\{1,\dots,k\}.\end{aligned}$$

If f is twice continuously differentiable and the Hessians $\nabla^2 f_i(x)$ are positive definite for all $i\in\{1,\dots,k\}$, then the Euclidean norm in the inequality constraints of this problem can be replaced by the norms induced by the respective hessian, i.e.,

$$\begin{aligned}\min_{(v,\beta)\in\mathbb{R}^{n+1}}\quad & \beta,\\ s.t.\quad & \nabla f_i(x)^\top v+\frac{1}{2}v^\top\nabla^2 f_i(x)v\le\beta\quad\forall i\in\{1,\dots,k\}.\end{aligned}$$

The direction obtained from solving this problem is known as the Newton direction *[FDS09].*

As in the single-objective case, it can be shown that the steepest descent direction $\bar{v}(x)$ is continuously depending on x (cf. [FS00], Lemma 1).

Lemma 2.3.3. *The function $x\mapsto\bar{v}(x)$ is continuous.*

Proof. Let $(x^j)_j\in\mathbb{R}^n$ with $\lim_{j\to\infty}x^j=\bar{x}\in\mathbb{R}^n$. Since ∇f_i is continuous for every $i\in\{1,\dots,k\}$, $(\bar{v}(x^j))_j$ must be bounded. In particular, $(\bar{v}(x^j))_j$ has an accumulation point $w\in\mathbb{R}^n$ with $w=-\sum_{i=1}^k\alpha_i'\nabla f_i(\bar{x})$ for some $\alpha'\in\Delta_k$. By construction it holds

$$\|\bar{v}(\bar{x})\|^2\le\|w\|^2. \tag{2.23}$$

If we have equality in (2.23), then $\bar{v}(\bar{x})=w$, since the solution of (2.19) is unique. Assume that $\|\bar{v}(\bar{x})\|^2<\|w\|^2$. Let $\bar{\alpha}\in\Delta_k$ such that $\bar{v}(\bar{x})=-\sum_{i=1}^k\bar{\alpha}_i\nabla f_i(\bar{x})$ and define

$$\varphi(x):=\left\|\sum_{i=1}^k\bar{\alpha}_i\nabla f_i(x)\right\|^2.$$

Then φ is continuous, $\varphi(x) \geq \|\bar{v}(x)\|^2$ for all $x \in \mathbb{R}^n$ and $\varphi(\bar{x}) = \|\bar{v}(\bar{x})\|^2$. Since $\varphi(\bar{x}) = \|\bar{v}(\bar{x})\|^2 < \|w\|^2$, there must be some open $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that

$$\|\bar{v}(x)\|^2 \leq \varphi(x) < \|\bar{v}(\bar{x})\|^2 + \frac{1}{2}(\|w\|^2 - \|\bar{v}(\bar{x})\|^2) < \|w\|^2 \quad \forall x \in U.$$

This is a contradiction to w being an accumulation point of $(\bar{v}(x^j))_j$. Thus we must have $w = \bar{v}(\bar{x})$. Since this holds for any accumulation point w of the bounded sequence $(\bar{v}(x^j))_j$, we must have $\lim_{j\to\infty} \bar{v}(x^j) = \bar{v}(\bar{x})$. □

To obtain a descent method for MOPs, the steepest descent direction has to be combined with a step length. To this end, for $c \in (0, 1)$, consider

$$\bar{t}(x) := \max(\{2^{-j} : j \in \mathbb{N},\ f_i(x + 2^{-j}\bar{v}) < f_i(x) + 2^{-j}c\langle \nabla f_i(x), \bar{v}\rangle\ \forall i \in \{1, \ldots, k\}\}). \tag{2.24}$$

By Lemma 2.2.2, $\bar{t}$ is well-defined. The inequality in (2.24) is called the *Armijo condition* [NW06]. The resulting descent method is

$$x^{j+1} = x^j + \bar{t}(x^j)\bar{v}(x^j), \quad j \in \mathbb{N} \tag{2.25}$$

for some initial point $x^1 \in \mathbb{R}^n$, $\bar{v}$ as in (2.19) and $\bar{t}$ as in (2.24). In terms of convergence, the following theorem holds (cf. [FS00], Theorem 1):

Theorem 2.3.4. *All accumulation points of the sequence $(x^j)_j$ generated by (2.25) are Pareto critical.*

Multiobjective proximal bundle method

The steepest descent method previously presented heavily relies on f being continuously differentiable. In the following, we will present a method that can be applied to MOPs where the objectives f_i are merely locally Lipschitz continuous.

In the single-objective case, *bundle methods* are regarded as the most efficient solution methods for nonsmooth problems [Kiw90; SZ92; BKM14]. As mentioned in Section 2.2.2, one of the challenges of nonsmooth optimization is the difficulty of approximating the subdifferential. Bundle methods are iterative methods where the idea is to only compute a single subgradient in each iteration while reusing subgradients computed in previous iterations. In this way, a *bundle* of subgradients is created which is used to approximate the subdifferential and compute descent directions. Similar to classical descent methods, these (approximated) descent directions are then combined with a line search method to generate a sequence of points along which f decreases. In [MKW14], this approach was generalized to the multiobjective setting, resulting in the *multiobjective proximal bundle method* (MPB). In the following, this method will be derived. In terms of evaluating the subdifferentials, we assume that at each point $x \in \mathbb{R}^n$, we can compute an (arbitrary) element from each $\partial f_i(x)$, $i \in \{1, \ldots, k\}$.

Let

$$H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad (x, y) \mapsto \max_{i\in\{1,\ldots,k\}} f_i(x) - f_i(y).$$

H is called the *improvement function* and has the following properties:

- If $H(x,y) < 0$, then $f_i(x) < f_i(y)$ for all $i \in \{1,\dots,k\}$, i.e., x is an improvement over y.
- $H(x,y) \geq 0$ for all $x \in \mathbb{R}^n$ if and only if y is weakly Pareto optimal. In particular, in this case, $y \in \arg\min_{x\in\mathbb{R}^n} H(x,y)$.

Thus, if x is not weakly Pareto optimal, then the solution of the problem

$$\min_{v\in\mathbb{R}^n} H(x+v,x) = \min_{v\in\mathbb{R}^n} \max_{i\in\{1,\dots,k\}} f_i(x+v) - f_i(x), \tag{2.26}$$

is a direction v in which all objective functions f_i decrease. To solve this problem, f will be approximated by a linearization based on subgradients. We will only consider the convex case here and refer to [MKW14] for the nonconvex case.

Assume that all f_i, $i \in \{1,\dots,k\}$, are convex. Then the function

$$v \mapsto \max_{i\in\{1,\dots,k\}} f_i(x+v) - f_i(x)$$

is convex as the maximum of convex functions. In particular, (2.26) is a convex problem. Note that (2.26) has two sources of nonsmoothness: the nonsmooth objectives f_i and the maximum function. To deal with the nonsmoothness of the f_i, we will use the bundle idea. To this end, assume that there are points y^j, $j \in \{1,\dots,N\}$, in which we already computed some subgradients $\xi_i^j \in \partial f_i(y^j)$, $i \in \{1,\dots,k\}$. Then the objectives can be linearized at these points via

$$\bar{f}_{i,j}(x) := f_i(y^j) + \langle \xi_i^j, x - y^j \rangle \quad \forall i \in \{1,..,k\}, j \in \{1,\dots,N\}.$$

Due to convexity of the objectives, each f_i can be approximated from below by the so-called *cutting-plane model*

$$f_i(x) \approx \hat{f}_i(x) = \max_{j\in\{1,\dots,N\}} \bar{f}_{i,j}(x).$$

Replacing the term $f_i(x+v)$ in the objective of (2.26) with the cutting-plane model yields

$$\begin{aligned} f_i(x+v) - f_i(x) \approx \hat{f}_i(x+v) - f_i(x) &= \max_{j\in\{1,\dots,N\}} \bar{f}_{i,j}(x+v) - f_i(x) \\ &= \max_{j\in\{1,\dots,N\}} f_i(y^j) + \langle \xi_i^j, x+v-y^j \rangle - f_i(x) \\ &= \max_{j\in\{1,\dots,N\}} \bar{f}_{i,j}(x) + \langle \xi_i^j, v \rangle - f_i(x) \\ &= \max_{j\in\{1,\dots,N\}} \alpha_{i,j}(x) + \langle \xi_i^j, v \rangle, \end{aligned}$$

where $\alpha_{i,j}(x) := \bar{f}_{i,j}(x) - f_i(x)$ is the so-called *linearization error*. Since the cutting-plane model is piecewise linear, a regularization term has to be added to (2.26) to make sure that the problem is bounded. The resulting problem is

$$\min_{v\in\mathbb{R}^n} \left(\max_{i\in\{1,\dots,k\}, j\in\{1,\dots,N\}} \alpha_{i,j}(x) + \langle \xi_i^j, v \rangle \right) + \frac{1}{2}\|v\|^2.$$

To get rid of the nonsmoothness in the objective function of this problem caused by the maximum function, we can rewrite it as a constrained smooth problem:

$$\begin{aligned} \min_{(v,\beta)\in\mathbb{R}^{n+1}} \quad & \beta + \frac{1}{2}\|v\|^2, \\ \text{s.t.} \quad & \alpha_{i,j}(x) + \langle \xi_i^j, v\rangle \leq \beta \quad \forall i \in \{1,\dots,k\}, j \in \{1,\dots,N\}. \end{aligned} \tag{2.27}$$

Since (2.27) is a quadratic problem (with linear inequality constraints), it can be solved efficiently.

It is important to note that since the problem (2.27) is only an approximation of the original problem (2.26), it cannot be guaranteed that the solution v of (2.27) actually yields a descent for f. If it does not, then the approximation of the current subdifferential is insufficient, and new subgradients close to the current point have to be computed, i.e., the current bundle has to be enriched. This can be handled in a line search method that either performs *serious steps*, if a descent of f can be achieved, or *null steps*, if the current direction v is not a descent direction. For details on the line search method and and the enrichment of the bundle, we refer to [MKW14]. In general, it can be shown that under some mild regularity assumption on f, all accumulation points of the sequence generated by the MPB are Pareto critical.

2.3.3 Set-based methods

All methods introduced up to this point are able to compute single Pareto optimal points, depending on an input parameter from some parameter space. That is, the result of the weighting method depends on the weighting vector $\alpha \in \Delta_k$, the result of the ε-constraint method depends on the $\varepsilon \in \mathbb{R}^{k-1}$ and the results of both the steepest descent method and the MPB depend on the initial points $x^1 \in \mathbb{R}^n$. In all these methods it is clear that different input parameters will (mostly) produce different solutions. But as emphasized in Section 2.1, the actual solution of an MOP is the Pareto set, i.e., the set of *all* Pareto optimal points. A naive approach to compute an approximation of the Pareto set would be to apply one of the above methods for many different input parameters. But since the relationship between the input parameter and the resulting solution is unknown, there is no way of knowing a priori how the parameter space should be discretized to obtain a good approximation. Thus, a more sophisticated approach is needed. To this end, this section will introduce the *subdivision method* [DSH05; DH97] and the *continuation method* [Hil01], which are set-based methods that are able to approximate the entire Pareto set.

Subdivision method

The steepest descent method (2.25) can be written as

$$x^{j+1} = g(x^j), \quad j \in \mathbb{N}, \tag{2.28}$$

for some initial point $x^1 \in \mathbb{R}^n$ and

$$g : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto x + \bar{t}(x)\bar{v}(x).$$

For a general $g : \mathbb{R}^n \to \mathbb{R}^n$, (2.28) is known as a *(discrete, autonomous) dynamical system*. In the following, some basic definitions for dynamical systems will be introduced.

Definition 2.3.5. *a) Let $x \in \mathbb{R}^n$. If $g(x) = x$, then x is called a* fixed point of (2.28)*. For a set $Q \subseteq \mathbb{R}^n$ let F_Q be the set of fixed points of g in Q.*

b) Let $Q \subseteq \mathbb{R}^n$ be compact and let g^j denote the j-times composition of g. Then

$$A_Q := Q \cap \bigcap_{j \in \mathbb{N}} g^j(Q)$$

is called the global attractor of (2.28) relative to Q.

Due to the construction of the steepest descent method, the Pareto critical set is the set of fixed points of (2.28). Clearly, all fixed points in a compact set Q are contained in the global attractor A_Q, i.e., $F_Q \subseteq A_Q$. Thus, if the Pareto critical set P_c is bounded and Q is chosen such that $P_c \subseteq Q$, then $P_c = F_Q \subseteq A_Q$.

To compute A_Q, the *subdivision method* can be used, which was originally introduced in [DH97] for general discrete dynamical systems. The idea is to iteratively divide Q into compact subsets of decreasing diameter (*subdivision step*) while removing subsets that have an empty intersection with A_Q (*selection step*). If g is a homeomorphism, then $A_Q \subseteq g(A_Q)$ (see [DH97]). In this case, to decide whether a set $B \subseteq \mathbb{R}^n$ has an empty intersection with A_Q, we can use the fact that if $Q' \subseteq \mathbb{R}^n$ is any superset of A_Q, then

$$B \cap A_Q \subseteq B \cap g(A_Q) \subseteq B \cap g(Q'),$$

so

$$B \cap A_Q \neq \emptyset \;\Rightarrow\; B \cap g(Q') \neq \emptyset. \tag{2.29}$$

In particular, the right-hand side of (2.29) can be used as a necessary condition for $B \cap A_Q \neq \emptyset$, and the condition gets more strict the smaller $Q' \setminus A_Q$. Since we start the subdivision method with the set Q and only remove subsets of $Q \setminus A_Q$ from Q, the current approximation of A_Q will always be a superset of A_Q and can thus be used for Q' in (2.29). For ease of implementation, it makes sense to choose Q as a hypercube (which will be referred to as a *box*) and to divide Q evenly into 2^n smaller boxes during the subdivision. (In theory, more general compact sets Q and subdivision schemes can be used, cf. [DH97].) Formally, the subdivision method can be denoted as in Algorithm 2.1.

For Q_i, $i \in \mathbb{N}$, as in Algorithm 2.1, let

$$Q_\infty := \bigcap_{i \in \mathbb{N}} Q_i.$$

By construction we have $Q_{i+1} \subseteq Q_i$ for all $i \in \mathbb{N}$, so in particular $Q_\infty \subseteq Q_i$ for all $i \in \mathbb{N}$. In other words, Q_∞ can be interpreted as the "limit" of the sequence $(Q_i)_{i \in \mathbb{N}}$. Regarding the convergence of Algorithm 2.1, the following result can be shown [DH97].

Theorem 2.3.6. *a) $F_Q \subseteq Q_\infty$.*

b) If g is a homeomorphism, then $A_Q = Q_\infty$.

Algorithm 2.1 Subdivision method

Given: Function $g : \mathbb{R}^n \to \mathbb{R}^n$, box $Q \subseteq \mathbb{R}^n$.

1: Initialize $\mathcal{B}_1 = \{Q\}$.
2: **for** $i \in \mathbb{N}$ **do**
3: *Subdivision step:* Divide each box in $\mathcal{B}_i$ evenly into 2^n smaller boxes (by halving along each dimension). Let $\mathcal{B}'_{i+1}$ be the resulting collection of boxes.
4: *Selection step:* Remove all boxes from $\mathcal{B}'_{i+1}$ that do not contain part of A_Q via (2.29), i.e., set

$$\mathcal{B}_{i+1} = \{B \in \mathcal{B}'_{i+1} : B \cap g(Q_i) \neq \emptyset\},$$

where $Q_i = \bigcup_{B \in \mathcal{B}_i} B$.
5: **end for**

By Theorem 2.3.6 a), if g is the dynamical system induced by the steepest descent method (2.25), then Q_∞ is a superset of the Pareto critical set P_c. Unfortunately, part b) of this theorem cannot be applied as in general, g is not a homeomorphism: although the descent direction $\bar{v}$ is continuous (cf. Lemma 2.3.3), g must neither be injective nor surjective. Thus, although P_c is contained in all Q_i, we can generally not expect that $P_c = Q_\infty$.

In practice, the image $g(Q_i)$ in the selection step cannot be computed exactly. Instead, $g(Q_i)$ is approximated by replacing Q_i with a finite set of test points in Q_i, e.g., by discretizing each box in $\mathcal{B}_i$. The resulting behavior of the practical implementation of the subdivision method for the solution of MOPs is discussed in the following remark.

Remark 2.3.7. *Intuitively, one might expect that replacing $g(Q_i)$ with the image of a finite number of test points would lower the quality of the result of Algorithm 2.1. On the one hand this is true, since it can happen that $B \cap g(Q_i) \neq \emptyset$ without there being a test point in $g^{-1}(B) \cap Q_i$. In other words, it can happen that boxes get falsely removed. On the other hand, note that we only have $P_c \subseteq Q_\infty$ and in general, there will be points in Q_∞ that are not Pareto critical. In the language of dynamical systems, these are, for example, points on the* unstable manifolds *(cf. [DH97]) of fixed points. In the language of optimization, these points arise due to the existence of Pareto critical points that are not locally weakly Pareto optimal. If x^* is such a point and if U is any open neighborhood of x^*, then part of U will always get mapped away from x^*, potentially into boxes that do not contain part of P_c. Thus, if the image of Q_i is merely approximated, then there is a chance that this behavior is prevented. Furthermore, with the same reasoning, it can happen that parts of P_c get removed which are not attractive in terms of the dynamics of the descent method, i.e., which are not Pareto optimal. Thus, even though the approximation of Q_i via test points lowers the quality of the approximation of Q_∞, it can improve the approximation of the actual Pareto set.*

The subdivision method can be seen as an approximation of the Pareto critical set "from the outside", since we start with a large box $Q \supseteq P_c$ and iteratively remove parts of Q that are not Pareto critical. In the following part, we will consider the opposite approach, i.e., an approximation "from the inside".

Continuation method

In the following, the *continuation method* (also known as the *homotopy method*) from [Hil01] will be introduced. The goal of this method is to approximate the manifold $\mathcal{M}$ from Theorem 2.2.10, i.e., the set

$$\mathcal{M} = \{(x, \alpha) \in \mathbb{R}^n \times \Delta_k : Df(x)^\top \alpha = 0,\ \alpha_i > 0\ \forall i \in \{1, \dots, k\}\}.$$

If the rank assumption in Theorem 2.2.10 is satisfied and $(x_0, \alpha_0) \in \mathcal{M}$, then $\mathcal{M}$ being a manifold means that there is a smooth local parametrization of $\mathcal{M}$ around (x_0, α_0) via $\mathbb{R}^{k-1}$, i.e., there is some open set $U \subseteq \mathbb{R}^{n+k}$ with $(x_0, \alpha_0) \in U$, an open set $V \subseteq \mathbb{R}^{k-1}$ and a C^1-diffeomorphism $\varphi : V \to \mathcal{M} \cap U$ (cf. [Lee12], Proposition 5.23). The idea of the continuation method is to construct a local parametrization that can be numerically evaluated. This allows for the computation of new points on $\mathcal{M}$ close to (x_0, α_0). By repeating this process in the newly found points, $\mathcal{M}$ can be further and further explored.

The basis of this method is the following theorem (cf. [Hil01], Theorem 5.7).

Theorem 2.3.8. *Let $(x_0, \alpha_0) \in \mathcal{M}$ with $rk(DF(x_0, \alpha_0)) = n+1$. Let $\{q_1, \dots, q_{n+k}\}$ be an orthonormal basis of $\mathbb{R}^{n+k}$ such that $\operatorname{span}(\{q_1, \dots, q_{k-1}\}) = T_{(x_0,\alpha_0)}\mathcal{M}$. Let $Q \in \mathbb{R}^{(n+k)\times(n+k)}$ be the orthogonal matrix with columns $q_1, \dots, q_{n+k}$. Then there is*

- *an open set $V \subseteq \mathbb{R}^{k-1}$ with $0 \in V$,*
- *an open set $U \subseteq \mathbb{R}^{k+n}$ with $(x_0, \alpha_0) \in U$,*
- *an open set $W \subseteq \mathbb{R}^{n+1}$ with $0 \in W$ and*
- *a continuously differentiable function $\eta : V \to W$ with $\eta(0) = 0$ and $D\eta(0) = 0$*

such that

$$\varphi : V \to \mathcal{M} \cap U, \quad \xi \mapsto (x_0, \alpha_0)^\top + Q \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix}$$

is a smooth local parametrization of $\mathcal{M}$ at (x_0, α_0).

In the following, we will discuss how φ in the previous theorem can be evaluated in practice. By Theorem 2.2.10, we have $T_{(x_0,\alpha_0)}\mathcal{M} = \ker(DF(x_0, \alpha_0))$. Thus, an orthonormal basis $\{q_1, \dots, q_{n+k}\}$ can be obtained by reordering the right-singular vectors of $DF(x_0, \alpha_0)$ (or via a QR-factorization of $DF(x_0, \alpha_0)^\top$ as in [Hil01]). (In particular, this can be used to check if the rank condition from Theorem 2.2.10 is satisfied.) Note that

$$\varphi(\xi) = (x_0, \alpha_0)^\top + Q \begin{pmatrix} \xi \\ \eta(\xi) \end{pmatrix} = (x_0, \alpha_0)^\top + \sum_{i=1}^{k-1} \xi_i q_i + \sum_{i=k}^{n+k} \eta_i(\xi) q_i$$

with

$$\sum_{i=1}^{k-1} \xi_i q_i \in T_{(x_0,\alpha_0)}\mathcal{M} \quad \text{and} \quad \sum_{i=k}^{n+k} \eta_i(\xi) q_i \in (T_{(x_0,\alpha_0)}\mathcal{M})^\perp.$$

So φ can be evaluated by first moving from (x_0, α_0) in the tangent direction induced by ξ, and then moving in the direction orthogonal to the tangent space induced by $\eta(\xi)$. Since Theorem 2.3.8 does not give an explicit expression for η, $\eta(\xi)$ has to be approximated. This can be done by recalling that $\varphi(\xi)$ has to satisfy

$$F(\varphi(\xi)) = 0$$

with F as in (2.11), so $\eta(\xi)$ is a root of the function

$$\mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \quad \eta \mapsto F\left((x_0, \alpha_0)^\top + \sum_{i=1}^{k-1} \xi_i q_i + \sum_{i=k}^{n+k} \eta_i q_i\right). \tag{2.30}$$

In [Hil01], $\eta(\xi)$ was approximated by applying Newton's method to solve this root-finding problem.

By choosing multiple different ξ, the above procedure to evaluate φ can be used to generate new points in $\mathcal{M}$. Since $\{q_1, \ldots, q_{k-1}\}$ was constructed as an orthonormal basis of $T_{(x_0,\alpha_0)}\mathcal{M}$, the unit vectors $\xi^j = e_j \in \mathbb{R}^{k-1}$, $j \in \{1, \ldots, k-1\}$, can be chosen, such that

$$\sum_{i=1}^{k-1} \xi_i^j q_i = \sum_{i=1}^{k-1} (e_j)_i q_i = q_j, \quad j \in \{1, \ldots, k-1\},$$

yields orthonormal directions in $T_{(x_0,\alpha_0)}\mathcal{M}$. Furthermore, in [Hil01] it was shown that the unit vectors can be adaptively scaled in a way that ensures that the individual distance of the images of the resulting Pareto critical points to the point $f(x_0)$ is constant, i.e., such that an even discretization of the Pareto front can be achieved.

In [SDD05], a version of the continuation method based on boxes was introduced, which will be discussed in Section 3.1.

2.3.4 Evolutionary methods

In theory, all methods introduced so far are able to compute points which are exactly Pareto optimal, and in practice, they are able to generate points which are arbitrarily close to Pareto optimal points. While convergence to Pareto optimal points is the key property of solutions methods from a theoretical point of view, there are applications where exact solutions are not needed and good approximations are sufficient. Furthermore, there are cases where the generation of exact solutions is infeasible due to the complexity of the problem. This creates the need for methods that are able to generate good approximations efficiently, even if they might not converge to exact solutions. Historically, the most popular methods in this area are *evolutionary algorithms* (EAs) [ZLB04; Deb01], which are based on stochastic optimization.

The idea of EAs is to consider a finite set of points in the variable space and iteratively modify it until its image forms a close and well-spread approximation of the Pareto front. Since EAs are inspired by natural evolution, each point is referred to as an *individual*, and the set of individuals is called the *population*. In each iteration, which is referred to as a *generation*, the population is modified through a *recombination* and a *mutation operator*. The recombination operator creates a number of new individuals (*children*) by combining features of a number of existing individuals (*parents*). The mutation operator takes a single individual and applies

a small change to it. To decide which individuals are used for recombination and mutation, *selection operators* are used. They are based on a *fitness function*, which assigns to each individual a scalar value based on their quality. For example, the fitness value of an individual could be the number of other individuals that have smaller objective function values for all objectives [FF93]. Based on these ideas, a general EA consists of the following steps:

1. Randomly initialize the population.
2. Compute the fitness function value of each individual.
3. Create new individuals through recombination and mutation.
4. Choose a subset of new and old individuals to create a new population.
5. Go to step 2.

Clearly, the above procedure is merely an abstract algorithm, and the specific choices of fitness function and evolution operators produce different EAs. Among the most popular EAs are *NSGA-II (Non-dominated Sorting Genetic Algorithm II)* [Deb+02] and *SPEA-2 (Strength Pareto Evolutionary Algorithm 2)* [ZLT01], but there is a vast amount of additional algorithms in the literature (see [Zho+11] for a survey).

3 Box-continuation methods for smooth problems

If the premise of Theorem 2.2.10 holds, then the set $\mathcal{M}$ is a manifold with a tangent space that is easy to compute. Since the projection of $\mathcal{M}$ onto its first n components is the set of Pareto critical points (with positive KKT vectors), this smoothness can be exploited for the solution of MOPs. As already seen in Section 2.3.3, if a single Pareto critical point is given, then new Pareto critical points in its vicinity can be computed by moving in tangent directions of $\mathcal{M}$. This is the basic idea of continuation methods [Hil01; SDD05; MS17; BCS20].

This chapter introduces new continuation methods which are inspired by the methods from [Hil01] (Section 2.3.3) and [SDD05]. We begin by deriving a method for the case where exact gradients are available (Section 3.1), which is similar to the method from [SDD05]. Afterwards, we consider the case where only inexact gradients are at hand together with an upper bound for the error, and construct an algorithm that is able to compute a tight superset of the Pareto critical set in this setting (Section 3.2).

Parts of Section 3.2 have been previously published in Section 2 of [Ban+19], to which the author of this thesis was the main contributor.

3.1 Exact gradients

Throughout this section, we assume that f satisfies the requirements of Theorem 2.2.10 a), such that $\mathcal{M}$ is a manifold. While from a local point of view, the continuation method from [Hil01] is able to produce well spread points on the set $\mathcal{M}$, there are certain difficulties that arise when trying to compute the complete set:

- By construction, the method is only able to compute the connected component of $\mathcal{M}$ that contains the starting point (x_0, α_0). Thus, if $\mathcal{M}$ consists of multiple connected components, then the method has to be applied multiple times with one starting point from each component. This either requires a priori information about the structure of $\mathcal{M}$, or a way to check if a new starting point lies in a component that was already computed. Since the approximation of $\mathcal{M}$ is based on a set of points, the latter approach requires a way of determining if a point lies within a smooth object (of known dimension) discretized by a finite set of points, which is a non-trivial task.

- Even if $\mathcal{M}$ is connected, it can happen that the method reaches the same area twice. Thus, as a stopping criterion, a mechanism is needed that recognizes

if an area of $\mathcal{M}$ has already been explored. Since the approximation of $\mathcal{M}$ is point based, this causes the same difficulties as mentioned above.

- In [Hil01], the system (2.30) is solved via Newton's method. This means that the resulting approximation (x^*, α^*) of $\varphi(\xi)$ can have negative multipliers α_i^*, $i \in \{1, \dots, k\}$. (This can, for example, be seen in Figure 7.2 in [Hil01].) As a result, the method can produce points which are not actually contained in $\mathcal{M}$.

In this section, a continuation method will be derived that resolves these difficulties by approximating the Pareto critical set using a collection of hypercubes (or *boxes*) instead of a set of points. The resulting *box-continuation method* will look similar to the method from [SDD05], but will have a key difference that we will discuss later (cf. Remark 3.1.4). We begin by deriving the idea of our method in Section 3.1.1 before presenting the resulting algorithm in Section 3.1.2. Afterwards, we apply our method to some numerical examples in Section 3.1.3. Finally, we discuss extensions to constrained MOPs (Section 3.1.4) and the computation of the actual Pareto set (Section 3.1.5)

3.1.1 Covering via boxes

In the following, we will introduce the division of the variable space $\mathbb{R}^n$ into boxes that forms the basis of the box-continuation method. To this end, for a *radius* $r > 0$ and an *anchor point* $a \in \mathbb{R}^n$ let

$$\mathcal{B} := \{[-r, r]^n + (2i_1 r, \dots, 2i_n r)^\top + a : (i_1, \dots, i_n) \in \mathbb{Z}^n\}.$$

Clearly, $\bigcup_{B \in \mathcal{B}} B = \mathbb{R}^n$ and two different boxes from $\mathcal{B}$ can only intersect in their lower dimensional faces. For $B_1, B_2 \in \mathcal{B}$, we say that B_1 is a *neighboring box of* B_2, if $B_1 \neq B_2$ and $B_1 \cap B_2 \neq \emptyset$.

The goal of the box-continuation method is to compute the collection of boxes in $\mathcal{B}$ with a nonempty intersection with the Pareto critical set P_c, i.e., to compute

$$\mathcal{B}_c := \{B \in \mathcal{B} : B \cap P_c \neq \emptyset\}.$$

To check if a box $B \in \mathcal{B}$ has a nonempty intersection with P_c, we can use the problem

$$\begin{aligned} \theta(B) := & \min_{x \in B, \alpha \in \Delta_k} \|Df(x)^\top \alpha\|^2 \\ = & \min_{(x,\alpha) \in \mathbb{R}^{n+k}} \alpha^\top Df(x) Df(x)^\top \alpha, \\ & s.t. \quad \sum_{i=1}^{k} \alpha_i = 1, \\ & \qquad \alpha_i \geq 0 \quad \forall i \in \{1, \dots, k\}, \\ & \qquad x \in B. \end{aligned} \tag{3.1}$$

Then

$$B \in \mathcal{B}_c \quad \Leftrightarrow \quad \theta(B) = 0. \tag{3.2}$$

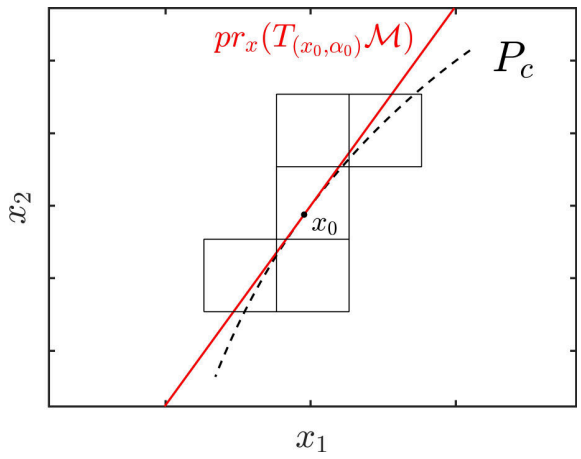


Figure 3.1: A sketch of the idea of the box-continuation method. The red line is the projection of the tangent space of $\mathcal{M}$ (attached to x_0) and the dashed line shows the Pareto critical set.

Given some initial box $B \in \mathcal{B}_c$, the idea of the box-continuation method is to compute all neighboring boxes of B in $\mathcal{B}_c$ and to then repeat this process with the new found boxes. Depending on the number of variables n and objectives k, we distinguish between the following two cases:

- $k \geq n+1$: In this case, the dimension of $\mathcal{M}$ is $k-1 \geq n$. Thus, as a superset of $\mathrm{pr}_x(\mathcal{M})$ (with pr_x as in (2.12)), the Pareto critical set is typically a set with positive measure in $\mathbb{R}^n$. Here, we just solve (3.1) for all neighboring boxes of B.

- $k < n+1$: In this case, the dimension of $\mathcal{M}$ is $k-1 < n$, so $\mathrm{pr}_x(\mathcal{M})$ and P_c are typically null sets in $\mathbb{R}^n$. In particular, solving (3.1) for all neighboring boxes is generally inefficient, since only few boxes will actually be contained in $\mathcal{B}_c$. Instead, we will first use the tangent space of the manifold $\mathcal{M}$ to obtain good candidates for neighboring boxes in $\mathcal{B}_c$, and then solve (3.1) only for those candidates.

The realization of the first case is straight-forward. (Note that for this case, $\mathcal{M}$ does not even have to be a manifold.) For the second case, we have to discuss the relationship between the tangent space of $\mathcal{M}$ and the "projected smooth structure" of $P_c \supseteq \mathrm{pr}_x(\mathcal{M})$.

Let $(x_0, \alpha_0) \in \mathcal{M}$. Since we are only interested in the Pareto critical set P_c (and not the augmented set $\mathcal{M}$), the plan is to only use the projection $\mathrm{pr}_x(T_{(x_0,\alpha_0)}\mathcal{M})$ of the tangent space onto the first n components to practically obtain the "tangent space" of P_c at x_0. A sketch of this idea is shown in Figure 3.1. This poses two theoretical problems:

(i) In general, the KKT vector α_0 of x_0 is not unique. This means that if we choose a different KKT vector α' of x_0, then we might have $T_{(x_0,\alpha_0)}\mathcal{M} \neq T_{(x_0,\alpha')}\mathcal{M}$. In other words, the tangent information we project onto the variable space $\mathbb{R}^n$ might depend on the chosen KKT vector.

(ii) When projecting from $\mathbb{R}^n \times \mathbb{R}^k$ onto $\mathbb{R}^n$, we lose tangent information that is orthogonal to $\mathbb{R}^n$, i.e., tangent vectors of the form $(0, \beta)^\top \in T_{(x_0,\alpha_0)}\mathcal{M}$ will get mapped to $0 \in \mathbb{R}^n$.

Both of these problems are addressed in the following lemma.

Lemma 3.1.1. *Let $x_0 \in P_c$ and $\alpha_0 \in \Delta_k$ be a KKT vector of x_0. If*

a) α_0 is not unique or

b) $(\alpha_0)_i > 0$ for all $i \in \{1, \dots, k\}$ and there is $v \in T_{(x_0,\alpha_0)}\mathcal{M}$ with $\mathrm{pr}_x(v) = 0$,

then

$$\mathrm{rk}(Df(x_0)) < k - 1.$$

Proof. a) Let $\alpha' \in \Delta_k$ be another KKT vector of x_0 and define $\beta = \alpha_0 - \alpha' \neq 0$. Then

$$Df(x_0)^\top \beta = Df(x_0)^\top \alpha_0 - Df(x_0)^\top \alpha' = 0,$$

so $\beta \in \ker(Df(x_0)^\top)$. Furthermore, β and α_0 are linearly independent, since for $\lambda \in \mathbb{R}$ we have

$$\beta = \lambda\alpha_0 \;\Rightarrow\; 0 = \sum_{i=1}^{k}(\alpha_0)_i - \sum_{i=1}^{k}(\alpha')_i = \sum_{i=1}^{k}\beta_i = \lambda\sum_{i=1}^{k}(\alpha_0)_i = \lambda,$$

which is a contradiction. So we must have $\dim(\ker(Df(x)^\top)) > 1$. By the rank-nullity theorem, this implies

$$k = \dim(\ker(Df(x_0)^\top)) + \mathrm{rk}(Df(x_0)^\top) > 1 + \mathrm{rk}(Df(x_0)^\top),$$

so $\mathrm{rk}(Df(x_0)) = \mathrm{rk}(Df(x_0)^\top) < k - 1$.
b) Let $v \in T_{(x_0,\alpha_0)}\mathcal{M}$ with $\mathrm{pr}_x(v) = 0$, i.e., $v = (0, \beta)^\top$ for some $\beta \in \mathbb{R}^k \setminus \{0\}$. By Theorem 2.2.10 we have $T_{(x_0,\alpha_0)}\mathcal{M} = \ker(DF(x_0, \alpha_0))$, so

$$0 = DF(x_0, \alpha_0)v = \begin{pmatrix} \sum_{i=1}^{k}(\alpha_0)_i\nabla^2 f_i(x_0) & Df(x_0)^\top \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ \beta \end{pmatrix} = \begin{pmatrix} Df(x_0)^\top\beta \\ \sum_{i=1}^{k}\beta_i \end{pmatrix}.$$

Thus, the proof follows as in a). □

By the definition of the Pareto critical set P_c, we have

$$P_c \subseteq \{x \in \mathbb{R}^n : \mathrm{rk}(Df(x)) \leq k - 1\}.$$

Before we discuss the implications of Lemma 3.1.1, we recall the following definition from [Sma73].

Definition 3.1.2. *A point $x \in \mathbb{R}^n$ is said to* satisfy the rank assumption *if* $\mathrm{rk}(Df(x)) = k-1$. *If the rank assumption is satisfied for all $x \in P_c$, then f* satisfies the rank assumption.

According to Lemma 3.1.1, if f satisfies the rank assumption, then the problems (i) and (ii) from above cannot occur. In [Sma73] it was argued that in a certain natural topology on the space of continuously differentiable functions, almost all points $x \in P_c$ satisfy the rank assumption for almost all f. If $\frac{n+4}{2} > k$, then we have the stronger result that almost all f actually satisfy the rank assumption. (For

example, if $k = 2$ with arbitrary $n \in \mathbb{N}$, then the rank assumption can only be violated in some $x \in \mathbb{R}^n$ if $\operatorname{rk}(Df(x)) = 0$, i.e., if the gradients of both objectives are zero in x.) Thus, we expect that the potential problems (i) and (ii) can be neglected in practice.

Note that the proof of Lemma 3.1.1 shows in particular that in both cases a) and b), there exists an element $(0, \beta)^\top \in T_{(x_0,\alpha_0)}\mathcal{M}$ with $\beta \neq 0$. In practice, this can be detected while computing the tangent space. So even if the algorithm reaches a point which violates the rank assumption, the potential problems could be prevented by just solving (3.1) for all neighboring boxes in that point.

In Section 5.1.2, we will further investigate the relationship between the tangent information of $\mathcal{M}$ and P_c by considering tangent vectors of P_c.

3.1.2 The algorithm

In the following, we will discuss the algorithm for our box-continuation method and its practical implementation. First, we need some additional notations. For $x \in \mathbb{R}^n$ let

$$B(x) := \{B \in \mathcal{B} : x \in B\},$$
$$N(x) := \{B \in \mathcal{B} \setminus B(x) : \exists B' \in B(x) \text{ with } B' \cap B \neq \emptyset\}.$$

Then $B(x)$ is the collection of boxes containing x and $N(x)$ is the collection of neighboring boxes of boxes containing x. Clearly, for a generic $x \in \mathbb{R}^n$, $B(x)$ contains only a single box and $N(x)$ contains $3^n - 1$ boxes.

The pseudo code for the box-continuation method is shown in Algorithm 3.1. The resulting collection of boxes $\mathcal{C}$ is an approximation of $\mathcal{B}_c$, and the set $X_\mathcal{C}$ contains one Pareto critical point from each box in $\mathcal{C}$. The following remark discusses the practical implementation of the algorithm.

Algorithm 3.1 Box-continuation method

Given: $(x_0, \alpha_0) \in \mathcal{M}$.

1: Initialize $\mathcal{C} = \{B(x_0)\}$, $X_\mathcal{C} = \{(x_0, \alpha_0)\}$, $\mathcal{C}_{out} = \emptyset$ and a queue $Q = \{(x_0, \alpha_0)\}$.
2: **while** $Q \neq \emptyset$ **do**
3: Remove the first element $(\bar{x}, \bar{\alpha})$ from Q.
4: If $k \geq n+1$ then set $T_\mathcal{B} := N(\bar{x})$. Otherwise, compute the neighboring boxes with a nonempty intersection with the projected tangent space, i.e.,

$$T_\mathcal{B} := \{B \in N(\bar{x}) : B \cap (\bar{x} + \operatorname{pr}_x(T_{(\bar{x},\bar{\alpha})}\mathcal{M})) \neq \emptyset\}.$$

5: **for** $B \in T_\mathcal{B} \setminus (\mathcal{C} \cup \mathcal{C}_{out})$ **do**
6: Compute the optimal value $\theta(B)$ and the solution (x_B, α_B) of (3.1).
7: If $\theta(B) = 0$, then set $Q = Q \cup \{(x_B, \alpha_B)\}$, $\mathcal{C} = \mathcal{C} \cup \{B\}$ and $X_\mathcal{C} = X_\mathcal{C} \cup \{(x_B, \alpha_B)\}$. Otherwise, set $\mathcal{C}_{out} = \mathcal{C}_{out} \cup \{B\}$.
8: **end for**
9: **end while**

Remark 3.1.3. *a) In Algorithm 3.1, the method is initialized with only a single starting point $(x_0, \alpha_0) \in \mathcal{M}$. Similar to what was discussed at beginning of this*

chapter, this means we can only approximate the connected component of P_c that contains x_0. To approximate all connected components of P_c, we merely have to add multiple points to Q during the initialization, such that Q contains at least one point from each component. A heuristic way of obtaining these points is to solve (3.1) *for all boxes in a coarse box covering of (some bounded subset of) $\mathbb{R}^n$ with large radius r.*

b) *Multiple starting points are also needed if $\mathcal{M}$ is only locally a manifold (in the sense of Theorem 2.2.10 b)), as there might be kinks in P_c that the method cannot move through smoothly.*

c) *By Theorem 2.2.10 we have $T_{(\bar{x},\bar{\alpha})}\mathcal{M} = \ker(DF(\bar{x},\bar{\alpha}))$. In practice, the kernel can be computed via a singular value decomposition of $DF(\bar{x},\bar{\alpha})$. More precisely, the kernel is spanned by the right-singular vectors corresponding to the singular values which are zero. (In particular, if the number of nonzero singular values is smaller than $n+1$, then $\mathcal{M}$ might not be a manifold around $(\bar{x},\bar{\alpha})$.)*

d) *Since $T_{(\bar{x},\bar{\alpha})}\mathcal{M}$ is a $(k-1)$-dimensional vector space, the set $T_{\mathcal{B}}$ can be obtained by computing the intersection of boxes in $N(\bar{x})$ with a pointwise discretization of $\mathrm{pr}_x(\ker(DF(\bar{x},\bar{\alpha})))$ if k is small. For larger k this becomes inefficient, and it might be better to consider a geometrical approach. For example, a $(k-1)$-dimensional subspace of $\mathbb{R}^n$ intersects a box if and only if it intersects one of its $(n-k+1)$-dimensional faces. Based on this, one could characterize the intersection for each neighboring box by a set of linear equations.*

e) *For the relation* (3.2) *to hold, it is important that we find global minimal points of* (3.1)*. If the box radius r is sufficiently small, then standard local methods for constrained nonlinear problems are typically satisfactory. In general, if* (3.1) *possesses local solutions, ideas from global optimization should additionally be applied (cf. [TŽ89]).*

f) *In practice, the condition $\theta(B) = 0$ has to be replaced by checking if $\theta(B)$ lies below a certain threshold, depending on the accuracy of the solver that is used for* (3.1)*. As accurate solutions of* (3.1) *are crucial for making sure that Algorithm 3.1 does not stop prematurely, it can be beneficial to apply multiple solvers in a staggered way: If for the first solver, $\theta(B)$ is larger than the chosen threshold, then a second solver is applied. Only if the second solver also returns a value larger than the threshold, the box B is discarded. On top of that, note that the existence of a vanishing convex combination of the gradients is independent of the scaling of the gradients. In some cases, rescaling can lead to better solutions, for example by normalizing the (nonzero) gradients.*

The box-continuation method avoids the problems of the classical continuation method from [Hil01] that were discussed at the beginning of Section 3.1: By using a covering with n-dimensional boxes instead of a set of points to approximate the Pareto critical set, we can easily check if we already explored a specific area of P_c. By using (2.12) instead of the corrector step from [Hil01], we can also assure that all boxes in the covering $\mathcal{C}$ actually contain part of P_c.

The convergence of Algorithm 3.1, i.e., its ability to compute $\mathcal{B}_c$, only depends on the ability of solving (3.1) and the assumption that $T_{\mathcal{B}}$ in step 4 actually contains all

neighboring boxes in $\mathcal{B}_c$, i.e., that $T_{\mathcal{B}} = \{B \in N(\bar{x}) : B \cap P_c \neq \emptyset\}$. The former was already discussed Remark 3.1.3. The latter depends on how well P_c is approximated by $\bar{x} + \mathrm{pr}_x(T_{(\bar{x},\bar{\alpha})}\mathcal{M})$ for different $(\bar{x}, \bar{\alpha}) \in \mathcal{M}$ with $\bar{x} \in B$. Since we assume $\mathcal{M}$ to be a manifold, this approximation becomes better the smaller the box radius r. Thus, if r is small enough and we are able to compute (global) solutions of (3.1), then the convergence of Algorithm 3.1 is guaranteed.

Finally, the following remark discusses the difference of Algorithm 3.1 to the continuation method from [SDD05].

Remark 3.1.4. *The main difference between Algorithm 3.1 and Algorithm* CONT-Recover *from [SDD05] is the way new boxes are added to the collection. In Algorithm 3.1, this is done by checking if $\theta(B)$ is (close to) zero for $B \in T_{\mathcal{B}}$. In [SDD05], it is done by using points in the tangent direction as starting points for a local method for minimizing F (cf.* (2.11)*) via least-squares, namely the* Gauss-Newton algorithm *(see, e.g., [NW06]). For this approach to work, one has to rely on the assumption that this local method naturally finds a point in P_c that is close to the starting point in a part of P_c that is not yet covered by the box covering. While this holds for well-behaved examples, it is not assured that it works in the general case. Furthermore, non-negativity of the KKT vector α at the result is not guaranteed.*

In [Sch04], where the method from [SDD05] was first suggested, the steepest descent method (cf. Section 2.3.2) was used instead. Clearly, this only allows for the computation of Pareto critical points which are at least locally Pareto optimal. But even then, this approach suffers from the same problems as the Gauss-Newton algorithm, as descent directions are not necessarily orthogonal to P_c, even in points that are close to P_c. This can be seen by considering the problem (2.10) *from Example 2.2.8 for*

$$c = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad Q = \begin{pmatrix} 4.1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Figure 3.2 shows the normalized descent direction resulting from (2.19) *for this problem for* 25 *points in $[-0.1, 0] \times [0.1, 0.2]$. It can be seen that the descent directions*

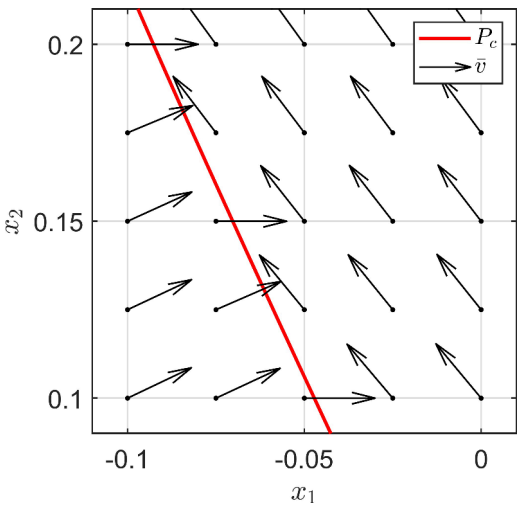


Figure 3.2: Normalized descent directions close to the Pareto critical set P_c for the problem in Remark 3.1.4.

in points to the right of P_c are not orthogonal to P_c, even in points which are almost Pareto critical.

3.1.3 Examples

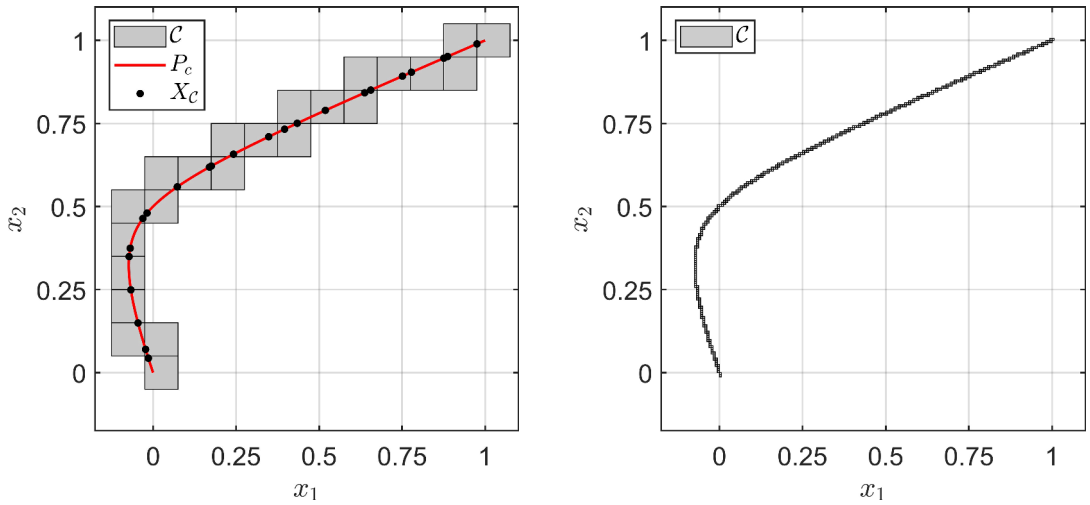
In this section, Algorithm 3.1 will be applied to some examples. For solving the problem (3.1) to decide if a box contains part of the Pareto critical set, the implementation of the *SQP method* (cf. [NW06]) in the MATLAB function `fmincon` is used. A box B is considered to contain part of P_c if $\theta(B) < 10^{-5}$. The initialization was done as described in Remark 3.1.3 a). In this section, the examples will be restricted to the case where $k \leq n$, since this is the more interesting case for our method (and the more realistic case in practice).

The first problem we consider is the simple MOP from Example 2.2.8.

Example 3.1.5. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} \|x\|_Q^2 \\ \|x-c\|^2 \end{pmatrix} = \begin{pmatrix} 5x_1^2 + 4x_1x_2 + x_2^2 \\ (x_1-1)^2 + (x_2-1)^2 \end{pmatrix}$$

from Example 2.2.8. Figure 3.3 shows the result of Algorithm 3.1 for different box radii. Figure 3.3(a) additionally shows the exact solution and the set $X_{\mathcal{C}}$, which contains one Pareto critical point in each box in $\mathcal{C}$. As expected, $\mathcal{C}$ is a tight covering of the Pareto critical set.



(a) $r = 0.05$ **(b)** $r = 0.3125 \cdot 10^{-2}$

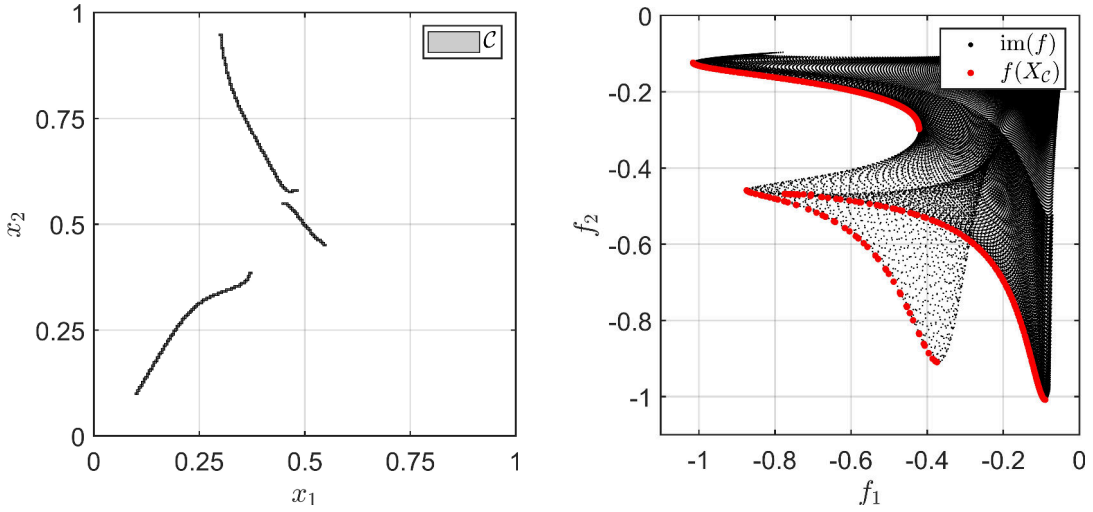
Figure 3.3: Result of Algorithm 3.1 for Example 3.1.5 for different box radii r. For comparison, (a) additionally shows the exact solution P_c and the set of points $X_{\mathcal{C}}$ resulting from Algorithm 3.1.

Next we consider Example 1.3 from [PŽŽ17], which was purposefully constructed to have a disconnected Pareto set. It consists of two so-called *Shekel functions*.

Example 3.1.6. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \textit{ with } f(x) = \begin{pmatrix} -\frac{0.1}{0.1+(x_1-0.1)^2+2(x_2-0.1)^2} - \frac{0.1}{0.14+20((x_1-0.45)^2+(x_2-0.55)^2)} \\ -\frac{0.1}{0.1+(x_1-0.3)^2+2(x_2-0.95)^2} - \frac{0.1}{0.15+40((x_1-0.55)^2+(x_2-0.45)^2)} \end{pmatrix}.$$

Figure 3.4 shows the result of Algorithm 3.1 with box radius $r = 2^{-9} \approx 0.1953 \cdot 10^{-2}$. *To obtain an approximation of the image of the Pareto critical set, we computed the image of* $X_{\mathcal{C}}$ *under* f. *In this case,* P_c *consists of three connected components, which*



(a) Variable space **(b)** Image space

Figure 3.4: **(a)** The result of Algorithm 3.1 for Example 3.1.6. **(b)** The image of the set $X_{\mathcal{C}}$ under f and a pointwise discretization of the image of f.

are all tightly approximated by the covering $\mathcal{C}$. *(This result coincides with the result from [PŽŽ17].)*

The following problem is Example 4 from [Lov11], where the Pareto critical set again possesses interesting topological properties.

Example 3.1.7. *Consider the problem* $\min_{x \in \mathbb{R}^2} f(x)$ *with*

$$f(x) = -\begin{pmatrix} -x_1^2 - x_2^2 - 4(\exp(-(x_1+2)^2 - x_2^2) + \exp(-(x_1-2)^2 - x_2^2)) \\ -(x_1-6)^2 - (x_2+0.5)^2 \end{pmatrix}.$$

Figure 3.5 shows the result of Algorithm 3.1 with box radius $r = 2^{-7} \approx 0.7812 \cdot 10^{-2}$. *Here,* P_c *consists of two connected components, one being a loop and one being a line. In the image space, we see that the loop gets mapped to a loop with two "cusps".*

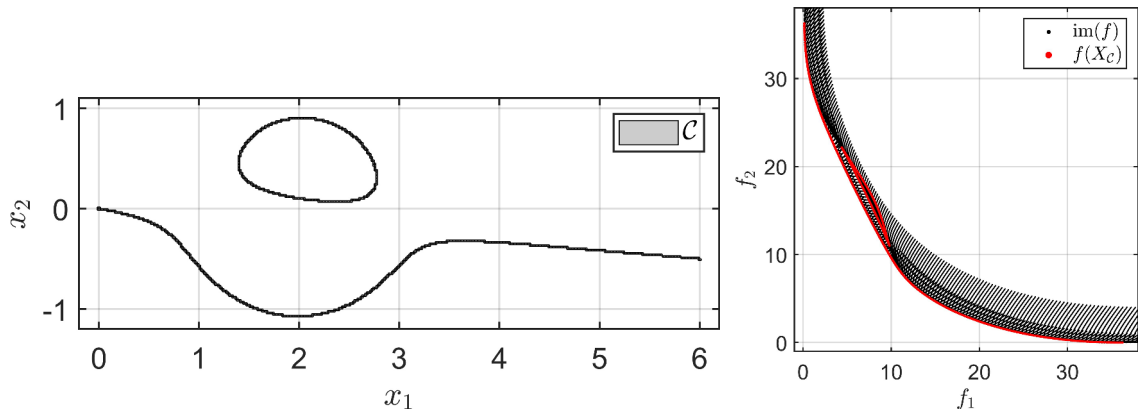
Finally, for a higher-dimensional example, we consider an unconstrained version of the problem $L\&H_{3\times 3}$ from [HL14].

Example 3.1.8. *Let* $\rho(\omega, \theta) \in \mathbb{R}^{3\times 3}$ *be the matrix corresponding to a rotation around the vector* $\omega \in \mathbb{R}^3$ *by an angle* $\theta \in [0, 2\pi)$. *Let*

$$b(x) := 0.075 g(x, (0, 0.15, 0)^\top, 0.3) + g(x, (0, -1.1, 0)^\top, 3),$$
$$g(x, p_0, \sigma) := \sqrt{\frac{2\pi}{\sigma}} \exp\left(-\frac{|x - p_0|^2}{\sigma^2}\right).$$

Consider the problem

$$\min_{x \in \mathbb{R}^3} f(x) \text{ with } f(x) = -\rho\left(\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, 0\right)^\top, \arctan(\sqrt{2})\right)(x_1, x_3, b(x))^\top.$$



(a) Variable space **(b)** Image space

Figure 3.5: **(a)** The result of Algorithm 3.1 for Example 3.1.7. **(b)** The image of the set $X_\mathcal{C}$ under f and a pointwise discretization of the image of f.

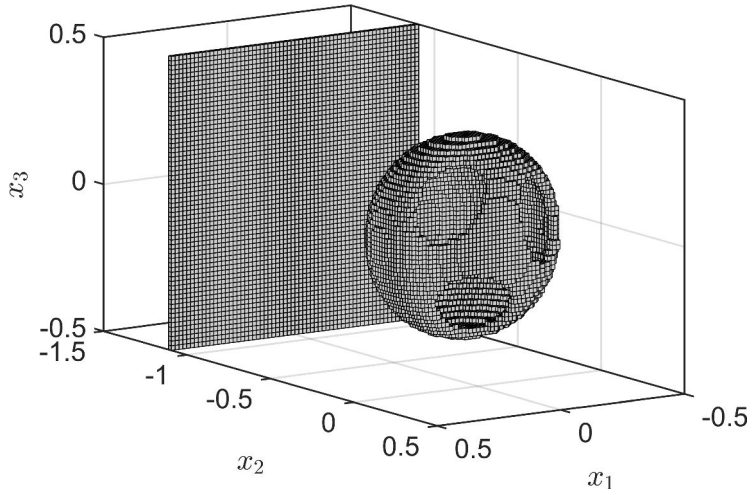


Figure 3.6: The result of Algorithm 3.1 for Example 3.1.8.

For details on the construction of this problem, see [HL14]. Figure 3.6 shows the result of Algorithm 3.1 with the box radius $r = 2^{-7}$. The Pareto critical set consists of two connected components, one being a two-dimensional surface and the other one being spherical with three holes.

3.1.4 Extension to constrained MOPs

In this section, we will discuss how Algorithm 3.1 can be generalized to MOPs with equality and inequality constraints. As we are more focused on the unconstrained case in this work, we will only discuss the basic ideas here. In particular, while the method will be able to compute Pareto critical sets of constrained MOPs, it will not necessarily be numerically efficient and should be seen as a starting point to develop more efficient methods.

For $f : \mathbb{R}^n \to \mathbb{R}^k$, $h : \mathbb{R}^n \to \mathbb{R}^{k_h}$ and $g : \mathbb{R}^n \to \mathbb{R}^{k_g}$, all twice continuously

differentiable, we now consider the MOP

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x), \\ \text{s.t.} \quad & h(x) = 0, \\ & g(x) \leq 0. \end{aligned} \tag{3.3}$$

For $x \in \mathbb{R}^n$ let

$$\mathcal{A}_g(x) := \{i \in \{1, \dots, k_g\} : g_i(x) = 0\} \tag{3.4}$$

be the *active set*. If $i \in \mathcal{A}_g(x)$, then we say the constraint g_i is *active* in x. We assume that for all feasible x, all elements in

$$\{\nabla h_i(x) : i \in \{1, \dots, k_h\}\} \cup \{\nabla g_i(x) : i \in \mathcal{A}_g(x)\}$$

are linearly independent. This condition is known as the *linear independence constraint qualification* (LICQ).

Since our continuation method is based on the KKT condition, we first have to generalize this condition to the constrained case. This is done in the following theorem (cf. [Hil01], Theorem 4.1).

Theorem 3.1.9. *Let x be a Pareto optimal point of* (3.3)*. Then there are $\alpha \in \Delta_k$, $\lambda \in \mathbb{R}^{k_h}$ and $\mu \in (\mathbb{R}^{\geq 0})^{k_g}$ such that*

$$\sum_{i=1}^{k} \alpha_i \nabla f_i(x) + \sum_{i=1}^{k_h} \lambda_i \nabla h_i(x) + \sum_{i\in\mathcal{A}_g(x)} \mu_i \nabla g_i(x) = 0, \tag{3.5}$$

$$\mu_i = 0 \quad \forall i \notin \mathcal{A}_g(x).$$

Based on the previous theorem, we can define the Pareto critical set for constrained MOPs.

Definition 3.1.10. *A feasible point $x \in \mathbb{R}^n$ is called* Pareto critical *if* (3.5) *holds in x. A corresponding $\alpha \in \Delta_k$ is called a* KKT vector of x*, containing* KKT multipliers *α_i, $i \in \{1, \dots, k\}$. The set P_c of all Pareto critical points is the* Pareto critical set*.*

In the unconstrained case, the continuation method was based on the manifold structure of the set $\mathcal{M}$ of Pareto critical points augmented by their corresponding KKT vectors (Theorem 2.2.10). If only equality constraints are present, then this result can be generalized in a straight-forward way. In the case of inequality constraints, the smoothness of the Pareto critical set is generally lost. The reason for this is the fact that a change in the active set causes a "discontinuous" change in the KKT condition. Visually speaking, this results in kinks in the Pareto critical set when the active set changes, i.e., when the solution hits the boundary of the feasible set. Nonetheless, we will show that it is possible to handle this nonsmoothness by transforming inequality into equality constraints in a certain way.

For the case where only equality constraints are present, consider the function

$$F : \mathbb{R}^n \times (\mathbb{R}^{\geq 0})^k \times \mathbb{R}^{k_h} \to \mathbb{R}^{n+k_h+1}, \quad (x, \alpha, \lambda) \mapsto \begin{pmatrix} Df(x)^\top \alpha + Dh(x)^\top \lambda \\ 1 - \sum_{i=1}^{k} \alpha_i \\ h(x) \end{pmatrix}$$

as a generalization of (2.11). The following result from [Hil01] generalizes Theorem 2.2.10.

Theorem 3.1.11. *Let*

$$\mathcal{M} := (F|_{\mathbb{R}^n \times (\mathbb{R}^{>0})^k \times \mathbb{R}^{k_h}})^{-1}(0). \tag{3.6}$$

If $\operatorname{rk}(DF(x,\alpha,\lambda)) = n+k_h+1$ *for all* $(x,\alpha,\lambda) \in \mathcal{M}$*, then* $\mathcal{M}$ *is a* $(k-1)$*-dimensional embedded submanifold of* $\mathbb{R}^{n+k+k_h}$ *with tangent space*

$$T_{(x,\alpha,\lambda)}\mathcal{M} = \ker(DF(x,\alpha,\lambda)).$$

Proof. Analogously to the proof of Theorem 2.2.10 a). □

Note that for all feasible points x with the same active set, the roles of the equality and inequality constraints in the KKT condition (3.5) are identical, except that the multipliers μ_i corresponding to the active indices must be non-negative. More formally, for $A \subseteq \{1,\dots,k_g\}$ let $g^A := (g_i)_{i\in A}$ and

$$P_c^A := \{x \in P_c : \mathcal{A}_g(x) = A\}.$$

Then the following lemma holds:

Lemma 3.1.12. *Let* $A \subseteq \{1,\dots,k_g\}$*. Then* P_c^A *is a subset of the Pareto critical set of the MOP*

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x), \\ s.t. \quad & h(x) = 0, \\ & g^A(x) = 0. \end{aligned} \tag{3.7}$$

Proof. Let $x \in P_c^A$. Then $h(x) = 0$ and $g^A(x) = 0$, i.e., x is a feasible point of (3.7). Furthermore, there are $\alpha \in \Delta_k$, $\lambda \in \mathbb{R}^{k_h}$ and $\mu \in (\mathbb{R}^{\geq 0})^{k_g}$ such that

$$\sum_{i=1}^{k} \alpha_i \nabla f_i(x) + \sum_{i=1}^{k_h} \lambda_i \nabla h_i(x) + \sum_{i\in A} \mu_i \nabla g_i(x) = 0.$$

In particular, x is Pareto critical for the problem (3.7) with the multipliers $\tilde{\alpha} = \alpha$ and $\tilde{\lambda} = (\lambda, (\mu_i)_{i\in A})$. □

Clearly,

$$P_c = \bigcup_{A\subseteq\{1,\dots,k_g\}} P_c^A.$$

The idea is now to construct a box-continuation method like Algorithm 3.1 based on Theorem 3.1.11 and Lemma 3.1.12 which is able to approximate P_c^A and then start that method for all possible $A \subseteq \{1,\dots,k_h\}$. (The potential practical inefficiency of this approach will be discussed later.) To this end, we need a way to compute the tangent directions of P_c^A (step 4 in Algorithm 3.1) and a way to check if a box B contains part of P_c^A (step 6 in Algorithm 3.1).

For $A \subseteq \{1, \dots, k_g\}$ let $\mathcal{M}^A$ be the manifold from Theorem 3.1.11 for the MOP (3.7) in Lemma 3.1.12. For the tangent directions, we consider the projection of the tangent space of $\mathcal{M}^A$, i.e., for

$$\begin{aligned} F^A &: \mathbb{R}^n \times (\mathbb{R}^{\geq 0})^k \times \mathbb{R}^{k_h} \times \mathbb{R}^{|A|} \to \mathbb{R}^{n+k_h+1}, \\ (x, \alpha, \lambda, \mu) &\mapsto \begin{pmatrix} Df(x)^\top \alpha + Dh(x)^\top \lambda + Dg^A(x)^\top \mu \\ 1 - \sum_{i=1}^k \alpha_i \\ h(x) \\ g^A(x) \end{pmatrix}, \end{aligned}$$

we consider $\mathrm{pr}_x(\ker(DF^A(x, \alpha, \lambda, \mu)))$. In this way, step 4 of Algorithm 3.1 can be generalized. In order to decide whether a box B has a nonempty intersection with P_c^A in step 6, we use the problem

$$\begin{aligned} \theta(B) := \min_{x \in B, \alpha \in \Delta_k, \lambda \in \mathbb{R}^{k_h}, \mu \in (\mathbb{R}^{\geq 0})^{|A|}} & \|Df(x)^\top \alpha + Dh(x)^\top \lambda + Dg^A(x)^\top \mu\|^2, \\ \text{s.t.} \quad & h(x) = 0, \\ & g^A(x) = 0, \\ & g_i(x) \leq 0 \quad \forall i \notin A. \end{aligned} \tag{3.8}$$

Then $B \cap P_c^A \neq \emptyset$ if and only if both $\theta(B) = 0$ and the inequality constraint in (3.8) is strict. (If $\theta(B) = 0$ and the inequality is not strict, then we still have at least $B \cap P_c \neq \emptyset$.) The resulting method is summarized in Algorithm 3.2.

Algorithm 3.2 Constrained box-continuation method

1: Initialize $\mathcal{C} = \emptyset$, $X_{\mathcal{C}} = \emptyset$.
2: **for** all subsets $A \subseteq \{1, \dots, k_g\}$ **do**
3: For $(x_0, \alpha_0, \lambda_0, \mu_0) \in \mathcal{M}^A$, initialize $\mathcal{C}^A = \{B(x_0)\}$, $X_{\mathcal{C}^A} = \{(x_0, \alpha_0, \lambda_0, \mu_0)\}$, $\mathcal{C}^A_{out} = \emptyset$ and a queue $Q = \{(x_0, \alpha_0, \lambda_0, \mu_0)\}$.
4: **while** $Q \neq \emptyset$ **do**
5: Remove the first element $(\bar{x}, \bar{\alpha}, \bar{\lambda}, \bar{\mu})$ from Q.
6: If $k \geq n+1$ then set $T_{\mathcal{B}} := N(\bar{x})$. Otherwise, compute the neighboring boxes with a nonempty intersection with the projected tangent space, i.e.,

$$T_{\mathcal{B}} := \{B \in N(\bar{x}) : B \cap (\bar{x} + \mathrm{pr}_x(\ker(DF^A(\bar{x}, \bar{\alpha}, \bar{\lambda}, \bar{\mu})))) \neq \emptyset\}.$$

7: **for** $B \in T_{\mathcal{B}} \setminus (\mathcal{C}^A \cup \mathcal{C}^A_{out})$ **do**
8: Compute the optimal value $\theta(B)$ and the solution $(x_B, \alpha_B, \lambda_B, \mu_B)$ of (3.8).
9: If $\theta(B) = 0$, then set $Q = Q \cup \{(x_B, \alpha_B, \lambda_B, \mu_B)\}$, $\mathcal{C}^A = \mathcal{C}^A \cup \{B\}$ and $X_{\mathcal{C}^A} = X_{\mathcal{C}^A} \cup \{(x_B, \alpha_B, \lambda_B, \mu_B)\}$. Otherwise, set $\mathcal{C}^A_{out} = \mathcal{C}^A_{out} \cup \{B\}$.
10: **end for**
11: **end while**
12: Set $\mathcal{C} = \mathcal{C} \cup \mathcal{C}^A$ and $X_{\mathcal{C}} = X_{\mathcal{C}} \cup X_{\mathcal{C}^A}$.
13: **end for**

In the following remark, some of the properties of Algorithm 3.2 will be discussed.

Remark 3.1.13. *a) For the initialization of each iteration of the outer loop in step 3 of Algorithm 3.2, a starting point* $(x_0, \alpha_0, \lambda_0, \mu_0) \in \mathcal{M}^A$ *is required. Since this has to be done for all* 2^{k_g} *possible choices of* $A \subseteq \{1, \dots, k_g\}$*, this part can be highly time consuming for MOPs with a large number of inequality constraints. In particular, there can be many* A *for which this step is unnecessary, since* $\{x \in \mathbb{R}^n : \mathcal{A}_g(x) = A\}$ *may be empty. (For example, due to the LICQ, this is the case when* $|A| > n - k_h$*.) To increase the efficiency, it would make sense to implement some mechanism that filters out all irrelevant active sets before the outer loop is started.*

b) In the unconstrained case, the condition $k \geq n + 1$ *in step 6 was added for when the projected tangent space is equal to the entire* $\mathbb{R}^n$*. In the constrained case, this condition could be sharpened by exploiting the fact that by the LICQ, the feasible set of the MOP* (3.7) *is a* $(n - k_h - |A|)$*-dimensional manifold (cf. Theorem 2.2.9). This implies that for* $k > n - k_h - |A|$*, the projected tangent space is equal to the tangent space of the feasible set of* (3.7)*.*

To conclude this section, we apply Algorithm 3.2 to an inequality constrained version of Example S4 from [Sch04]. The practical implementation of the algorithm is analogous to the implementation of Algorithm 3.1 in Section 3.1.3, except that we additionally use the *interior-point method* (cf. [Wal+05]) of `fmincon` as a second solver (as discussed in Remark 3.1.3 f)).

Example 3.1.14. *For*

$$f(x) := \begin{pmatrix} (x_1 - 1)^4 + (x_2 - 1)^2 + (x_3 - 1)^2 \\ (x_1 + 1)^2 + (x_2 + 1)^4 + (x_3 + 1)^2 \\ (x_1 - 1)^2 + (x_2 + 1)^2 + (x_3 - 1)^4 \end{pmatrix}$$

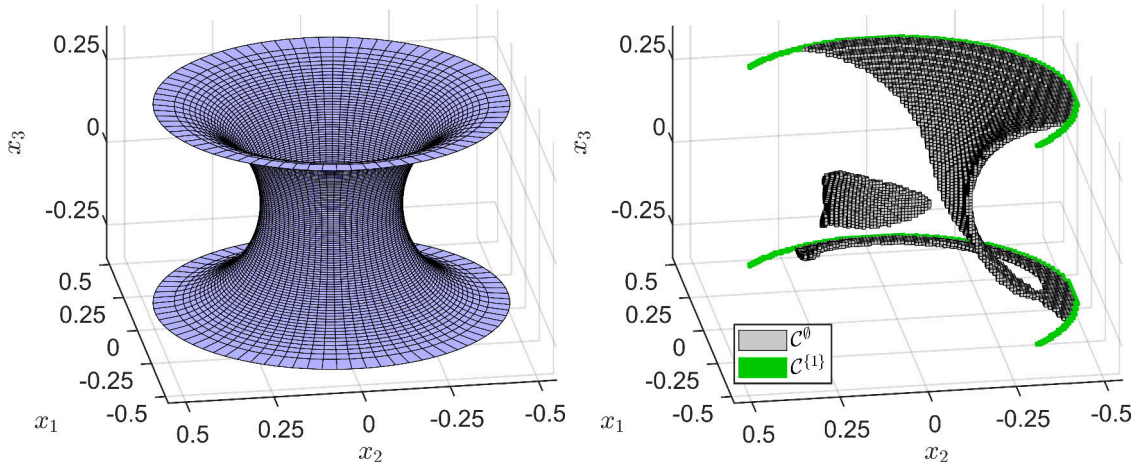
consider the problem

$$\begin{aligned} \min_{x \in \mathbb{R}^3} \quad & f(x), \\ s.t. \quad & (x_1^2 + x_2^2 + x_3^2 - (R^2 + r^2))^2 - 4R^2(r^2 - x_3^2) = 0, \\ & x_1^2 + x_2^2 - 0.5^2 \leq 0, \end{aligned}$$

for $r = 0.3$ *and* $R = 0.5$*. The set of points satisfying the equality constraint forms a torus with major radius* R *and minor radius* r*. The set of points satisfying the inequality constraint is the interior (and boundary) of a cylinder with radius* 0.5 *around the* x_3*-axis. The resulting feasible set is shown in Figure 3.7(a). The result of Algorithm 3.2 with a box radius* $r = 1.2 \cdot 2^{-8} = 0.46875 \cdot 10^{-2}$ *is shown in Figure 3.7(b). Since* $k_h = 1$*, the only choices for* A *in step 2 of Algorithm 3.2 are* $\emptyset$ *and* $\{1\}$*. The corresponding coverings* $\mathcal{C}^\emptyset$ *and* $C^{\{1\}}$ *are highlighted by different colors.*

3.1.5 Obtaining the Pareto set

In this work, we are mainly interested in computing the Pareto *critical* set of MOPs. But in the general context of multiobjective optimization, the actual Pareto set is of more interest. Thus, in this section, we will briefly discuss a simple way of obtaining the Pareto set from the Pareto critical set. For more advanced techniques, see [Gup+97; Sch03].



(a) **(b)**

Figure 3.7: **(a)** Feasible set of the MOP in Example 3.1.14. **(b)** Result of Algorithm 3.2, colored according to the different active sets.

In the generic single-objective case, if a global minimal point exists and all critical points (with $\nabla f(x) = 0$) are known, then all global minimal points can be easily obtained by filtering out the critical points with the smallest function values. In the multiobjective case, it works in a similar way by replacing the notion of "minimal" in $\mathbb{R}$ with the vector valued version from (2.1). The basis is the following lemma.

Lemma 3.1.15. *Let $X \subseteq \mathbb{R}^n$ be compact, $f : X \to \mathbb{R}^k$ be continuous and P be the Pareto set of $\min_{x \in X} f(x)$. Let $C \subseteq X$ with $P \subseteq C$. Let*

$$C_{nd} := \{x \in C : \nexists x' \in C \text{ with } f(x') \leq f(x),\ f(x') \neq f(x)\}.$$

Then $P = C_{nd}$.

Proof. The relation $P \subseteq C_{nd}$ is easy to see. For the relation $C_{nd} \subseteq P$, assume that there is some $x \in C_{nd}$ with $x \notin P$. Then there must be some $x' \in X \setminus C$ with $f(x') \leq f(x)$ and $f(x') \neq f(x)$. Let $Z = f^{-1}(\{y \in \mathbb{R}^k : y \leq f(x')\})$. Since f is continuous, Z is closed. Additionally, Z is compact as a closed subset of X. By [Ehr05], Theorem 2.19, this means that the Pareto set of $\min_{z \in Z} f(z)$ is nonempty. By the definition of Z, this implies that $\emptyset \neq Z \cap P \subseteq Z \cap C$. Let $\bar{x} \in Z \cap C$. Then $f(\bar{x}) \leq f(x') \leq f(x)$ and $f(\bar{x}) \leq f(x') \neq f(x)$, which is a contradiction to $x \in C_{nd}$. □

Since we mostly consider the unconstrained case (where $X = \mathbb{R}^n$ is not compact), we also need the following result.

Lemma 3.1.16. *Let $f : \mathbb{R}^n \to \mathbb{R}^k$ be continuous and P be the Pareto set of $\min_{x \in \mathbb{R}^n} f(x)$. Let $C \subseteq \mathbb{R}^n$ with $P \subseteq C$. If the sublevel sets $\{z \in \mathbb{R}^n : f(z) \leq f(x)\}$ are bounded for all $x \in C$, then $P = C_{nd}$.*

Proof. The proof is analogously to the proof of Lemma 3.1.15, except that the compactness of Z follows from

$$Z = \{z \in \mathbb{R}^n : f(z) \leq f(x')\} \subseteq \{z \in \mathbb{R}^n : f(z) \leq f(x)\}$$

and from the latter set being bounded by assumption. □

For $x \in \mathbb{R}^n$, the sublevel set in the previous lemma can be written as

$$\{z \in \mathbb{R}^n : f(z) \leq f(x)\} = \bigcap_{i=1}^{k} \{z \in \mathbb{R}^n : f_i(z) \leq f_i(x)\},$$

which means that it is bounded if one of the sublevel sets of the f_i, $i \in \{1, \ldots, k\}$, is bounded. By Lemma 3.1.15 and 3.1.16, if we have some (bounded) superset C of the Pareto set, then the values of f in C are sufficient to decide which points in C are Pareto optimal.

In practice, this means that if we have a finite set $C \subseteq \mathbb{R}^n$ including a pointwise discretization of P, then $C_{nd} \approx P$. A straight-forward way of computing C_{nd} is shown in Algorithm 3.3. In our case, the obvious choice for the input C of Algorithm 3.3 would be the output $X_{\mathcal{C}}$ from Algorithm 3.1 (or 3.2), containing a Pareto critical point from each box in $\mathcal{C}$.

Algorithm 3.3 Nondominance algorithm

Given: Finite set $C = \{x^1, \ldots, x^N\} \subseteq \mathbb{R}^n$.
1: Compute the image $y^i = f(x^i)$ for all $i \in \{1, \ldots, N\}$.
2: Initialize $I_{dom} = \emptyset$.
3: **for** $i \in \{1, \ldots, N\}$ **do**
4: **if** $i \notin I_{dom}$ **then**
5: **for** $j \in \{1, \ldots, N\} \setminus I_{dom}$ **do**
6: If $y^i \leq y^j$ and $y^i \neq y^j$, then $I_{dom} = I_{dom} \cup \{j\}$.
7: **end for**
8: **end if**
9: **end for**
10: Set $C_{nd} = \{x^i \in C : i \notin I_{dom}\}$.

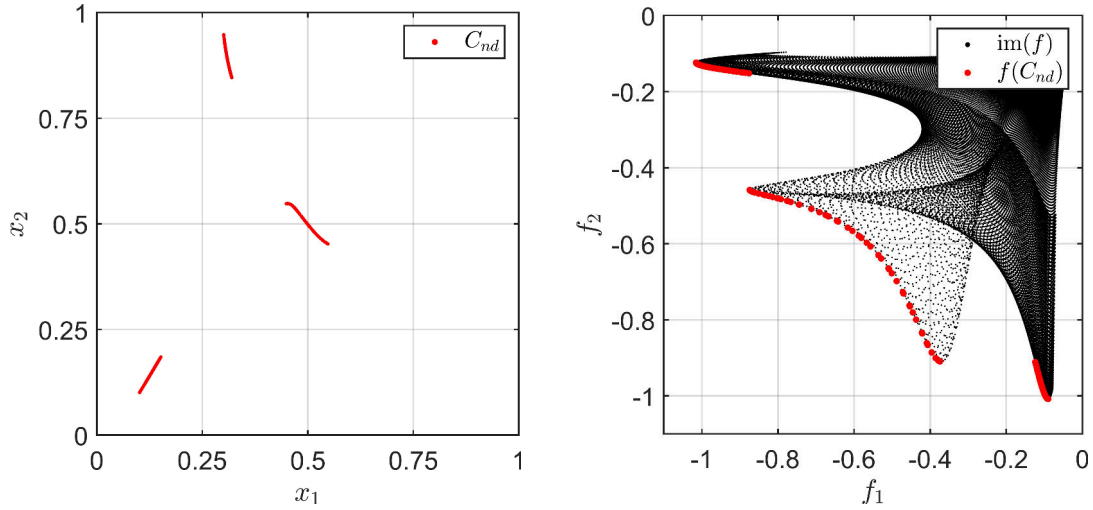
Remark 3.1.17. *Lemma 3.1.15, Lemma 3.1.16 and Algorithm 3.3 hold for (almost) arbitrary supersets of the Pareto set and are not restricted to Pareto critical points. Thus, it would also be possible to use a pointwise discretization of the box covering $\mathcal{C}$, by choosing multiple points in each box in $\mathcal{C}$. This may result in a finer approximation of the Pareto set, at the cost of (potentially) having non Pareto critical points in the approximation.*

To conclude this section, we apply Algorithm 3.3 to the MOP in Example 3.1.6.

Example 3.1.18. *Consider the MOP from Example 3.1.6. For the input of Algorithm 3.3, we choose the output $X_{\mathcal{C}}$ from Algorithm 3.1. Figure 3.8 shows the result of Algorithm 3.3 in variable and image space. Compared to Figure 3.4, we see that the algorithm correctly discarded Pareto critical points which are not Pareto optimal. (In particular, this result coincides with the result in [PŽŽ17].)*

3.2 Inexact gradients

In this section, we will investigate the influence of inexactness (or errors) in the gradients of the objective functions f_i, $i \in \{1, \ldots, k\}$, on the Pareto critical set.



(a) Variable space **(b)** Image space

Figure 3.8: The result of Algorithm 3.3 applied to the result of Algorithm 3.1 for Example 3.1.6.

More formally, we assume that we have an *inexact objective vector* $f^r : \mathbb{R}^n \to \mathbb{R}^k$, which is continuously differentiable, and upper bounds $\varepsilon \in (\mathbb{R}^{\geq 0})^k$ such that

$$\sup_{x \in \mathbb{R}^n} \|\nabla f_i(x) - \nabla f_i^r(x)\| \leq \varepsilon_i \quad \forall i \in \{1, \ldots, k\}. \tag{3.9}$$

Our goal is to approximate the Pareto critical set P_c of the original function f using only the inexact function f^r and the error bounds ε.

This scenario typically occurs when working with objective functions which are computationally expensive to evaluate, and are therefore replaced by cheaper surrogate models. As an example, in the context of parameter optimization of physical systems, every evaluation of f (and its gradients) might require the solution of some *partial differential equation* (PDE). In this case, *reduced-order modeling* (ROM) can be used to obtain a cheaper surrogate model as in (3.9). See [ASG01; Que+05] for an overview. We will later consider an explicit example from this problem class.

If the error bound ε is small, the naive approach for obtaining an approximation of P_c would be to just ignore the error bounds and compute the Pareto critical set P_c^r of the inexact function, i.e., the Pareto critical set of the MOP

$$\min_{x \in \mathbb{R}^n} f^r(x),$$

and hope that $P_c^r \approx P_c$. But as the following example shows, while P_c^r may be close to P_c in a Hausdorff sense, it might not actually be a good approximation in terms of its structure.

Example 3.2.1. *Consider the MOP*

$$\min_{x \in \mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} x_1^4 + \frac{9}{5}x_1^3 - \frac{2}{5}x_1^2 + \frac{1}{10}x_1 + x_2^2 \\ x_1^2 - \frac{19}{10}x_1 + (x_2 - 1)^2 + 1 \end{pmatrix}$$

and the inexact objective vector

$$f^r : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto f(x) - \frac{1}{10}\begin{pmatrix} x_1 \\ x_1 \end{pmatrix}.$$

It is easy to see that f and f^r satisfy condition (3.9) for $\varepsilon = (\frac{1}{10}, \frac{1}{10})^\top$. Figure 3.9 shows the Pareto critical set P_c of f and P_c^r of f^r (both computed via X_C from Algorithm 3.1). While the Hausdorff distance between P_c and P_c^r is relatively small

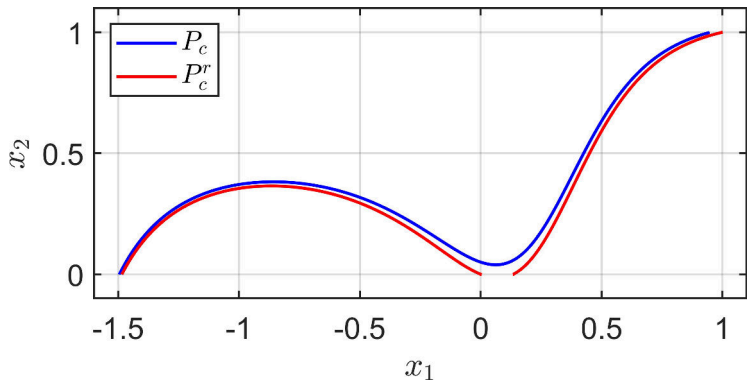


Figure 3.9: The sets P_c and P_c^r in Example 3.2.1.

($\approx 7.4 \cdot 10^{-2}$), there is an obvious difference in the structure of both sets: P_c is connected while P_c^r is disconnected.

Due to the previous example, a more sophisticated approach that explicitly takes the error bounds ε into account is needed. To this end, we begin by deriving a tight superset P_ε^r of P_c that only depends on f^r and ε (Section 3.2.1). Afterwards, to be able to efficiently compute it in practice, we analyze the structure of P_ε^r (Section 3.2.2). It will turn out that P_ε^r has the same dimension as the variable space with a boundary that seems to be piecewise smooth. We exploit this structure by constructing a box-continuation method which approximates the boundary ∂P_ε^r of P_ε^r (Section 3.2.3). Finally, we consider some simple examples (Section 3.2.4) before using our method for the solution a PDE-constrained MOP (Section 3.2.5).

3.2.1 A tight superset of the Pareto critical set and its structure

In this section, we will derive a set which only depends on f^r and ε (as in (3.9)) and is guaranteed to contain P_c. The basis for our approch is the following, simple result (which is similar to Lemma 3.2 in [PD17]).

Lemma 3.2.2. *Let $\bar{x} \in P_c$ with a KKT vector $\bar{\alpha} \in \Delta_k$. Then*

$$\|Df^r(\bar{x})^\top \bar{\alpha}\| \leq \varepsilon^\top \bar{\alpha} \leq \|\varepsilon\|_\infty.$$

Proof. By assumption we have $Df(\bar{x})^\top \bar{\alpha} = 0$. This implies

$$\begin{aligned}\|Df^r(\bar{x})^\top \bar{\alpha}\| = \|Df^r(\bar{x})^\top \bar{\alpha} - Df(\bar{x})^\top \bar{\alpha}\| &= \left\| \sum_{i=1}^k (\nabla f_i^r(\bar{x}) - \nabla f_i(\bar{x}))^\top \bar{\alpha}_i \right\| \\ &\leq \sum_{i=1}^k \|\nabla f_i^r(\bar{x}) - \nabla f_i(\bar{x})\| \bar{\alpha}_i \leq \sum_{i=1}^k \varepsilon_i \bar{\alpha}_i = \varepsilon^\top \bar{\alpha} \leq \|\varepsilon\|_\infty.\end{aligned}$$

□

By the previous lemma, we are able to obtain a superset of P_c by weakening the KKT condition of the inexact objective vector using the given error bounds. More precisely, we define

$$P_\varepsilon^r := \{x \in \mathbb{R}^n : \min_{\alpha \in \Delta_k} (\|Df^r(x)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2) \leq 0\}. \tag{3.10}$$

(This set is well-defined since Δ_k is compact and the objective is continuous, i.e., the minimum always exists.) It is easy to see that $P_c \subseteq P_\varepsilon^r$ and $P_c^r \subseteq P_\varepsilon^r$. Furthermore, if $\varepsilon = 0 \in \mathbb{R}^k$, then $P_\varepsilon^r = P_c = P_c^r$.

Remark 3.2.3. *By Lemma 3.2.2, instead of $\varepsilon^\top \alpha$, we could also take $\|\varepsilon\|_\infty$ as an upper bound for norm of the KKT condition of the inexact objective vector to obtain a coarser superset of P_c, as was done in [PD17]. More formally, if we define*

$$\tilde{P} := \{x \in \mathbb{R}^n : \min_{\alpha \in \Delta_k} \|Df^r(x)^\top \alpha\|^2 \leq \|\varepsilon\|_\infty^2\},$$

then $P_c \subseteq P_\varepsilon^r \subseteq \tilde{P}$ (and $P_c^r \subseteq P_\varepsilon^r \subseteq \tilde{P}$). Note that by (2.21), $\tilde{P}$ can be interpreted as the set of points in which the length of the steepest descent direction for f^r is less or equal to $\|\varepsilon\|_\infty$. This makes it relatively easy to compute $\tilde{P}$ via descent-based methods. For example, in [PD17], $\tilde{P}$ was approximated via the subdivision method (cf. Section 2.3). The disadvantage of working with $\tilde{P}$ is the fact that it is a less tight covering of P_c. In particular, all error bounds except the maximal bound $\|\varepsilon\|_\infty$ are ignored.

In the proof of Lemma 3.2.2, we only used the subadditivity of the norm and the inequality (3.9). As the following lemma shows, these estimations are sufficiently tight for P_ε^r to be a tight superset of P_c.

Lemma 3.2.4. *Let $\tilde{x} \in P_\varepsilon^r$. Then there is a continuously differentiable function $\tilde{f} : \mathbb{R}^n \to \mathbb{R}^k$ with*

$$\sup_{x \in \mathbb{R}^n} \|\nabla \tilde{f}_i(x) - \nabla f_i^r(x)\| \leq \varepsilon_i \quad \forall i \in \{1, \ldots, k\}$$

such that $\tilde{x}$ is Pareto critical for $\tilde{f}$.

Proof. Let

$$\tilde{\alpha} \in \arg\min_{\alpha \in \Delta_k} (\|Df^r(\tilde{x})^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2).$$

If $\varepsilon^\top \tilde{\alpha} = 0$, then we have $0 \geq \|Df^r(\tilde{x})^\top \tilde{\alpha}\|^2 - (\varepsilon^\top \tilde{\alpha})^2 = \|Df^r(\tilde{x})^\top \tilde{\alpha}\|^2$ since $\tilde{x} \in P_\varepsilon^r$, so $Df^r(\tilde{x})^\top \tilde{\alpha} = 0$ and the statement holds for $\tilde{f} = f^r$. If $\varepsilon^\top \tilde{\alpha} \neq 0$, we define

$$\begin{aligned} \nu &:= Df^r(\tilde{x})^\top \tilde{\alpha} \in \mathbb{R}^n, \\ g(x) &:= -\left(\frac{1}{\varepsilon^\top \tilde{\alpha}} \sum_{j=1}^n \nu_j x_j\right) \varepsilon \in \mathbb{R}^k, \\ \tilde{f}(x) &:= f^r(x) + g(x) \in \mathbb{R}^k. \end{aligned}$$

Since $\tilde{x} \in P_\varepsilon^r$ we have $\|Df^r(\tilde{x})^\top \tilde{\alpha}\|^2 - (\varepsilon^\top \tilde{\alpha})^2 \leq 0$, so $\|\nu\| \leq \varepsilon^\top \tilde{\alpha}$. Thus,

$$\|\nabla \tilde{f}_i(x) - \nabla f_i^r(x)\| = \|\nabla g_i(x)\| = \frac{\varepsilon_i}{\varepsilon^\top \tilde{\alpha}} \|\nu\| \leq \varepsilon_i \quad \forall x \in \mathbb{R}^n,\ i \in \{1, \ldots, k\}$$

and

$$D\tilde{f}(\tilde{x})^\top \tilde{\alpha} = Df^r(\tilde{x})^\top \tilde{\alpha} + Dg(\tilde{x})^\top \tilde{\alpha} = \nu + \sum_{i=1}^{k} \tilde{\alpha}_i \nabla g_i(\tilde{x})$$
$$= \nu - \sum_{i=1}^{k} \tilde{\alpha}_i \frac{\varepsilon_i}{\varepsilon^\top \tilde{\alpha}} \nu = \nu - \frac{\varepsilon^\top \tilde{\alpha}}{\varepsilon^\top \tilde{\alpha}} \nu = 0,$$

which completes the proof. □

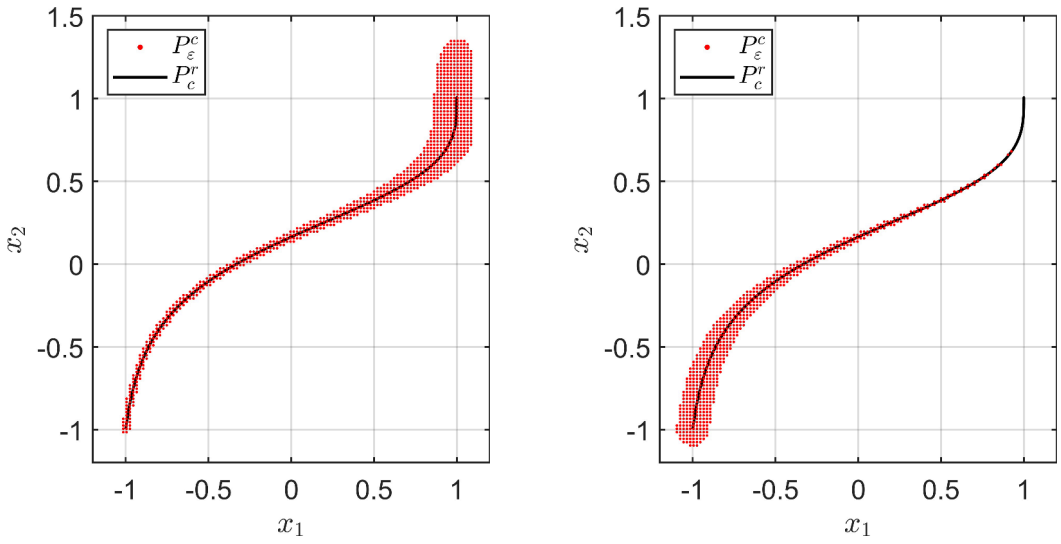
The previous lemma shows that for any element $\tilde{x}$ of our superset P_ε^r, there is some $\tilde{f}$ that satisfies our error bounds (3.9) (for $f = \tilde{f}$), such that $\tilde{x}$ is Pareto critical for $\tilde{f}$. In other words, any element of P_ε^r can potentially be Pareto critical for a function satisfying our error bounds. In this sense, using only the inexact function and the error bounds, P_ε^r is the tightest superset of P_c that we can hope for.

Before analyzing the structure of P_ε^r, we will consider a simple example (from [PD17]).

Example 3.2.5. *Consider the inexact function*

$$f^r : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} (x_1 - 1)^2 + (x_2 - 1)^4 \\ (x_1 + 1)^2 + (x_2 + 1)^2 \end{pmatrix}.$$

To obtain a rough approximation of P_ε^r, we solve the optimization problem in the definition (3.10) *of P_ε^r in different points in $\mathbb{R}^2$ and then check if the optimal value is non-positive. Figure 3.10 shows the result for different error bounds ε and the Pareto critical set P_c^r of f^r. By our previous considerations, the Pareto critical set of any objective vector f satisfying the given error bounds as in* (3.9) *must lie in the area indicated by the red dots.*



(a) $\varepsilon = (0.2, 0.05)^\top$ **(b)** $\varepsilon = (0, 0.2)^\top$

Figure 3.10: Pointwise approximation of P_ε^r for Example 3.2.5 for different error bounds.

By comparing (a) and (b) in Figure 3.10, it is easy to see the influence of ε on P_ε^r. In this example, the critical (and optimal) points of f_1^r and f_2^r are given by

$(1,1)^\top$ *and* $(-1,-1)^\top$*, respectively. The corresponding KKT vectors are* $(1,0)^\top$ *and* $(0,1)^\top$*. Since in* (3.10)*, the error bounds are weighted with the KKT vector, this means that* ε_2 *has almost no influence close to* $(1,1)^\top$ *and* ε_1 *has almost no influence close to* $(-1,-1)^\top$*. In (a), the error bound* ε_2 *for* f_2^r *is smaller than the error bound* ε_1 *for* f_1^r*. Thus,* P_ε^r *is thinner close to* $(-1,-1)^\top$ *and thicker close to* $(1,1)^\top$*. In (b) the situation is reversed. In particular, since* $\varepsilon_1 = 0$ *in (b),* P_ε^r *approaches* P_c^r *close to* $(1,1)$*.*

In the previous example, P_ε^r had the same "dimension" as the variable space, i.e., it was not a null set. The following lemma shows that if f^r is twice continuously differentiable, then this is what we can expect in the general case.

Lemma 3.2.6. *Let* f^r *be twice continuously differentiable.*

a) P_ε^r *is closed.*

b) Assume that $P_c^r \neq \emptyset$ *and that there exists some* $x^r \in P_c^r$ *with a KKT vector* α^r *such that* $\varepsilon^\top \alpha^r > 0$*. Let*

$$A := \{x \in \mathbb{R}^n : \min_{\alpha \in \Delta_k} (\|Df^r(x)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2) < 0\}.$$

Then $\emptyset \neq A \subseteq P_\varepsilon^r$ *and* A *is open in* $\mathbb{R}^n$*.*

Proof. a) The case $P_\varepsilon^r = \emptyset$ is trivial, so assume that $P_\varepsilon^r \neq \emptyset$. Let $\bar{x} \in \overline{P_\varepsilon^r}$. Then there is a sequence $(x^i)_i \in P_\varepsilon^r$ with $\lim_{i\to\infty} x^i = \bar{x}$. Let $(\alpha^i)_i \in \Delta_k$ be a sequence with

$$\alpha^i \in \arg\min_{\alpha \in \Delta_k} (\|Df^r(x^i)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2).$$

By compactness of Δ_k, we can assume w.l.o.g. that there is some $\bar{\alpha} \in \Delta_k$ with $\lim_{i\to\infty} \alpha^i = \bar{\alpha}$. Let

$$\Psi : \mathbb{R}^n \times \Delta_k \to \mathbb{R}, \quad (x,\alpha) \mapsto \|Df^r(x)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2.$$

By our assumption, Ψ is continuous and $\Psi(x^i, \alpha^i) < 0$ for all $i \in \mathbb{N}$. Thus, it holds $\Psi(\bar{x}, \bar{\alpha}) \leq 0$, which yields $\bar{x} \in P_\varepsilon^r$.
b) The relation $A \subseteq P_\varepsilon^r$ holds trivially. By assumption, we have

$$\min_{\alpha \in \Delta_k} (\|Df^r(x^r)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2) \leq \|Df^r(x^r)^\top \alpha^r\|^2 - (\varepsilon^\top \alpha^r)^2 = -(\varepsilon^\top \alpha^r)^2 < 0,$$

so $x^r \in A$ and $A \neq \emptyset$. To show that A is open, let $\bar{x} \in A$ and

$$\bar{\alpha} \in \arg\min_{\alpha \in \Delta_k} (\|Df^r(\bar{x})^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2).$$

Let $\Phi : \mathbb{R}^n \to \mathbb{R}$, $x \mapsto \|Df^r(x)^\top \bar{\alpha}\|^2 - (\varepsilon^\top \bar{\alpha})^2$. Then Φ is continuous and $\Phi(\bar{x}) < 0$. Thus, there must be some open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that $\Phi(y) < 0$ for all $y \in U$. Since

$$\min_{\alpha \in \Delta_k} (\|Df^r(y)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2) \leq \Phi(y) < 0 \quad \forall y \in U$$

we have $U \subseteq A$, implying that A is open. □

The assumptions in b) of the previous lemma are relatively weak. For example, they are satisfied if $P_c^r \neq \emptyset$ and

- $\varepsilon_i > 0$ for all $i \in \{1, \dots, k\}$ or
- $\varepsilon \neq 0 \in \mathbb{R}^k$ and there is some KKT vector α^r with $\alpha_i^r > 0$ for all $i \in \{1, \dots, k\}$.

This leads us to the conclusion that the superset P_ε^r, in contrast to the Pareto critical set P_c^r, will generally not be a null set in $\mathbb{R}^n$, as it contains an open subset of $\mathbb{R}^n$. The implications of this will be discussed in the following section.

3.2.2 Efficient computation

Assume from now on that f^r is twice continuously differentiable. As shown in the previous section, the superset P_ε^r has generally the same "dimension" as the variable space $\mathbb{R}^n$. So in contrast to Pareto critical sets, which are typically lower-dimensional objects in $\mathbb{R}^n$ (for $k \leq n$), P_ε^r is always relatively large. In particular, for problems with a large number of variables, the computation of P_ε^r suffers from the curse of dimensionality. But recall that only using the information in (3.9), Lemma 3.2.4 showed that P_ε^r is the tightest superset of P_c that we can hope for. So the high dimension of P_ε^r is a natural result from the available information.

In this section, we will derive an approach that makes it possible to compute P_ε^r in a (relatively) efficient manner. The idea is to only compute the topological boundary ∂P_ε^r of P_ε^r instead of the entire set. After ∂P_ε^r is computed, P_ε^r can be obtained by checking the condition in (3.10) for one point from each connected component of $\mathbb{R}^n \setminus \partial P_\varepsilon^r$, since each of these components lies either completely inside or outside P_ε^r.

We will begin by describing ∂P_ε^r as a level set. Let

$$\varphi : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \min_{\alpha \in \Delta_k} (\|Df^r(x)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2).$$

By Lemma 3.2.6, $\varphi^{-1}(\mathbb{R}^{\leq 0}) = P_\varepsilon^r$ is closed and $\varphi^{-1}(\mathbb{R}^{<0})$ is a subset of the interior $(P_\varepsilon^r)^\circ$ of P_ε^r, so

$$\partial P_\varepsilon^r = \overline{P_\varepsilon^r} \setminus (P_\varepsilon^r)^\circ \subseteq \varphi^{-1}(\mathbb{R}^{\leq 0}) \setminus \varphi^{-1}(\mathbb{R}^{<0}) = \varphi^{-1}(0).$$

In other words, ∂P_ε^r is contained in the level set of φ corresponding to the value 0. By the level set theorem (Theorem 2.2.10), if we were able to show that φ is differentiable and $D\varphi$ has full rank (i.e., $D\varphi \neq 0 \in \mathbb{R}^{1 \times n}$), then $\varphi^{-1}(0)$ would be an $(n-1)$-dimensional manifold and we could compute it via continuation. Unfortunately, this does not hold in general. To see this, consider the optimization problem in the definition of φ:

$$\begin{aligned} \min_{\alpha \in \mathbb{R}^k} \quad & \omega(\alpha), \\ \text{s.t.} \quad & \sum_{i=1}^k \alpha_i = 1, \\ & \alpha_i \geq 0 \quad \forall i \in \{1, \dots, k\}, \end{aligned} \tag{3.11}$$

with

$$\omega(\alpha) := \|Df^r(x)^\top \alpha\|^2 - (\alpha^\top \varepsilon)^2 = \alpha^\top (Df^r(x)Df^r(x)^\top - \varepsilon\varepsilon^\top)\alpha.$$

Due to the inequality constraints $\alpha_i \geq 0$ for all $i \in \{1, \ldots, k\}$, the optimal value $\varphi(x)$ of (3.11) is generally nonsmooth in points $x \in \mathbb{R}^n$ where the active set (cf. (3.4)) changes. This will be visualized in the following example:

Example 3.2.7. *Consider the function*

$$f^r : \mathbb{R}^2 \to \mathbb{R}^3, \quad x \mapsto \begin{pmatrix} -6x_1^2 + x_1^4 + 3x_2^2 \\ (x_1 - \frac{1}{2})^2 + 2(x_2 - 1)^2 \\ (x_1 - 1)^2 + 2(x_2 - \frac{1}{2})^2 \end{pmatrix}$$

from Example 4.1.5 in [Pei17]. Let $\varepsilon = (1, 1, 1)^\top$. Figure 3.11 shows an approximation of P_ε^r in $[0.25, 0.3] \times [0.65, 0.7]$, computed as in Example 3.2.5. It suggests that there is a kink in the boundary of P_ε^r. Figure 3.12 shows the graph of φ and the first

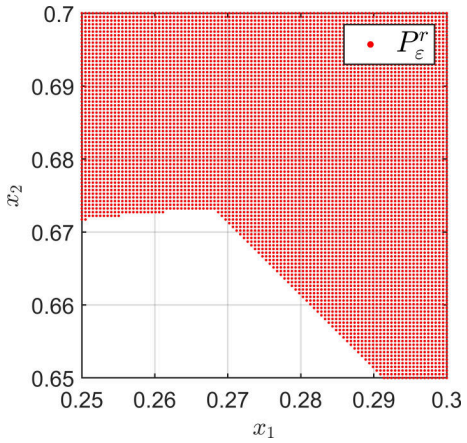


Figure 3.11: Pointwise approximation of P_ε^r for Example 3.2.7.

component α_1 of the solution of problem (3.11) *in the same area. We see that φ has*

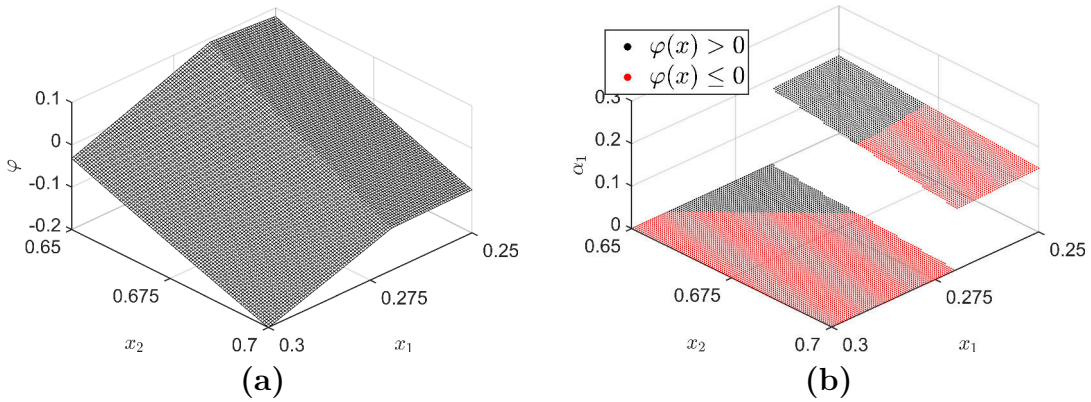


Figure 3.12: **(a)** Graph of φ in Example 3.2.7. **(b)** Second component α_2 of the solution of problem (3.11).

a "one-dimensional" nonsmoothness which intersects ∂P_ε^r in the kink from Figure 3.11. Considering the solution α of problem (3.11)*, we see that the nonsmoothness of φ lines up with the (discontinuous) activation of α_1.*

There are multiple reasons why the analysis of φ is challenging:

- The function φ is defined as the optimal value of a parameter-dependent optimization problem.
- Although the problem (3.11) is quadratic, the points in $\varphi^{-1}(0)$ are points where the matrix in the objective function ω is (at best) semidefinite.
- The solution α of (3.11) does not continuously depend on x.

Due to these difficulties, we will not present a thorough analysis of φ here. Instead, in the following, we will show some basic results about the properties of φ and afterwards give a rough sketch of how stronger results could be achieved.

Lemma 3.2.8. *The function φ is continuous.*

Proof. Let $(x^i)_i \in \mathbb{R}^n$ be a sequence with $\lim_{i\to\infty} x^i = \bar{x}$ for some $\bar{x} \in \mathbb{R}^n$. Note that

$$\varphi(x) \geq -\left(\sum_{i=1}^{k} \varepsilon_i\right)^2 \quad \forall x \in \mathbb{R}^n$$

and

$$\begin{aligned}\varphi(x) &\leq \|Df^r(x)^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2 \leq \|Df^r(x)^\top \alpha\|^2 \\ &\leq \left(\sum_{i=1}^{k} \alpha_i \|\nabla f_i^r(x)\|\right)^2 \leq \left(\sum_{i=1}^{k} \|\nabla f_i^r(x)\|\right)^2 \quad \forall x \in \mathbb{R}^n, \alpha \in \Delta_k.\end{aligned}$$

Since all ∇f_i^r, $i \in \{1, \dots, k\}$, are continuous, this implies that $(\varphi(x^i))_i$ is contained in a compact set and thus possesses an accumulation point. We will show that all accumulation points of $(\varphi(x^i))_i$ are equal to $\varphi(\bar{x})$, which will complete the proof.

Let $y \in \mathbb{R}$ be an accumulation point of $(\varphi(x^i))_i$. Assume w.l.o.g. that $\lim_{i\to\infty} \varphi(x^i) = y$. Let $(\alpha^i)_i \in \Delta_k$ such that α^i is the solution of (3.11) in x^i. Since Δ_k is compact, we can assume w.l.o.g. that $\lim_{i\to\infty} \alpha^i = \bar{\alpha}$ for some $\bar{\alpha} \in \Delta_k$. Since Df^r is continuous by assumption, we have

$$y = \lim_{i\to\infty} \varphi(x^i) = \lim_{i\to\infty} \|Df^r(x^i)^\top \alpha^i\|^2 - (\varepsilon^\top \alpha^i)^2 = \|Df^r(\bar{x})^\top \bar{\alpha}\|^2 - (\varepsilon^\top \bar{\alpha})^2. \quad (3.12)$$

Assume that $y \neq \varphi(\bar{x})$, i.e.,

$$\bar{\alpha} \notin \operatorname*{arg\,min}_{\alpha \in \Delta_k} \|Df^r(\bar{x})^\top \alpha\|^2 - (\varepsilon^\top \alpha)^2.$$

Then there must be some $\tilde{\alpha} \in \Delta_k$ with

$$\|Df^r(\bar{x})^\top \tilde{\alpha}\|^2 - (\varepsilon^\top \tilde{\alpha})^2 < \|Df^r(\bar{x})^\top \bar{\alpha}\|^2 - (\varepsilon^\top \bar{\alpha})^2. \quad (3.13)$$

Let

$$\Phi : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \|Df^r(x)^\top \tilde{\alpha}\|^2 - (\varepsilon^\top \tilde{\alpha})^2.$$

Then Φ is continuous due to continuity of Df^r. By (3.13), there must be some open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ and some $\delta > 0$ such that

$$\Phi(x) < \|Df^r(\bar{x})^\top \bar{\alpha}\|^2 - (\varepsilon^\top \bar{\alpha})^2 - \delta \quad \forall x \in U.$$

In particular, since $\lim_{i\to\infty} x^i = \bar{x}$, there must be some $N \in \mathbb{N}$ such that

$$\|Df^r(x^i)^\top \alpha^i\|^2 - (\varepsilon^\top \alpha^i)^2 \leq \Phi(x^i) < \|Df^r(\bar{x})^\top \bar{\alpha}\|^2 - (\varepsilon^\top \bar{\alpha})^2 - \delta \quad \forall i > N.$$

This is a contradiction to (3.12). □

Example 3.2.7 suggests that the nonsmoothness of φ is related to discontinuous "jumps" in the solution of (3.11). Since φ is still continuous, these jumps correspond to points $x \in \mathbb{R}^n$ in which the solution of (3.11) is non-unique. As we are mainly interested in the boundary $\partial P_\varepsilon^r \subseteq \varphi^{-1}(0)$, the following lemma investigates the uniqueness of solutions of (3.11) for $x \in \varphi^{-1}(0)$.

Lemma 3.2.9. *Let $x \in \varphi^{-1}(0)$ and let $\alpha^1, \alpha^2 \in \Delta_k$ be two solutions of (3.11) with $\alpha^1 \neq \alpha^2$. Then*

a) $\omega(\alpha) = 0 \quad \forall \alpha \in \operatorname{span}(\{\alpha^1, \alpha^2\})$,

b) there is some $\tilde{\alpha} \in \Delta_k$ with $\omega(\tilde{\alpha}) = 0$ and $\tilde{\alpha}_j = 0$ for some $j \in \{1, \ldots, k\}$,

c) there is some $\bar{\alpha} \in \operatorname{span}(\{\alpha^1, \alpha^2\})$ with $Df(x)^\top \bar{\alpha} = 0$.

Proof. a) For all $c_1, c_2 \in \mathbb{R}$ we have

$$\begin{aligned}
&\omega(c_1\alpha^1 + c_2\alpha^2) \\
&= (c_1\alpha^1 + c_2\alpha^2)^\top (Df^r(x)Df^r(x)^\top - \varepsilon\varepsilon^\top)(c_1\alpha^1 + c_2\alpha^2) \\
&= c_1^2\omega(\alpha^1) + 2(c_1(\alpha^1)^\top Df^r(x)Df^r(x)^\top \alpha^2 c_2 - c_1(\alpha^1)^\top \varepsilon\varepsilon^\top \alpha^2 c_2) + c_2^2\omega(\alpha^2) \quad (3.14) \\
&= 2(c_1(\alpha^1)^\top Df^r(x)Df^r(x)^\top \alpha^2 c_2 - c_1(\alpha^1)^\top \varepsilon\varepsilon^\top \alpha^2 c_2) \\
&= 2c_1c_2((Df^r(x)^\top \alpha^1)^\top (Df^r(x)^\top \alpha^2) - (\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2)).
\end{aligned}$$

Note that $\omega(\alpha^1) = \omega(\alpha^2) = 0$ implies $\|Df^r(x)^\top \alpha^i\| = \varepsilon^\top \alpha^i$, $i \in \{1, 2\}$. For $c_1, c_2 \geq 0$ we can apply the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
\omega(c_1\alpha^1 + c_2\alpha^2) &\leq 2c_1c_2(\|Df^r(x)^\top \alpha^1\|\|Df^r(x)^\top \alpha^2\| - (\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2)) \\
&= 2c_1c_2((\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2) - (\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2)) \\
&= 0.
\end{aligned}$$

Furthermore we must have $\omega(c_1\alpha^1 + c_2\alpha^2) \geq 0$ for all $c_1, c_2 \geq 0$ with $c_1 + c_2 = 1$, since $\varphi(x) = 0$ and $c_1\alpha^1 + c_2\alpha^2 \in \Delta_k$. Via scaling of $c_1\alpha^1 + c_2\alpha^2$, this implies

$$\omega(c_1\alpha^1 + c_2\alpha^2) = 0 \quad \forall c_1, c_2 \geq 0. \tag{3.15}$$

Combined with (3.14) we obtain

$$(Df^r(x)^\top \alpha^1)^\top (Df^r(x)^\top \alpha^2) - (\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2) = 0, \tag{3.16}$$

so (3.15) holds for all $c_1, c_2 \in \mathbb{R}$.
b) Consider the map $\psi(\lambda) := \lambda\alpha^1 + (1-\lambda)\alpha^2$. Due to a), we have $\omega(\psi(\lambda)) = 0$ for all $\lambda \in \mathbb{R}$. Since $\alpha^1, \alpha^2 \in \Delta_k$ with $\alpha^1 \neq \alpha^2$ and Δ_k is convex and bounded, there must be some $\tilde{\lambda}$ such that $\tilde{\alpha} := \psi(\tilde{\lambda}) \in \Delta_k$ and $\tilde{\alpha}_j = 0$ for some $j \in \{1, \ldots, k\}$.
c) By (3.16) we have

$$\begin{aligned}
(\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2) &= (Df^r(x)^\top \alpha^1)^\top (Df^r(x)^\top \alpha^2) \leq \|Df^r(x)^\top \alpha^1\|\|Df^r(x)^\top \alpha^2\| \\
&= (\varepsilon^\top \alpha^1)(\varepsilon^\top \alpha^2),
\end{aligned}$$

so the Cauchy-Schwarz inequality holds with equality. This is equivalent to $Df^r(x)^\top \alpha^1$ and $Df^r(x)^\top \alpha^2$ being linearly dependent. Thus, there are $c_1, c_2 \in \mathbb{R}\backslash\{0\}$ such that

$$Df^r(x)^\top \bar{\alpha} = 0$$

for $\bar{\alpha} = c_1\alpha^1 + c_2\alpha^2$. Since $\alpha^1 \neq \alpha^2$ by assumption and all elements of Δ_k are pairwise linearly independent, we have $\bar{\alpha} \neq 0$, completing the proof. □

By the previous lemma, for $x \in \varphi^{-1}(0)$, the solution of (3.11) can only be non-unique in points where the kernel of $Df(x)^\top$ has a non-trivial intersection with $\operatorname{span}(\{\alpha^1, \alpha^2\})$ for any two solutions α^1, α^2. If we assume $Df(x)^\top \in \mathbb{R}^{n\times k}$ to be a generic matrix, then we expect $\dim(\ker(Df(x)^\top)) = \max(\{0, k-n\})$. So in a generic setting, if (3.11) has a non-unique solution, then we must have

$$\begin{aligned} & \dim(\operatorname{span}(\{\alpha^1, \alpha^2\})) + \dim(\ker(Df(x)^\top)) > k \\ \Leftrightarrow \quad & 2 + \max(\{0, k-n\}) > k \\ \Leftrightarrow \quad & \begin{cases} k < 2, & \text{if } k \leq n, \\ n < 2, & \text{if } k > n. \end{cases} \end{aligned}$$

Thus, we expect that the set of all $x \in \varphi^{-1}(0)$ with (potentially) non-unique solutions of (3.11) is small compared to $\varphi^{-1}(0)$ itself. Combined with our earlier considerations, this would imply that the set of nonsmooth points of φ in $\varphi^{-1}(0)$ is small as well.

Recall that our goal is to use numerical continuation to compute $\varphi^{-1}(0)$ as a superset of ∂P^r_ε. Since φ is nonsmooth, the best we can hope for is that $\varphi^{-1}(0)$ is at least piecewise smooth, where "kinks" arise if φ is nonsmooth or if $D\varphi = 0$. While we will not actually prove that this is true, the following remark gives a rough sketch of how a proof might be possible.

Remark 3.2.10. *The first-order necessary optimality conditions of* (3.11) *are*

$$\begin{aligned} (Df^r(x)Df^r(x)^\top - \varepsilon\varepsilon^\top)\alpha + \begin{pmatrix} \lambda - \mu_1 \\ \vdots \\ \lambda - \mu_k \end{pmatrix} &= 0, \\ \sum_{i=1}^{k} \alpha_i - 1 &= 0, \\ \alpha_i \geq 0 \quad \forall i \in \{1, \dots, k\}, & \\ \mu_i \geq 0 \quad \forall i \in \{1, \dots, k\}, & \\ \mu_i\alpha_i = 0 \quad \forall i \in \{1, \dots, k\}, & \end{aligned} \tag{3.17}$$

for $\lambda \in \mathbb{R}^k$ and $\mu \in \mathbb{R}^k$. Let $\bar{x} \in \varphi^{-1}(0)$ with multipliers $\bar{\alpha}$, $\bar{\lambda}$ and $\bar{\mu}$ in (3.17). *Assume that there is an open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that that the solution of* (3.11) *on U is unique. Furthermore, assume that the activity of the inequality constraints of* (3.11) *is the same for any solution on U. Note that if $\alpha_j = 0$ for some $j \in \{1, \dots, k\}$ (i.e., the j-th inequality constraint is active), then f^r_j has no*

influence on (3.11) *and can be ignored. W.l.o.g. we assume that there are no active indices, i.e.,* $\alpha_i > 0$ *for all* $i \in \{1, \dots, k\}$*. Then on* U*,* (3.17) *is equivalent to*

$$\begin{aligned}(Df^r(x)Df^r(x)^\top - \varepsilon\varepsilon^\top)\alpha + \begin{pmatrix}\lambda\\ \vdots \\ \lambda\end{pmatrix} &= 0,\\ \sum_{i=1}^{k} \alpha_i - 1 &= 0,\\ \alpha_i \geq 0 \quad \forall i \in \{1, \dots, k\}.&\end{aligned} \tag{3.18}$$

If there is some open set $V \subseteq \mathbb{R}^{k+1}$ *with* $(\bar{\alpha}, \bar{\lambda}) \in V$ *such that* α *and* λ *in* (3.18) *are unique in* V*, then this system can be rewritten as* $G(x, (\alpha, \lambda)) = 0$ *for*

$$G : U \times V \to \mathbb{R}^{k+1}, \quad (x, (\alpha, \lambda)) \mapsto \begin{pmatrix}(Df^r(x)Df^r(x)^\top - \varepsilon\varepsilon^\top)\alpha + (\lambda, \dots, \lambda)^\top \\ \sum_{i=1}^{k} \alpha_i - 1\end{pmatrix}.$$

If the Jacobian

$$D_{(\alpha,\lambda)}G(x, (\alpha, \lambda)) = \begin{pmatrix}Df^r(\bar{x})Df^r(\bar{x})^\top - \varepsilon\varepsilon^\top & 1\\ 1 & 0\end{pmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}$$

of G *with respect to* (α, λ) *is invertible in* $(\bar{x}, (\bar{\alpha}, \bar{\lambda}))$*, then we could apply the implicit function theorem to obtain open sets* $U' \subseteq U$ *and* $V' \subseteq \mathbb{R}^{k+1}$ *with* $\bar{x} \in U'$ *and* $(\bar{\alpha}, \bar{\lambda}) \in V'$*, and a continuously differentiable function* $\phi = (\phi_\alpha, \phi_\lambda) : U' \to V'$ *with*

$$G(x, (\alpha, \lambda)) = 0 \;\Leftrightarrow\; (\alpha, \lambda) = \phi(x) \quad \forall x \in U', (\alpha, \lambda) \in V'.$$

Finally, on U'*,* φ *could be written as* $\varphi(x) = \omega(\phi_\alpha(x))$*, which would imply differentiability of* φ*. (In particular,* $D\varphi$ *could be computed via the chain rule and implicit differentiation.)*

In the following section, we will construct a box-continuation method which is able to approximate $\varphi^{-1}(0)$ similarly to the exact box-continuation method in Section 3.1. Due to the discussion in this section, we will assume that $\varphi^{-1}(0)$ is piecewise smooth. Since we did not actually prove this, we will also discuss how the method can be modified to be able to deal with problems where $\varphi^{-1}(0)$ does not have this structure.

3.2.3 The algorithm

The ideas of the box-continuation algorithm (Algorithm 3.1) can generally be used to compute any smooth manifold $M \subseteq \mathbb{R}^n$. The only ingredients needed are

- a way to compute the tangent space T_xM of M for any $x \in M$ and
- a way to check if $B \cap M \neq \emptyset$ for any box $B \in \mathcal{B}$.

If M can be written as a (zero) level set $M = h^{-1}(0)$ as in Theorem 2.2.9, then the tangent space can be computed via $T_xM = \ker(Dh(x))$. To check if $B \cap M \neq \emptyset$, we can minimize the sum of squares of the components of h in B, i.e., $\min_{x\in B} \|h(x)\|^2$.

Thus, in case of the boundary $\partial P_\varepsilon^r \subseteq \varphi^{-1}(0)$ of P_ε^r, we use $\ker(D\varphi(x))$ to compute the tangent space. Since we might have $D\varphi(x) = 0$, in which case $\varphi^{-1}(0)$ might not be a manifold around x, we consider $\mathbb{R}^n$ as the tangent space if the norm of $D\varphi(x)$ is small. To check if a box B contains part of $\varphi^{-1}(0)$, we solve the problem

$$\theta(B) := \min_{x \in B} \varphi(x)^2. \tag{3.19}$$

(Note that compared to (3.1) and (3.8), this problem is more difficult to solve since the problem (3.11) has to be solved for every evaluation of φ.) The resulting algorithm is Algorithm 3.4. In the following remark, we will discuss some of its properties.

Algorithm 3.4 Box-continuation method for inexact gradients

Given: $x_0 \in \varphi^{-1}(0)$.

1: Initialize $\mathcal{C} = \{B(x_0)\}$, $X_\mathcal{C} = \{x_0\}$, $\mathcal{C}_{out} = \emptyset$ and a queue $Q = \{x_0\}$.
2: **while** $Q \neq \emptyset$ **do**
3: Remove the first element $\bar{x}$ from Q.
4: Compute $D\varphi(\bar{x})$.
5: If $\|D\varphi(\bar{x})\|$ is small then set $T_\mathcal{B} := N(\bar{x})$. Otherwise, compute the neighboring boxes with a nonempty intersection with the tangent space, i.e.,

$$T_\mathcal{B} := \{B \in N(\bar{x}) : B \cap (\bar{x} + \ker(D\varphi(\bar{x}))) \neq \emptyset\}.$$

6: **for** $B \in T_\mathcal{B} \setminus (\mathcal{C} \cup \mathcal{C}_{out})$ **do**
7: Compute the optimal value $\theta(B)$ and the solution x_B of (3.19).
8: If $\theta(B) = 0$, then set $Q = Q \cup \{x_B\}$, $\mathcal{C} = \mathcal{C} \cup \{B\}$ and $X_\mathcal{C} = X_\mathcal{C} \cup \{x_B\}$. Otherwise, set $\mathcal{C}_{out} = \mathcal{C}_{out} \cup \{B\}$.
9: **end for**
10: **end while**

Remark 3.2.11. *a) As discussed earlier, we can only expect $\varphi^{-1}(0)$ to be piecewise smooth. Since we cannot expect the method to overcome "kinks" in $\varphi^{-1}(0)$, Algorithm 3.4 potentially has to be started multiples times to obtain the complete set, even if it is connected. This can be done as discussed in Remark 3.1.3 a).*

b) If φ is nonsmooth in $B \in \mathcal{B}$, then a nonsmooth solver is required to reliably solve (3.19)*. Otherwise, the algorithm may stop prematurely if a box is falsely identified to not contain part of $\varphi^{-1}(0)$.*

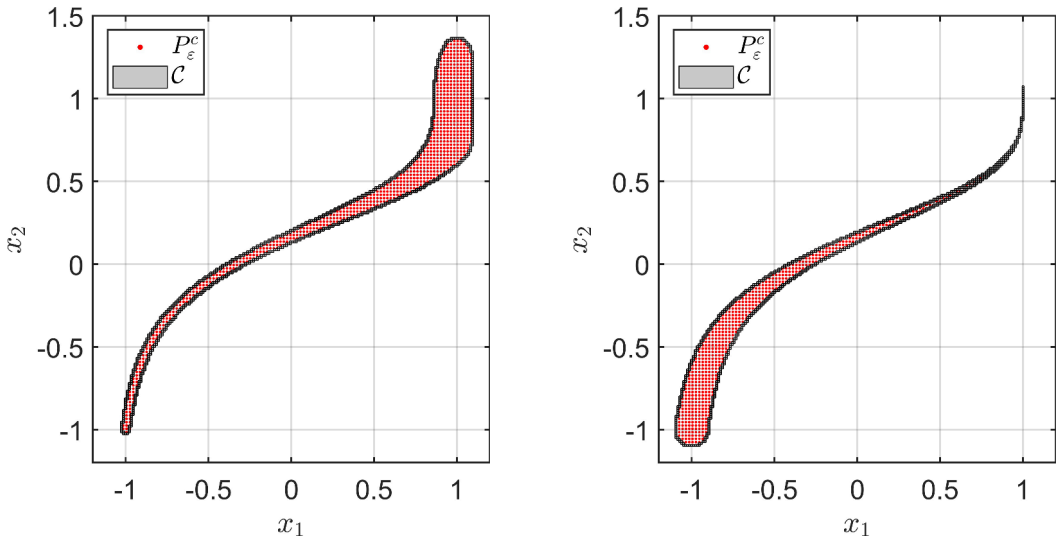
c) By construction, only boxes with $\theta(B) = 0$ (or, in practice, with $\theta(B)$ being small) are added to the covering $\mathcal{C}$. Thus, if our assumption on the piecewise smooth structure of ∂P_ε^r does not hold, then the effect is that our method might stop prematurely. In theory, this problem could be solved by always choosing $T_\mathcal{B} = N(\bar{x})$ in step 5 (and therefore not relying on any smoothness of $\varphi^{-1}(0)$), but this would clearly significantly decrease the efficiency of the algorithm.

3.2.4 Examples

In this section, we will apply Algorithm 3.4 to some examples. For the evaluation of φ via the solution of (3.11), we use the SQP method of the MATLAB function `fmincon`. (The MATLAB function `quadprog` for quadratic problems requires positive definiteness of the matrix in the objective, which is not satisfied in our case.) The derivative $D\varphi$ is approximated via finite differences. For the solution of (3.19) we also use the SQP method. To decide whether $B \cap \varphi^{-1}(0) \neq \emptyset$ in step 8, we use the threshold $\theta(B) < 10^{-10}$.

We will begin by revisiting Example 3.2.5 to compare the result $\mathcal{C}$ of Algorithm 3.4 with the simple, pointwise approximation of P_ε^r that we computed earlier.

Example 3.2.12. *Let f^r be defined as in Example 3.2.5. Figure 3.13 shows the result of Algorithm 3.4 for the box radius $r = 3 \cdot 2^{-9} \approx 0.5859 \cdot 10^{-3}$ and two different error bounds ε. For both error bounds, our method correctly computed the boundary of P_ε^r.*



(a) $\varepsilon = (0.2, 0.05)^\top$ **(b)** $\varepsilon = (0, 0.2)^\top$

Figure 3.13: Result $\mathcal{C}$ of Algorithm 3.4 compared to a pointwise approximation of P_ε^r for Example 3.2.12 for different error bounds.

Next, we will revisit Example 3.2.1, where it was highlighted why the Pareto critical set P_c^r of the inexact objective vector might not be a good approximation of P_c.

Example 3.2.13. *Let f and f^r be defined as in Example 3.2.1. Then (3.9) is satisfied for $\varepsilon = (\frac{1}{10}, \frac{1}{10})^\top$. Figure 3.14 shows the result of Algorithm 3.4 for the box radius $r = 3 \cdot 2^{-9} \approx 0.5859 \cdot 10^{-3}$. In this example, the Pareto critical set P_c^r of the inexact objective function is disconnected, while both the actual Pareto critical set P_c and the superset P_ε^r (as given by its boundary) are connected. This shows that P_ε^r may be better suited for making predictions about the (topological) properties of P_c than P_c^r.*

So far, we only considered problems with $n = 2$ variables since they are easier to illustrate. As a higher-dimensional example, we will now consider a problem with $k = 3$ objective functions in $n = 3$ variables.

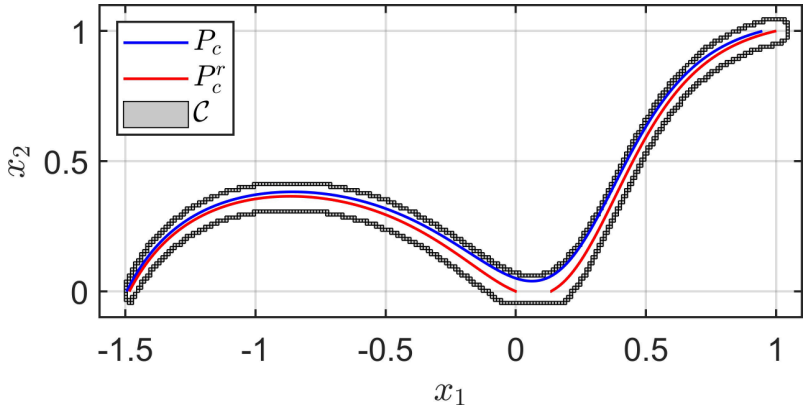


Figure 3.14: The sets P_c and P_c^r and the result of Algorithm 3.4 in Example 3.2.13.

Example 3.2.14. *Consider the MOP* $\min_{x\in\mathbb{R}^3} f(x)$ *with*

$$f(x) = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 \\ x_1^2 + (x_2-1)^2 + (x_3-1)^2 \\ (x_1-1)^2 + (x_2-1)^2 + (x_3-\frac{1}{2})^2 + \frac{\delta}{c\sqrt{2}}\sin(c(x_1+x_3)) \end{pmatrix} \tag{3.20}$$

for $c, \delta > 0$ *and the inexact objective vector*

$$f^r : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 \\ x_1^2 + (x_2-1)^2 + (x_3-1)^2 \\ (x_1-1)^2 + (x_2-1)^2 + (x_3-\frac{1}{2})^2 \end{pmatrix}.$$

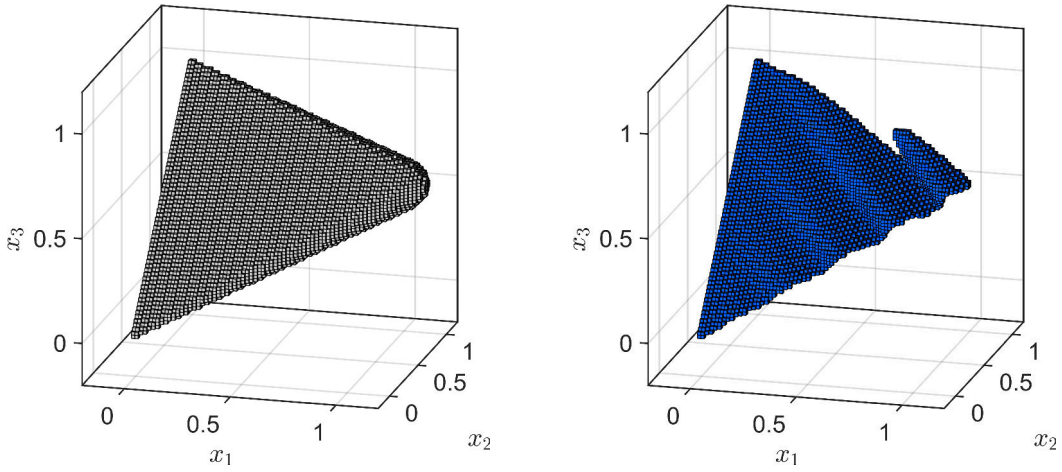
Then $\|\nabla f_1(x) - \nabla f_1^r(x)\| = \|\nabla f_2(x) - \nabla f_2^r(x)\| = 0$ *for all* $x \in \mathbb{R}^n$ *and*

$$\begin{aligned} \|\nabla f_3(x) - \nabla f_3^r(x)\| &= \left\| \left(\frac{\delta}{\sqrt{2}}\cos(c(x_1+x_3)), 0, \frac{\delta}{\sqrt{2}}\cos(c(x_1+x_3)) \right)^\top \right\| \\ &= \delta|\cos(c(x_1+x_3))| \le \delta \quad \forall x \in \mathbb{R}^n, \end{aligned}$$

so (3.9) *is satisfied for the error bounds* $\varepsilon = (0,0,\delta)^\top$. *For our numerical experiment, we choose* $c = 20$ *and* $\delta = \frac{1}{5}$. *Due to the simple structure of* f^r, *it is easy to see that its Pareto critical set* P_c^r *is the convex hull of the points* $(0,0,0)^\top$, $(0,1,1)^\top$ *and* $(1,1,\frac{1}{2})^\top$ *(cf.* (2.4)*). Figure 3.15 shows the approximation* $\mathcal{C}$ *of* ∂P_ε^r *from Algorithm 3.4 and the Pareto critical set* P_c *of* (3.20) *(computed via Algorithm 3.1). Figure 3.16 shows two* x_2-x_3*-slices of both sets, showing that* P_c *lies within* ∂P_ε^r *as expected.*

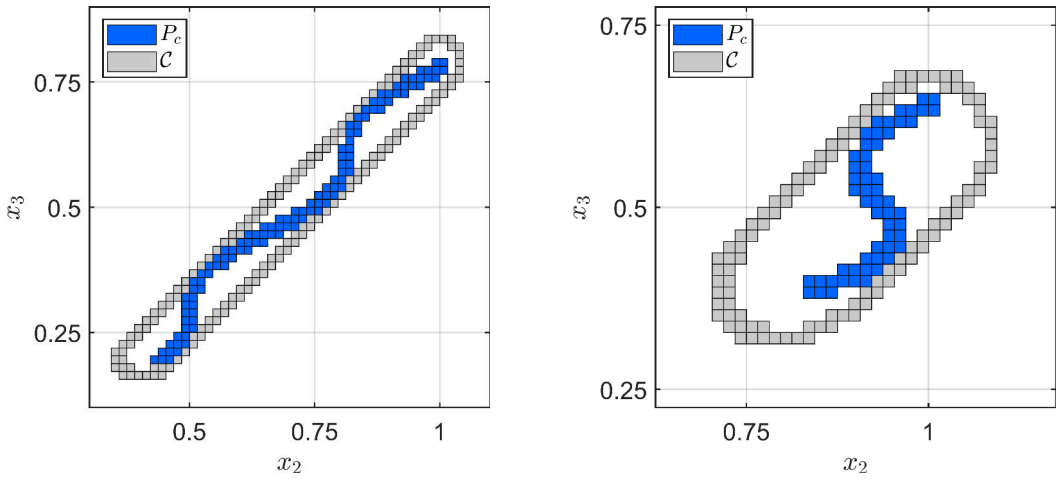
3.2.5 Application to PDE-constrained MOPs

A typical situation in practice where only inexact objective functions are available is the situation where the original objective functions are too computationally expensive to evaluate and are therefore replaced by cheaper surrogate models. For example, this is done in the area of *PDE-constrained multiobjective optimization* (also sometimes referred to as *multiobjective optimal control*), which we will consider in this section. Only a short introduction to this topic will be given here. For the details, see [Trö10; IUV17; Pei17].



(a) $\mathcal{C}$ **(b)** P_c

Figure 3.15: Superset P_ε^r (via $\mathcal{C}$) and the Pareto critical set P_c of the MOP (3.20) in Example 3.2.14.



(a) $x_1 = 53 \cdot 2^{-7} \approx 0.4141$ **(b)** $x_1 = 105 \cdot 2^{-7} \approx 0.8203$

Figure 3.16: Comparison of $\mathcal{C}$ (gray) and P_c (blue) via x_2-x_3-slices for two different x_1 in Example 3.2.13.

We consider MOPs of the (abstract) form

$$\begin{aligned} &\min_{x \in U, y \in Y} J(x, y), \\ \text{s.t.} \quad &e(x, y) = 0, \end{aligned} \tag{MOCP}$$

where

- U is the *control space*, containing the *controls* $x \in U$,
- Y is the *state space*, containing the *states* $y \in Y$,
- $J : U \times Y \to \mathbb{R}^k$ is the vector of (state-dependent) objectives,
- e is a PDE called the *state equation*.

Many physical systems can be modeled in this form. For example, if a room is to be heated, then the control x is the dial on the heating device, e is the *heat equation* and y is the resulting temperature distribution in the room. The objective functions could be the distance to some target temperature distribution (known as *tracking type* functions) and the energy cost of the heating process.

If the state is uniquely determined by the control and the state equation, i.e., if there is a *solution operator* $\mathcal{S} : U \to Y$ with

$$e(x, y) = 0 \;\Leftrightarrow\; y = \mathcal{S}(x)$$

for all $x \in U$, then (MOCP) can be written as

$$\min_{x \in U} f(x) \quad \text{for} \quad f(x) := J(x, \mathcal{S}(x)). \tag{3.21}$$

In theory, the solution of (MOCP) could be obtained by simply solving (3.21) with a standard MOP solver. But every evaluation of f involves the computation of the state of the system via the solution operator $\mathcal{S}$. In practice, $\mathcal{S}$ can be evaluated numerically via the *finite element method* (FEM), but it is computationally expensive to do so. So instead of solving (3.21) directly using the exact solution operator $\mathcal{S}$, an approximated operator $\mathcal{S}^r$ is used. This approximated operator can be obtained via reduced-order modeling techniques like *proper orthogonal decomposition* (POD) [KV01; KV02; BBV17] or the *reduced basis method* (RB) [QMN16]. Using $\mathcal{S}^r$, the objective vector of (3.21) can then be replaced by the cheaper, but inexact objective vector

$$f^r : U \to \mathbb{R}^k, \quad x \mapsto J(x, \mathcal{S}^r(x)).$$

Furthermore, it is possible to derive error estimators to obtain error bounds ε for f^r as in (3.9).

As an explicit example, we consider the problem (4.2) in [Ban+19], where the state equation is an elliptic advection-diffusion-reaction equation and the two dimensional control consists of the diffusivity in the domain and the strength and orientation of the advection field. There are four objective functions, of which the first three are of tracking type and the last one is the energy cost. For reference, the exact solution of this problem (computed via Algorithm 3.1 with the box radius $r = 2.5 \cdot 2^{-9} \approx 0.4883 \cdot 10^{-2}$ and FEM for the evaluation of the solution operator) is the blue set shown in Figure 3.17. For the generation of the surrogate model, the RB method from Section 3.2 in [Ban+19] is used. The corresponding error estimation is done in Section 3.3. For our inexact approach, a surrogate model is generated such that the resulting inexact objective function f^r satisfies the error bounds $\varepsilon = (0.03, 0.03, 0.01, 0.01)^\top$. The result of Algorithm 3.4 for this f^r is shown in red in Figure 3.17. Table 3.1 shows the performance of both the exact solution via Algorithm 3.1 and the inexact solution via Algorithm 3.4. More specifically, it shows the number of boxes in the box covering, the number of times the subproblems (3.1) and (3.19) had to be solved, and the total runtime of the algorithms. When comparing the runtime, we see that Algorithm 3.4 only needs about 1.6% of the runtime of Algorithm 3.1. (While this is a relatively large increase in efficiency, we have to note that it is strongly influenced by the lower dimension of ∂P_ε^r compared to P_c, which is only the case when $k > n$. Due to this, far less boxes are needed to cover ∂P_ε^r.)

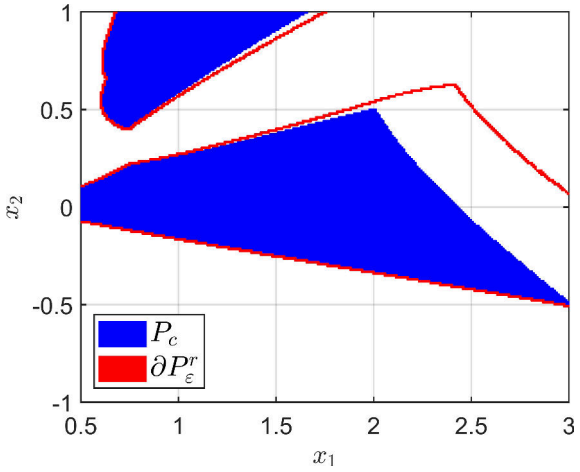


Figure 3.17: Exact solution via FEM and the boundary of the superset P_ε^r for the error bounds $\varepsilon = (0.03, 0.03, 0.01, 0.01)^\top$.

Table 3.1: Comparison of the performance of the exact solution via Algorithm 3.1 and the inexact solution via Algorithm 3.4 from Figure 3.17. The number of subproblems is split up into subproblems for the continuation and initialization (cf. Remark 3.1.3 a)).

Algorithm	# Boxes	# Subproblems	Runtime (in seconds)
Algo. 3.1	15916	$18721 + 25$	$17501s$
Algo. 3.4	899	$1027 + 225$	$276s$

When we consider the accuracy of the approximation of P_c by ∂P_ε^r in Figure 3.17, we see that it is relatively high everywhere except on the right-hand side of the lower connected component of P_c. To improve the quality, a more accurate surrogate model is computed, this time such that the error bounds $\varepsilon = (0.03, 0.01, 0.01, 0.01)^\top$ are satisfied. The corresponding result (with a box radius $2.5 \cdot 2^{-10} \approx 0.2441 \cdot 10^{-2}$) is shown in Figure 3.18, where the size of the gap on the right-hand side of the lower connected component significantly decreased.

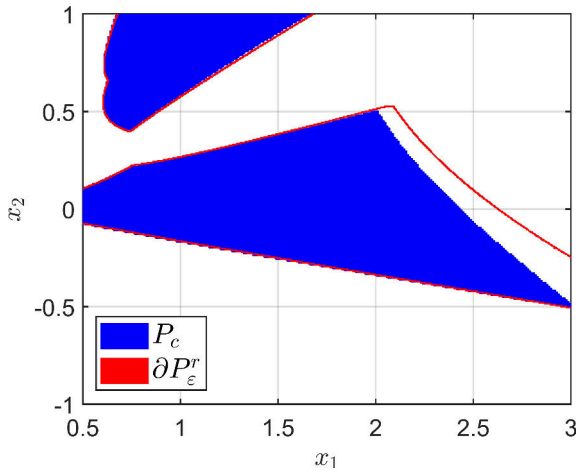


Figure 3.18: Exact solution via FEM and the boundary of the superset P_ε^r for the error bounds $\varepsilon = (0.03, 0.01, 0.01, 0.01)^\top$.

4 An efficient descent method for nonsmooth problems

There are many single-objective optimization problems that are naturally nonsmooth. Popular examples are obstacle problems in optimal shape design, clustering problems with nonsmooth similarity measures and problems involving exact penalty functions in constrained optimization. See [BKM14] for an overview. Furthermore, there are cases where these problems are equipped with multiple objectives. For example, in [MM93], the authors consider an obstacle problem where an elastic string is pushed against a rigid obstacle by some vertical force. The objectives are to maximize the contact area between the obstacle and the string while using minimal force. This creates the need for solution methods for nonsmooth MOPs. As already discussed at the beginning of Section 2.2.2, the class of locally Lipschitz continuous functions is a suitable class of "nonsmooth" objective functions.

This chapter will introduce a descent method for computing Pareto critical points of locally Lipschitz MOPs. In the smooth case in Section 2.3.2, the descent method was based on a direction $\bar{v}$, along which all objectives f_i decrease, and a step length $\bar{t}$, which assures a sufficient amount of improvement for all f_i. This will be generalized to the nonsmooth case via the concepts and tools from nonsmooth analysis introduced in Section 2.2.2. By Lemma 2.2.20, if we have a direction v that has a negative scalar product with all subgradients of all objective functions, then we have a guaranteed descent along v. As shown in [AGG15], such a direction can be computed as the element with the smallest norm in the negative convex hull of all subdifferentials. In the following, we begin by formally introducing and discussing the resulting descent direction in Section 4.1. As Clarke subdifferentials are difficult to compute in practice, we then introduce the so-called *Goldstein ε-subdifferential* [Gol77] in Section 4.2, and discuss its efficient approximation (which is based on the ideas in [MAY12]) in Section 4.3. Combined with an Armijo-like step length, we obtain a descent method for which we prove convergence to points that satisfy a necessary condition for Pareto criticality in Section 4.4. Additionally, for a small extension of our method, we show convergence to actual Pareto critical points. In Section 4.5, our method is applied to numerical examples. Using a set of test problems, a comparison to the multiobjective proximal bundle method from Section 2.3.2 indicates that the performance of our method is competitive. Furthermore, we investigate the practical behavior of our method via examples and discuss an extension to approximate entire Pareto sets.

Throughout this chapter, we consider the unconstrained minimization of an objective vector $f : \mathbb{R}^n \to \mathbb{R}^k$ which is locally Lipschitz continuous (cf. (2.15)). Although we will prove convergence of our method to Pareto critical points, it is much more likely that our method finds points that are at least locally Pareto optimal,

since we enforce a decrease of the objectives in each iteration. Thus, in contrast to the previous chapter, the method in this chapter is better suited for computing the actual Pareto set and only theoretically able to compute the Pareto critical set of nonsmooth MOPs.

Parts of this chapter have been previously published in [GP21a], to which the author of this thesis was the main contributor.

4.1 Theoretical descent direction

By Lemma 2.2.20, if we have a direction $v \in \mathbb{R}^n$ with

$$\max_{\xi \in \partial^{\cup} f(x)} \langle \xi, v \rangle = \max_{\xi \in \partial f_i(x), i \in \{1,\dots,k\}} \langle \xi, v \rangle < 0,$$

where $\partial^{\cup} f(x)$ is defined as in (2.16), then we have guaranteed descent along v. In this case, we say that v is a *descent direction of f in x*. By choosing $W = \partial^{\cup} f(x)$ in Lemma 2.2.1, a descent direction can be obtained via

$$v(x) := \operatorname*{arg\,min}_{\xi \in -\partial^{\cup} f(x)} \|\xi\|^2, \tag{4.1}$$

as suggested in [AGG15]. In particular, if x is not Pareto critical, it holds

$$\langle v(x), \xi \rangle \leq -\|v(x)\|^2 < 0 \quad \forall \xi \in \partial^{\cup} f(x). \tag{4.2}$$

The following examples visualizes this descent direction.

Example 4.1.1. *Consider the simple MOP from Example 2.2.25. Figure 4.1 shows the (uniformly scaled) direction resulting from* (4.1) *in a grid of points. It coincides with the descent direction* (2.19) *from the smooth case everywhere except on the nondifferentiable set $\Omega = (\mathbb{R} \times \{0\}) \cup (\{0\} \times \mathbb{R})$ of f. We see that Ω causes discontinuous transitions in v, which cannot occur in the smooth case (cf. Lemma 2.3.3). For example, when considering the local behavior around $(0,0)^\top$, we see that v does not vanish close to the Pareto critical set.*

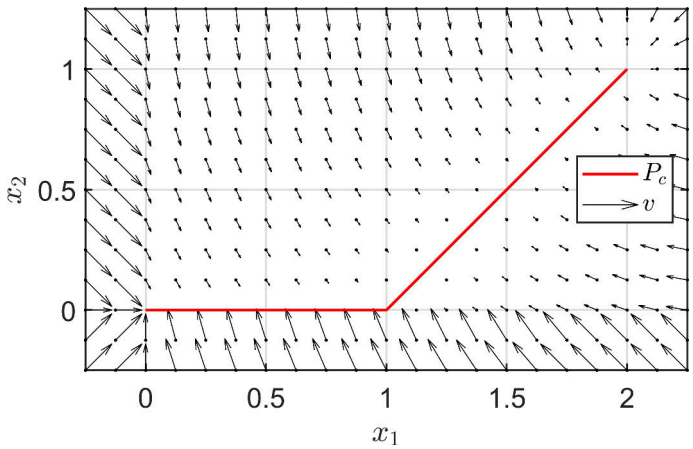


Figure 4.1: The descent direction resulting from (4.1) for Example 4.1.1.

While solving (4.1) is a theoretically sound way of computing descent directions and a natural generalization of the smooth case, there are mainly two reasons why it is problematic from a practical point of view:

(i) The first reason is the difficulty of obtaining the information required to solve (4.1). As $\partial^\cup f(x)$ is the union of all subdifferentials of the objectives in x, all these subdifferentials have to be computed. As discussed in Section 2.2.2, there is generally no easy way to do so. Furthermore, even if $\partial^\cup f(x)$ is known, it might not be possible to represent it in a simple way, e.g., as the convex hull of a finite number of points, making the solution of (4.1) even more involved.

(ii) The second reason is the fact that $\partial^\cup f(x)$ only captures the nonsmooth behavior of f if any of the components of f are nondifferentiable in x. By Rademacher's theorem (Theorem 2.2.14), the set of nondifferentiable points Ω of f is a null set in $\mathbb{R}^n$. Thus, in practice, we can generally not expect to actually encounter an $x \in \Omega$. In other words, any method that relies on $\partial^\cup f(x)$ to capture the nonsmoothness of f would behave like a smooth method applied to a nonsmooth problem, which generally fails to work (cf. [AO19]).

In the single-objective case, the result of (4.1) reduces to the direction from Proposition 6.2.4 in [Cla90], suffering from the same issues as mentioned above. To solve these issues in the single-objective case, different methods have been suggested. In *subgradient methods* [Sho85], a single subgradient $\xi \in \partial f(x)$ is used to approximate $\partial f(x)$. This implies that the solution of (4.1) is trivially given by $-\xi$, which makes it easy to implement in practice. The clear downside is that a single subgradient is not always sufficient for a good approximation of the local behavior of a nonsmooth function, so it cannot be guaranteed that $-\xi$ is actually a direction that yields (sufficient) descent. In *gradient sampling methods* [BLO05; Bur+20], $\partial f(x)$ is approximated by the convex hull of a finite number of gradients of f in randomly sampled, differentiable points around x. Due to the randomness, it can again not be guaranteed that the resulting direction actually yields sufficient descent. Furthermore, a check for differentiability of the objective functions in the randomly sampled points is needed, which is non-trivial in practice (cf. [HSS16]).

In the multiobjective case, the only methods that directly deal with (4.1) are generalizations of the subgradient method proposed in [Cru13; Net+13], but they were reported to be unsuitable for real life applications. In the multiobjective proximal bundle method from Section 2.3.2, (4.1) only occurs implicitly in (2.26). The problem (2.26) is solved via a linearization based on subgradients which were computed in earlier iterations, resulting in problem (2.27). This approach is comparable to the solution of (4.1) via approximation of $\partial^\cup f(x)$ with a finite number of subgradients (which are computed in a specific way).

In the following, we will describe a way to efficiently compute descent directions of locally Lipschitz MOPs by systematically computing an approximation of $\partial^\cup f(x)$ that is sufficient to obtain "good" descent directions via (4.1). As for the multiobjective proximal bundle method, we assume that for each $x \in \mathbb{R}^n$, we can compute a single subgradient for each objective f_i, $i \in \{1, \dots, k\}$. (This is sometimes referred to as *oracle information* [OS14].)

4.2 The Goldstein ε-subdifferential

By definition, $\partial f_i(x)$ is the convex hull of all (existing) limits $\lim_{j\to\infty} \nabla f_i(x^j)$ for all sequences $(x^j)_j \in \mathbb{R}^n$ converging to x. Thus, if we evaluate ∇f_i in points which are close to x (where f is differentiable) and then take the convex hull of the results, we expect to have some kind of approximation of $\partial f_i(x)$. To formalize this idea, we introduce the so-called *Goldstein ε-subdifferential* [Kiw10; Gol77]. To this end, for $\varepsilon \geq 0$ and $x \in \mathbb{R}^n$ let $\overline{B_\varepsilon}(x) := \{y \in \mathbb{R}^n : \|x - y\| \leq \varepsilon\}$.

Definition 4.2.1. *Let $\varepsilon \geq 0$ and $x \in \mathbb{R}^n$. Then*

$$\partial_\varepsilon f_i(x) := \operatorname{conv}\left(\bigcup_{y\in\overline{B_\varepsilon}(x)} \partial f_i(y)\right)$$

is the (Goldstein) ε-subdifferential of f_i in x. *An element $\xi \in \partial_\varepsilon f_i(x)$ is an* ε-subgradient.

Clearly, $\partial_0 f_i(x) = \partial f_i(x)$ and $\partial f_i(x) \subseteq \partial_\varepsilon f_i(x)$. The ε-subdifferential can be interpreted as a "stabilized" version of the Clarke subdifferential, since it not only contains the differential information of f_i in x, but from a (local) neighborhood of x. This deals with the problem (ii) from above, since the ε-subdifferential in x captures the nonsmoothness of f_i as soon as the distance of x to Ω_i is less or equal ε.

At first glance, the ε-subdifferential may seem even more difficult to compute than the Clarke subdifferential, since it is defined as the convex hull of the union of infinitely many Clarke subdifferentials. But note that since we assumed that we are able to compute a single Clarke subgradient at each point in $\mathbb{R}^n$, we are actually able to compute multiple ε-subgradients in x by simply computing multiple Clarke subgradients in $\overline{B_\varepsilon}(x)$. By systematically "sampling" ε-subgradients in this way, we will later be able to compute an approximation of the ε-subdifferential that is sufficient to compute descent directions. This will deal with problem (i) from the above discussion.

In the following, we will show some technical results about the ε-subdifferential. Analogously to (2.16), let

$$\partial_\varepsilon^\cup f(x) := \operatorname{conv}\left(\bigcup_{i=1}^k \partial_\varepsilon f_i(x)\right).$$

The following lemma shows that the ε-subdifferential has the same basic properties as the Clarke subdifferential. (As there are different ways to define ε-subdifferentials, we will give the full proof here for the sake of completeness.)

Lemma 4.2.2. *Let $x \in \mathbb{R}^n$ and $\varepsilon \geq 0$. The set $\partial_\varepsilon f_i(x)$ is nonempty, convex and compact. In particular, the same holds for $\partial_\varepsilon^\cup f(x)$.*

Proof. Non-emptiness and convexity of $\partial_\varepsilon f_i(x)$ and $\partial_\varepsilon^\cup f(x)$ are trivial. As the convex hull of a compact set is compact (cf. [BV10], Exercise 2.4.11) and the union of finitely many compact sets is compact, we only have to show compactness of

$$K := \bigcup_{y\in\overline{B_\varepsilon}(x)} \partial f_i(y).$$

To this end, let $(V_j)_{j\in J}$ be an open covering of K, i.e., V_j is open for all $j \in J$ and

$$K \subseteq \bigcup_{j\in J} V_j.$$

Let $y \in \overline{B_\varepsilon}(x)$. Since $\partial f_i(y)$ is compact and a subset of K, there is some finite subset $J(y)$ of the index set J with

$$\partial f_i(y) \subseteq \bigcup_{j\in J(y)} V_j =: V(y). \tag{4.3}$$

Since $V(y)$ is open and ∂f_i is upper semicontinuous, there is an open set $U(y) \subseteq \mathbb{R}^n$ with $y \in U(y)$ such that

$$\partial f_i(z) \subseteq V(y) \quad \forall z \in U(y). \tag{4.4}$$

Since $\overline{B_\varepsilon}(x)$ is compact and $(U(y))_{y\in\overline{B_\varepsilon}(x)}$ is an open covering of $\overline{B_\varepsilon}(x)$, there is a finite set of points $\{y^1, \ldots, y^N\}$ such that

$$\overline{B_\varepsilon}(x) \subseteq \bigcup_{l=1}^{N} U(y^l). \tag{4.5}$$

Now let $J' := \bigcup_{l=1}^{N} J(y^l)$ and $V' := \bigcup_{j\in J'} V_j$. By construction, $|J'|$ is finite. We will show that $K \subseteq V'$, implying that K is compact. Let $\xi \in K$. By definition, there is some $y' \in \overline{B_\varepsilon}(x)$ with $\xi \in \partial f_i(y')$. By (4.5), there is some $l \in \{1, \ldots, N\}$ with $y' \in U(y^l)$. By (4.3) and (4.4),

$$\xi \in \partial f_i(y') \subseteq V(y^l) = \bigcup_{j\in J(y^l)} V_j \subseteq \bigcup_{j\in J'} V_j = V',$$

completing the proof. □

Combination of the previous lemma with Theorem 2.2.22 and Lemma 2.2.1 yields the following corollary.

Corollary 4.2.3. *Let $x \in \mathbb{R}^n$ and $\varepsilon \geq 0$.*

a) If x is Pareto optimal, then

$$0 \in \partial_\varepsilon^\cup f(x). \tag{4.6}$$

b) Let

$$\bar{v}(x) := \operatorname*{arg\,min}_{\xi\in-\partial_\varepsilon^\cup f(x)} \|\xi\|^2. \tag{4.7}$$

Then either $\bar{v}(x) \neq 0$ and

$$\langle \bar{v}(x), \xi\rangle \leq -\|\bar{v}(x)\|^2 < 0 \quad \forall \xi \in \partial_\varepsilon^\cup f(x), \tag{4.8}$$

or $\bar{v}(x) = 0$ and there is no $v \in \mathbb{R}^n$ with $\langle v, \xi\rangle < 0$ for all $\xi \in \partial_\varepsilon^\cup f(x)$.

The previous corollary states that if we work with the ε-subdifferential instead of the Clarke subdifferential, we still have a necessary optimality condition and a way to compute descent directions. The difference is that since $\partial^\cup f(x) \subseteq \partial^\cup_\varepsilon f(x)$, the optimality condition (4.6) is weaker than the original condition (2.17). Furthermore, by (4.2) and (4.8), $-\|v(x)\|^2$ and $-\|\bar{v}(x)\|^2$ are upper bounds for the generalized directional derivatives $\partial^\circ f_i(x, v(x))$ and $\partial^\circ f_i(x, \bar{v}(x))$, respectively (cf. Remark 2.2.21). Since $-\|v(x)\|^2 \leq -\|\bar{v}(x)\|^2$, this means that $v(x)$ potentially yields a steeper descent than $\bar{v}(x)$. We compare both descent directions in the following example.

Example 4.2.4. *Consider the locally Lipschitz problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} (x_1-1)^2 + (x_2-1)^2 \\ x_1^2 + |x_2| \end{pmatrix}.$$

The set of nondifferentiable points of f is $\Omega = \mathbb{R} \times \{0\}$. Let $\varepsilon \geq 0$. For any $x \in \mathbb{R}^2$ we have

$$\partial f_1(x) = \{\nabla f_1(x)\} = \left\{ \begin{pmatrix} 2x_1 - 2 \\ 2x_2 - 2 \end{pmatrix} \right\} \quad \textit{and} \quad \partial_\varepsilon f_1(x) = 2\overline{B_\varepsilon}(x) - \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

For $x \in \Omega$ we have

$$\partial f_2(x) = \{2x_1\} \times [-1, 1] \quad \textit{and} \quad \partial_\varepsilon f_2(x) = (2x_1 + [-2\varepsilon, 2\varepsilon]) \times [-1, 1].$$

Figure 4.2 shows the Clarke subdifferentials (a) and the ε-subdifferentials (b) for $\varepsilon = 0.2$ and $x^* = (1.5, 0)^\top$. Additionally, the resulting descent directions are shown. In this case, the predicted descent is $-\|v\|^2 \approx -3.7692$ in (a) and $-\|\bar{v}\|^2 \approx -2.4433$ in (b).

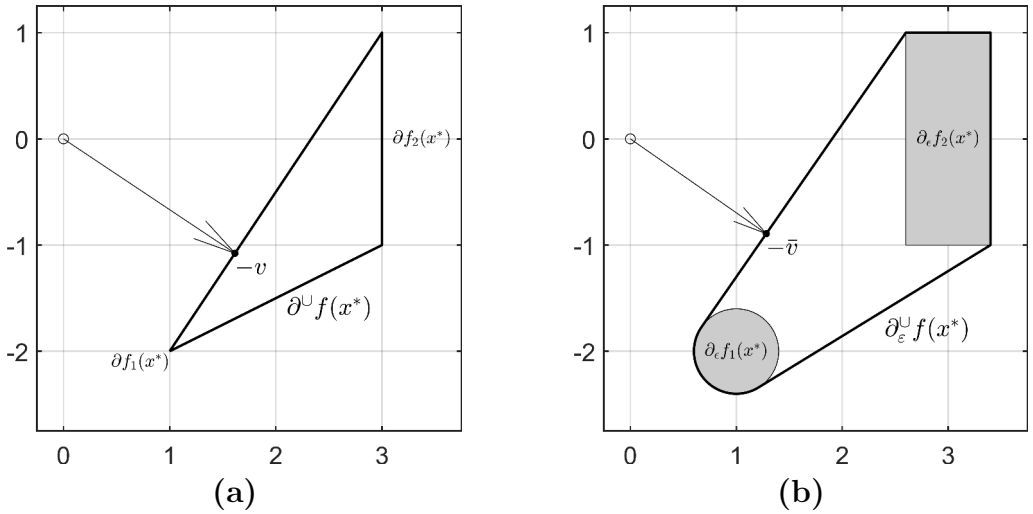


Figure 4.2: Clarke subdifferentials (a), ε-subdifferentials (b) and the corresponding descent directions for $\varepsilon = 0.2$ and $x^* = (1.5, 0)^\top$ in Example 4.2.4.

Figure 4.3 shows the same scenario for $x^* = (0.5, 0)^\top$. Here, the Clarke subdifferential still yields a descent, while the descent direction for the ε-subdifferential is zero. In other words, the weaker optimality condition (4.6) is satisfied, while the original KKT condition (2.17) is not satisfied.

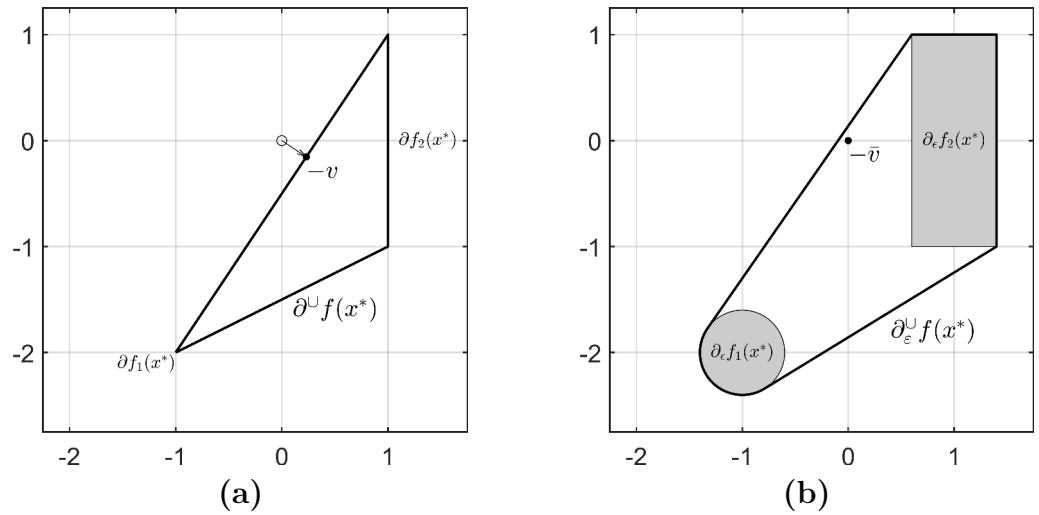


Figure 4.3: Clarke subdifferentials (a), ε-subdifferentials (b) and the corresponding (potentially vanishing) descent directions for $\varepsilon = 0.2$ and $x^* = (0.5, 0)^\top$ in Example 4.2.4.

By Lemma 2.2.20 and inequality (4.8) we know that if $\bar{v}(x) \neq 0$, then there is some $T > 0$ such that there is a guaranteed descent along $\bar{v}(x)$ for all step lengths $t \in (0, T]$. Due to the definition of $\bar{v}(x)$, it turns out that it is even possible to derive a lower bound for the largest possible choice for T in this case.

Lemma 4.2.5. *Let $x \in \mathbb{R}^n$ and $\varepsilon \geq 0$. If $\bar{v}(x) \neq 0$, then*

$$f_i(x + t\bar{v}(x)) \leq f_i(x) - t\|\bar{v}(x)\|^2 \quad \forall t \in \left(0, \frac{\varepsilon}{\|\bar{v}(x)\|}\right], i \in \{1, \ldots, k\}.$$

In particular,

$$f_i\left(x + \frac{\varepsilon}{\|\bar{v}(x)\|}\bar{v}(x)\right) \leq f_i(x) - \varepsilon\|\bar{v}(x)\| \quad \forall i \in \{1, \ldots, k\}.$$

Proof. For the sake of brevity we write $\bar{v} = \bar{v}(x)$. Let $t \in (0, \frac{\varepsilon}{\|\bar{v}\|}]$ and $i \in \{1, \ldots, k\}$. By applying the mean value theorem (Theorem 2.2.18) to x and $x + t\bar{v}$, we obtain

$$f_i(x + t\bar{v}) - f_i(x) = \langle \xi, t\bar{v} \rangle = t\langle \xi, \bar{v} \rangle$$

for some $\xi \in \partial f_i(x + r\bar{v})$ with $r \in (0, t)$. Since $\|x - (x + r\bar{v})\| = r\|\bar{v}\| < t\|\bar{v}\| \leq \varepsilon$, it follows that $\xi \in \partial_\varepsilon f_i(x)$. By (4.8), this implies $\langle \xi, \bar{v} \rangle \leq -\|\bar{v}\|^2$, so

$$\begin{aligned} & f_i(x + t\bar{v}) - f_i(x) \leq -t\|\bar{v}\|^2 \\ \Leftrightarrow \quad & f_i(x + t\bar{v}) \leq f_i(x) - t\|\bar{v}\|^2. \end{aligned}$$

□

4.3 Efficient computation of descent directions

In this section, we will derive a way to efficiently compute approximations of the descent direction from (4.7) in practice, i.e., when only a single Clarke subgradient

can be computed in each point. Similar to the gradient sampling method, the idea of our method is to approximate $\partial_\varepsilon^\cup f(x)$ by the convex hull of a finite number $m \in \mathbb{N}$ of ε-subgradients $\xi_1, \dots, \xi_m \in \partial_\varepsilon^\cup f(x)$. Since it is impossible to know a priori how many and which ε-subgradients are required to obtain a good descent direction, we solve (4.7) multiple times while enriching our approximation with new ε-subgradients, until an acceptable descent direction is found. To be able to do this, we have to derive a way to improve our approximation $\operatorname{conv}(\{\xi_1, \dots, \xi_m\})$ of $\partial_\varepsilon^\cup f(x)$ by adding new ε-subgradients and we have to characterize acceptable descent directions.

To this end, let $W = \{\xi_1, \dots, \xi_m\} \subseteq \partial_\varepsilon^\cup f(x)$ and define

$$\tilde{v} := \operatorname*{arg\,min}_{\xi \in -\operatorname{conv}(W)} \|\xi\|^2. \tag{4.9}$$

Let $c \in (0,1)$. Motivated by Lemma 4.2.5, we say that $\tilde{v}$ is an *acceptable* descent direction if

$$f_i\left(x + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}\right) \leq f_i(x) - c\varepsilon\|\tilde{v}\| \quad \forall i \in \{1, \dots, k\}. \tag{4.10}$$

The parameter c is similar to an Armijo parameter (cf. (2.24)). If $\tilde{v}$ is not an acceptable direction, i.e., if the index set $I \subseteq \{1, \dots, k\}$ for which (4.10) is violated is nonempty, then we want to compute a new ε-subgradient $\xi' \in \partial_\varepsilon^\cup f(x)$ such that $W \cup \{\xi'\}$ yields a steeper descent direction. Intuitively, (4.10) being violated for some $i \in \{1, \dots, k\}$ means that the local behavior of f_i in the direction $\tilde{v}$ is not sufficiently captured in W. Thus, for each $i \in I$, we expect there to be some $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ and $\xi' \in \partial f_i(x + t'\tilde{v})$ such that $W \cup \{\xi'\}$ is a better approximation of $\partial_\varepsilon^\cup f(x)$ than W. This is proven in the following lemma.

Lemma 4.3.1. *Let $c \in (0,1)$, $W = \{\xi_1, \dots, \xi_m\} \subseteq \partial_\varepsilon^\cup f(x)$ and $\tilde{v}$ be the solution of (4.9). If $\tilde{v}$ is not an acceptable direction, i.e., if (4.10) is violated for some $i \in \{1, \dots, k\}$, then there is some $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ and $\xi' \in \partial f_i(x + t'\tilde{v})$ such that*

$$\langle \tilde{v}, \xi' \rangle > -c\|\tilde{v}\|^2. \tag{4.11}$$

Furthermore,

$$\xi' \in \partial_\varepsilon^\cup f(x) \setminus \operatorname{conv}(W).$$

Proof. Assume that there is some $i \in \{1, ..., k\}$ such that (4.10) is violated and the statement does not hold, i.e., for all $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ and all $\xi' \in \partial f_i(x + t'\tilde{v})$ we have

$$\langle \tilde{v}, \xi' \rangle \leq -c\|\tilde{v}\|^2. \tag{4.12}$$

By applying the mean value theorem to x and $x + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}$ as in the proof of Lemma 4.2.5, we obtain some $r \in (0, \frac{\varepsilon}{\|\tilde{v}\|})$ and $\bar{\xi} \in \partial f_i(x + r\tilde{v})$ such that

$$f_i\left(x + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}\right) - f_i(x) = \frac{\varepsilon}{\|\tilde{v}\|}\langle \bar{\xi}, \tilde{v} \rangle.$$

With (4.12) it follows that

$$f_i\left(x + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}\right) = f_i(x) + \frac{\varepsilon}{\|\tilde{v}\|}\langle \bar{\xi}, \tilde{v} \rangle \leq f_i(x) - c\varepsilon\|\tilde{v}\|,$$

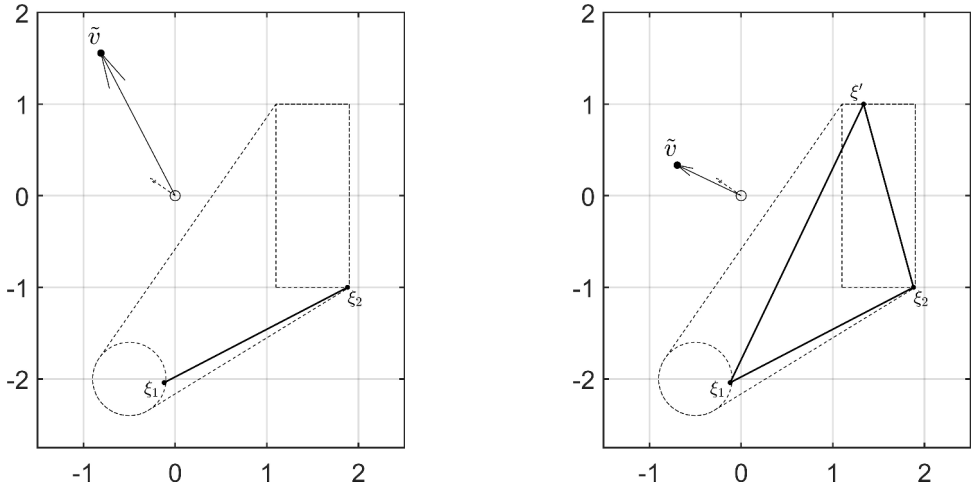
which is a contradiction to (4.10) being violated.
So there must be some $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ and $\xi' \in \partial f_i(x + t'\tilde{v}) \subseteq \partial_\varepsilon^\cup f(x)$ such that $\langle \tilde{v}, \xi' \rangle > -c\|\tilde{v}\|^2 > -\|\tilde{v}\|^2$. Recall that by Lemma 2.2.1 and by the definition of $\tilde{v}$, we have

$$\langle \tilde{v}, \xi \rangle \leq -\|\tilde{v}\|^2 < 0 \quad \forall \xi \in \operatorname{conv}(W),$$

so $\xi' \notin \operatorname{conv}(W)$, completing the proof. □

The following example visualizes the previous lemma.

Example 4.3.2. *We revisit the MOP from Example 4.2.4 and consider* $\varepsilon = 0.2$ *and* $x^* = (0.75, 0)^\top$. *The dashed lines in Figure 4.4 show the* ε*-subdifferentials,* $\partial_\varepsilon^\cup f(x^*)$ *and the direction* $\bar{v}$ *(cf. Figure 4.2 and 4.3).*



(a) $\operatorname{conv}(\{\xi_1, \xi_2\})$ **(b)** $\operatorname{conv}(\{\xi_1, \xi_2, \xi'\})$

Figure 4.4: Approximations of $\partial_\varepsilon^\cup f(x^*)$ for $\varepsilon = 0.2$ and $x^* = (0.75, 0)^\top$ in Example 4.3.2.

Let $y = (0.94, -0.02)^\top$. *Then* $\|x^* - y\| \approx 0.191 \leq \varepsilon$, *so* $y \in \overline{B_\varepsilon}(x^*)$ *and*

$$\partial_\varepsilon f_1(x^*) \supseteq \partial f_1(y) = \left\{ \begin{pmatrix} -0.12 \\ -2.04 \end{pmatrix} \right\} =: \{\xi_1\},$$
$$\partial_\varepsilon f_2(x^*) \supseteq \partial f_2(y) = \left\{ \begin{pmatrix} 1.88 \\ -1 \end{pmatrix} \right\} =: \{\xi_2\}.$$

Let $W := \{\xi_1, \xi_2\}$ *and* $\operatorname{conv}(W)$ *be the current approximation of* $\partial_\varepsilon^\cup f(x^*)$, *shown as the solid line in Figure 4.4(a). We choose* $c = 0.25$ *and consider the direction* $\tilde{v}$ *resulting from* (4.9), *also shown in Figure 4.4(a). Checking if* $\tilde{v}$ *is an acceptable direction, we see that*

$$f_2\left(x^* + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}\right) \approx 0.6101 > 0.4748 \approx f_2(x^*) - c\varepsilon\|\tilde{v}\|.$$

By Lemma 4.3.1, this means that there have to be $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ *and* $\xi' \in \partial f_2(x^* + t'\tilde{v})$ *such that* $\langle \tilde{v}, \xi' \rangle > -c\|\tilde{v}\|^2$. *For example, in this case, we can choose* $t' = \frac{1}{2}\frac{\varepsilon}{\|\tilde{v}\|}$,

resulting in

$$\partial f_2(x^* + t'\tilde{v}) \approx \left\{ \begin{pmatrix} 1.4077 \\ 1 \end{pmatrix} \right\} =: \{\xi'\},$$
$$\langle \tilde{v}, \xi' \rangle \approx 0.4172 > -0.7696 \approx -c\|\tilde{v}\|^2.$$

Figure 4.4(b) shows the enriched approximation $W \cup \{\xi'\}$ and the new direction $\tilde{v}$. By checking (4.10)*, we see that $\tilde{v}$ is acceptable. (Note that in general, multiple improvement steps are needed to obtain an acceptable direction.)*

While Lemma 4.3.1 shows the existence of a new ε-subgradient ξ' that improves the approximation of $\partial_\varepsilon^\cup f(x)$, its proof is non-constructive. In the following, we will discuss how such a ξ' can be computed (almost reliably) in practice. To this end, assume that $\tilde{v}$ from (4.9) is not an acceptable direction and let $i \in \{1, \dots, k\}$ be an index such that (4.10) is violated. Our goal is to find an ε-subgradient ξ' as in (4.11). Consider the function

$$h_i : \mathbb{R} \to \mathbb{R}, \quad t \mapsto f_i(x + t\tilde{v}) - f_i(x) + ct\|\tilde{v}\|^2.$$

Then $h_i(0) = 0$ and, since $\tilde{v}$ is not acceptable, $h_i(\frac{\varepsilon}{\|\tilde{v}\|}) > 0$. If f_i would be continuously differentiable, then

$$h_i'(t) := \nabla h_i(t) = \langle \nabla f_i(x + t\tilde{v}), \tilde{v} \rangle + c\|\tilde{v}\|^2,$$

and (4.11) would be equivalent to $h_i'(t') > 0$, i.e., h_i' being monotonically increasing in t'. In the nonsmooth case, the idea is to borrow this result by searching for a $t' \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ such that h_i is increasing around t'. This can be done via Algorithm 4.1, which performs bisections while simultaneously checking (4.11) until a new ε-subgradient is found.

Algorithm 4.1 Computing new ε-subgradients

Given: Current point $x \in \mathbb{R}^n$, direction $\tilde{v} \in \mathbb{R}^n$, index $i \in \{1, \dots, k\}$ violating (4.10), tolerance $\varepsilon > 0$, Armijo parameter $c \in (0, 1)$.

1: Initialize $a = 0$, $b = \frac{\varepsilon}{\|\tilde{v}\|}$ and $t = \frac{a+b}{2}$.
2: Compute $\xi' \in \partial f_i(x + t\tilde{v})$.
3: If $\langle \tilde{v}, \xi' \rangle > -c\|\tilde{v}\|^2$ then stop.
4: If $h_i(b) > h_i(t)$ then set $a = t$. Otherwise set $b = t$.
5: Set $t = \frac{a+b}{2}$ and go to step 2.

In terms of its convergence, we have the following result.

Lemma 4.3.3. *Let $(t_j)_j$ be the sequence generated by Algorithm 4.1.*

a) If $(t_j)_j$ is finite, then some ξ' was found satisfying (4.11)*.*

b) If $(t_j)_j$ is infinite, then it converges to some $\bar{t} \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$ with $h_i(\bar{t}) \geq h_i(\frac{\varepsilon}{\|\tilde{v}\|})$ such that

(i) there is some $\xi' \in \partial f_i(x + \bar{t}\tilde{v})$ satisfying (4.11) *or*

(ii) $0 \in \partial h_i(\bar{t})$, i.e., $\bar{t}$ is a critical point *of h_i.*

Proof. The case where $(t_j)_j$ is finite is trivial. So assume that $(t_j)_j$ is infinite. By construction, $(t_j)_j$ is a Cauchy sequence in the compact set $[0, \frac{\varepsilon}{\|\tilde{v}\|}]$, so it has to converge to some $\bar{t} \in [0, \frac{\varepsilon}{\|\tilde{v}\|}]$. Since we assumed that (4.10) is violated for the index i, we have

$$h_i(0) = 0 \quad \text{and} \quad h_i\left(\frac{\varepsilon}{\|\tilde{v}\|}\right) > 0.$$

Let $(a_j)_j$ and $(b_j)_j$ be the sequences corresponding to a and b in Algorithm 4.1 (at the start of each iteration in step 2, i.e., $a_1 = 0$ and $b_1 = \frac{\varepsilon}{\|\tilde{v}\|}$).
Note that $\lim_{j\to\infty} b_j = \bar{t}$ and $(h_i(b_j))_j$ is non-decreasing (due to step 4), so the continuity of h_i implies

$$h_i(\bar{t}) = \lim_{j\to\infty} h_i(b_j) \geq h_i(b_1) = h_i\left(\frac{\varepsilon}{\|\tilde{v}\|}\right).$$

In particular we must have $\bar{t} > 0$, since $h_i(0) = 0$ and $h_i(\frac{\varepsilon}{\|\tilde{v}\|}) > 0$. By construction,

$$h_i(a_j) < h_i(b_j) \quad \forall j \in \mathbb{N}.$$

Thus, by the mean value theorem (Theorem 2.2.18), there has to be some $r_j \in (a_j, b_j)$ such that (in set notation)

$$0 < h_i(b_j) - h_i(a_j) \in \langle \partial h_i(r_j), b_j - a_j \rangle = \partial h_i(r_j)(b_j - a_j).$$

In particular, $\lim_{j\to\infty} r_j = \bar{t}$ and since $a_j < b_j$, it follows $\partial h_i(r_j) \cap \mathbb{R}^{>0} \neq \emptyset$ for all $j \in \mathbb{N}$. Due to the upper semicontinuity of ∂h_i, this implies

$$\partial h_i(\bar{t}) \cap \mathbb{R}^{\geq 0} \neq \emptyset. \tag{4.13}$$

By the chain rule for locally Lipschitz functions (cf. Theorem 3.19 in [BKM14]), we have

$$\partial h_i(t) \subseteq \operatorname{conv}\left(\left\{\xi^\top \tilde{v} + c\|\tilde{v}\|^2 : \xi \in \partial f_i(x + t\tilde{v})\right\}\right) = \langle \tilde{v}, \partial f_i(x + t\tilde{v})\rangle + c\|\tilde{v}\|^2 \tag{4.14}$$

for all $t \in \mathbb{R}$. By (4.13), this means there is some $\xi \in \partial f_i(x + \bar{t}\tilde{v})$ such that

$$\langle \tilde{v}, \xi \rangle \geq -c\|\tilde{v}\|^2.$$

If there exists $\xi' \in \partial f_i(x + \bar{t}\tilde{v})$ with $\langle \tilde{v}, \xi' \rangle > -c\|\tilde{v}\|^2$ then we are in case b)(i). Otherwise, by (4.14), $\langle \tilde{v}, \xi \rangle \leq -c\|\tilde{v}\|^2$ for all $\xi \in \partial f_i(x + \bar{t}\tilde{v})$ implies $\partial h_i(\bar{t}) \subseteq \mathbb{R}^{\leq 0}$. By (4.13) and by convexity of $\partial h_i(\bar{t})$, this implies $0 \in \partial h_i(\bar{t})$, i.e., case b)(ii). □

The following example shows that there are indeed functions for which Algorithm 4.1 does not stop.

Example 4.3.4. *We will construct a function h for which Algorithm 4.1 does not stop. To this end, for $t \in (0, 1)$, let*

$$\begin{aligned} j(t) &:= \lfloor 1 - \log_2(1 - t) \rfloor, \\ q(t) &:= 1 - 2^{-(j(t)-1)}, \\ g(t) &:= -2^{-(j(t)+1)}(\cos(2^{j(t)}\pi(t - q(t))) - 1) + q(t), \end{aligned}$$

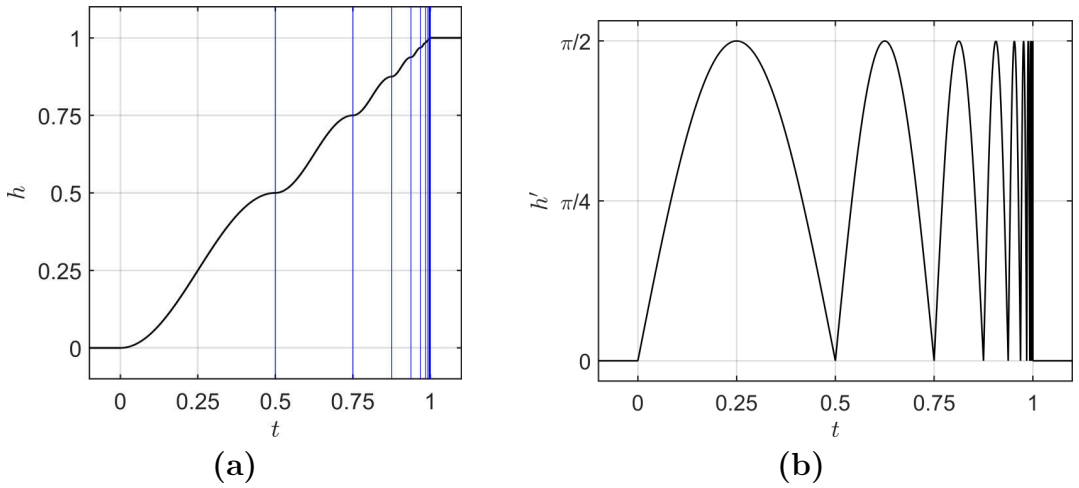
where $\lfloor\cdot\rfloor$ is the floor function, i.e., $\lfloor t\rfloor$ is the smallest integer less or equal t. Define

$$h : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \begin{cases} 0, & t \leq 0, \\ g(t), & t \in (0,1), \\ 1, & t \geq 1. \end{cases}$$

The graph of h is shown in Figure 4.5(a). The local Lipschitz continuity of h outside of 1 is obvious. For $t \in (0,1)$, we have

$$\begin{aligned} |1 - h(t)| &= 1 + 2^{-(j(t)+1)}(\cos(2^i \pi(t - q(t))) - 1) - q(t) \\ &\leq 1 - q(t) = 2^{-(j(t)-1)} = 2^{-\lfloor 1 - \log_2(1-t)\rfloor + 1} \\ &= 2^{-(\lfloor 1-\log_2(1-t)\rfloor - (1-\log_2(1-t)))} \cdot 2^{-((1-\log_2(1-t))-1)} \\ &= 2^{(1-\log_2(1-t)) - \lfloor 1-\log_2(1-t)\rfloor}(1-t) \\ &\leq 2(1-t) = 2|1-t|, \end{aligned}$$

implying that h is locally Lipschitz in 1 as well. The graph of the derivative of h (where defined) is shown in Figure 4.5(b). It is possible to show that $\partial h(0) = \{0\}$



(a) **(b)**

Figure 4.5: **(a)** Graph of h (black) and the sequence $(t_j)_j$ (blue) resulting from Algorithm 4.1 in Example 4.3.4. **(b)** Graph of h' (where defined).

and $\partial h(1) = [0, \frac{\pi}{2}]$. Assuming $\frac{\varepsilon}{\|\tilde{v}\|} = 1$, application of Algorithm 4.1 produces the infinite sequence

$$t_j = 1 - 2^{-j}$$

converging to $\bar{t} = 1$. In particular, $h'(t_j) = 0$ for all $t \in \mathbb{N}$ and $0 \in \partial h(\bar{t})$, so we are in case b)(ii) in Lemma 4.3.3.

The previous example is clearly non-generic. The implication of Lemma 4.3.3 for the practical application of Algorithm 4.1 will be discussed in the following remark.

Remark 4.3.5. *a) Assume that $(t_j)_j$ is infinite with limit $\bar{t} \in (0, \frac{\varepsilon}{\|\tilde{v}\|}]$. In the proof of Lemma 4.3.3, we showed that there is a sequence $(r_j)_j \in (0, \frac{\varepsilon}{\|\tilde{v}\|})$ with $\lim_{j\to\infty} r_j = \bar{t}$ such that $\partial h_i(r_j) \cap \mathbb{R}^{>0} \neq \emptyset$ for all $j \in \mathbb{N}$. By the definition of the Clarke subdifferential, this implies that there is a sequence $(s_j)_j \in (0, \frac{\varepsilon}{\|\tilde{v}\|})$*

with $\lim_{j\to\infty} s_j = \bar{t}$ *such that* h_i *is differentiable in* s_j *and* $h_i'(s_j) > 0$ *for all* $j \in \mathbb{N}$. *Due to the upper semicontinuity of the Clarke subdifferential, each* s_j *has an open neighborhood* U_j *with*

$$\partial h_i(s) \subseteq \mathbb{R}^{>0} \quad \forall s \in U_j, j \in \mathbb{N}. \tag{4.15}$$

Thus, roughly speaking, even if we are in case b)(ii) in Lemma 4.3.3, there are open sets arbitrary close to $\bar{t}$ *on which we can potentially find new* ε*-subgradients (as in case b)(i)).*

b) *Note that if* h_i *is nonsmooth in some* $t \in \mathbb{R}$, *then* f_i *has to be nonsmooth in* $x + t\tilde{v}$. *In other words, the set of nonsmooth points of* h_i, *embedded into* $\mathbb{R}^n$ *via* $t \mapsto x + t\tilde{v}$, *is a subset of* $\Omega_i \cap (x + \mathbb{R}\tilde{v})$. *Since* Ω_i *is a null set in* $\mathbb{R}^n$ *(cf. Theorem 2.2.14) and* $x + \mathbb{R}\tilde{v}$ *is a one-dimensional affine linear subspace of* $\mathbb{R}^n$, *we expect that in the generic case,* $\Omega_i \cap (x + \mathbb{R}\tilde{v})$ *only contains finitely many points. In this case, there has to be some* $N \in \mathbb{N}$ *such that* $x + [a_j, b_j]\tilde{v}$ *(from the proof of Lemma 4.3.3) contains at most a single nonsmooth point of* f_i *for all* $j > N$. *If the algorithm does not stop, then this point has to be the limit* $\bar{t}$ *of* $(t_j)_j$. *In particular, for* $j > N$ *and* $t \in [a_j, b_j] \setminus \{\bar{t}\}$, (4.14) *turns into*

$$h_i'(t) = \langle \tilde{v}, \nabla f_i(x + t\tilde{v}) \rangle + c\|\tilde{v}\|^2.$$

Combined with (4.15), *this means that there are open sets arbitrarily close to* $\bar{t}$ *in which we are guaranteed to find new* ε*-subgradients.*

c) *If* f_i *is differentiable on* $x + (0, \frac{\varepsilon}{\|\tilde{v}\|}]\tilde{v}$, *then* (4.14) *turns into*

$$h_i'(t) = \langle \tilde{v}, \nabla f_i(x + t\tilde{v}) \rangle + c\|\tilde{v}\|^2 \quad \forall t \in \left(0, \frac{\varepsilon}{\|\tilde{v}\|}\right].$$

If case b)(i) of Lemma 4.3.3 would hold, then we would have $h_i'(\bar{t}) > 0$. *Due to upper semicontinuity of* h_i', *this would imply that there is an open set* $U \subseteq \mathbb{R}^{>0}$ *with* $\bar{t} \in U$ *such that* $h_i'(t) > 0$ *for all* $t \in U$. *In particular, there would be some* $N \in \mathbb{N}$ *such that* $h_i'(t_j) > 0$ *for all* $j > N$. *This is a contradiction to the algorithm not stopping. Thus, case b)(ii) must hold, i.e.,* $h_i'(\bar{t}) = 0$.

d) *In practice, if Algorithm 4.1 appears to produce an infinite sequence* $(t_j)_j$ *converging to* $\bar{t}$, *then we can restart the algorithm with some* $b \in (0, \bar{t})$. *If the number of critical points of* h_i *in* $[0, \frac{\varepsilon}{\|\tilde{v}\|}]$ *and the number of nonsmooth points of* f_i *in* $x + [0, \frac{\varepsilon}{\|\tilde{v}\|}]\tilde{v}$ *is finite, then after a finite number of restarts, the algorithm will be restarted on an interval* $[0, b]$ *such that* $(0, b]$ *does not contain any nonsmooth or critical points. By c) of this remark, this means that the algorithm stops and a new* ε*-subgradient is found.*

Motivated by the previous remark, we will from now on assume that Algorithm 4.1 stops after finitely many iterations and successfully finds a new ε-subgradient. Based on this, we can compute (acceptable) descent directions for nonsmooth MOPs by iteratively improving the approximation of the ε-subdifferentials, as shown in Algorithm 4.2.

The following theorem shows that Algorithm 4.2 stops after a finite number of iterations and finds a direction $\tilde{v}$ that is an acceptable descent direction (or meets the stopping criterion $\|\tilde{v}\| \leq \delta$, suggesting that x is "almost" Pareto critical).

Algorithm 4.2 Compute descent directions

Given: Current point $x \in \mathbb{R}^n$, tolerances $\varepsilon > 0$, $\delta > 0$, Armijo parameter $c \in (0,1)$.

1: Compute $\xi_1^i \in \partial_\varepsilon f_i(x)$ for all $i \in \{1,\dots,k\}$. Set $W_1 = \{\xi_1^1,\dots,\xi_1^k\}$ and $l = 1$.

2: Compute $v_l = \arg\min_{v \in -\operatorname{conv}(W_l)} \|v\|^2$.

3: If $\|v_l\| \leq \delta$ then stop.

4: Find all objective functions for which there is insufficient descent, i.e., for which (4.10) is violated:

$$I_l = \left\{ j \in \{1,\dots,k\} : f_j\left(x + \frac{\varepsilon}{\|v_l\|} v_l\right) > f_j(x) - c\varepsilon\|v_l\| \right\}.$$

If $I_l = \emptyset$ then stop.

5: For each $j \in I_l$, compute $t \in (0, \frac{\varepsilon}{\|v_l\|}]$ and $\xi_l^j \in \partial f_j(x + tv_l)$ such that

$$\langle v_l, \xi_l^j \rangle > -c\|v_l\|^2$$

via Algorithm 4.1.

6: Set $W_{l+1} = W_l \cup \{\xi_l^j : j \in I_l\}$, $l = l + 1$ and go to step 2.

Theorem 4.3.6. *Algorithm 4.2 stops after a finite number of iterations. In particular, if $\tilde{v}$ is the final element of $(v_l)_l$, then $\|\tilde{v}\| \leq \delta$ or $\tilde{v}$ is an acceptable descent direction, i.e.,*

$$f_i\left(x + \frac{\varepsilon}{\|\tilde{v}\|}\tilde{v}\right) \leq f_i(x) - c\varepsilon\|\tilde{v}\| \quad \forall i \in \{1,\dots,k\}.$$

Proof. Assume that Algorithm 4.2 does not stop, i.e., $(v_l)_l$ is an infinite sequence. Let $l > 1$ and $j \in I_{l-1}$. Since $\xi_{l-1}^j \in W_l$, $-v_{l-1} \in \operatorname{conv}(W_{l-1}) \subseteq \operatorname{conv}(W_l)$ and v_l is the minimal point in step 2, we have

$$\begin{aligned} \|v_l\|^2 &\leq \| - v_{l-1} + s(\xi_{l-1}^j + v_{l-1})\|^2 \\ &= \|v_{l-1}\|^2 - 2s\langle v_{l-1}, \xi_{l-1}^j + v_{l-1}\rangle + s^2\|\xi_{l-1}^j + v_{l-1}\|^2 \\ &= \|v_{l-1}\|^2 - 2s\langle v_{l-1}, \xi_{l-1}^j\rangle - 2s\|v_{l-1}\|^2 + s^2\|\xi_{l-1}^j + v_{l-1}\|^2 \end{aligned} \tag{4.16}$$

for all $s \in [0,1]$. Since $j \in I_{l-1}$ we must have

$$\langle v_{l-1}, \xi_{l-1}^j \rangle > -c\|v_{l-1}\|^2 \tag{4.17}$$

by step 5. By Remark 2.2.19, there has to be a common Lipschitz constant L of all f_i on $\overline{B_\varepsilon}(x)$. By Lemma 2.2.17 and the definition of the ε-subdifferential, we must have $\|\xi\| \leq L$ for all $\xi \in \partial_\varepsilon^\cup f(x)$. So in particular,

$$\|\xi_{l-1}^j + v_{l-1}\| \leq 2L. \tag{4.18}$$

Combining (4.16) with (4.17) and (4.18) yields

$$\begin{aligned} \|v_l\|^2 &< \|v_{l-1}\|^2 + 2sc\|v_{l-1}\|^2 - 2s\|v_{l-1}\|^2 + 4s^2L^2 \\ &= \|v_{l-1}\|^2 - 2s(1-c)\|v_{l-1}\|^2 + 4s^2L^2. \end{aligned}$$

Choose $s := \frac{1-c}{4L^2}\|v_{l-1}\|^2$. Since $1-c \in (0,1)$ and $\|v_{l-1}\| \leq L$, we have $s \in (0,1)$. We obtain

$$\begin{aligned}\|v_l\|^2 &< \|v_{l-1}\|^2 - 2\frac{(1-c)^2}{4L^2}\|v_{l-1}\|^4 + \frac{(1-c)^2}{4L^2}\|v_{l-1}\|^4 \\ &= \left(1 - \frac{(1-c)^2}{4L^2}\|v_{l-1}\|^2\right)\|v_{l-1}\|^2.\end{aligned}$$

Since the algorithm did not stop, it holds $\|v_{l-1}\| > \delta$. It follows that

$$\|v_l\|^2 < \left(1 - \left(\frac{1-c}{2L}\delta\right)^2\right)\|v_{l-1}\|^2.$$

Let $r = 1 - (\frac{1-c}{2L}\delta)^2$. Recall that we have $\delta < \|v_l\| \leq L$ for all $j \in \mathbb{N}$, so $r \in (0,1)$. Additionally, r does not depend on l, so we have

$$\|v_l\|^2 < r\|v_{l-1}\|^2 < r^2\|v_{l-2}\|^2 < \ldots < r^{l-1}\|v_1\|^2 \leq r^{l-1}L^2. \tag{4.19}$$

In particular, there is some l such that $\|v_l\| \leq \delta$, which is a contradiction. □

Remark 4.3.7. *a) The proof of Theorem 4.3.6 shows that Algorithm 4.2 would still work if we only considered a single $j \in I_l$ in step 5. Similarly, a single element from $\partial_\varepsilon f_i(x)$ for any $i \in \{1,\ldots,k\}$ would be sufficient for the initialization of W_1 in step 1. A modification of either step can potentially reduce the executions of Algorithm 4.1 in step 5 if the ε-subdifferentials of multiple objective functions are similar. But both modifications would introduce a bias towards certain objectives into the algorithm, which we want to avoid for this work.*

b) Inequality (4.19) *in the proof of Theorem 4.3.6 shows that we have (at least) linear convergence of the sequence $(\|v_l\|^2)_l$ to zero with a rate of*

$$r = 1 - \left(\frac{1-c}{2L}\delta\right)^2 \in (0,1).$$

Thus, the smaller r, the faster the decrease of $(\|v_l\|)_l$. We can influence r by our choice of $c \in (0,1)$ and $\delta > 0$. More precisely, r becomes smaller the smaller c and the larger δ. In terms of c this is to be expected, since the condition (4.10) *for acceptable descent directions becomes stricter the larger c. In turn, smaller c results in less descent. In terms of δ this is obvious as well, since larger δ means that the stopping criterion in step 3 becomes easier to satisfy. In particular, this means that points in which the algorithm stops via step 3 are potentially "less critical" for larger δ. Thus, for both c and δ, a balance has to be found.*

To highlight the strengths of Algorithm 4.2, we consider an example where classical gradient sampling approaches may struggle to obtain a useful descent direction.

Example 4.3.8. *For $a, b \in \mathbb{R} \setminus \{0\}$ consider the locally Lipschitz problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix}(x_1-1)^2 + (x_2-1)^2 \\ |x_2 - a|x_1|| + bx_2\end{pmatrix}.$$

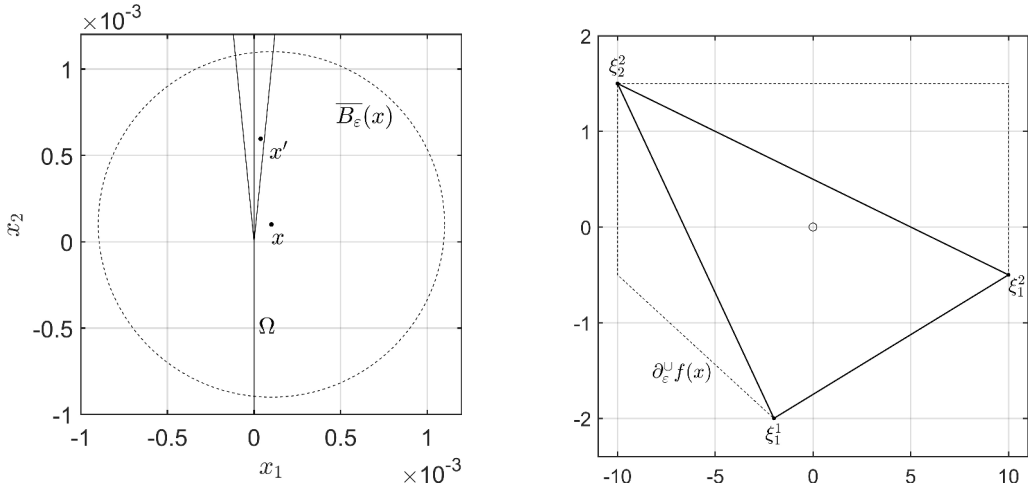


Figure 4.6: **(a)** The nonsmooth set Ω of f (solid), the closed ball $\overline{B_\varepsilon}(x)$ (dashed) for the ε-subdifferential and the points in which subgradients where evaluated in Algorithm 4.2 in Example 4.3.8. **(b)** The ε-subdifferential $\partial_\varepsilon^\cup f(x)$ (dashed) and its approximation (solid) in Algorithm 4.2.

The set of nondifferentiable points is

$$\Omega = (\{0\} \times \mathbb{R}) \cup \{(\lambda, a|\lambda|)^\top : \lambda \in \mathbb{R}\},$$

seperating $\mathbb{R}^2$ *into four "smooth areas" (cf. Figure 4.6(a)). For large* $a > 0$*, the two areas above the graph of* $\lambda \mapsto a|\lambda|$ *become small. Thus, due to the random sampling in gradient sampling methods, it is unlikely that these methods sample gradients in these areas. But as we will see in the following, the behavior of* f *in these areas is crucial for approximating the* ε*-subdifferential in this example.*

Let $a = 10$*,* $b = 0.5$*,* $\varepsilon = 10^{-3}$ *and* $x = (10^{-4}, 10^{-4})^\top$*. Then* $(0,0)^\top$ *is the (unique) minimal point of* f_2 *and*

$$\begin{aligned}\partial_\varepsilon f_2(x) &= \operatorname{conv}\left(\left\{\begin{pmatrix}-a\\ b-1\end{pmatrix}, \begin{pmatrix}a\\ b+1\end{pmatrix}, \begin{pmatrix}a\\ b-1\end{pmatrix}, \begin{pmatrix}-a\\ b+1\end{pmatrix}\right\}\right)\\ &= \operatorname{conv}\left(\left\{\begin{pmatrix}-10\\ -0.5\end{pmatrix}, \begin{pmatrix}10\\ 1.5\end{pmatrix}, \begin{pmatrix}10\\ -0.5\end{pmatrix}, \begin{pmatrix}-10\\ 1.5\end{pmatrix}\right\}\right).\end{aligned}$$

In particular, $0 \in \partial_\varepsilon f_2(x)$*, so the descent direction from* (4.7) *based on the exact* ε*-subdifferentials is zero. When applying Algorithm 4.2 at* x *with* $\delta = 10^{-3}$*, the method stops in the second iteration as*

$$\tilde{v} = v_2 \approx (0.118, 1.185)^\top \cdot 10^{-9},$$

i.e., $\|\tilde{v}\| \approx 1.191 \cdot 10^{-11} < \delta$*. Thus,* x *is correctly identified as "almost" Pareto critical. The final approximation* W_2 *of* $\partial_\varepsilon^\cup f(x)$ *is*

$$W_2 = \{\xi_1^1, \xi_1^2, \xi_2^2\} = \left\{\begin{pmatrix}-1.9998\\ -1.9998\end{pmatrix}\begin{pmatrix}10\\ -0.5\end{pmatrix}, \begin{pmatrix}-10\\ 1.5\end{pmatrix}\right\},$$

as shown in Figure 4.6(b). The first two elements ξ_1^1 *and* ξ_1^2 *are the gradients of* f_1 *and* f_2 *in* x *from the initialization of Algorithm 4.2, and the final element* ξ_2^2 *is the*

gradient of f_2 in $x' = x + tv = (0.038, 0.596)^\top \cdot 10^{-3} \in \overline{B_\varepsilon}(x)$ (computed during the first iteration). Figure 4.6(a) shows that ξ_2^2 may be difficult to obtain via gradient sampling approaches, but Figure 4.6(b) shows that it is crucial for the approximation of the ε-subdifferential.

4.4 The algorithm

In the previous section, we derived a method (Algorithm 4.2) that is able to compute acceptable descent directions (cf. (4.10)) in arbitrary points $x \in \mathbb{R}^n$. Based on this method, it is now straight-forward to construct the descent method for locally Lipschitz MOPs shown in Algorithm 4.3. For the step length, we use a slightly

Algorithm 4.3 Descent method for locally Lipschitz MOPs

Given: Initial point $x^1 \in \mathbb{R}^n$, tolerances $\varepsilon > 0$, $\delta > 0$, Armijo parameter $c \in (0, 1)$, initial step length $t_0 > 0$.

1: Initialize $j = 1$.
2: Compute a descent direction v^j via Algorithm 4.2.
3: If $\|v^j\| \leq \delta$ then stop.
4: Compute

$$\bar{s} = \inf(\{s \in \mathbb{N} \cup \{0\} : f_i(x^j + 2^{-s} t_0 v^j) \leq f_i(x^j) - 2^{-s} t_0 c \|v^j\|^2 \ \forall i \in \{1, \dots, k\}\})$$

and set $\bar{t} = \max(\{2^{-\bar{s}} t_0, \frac{\varepsilon}{\|v^j\|}\})$.

5: Set $x^{j+1} = x^j + \bar{t} v^j$, $j = j + 1$ and go to step 2.

modified Armijo step length, which is similar to the one used in the smooth case in (2.24).

By construction, Algorithm 4.3 stops if $\|v^j\| \leq \delta$ in step 3, i.e., if there is some $v \in -\partial_\varepsilon^\cup f(x)$ with $\|v\| \leq \delta$. Thus, in terms of optimality, we can only expect the result of the algorithm to be "almost" Pareto critical in the following sense.

Definition 4.4.1. *Let $x \in \mathbb{R}^n$, $\varepsilon > 0$ and $\delta > 0$. Then x is called* (ε, δ)-critical *if*

$$\min_{v \in \partial_\varepsilon^\cup f(x)} \|v\| \leq \delta.$$

Clearly, (ε, δ)-criticality is a necessary condition for Pareto criticality (cf. Definition 2.2.23) and equivalent to Pareto criticality for $\varepsilon = \delta = 0$. The following theorem shows that Algorithm 4.3 always stops in an (ε, δ)-critical point.

Theorem 4.4.2. *Let $(x^j)_j$ be the sequence generated by Algorithm 4.3. Then either $(f_i(x^j))_j$ is unbounded below for all $i \in \{1, \dots, k\}$, or $(x^j)_j$ is finite with the last element being (ε, δ)-critical.*

Proof. Assume that $(x^j)_j$ is infinite. Then $\|v^j\| > \delta$ for all $j \in \mathbb{N}$. Let $j \in \mathbb{N}$. If $\bar{t} = 2^{-\bar{s}} t_0 \geq \frac{\varepsilon}{\|v^j\|}$ in step 4 of Algorithm 4.3, then

$$\begin{aligned} f_i(x^j + \bar{t} v^j) - f_i(x^j) &= f_i(x^j + 2^{-\bar{s}} t_0 v^j) - f_i(x^j) \\ &\leq -2^{-\bar{s}} t_0 c \|v^j\|^2 \leq -c\varepsilon \|v^j\| < -c\varepsilon\delta < 0 \end{aligned} \tag{4.20}$$

for all $i \in \{1, \dots, k\}$. If instead $\bar{t} = \frac{\varepsilon}{\|v^j\|}$ in step 4, then the same inequality holds due to Theorem 4.3.6. This implies that in each iteration, the value of each objective function is decreased by at least $c\varepsilon\delta$, which is constant with respect to the iteration j. Thus, $(f_i(x^j))_j$ must be unbounded below for each $i \in \{1, \dots, k\}$.
Now assume that $(x^j)_j$ is finite, with $\bar{x}$ and $\bar{v}$ being the final elements of $(x^j)_j$ and $(v^j)_j$, respectively. Since the algorithm stopped, we must have $\|\bar{v}\| \leq \delta$. From the application of Algorithm 4.2 in step 2, we know that there must be some $\overline{W} \subseteq \partial_\varepsilon^\cup f(\bar{x})$ such that $\bar{v} = \arg\min_{v \in -\operatorname{conv}(\overline{W})} \|v\|^2$. This implies

$$\min_{v \in -\partial_\varepsilon^\cup f(\bar{x})} \|v\| \leq \min_{v \in -\operatorname{conv}(\overline{W})} \|v\| = \|\bar{v}\| \leq \delta.$$

□

Remark 4.4.3. *a) If we compare the proof of Theorem 4.4.2 to the proof of convergence of the steepest descent method in the smooth case from [FS00], Theorem 1, we see that the proof here is a lot shorter. The reason for this is the fact that usage of the ε-subdifferential gives us the lower bound $c\varepsilon\delta$ for the decrease in each objective function in each step (cf. (4.20)). The disadvantage is that the assertion of our convergence result is weaker, since (even in the smooth case) a point is already guaranteed to be (ε, δ)-critical when its distance to the Pareto critical set is less or equal ε.*

b) In theory, we can only expect the result $\bar{x}$ of Algorithm 4.3 to be Pareto critical at best. But note that our line search in step 4 enforces a descent in the sequence $(f(x^j))_j$. Thus, Pareto critical points which are not actually Pareto optimal are "less attractive" for our algorithm.

To obtain a stronger convergence result than Theorem 4.4.2, we consider the following lemma.

Lemma 4.4.4. *Let $(\bar{x}^j)_j \in \mathbb{R}^n$ be a sequence with $\lim_{j\to\infty} \bar{x}^j = \bar{x} \in \mathbb{R}^n$. Let $(\varepsilon_j)_j \in \mathbb{R}^{>0}$ and $(\delta_j)_j \in \mathbb{R}^{>0}$ with $\lim_{j\to\infty} \varepsilon_j = 0$ and $\lim_{j\to\infty} \delta_j = 0$. If $\bar{x}^j$ is $(\varepsilon_j, \delta_j)$-critical for all $j \in \mathbb{N}$, then $\bar{x}$ is Pareto critical.*

Proof. Assume that $\bar{x}$ is not Pareto critical, i.e., $0 \notin \partial^\cup f(\bar{x})$. Let

$$v := \arg\min_{\xi \in -\partial^\cup f(\bar{x})} \|\xi\|^2 \neq 0.$$

Then by Lemma 2.2.1, $\langle v, \xi \rangle \leq -\|v\|^2$ for all $\xi \in \partial^\cup f(\bar{x})$, so

$$\partial f_i(\bar{x}) \subseteq \partial^\cup f(\bar{x}) \subseteq \left\{ \xi \in \mathbb{R}^n : \langle v, \xi \rangle < -\frac{\|v\|^2}{2} \right\} =: V \quad \forall i \in \{1, \dots, k\}.$$

Note that V is open, convex and $\min_{\xi \in \overline{V}} \|\xi\| = \frac{\|v\|}{2}$. By upper semicontinuity of ∂f_i, $i \in \{1, \dots, k\}$, there has to be some open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that

$$\partial f_i(y) \subseteq V \quad \forall y \in U, i \in \{1, \dots, k\}.$$

In particular, by convexity of V, we have

$$\partial^\cup f(y) \subseteq V \quad \forall y \in U. \tag{4.21}$$

Now let $N \in \mathbb{N}$ such that $\overline{B_{\varepsilon_j}}(\bar{x}^j) \subseteq U$ and $\delta_j < \frac{\|v\|}{2}$ for all $j > N$. Then by (4.21) we have $\partial^{\cup}_{\varepsilon_j} f(\bar{x}^j) \subseteq V \subseteq \overline{V}$ for all $j > N$, so

$$\min_{v \in \partial^{\cup}_{\varepsilon_j} f(\bar{x}^j)} \|v\| \geq \min_{\xi \in \overline{V}} \|\xi\| = \frac{\|v\|}{2} > \delta_j \quad \forall j > N,$$

which is a contradiction to $\bar{x}^j$ being $(\varepsilon_j, \delta_j)$-critical. □

By the previous lemma, if we iteratively decrease ε and δ to zero for Algorithm 4.3, we obtain a point which is actually Pareto critical. This is done in Algorithm 4.4, for which we assume that Algorithm 4.3 always stops after a finite number of iterations (cf. Theorem 4.4.2).

Algorithm 4.4 ε-δ-decreasing nonsmooth descent method

Given: Initial point $x^1 \in \mathbb{R}^n$, tolerance sequences $(\varepsilon_j)_j \in \mathbb{R}^{>0}$, $(\delta_j)_j \in \mathbb{R}^{>0}$ converging to 0, Armijo parameter $c \in (0,1)$, initial step length $t_0 > 0$.

1: Initialize $\bar{x}^0 = x^1$.
2: **for** $j = 0, 1, \ldots, \infty$ **do**
3: Apply Algorithm 4.3 with initial point $\bar{x}^j$ and tolerances $\varepsilon = \varepsilon_{j+1}$, $\delta = \delta_{j+1}$. Let $\bar{x}^{j+1}$ be the final element in the generated sequence.
4: **end for**

From Lemma 4.4.4, we immediately get the following corollary.

Corollary 4.4.5. *Let $(\bar{x}^j)_j$ be a sequence generated by Algorithm 4.4. Then all accumulation points of $(\bar{x}^j)_j$ are Pareto critical.*

To additionally obtain existence of accumulation points in the previous corollary, we could employ standard assumptions like boundedness (and thus compactness) of the sublevel set $\{x \in \mathbb{R}^n : f(x) \leq f(\bar{x}^1)\}$. Finally, in practice, we will use finite sequences $(\varepsilon_j)_{j \in \{1,\ldots,N\}}$ and $(\delta_j)_{j \in \{1,\ldots,N\}}$ for the input of Algorithm 4.4 and only loop over $j \in \{0, 1, \ldots, N\}$ in step 2. This will cause the final point $\bar{x}^{N+1}$ to be $(\varepsilon_N, \delta_N)$-critical.

4.5 Numerical experiments

In this section, we will investigate the practical behavior of our nonsmooth descent method (Algorithm 4.3) and its extension (Algorithm 4.4). We will begin by visualizing and discussing its typical behavior before comparing its performance to the multiobjective proximal bundle method (MPB) from Section 2.3.2. Finally, we will combine our method with the subdivision method (as in Section 2.3.3) to approximate entire Pareto sets of nonsmooth MOPs.

4.5.1 Typical behavior

In areas where all objective functions are differentiable the behavior of our method is almost identical to the behavior of the steepest descent method from Section 2.3.2.

They are not completely identical since ε-subdifferentials, unlike Clarke subdifferentials, do not reduce to the classical gradients when the objectives are differentiable. More precisely, we have $\partial_\varepsilon f_i(x) = \nabla f_i(\overline{B_\varepsilon}(x)) \ni \nabla f_i(x)$, so if

$$\max_{y \in \overline{B_\varepsilon}(x)} \|\nabla f_i(x) - \nabla f_i(y)\|$$

is large for some $i \in \{1, \dots, k\}$, then the methods may behave differently. But since ε is typically chosen to be small, we will ignore this difference here and focus on the behavior with respect to the nondifferentiable set.

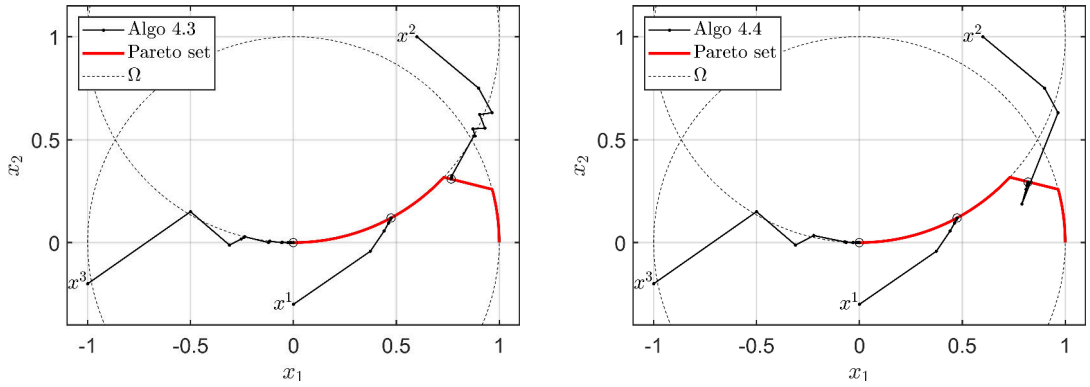
To visualize the typical behavior of Algorithm 4.3, we consider the problem $\min_{x \in \mathbb{R}^2} f(x)$ for the locally Lipschitz function

$$f(x) := \begin{pmatrix} \max\{x_1^2 + (x_2 - 1)^2 + x_2 - 1, -x_1^2 - (x_2 - 1)^2 + x_2 + 1\} \\ -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1| \end{pmatrix} \tag{4.22}$$

from [MKW14] (combining *Crescent* from [Kiw85] and *Mifflin 2* from [MN92]). It is easy to see that the sets of nondifferentiable points of f_1 and f_2 are $S^1 + (0,1)^\top$ and S^1, respectively. Thus, f is nondifferentiable in $\Omega = S^1 \cup (S^1 + (0,1)^\top)$. We consider the initial points

$$x^1 = (0, -0.3)^\top, \quad x^2 = (0.6, 1)^\top, \quad x^3 = (-1, -0.2)^\top,$$

with the fixed parameters $\varepsilon = 10^{-3}$, $\delta = 10^{-3}$, $c = 0.25$ and $t_0 = 1$. Figure 4.7(a) shows the sequences generated by Algorithm 4.3, the nondifferentiable set Ω and the Pareto critical set of the MOP (which can be derived by hand). We will briefly



(a) **(b)**

Figure 4.7: The results of Algorithm 4.3 (a) and Algorithm 4.4 (b) for three different initial points, the Pareto set of (4.22) and the set of nondifferentiable points Ω.

discuss the behavior for each initial point:

- For x^1, the sequence moves through the smooth area like the steepest descent method until a point is reached with a distance to Ω that is less or equal ε. In that point, the method computes an approximation of the ε-subdifferentials using multiple ε-subgradients. Since this part of Ω is Pareto critical, no (acceptable) descent direction is found and the algorithm stops with the final point being (ε, δ)-critical.

Table 4.1: Test problems (using objectives from [MKW14])

Nr.	f_i	Area	Nr.	f_i	Area
1.	CB3, DEM	$[-3,3]^2$	10.	QL, LQ	$[-3,3]^2$
2.	CB3, QL	$[-3,3]^2$	11.	QL, Mifflin 1	$[-3,3]^2$
3.	CB3, LQ	$[0.5,1.5]^2$	12.	QL, Wolfe	$[-3,3]^2$
4.	CB3, Mifflin 1	$[-3,3]^2$	13.	LQ, Mifflin 1	$[0.5,1.5]\times[-0.5,1]$
5.	CB3, Wolfe	$[-3,3]^2$	14.	LQ, Wolfe	$[-3,3]^2$
6.	DEM, QL	$[-3,3]^2$	15.	Mifflin 1, Wolfe	$[-3,3]^2$
7.	DEM, LQ	$[-3,3]^2$	16.	Crescent, Mifflin 2	$[-0.5,1.5]^2$
8.	DEM, Mifflin 1	$[-3,3]^2$	17.	Mifflin 2, WF	$[-3,3]^2$
9.	DEM, Wolfe	$[-3,3]^2$	18.	Mifflin 2, SPIRAL	$[-3,3]^2$

- For x^2, the sequence starts zig-zagging around the non-critical part of Ω. The reason for this behavior is the fact that in this area, the information from both "sides" of the nondifferentiable set is required to obtain a significant reduction of the objective functions. Only when the distance of the sequence to Ω is less or equal ε, the algorithm actually notices the nonsmoothness and computes a descent direction which breaks the zig-zagging motion.
- For x^3, the sequence performs a similar zig-zagging motion to the previous case. The difference is that this time, this motion does not get broken when the sequence gets close to Ω. Thus, it moves along Ω until an (ε,δ)-critical point is found.

In the case of x^2, the zig-zagging motion is broken as soon as the distance of the descent sequence to Ω is less or equal ε. Thus, choosing larger ε would stop the zig-zagging earlier. This is another reason for using Algorithm 4.3 with dynamically decreasing ε instead of a fixed one, i.e., another reason for using Algorithm 4.4. For example, Figure 4.7(b) shows the result of Algorithm 4.4 for $(\varepsilon_j)_j = (\varepsilon_1,\varepsilon_2) = (10^{-1},10^{-3})$ and constant $\delta = 10^{-3}$. (More precisely, it shows all the descent sequences produced by calls of Algorithm 4.3 during Algorithm 4.4.) In this case, we see that the zig-zagging for x^2 is completely avoided.

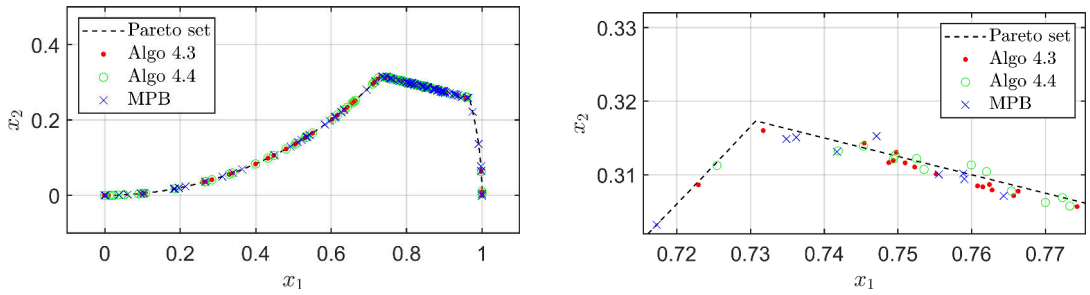
4.5.2 Comparison to the MPB

We will now compare the performance of Algorithms 4.3 and 4.4 to the multiobjective proximal bundle method by Mäkelä, Karmitsa and Wilppu from [MKW14] (see also [Mäk03]), which was discussed in Section 2.3.2. The MPB currently seems to be the most efficient (deterministic) solver for locally Lipschitz MOPs. (Bundle methods are also regarded as the most effective and reliable method for the nonsmooth, single-objective case, cf. [BKM14].) As test problems, we consider the 18 MOPs in Table 4.1. They were created via combination of classical test problems from nonsmooth, single-objective optimization (cf. [MKW14], Appendix A, and [MKM18]). Problems 1 to 15 are convex and problems 16 to 18 are nonconvex. Due to their piecewise smooth structure, it is possible to differentiate each objective vector by hand to obtain exact subgradients. For each test problem, we choose 100 starting points on a 10×10 grid in the corresponding area given in Table 4.1.

For the MPB, we use the Fortran implementation from [Mäk03]. For the stopping criterion, the second component β of the solution (v,β) of the direction finding

problem (2.27) is used. The method stops when $\beta > -2\varepsilon_s$ for some $\varepsilon_s > 0$. By our discussion of duality in Remark 2.3.2 b), problem (2.27) is related to problem (4.9) in our algorithms. In particular, β is related to $\max_{\xi \in -\operatorname{conv}(W)} \langle \xi, \tilde{v} \rangle$, which is bounded by $-\|\tilde{v}\|^2$. To obtain similar stopping conditions for all methods, this motivates us to choose $\delta = \sqrt{\varepsilon_s}$ in our algorithms. More precisely, we choose $\varepsilon_s = 10^{-6}$ for the MPB (and keep the default values for the other parameters) and $\delta = 10^{-3}$ for our algorithms. (We will verify below in Figure 4.8 that this indeed leads to results of similar approximation quality.)

For the implementation of our methods we use MATLAB. The optimization problem in step 2 of Algorithm 4.3 is solved via the function `quadprog`. As parameters for Algorithm 4.3, we choose $\varepsilon = 10^{-3}$, $\delta = 10^{-3}$, $c = 0.25$ and $t_0 = \max(\{\|v^j\|^{-1}, 1\})$ (i.e., the initial step size t_0 is chosen depending on the norm of the descent direction v^j in the current interation). For Algorithm 4.4, we replace the fixed ε by $(\varepsilon_j)_j = (\varepsilon_1, \varepsilon_2, \varepsilon_3) = (10^{-1}, 10^{-2}, 10^{-3})$ and keep δ constant. By this choice of parameters, all three methods produce results of equal approximation quality. For example, Figure 4.8 shows the result of Algorithm 4.3, Algorithm 4.4 and the MPB for problem 16 from Table 4.1 (which we also previously considered in (4.22)).



(a) **(b)**

Figure 4.8: **(a)** Results of Algorithm 4.3, Algorithm 4.4 and the MPB for problem 16 from Table 4.1. **(b)** Same as (a) but zoomed in closer.

To compare the performance of the methods, we count the number of evaluations of objectives f_i, the number of evaluations of subgradients $\xi \in \partial f_i$ and the number of iterations (i.e., descent steps) needed. (For example, by this method of counting, one evaluation of the objective vector f accounts for k evaluations of objectives.) By construction, the MPB always evaluates all objectives and subgradients for all objectives in a point, so the value for the objectives and the subgradients are equal for this method. In case of the number of iterations of Algorithm 4.4, we sum up all iterations of each application of Algorithm 4.3 in step 3. The results are shown in Table 4.2 and are discussed in the following.

- **Function evaluations:** In our methods, function evaluations are used to compute the descent directions via Algorithm 4.2 and to compute the step length in step 4 of Algorithm 4.3. When comparing the MPB to our methods, we see that the MPB needs far less evaluations. One possible reason for this discrepancy is the simplicity of the line search that we use, which is basically the default Armijo line search. Since there are more advanced line search methods in the smooth case (cf. [NW06], Chapter 3), it is likely that the efficiency of this part of our methods can be improved by generalizing some of the concepts from the smooth case. When comparing our methods to each

Table 4.2: Performance of MPB, Algorithm 4.3 and Algorithm 4.4 for the test problems in Table 4.1 for 100 starting points. The last two rows contain the rounded average over the respective column and the percentage compared to the MPB result

	$\#f_i$			$\#\partial f_i$			# Iter		
Nr.	MPB	Algo. 4.3	Algo. 4.4	MPB	Algo. 4.3	Algo. 4.4	MPB	Algo. 4.3	Algo. 4.4
1.	**1846**	6924	7801	1846	**1102**	1751	794	**492**	695
2.	**2652**	14688	12263	2652	**1906**	2351	1214	**842**	914
3.	**888**	5625	6447	**888**	921	1534	**344**	448	662
4.	**6832**	103826	17664	6832	11774	**3415**	2385	4644	**1242**
5.	**3292**	30457	16877	3292	3479	**3037**	1545	1616	**1161**
6.	**7634**	8357	8684	7634	**1209**	1802	1089	**552**	736
7.	**1806**	8736	8483	1806	**1307**	1832	775	**595**	739
8.	8410	**8283**	8620	8410	**1318**	1914	1481	**582**	759
9.	**1894**	8201	8794	1894	**1194**	1805	826	**536**	732
10.	11686	**6799**	7201	11686	**1101**	1722	1275	**543**	733
11.	30436	52096	**17594**	30436	6311	**3189**	2385	2442	**1206**
12.	30496	15146	**12446**	30496	**1992**	2401	2015	**967**	1010
13.	**3090**	36570	9513	3090	4958	**2247**	1423	1692	**787**
14.	**5086**	95303	12227	5086	9524	**2571**	2113	4379	**921**
15.	**4104**	85936	15669	4104	9329	**3124**	1647	3963	**1125**
16.	**2628**	20372	11094	2628	2596	**2400**	1063	1194	**947**
17.	**3208**	6168	5262	3208	**1064**	1459	1170	**520**	652
18.	**15506**	166707	31528	15506	16676	**6902**	3595	8291	**2412**
Avg.	**7860.8**	37788.6	12120.4	7860.8	4320.1	**2525.3**	1507.7	1905.4	**968.5**
	100%	480.7%	154.2%	100%	55.0%	32.1%	100%	126.4%	64.2%

other, we see that Algorithm 4.4 is much more efficient than Algorithm 4.3 when the overall number of evaluations is high (e.g., problem 14), and slightly less efficient when the number of evaluations is low (e.g., problem 9). The reason for this is the fact that for easier problems (i.e., where the number of evaluations is low), some of the iterations of Algorithm 4.4 will be redundant, because the $(\varepsilon_{i-1}, \delta)$-critical point of the previous iterations is already (ε_i, δ)-critical. When summing up the evaluations for all 18 problems, we see that Algorithm 4.4 is superior to Algorithm 4.3.

- **Subgradient evaluations:** When comparing the number of subgradient evaluations, we see that the MPB is superior on problem 3, but inferior to our methods in all other cases. When comparing our methods to each other, we see the same pattern as for the function evaluations: Algorithm 4.4 is a lot more efficient for complex problems, and slightly less efficient for easier problems.

- **Iterations:** For the number of iterations, we see the same pattern as for the subgradient evaluations, with Algorithm 4.4 again being the most efficient on average. It is important to note that for the MPB, this number includes so-called *null steps*, which are iterations where the bundle (i.e., the set of subgradients computed during all previous iterations) is enriched, but the current point in the descent sequence does not get updated.

For the set of test problems we considered here, this leads us to the conclusion that the MPB is superior to our methods in terms of function evaluations, but inferior in terms of subgradient evaluations and number of iterations. When comparing our methods with each other, Algorithm 4.3 is slightly more efficient than Algorithm 4.4 for simple problems, but a lot less efficient for more complex problems.

Finally, it is noteworthy that one might find Algorithm 4.3 and Algorithm 4.4 easier to implement than the MPB, due to their simplicity and similarities to classical descent methods from smooth optimization. In particular, each iteration of

Algorithm 4.3 is independent from any other iteration, whereas each iteration of the MPB relies on information computed in previous iterations.

4.5.3 Combination with the subdivision algorithm

So far we have constructed two algorithms (Algorithm 4.3 and Algorithm 4.4) that take an initial point $x^1 \in \mathbb{R}^n$ and compute one (almost) Pareto critical point of our nonsmooth MOP. To actually *solve* the MOP, i.e., to compute the entire Pareto (critical) set, the easiest approach would be to just apply our methods to many different initial points and hope that the resulting set of points is a good approximation of the Pareto set. As an example, Figure 4.8 shows the results of our two algorithms (and the MPB) applied the MOP (4.22) for 100 initial points on a quadratic 10×10 grid. While it is possible to get an idea of the Pareto set by looking at the individual results of the algorithms, we did not obtain satisfying approximations in a set-wise sense, as the distribution of the points is uneven. The reason for this is the fact that some parts of the Pareto set are more attractive to our methods than others. The problem is that there is no way of knowing a priori which initial points we have to choose to obtain an even discretization of the Pareto set.

For the steepest descent method in the smooth case (cf. Section 2.3.2), this problem can be solved via combination with the subdivision algorithm (Algorithm 2.1), as presented in Section 2.3.3. To be able to do the same in the nonsmooth case, we have to write our methods as discrete dynamical systems

$$x^{j+1} = g(x^j), \quad j \in \mathbb{N},$$

for a map $g : \mathbb{R}^n \to \mathbb{R}^n$ and some initial point $x^1 \in \mathbb{R}^n$. For Algorithm 4.4 (and the MPB) this is not directly possible, since each iteration depends on information from previous iterations (i.e., the dynamical system would not be *autonomous*). For Algorithm 4.3, it can be done by defining

$$g(x) := x + \bar{t}(x)\tilde{v}(x), \tag{4.23}$$

where $\tilde{v}$ is the result of Algorithm 4.2 and $\bar{t}$ is the step length from step 4 of Algorithm 4.3 in x. If we apply Algorithm 2.1 to the dynamical system induced by (4.23), we know that by Theorem 2.3.6 a), the Pareto critical set P_c is contained in the result Q_∞ of the subdivision method.

Since Algorithm 4.3 behaves like the steepest descent method in the smooth case (for small $\varepsilon > 0$), we generally cannot expect to have a better convergence behavior of the subdivision method in the nonsmooth case. In particular, with the same reasoning as in the smooth case, we cannot expect to have $P_c = Q_\infty$, since g is not a homeomorphism. In the nonsmooth case, g is even more irregular, since the descent direction $\tilde{v} : \mathbb{R}^n \to \mathbb{R}^n$ is inherently discontinuous close to the set of nondifferentiable points of f. The problems caused by this are highlighted in the following example.

Example 4.5.1. *Consider the problem*

$$\min_{x \in \mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} \lambda x_1^2 + |x_2| \\ (x_1 - 1)^2 + (x_2 - 1)^2 \end{pmatrix} \tag{4.24}$$

for $\lambda > 0$. It is easy to see that f is locally Lipschitz continuous. It is differentiable everywhere except on the x_1-axis $\mathbb{R} \times \{0\} = \Omega$. For this example we choose $\lambda = 0.5$. The Pareto critical set P_c of (4.24) can be computed by hand and is shown in Figure 4.9(b). Since f is convex, it coincides with the (weak) Pareto set (cf. Remark 2.2.24).

Let $\varepsilon = 10^{-3}$, $\delta = 10^{-3}$, $c = 0.25$ and $t_0 = 1$ be the parameters for the nonsmooth descent method. Figure 4.9(a) shows the result $\mathcal{B}_8$ of applying 8 iterations of the subdivision algorithm to the function g from (4.23) for the initial box $Q = [-0.25, 1.25]^2$ and 4 sample points per box (on an even 2×2 grid). The resulting radius of boxes is $r = 1.5 \cdot 2^{-9} \approx 0.2930 \cdot 10^{-2}$. As expected, $\mathcal{B}_8$ is a covering of P_c. But unfor-

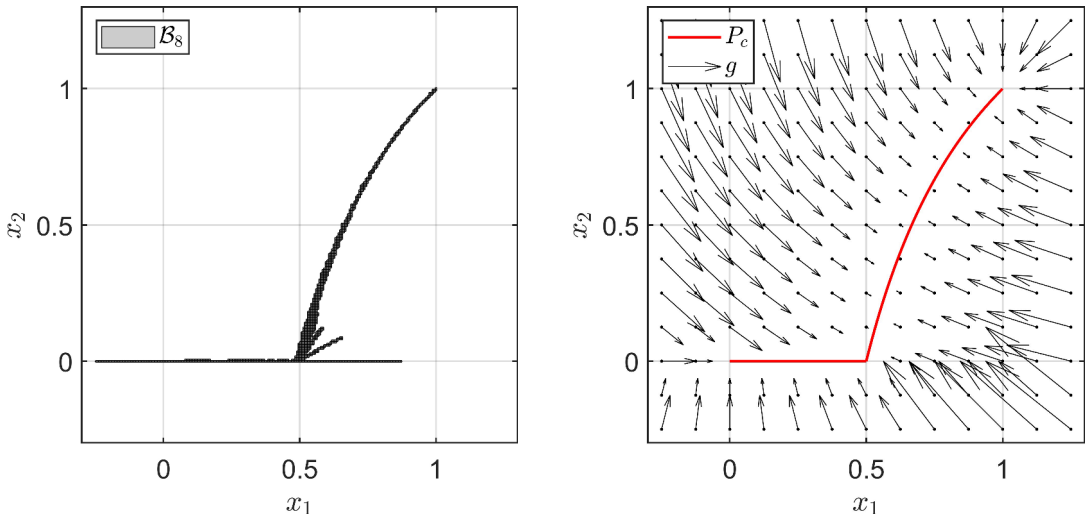


Figure 4.9: **(a)** Result after 8 iterations of Algorithm 2.1 in Example 4.5.1. **(b)** Visualization of the dynamical system g from (4.23). (The length of each vector is halved for better visibility.)

tunately, it contains significantly more than just P_c: A large part of the x_1-axis is unnecessarily covered and there are two unnecessary "branches" of boxes emerging from $(0.5, 0)^\top$. This can be explained by considering the visualization of g in Figure 4.9(b). All points $x \in Q$ with $x_2 < -\varepsilon < 0$ are strongly attracted by the x_1-axis and only slightly attracted by the actual Pareto critical part on the x_1-axis. Thus, in the first iterations of the subdivision algorithm, if we have a box B that has a nonempty intersection with the x_1-axis and the lower half-space, then many points in B will get mapped again into B such that B does not get removed in the selection step. The additional boxes on the x_1-axis then induce the two unecessary branches above it: The lower branch is the set of boxes that contain the image of the boxes to the right of P_c, and the upper branch is the set of boxes that contain the image of the boxes from the lower branch.

Roughly speaking, the reason for the convergence issues in the previous example is the fact that the behavior of the dynamical system on the set of nondifferentiable points Ω is required for the descent method to reach the Pareto critical set. Since we discretize boxes with a set of test points to approximate their image under the dynamical system (cf. Remark 2.3.7), this behavior can only be captured if one of these test points has a distance of ε or less from Ω. Since Ω is a null set (cf. Theorem 2.2.14) and we typically only want to use a few test points per box, this practically

only happens when the box radius r is of the same magnitude as ε (or smaller than ε). In theory, one could apply iterations of the subdivision algorithm until this is the case. But since we can only approximate the actual Pareto critical with an accuracy of ε (cf. Theorem 4.4.2), it would be highly inefficient to cover it using boxes with such a small radius.

Fortunately, there is a more efficient way of solving this issue. To this end, we can artificially increase the area in which the subdivision algorithm incorporates the behavior of g on Ω by applying multiple descent steps (i.e., multiple iterations of Algorithm 4.3) at once in the dynamical system g. More formally, if we define

$$g^N : \mathbb{R}^n \to \mathbb{R}^n \tag{4.25}$$

to be the function which applies $N \in \mathbb{N}$ iterations of Algorithm 4.3 starting at $x \in \mathbb{R}^n$, then it is sufficient for any of the N iterates $x, g(x), g^2(x), \ldots, g^{N-1}(x)$ to have a distance less or equal ε to Ω for the dynamical system to notice the nonsmoothness. Furthermore, since the Pareto critical set is still contained in the global attractor of g^N, we obtain the same theoretical convergence result if we replace the original g from (4.23) by g^N. The behavior of the resulting subdivision algorithm is shown in the following example.

Example 4.5.2. *Consider again problem* (4.24). *Figure 4.10(a) shows the result of applying the subdivision algorithm as in Example 4.5.1, except that we replace g by g^N for $N = 5$, i.e., 5 descent steps are applied in each evaluation of the dynamical system. If we compare the result with the actual Pareto critical set P_c*

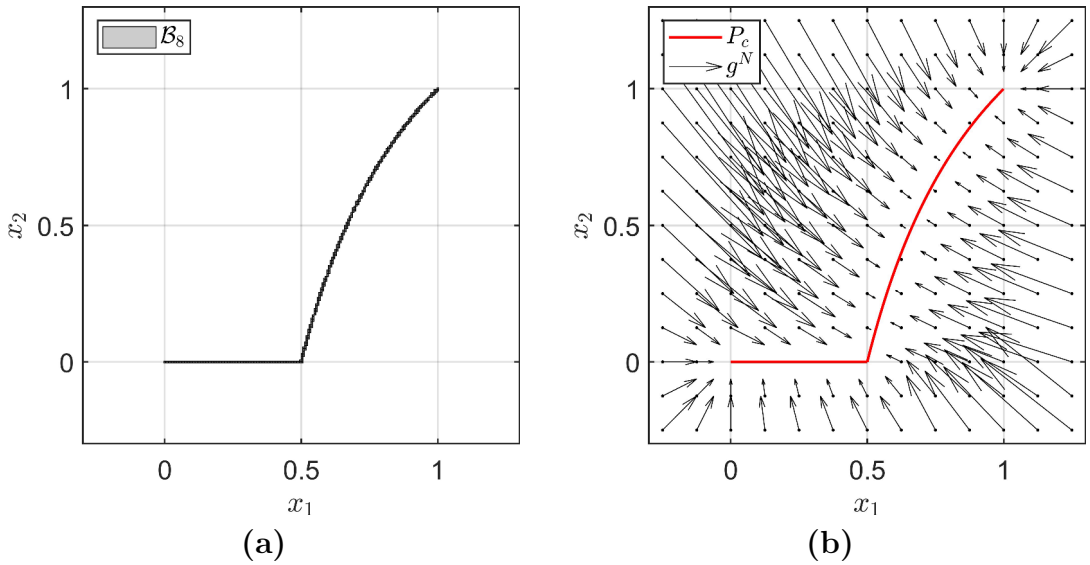
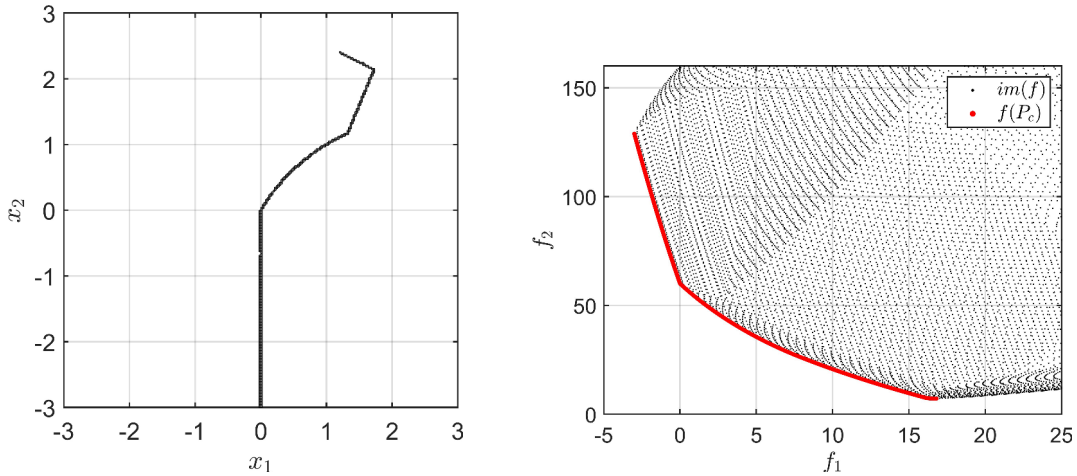


Figure 4.10: Same as Figure 4.9, except that g from (4.23) is replaced by g^N from (4.25) for $N = 5$.

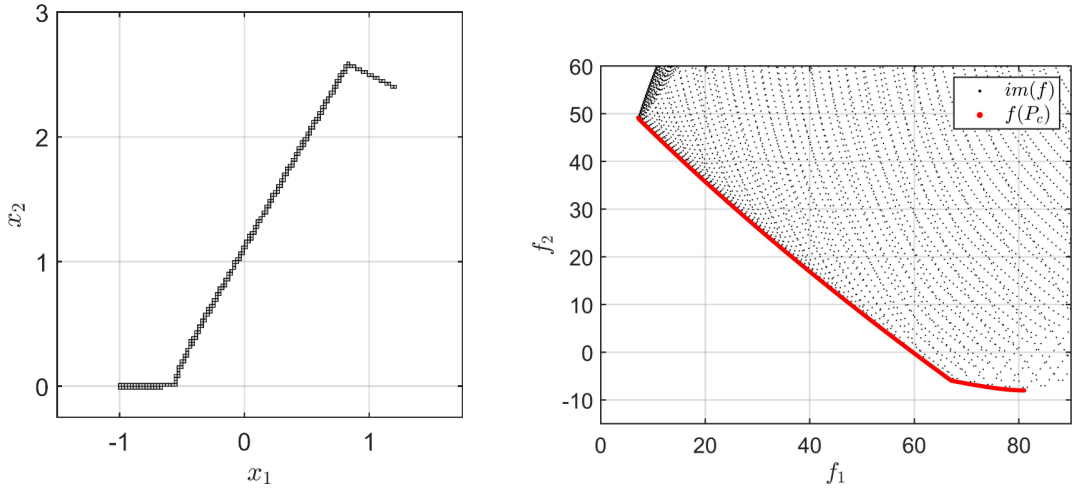
from Figure 4.10(b), we see that this did not only improve the convergence behavior with respect to the nonsmoothness, but also the behavior in the smooth areas above the x_1-axis, due to the sped up dynamic of the system. The reason for this improved convergence behavior can be seen by considering the visualization of g^N in Figure 4.10(b). Compared to Figure 4.9(b), points are more attracted by P_c, especially points which lie below the x_1-axis.

To conclude this section, we apply the subdivision algorithm (with the dynamical system (4.25)) to some of the problems from Table 4.1. For the descent method, we use the same parameters as in Example 4.5.1, and for the initial boxes Q, we use the areas given in Table 4.1. The results are shown in Figure 4.11, 4.12 and 4.13. For the approximation of the image $f(P_c)$ of P_c, we evaluated f in the image points of g that were computed in the final selection step (cf. Algorithm 2.1) and also lie in the final collection of boxes $\mathcal{B}_8$. For all three examples, we obtain a tight covering of the Pareto critical set.



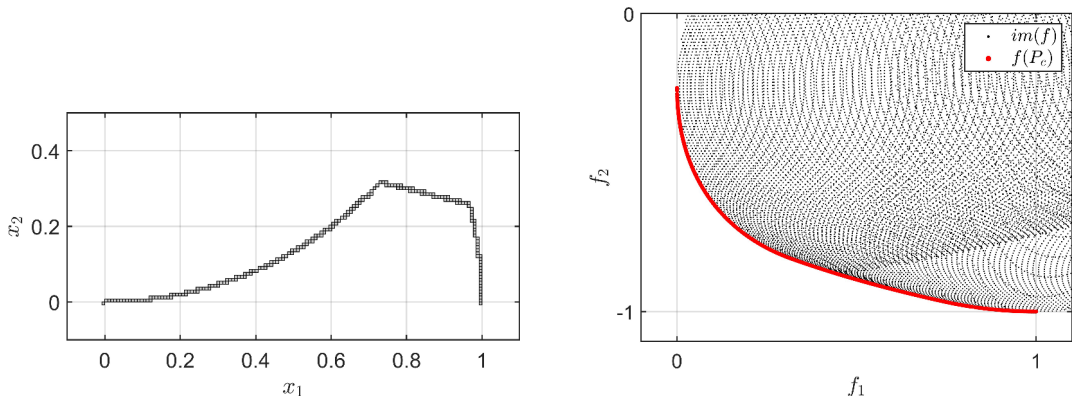
(a) **(b)**

Figure 4.11: **(a)** Result of the subdivision algorithm for problem 6 from Table 4.1 with $N = 5$ in (4.25). **(b)** Image of the center points of all boxes in (a) (red) and an approximation of the image of f (black).



(a) **(b)**

Figure 4.12: **(a)** Result of the subdivision algorithm for problem 12 from Table 4.1 with $N = 5$ in (4.25). **(b)** Image of the center points of all boxes in (a) (red) and an approximation of the image of f (black).



(a) **(b)**

Figure 4.13: **(a)** Result of the subdivision algorithm for problem 12 from Table 4.1 with $N = 10$ in (4.25). **(b)** Image of the center points of all boxes in (a) (red) and an approximation of the image of f (black).

5 Hierarchical structure of Pareto critical sets

Arguably, the main difference between scalar and multiobjective optimization is the fact that the solution of an MOP is (typically) a *set* instead of a single point. Thus, to better understand a multiobjective problem and to be able to efficiently solve it, it is vital to have a good understanding of the structure of the solution set, i.e., of the Pareto set. Unfortunately, the structure of the Pareto set itself (as defined in Definition 2.1.2 via the order relation on the image space $\mathbb{R}^k$) is difficult to analyze. The reason for this is the fact that it is not possible to decide whether a point $x \in \mathbb{R}^n$ is Pareto optimal using only local information of f around x, like gradients or higher-order derivatives. Fortunately, with the Pareto critical set P_c, we have a superset of the Pareto set that only depends on local information and is much easier to work with. Furthermore, in many cases (cf. Remark 2.2.6 e)), P_c is a tight enough superset to allow for the derivation of relevant structural properties of the actual Pareto set from it. Thus, in this chapter, we will investigate the structure of the Pareto critical set.

Despite its importance, there are relatively few existing results in this area. The first definition of the Pareto critical set was given by Smale in [Sma73] (for the more general case of objective vectors which are defined on smooth manifolds). There, it was proposed (without proof) that if a C^∞ function $f : M \to \mathbb{R}^k$ on an n-dimensional C^∞ manifold M with $k \leq n$ satisfies the rank assumption (cf. Definition 3.1.2) and certain assumptions on the transversality with respect to M, then P_c is a *stratification* [Mat12]. Stratifications can be thought of as "manifolds with boundaries and corners" [LP14]. Similarly, in [Mel76], de Melo showed that there is an open and dense subset of all C^∞ functions defined on compact C^∞ manifolds for which P_c is a stratification. If $M = \mathbb{R}^n$, $k \leq n$ and f is C^∞, satisfies the rank assumption and is convex, then P_c is diffeomorphic to a $(k-1)$-simplex [Sma73; LP14]. In this case, each $(j-1)$-dimensional facet of P_c is the Pareto (critical) set of a subset of the objective functions $\{f_1, \ldots, f_k\}$ of size j. More recently, in [Hil01], Hillermeier showed that under a certain regularity assumption involving the Hessians of the objectives, part of the Pareto critical set is the projection of a $(k-1)$-dimensional, differentiable manifold from $\mathbb{R}^{n+k}$ onto $\mathbb{R}^n$ (cf. Theorem 2.2.10). In [LP14], Lovison and Pecci showed that there is a dense subset of the set of differentiable functions for which the local Pareto set is a *Whitney stratification* [Mat12]. In [Pei17], the structure of the boundary of P_c was investigated in a more practical setting and a first solution method that exploits this structure was proposed. All results mentioned so far require some degree of differentiability of f. Without requiring differentiability, Lowe et al. showed in [Low+84] that convex MOPs are *Pareto reducible*, which means that the set of weak Pareto optimal points

is the union of Pareto optimal points of subsets of objectives. This result was later extended by Malivert and Boissard [MB94] and Popovici [Pop05]. Finally, since much of the research in multiobjective optimization is focused on the image space, it is of equal interest to investigate the structure of the Pareto front. This has been done by Mueller-Gritschneider, Graeb and Schlichtmann in [MGGS09], where it was shown that for certain well-behaved objective vectors, the "boundary" of the Pareto front is given by Pareto fronts of subsets of objectives.

In this chapter, the focus will lie on the structure of P_c with respect to Pareto critical sets of so-called *subproblems*, which are MOPs where only a subset of all objective functions is considered. As mentioned above, if $k \leq n$ and f is smooth, convex and satisfies the rank assumption, then P_c is diffeomorphic to a $(k-1)$-simplex such that the Pareto critical sets of the subproblems describe the facets of P_c. In particular, considering subproblems of different sizes induces a *hierarchy* of boundaries of P_c. Here, we will generalize this result by dropping both the convexity assumption and the restriction on the number of variables n and objectives k. From a theoretical point of view, we will do this by investigating the relationship between Pareto critical points and their corresponding KKT vectors, while exploiting the fact that the set $\mathcal{M}$ (as defined in Theorem 2.2.10) is a manifold.

In Section 5.1, we begin by considering smooth, unconstrained MOPs and show that the boundary of P_c can be covered by the Pareto critical sets of all subproblems where one objective is ignored. Since it can happen that this covering is not very tight (especially for $k > n+1$), we then investigate how many objective functions are actually required to still obtain a covering of the boundary in Section 5.2. It will turn out that the number of required objectives is given by the maximal rank of the Jacobian of the objective vector f on P_c. Afterwards, we visualize these results in examples (Section 5.3) before considering extensions to the constrained (Section 5.4) and nonsmooth case (Section 5.5).

We conclude the introduction to this chapter with a simple example that highlights the structure we want to investigate. Consider the problem

$$\min_{x \in \mathbb{R}^2} f(x) \quad \text{with} \quad f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} = \begin{pmatrix} (x_1-1)^2 + (x_2+1)^2 \\ x_1^2 + (x_2-1)^2 \\ (x_1+1)^2 + (x_2+1)^2 \end{pmatrix}. \tag{5.1}$$

As already discussed in Example 2.1.6, the Pareto critical set of this problem is given by the convex hull of $(1,-1)^\top$, $(0,1)^\top$ and $(-1,-1)^\top$, as shown in Figure 5.1(a). Figure 5.1(b) shows the Pareto critical sets of all subproblems. For example, if we ignore the third objective function f_3 and only consider f_1 and f_2, the resulting Pareto critical set is given by the red line connecting $(1,-1)^\top$ and $(0,1)^\top$, which is part of the boundary of P_c. If we additionally ignore f_2 and only consider f_1, the Pareto critical set of the resulting single-objective problem is the lower-right corner $(1,-1)^\top$ of the original Pareto critical set. As we see, in this case, the boundary of P_c is given by all subproblems with two objective functions. The simplicity of the structure of P_c in this example comes down to two reasons:

- The relationship between Pareto critical points and their corresponding KKT vectors is simple.
- The topological boundary of P_c as a subset of $\mathbb{R}^n$ can be used to describe the "boundary" of P_c.

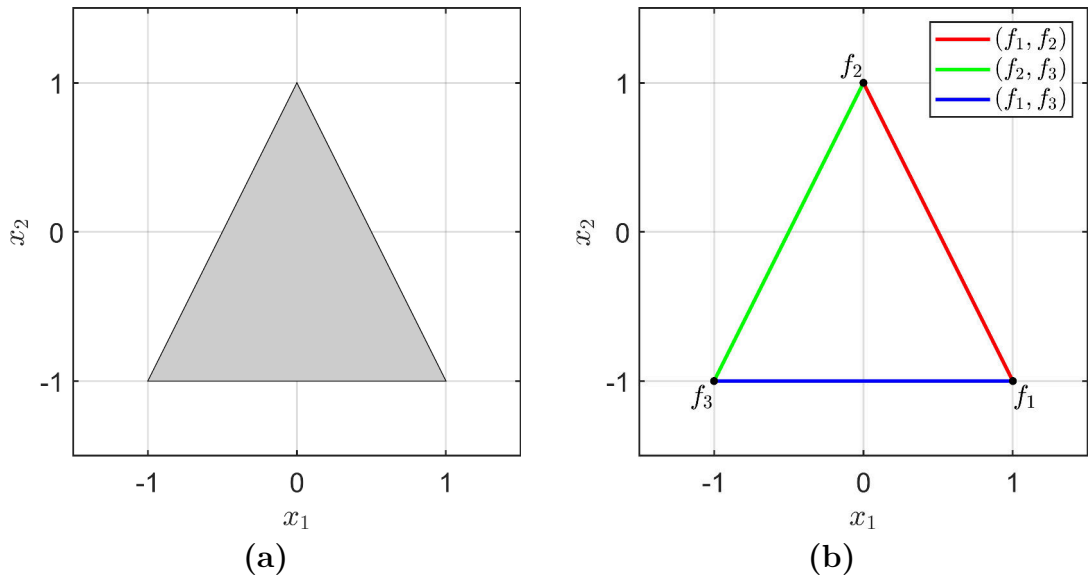


Figure 5.1: **(a)** Pareto critical set of (5.1). **(b)** Pareto critical sets of all subproblems.

In the more general setting we will consider in the following, both of these properties are lost. For example, when P_c consists of multiple connected components, the relationship between P_c and the KKT vectors is generally more complicated. Furthermore, for $k < n$, P_c is typically a null set (or "lower-dimensional" set) in $\mathbb{R}^n$, in which case the topology of $\mathbb{R}^n$ does not yield a useful definition for the "boundary" of P_c.

For the computation of the Pareto critical sets in the examples in this chapter, we will use the methods presented in Chapter 3 and 4. (To better visualize the structure of the sets, we will use the approximately Pareto critical points generated in the algorithms instead of the box coverings.)

Parts of the Sections 5.1, 5.2 and 5.3 have been previously published in [GPD19], to which the author of this thesis was the main contributor.

5.1 Topological and geometrical properties of the Pareto critical set

In this section, we will investigate the structure of the Pareto critical set with respect to the corresponding KKT vectors. Clearly, $x \in P_c$ having a KKT vector $\alpha \in \Delta_k$ with $\alpha_i = 0$ for some $i \in \{1, \dots, k\}$ is equivalent to x being Pareto critical for the subproblem where f_i is ignored. Thus, we will begin by classifying Pareto critical points with respect to zero entries in the corresponding KKT vectors, resulting in the sets P_0 and P_{int}. We will discuss some of their properties and relate them to the manifold structure of the set $\mathcal{M}$ from Theorem 2.2.10. As mentioned in the introduction, it is not yet clear what we mean by the "boundary" of P_c in the general case. To be able to define it, we will consider tangent cones with respect to P_c and investigate their properties.

5.1.1 Classifying Pareto critical points via KKT multipliers

Let the objective vector $f : \mathbb{R}^n \to \mathbb{R}^k$ be twice continuously differentiable and $k > 1$. Recall from Section 2.2 that $x \in \mathbb{R}^n$ being Pareto critical with KKT vector $\alpha \in \Delta_k$ is equivalent to $F(x, \alpha) = 0$ for

$$F : \mathbb{R}^n \times (\mathbb{R}^{\geq 0})^k \to \mathbb{R}^{n+1}, \quad (x, \alpha) \mapsto \begin{pmatrix} \sum_{i=1}^k \alpha_i \nabla f_i(x) \\ 1 - \sum_{i=1}^k \alpha_i \end{pmatrix}.$$

For ease of notation, we define

$$A(x) := \{\alpha \in \Delta_k : F(x, \alpha) = 0\} \tag{5.2}$$

as the set of KKT vectors of $x \in \mathbb{R}^n$. Clearly, $x \in P_c$ is equivalent to $A(x) \neq \emptyset$. We classify Pareto critical points in the following way.

Definition 5.1.1. *Define*

$$P_{int} := \{x \in P_c : \exists \alpha \in A(x) \text{ with } \alpha_i > 0 \ \forall i \in \{1, \dots, k\}\},$$
$$P_0 := P_c \setminus P_{int} = \{x \in P_c : \forall \alpha \in A(x) \exists j \in \{1, \dots, k\} \text{ with } \alpha_j = 0\}.$$

In words, P_{int} is the set of Pareto critical points that have a KKT vector that is strictly positive and P_0 is the set of Pareto critical points where every KKT vector has a zero component. The following example derives both sets for the simple MOP (5.1).

Example 5.1.2. *Let f be as in (5.1) and $x \in P_c$. Then $\alpha \in A(x)$ implies*

$$\begin{aligned} 0 = Df(x)^\top \alpha &= 2 \begin{pmatrix} x_1 - 1 & x_1 & x_1 + 1 \\ x_2 + 1 & x_2 - 1 & x_2 + 1 \end{pmatrix} \alpha \\ &= 2 \begin{pmatrix} (\alpha_1 + \alpha_2 + \alpha_3)x_1 - \alpha_1 + \alpha_3 \\ (\alpha_1 + \alpha_2 + \alpha_3)x_2 + \alpha_1 - \alpha_2 + \alpha_3 \end{pmatrix} = 2 \begin{pmatrix} x_1 - \alpha_1 + \alpha_3 \\ x_2 + \alpha_1 - \alpha_2 + \alpha_3 \end{pmatrix}, \end{aligned}$$

i.e.,

$$\begin{aligned} x_1 &= \alpha_1 - \alpha_3, \\ x_2 &= -\alpha_1 + \alpha_2 - \alpha_3, \end{aligned}$$

which is equivalent to

$$\begin{aligned} \alpha_1 &= \frac{1}{4}(2x_1 - x_2 + 1), \\ \alpha_2 &= \frac{1}{2}(x_2 + 1), \\ \alpha_3 &= \frac{1}{4}(-2x_1 - x_2 + 1). \end{aligned}$$

It is easy to see that P_{int} is the interior of the triangle in Figure 5.1 and P_0 is its boundary, i.e., $P_{int} = P_c^\circ$ and $P_0 = \partial P_c$.

In the following, we will investigate some of the topological and differential geometrical properties of P_{int} and P_0. A useful strategy for doing so is the local parametrization of P_c via the KKT vectors, as was done in the previous example. In the general case, this can be done by applying the implicit function theorem to the equation $F(x, \alpha) = 0$. Note that due to the dimensions of the domain and the image of F, we cannot apply it directly. Instead, we first have to modify F by incorporating the fact that if $\alpha \in \Delta_k$, then any component of α is uniquely determined by the other $k-1$ components. More precisely, we have the following, technical result.

Lemma 5.1.3. *Let f be m-times continuously differentiable for $m \geq 2$. Let $x^0 \in P_c$ and $\alpha^0 = (\alpha_1^0, \ldots, \alpha_k^0)^\top \in A(x^0)$ such that $\sum_{i=1}^k \alpha_i^0 \nabla^2 f_i(x^0)$ is regular. Then there is an open set $\tilde{U} \subseteq \mathbb{R}^{k-1}$ with $(\alpha_1^0, \ldots, \alpha_{k-1}^0) \in \tilde{U}$, an open set $V \subseteq \mathbb{R}^n$ with $x^0 \in V$ and an $(m-1)$-times continuously differentiable function $\phi : \tilde{U} \to V$ with $\phi((\alpha_1^0, \ldots, \alpha_{k-1}^0)) = x^0$ and*

$$F(x, \alpha) = 0 \;\Leftrightarrow\; x = \phi((\alpha_1, \ldots, \alpha_{k-1}))$$

for all $x \in V$ and $\alpha \in \mathbb{R}^k$ with $\sum_{i=1}^k \alpha_i = 1$ and $(\alpha_1, \ldots, \alpha_{k-1}) \in \tilde{U}$. Furthermore,

$$D\phi(\tilde{\alpha}) = -D_1(x, \tilde{\alpha})^{-1} D_2(x) \in \mathbb{R}^{n \times (k-1)} \quad \forall \tilde{\alpha} \in \tilde{U},$$

where

$$\begin{aligned} D_1(x, \tilde{\alpha}) &:= \left(\sum_{i=1}^{k-1} \tilde{\alpha}_i \nabla^2 f_i(x) \right) + \left(1 - \sum_{i=1}^{k-1} \tilde{\alpha}_i \right) \nabla^2 f_k(x) \in \mathbb{R}^{n \times n}, \\ D_2(x) &:= (\nabla f_1(x) - \nabla f_k(x), \ldots, \nabla f_{k-1}(x) - \nabla f_k(x)) \in \mathbb{R}^{n \times (k-1)}. \end{aligned}$$

Proof. Let $\tilde{F} : \mathbb{R}^n \times \mathbb{R}^{k-1} \to \mathbb{R}^n$,

$$\begin{aligned} \tilde{F}(x, \tilde{\alpha}) &:= \left(\sum_{i=1}^{k-1} \tilde{\alpha}_i \nabla f_i(x) \right) + \left(1 - \sum_{i=1}^{k-1} \tilde{\alpha}_i \right) \nabla f_k(x) \\ &= \left(\sum_{i=1}^{k-1} \tilde{\alpha}_i (\nabla f_i(x) - \nabla f_k(x)) \right) + \nabla f_k(x), \end{aligned}$$

and

$$\tilde{\Delta}_{k-1} := \left\{ \tilde{\alpha} \in \mathbb{R}^{k-1} : \tilde{\alpha}_i \geq 0 \; \forall i \in \{1, \ldots, k-1\}, \; \sum_{i=1}^{k-1} \tilde{\alpha}_i \leq 1 \right\}.$$

Then by construction,

$$\begin{aligned} x \in P_c &\Leftrightarrow \exists \alpha \in \Delta_k : F(x, \alpha) = 0 \\ &\Leftrightarrow \exists \tilde{\alpha} \in \tilde{\Delta}_{k-1} : \tilde{F}(x, \tilde{\alpha}) = 0. \end{aligned}$$

Furthermore, the Jacobian of $\tilde{F}$ consists of the matrices $D_x \tilde{F}(x, \tilde{\alpha}) = D_1(x, \tilde{\alpha})$ and $D_{\tilde{\alpha}} \tilde{F}(x, \tilde{\alpha}) = D_2(x)$ as defined above. Now let $\tilde{\alpha}^0 := (\alpha_1^0, \ldots, \alpha_{k-1}^0)$. From our assumptions it follows that $\tilde{F}$ is $(m-1)$-times continuously differentiable and $D_x \tilde{F}(x^0, \tilde{\alpha}^0)$ is regular. Thus, we can apply the implicit function theorem (see, e.g.,

[KP13], Theorem 3.3.1) to $\tilde{F}$ to obtain an open set $\tilde{U} \subseteq \mathbb{R}^{k-1}$ with $\tilde{\alpha}^0 \in \tilde{U}$, an open set $V \subseteq \mathbb{R}^n$ with $x^0 \in V$ and a C^{m-1} function $\phi : \tilde{U} \to V$ with $\phi(\tilde{\alpha}^0) = x^0$ and

$$\tilde{F}(x, \tilde{\alpha}) = 0 \;\Leftrightarrow\; x = \phi(\tilde{\alpha})$$

for each $x \in V$ and $\tilde{\alpha} \in \tilde{U}$. In particular, implicit differentiation of $\tilde{F}(\phi(\tilde{\alpha}), \tilde{\alpha}) = 0$ shows that

$$D\phi(\tilde{\alpha}) = -D_x\tilde{F}(x, \tilde{\alpha})^{-1} D_{\tilde{\alpha}}\tilde{F}(x, \tilde{\alpha}) = -D_1(x, \tilde{\alpha})^{-1} D_2(x),$$

completing the proof. □

The previous lemma has some similarities to Theorem 2.2.10, which gave a condition for when the set

$$\mathcal{M} = \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R}^k : \alpha \in \Delta_k \cap (\mathbb{R}^{>0})^k, F(x, \alpha) = 0\}$$

is a $(k-1)$-dimensional manifold. $\mathcal{M}$ being a manifold means that it is locally diffeomorphic to an open subset of $\mathbb{R}^{k-1}$. In particular, by ignoring the α component of $\mathcal{M}$, this allows us to locally express the set of Pareto critical points (with positive KKT vectors) as the image of a differentiable function in $k-1$ variables. By Lemma 2.2.12, the requirements for Lemma 5.1.3 are stronger than the requirements for Theorem 2.2.10 b). So in the setting of Lemma 5.1.3, $\mathcal{M}$ is also a manifold around (x^0, α^0) (for $\alpha^0 \in (\mathbb{R}^{>0})^k$), but we have the even stronger result that we can parametrize P_c specifically as the graph of a function ϕ which depends on (the first $k-1$ components of the) KKT vectors. The difference is shown in the following example.

Example 5.1.4. *Consider the problem*

$$\min_{x \in \mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} x_1^3 - x_1^2 - x_1x_2 + \frac{1}{2}x_2^2 \\ \frac{23}{3}x_1^3 - x_1^2x_2 - \frac{21}{2}x_1^2 + 6x_1 + \frac{1}{3}x_2^3 \end{pmatrix}.$$

The Pareto critical set is the line connecting $(0,0)^\top$ and $(1,1)^\top$, as shown in Figure 5.2(a). Figure 5.2(b) shows the relationship between α_1 and x_1 for Pareto critical points $x = (x_1, x_2)^\top$ with KKT vector $\alpha = (\alpha_1, \alpha_2)^\top$. (Note that $x_2 = x_1$ for $x \in P_c$ and $\alpha_2 = 1 - \alpha_1$.) We see that no Pareto critical point with $\alpha_1 \in [0, \frac{1}{2})$ exists, the point $(\frac{1}{2}, \frac{1}{2})^\top$ is Pareto critical with KKT vector $(\frac{1}{2}, \frac{1}{2})^\top$ and for $\alpha_1 \in (\frac{1}{2}, 1]$, there are two Pareto critical points for each KKT vector. In particular, around $x^0 = (\frac{1}{2}, \frac{1}{2})^\top$, the Pareto critical set cannot be expressed as the graph of a function in α.
If we consider the derivatives of f, we obtain

$$\nabla f_1(x) = \begin{pmatrix} 3x_1^2 - 2x_1 - x_2 \\ x_2 - x_1 \end{pmatrix}, \; \nabla f_2(x) = \begin{pmatrix} 23x_1^2 - 2x_1x_2 - 21x_1 + 6 \\ x_2^2 - x_1^2 \end{pmatrix},$$
$$\nabla^2 f_1(x) = \begin{pmatrix} 6x_1 - 2 & -1 \\ -1 & 1 \end{pmatrix}, \; \nabla^2 f_2(x) = \begin{pmatrix} 46x_1 - 2x_2 - 21 & -2x_1 \\ -2x_1 & 2x_2 \end{pmatrix}.$$

Thus, for $x^0 = (\frac{1}{2}, \frac{1}{2})^\top$ (with KKT vector $\alpha^0 = (\frac{1}{2}, \frac{1}{2})^\top$) we have

$$\sum_{i=1}^{2} \alpha_i^0 \nabla^2 f_i(x^0) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$
$$DF(x^0, \alpha^0) = \begin{pmatrix} \sum_{i=1}^{2} \alpha_i^0 \nabla^2 f_i(x^0) & Df(x^0)^\top \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & -\frac{3}{4} & \frac{3}{4} \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

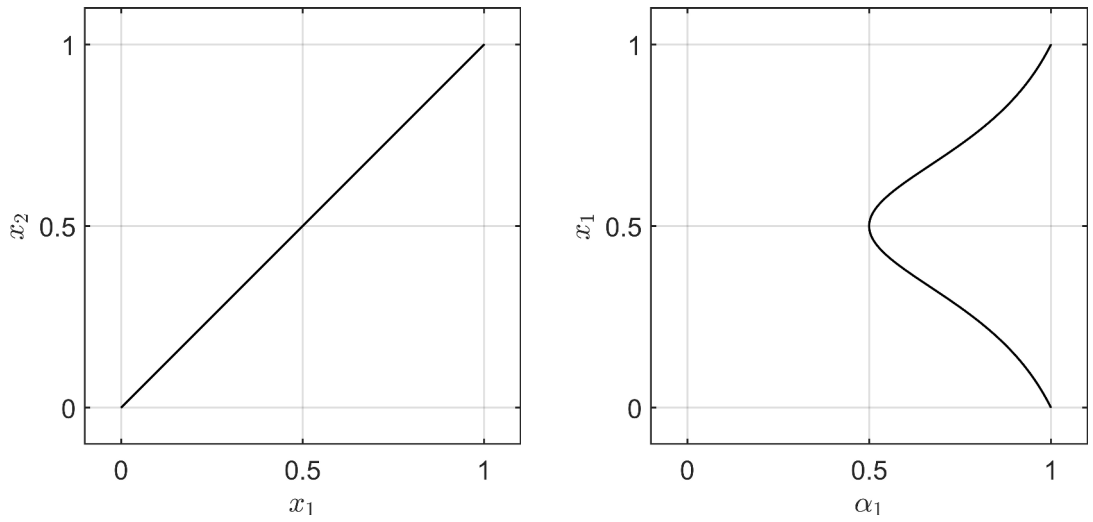


Figure 5.2: **(a)** Pareto critical set in Example 5.1.4. **(b)** Relationship between x_1 and the first entry of the KKT vector α_1.

Since the weighted Hessian $\sum_{i=1}^{2} \alpha_i^0 \nabla^2 f_i(x^0)$ is not regular, Lemma 5.1.3 cannot be applied in x^0, as expected. But since the rank of $DF(x^0, \alpha^0)$ is full nonetheless, $\mathcal{M}$ is still a manifold around $(x^0, \alpha^0) \in \mathcal{M}$.

In [Wit12], Pareto critical points with an irregular weighted Hessian like in the previous example were called *dent border points*, and their relationship to the curvature of the Pareto front was investigated.

Lemma 5.1.3 allows us to show a topological result about P_{int} and P_0.

Lemma 5.1.5. *a) P_c is closed.*

b) If for all $x^0 \in P_0$ there is some $\alpha^0 \in A(x^0)$ such that $\sum_{i=1}^{k} \alpha_i^0 \nabla^2 f_i(x^0)$ is regular, then

$$\overline{P_{int}} = P_c,$$

where $\overline{P_{int}}$ is the closure of P_{int} in $\mathbb{R}^n$.

Proof. a) Let $(x^j)_j \in P_c$ be a sequence with $\lim_{j\to\infty} x^j = \bar{x} \in \mathbb{R}^n$. Let $(\alpha^j)_j \in \Delta_k$ be a sequence with $\alpha^j \in A(x^j)$ for all $j \in \mathbb{N}$, i.e.,

$$F(x^j, \alpha^j) = 0 \quad \forall j \in \mathbb{N}. \tag{5.3}$$

Since Δ_k is compact, we can assume w.l.o.g. that $\lim_{j\to\infty} \alpha^j = \bar{\alpha} \in \Delta_k$. Since f is twice continuously differentiable by assumption, F is continuous. Thus, (5.3) implies $F(\bar{x}, \bar{\alpha}) = 0$, i.e., $\bar{x} \in P_c$ (with KKT vector $\bar{\alpha}$).
b) By a) we have $\overline{P_{int}} \subseteq \overline{P_c} = P_c$, so it remains to show that $P_c \subseteq \overline{P_{int}}$. Trivially, $P_{int} \subseteq \overline{P_{int}}$, so let $x^0 \in P_c \setminus P_{int} = P_0$ and $\alpha^0 \in A(x^0)$ such that the weighted Hessian is regular. Then we can apply Lemma 5.1.3 in (x^0, α^0) to obtain a function $\phi : \tilde{U} \to V$ with the stated properties. Since $\alpha^0 \in \Delta_k$ and by the structure of Δ_k, we can write α^0 as the limit of a sequence $\alpha^j \in \Delta_k \cap (\mathbb{R}^{>0})^k$. Let $\tilde{\alpha}^j \in \mathbb{R}^{k-1}$ be the projection of α^j onto the first $k-1$ components. We can assume w.l.o.g. that $\tilde{\alpha}^j \in \tilde{U}$. Let $x^j := \phi(\tilde{\alpha}^j)$. Then by construction, $x^j \in P_{int}$ for all $j \in \mathbb{N}$. Continuity of ϕ implies $\lim_{j\to\infty} x^j = x^0$, i.e., $x^0 \in \overline{P_{int}}$. □

The following examples shows how P_{int} and P_0 may look when the premise of Lemma 5.1.5 b) is false.

Example 5.1.6. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} x_2^2 \\ x_1^2 + (x_2-1)^2 \end{pmatrix}.$$

Then for $x \in \mathbb{R}^2$ and $\alpha \in \Delta_2$ we have

$$\begin{aligned} 0 = Df(x)^\top \alpha &= \begin{pmatrix} 2\alpha_2 x_1 \\ 2\alpha_1 x_2 + 2\alpha_2(x_2-1) \end{pmatrix} = \begin{pmatrix} 2\alpha_2 x_1 \\ 2x_2 - 2\alpha_2 \end{pmatrix} \\ \Leftrightarrow \quad x &\in \begin{cases} \mathbb{R}\times\{0\}, & \textit{if } \alpha = (1,0)^\top, \\ \{0\}\times(0,1], & \textit{otherwise.} \end{cases} \end{aligned}$$

In particular, $P_0 = \{(0,1)^\top\} \cup (\mathbb{R}\times\{0\})$ and $P_{int} = \{0\}\times(0,1)$. The Pareto critical set is shown in Figure 5.3 with colors highlighting P_{int} and P_0.

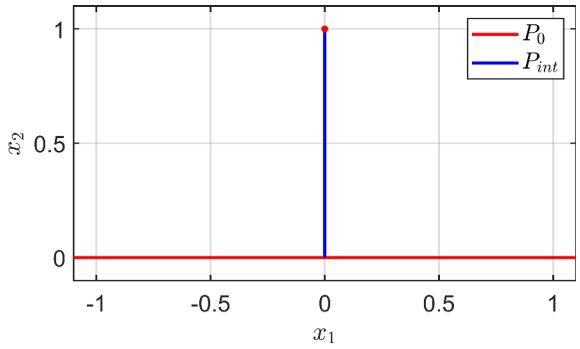


Figure 5.3: P_{int} and P_0 for Example 5.1.6.

Clearly,

$$\overline{P_{int}} = \{0\}\times[0,1] \neq (\{0\}\times[0,1]) \cup (\mathbb{R}\times\{0\}) = P_c.$$

If we consider the weighted Hessian for $x \in P_0$ and $\alpha = (1,0) \in A(x)$, we obtain

$$\sum_{i=1}^{2} \alpha_i \nabla^2 f_i(x) = \nabla^2 f_1(x) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

As expected, this matrix is not regular.

If the premise of Lemma 5.1.5 b) is true, then P_0 can be thought of as "small" compared to P_{int}. Recall that our goal is to show that on the "boundary" of P_c, there is a component of the KKT vector which is zero. As mentioned in the introduction, for well-behaved problems (cf. [Sma73; LP14]), P_0 is a suitable (and intuitive) description of the boundary. Unfortunately, for more complex problems, the definition of the boundary of P_c is more difficult, since the structure of P_c is in general unclear. For example, even if $\mathcal{M}$ is a manifold (cf. Theorem 2.2.10), $\mathrm{pr}_x(\mathcal{M}) = P_{int}$ might not be a manifold. This is visualized in the following example.

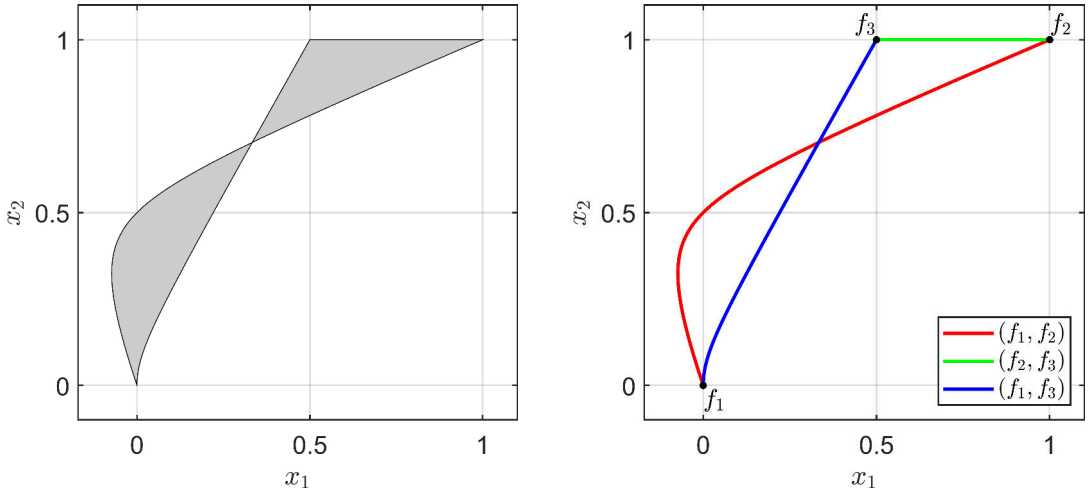
Example 5.1.7. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \text{with} \quad f(x) = \begin{pmatrix} \|x\|_{Q_1}^2 \\ \|x - c^1\|^2 \\ \|x - c^2\|_{Q_2}^2 \end{pmatrix},$$

where $\|\cdot\|_{Q_i}$ *is the norm induced by* Q_i *and*

$$Q_1 = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, c^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{pmatrix}, c^2 = \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}.$$

It is easy to see that all objective functions are strongly convex, so $\mathcal{M}$ *is a manifold by Lemma 2.2.12. The Pareto critical set of this problem can be computed with the same technique as in Example 2.2.8 and is shown in Figure 5.4(a). Figure 5.4(b) shows the solution of all subproblems with one or two objective functions. Compared*



(a) **(b)**

Figure 5.4: **(a)** Pareto critical set in Example 5.1.7. **(b)** Pareto critical sets of all subproblems.

to the simple example in Figure 5.1 at the beginning of this chapter, we see that in this case, the Pareto critical set is "folded" such that the Pareto critical sets of (f_1, f_2) *and* (f_1, f_3) *have two intersections. The first intersection trivially occurs at the critical point of* f_1*, but a second intersection occurs at*

$$x^0 = \left(\frac{1}{3}, \frac{1}{12}(\sqrt{89} - 1)\right)^\top \approx (0.3333, 0.7028)^\top.$$

Due to the linearity of the KKT condition $Df(x)^\top\alpha = 0$ *in* α*, all convex combinations of the KKT vectors of the subproblems (lifted into* Δ_3*) are KKT vectors of* x^0*. In particular, there is a KKT vector where all entries are positive, showing that* $x^0 \in P_{int}$*. The remaining part of* P_{int} *is the gray area shown in Figure 5.4(a). We see that* P_{int} *cannot be diffeomorphic to an open subset of* $\mathbb{R}^2$ *locally around* x^0*. (For example, if we remove the point* x^0 *from* P_{int}*, then* P_{int} *consists of two connected components. But we cannot remove a single point from an open, connected subset of* $\mathbb{R}^2$ *such that the remaining set has more than one connected component.) Thus,* P_{int} *is not a manifold.*

5.1.2 Tangent cones and the uniqueness of KKT vectors

As discussed in the introduction of this chapter, a proper definition of the boundary of the Pareto critical set P_c is non-trivial. Since P_c is typically a null set in $\mathbb{R}^n$ for $k \leq n$, we cannot use the topological boundary with respect to the topology of $\mathbb{R}^n$. As Example 5.1.7 showed, we also cannot expect P_c (or the subset P_{int}) to be a manifold (even if f is strongly convex), so we cannot directly apply tools or use definitions from differential geometry. Instead, we will use a notion of the boundary which is based on *tangent cones* of the Pareto critical set, which we will investigate in the following.

Definition 5.1.8. *Let $Y \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$. Then*

$$\mathrm{Tan}(Y,x) := \Bigg\{ v \in \mathbb{R}^n : \exists (v^i)_i \in \mathbb{R}^n \setminus \{0\} \text{ with } \lim_{i\to\infty} v^i = 0,\ x + v^i \in Y,\ \lim_{i\to\infty} \frac{v^i}{\|v^i\|} = \frac{v}{\|v\|} \Bigg\} \cup \{0\}$$

is the tangent cone of Y at x.

There are many equivalent ways of defining the tangent cone. For a detailed overview, see [GG92]. For $x \in Y$, the tangent cone $\mathrm{Tan}(Y,x)$ contains all directions emanating from x that point into or alongside Y. As trivial cases, for $x \in Y^\circ$ we have $\mathrm{Tan}(Y,x) = \mathbb{R}^n$, and for $x \notin \overline{Y}$ we have $\mathrm{Tan}(Y,x) = \{0\}$. Note that in contrast to the tangent space of a manifold, the tangent cone can be defined for any subset $Y \subseteq \mathbb{R}^n$ without requiring any geometrical structure. The downside is that there are fewer structural results for tangent cones than for tangent spaces.

From now on, we will always consider the tangent cone with respect to sets of Pareto critical points. We will begin by showing that if the weighted Hessian is regular on P_0, then it does not matter if we take the tangent cone with respect to P_c or P_{int}.

Lemma 5.1.9. *If for all $x^0 \in P_0$ there is some $\alpha^0 \in A(x^0)$ such that the weighted Hessian $\sum_{i=1}^k \alpha_i^0 \nabla^2 f_i(x^0)$ is regular, then*

$$\mathrm{Tan}(P_{int}, x) = \mathrm{Tan}(P_c, x) \quad \forall x \in P_c.$$

Proof. As shown in [GG92], the tangent cone can equivalently be defined as

$$\mathrm{Tan}(Y,x) = \left\{ v \in \mathbb{R}^n : \liminf_{h \searrow 0} \frac{d_Y(x+hv)}{h} = 0 \right\} \cup \{0\}$$

with $d_Y(x) := \inf_{y\in Y} \|y - x\|$. Clearly, $d_Y(x) = d_{\overline{Y}}(x)$ for all $x \in \mathbb{R}^n$, so we have $\mathrm{Tan}(Y,x) = \mathrm{Tan}(\overline{Y},x)$. By Lemma 5.1.5 b) it holds $\overline{P_{int}} = P_c$, so

$$\mathrm{Tan}(P_{int}, x) = \mathrm{Tan}(\overline{P_{int}}, x) = \mathrm{Tan}(P_c, x) \quad \forall x \in P_c.$$

□

By revisiting Example 5.1.6, we see that the assertion of the previous lemma may fail to hold if the weighted Hessian is not regular.

Although P_{int} is in general not a manifold, recall that we can write it as the projection of the set $\mathcal{M} \subseteq \mathbb{R}^n \times \mathbb{R}^k$ onto $\mathbb{R}^n$. If $\mathcal{M}$ is a manifold (cf. Theorem

2.2.10), this will allow us to convey some of the smooth structure of $\mathcal{M}$ to P_{int}. More precisely, we will investigate the relationship between the tangent cone of P_{int}, i.e., $\mathrm{Tan}(P_{int}, x) = \mathrm{Tan}(\mathrm{pr}_x(\mathcal{M}), x)$, and the projection of the tangent space of $\mathcal{M}$ onto the first n components, i.e., $\mathrm{pr}_x(T_{(x,\alpha)}\mathcal{M})$, for $(x, \alpha) \in \mathcal{M}$. Our goal is to show that under certain assumptions, both sets coincide. We will begin by showing that any tangent vector in the tangent space of $\mathcal{M}$ is projected onto an element of the tangent cone of P_{int}.

Lemma 5.1.10. *Let $(x, \alpha) \in \mathcal{M}$ with $\mathrm{rk}(DF(x, \alpha)) = n + 1$. Then*

$$\mathrm{pr}_x(T_{(x,\alpha)}\mathcal{M}) \subseteq \mathrm{Tan}(P_{int}, x).$$

Proof. By Theorem 2.2.10 b), $\mathcal{M}$ is a manifold around (x, α). Let $v \in T_{(x,\alpha)}\mathcal{M}$. Since $0 \in \mathrm{Tan}(P_{int}, x)$ by definition, assume that $\mathrm{pr}_x(v) \neq 0$. By definition of the tangent space as the set of equivalence classes of curves (cf. [Lee12], Chapter 3), there is a C^1 function $\gamma : (-1, 1) \to \mathcal{M}$ with $\gamma(0) = (x, \alpha)$ and $\gamma'(0) = v$. For $i \in \mathbb{N}$ let

$$v^i := \mathrm{pr}_x(\gamma(i^{-1}) - \gamma(0)) = \mathrm{pr}_x(\gamma(i^{-1})) - x.$$

W.l.o.g. assume that $v^i \neq 0$ for all $i \in \mathbb{N}$. Then

$$\begin{aligned}\frac{v^i}{\|v^i\|} &= \frac{\mathrm{pr}_x(\gamma(i^{-1}) - \gamma(0))}{\|\mathrm{pr}_x(\gamma(i^{-1}) - \gamma(0))\|} = \frac{\mathrm{pr}_x(\gamma(i^{-1}) - \gamma(0))}{i^{-1}} \frac{i^{-1}}{\|\mathrm{pr}_x(\gamma(i^{-1}) - \gamma(0))\|} \\ &= \mathrm{pr}_x\left(\frac{\gamma(i^{-1}) - \gamma(0)}{i^{-1}}\right) \left\|\mathrm{pr}_x\left(\frac{\gamma(i^{-1}) - \gamma(0)}{i^{-1}}\right)\right\|^{-1},\end{aligned}$$

so

$$\lim_{i\to\infty} \frac{v^i}{\|v^i\|} = \frac{\mathrm{pr}_x(v)}{\|\mathrm{pr}_x(v)\|}$$

and $\mathrm{pr}_x(v) \in \mathrm{Tan}(P_{int}, x)$. □

To show the opposite implication, i.e., $\mathrm{Tan}(P_{int}, x) \subseteq \mathrm{pr}_x(T_{(x,\alpha)}\mathcal{M})$, we first need some technical results about the uniqueness of KKT vectors.

Lemma 5.1.11. *Let $x \in P_c$.*

a) If $\mathrm{rk}(Df(x)) = k - 1$ then $|A(x)| = 1$.

b) If $x \in P_{int}$ and $|A(x)| = 1$, then $\mathrm{rk}(Df(x)) = k - 1$.

c) If $x \in P_{int}$ and $\mathrm{rk}(Df(x)) < k-1$, then $|A(x)| > 1$ and there is some $\alpha \in A(x)$ with $\alpha_j = 0$ for some $j \in \{1, \ldots, k\}$.

Proof. a) This follows directly from Lemma 3.1.1.
b) Since $x \in P_{int}$ there is some $\alpha \in A(x)$ with $\alpha_i > 0$ for all $i \in \{1, \ldots, k\}$. In particular, $\alpha \in \ker(Df(x)^\top)$. Since $x \in P_c$ we must have $\mathrm{rk}(Df(x)) = \mathrm{rk}(Df(x)^\top) < k$. Assume that $\mathrm{rk}(Df(x)) \leq k - 2$. By the rank-nullity theorem it follows that $\dim(\ker(Df(x)^\top)) \geq 2$. Let $\beta \in \ker(Df(x)^\top)$ such that α and β are linearly independent. If $\sum_{i=1}^k \beta_i \neq 0$ then assume w.l.o.g. that $\sum_{i=1}^k \beta_i = 1$ and define $\nu := \alpha - \beta$. If $\sum_{i=1}^k \beta_i = 0$ define $\nu := \beta$. In both cases, consider the curve $\gamma(t) := \alpha + t\nu$. By

construction, $Df(x)^\top \gamma(t) = 0$ and $\sum_{i=1}^k (\gamma(t))_i = 1$ for all $t \in \mathbb{R}^n$. Since $\gamma(0) = \alpha$ and $\alpha_i > 0$ for all $i \in \{1, \dots, k\}$, there has to be some $t' > 0$ such that $\gamma(t') \in \Delta_k$. This implies $\alpha \neq \gamma(t') \in A(x)$, which is a contradiction.
c) Since $x \in P_{int}$ there is some $\alpha \in A(x)$ with $\alpha_i > 0$ for all $i \in \{1, \dots, k\}$. As in b), there must be some $\nu \in \mathbb{R}^k \setminus \{0\}$ such that $\gamma(t) = \alpha + t\nu$ satisfies $\gamma(0) = \alpha$, $Df(x)^\top \gamma(t) = 0$ and $\sum_{i=1}^k (\gamma(t))_i = 1$ for all $t \in \mathbb{R}^n$. Since Δ_k is bounded and $\alpha_i > 0$ for all $i \in \{1, \dots, k\}$, there must be some $t' > 0$ with $\gamma(t') \in A(x)$ and $\gamma(t')_j = 0$ for some $j \in \{1, \dots, k\}$, completing the proof. □

The previous lemma has the following implications.

Corollary 5.1.12. *a) Let $x \in P_{int}$. Then $\operatorname{rk}(Df(x)) = k - 1 \Leftrightarrow |A(x)| = 1$.*

b) Let $x \in P_c$. If $\operatorname{rk}(Df(x)) < k - 1$ then there is some $\alpha \in A(x)$ with $\alpha_j = 0$ for some $j \in \{1, \dots, k\}$.

Although Corollary 5.1.12 b) is a relatively simple result, it has an important implication for MOPs with more than $n + 1$ objectives: If $k > n + 1$, then

$$\operatorname{rk}(Df(x)) \leq \min(\{n, k\}) = n < k - 1$$

for any $x \in \mathbb{R}^n$. So by Corollary 5.1.12 b), *every* Pareto critical point is also Pareto critical for a subproblem, and we do not only cover the boundary when we compute the Pareto critical sets of all subproblems. This will be further discussed in Section 5.2.

We now use the result about the uniqueness of KKT vectors to show that if f satisfies the rank assumption (Definition 3.1.2) and the weighted Hessian is regular, then every element of the tangent cone of P_{int} is the projection of a tangent vector of $\mathcal{M}$.

Lemma 5.1.13. *Let $x^0 \in P_{int}$ with $\operatorname{rk}(Df(x^0)) = k - 1$. Let $\alpha^0 \in A(x^0)$. If $\sum_{i=1}^k \alpha_i^0 \nabla^2 f_i(x^0)$ is regular, then*

$$\operatorname{Tan}(P_{int}, x^0) \subseteq \operatorname{pr}_x(T_{(x^0, \alpha^0)}\mathcal{M}).$$

Proof. If $\operatorname{Tan}(P_{int}, x^0)$ is empty then the assertion trivially holds, so assume that $\operatorname{Tan}(P_{int}, x^0)$ is nonempty. Let $v \in \operatorname{Tan}(P_{int}, x^0)$ and $(v^i)_i \in \mathbb{R}^n \setminus \{0\}$ be a corresponding sequence as in Definition 5.1.8. Since $x^0 + v^i \in P_{int}$ for all $i \in \mathbb{N}$, this induces a sequence $(\alpha^i)_i \in \Delta_k$ with $\alpha^i \in A(x^0 + v^i)$ for all $i \in \mathbb{N}$. Since Δ_k is compact, we can assume w.l.o.g. that $\lim_{i\to\infty} \alpha^i = \bar{\alpha} \in \Delta_k$. This implies $F(x^0, \bar{\alpha}) = 0$ by continuity of F, so $\bar{\alpha} \in A(x^0)$. By Corollary 5.1.12, α^0 is the unique KKT vector of x^0, so $\lim_{i\to\infty} \alpha^i = \alpha^0$.
Since $\sum_{i=1}^k \alpha_i^0 \nabla^2 f_i(x^0)$ is regular, we can apply Lemma 5.1.3 in (x^0, α^0) to obtain a C^1 function $\phi : \tilde{U} \to V$ with the stated properties. Let $(\tilde{\alpha}^i)_i \in \mathbb{R}^{k-1}$ and $\tilde{\alpha}^0 \in \mathbb{R}^{k-1}$ be the sequence $(\alpha^i)_i$ and the vector α^0 without the last component, respectively. By construction, $\tilde{U}$ is open, $\tilde{\alpha}^0 \in \tilde{U}$ and $\lim_{i\to\infty} \tilde{\alpha}^i = \tilde{\alpha}^0$, so we can assume w.l.o.g. that $\tilde{\alpha}^i \in \tilde{U}$ for all $i \in \mathbb{N}$. Furthermore, $x^0 \in V$, V is open and $\lim_{i\to\infty} x^0 + v^i = x^0$, so we can assume w.l.o.g. that $x^0 + v^i \in V$ for all $i \in \mathbb{N}$. Thus, $\phi(\tilde{\alpha}^i) = x^0 + v^i$ for all

$i \in \mathbb{N}$, and with a first degree Taylor approximation ([Kön04], p. 66) of ϕ we obtain

$$\begin{aligned}\frac{v}{\|v\|} &= \lim_{i\to\infty} \frac{v^i}{\|v^i\|} = \lim_{i\to\infty} \frac{\phi(\tilde{\alpha}^i) - \phi(\tilde{\alpha}^0)}{\|v^i\|} = \lim_{i\to\infty} \frac{\phi(\tilde{\alpha}^0 + (\tilde{\alpha}^i - \tilde{\alpha}^0)) - \phi(\tilde{\alpha}^0)}{\|v^i\|} \\ &= \lim_{i\to\infty} \frac{\phi(\tilde{\alpha}^0) + D\phi(\tilde{\alpha}^0)(\tilde{\alpha}^i - \tilde{\alpha}^0) + o(\|\tilde{\alpha}^i - \tilde{\alpha}^0\|) - \phi(\tilde{\alpha}^0)}{\|v^i\|} \\ &= \lim_{i\to\infty} D\phi(\tilde{\alpha}^0)\frac{\tilde{\alpha}^i - \tilde{\alpha}^0}{\|v^i\|} + \frac{o(\|\tilde{\alpha}^i - \tilde{\alpha}^0\|)}{\|v^i\|} \\ &= \lim_{i\to\infty} \frac{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|}{\|v^i\|} \left(D\phi(\tilde{\alpha}^0)\frac{\tilde{\alpha}^i - \tilde{\alpha}^0}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|} + \frac{o(\|\tilde{\alpha}^i - \tilde{\alpha}^0\|)}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|} \right) \end{aligned} \tag{5.4}$$

with $\lim_{i\to\infty} \frac{o(\|\tilde{\alpha}^i - \tilde{\alpha}^0\|)}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|} = 0$. Since $(\frac{\tilde{\alpha}^i - \tilde{\alpha}^0}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|})_i$ is bounded, we can assume w.l.o.g. that $\lim_{i\to\infty} \frac{\tilde{\alpha}^i - \tilde{\alpha}^0}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|} = \tilde{w} \in \mathbb{R}^{k-1}$.
If $(\frac{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|}{\|v^i\|})_i$ would be unbounded, then we would have

$$0 = \lim_{i\to\infty} D\phi(\tilde{\alpha}^0)\frac{\tilde{\alpha}^i - \tilde{\alpha}^0}{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|} = D\phi(\tilde{\alpha}^0)\tilde{w},$$

since the right-hand side of (5.4) converges. Using the formula for $D\phi$ from Lemma 5.1.3, we would obtain

$$-\left(\sum_{i=j}^{k} \alpha_j^0 \nabla^2 f_j(x^0)\right)^{-1} Df(x)^\top w = 0$$

for $w := (\tilde{w}_1, \ldots, \tilde{w}_{k-1}, -\sum_{j=1}^{k-1} \tilde{w}_j)^\top \in \mathbb{R}^k$, implying $Df(x)^\top w = 0$. This is a contradiction, since we assumed that $\operatorname{rk}(Df(x^0)) = k - 1$, which means $\dim(\ker(Df(x^0)^\top)) = k - \operatorname{rk}(Df(x^0)) = 1$ and α^0 and w must be linearly independent (since $\sum_{j=1}^{k} w_j = 0$ and $\sum_{j=1}^{k} \alpha_j = 1$).
Thus, $(\frac{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|}{\|v^i\|})_i$ must be bounded and we can assume w.l.o.g. that there is some $\theta \in \mathbb{R}$ such that $\lim_{i\to\infty} \frac{\|\tilde{\alpha}^i - \tilde{\alpha}^0\|}{\|v^i\|} = \theta$. Combined with (5.4) (and (2.13)) we obtain

$$\begin{aligned} & \frac{v}{\|v\|} = \theta D\phi(\alpha^0)\tilde{w} = -\theta \left(\sum_{i=j}^{k} \alpha_j^0 \nabla^2 f_j(x^0)\right)^{-1} Df(x)^\top w \\ \Leftrightarrow \quad & \left(\sum_{i=j}^{k} \alpha_j^0 \nabla^2 f_j(x^0)\right) v + \theta\|v\| Df(x)^\top w = 0 \\ \Rightarrow \quad & DF(x^0, \alpha^0) \begin{pmatrix} v \\ \theta\|v\| w \end{pmatrix} = \begin{pmatrix} \sum_{i=j}^{k} \alpha_j^0 \nabla^2 f_j(x^0) & Df(x^0)^\top \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v \\ \theta\|v\| w \end{pmatrix} = 0. \end{aligned}$$

So

$$\begin{pmatrix} v \\ \theta\|v\| w \end{pmatrix} \in \ker(DF(x^0, \alpha^0)) = T_{(x^0,\alpha^0)}\mathcal{M}$$

and, in particular, $v \in \operatorname{pr}_x(T_{(x^0,\alpha^0)}\mathcal{M})$. This completes the proof. □

5.1.3 The boundary of the Pareto critical set

We will now use the results about tangent cones of the Pareto critical set from the previous part to show the main result of this section, stating that on the boundary of P_c, there is a KKT multiplier which is zero. As discussed in the introduction of this chapter, the topological boundary of P_c with respect to the natural topology on $\mathbb{R}^n$ is not a useful concept for the boundary in this case. Instead, we make the following definition:

Definition 5.1.14. *The set*

$$\partial_T P_c := \{x \in P_c : \operatorname{Tan}(P_{int}, x) \neq -\operatorname{Tan}(P_{int}, x)\}$$

is called the boundary of P_c.

Since this is not a standard definition, we first have to argue why $\partial_T P_c$ is actually a reasonable definition for what we mean by the "boundary" of the Pareto critical set. In words, $\partial_T P_c$ consists of all points $x \in P_c$ that have a tangent vector v (with respect to P_{int}) whose negation $-v$ is not a tangent vector. Roughly speaking, this means that we can reach x from within P_c on a path with direction v, which cannot be smoothly continued after reaching x. In a visual sense, this path hits the "boundary" of P_c. A more formal argument is given in the following remark.

Remark 5.1.15. *Assume that for all $x^0 \in P_0$ there is some $\alpha^0 \in A(x^0)$ such that the weighted Hessian is regular, so that $\operatorname{Tan}(P_{int}, x) = \operatorname{Tan}(P_c, x)$ for all $x \in P_c$ by Lemma 5.1.9.*

a) If the interior P_c° is nonempty and $x \in P_c^\circ$, then it is easy to see that

$$\operatorname{Tan}(P_{int}, x) = \operatorname{Tan}(P_c, x) = \mathbb{R}^n.$$

In particular, $x \notin \partial_T P_c$, so $\partial_T P_c \subseteq P_c \setminus P_c^\circ = \overline{P_c} \setminus P_c^\circ = \partial P_c$.

b) If $x \in P_c$ and P_c is a manifold around x, then $\operatorname{Tan}(P_c, x)$ is related to the tangent space of P_c at x. For example, by Lyusternik's Theorem ([Gül10], Theorem 2.29), if P_c is a level set as in Theorem 2.2.9 (locally around x), then the tangent cone $\operatorname{Tan}(P_c, x)$ is equal to the tangent space of P_c at x. Since the tangent space has the structure of a vector space, this implies $x \notin \partial_T P_c$.

From now on, the "boundary" of P_c will always refer to $\partial_T P_c$, and we will use the term "topological boundary" to refer to the topological boundary of ∂P_c as a subset of $\mathbb{R}^n$. By combining Lemma 5.1.10 and 5.1.13, we can now show the main result of this section.

Theorem 5.1.16. *Let $x^0 \in \partial_T P_c$ and $\alpha^0 \in A(x^0)$ such that $\sum_{i=1}^k \alpha_i^0 \nabla^2 f_i(x^0)$ is regular. Then there is some $\alpha \in A(x^0)$ such that $\alpha_j = 0$ for some $j \in \{1, \dots, k\}$.*

Proof. Assume that the assertion does not hold, so $\alpha_i > 0$ for all $\alpha \in A(x^0)$ and $i \in \{1, \dots, k\}$. Then $x^0 \in P_{int}$ and we can apply Lemma 5.1.11 a) and c) to see that $\operatorname{rk}(Df(x)) = k - 1$ and that α^0 is the unique KKT vector of x^0. By applying Lemma 5.1.10 and 5.1.13, we obtain

$$\operatorname{Tan}(P_{int}, x^0) = \operatorname{pr}_x(T_{(x^0, \alpha^0)}\mathcal{M}).$$

Since $\operatorname{pr}_x(T_{(x^0,\alpha^0)}\mathcal{M})$ is a linear subspace of $\mathbb{R}^n$, this implies

$$\operatorname{Tan}(P_{int}, x^0) = \operatorname{pr}_x(T_{(x^0,\alpha^0)}\mathcal{M}) = -\operatorname{pr}_x(T_{(x^0,\alpha^0)}\mathcal{M}) = -\operatorname{Tan}(P_{int}, x^0),$$

which is a contradiction to $x^0 \in \partial_T P_c$. □

By the previous theorem, all points on the boundary $\partial_T P_c$ of P_c with a regular weighted Hessian are still Pareto critical if we ignore one of the objectives. In case f satisfies the rank assumption on P_{int}, it implies the following, slightly stronger result.

Corollary 5.1.17. *If* $\operatorname{rk}(Df(x)) = k-1$ *and* $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ *is regular for all* $x \in P_{int}$ *and* $\alpha \in A(x)$*, then* $\partial_T P_c \subseteq P_0$*.*

To show what may occur if the weighted Hessian is not regular, consider the following example.

Example 5.1.18. *Consider the problem* $\min_{x\in\mathbb{R}^2} f(x)$ *with*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{cases} \begin{pmatrix} x_1^p + (x_2-1)^p \\ x_1^p + (x_2+1)^p \end{pmatrix}, & \text{if } x_1 \geq 0, \\ \begin{pmatrix} (x_2-1)^p \\ (x_2+1)^p \end{pmatrix}, & \text{if } x_1 < 0, \end{cases}$$

for even $p \in \mathbb{N}$*,* $p > 2$*. Then* f *is* $(p-1)$*-times continuously differentiable and*

$$\nabla f_1(x) = \begin{cases} (px_1^{p-1}, p(x_2-1)^{p-1})^\top, & \text{if } x_1 \geq 0, \\ (0, p(x_2-1)^{p-1})^\top, & \text{if } x_1 < 0, \end{cases}$$
$$\nabla f_2(x) = \begin{cases} (px_1^{p-1}, p(x_2+1)^{p-1})^\top, & \text{if } x_1 \geq 0, \\ (0, p(x_2+1)^{p-1})^\top, & \text{if } x_1 < 0. \end{cases}$$

For $x_1 \geq 0$*,* $x \in P_c$ *is equivalent to the existence of* $\alpha \in \Delta_2$ *such that*

$$\begin{aligned} \alpha_1 \nabla f_1(x) + \alpha_2 \nabla f_2(x) = 0 &\Leftrightarrow \begin{pmatrix} \alpha_1 p x_1^{p-1} + \alpha_2 p x_1^{p-1} \\ \alpha_1 p (x_2-1)^{p-1} + \alpha_2 p (x_2+1)^{p-1} \end{pmatrix} = 0 \\ &\Leftrightarrow x_1 = 0 \text{ and } \operatorname{sgn}(x_2+1) \neq \operatorname{sgn}(x_2-1) \\ &\Leftrightarrow x \in \{0\} \times [-1,1]. \end{aligned}$$

(Here the convention $\operatorname{sgn}(0) = 0$ *was used.) Analogously, for* $x_1 < 0$*,* $x \in P_c$ *is equivalent to* $x \in \mathbb{R}^{<0} \times [-1,1]$*. For both cases, the (unique) corresponding KKT vector* $\alpha \in \Delta_2$ *is given by*

$$\alpha = \left(\frac{(x_2+1)^{p-1}}{(x_2+1)^{p-1} - (x_2-1)^{p-1}}, -\frac{(x_2-1)^{p-1}}{(x_2+1)^{p-1} - (x_2-1)^{p-1}} \right)^\top.$$

Thus, we have

$$\begin{aligned} P_c &= \mathbb{R}^{\leq 0} \times [-1,1], \\ P_{int} &= \mathbb{R}^{\leq 0} \times (-1,1), \\ P_0 &= (\mathbb{R}^{\leq 0} \times \{-1\}) \cup (\mathbb{R}^{\leq 0} \times \{1\}), \\ \partial_T P_c &= (\mathbb{R}^{\leq 0} \times \{-1\}) \cup (\mathbb{R}^{\leq 0} \times \{1\}) \cup (\{0\} \times [-1,1]), \end{aligned}$$

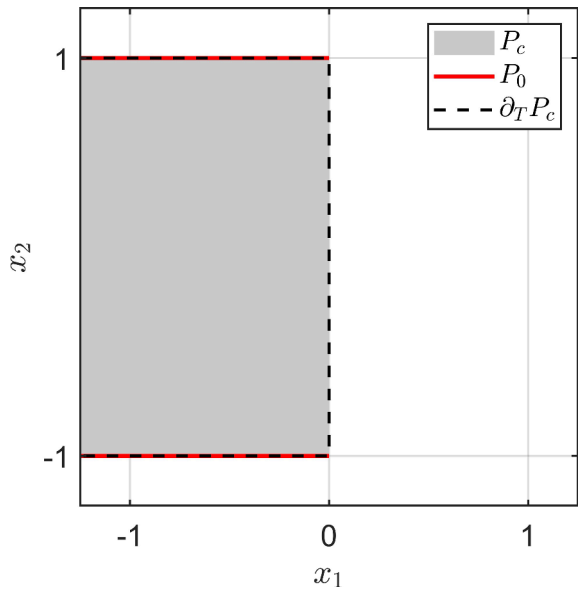


Figure 5.5: The sets P_{int}, P_0 and $\partial_T P_c$ in Example 5.1.18.

as shown in Figure 5.5. For $x \in \{0\} \times (-1,1) \subseteq \partial_T P_c$ *we have* $x \notin P_0$. *Since* $rk(Df(x)) = k-1$ *for all* $x \in P_c$, *this means that the assertion of Corollary 5.1.17 (and Theorem 5.1.16) does not hold. The reason for this can be seen by considering the Hessian matrices of* f_1 *and* f_2. *For* $x \in \{0\} \times (-1,1)$, *they are given by*

$$\nabla^2 f_1(x) = \begin{pmatrix} 0 & 0 \\ 0 & p(p-1)(x_2-1)^{p-2} \end{pmatrix}, \quad \nabla^2 f_2(x) = \begin{pmatrix} 0 & 0 \\ 0 & p(p-1)(x_2+1)^{p-2} \end{pmatrix}.$$

Clearly, the weighted Hessians cannot be regular.

5.2 Decomposing an MOP into lower-dimensional subproblems

In the previous section, we showed that points on the boundary $\partial_T P_c$ of the Pareto critical set (with a regular weighted Hessian) possess a KKT multiplier which is zero. Thus, if we unite the Pareto critical sets of all subproblems of the original MOP where one objective function is ignored, then we obtain a covering of $\partial_T P_c$. Unfortunately, this covering is not necessarily tight. For example, if $k > n+1$, then we have $\operatorname{rk}(Df(x)) < k-1$ for all $x \in P_c$. By Corollary 5.1.12 b), this implies that the union of the Pareto critical sets of the subproblems contains *every* Pareto critical point of the original MOP instead of just the boundary points. In such a case, more than one objective function can be neglected for each subproblem to still obtain a covering of $\partial_T P_c$. In this section, we will show that the number of objective functions that can be neglected is related to the rank of the Jacobian Df of the objective vector.

We begin by formalizing what is meant by the term "subproblem". To this end, denote by $\mathcal{P}(\{1,\dots,k\})$ the *power set* of $\{1,\dots,k\}$, i.e., the set of all possible subsets of $\{1,\dots,k\}$ (including $\{1,\dots,k\}$ itself). For $\emptyset \neq I \in \mathcal{P}(\{1,\dots,k\})$ with $I = \{i_1,\dots,i_{|I|}\}$, $i_1 < \dots < i_{|I|}$, let

$$f^I : \mathbb{R}^n \to \mathbb{R}^{|I|}, \quad x \mapsto (f_{i_1}(x),\dots,f_{i_{|I|}}(x))^\top.$$

Then the *subproblem corresponding to* I is the MOP

$$\min_{x\in\mathbb{R}^n} f^I(x). \tag{MOP^I}$$

Let P_c^I be its Pareto critical set and let P_{int}^I and P_0^I be defined as in Definition 5.1.1 for (MOP^I). Let $A^I(x) \subseteq \Delta_{|I|}$ be the set of KKT vectors of x for (MOP^I). Since we only consider nonempty sets I for (MOP^I), we will tacitly assume $I \neq \emptyset$ for any $I \in \mathcal{P}(\{1,\dots,k\})$ for ease of notation.

The following lemma shows that every Pareto critical point of a subproblem is also Pareto critical for the original problem.

Lemma 5.2.1. *It holds* $P_c^I \subseteq P_c$ *for all* $I \in \mathcal{P}(\{1,\dots,k\})$.

Proof. Let $I = \{i_1,\dots,i_{|I|}\}$ and $x \in P_c^I$. Then there is some $\alpha^I \in \Delta_{|I|}$ with

$$\alpha_1^I \nabla f_{i_1}(x) + \ldots + \alpha_{|I|}^I \nabla f_{i_{|I|}}(x) = 0.$$

Define

$$\alpha_i := \begin{cases} \alpha_j^I, & \text{if } i = i_j, \\ 0, & \text{otherwise.} \end{cases}$$

Then, by construction, $\alpha \in \Delta_k$ and $Df(x)^\top \alpha = 0$, so $x \in P_c$. □

Remark 5.2.2. *a) Analogously to Lemma 5.2.1, it holds*

$$P_c^{I_1} \subseteq P_c^{I_2}$$

for $I_1, I_2 \in \mathcal{P}(\{1,\dots,k\})$ *with* $I_1 \subseteq I_2$.

b) Lemma 5.2.1 still holds if we replace Pareto criticality by weak Pareto optimality (cf. Definition 2.1.4): Let x *be weakly Pareto optimal for* (MOP^I)*, i.e.,*

$$\nexists y \in \mathbb{R}^n : f_i(y) < f_i(x) \quad \forall i \in I.$$

Clearly, this still holds if we replace I *by* $\{1,\dots,k\}$*, such that* x *is weakly Pareto optimal for the original problem.*

c) Lemma 5.2.1 does not hold if we replace Pareto criticality by Pareto optimality. To see this, consider the problem

$$\min_{x\in\mathbb{R}} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} 0 \\ x^2 \end{pmatrix}.$$

Then any $x \in \mathbb{R}$ *is Pareto optimal for* (MOP^I) *with* $I = \{1\}$*, but only* $0 \in \mathbb{R}$ *is Pareto optimal for the original problem.*

Throughout Section 5.1, we saw that the rank assumption (cf. Definition 3.1.2) is an important property of the objective vector when investigating the structure of the Pareto critical set. If the rank assumption is violated in some $x^0 \in \mathbb{R}^n$, then the image of $Df(x^0)^\top$, i.e., the span of the gradients of the k objectives, has a dimension less than $k-1$. This implies that we can remove one of the objectives to obtain a subproblem (MOP^I) with $|I| < k$ and $\operatorname{rk}(Df^I(x^0)) = \operatorname{rk}(Df(x^0)) \leq k-1$. The following lemma shows that we can choose $I \in \mathcal{P}(\{1,\dots,k\})$ in a way that additionally ensures $x^0 \in P_c^I$ and that the removal of indices can be repeated until the rank assumption is satisfied for the subproblem (MOP^I).

Lemma 5.2.3. *Let $x^0 \in P_c$. Then there is some $I \in \mathcal{P}(\{1,\dots,k\})$ such that*

(i) *x_0 satisfies the rank assumption for f^I, i.e., $\operatorname{rk}(Df^I(x^0)) = |I| - 1$,*

(ii) *$\operatorname{rk}(Df^I(x^0)) = \operatorname{rk}(Df(x^0))$,*

(iii) *$x^0 \in P_c^I$.*

Furthermore, if $x^0 \in P_0$ then $x^0 \in P_0^I$.

Proof. Let $x^0 \in P_c$. If $\operatorname{rk}(Df(x^0)) = k-1$ then we can choose $I = \{1,\dots,k\}$. So assume that $\operatorname{rk}(Df(x^0)) < k-1$. Let

$$J := \left\{ j \in \{1,\dots,k\} : \operatorname{rk}(Df^{\{1,\dots,k\}\setminus\{j\}}(x^0)) < \operatorname{rk}(Df(x^0)) \right\}$$

be the set of indices of gradients of objectives which cannot be expressed as a linear combination of other gradients. Then $\operatorname{rk}(Df^J(x^0)) = |J|$ and since we assumed $\operatorname{rk}(Df(x^0)) < k-1$, we must have $J \subsetneq \{1,\dots,k\}$. Let $K := \{1,\dots,k\} \setminus J \neq \emptyset$. Since we must have $\alpha_j = 0$ for all $j \in J$ and $\alpha \in A(x^0)$, it follows that $x^0 \in P_c^K$. Furthermore,

$$\begin{aligned} & k-1 > \operatorname{rk}(Df(x^0)) = \operatorname{rk}(Df^K(x^0)) + |J| \\ \Leftrightarrow \quad & \operatorname{rk}(Df^K(x^0)) < k - |J| - 1 = |K| - 1. \end{aligned}$$

Thus, application of Corollary 5.1.12 to f^K in x^0 yields the existence of $\alpha' \in A(x^0)$ and $l \in K$ with $\alpha'_l = 0$. Let $I := \{1,\dots,k\} \setminus \{l\}$. Since $l \notin J$ we must have $\operatorname{rk}(Df^I(x^0)) = \operatorname{rk}(Df(x^0))$, i.e., (ii) holds. Furthermore, (iii) holds since $\alpha'_l = 0$. If $|I| - 1 = k - 2 = \operatorname{rk}(Df(x^0))$, then (i) holds as well and we are done. Otherwise, we make the same construction as above for objective vector f^I. Since the size of the index set I decreases every time we repeat this process, we find an index set satisfying (i) after a finite number of repetitions.
Now assume that $x^0 \in P_0$. We will show that after we set $I = \{1,\dots,k\} \setminus \{l\}$ as above, there is still some $l' \in \{1,\dots,|I|\}$ with $\alpha_{l'} = 0$ for all $\alpha \in A^I(x^0)$, such that $x^0 \in P_0^I$. Note that if there are two different $l^1, l^2 \in \{1,\dots,k\}$ with $\alpha_{l^1} = \alpha_{l^2} = 0$ for all $\alpha \in A(x^0)$ then there is nothing to show, since only one index is removed from $\{1,\dots,k\}$. So assume that there is a unique $l' \in \{1,\dots,k\}$ with $\alpha_{l'} = 0$ for all $\alpha \in A(x^0)$. For l' to be removed from $\{1,\dots,k\}$ via the above procedure, we would have to have $l' \in K$, i.e.,

$$\operatorname{rk}(Df^{\{1,\dots,k\}\setminus\{l'\}}(x^0)) = \operatorname{rk}(Df(x^0)). \tag{5.5}$$

This implies that there is some $\beta \in \mathbb{R}^k$ with $\beta_{l'} \neq 0$ such that $Df(x^0)^\top \beta = 0$. W.l.o.g. assume $\beta_{l'} > 0$. From the uniqueness of l' (and the structure of $A(x^0)$) it follows that there is some $\alpha' \in A(x^0)$ with $\alpha'_j > 0$ for all $j \in \{1,\dots,k\}$, $j \neq l'$. Let

$$\gamma(t) := \frac{\alpha' + t\beta}{\sum_{i=1}^k \alpha'_i + t\beta_i} = \frac{\alpha' + t\beta}{1 + t\sum_{i=1}^k \beta_i}.$$

Then by construction, there has to be some $t' > 0$ such that $\gamma(t') \in A(x^0)$ and $(\gamma(t'))_{l'} > 0$. This is a contradiction to the assumption that $\alpha_{l'} = 0$ for all $\alpha \in A(x^0)$. Thus, (5.5) cannot hold, i.e., l' cannot be removed from $\{1,\dots,k\}$. □

By applying the previous lemma in all $x \in P_c$, we immediately obtain the following corollary.

Corollary 5.2.4. *There is some $\mathcal{J} \subseteq \mathcal{P}(\{1,\dots,k\})$ with*

$$P_c = \bigcup_{I\in\mathcal{J}} P_c^I \quad \text{and} \quad P_0 \subseteq \bigcup_{I\in\mathcal{J}} P_0^I$$

such that for all $x \in P_c$, there is a subproblem corresponding to $I \in \mathcal{J}$ with $x \in P_c^I$ and $|I| = \operatorname{rk}(Df(x)) + 1$ for which x satisfies the rank assumption.

The previous corollary shows that if the rank of the Jacobian of the objective vector in all $x \in P_c$ is small compared to k, then it suffices to solve a number of subproblems with fewer objective functions to still obtain the entire Pareto critical set of the original MOP. In particular, as an upper bound for the size of the subproblems, we obtain $|I| = \operatorname{rk}(Df(x)) + 1 \leq n + 1$. Thus, due to Remark 5.2.2 a), if no knowledge about the rank of the Jacobian is available, we can just solve all subproblems of size $n + 1$ to obtain a covering. Clearly, in practical applications, we would be interested in finding the smallest set of subproblems that have to be solved. Unfortunately, Corollary 5.2.4 gives no insight into how such a smallest set of subproblems can be obtained.

Remark 5.2.5. *Recall that if all objective functions are paraboloids, i.e., if there are $c^i \in \mathbb{R}^n$, $i \in \{1,\dots,k\}$, with $f_i(x) = \|x - c^i\|_2^2$, then $P_c = \operatorname{conv}(\{c^1,\dots,c^k\})$ (cf. Example 2.1.6). In this case, Lemma 5.2.3 implies the well-known* Carathéodory's theorem *from convex geometry [DGK63]: For every $x^0 \in P_c \subseteq \mathbb{R}^n$, there is some $I \in \mathcal{P}(\{1,\dots,k\})$ with $|I| \leq n + 1$ such that $x^0 \in P_c^I = \operatorname{conv}(\{c^i : i \in I\})$.*

In the following example, Corollary 5.2.4 is applied to two simple problems.

Example 5.2.6. *a) Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \text{with} \quad f(x) = \begin{pmatrix} x_1^2 + x_2^2 \\ (x_1-1)^2 + x_2^2 \\ (x_1-1)^2 + (x_2-1)^2 \\ x_1^2 + (x_2-1)^2 \end{pmatrix}.$$

As discussed in Example 2.1.6, the Pareto critical set of this problem is given by

$$P_c = \operatorname{conv}(\{(0,0)^\top, (1,0)^\top, (1,1)^\top, (0,1)^\top\}),$$

which is the unit square in $\mathbb{R}^2$. In this case, we have

$$\operatorname{rk}(Df(x)) + 1 \leq \min(\{n,k\}) + 1 = 3$$

for all $x \in P_c \subseteq \mathbb{R}^2$. Thus, we can write P_c as the union of all P_c^I with $|I| = 3$. Note that not all such subproblems are needed to cover P_c. For example, it is sufficient to only consider the subproblems corresponding to sets in $\{\{1,2,3\},\{1,3,4\}\}$ or $\{\{1,2,4\},\{2,3,4\}\}$, respectively. The two coverings are shown in Figure 5.6.

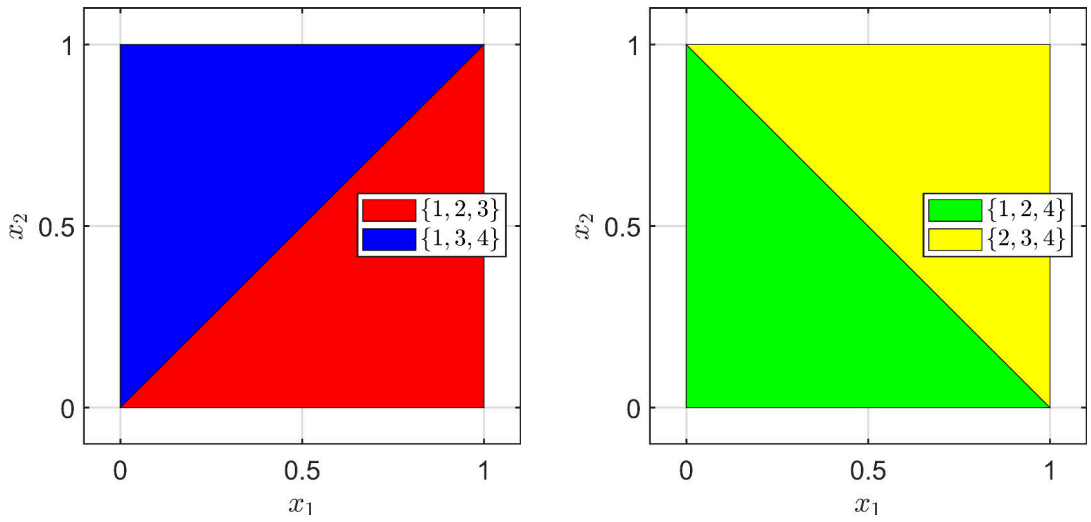


Figure 5.6: Covering of P_c for different sets of subproblems in Example 5.2.6 a).

b) Consider the problem $\min_{x\in\mathbb{R}^2} f(x)$ *with* $f_i(x) = \|x - c^i\|_{Q_i}^2$, $i \in \{1,\dots,4\}$, *for*

$$Q_1 = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, c^1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$Q_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, c^2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$
$$Q_3 = \begin{pmatrix} 5 & 0 \\ 0 & 1 \end{pmatrix}, c^3 = \begin{pmatrix} 0.4 \\ 0.3 \end{pmatrix},$$
$$Q_4 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}, c^4 = \begin{pmatrix} 0.7 \\ 0.7 \end{pmatrix}.$$

The Pareto critical P_c set of this problem can be computed as in Example 2.2.8 and is shown as the gray set in Figure 5.7. As in a), it is sufficient to consider all subproblems with $|I| = 3$ objective functions to cover P_c. The corresponding Pareto critical sets are shown in Figure 5.7. We see that in contrast to a), there is no covering of P_c with two or less subproblems. Instead, we have to choose any combination of three out of the four subproblems to obtain a covering. For example, Figure 5.8 shows a covering of P_c for the choice $\{\{1,2,3\},\{1,2,4\},\{1,3,4\}\}$.

So far, we have analyzed how the complete Pareto critical set P_c can be covered by Pareto critical sets of smaller subproblems. Obviously, these subproblems also cover the boundary $\partial_T P_c$ of P_c, but not in a tight way. In the following, we will analyze how a tight covering of the boundary of P_c can be obtained.

Lemma 5.2.7. *Let $x^0 \in \partial_T P_c$ and let $\mathcal{J} \subseteq \mathcal{P}(\{1,\dots,k\})$ be a set of index sets as in Corollary 5.2.4. Then there is some $I \in \mathcal{J}$ such that*

(a) there is some $\alpha^I \in A^I(x^0)$ with $\alpha_i^I = 0$ for some $i \in \{1,\dots,|I|\}$ or

(b) $x^0 \in \partial_T P_c^I$.

Proof. We distinguish between two cases:
Case 1: $\mathrm{Tan}(P_{int}, x^0) \neq \mathrm{Tan}(P_c, x^0)$. Since, by definition, $P_c \setminus P_{int} = P_0$, there must

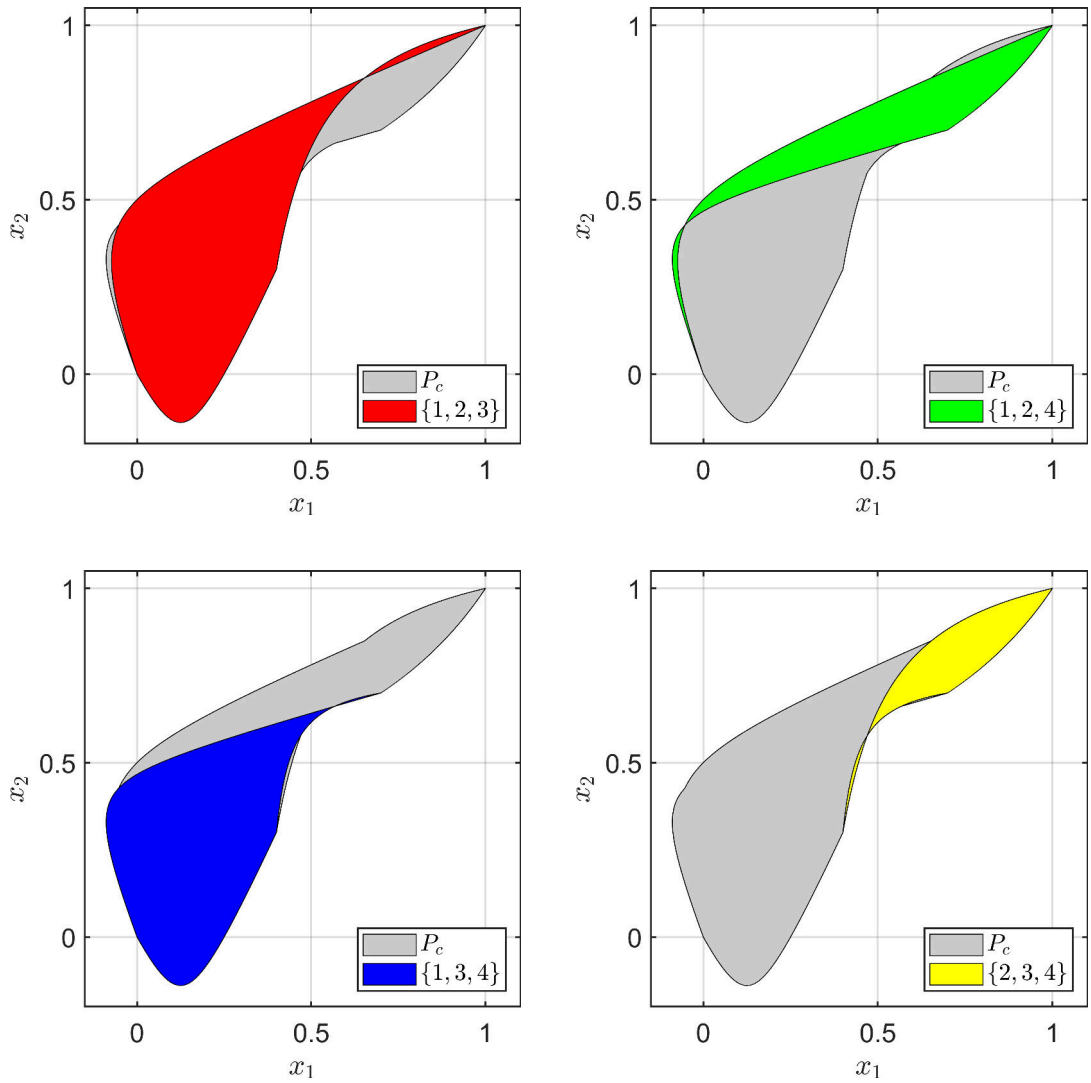


Figure 5.7: Pareto critical sets of the original problem and all subproblems with $|I| = 3$ in Example 5.2.6 b).

be a sequence $(v^i)_i \in \mathbb{R}^n \setminus \{0\}$ (as in Definition 5.1.8 of the tangent cone) with $\lim_{i\to\infty} v^i = 0$ and

$$x^0 + v^i \in P_0 \subseteq \bigcup_{I\in\mathcal{J}} P_0^I.$$

Since $|\mathcal{J}|$ is finite, we can assume w.l.o.g. that $x^0 + v^i \in P_0^I$ for a fixed $I \in \mathcal{J}$ and all $i \in \mathbb{N}$. In particular we have $x^0 \in \overline{P_0^I}$, which implies (a) (due to continuity of Df).
Case 2: $\operatorname{Tan}(P_{int}, x^0) = \operatorname{Tan}(P_c, x^0)$. Let $v \in \operatorname{Tan}(P_c, x^0) = \operatorname{Tan}\left(\bigcup_{I\in\mathcal{J}} P_c^I, x^0\right)$ and let $(v^i)_i$ be the corresponding sequence with $x^0 + v^i \in \bigcup_{I\in\mathcal{J}} P_c^I$ for all $i \in \mathbb{N}$. Since $|J|$ is finite, there must be a fixed $I \in \mathcal{J}$ with $x^0 + v^i \in P_c^I$ infinitely many times, such that $v \in \operatorname{Tan}(P_c^I, x^0)$. Furthermore, $P_c^I \subseteq P_c$ implies $\operatorname{Tan}(P_c^I, x^0) \subseteq \operatorname{Tan}(P_c, x^0)$ for all $I \in \mathcal{J}$. Thus, we have

$$\bigcup_{I\in\mathcal{J}} \operatorname{Tan}(P_c^I, x^0) = \operatorname{Tan}(P_c, x^0) = \operatorname{Tan}(P_{int}, x^0). \tag{5.6}$$

Since $x^0 \in \partial_T P_c$ we have $\operatorname{Tan}(P_{int}, x^0) \neq -\operatorname{Tan}(P_{int}, x^0)$, so (5.6) implies that there is some $I \in \mathcal{J}$ with $\operatorname{Tan}(P_c^I, x^0) \neq -\operatorname{Tan}(P_c^I, x^0)$. If $\operatorname{Tan}(P_c^I, x^0) = \operatorname{Tan}(P_{int}^I, x^0)$

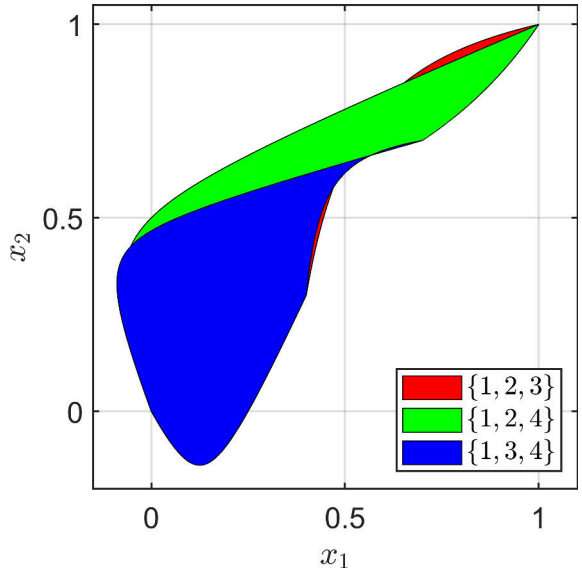


Figure 5.8: A possible covering of P_c in Example 5.2.6 b).

then $x^0 \in \partial_T P_c^I$, so (b) holds. Otherwise, it follows as in case 1 that $x^0 \in \overline{P_0^I}$, implying (a). □

The following corollary shows that if the weighted Hessian of the subproblem in the previous lemma is regular, then (a) must hold.

Corollary 5.2.8. *Let $x^0 \in \partial_T P_c$ and let $I = \{i_1, \ldots, i_{|I|}\} \in \mathcal{P}(\{1, \ldots, k\})$ be as in Lemma 5.2.7. If there is some $\alpha^0 \in A^I(x^0)$ such that the weighted Hessian $\sum_{j=1}^{|I|} \alpha_j^0 \nabla^2 f_{i_j}(x^0)$ of the subproblem is regular, then there is some $\alpha \in A^I(x^0)$ with $\alpha_j = 0$ for some $j \in \{1, \ldots, |I|\}$.*

Proof. If $\operatorname{rk}(Df^I(x^0)) < |I| - 1$, then we can apply Corollary 5.1.12 to obtain the result. Otherwise we must have $\operatorname{rk}(Df^I(x^0)) = |I| - 1$, in which case the result follows from Theorem 5.1.16 (applied to the subproblem corresponding to I). □

By applying Corollary 5.2.8 to all $x^0 \in \partial_T P_c$ and making a stronger assumption on the weighted Hessians, we obtain the following result.

Corollary 5.2.9. *Let $m = \max_{x \in P_c} \operatorname{rk}(Df(x))$. If $m > 0$ and $\sum_{i=1}^k \alpha_i \nabla^2 f_i(x)$ is regular for all $x \in \partial_T P_c$ and $\alpha \in A(x)$, then*

$$\partial_T P_c \subseteq \bigcup_{I \in \mathcal{P}(\{1,\ldots,k\}), |I|=m} P_c^I.$$

Since the rank of the Jacobian is bounded by the number of variables n, we only have to consider subproblems of size (at most) n to obtain a covering of the boundary of the Pareto critical set in Corollary 5.2.9. We will demonstrate this via several examples in the following section.

5.3 Examples

In this section, we will consider some examples to show how the results from the previous section can be used to analyze the structure of Pareto critical sets. We will

begin by revisiting the MOPs in Example 5.2.6. Due to their strict convexity, both problems satisfy the assumptions of Corollary 5.2.9.

Example 5.3.1. *a) Consider the MOP from Example 5.2.6 a). By Corollary 5.2.9 and since* $rk(Df(x)) \leq min(\{n, k\}) = 2$ *for all* $x \in \mathbb{R}^2$*, we only have to consider the subproblems of size* $|I| = 2$ *to obtain a covering of the boundary of* P_c*, as shown in Figure 5.9(a). As in Example 5.2.6 a), not all subproblems are*

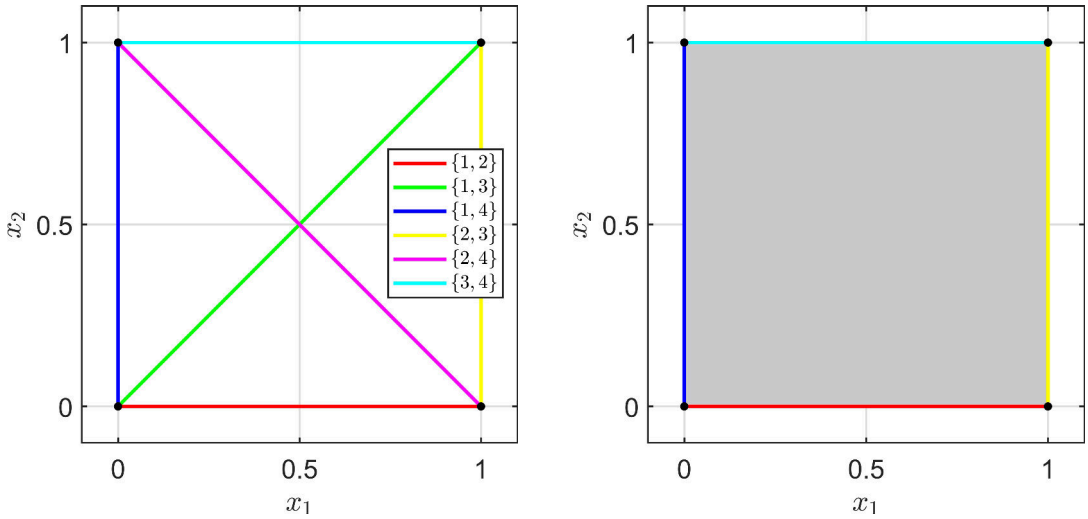


Figure 5.9: **(a)** All subproblems of size $|I| = 2$ in Example 5.3.1 a). The black dots show the critical points of the individual objective functions. **(b)** The Pareto critical set (gray) and a possible covering of its boundary.

needed for this. In this case, a (perfect) covering of $\partial_T P_c$ *can be obtained via the subproblems corresponding to the index sets* $\{\{1, 2\}, \{1, 4\}, \{2, 3\}, \{3, 4\}\}$*, as shown in Figure 5.9(b).*

b) Consider the MOP from Example 5.2.6 b). As in a), we only have to consider the subproblems of size $|I| = 2$ *to obtain a covering of* $\partial_T P_c$*. The corresponding Pareto critical sets are shown in Figure 5.10(a). We see that in contrast to a), all subproblems are needed to cover* $\partial_T P_c$ *and the covering is only a superset of* $\partial_T P_c$*. Figure 5.10(b) shows the boundary of* P_c*, where each boundary point is given a color depending on the subproblem for which it is Pareto critical. We see that kinks in the boundary arise whenever the Pareto critical sets of two subproblems intersect, i.e., whenever a boundary point is Pareto critical for more than one subproblem.*

While the assumption of the regularity of the weighted Hessian for all KKT vectors in Corollary 5.2.9 is strong (and hard to verify in practice), the premise of Corollary 5.2.8 is much weaker. Without giving a proof (or a precise statement), we expect that for a "generic" objective vector f, regularity of the weighted Hessian $\sum_{i=1}^{k} \alpha_i \nabla^2 f(x)$ for a single KKT vector $\alpha \in A(x)$ is a "generic" property in x. In other words, we expect that points where this assumption is violated can be ignored in practice.

In Example 5.3.1, both MOPs were strictly convex. In the following, we will consider nonconvex MOPs where the Pareto critical sets have some additional features.

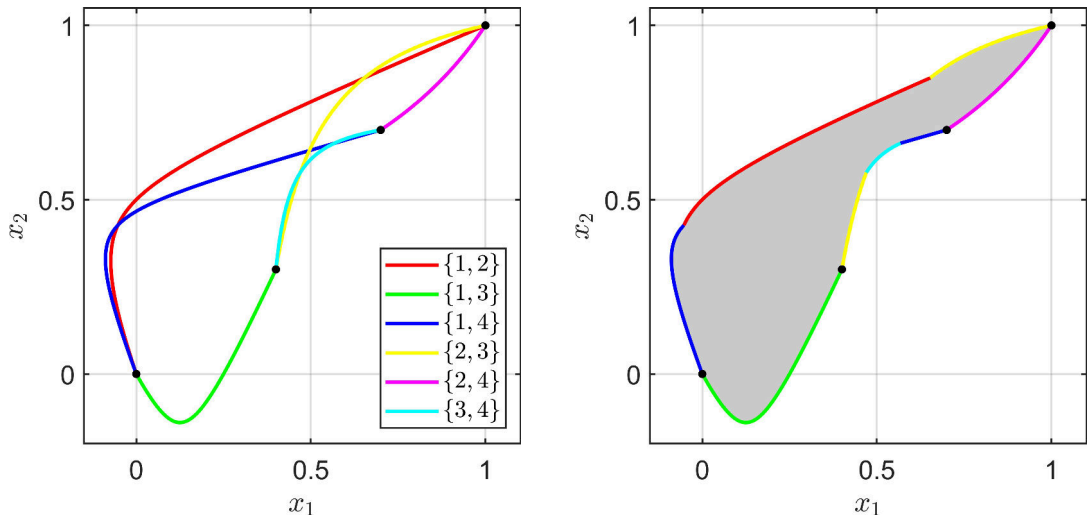


Figure 5.10: **(a)** All subproblems of size $|I| = 2$ in Example 5.3.1 b). The black dots show the critical points of the individual objective functions. **(b)** The Pareto critical set (gray) with its boundary colored depending on the subproblems.

Example 5.3.2. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} -6x_1^2 + x_1^4 + 3x_2^2 \\ (x_1 - \frac{1}{2})^2 + 2(x_2 - 1)^2 \\ (x_1 - 1)^2 + 2(x_2 - \frac{1}{2})^2 \end{pmatrix},$$

which was also considered in [Pei17], Example 4.1.5. The Pareto critical set of this problem is shown in Figure 5.11(a) and consists of two connected components. The

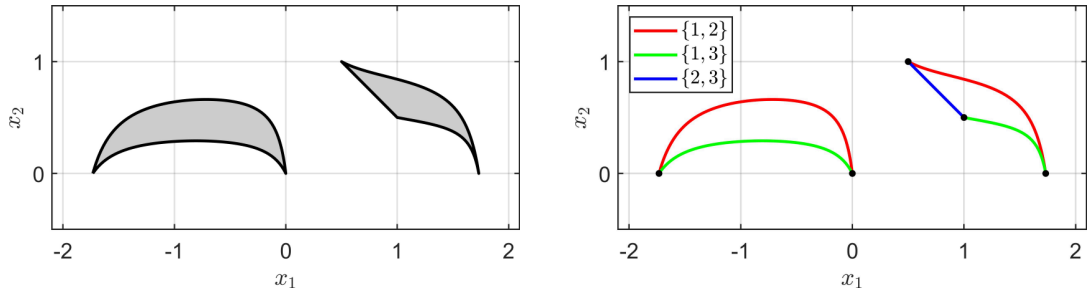


Figure 5.11: **(a)** Pareto critical set of the MOP in Example 5.3.2. **(b)** The Pareto critical sets of all subproblems with $|I| = 2$ objective functions. The black dots show the critical points of the individual objective functions.

right component is a curved triangle and the left component is a curved digon (i.e., a curved polygon with two sides and two corners). Figure 5.11(b) shows the Pareto critical sets of all subproblems with $|I| = 2$. We see that the left component only has two sides and two corners because its boundary consists of the Pareto critical sets of only two subproblems.

Example 5.3.3. *Consider the problem*

$$\min_{x\in\mathbb{R}^2} f(x) \quad \textit{with} \quad f(x) = \begin{pmatrix} \frac{1}{2}(x_1-1)^2 + x_2^2 \\ 2(x_1+\frac{1}{2})^2 + 2(x_2-1)^2 \\ 2(x_1+1)^2 + \frac{7}{2}(x_2-\frac{1}{10})^3 \\ -2x_1^3 + 2(x_2+\frac{1}{4})^2 \end{pmatrix}.$$

The Pareto critical set of this problem is shown in Figure 5.12(a). (Since the Pareto critical set is unbounded, only the relevant part is depicted.) In contrast to our

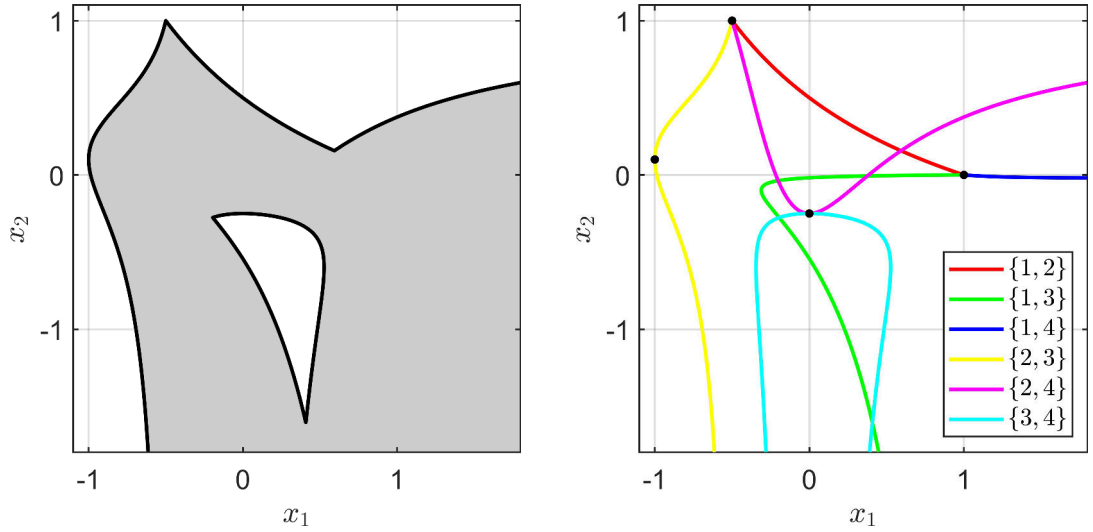


Figure 5.12: **(a)** Pareto critical set of the MOP in Example 5.3.3. **(b)** The Pareto critical sets of all subproblems with $|I| = 2$ objective functions. The black dots show the critical points of the individual objective functions.

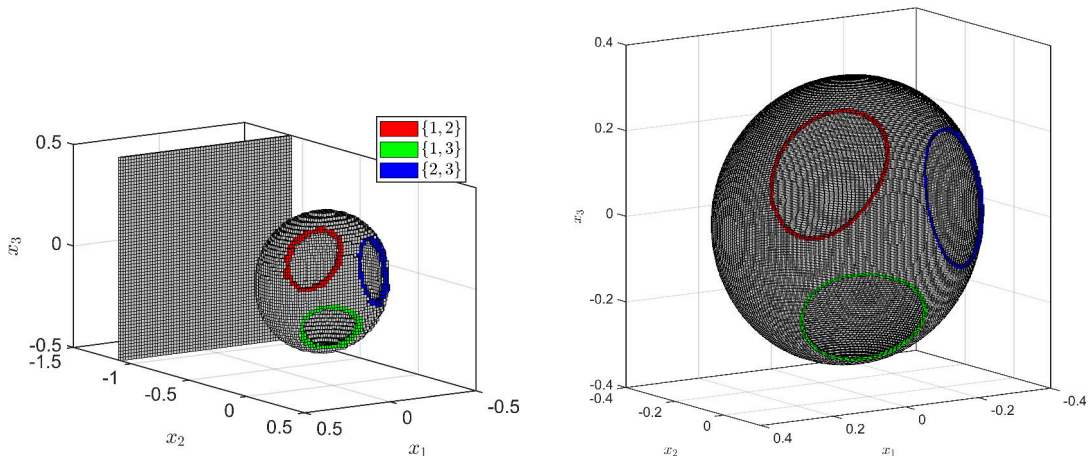
previous examples, we see that the Pareto critical set possesses a hole (i.e., it is not simply connected). This structure can be analyzed by considering the Pareto critical sets of all subproblems with $|I| = 2$, as shown in Figure 5.12(b). We see that the hole is given by the area that is enclosed by the subproblems corresponding to $\{1, 3\}$ and $\{3, 4\}$.

So far, in all examples in this section, we had $k > 2$ objective functions in $n = 2$ variables, i.e., $k \geq n + 1$. Due to this, the Pareto critical sets were of the same "dimension" as the variable space itself and $\partial_T P_c$ coincided with the topological boundary. Next, we will consider an example where this is not the case.

Example 5.3.4. *Consider the problem $L\&H_{3\times 3}$ from Example 3.1.8, where have $k = n = 3$. Since $Df(x)$ cannot have full rank for $x \in P_c$, we must have*

$$\operatorname{rk}(Df(x)) < \min(\{k, n\}) = 3.$$

Thus, by Corollary 5.2.8, it is sufficient to consider all problems of size $|I| = 2$ to cover the boundary $\partial_T P_c$ of P_c. The corresponding Pareto critical sets and the original Pareto critical set (computed via the continuation method) are shown in Figure 5.13. As expected, the Pareto critical sets of the subproblems cover the "edges" of the holes in the spherical component of P_c. (Note that Figure 5.13(a) only shows



(a) **(b)**

Figure 5.13: **(a)** The Pareto critical sets of all subproblems with $|I| = 2$ and the original Pareto critical set for Example 5.3.4. **(b)** Same as (a), but with a finer box covering of P_c and zoomed in on the spherical component.

$P_c \cap ([-0.5, 0.5] \times [-1.5, 0.5] \times [-0.5, 0.5])$. *The two-dimensional surface in the back expands further outside of this region, which is why its boundary in Figure 5.13(a) is not covered by subproblems.)*

5.4 Extension to constrained MOPs

In this section, we will discuss the generalization of the results from Sections 5.1 and 5.2 about the boundary of Pareto critical sets to the case of constrained MOPs. Recall that in the constrained case, the Pareto critical set P_c was defined in Definition 3.1.10 as the set of points satisfying the condition (3.5).

There are two different approaches for generalizing our previous results to the constrained case: The first, straight-forward approach is to consider all theoretical results from Sections 5.1 and 5.2 and adapt each result to the constrained version of the KKT condition. The second approach is to transform the constrained MOP into an unconstrained MOP by interpreting the constraints as additional objective functions (cf. [KJ06]), and to then apply our results from the unconstrained case to the resulting MOP. Here, we will consider the latter of the two approaches, as it gives an interesting general view on constrained MOPs. We will only develop the basic ideas behind this approach instead of an exhaustive theory, as the constrained case will turn out to be more complicated than the unconstrained case. In particular, we will omit a rigorous definition of the "boundary" of P_c in the constrained case and simply define

$$\partial_T P_c := \{x \in P_c : \operatorname{Tan}(P_c, x) \neq -\operatorname{Tan}(P_c, x)\}. \tag{5.7}$$

This definition and the definition of the boundary in the unconstrained case (Definition 5.1.14) are slightly inconsistent because of the usage of P_{int} in the unconstrained case. (They are consistent when Lemma 5.1.9 is applicable.) But the simpler definition (5.7) in the constrained case will be sufficient for our discussion and allows us to omit the generalization of P_{int}. We will begin by considering equality constraints and inequality constraints separately.

5.4.1 Equality constraints

Consider the MOP

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x), \\ \text{s.t.} \quad & h(x) = 0, \end{aligned} \tag{5.8}$$

where $f : \mathbb{R}^n \to \mathbb{R}^k$ and $h : \mathbb{R}^n \to \mathbb{R}^{k_h}$ are twice continuously differentiable and h satisfies the LICQ, i.e., the gradients $\nabla h_i(x)$, $i \in \{1, \dots, k_h\}$, are linearly independent for all $x \in \mathbb{R}^n$ with $h(x) = 0$. By Definition 3.1.10, a point $\bar{x} \in \mathbb{R}^n$ is Pareto critical for (5.8) if $h(\bar{x}) = 0$ and there are $\alpha \in \Delta_k$ and $\lambda \in \mathbb{R}^{k_h}$ such that

$$\sum_{i=1}^{k} \alpha_i \nabla f_i(\bar{x}) + \sum_{i=1}^{k_h} \lambda_i \nabla h_i(\bar{x}) = 0.$$

With the convention $\operatorname{sgn}(0) = 0$, this can be written as

$$\sum_{i=1}^{k} \alpha_i \nabla f_i(\bar{x}) + \sum_{i=1}^{k_h} |\lambda_i| \nabla(\operatorname{sgn}(\lambda_i) h_i)(\bar{x}) = 0. \tag{5.9}$$

Let

$$\tilde{\alpha} := \frac{1}{\sum_{i=1}^{k} \alpha_i + \sum_{i=1}^{k_h} |\lambda_i|} (\alpha, |\lambda|),$$

where $|\lambda| := (|\lambda_1|, \dots, |\lambda_{k_h}|)^\top$. Then $\tilde{\alpha} \in \Delta_{k+k_h}$ and (5.9) implies that $\bar{x}$ is Pareto critical with KKT vector $\tilde{\alpha}$ for the unconstrained MOP

$$\min_{x\in\mathbb{R}^n} \begin{pmatrix} f(x) \\ \operatorname{sgn}(\lambda) \odot h(x) \end{pmatrix},$$

where $\operatorname{sgn}(\lambda) := (\operatorname{sgn}(\lambda_1), \dots, \operatorname{sgn}(\lambda_{k_h}))^\top$ and $\odot$ is the *Hadamard product*, i.e., $\operatorname{sgn}(\lambda) \odot h(x) = (\operatorname{sgn}(\lambda_i) h_i(x))_{i\in\{1,\dots,k_h\}}$. Conversely, it is easy to see that for every $s \in \{-1, 1\}^{k_h}$, every Pareto critical point of the unconstrained MOP

$$\min_{x\in\mathbb{R}^n} \begin{pmatrix} f(x) \\ s \odot h(x) \end{pmatrix} \tag{5.10}$$

that satisfies $h(x) = 0$ is also Pareto critical for the original MOP (5.8). (Note that here we need the LICQ to assure that not all multipliers of f in the KKT conditions of (5.10) are zero.) For $s \in \{-1, 1\}^{k_h}$ let P_c^s be the Pareto critical set of (5.10). Then our previous considerations can be summarized as

$$P_c = \left(\bigcup_{s\in\{-1,1\}^{k_h}} P_c^s \right) \cap h^{-1}(0). \tag{5.11}$$

This relationship is visualized in the following example.

Example 5.4.1. *Consider the MOP* (5.8) *for*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \end{pmatrix},$$
$$h : \mathbb{R}^2 \to \mathbb{R}, \quad x \mapsto \|x - c^2\|^2 - \left(\frac{3}{8}\right)^2$$

with

$$Q = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, \; c^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; c^2 = \begin{pmatrix} \frac{5}{8} \\ \frac{1}{2} \end{pmatrix}.$$

The objective vector f was already considered in Example 2.2.8. We now constrain it to $h^{-1}(0)$, which is a circle with center c^2 and radius $\frac{3}{8}$. By (5.11)*, the Pareto critical set of this MOP is given by the intersection of $h^{-1}(0)$ with the union of the Pareto critical sets P_c^{+1} and P_c^{-1} of the MOPs*

$$\min_{x \in \mathbb{R}^2} \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \\ \|x - c^2\|^2 - \left(\frac{3}{8}\right)^2 \end{pmatrix} \quad \text{and} \quad \min_{x \in \mathbb{R}^2} \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \\ -(\|x - c^2\|^2 - \left(\frac{3}{8}\right)^2) \end{pmatrix},$$

respectively. Due to the structure of the objective vectors of these MOPs, P_c^{+1} and P_c^{-1} can be computed with the same technique as in Example 2.2.8. The result is shown in Figure 5.14.

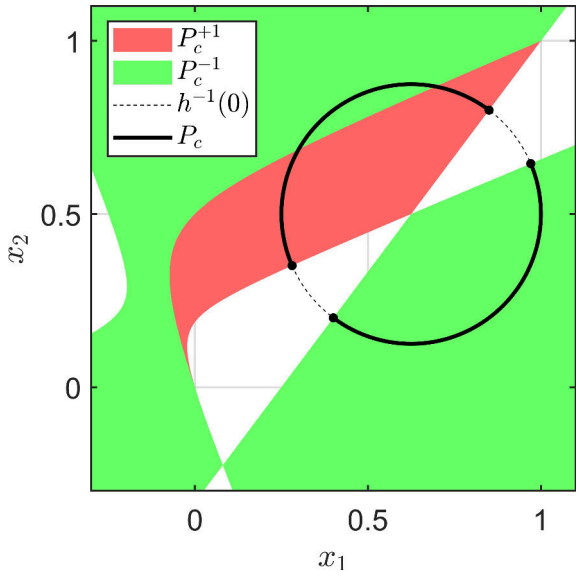


Figure 5.14: Deriving P_c by interpreting the equality constraint as an additional objective in Example 5.4.1. The black dots show the critical points of the individual objectives (as constrained scalar problems).

We will now discuss how the structure of the boundary of P_c can be derived from equality (5.11). As mentioned at the beginning of this section, we will only sketch the derivation here for the generic case and leave the details for future work. We will first consider the boundary of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$ (which can formally be defined as in (5.7) via tangent cones). To this end, let $s \in \{-1, 1\}^{k_h}$ and consider the Pareto

critical set P_c^s of the MOP (5.10). Since this is an unconstrained MOP, we can apply our results from Section 5.1 (namely Theorem 5.1.16) to see that if x lies on the boundary of P_c^s (and the assumption about the weighted Hessian is satisfied), then one of the KKT vectors $\tilde{\alpha} \in \Delta_{k+k_h}$ of x must have a zero entry $\tilde{\alpha}_j = 0$, $j \in \{1, \dots, k+k_h\}$. Due to the structure of the objective vector, this entry either belongs to f (if $j \in \{1, \dots, k\}$) or to $s \odot h$ (if $j \in \{k+1, \dots, k+k_h\}$). In the latter case, it is easy to see that x is also Pareto critical for the MOP (5.10) corresponding to $s' \in \Delta_{k+k_h}$ with

$$s'_i = \begin{cases} s_i, & \text{if } i \neq j, \\ -s_i, & \text{if } i = j, \end{cases}$$

i.e., we have $x \in P_c^{s'}$. Furthermore, by recalling Theorem 2.2.10 and equation (2.13) about the projected smooth structure of Pareto critical sets, we see that $\operatorname{Tan}(P_c^s, x) = \operatorname{Tan}(P_c^{s'}, x)$, which suggests that the union $P_c^s \cup P_c^{s'}$ is "smooth" around x. Thus, we expect that points that have a KKT multiplier which is zero for $s \odot h$ lie on the boundary of P_c^s, but not on the boundary of the union $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$. This would imply that the boundary of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$ consists of the points that have a zero multiplier with respect to f. This can also be seen in Example 5.4.1: The points in P_c^{+1} and P_c^{-1} where a KKT multiplier with respect to $s \odot h$ is zero are precisely the points in which the red and the green sets touch. Clearly, these points do not lie on the boundary of $P_c^{+1} \cup P_c^{-1}$ (unless an additional multiplier of f is zero).

Based on the previous considerations, we can now derive the structure of the boundary of P_c from the boundary of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$. To this end, we again consider equation (5.11). The LICQ implies that we can apply the level set theorem (Theorem 2.2.9) to see that the feasible set $h^{-1}(0)$ is an $(n-k_h)$-dimensional manifold. Thus, by (5.11), we expect that in the generic case, the tangent cone $\operatorname{Tan}(P_c, x)$ of P_c is the intersection of the tangent space of $h^{-1}(0)$ and the tangent cone of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$. (In the context of the intersection of tangent spaces of manifolds, this relationship is connected to transversality, cf. Theorem 6.30 in [Lee12].) Now let $x \in \partial_T P_c$, i.e., let there be some $v \in \operatorname{Tan}(P_c, x)$ with $-v \notin \operatorname{Tan}(P_c, x)$. Since the tangent space of $h^{-1}(0)$ is a vector space, this implies that $-v$ is not contained in the tangent cone of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$. Thus, x must lie on the boundary of $\bigcup_{s \in \{-1,1\}^{k_h}} P_c^s$. By our previous discussion, this implies that, as a Pareto critical point of (5.8), x possesses a KKT vector with a zero component. In other words, we expect that also in the equality-constrained case, the boundary of P_c consists of Pareto critical sets of (constrained) subproblems.

5.4.2 Inequality constraints

We will now consider inequality constrained MOPs of the form

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x), \\ \text{s.t.} \quad & g(x) \leq 0, \end{aligned} \tag{5.12}$$

where $f : \mathbb{R}^n \to \mathbb{R}^k$ and $g : \mathbb{R}^n \to \mathbb{R}^{k_g}$ are twice continuously differentiable and g satisfies the LICQ, i.e., for all $x \in g^{-1}(\mathbb{R}^{\leq 0})$, the gradients $\nabla g_i(x)$ of all active

indices $i \in \mathcal{A}_g(x) = \{i \in \{1, \dots, k_g\} : g_i(x) = 0\}$ are linearly independent. By Definition 3.1.10, a point $\bar{x} \in \mathbb{R}^n$ is Pareto critical for (5.12) if $g(\bar{x}) \leq 0$ and there are $\alpha \in \Delta_k$ and $\mu \in (\mathbb{R}^{\geq 0})^{k_g}$ such that

$$\sum_{i=1}^{k} \alpha_i \nabla f_i(\bar{x}) + \sum_{i \in \mathcal{A}_g(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0,$$
$$\mu_i = 0 \quad \forall i \notin \mathcal{A}_g(\bar{x}).$$

This implies that $\bar{x}$ is Pareto critical for the unconstrained MOP

$$\min_{x \in \mathbb{R}^n} \begin{pmatrix} f(x) \\ (g_i(x))_{i \in \mathcal{A}_g(\bar{x})} \end{pmatrix}. \tag{5.13}$$

(In contrast to the equality constrained case, we do not have to consider the sign of the KKT multipliers to obtain the unconstrained MOP, since $\mu \in (\mathbb{R}^{\geq 0})^{k_g}$. But instead, we now have the dependency on the active set $\mathcal{A}_g(\bar{x})$.) Conversely, for every $A \subseteq \{1, \dots, k_g\}$, every Pareto critical point of the unconstrained MOP

$$\min_{x \in \mathbb{R}^n} \begin{pmatrix} f(x) \\ (g_i(x))_{i \in A} \end{pmatrix} \tag{5.14}$$

that satisfies $g(x) \leq 0$ and $\mathcal{A}_g(x) = A$ is Pareto critical for the original MOP (5.12) (by setting $\mu_i = 0$ for all $i \notin A$). For $A \subseteq \{1, \dots, k_g\}$ let P_c^A be the Pareto critical set of (5.14). Then we can summarize our above considerations as

$$P_c = \bigcup_{A \subseteq \{1, \dots, k_g\}} P_c^A \cap \{x \in g^{-1}(\mathbb{R}^{\leq 0}) : \mathcal{A}_g(x) = A\}. \tag{5.15}$$

This relationship is visualized in the following example.

Example 5.4.2. *Consider the MOP* (5.12) *for*

$$f : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \end{pmatrix},$$
$$g : \mathbb{R}^2 \to \mathbb{R}, \quad x \mapsto \|x - c^2\|^2 - \left(\frac{3}{8}\right)^2$$

with

$$Q = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}, \; c^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; c^2 = \begin{pmatrix} \frac{5}{8} \\ \frac{1}{2} \end{pmatrix}.$$

This is the same objective vector and constraint function as in Example 5.4.1, but this time we treat the constraint as an inequality. In this way, the feasible set is now the (closed) sphere centered at c^2 with radius $\frac{3}{8}$. By (5.15)*, the Pareto critical set of this MOP can be expressed via the Pareto critical sets $P_c^{\emptyset}$ and $P_c^{\{1\}}$ of the unconstrained MOPs*

$$\min_{x \in \mathbb{R}^2} \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \end{pmatrix} \quad \textit{and} \quad \min_{x \in \mathbb{R}^2} \begin{pmatrix} \|x\|_Q^2 \\ \|x - c^1\|^2 \\ \|x - c^2\|^2 - \left(\frac{3}{8}\right)^2 \end{pmatrix},$$

respectively. More precisely, in this case we have

$$\begin{aligned} P_c &= (P_c^{\emptyset} \cap \{x \in g^{-1}(\mathbb{R}^{\leq 0}) : \mathcal{A}_g(x) = \emptyset\}) \cup (P_c^{\{1\}} \cap \{x \in g^{-1}(\mathbb{R}^{\leq 0}) : \mathcal{A}_g(x) = \{1\}\}) \\ &= (P_c^{\emptyset} \cap g^{-1}(\mathbb{R}^{<0})) \cup (P_c^{\{1\}} \cap g^{-1}(0)). \end{aligned}$$

This equation is visualized in Figure 5.15. ($P_c^{\emptyset}$ and $P_c^{\{1\}}$ can again be computed as in Example 2.2.8.)

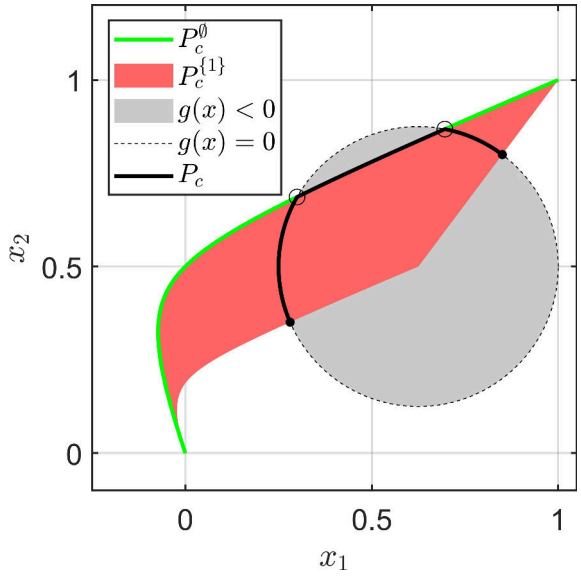


Figure 5.15: Deriving P_c by interpreting the inequality constraint as an additional objective in Example 5.4.2. The black dots show the critical points of the individual objectives (as constrained scalar problems). The black circles show the Pareto critical points where the active set changes.

We will now discuss the structure of the boundary of P_c in the inequality constrained case. Since g is continuous, the set $\{x \in g^{-1}(\mathbb{R}^{\leq 0}) : \mathcal{A}_g(x) = \emptyset\} = g^{-1}(\mathbb{R}^{<0})$ of points where all constraints are inactive is open. This means that for $A = \emptyset$ in (5.15), we can simply apply the (local) results from Section 5.1. Thus, we will focus on the case where $A \neq \emptyset$. In general, (5.15) can also be written as

$$P_c = \bigcup_{A \subseteq \{1,\dots,k_g\}} P_c^A \cap ((g_i)_{i \in A})^{-1}(0) \cap ((g_i)_{i \notin A})^{-1}(\mathbb{R}^{<0}).$$

Note that in the expression $P_c^A \cap ((g_i)_{i \in A})^{-1}(0)$, $(g_i)_{i \in A}$ acts like an equality constraint. Furthermore, $((g_i)_{i \notin A})^{-1}(\mathbb{R}^{<0})$ is an open set, such that it can be ignored for the consideration of the boundary of $P_c^A \cap ((g_i)_{i \in A})^{-1}(0)$. Thus, we can argue as in the equality constrained case to see that the boundary of $P_c^A \cap ((g_i)_{i \in A})^{-1}(0)$ consists of Pareto critical points of (5.14) where a KKT multiplier corresponding to either f or $(g_i)_{i \in A}$ is zero. But since we do not switch the sign of the constraint functions in the MOP (5.14) in the inequality constrained case, we cannot argue that the latter points can be ignored when considering the boundary of P_c. This was to be expected, as points with a vanishing KKT multiplier with respect to $(g_i)_{i \in A}$ are precisely the points in which the active set may change. Since this causes a "discontinuous change" in the KKT conditions, these points generically cause kinks in P_c.

Furthermore, for different sets $A \subseteq \{1, \dots, k_g\}$, the sets in the union in (5.15) may have different "dimensions", which can lead to additional boundary points where they touch. Thus, in the inequality constrained case, we expect the boundary of P_c to consist of points where a KKT multiplier of f is zero and points where the active set changes. This can also be seen in Example 5.4.2.

5.4.3 Examples

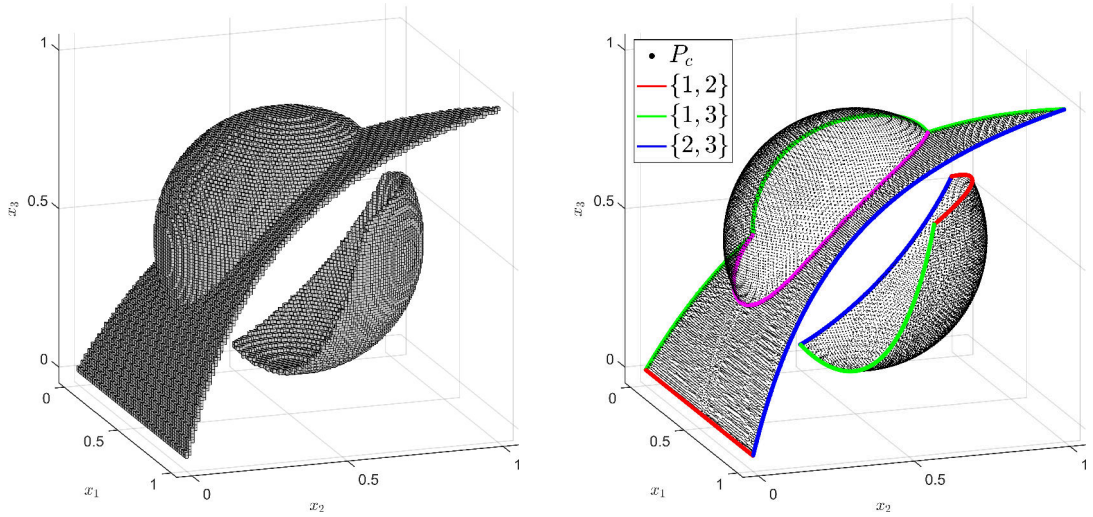
We will now visualize the structure of Pareto critical sets for equality and inequality constrained MOPs using some examples. For the computation of the Pareto critical sets, we use the continuation method for the constrained case from Section 3.1.4. The first example is an inequality constrained MOP with $n = 3$ variables and $k = 3$ objectives.

Example 5.4.3. *Consider the problem* (5.12) *for*

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto \begin{pmatrix} 2x_1^2 + x_2^2 + x_3^2 \\ (x_1 - 1)^2 + 2x_2^2 + x_3^2 \\ (x_1 - 1)^2 + (x_2 - 1)^2 + 2(x_3 - 1)^2 \end{pmatrix},$$

$$g : \mathbb{R}^3 \to \mathbb{R}, \quad x \mapsto -\left(\left(x_1 - \frac{1}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 + \left(x_2 - \frac{1}{2}\right)^2 - \left(\frac{2}{5}\right)^2\right).$$

The feasible set $g^{-1}(\mathbb{R}^{\leq 0})$ is the complement of the open sphere with radius $\frac{2}{5}$ and center $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})^\top$. The Pareto critical set of this problem is shown in Figure 5.16(a). Figure 5.16(b) shows the Pareto critical sets of all (constrained) subproblems with $|I| = 2$. Furthermore, all points where the active set changes are shown. As predicted, we see that the boundary of P_c consists of these two types of points.



(a) **(b)**

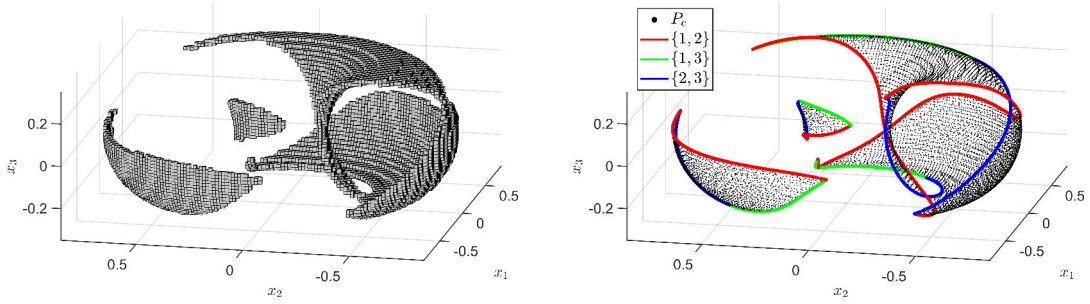
Figure 5.16: **(a)** A covering of the Pareto critical set P_c in Example 5.4.3 via boxes. **(b)** A pointwise discretization of P_c (black), the Pareto critical sets of all subproblems with $|I| = 2$ (red, green and blue) and the points in which the active set changes (magenta).

Next, we consider an equality constrained MOP which was previously considered in [Sch04], Example S4.

Example 5.4.4. *Consider the problem* (5.8) *for*

$$f : \mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto \begin{pmatrix} (x_1-1)^4 + (x_2-1)^2 + (x_3-1)^2 \\ (x_1+1)^2 + (x_2+1)^4 + (x_3+1)^2 \\ (x_1-1)^2 + (x_2+1)^2 + (x_3-1)^4 \end{pmatrix},$$
$$h : \mathbb{R}^3 \to \mathbb{R}, \quad x \mapsto -((x_1^2 + x_2^2 + x_3^2 - (R^2 + r^2))^2 - 4R^2(r^2 - x_3^2))$$

with $r = 0.3$ *and* $R = 0.5$. *The feasible set* $h^{-1}(0)$ *is the torus around the* x_3*-axis with minor radius* r *and major radius* R. *(In [Sch04], a different equality constraint was chosen, but the feasible set is identical.) The Pareto critical set* P_c *is shown in Figure 5.17(a) and Figure 5.17(b) shows the Pareto critical sets of all subproblems. We see that* P_c *consists of three connected components, all bounded by the Pareto*



(a) **(b)**

Figure 5.17: **(a)** A covering of the Pareto critical set P_c in Example 5.4.4 via boxes. **(b)** A pointwise discretization of P_c (black) and the Pareto critical sets of all subproblems with $|I| = 2$ (red, green and blue).

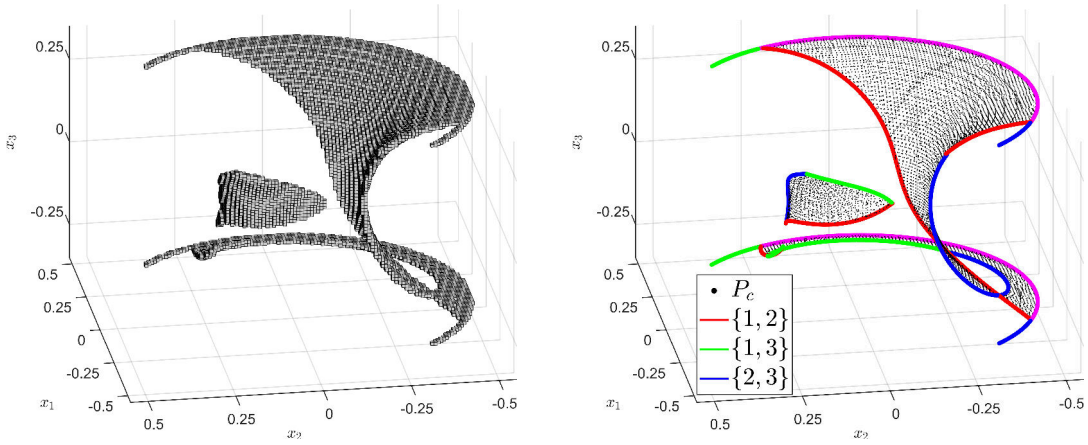
critical sets of the subproblems.

Finally, we will consider the MOP from Example 3.1.14, which has both equality and inequality constraints. In our earlier theoretical considerations, we only discussed the two types of constraints separately. The results here will suggest that our theoretical results also hold in the case where we have both types of constraints simultaneously.

Example 5.4.5. *Consider the MOP from Example 3.1.14. Figure 5.18(a) shows the Pareto critical set of this MOP. Figure 5.18(b) shows the Pareto critical set of all subproblems and the points in which the active set changes. We again see that the boundary of* P_c *consists of these two types of points. (In this example, the KKT vectors of the MOP* (5.14) *(constrained with* h*) are not unique in points where the inequality constraint is active. Thus, all Pareto critical points on the boundary of the feasible set are Pareto critical for multiple subproblems, such that their colors in Figure 5.18(b) are not unique.)*

5.5 Extension to the nonsmooth case

In this section, we will analyze the structure of the Pareto critical set in the nonsmooth (i.e., locally Lipschitz continuous) case (cf. Section 2.2.2). In the smooth



(a) **(b)**

Figure 5.18: **(a)** A covering of the Pareto critical set P_c in Example 5.4.5 via boxes. **(b)** A pointwise discretization of P_c (black), the Pareto critical sets of all subproblems with $|I| = 2$ (red, green and blue) and the points in which the active set changes (magenta).

case, we heavily exploited the fact that the set of Pareto critical points with a positive KKT vector can be written as the projection of the manifold $\mathcal{M}$ from Theorem 2.2.10 onto the variable space $\mathbb{R}^n$. Unfortunately, in the nonsmooth case, this smoothness property of P_c is lost. This can already be seen in the simple Example 2.2.25, where P_c has a kink (that is not related to the Pareto critical points of the individual objectives). Furthermore, with a definition of the boundary based on tangent cones (as in Definition 5.1.14), this kink would be identified as a boundary point. This shows that our definition of the boundary from the smooth case does not yield a "useful" characterization of the boundary of P_c in the nonsmooth case. Nonetheless, in the following, we will show that some basic results about the structure of P_c can be carried over. As in Section 2.2.2 and Chapter 4, we will assume for this section that all f_i, $i \in \{1, \dots, k\}$, are locally Lipschitz continuous.

We will begin by showing that the Pareto critical set is still a closed subset of $\mathbb{R}^n$.

Lemma 5.5.1. *P_c is closed.*

Proof. We will show that $\mathbb{R}^n \setminus P_c$ is open. To this end, let $x \in \mathbb{R}^n \setminus P_c$, so $0 \notin \partial^\cup f(x)$. Since $\partial^\cup f(x)$ is compact (as the convex hull of a compact set) and convex, there has to be an open, convex set $V \subseteq \mathbb{R}^n$ with $\partial^\cup f(x) \subseteq V$ and $0 \notin V$. In particular,

$$\partial f_i(x) \subseteq \partial^\cup f(x) \subseteq V \quad \forall i \in \{1, \dots, k\}.$$

Since ∂f_i is upper semicontinuous for all $i \in \{1, \dots, k\}$ (cf. Lemma 2.2.17), there has to be an open set $U \subseteq \mathbb{R}^n$ with $x \in U$ such that

$$\partial f_i(y) \subseteq V \quad \forall y \in U, i \in \{1, \dots, k\}.$$

Since V is convex, this implies

$$\partial^\cup f(y) = \operatorname{conv}\left(\bigcup_{i=1}^{k} \partial f_i(y)\right) \subseteq V \quad \forall y \in U.$$

In particular, $0 \notin \partial^{\cup} f(y)$ for all $y \in U$, so $U \subseteq \mathbb{R}^n \setminus P_c$, showing that $\mathbb{R}^n \setminus P_c$ is open. □

One of the basic concepts in the smooth case was the relationship between Pareto critical points and their KKT vectors, which are implicitly defined by the KKT condition. More precisely, KKT vectors contain the coefficients of a vanishing convex combination of the gradients $\nabla f_i(x)$, $i \in \{1, \ldots, k\}$, of the objectives. Although in the nonsmooth case, the KKT condition

$$0 \in \partial^{\cup} f(x) = \operatorname{conv}\left(\bigcup_{i=1}^{k} \partial f_i(x)\right)$$

still involves a vanishing convex combination, the coefficients now also depend on which elements from the subdifferentials $\partial f_i(x)$, $i \in \{1, \ldots, k\}$, are chosen. To obtain a well-defined characterization of KKT vectors in this case, we consider the following lemma.

Lemma 5.5.2. *Let $x \in \mathbb{R}^n$. Then*

$$\partial^{\cup} f(x) = \left\{ \sum_{i=1}^{k} \alpha_i \xi_i : \alpha \in \Delta_k, \xi_i \in \partial f_i(x), i \in \{1, \ldots, k\} \right\}.$$

Proof. Lemma 5.29 in [AB06]. □

By this lemma, it is sufficient to only consider a single element from each $\partial f_i(x)$, $i \in \{1, \ldots, k\}$, for the representation of $\partial^{\cup} f(x)$. This allows us to make the following definition.

Definition 5.5.3. *Let $x \in P_c$ and $\alpha \in \Delta_k$ such that there are elements $\xi_i \in \partial f_i(x)$, $i \in \{1, \ldots, k\}$, with $\sum_{i=1}^{k} \alpha_i \xi_i = 0$. Then α is a* KKT vector *of x. We denote by $A(x) \subseteq \Delta_k$ the set of all KKT vectors of x, i.e.,*

$$A(x) := \left\{ \alpha \in \Delta_k : \exists \xi_i \in \partial f_i(x), i \in \{1, \ldots, k\}, \text{ with } \sum_{i=1}^{k} \alpha_i \xi_i = 0 \right\}.$$

Recall that if all objectives f_i, $i \in \{1, \ldots, k\}$, are actually continuously differentiable, then each subdifferential just reduces to the set that only contains the respective gradient. In particular, the choice of $\xi_i \in \partial f_i(x) = \{\nabla f_i(x)\}$ is fixed and the previous definition reduces to the definition of KKT vectors from the smooth case.

In the smooth case, the KKT vectors can be used to locally parameterize the Pareto critical set, which was an important tool for our theoretical results (cf. Lemma 5.1.3). The parameterization was based on the application of the implicit function theorem to the system of equations $F(x, \alpha) = 0$ from (2.11). In the nonsmooth case this relationship is lost, since F would also depend on the choice of $\xi_i \in \partial f_i(x)$ and would generally be nonsmooth, as there is no smooth dependancy of $\xi_i \in \partial f_i(x)$ on x. Nonetheless, it is possible to derive some structural properties of P_c from the KKT vectors by considering the set-valued mapping A which maps x onto $A(x)$. To this end, we first need a technical result about sequences of subgradients (similar to [Cla90], Proposition 2.1.5).

Lemma 5.5.4. *Let $h : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitz continuous and $(x^i)_i \in \mathbb{R}^n$ with $\lim_{i\to\infty} x^i = \bar{x} \in \mathbb{R}^n$. Let $(v^i)_i \in \mathbb{R}^n$ with $v^i \in \partial h(x^i)$ for all $i \in \{1, \dots, k\}$. Then $(v^i)_i$ has an accumulation point and all accumulation points of $(v^i)_i$ lie in $\partial h(\bar{x})$.*

Proof. Let L be a Lipschitz constant of h at $\bar{x}$. (By definition, this implies that there is some open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that L is a Lipschitz constant for all $y \in U$.) By Lemma 2.2.17 and since $\lim_{i\to\infty} x^i = \bar{x}$, we can assume w.l.o.g. that

$$v^i \in \partial h(x^i) \subseteq \overline{B_L(0)} \quad \forall i \in \mathbb{N}.$$

Since $\overline{B_L(0)}$ is compact, $(v^i)_i$ must have an accumulation point in $\overline{B_L(0)}$. Let $\bar{v}$ be any accumulation point of $(v^i)_i$ and assume that $\bar{v} \notin \partial h(\bar{x})$. Since $\partial h(\bar{x})$ is compact, there must be some open set $V \subseteq \mathbb{R}^n$ with $\partial h(\bar{x}) \subseteq V$ such that $\bar{v} \notin \overline{V}$. Since ∂h is upper semicontinuous, there is an open set $W \subseteq \mathbb{R}^n$ with $\bar{x} \in W$ and $\partial h(y) \subseteq \overline{V}$ for all $y \in W$. In particular, there is some $N \in \mathbb{N}$ with $x^i \in W$ for all $i > N$ such that $\partial h(x^i) \subseteq \overline{V} \not\ni \bar{v}$. This is a contradiction to $\bar{v}$ being an accumulation point of $(v^i)_i$. □

The previous lemma can be used to show that $A(x)$ is compact (as a subset of $\mathbb{R}^k$) and that the set-valued map

$$A : \mathbb{R}^n \to \mathcal{P}(\mathbb{R}^k), \quad x \mapsto A(x)$$

is upper semicontinuous.

Lemma 5.5.5. *a) The set $A(x)$ is compact for all $x \in \mathbb{R}^n$.*

b) The map A is upper semicontinuous, i.e., for all $x \in \mathbb{R}^n$ and all open sets $V \subseteq \mathbb{R}^k$ with $A(x) \subseteq V$, there is some open set $U \subseteq \mathbb{R}^n$ with $x \in U$ such that

$$A(y) \subseteq V \quad \forall y \in U.$$

Proof. a) $A(x)$ is bounded as a subset of Δ_k. To show that it is also closed, let $(\alpha^i)_i \in A(x)$ with $\lim_{i\to\infty} \alpha^i = \bar{\alpha} \in \mathbb{R}^k$. Since Δ_k is compact we have $\bar{\alpha} \in \Delta_k$. By definition, there is a sequence $(\xi_j^i)_i \in \partial f_j(x)$ for each $j \in \{1, .., k\}$ with $\sum_{j=1}^k \alpha_j^i \xi_j^i = 0$ for all $i \in \mathbb{N}$. By Lemma 5.5.4 we can assume w.l.o.g. that for each $j \in \{1, \dots, k\}$, there is some $\bar{\xi}_j \in \partial f_j(x)$ with $\lim_{i\to\infty} \xi_j^i = \bar{\xi}_j$. It follows that

$$0 = \lim_{i\to\infty} \sum_{j=1}^k \alpha_j^i \xi_j^i = \sum_{j=1}^k \bar{\alpha}_j \bar{\xi}_j,$$

so $\bar{\alpha} \in A(x)$.
b) Let $\bar{x} \in \mathbb{R}^n$ and $V \subseteq \mathbb{R}^k$ open with $A(\bar{x}) \subseteq V$. Assume that for all open sets $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ there is some $y \in U$ such that $A(y) \not\subseteq V$. Then there must be a sequence $(x^i)_i \in \mathbb{R}^n$ with $\lim_{i\to\infty} x^i = \bar{x}$ and $A(x^i) \not\subseteq V$ for all $i \in \mathbb{N}$. In particular, there must be a sequence $(\alpha^i)_i \in \Delta_k$ with $\alpha^i \in A(x^i)$ and $\alpha^i \notin V$ for all $i \in \mathbb{N}$. By definition, for each $i \in \mathbb{N}$ there are $\xi_j^i \in \partial f_j(x^i)$, $j \in \{1, \dots, k\}$, such that

$$\sum_{j=1}^k \alpha_j^i \xi_j^i = 0.$$

Since $(\alpha^i)_i \in \Delta_k$ and Δ_k is compact, we can assume w.l.o.g. that there is some $\bar{\alpha} \in \Delta_k$ with $\lim_{i\to\infty} \alpha^i = \bar{\alpha}$. Furthermore, by Lemma 5.5.4 we can assume w.l.o.g. that for each $j \in \{1, \dots, k\}$, there is some $\bar{\xi}_j \in \partial f_j(\bar{x})$ with $\lim_{i\to\infty} \xi_j^i = \bar{\xi}_j$. This implies

$$0 = \lim_{i\to\infty} \sum_{j=1}^{k} \alpha_j^i \xi_j^i = \sum_{j=1}^{k} \bar{\alpha}_j \bar{\xi}_j.$$

This means that $\bar{\alpha} \in A(\bar{x}) \subseteq V$, which is a contradiction to $\lim_{i\to\infty} \alpha^i = \bar{\alpha}$ since V is open and $\alpha^i \notin V$ for all $i \in \mathbb{N}$. □

The upper semicontinuity of A can now be used to show a first result about the relationship between P_c and the Pareto critical sets of subproblems. In the nonsmooth case, we can define subproblems analogously to the smooth subproblems (MOP^I) for $I \in \mathcal{P}(\{1, \dots, k\})$. For ease of notation, let

$$\begin{aligned} \Delta_k^\circ &:= \{\alpha \in \Delta_k : \alpha_i > 0 \ \forall i \in \{1, \dots, k\}\}, \\ \partial\Delta_k &:= \{\alpha \in \Delta_k : \exists j \in \{1, \dots, k\} \text{ with } \alpha_j = 0\}, \end{aligned}$$

i.e., Δ_k° is the interior and $\partial\Delta_k$ is the boundary of Δ_k as a subset of the affine linear space $\{\alpha \in \mathbb{R}^k : \sum_{i=1}^k \alpha_i = 1\}$ endowed with the subspace topology of $\mathbb{R}^k$.

Lemma 5.5.6. *Let $\bar{x} \in P_c$. If $A(\bar{x}) \subseteq \Delta_k^\circ$ then there is some open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that*

$$A(x) \subseteq \Delta_k^\circ \quad \forall x \in U.$$

Proof. Since both $A(\bar{x})$ and $\partial\Delta_k$ are compact (cf. Lemma 5.5.5) and $A(\bar{x}) \cap \partial\Delta_k = \emptyset$, there is an open set $V \subseteq \mathbb{R}^k$ with $A(\bar{x}) \subseteq V$ and $V \cap \partial\Delta_k = \emptyset$. By Lemma 5.5.5 there has to be an open set $U \subseteq \mathbb{R}^n$ with $\bar{x} \in U$ such that $A(x) \subseteq V$ for all $x \in U$. In particular,

$$A(x) \cap \partial\Delta_k \subseteq V \cap \partial\Delta_k = \emptyset \quad \forall x \in U,$$

so $A(x) \subseteq \Delta_k^\circ$ for all $x \in U$. □

By the previous lemma, if $x \in P_c$ is a Pareto critical point that is not Pareto critical for any subproblem (with less than k objectives), then there are also no Pareto critical points of subproblems in an open neighborhood (in $\mathbb{R}^n$) around x. Thus, the set of Pareto critical points with positive KKT vectors is open as a subset of P_c with respect to the subspace topology of P_c from $\mathbb{R}^n$.

An important result in the smooth case was the fact that every $x \in P_c$ is Pareto critical for a subproblem with at most $\operatorname{rk}(Df(x)) + 1$ objectives (cf. Lemma 5.2.3). As there is no obvious concept for the "rank of the Jacobian" in the nonsmooth case, this result cannot directly be generalized. Nonetheless, a simpler version of the result can still be shown in the special case where $k > n + 1$ by applying Carathéodory's theorem (cf. Remark 5.2.5). To this end, let P_c^I again be the Pareto critical set of the subproblem (MOP^I) (for a locally Lipschitz continuous f) with $I \in \mathcal{P}(\{1, \dots, k\})$.

Lemma 5.5.7. *Let $k > n + 1$. Then for all $x \in P_c$ there is some $I \in \mathcal{P}(\{1, \dots, k\})$ with $|I| \leq n + 1$ such that $x \in P_c^I$.*

Proof. By Lemma 5.5.2 there are subgradients $\xi_i \in \partial f_i(x)$, $i \in \{1, \dots, k\}$, such that $0 \in \operatorname{conv}(\{\xi_1, \dots, \xi_k\})$. By Carathéodory's theorem [DGK63] and since $\xi_i \in \mathbb{R}^n$, there is some $I \in \mathcal{P}(\{1, \dots, k\})$ with $|I| \leq n+1$ such that

$$0 \in \operatorname{conv}(\{\xi_i : i \in I\}) \subseteq \bigcup_{i \in I} \partial f_i(x) = \partial^{\cup} f^I(x).$$

Thus, $x \in P_c^I$. □

The following example shows an application of the previous lemma for the case of $k = 4$ locally Lipschitz objectives in $n = 2$ variables.

Example 5.5.8. *Consider the problem* $\min_{x \in \mathbb{R}^2} f(x)$ *with*

$$f(x) := \begin{pmatrix} \max\{x_1^2 + (x_2 - 1)^2 + x_2 - 1, -x_1^2 - (x_2 - 1)^2 + x_2 + 1\} \\ -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1| \\ |x_1 - 0.25| + 2|x_2 - 0.5| \\ |x_1 - 0.8| + |x_2 - 0.8| \end{pmatrix}.$$

(The MOP consisting of only the first two objective functions was already considered in (4.22).) Since all subdifferentials can be computed by hand, it is possible to show that the Pareto critical set is the gray set in Figure 5.19. (For better visualization, additional components of the Pareto critical set outside of the shown area are ignored in this example.) Furthermore, the Pareto critical sets of all subproblems with $|I| = n + 1 = 3$ *are shown. As proven in Lemma 5.5.7, the original Pareto critical set can be written as the union of all Pareto critical sets of these subproblems.*

By Lemma 5.5.7, a problem with $k > n + 1$ can be reduced to a number of subproblems with $k = n+1$. For these subproblems, we can obtain a result about the structure of their Pareto critical sets by using the fact that generically, the interior of the convex hull of $n + 1$ elements in $\mathbb{R}^n$ is open (in $\mathbb{R}^n$) and the coefficients of a vanishing convex combination of $n + 1$ elements are unique.

Theorem 5.5.9. *Let* $k = n + 1$ *and* $\bar{x} \in P_c$. *If there is no* $I \in \mathcal{P}(\{1, \dots, k\})$ *with* $|I| < k$ *and* $\bar{x} \in P_c^I$ *(i.e., if* $\alpha_j > 0$ *for all* $\alpha \in A(\bar{x})$, $j \in \{1, \dots, k\}$*), then there is an open set* $U \subseteq \mathbb{R}^n$ *with* $\bar{x} \in U$ *and* $U \subseteq P_c$.

Proof. Assume that there is no such U. Then there is a sequence $(x^i)_i \in \mathbb{R}^n$ with $x^i \notin P_c$ and $\lim_{i \to \infty} x^i = \bar{x}$. For each $j \in \{1, \dots, k\}$, let $(\xi_j^i)_i \in \mathbb{R}^n$ be a sequence with $\xi_j^i \in \partial f_j(x^i)$ for all $i \in \mathbb{N}$. Since $x^i \notin P_c$ we have $0 \notin \operatorname{conv}(\{\xi_1^i, \dots, \xi_k^i\})$ for all $i \in \mathbb{N}$. By Lemma 5.5.4 we can assume w.l.o.g. that there are $\bar{\xi}_j \in \partial f_j(\bar{x})$ with $\lim_{i \to \infty} \xi_j^i = \bar{\xi}_j$ for each $j \in \{1, \dots, k\}$. We distinguish between two cases:
Case 1: $0 \in \operatorname{conv}(\{\bar{\xi}_1, \dots, \bar{\xi}_k\})$. Let $\bar{\alpha} \in A(\bar{x})$ with $\sum_{j=1}^k \bar{\alpha}_j \bar{\xi}_j = 0$. By assumption we have $\bar{\alpha}_j > 0$ for all $j \in \{1, \dots, k\}$. Note that we can write

$$\begin{aligned} 0 &= \sum_{j=1}^{k} \bar{\alpha}_j \bar{\xi}_j = \sum_{j=1}^{k-1} \bar{\alpha}_j \bar{\xi}_j + \bar{\alpha}_k \bar{\xi}_k = \sum_{j=1}^{k-1} \bar{\alpha}_j \bar{\xi}_j + \left(1 - \sum_{j=1}^{k-1} \bar{\alpha}_j\right) \bar{\xi}_k \\ &= \sum_{j=1}^{k-1} \bar{\alpha}_j \bar{\xi}_j + \bar{\xi}_k - \sum_{j=1}^{k-1} \bar{\alpha}_j \bar{\xi}_k = \sum_{j=1}^{k-1} \bar{\alpha}_j (\bar{\xi}_j - \bar{\xi}_k) + \bar{\xi}_k \\ &= M(\bar{\alpha}_1, \dots, \bar{\alpha}_{k-1})^{\top} + \bar{\xi}_k, \end{aligned} \tag{5.16}$$

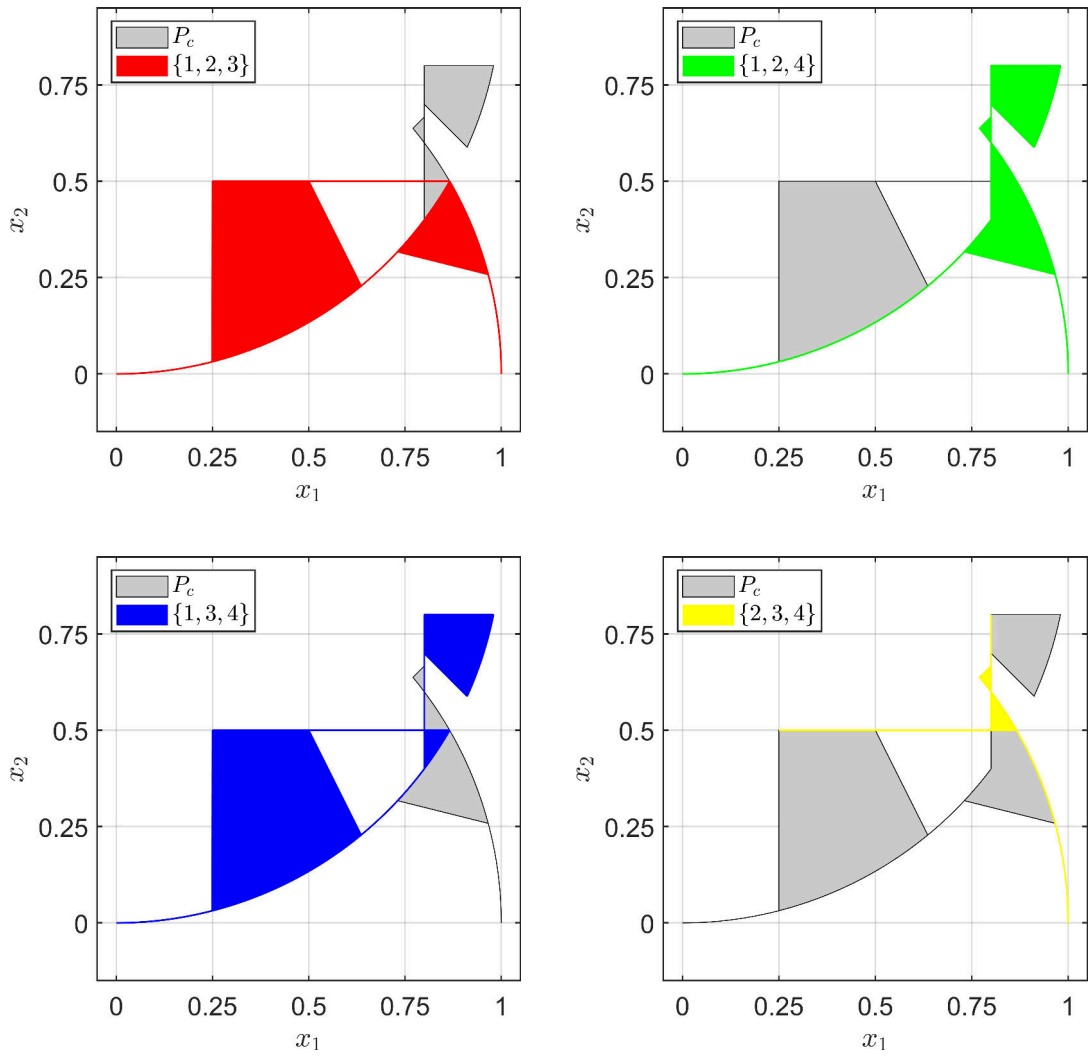


Figure 5.19: Pareto critical sets of the original problem and all subproblems with $|I| = 3$ in Example 5.5.8.

where $M \in \mathbb{R}^{n\times k-1} = \mathbb{R}^{n\times n}$ is the square matrix with columns $(\bar{\xi}_j - \bar{\xi}_k)_j \in \mathbb{R}^n$, $j \in \{1,\dots,k-1\}$. If M would not be regular, then there would be some $\beta \in \mathbb{R}^n \setminus \{0\}$ with

$$0 = M\beta = \sum_{j=1}^{k-1} \beta_j(\bar{\xi}_j - \bar{\xi}_k) = \sum_{j=1}^{k-1} \beta_j \bar{\xi}_j + \left(-\sum_{j=1}^{k-1} \beta_j\right) \bar{\xi}_k.$$

Defining $\beta' := (\beta_1,\dots,\beta_{k-1}, -\sum_{j=1}^{k-1}\beta_j)^\top$ would lead to $\sum_{j=1}^{k}\beta'_j\bar{\xi}_j = 0$ and $\sum_{j=1}^{k}\beta'_j = 0$. Analogously to the proof of Lemma 5.1.11 c), it would then follow that there is a KKT vector (for this choice of $\bar{\xi}_j \in \partial f_j(\bar{x})$, $j \in \{1,\dots,k\}$) with a zero component, which is a contradiction to our assumption. Thus, M is regular. This means that $\bar{\alpha}$ is uniquely determined by (5.16) via

$$(\bar{\alpha}_1,\dots,\bar{\alpha}_{k-1})^\top = -M^{-1}\bar{\xi}_k \tag{5.17}$$

and $\bar{\alpha}_k = 1 - \sum_{j=1}^{k-1}\bar{\alpha}_j$. Since the set of regular matrices is open, there is an open neighborhood of M in which all matrices are regular. Furthermore, the right-hand

side of (5.17) is continuous in $\bar{\xi}_1, \dots, \bar{\xi}_k$. Since $\bar{\alpha}_j > 0$ for all $j \in \{1, \dots, k\}$ and $\lim_{i\to\infty} \xi_j^i = \bar{\xi}_j$ for all $j \in \{1, \dots, k\}$, this means that there is some $N \in \mathbb{N}$ such that $A(x^i) \neq \emptyset$, i.e., $x^i \in P_c$ for all $i > N$, which is a contradiction.
Case 2: $0 \notin \operatorname{conv}(\{\bar{\xi}_1, \dots, \bar{\xi}_k\})$. Since $\bar{x} \in P_c$, there are $\alpha' \in \Delta_k$ and $\xi'_j \in \partial f_j(\bar{x})$, $j \in \{1, \dots, k\}$, such that $\sum_{j=1}^k \alpha'_j \xi'_j = 0$ and (by our assumption) $\alpha'_j > 0$ for all $j \in \{1, \dots, k\}$. For each $j \in \{1, \dots, k\}$, let

$$\xi_j : [0,1] \to \mathbb{R}^n, \quad t \mapsto \xi'_j + t(\bar{\xi}_j - \xi'_j).$$

Since $\partial f_j(\bar{x})$ is convex, we have $\xi_j(t) \in \partial f_j(\bar{x})$ for all $j \in \{1, \dots, k\}$, $t \in [0,1]$. By our construction, it holds $0 \in \operatorname{conv}(\{\xi_1(0), \dots, \xi_k(0)\})$ and $0 \notin \operatorname{conv}(\{\xi_1(1), \dots, \xi_k(1)\})$. Consider the continuous function

$$M : [0,1] \to \mathbb{R}^{n\times n}, \quad t \mapsto (\xi_1(t) - \xi_k(t), \dots, \xi_{k-1}(t) - \xi_k(t)).$$

With the same argument as in Case 1, we know that $M(t)$ must be regular for all $t \in [0,1]$ for which there is a positive KKT vector corresponding to $\xi_1(t), \dots, \xi_k(t)$. Since α' is positive, $M(t)$ is regular in an open neighborhood around 0. Since the set $\{t \in [0,1] : M(t) \text{ is irregular}\}$ is closed (and therefore compact), it is either empty or there is a smallest $t' \in (0,1]$ such that $M(t')$ is irregular. In the latter case, since $M(t')$ is irregular, there cannot be a positive KKT vector corresponding to $\xi_1(t'), \dots, \xi_k(t')$. Furthermore, there cannot be a KKT vector with a zero component by assumption. Thus, we must have $0 \notin \operatorname{conv}(\{\xi_1(t'), \dots, \xi_k(t')\})$. Due to the continuity of $\xi_1, \dots, \xi_k$ and compactness of $\operatorname{conv}(\{\xi_1(t'), \dots, \xi_k(t')\})$, this implies that there is some $t_0 < t'$ with $0 \notin \operatorname{conv}(\{\xi_1(t_0), \dots, \xi_k(t_0)\})$ such that $M(t)$ is regular for all $t \in [0, t_0)$. If $M(t)$ is regular, then the KKT vector corresponding to $\xi_1(t), \dots, \xi_k(t)$ is uniquely determined by

$$(\alpha_1, \dots, \alpha_{k-1})^\top = -M(t)^{-1}\xi_k(t) \tag{5.18}$$

and $\alpha_k = 1 - \sum_{j=1}^{k-1} \alpha_j$. In particular, the right-hand side of (5.18) is well defined and continuous in t for all $t \in [0, t_0)$. Let $\alpha^0 = (\alpha_1^0, \dots, \alpha_k^0)^\top \in \mathbb{R}^k$ with $(\alpha_1^0, \dots, \alpha_{k-1}^0)^\top := -M(t_0)^{-1}\xi_k(t_0)$ and $\alpha_k^0 := 1 - \sum_{j=1}^{k-1} \alpha_j^0$. Since we know that $0 \notin \operatorname{conv}(\{\xi_1(t_0), \dots, \xi_k(t_0)\})$, α^0 cannot be a KKT vector of $\bar{x}$, so $\alpha^0 \notin \Delta_k$. As the right-hand side of (5.18) is continuous, we can apply the intermediate value theorem to obtain the existence of some $\bar{t} \in [0, t_0)$ with a KKT vector $\bar{\alpha} \in \partial\Delta_k$. This implies that $\bar{x}$ is Pareto critical for a subproblem with $|I| < k$, which is a contradiction. □

By the previous theorem, for $k = n+1$ the set of Pareto critical points with only strictly positive KKT vectors is contained in the interior P_c° of P_c (with respect to the topology of $\mathbb{R}^n$). As P_c is closed (cf. Lemma 5.5.1), we have $P_c \setminus P_c^\circ = \overline{P_c} \setminus P_c^\circ = \partial P_c$. Thus, the topological boundary ∂P_c of P_c is contained in the set of Pareto critical points with at least one non-positive KKT vector. More formally, we have the following corollary.

Corollary 5.5.10. *If $k = n + 1$, then*

$$\partial P_c \subseteq \bigcup_{I\in\mathcal{P}(\{1,\dots,k\}),|I|=n} P_c^I.$$

In the following example, we apply the previous corollary to the subproblem corresponding to $I = \{1, 2, 4\}$ of the MOP considered in Example 5.5.8.

Example 5.5.11. *Consider the problem* $\min_{x\in\mathbb{R}^2} f(x)$ *with*

$$f(x) := \begin{pmatrix} \max\{x_1^2 + (x_2-1)^2 + x_2 - 1, -x_1^2 - (x_2-1)^2 + x_2 + 1\} \\ -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1| \\ |x_1 - 0.8| + |x_2 - 0.8| \end{pmatrix}.$$

The Pareto critical set of this MOP is the green set shown in Figure 5.19. Figure 5.20 shows the Pareto critical sets of all subproblems with $|I| = 2$*. As expected, we*

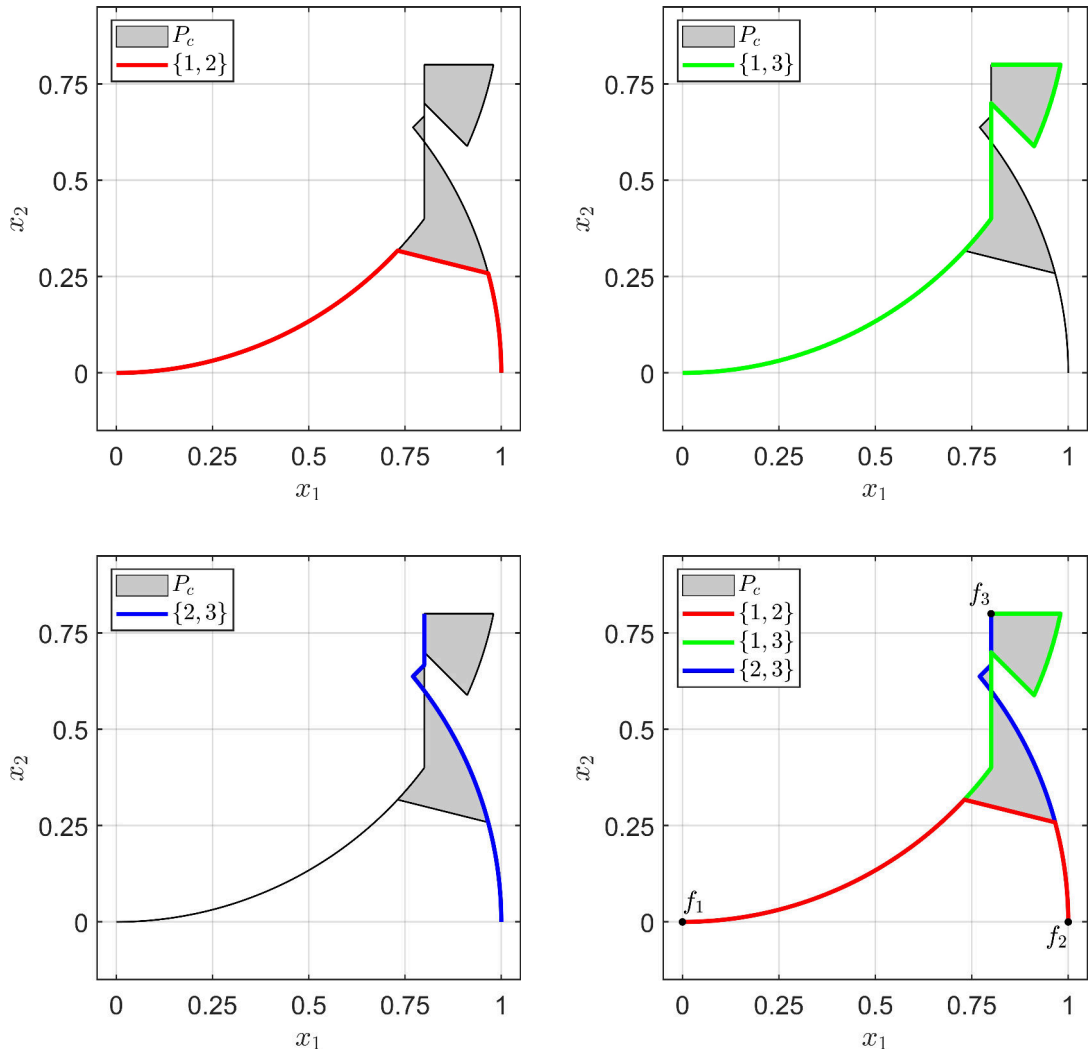


Figure 5.20: Pareto critical sets of the original problem and all subproblems with $|I| = 2$ in Example 5.5.11. The bottom right image also shows the critical points of the individual objectives as black dots.

see that the topological boundary ∂P_c *of* P_c *is covered by the Pareto critical sets of these subproblems.*

Summarizing our results in the nonsmooth case, we see that for $k \geq n+1$, we have a similar structure of the Pareto critical set with respect to subproblems as in the smooth case. For $k < n+1$, we basically only have Lemma 5.5.6, showing that if we have a Pareto critical point where all KKT vectors are strictly positive, then there is an open neighborhood around that point where all Pareto critical points have the same property. Clearly, in the smooth case, we had much stronger results.

As mentioned at the beginning of this section, the problem with generalizing these results is the fact that the nonsmoothness of the objective vector translates into a nonsmoothness of the Pareto critical set, which we cannot analyze with the tools we have used so far.

6 Inferring objective vectors from Pareto critical data

So far, throughout this thesis, we were always in the situation where the objective vector f of an MOP was given and we were trying to compute or analyze the structure of its Pareto (critical) set. While this is the typical setting of multiobjective optimization, there are cases where it makes sense to instead consider the objective vector as unknown and the solution as given. For example, for the generation of test problems for MOP solvers, one might search for an objective vector where the Pareto set has certain properties. This motivates us to consider the following problem, which can be regarded as the *inverse problem of multiobjective optimization*:

> Given a set $P \subseteq \mathbb{R}^n$, find an objective vector for which P is the Pareto set. (IMOP)

While it is possible to state this problem in this general form, it has many degenerate and trivial solutions. For example, for any $P \subseteq \mathbb{R}^n$, the indicator function

$$\mathbf{1}_{\mathbb{R}^n \setminus P} : \mathbb{R}^n \to \mathbb{R}, \quad x \mapsto \begin{cases} 0, & \text{if } x \in P, \\ 1, & \text{otherwise,} \end{cases}$$

trivially solves (IMOP). To deal with this ill-posedness, we will instead consider an inverse problem that is based on the concept of Pareto criticality. Since our understanding of the Pareto critical set in the smooth case is much better than in the nonsmooth case, we will also restrict our search to smooth objectives. While replacing the Pareto set P in (IMOP) by the Pareto critical set P_c and the restriction on smooth objectives would suffice to rule out the trivial solution given above, the solution of the inverse problem would still be highly non-unique. In other words, there can be vastly different objective vectors that correspond to the same Pareto critical set, as the following example shows.

Example 6.0.1. *Consider the problems* $\min_{x \in \mathbb{R}^2} f^i(x)$, $i \in \{1, 2, 3\}$, *with the following objective vectors:*

$$f^1 : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} x_1^2 + x_2^2 \\ (x_1 - 1)^2 + x_2^2 \end{pmatrix},$$

$$f^2 : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} x_1(x_1 - 1) + x_2^2 \\ x_1^2(x_1 - 1)^2 + x_2^2 \end{pmatrix},$$

$$f^3 : \mathbb{R}^2 \to \mathbb{R}^3, \quad x \mapsto \begin{pmatrix} x_1^2 + x_2^2 \\ (x_1 - \frac{1}{2})^2 + x_2^2 \\ (x_1 - 1)^2 + x_2^2 \end{pmatrix}.$$

One can show that all three MOPs have the same Pareto critical set $P_c = [0,1] \times \{0\}$, but vastly different properties: f^1 is convex while f^2 is nonconvex. Furthermore, the two boundary points of P_c (as defined in Chapter 5) correspond to critical points of different objectives for f^1, but to critical points of the same objective for f^2. Finally, f^3 shows that not even the number of objectives is uniquely determined by the Pareto critical set. The images of the f^i, $i \in \{1,2,3\}$, and the corresponding images of the Pareto critical sets are shown in Figure 6.1.

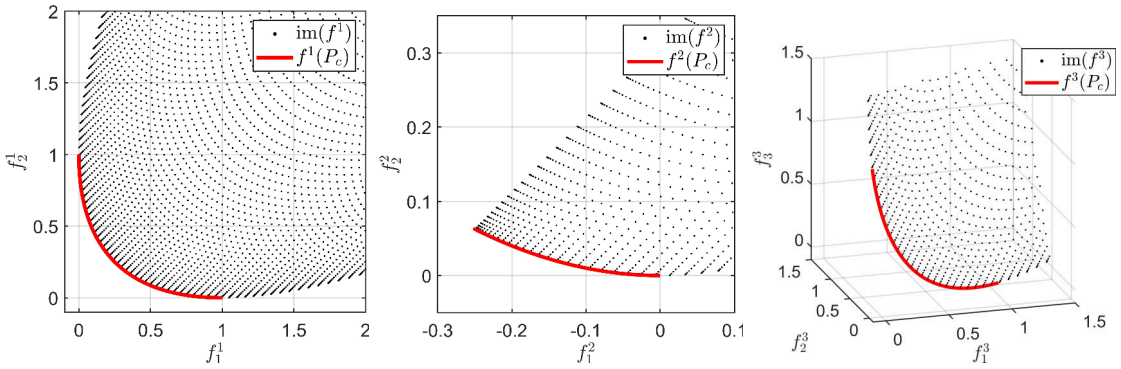


Figure 6.1: Image of the Pareto critical set under f^i and a pointwise discretization of $\mathrm{im}(f^i)$ for f^1, f^2 and f^3 from Example 6.0.1. (The image of f^3 lies in a two-dimensional affine subspace of $\mathbb{R}^3$.)

To solve this issue, the idea is to not only prescribe the Pareto critical points, but also the corresponding KKT vectors. To this end, we make the following definition.

Definition 6.0.2. *The set*

$$P_{\mathcal{M}} := \{(x, \alpha) \in \mathbb{R}^n \times \Delta_k : x \in P_c \text{ with KKT vector } \alpha\}$$

is the extended Pareto critical set. *A point $(x, \alpha) \in P_{\mathcal{M}}$ is an* extended Pareto critical point.

Note that compared to the manifold $\mathcal{M}$ from Theorem 2.2.10, we allow KKT multipliers to be zero in $P_{\mathcal{M}}$. Using this definition, we can now state a more well-behaved version of (IMOP):

Given a finite *data set* $\mathcal{D} = (\mathcal{D}_x, \mathcal{D}_\alpha) \subseteq \mathbb{R}^n \times \Delta_k$, find an objective vector $f \in C^1(\mathbb{R}^n, \mathbb{R}^k)$ whose extended Pareto critical set contains $\mathcal{D}$. (IMOPc)

Since our goal is to solve the inverse problem numerically, we have additionally imposed that the data set we prescribe is finite (and that it is only a subset of $P_{\mathcal{M}}$ instead of $P_{\mathcal{M}}$ itself). As the space of continuously differentiable functions $C^1(\mathbb{R}^n, \mathbb{R}^k)$ is infinite-dimensional, we will consider finite-dimensional subspaces of $C^1(\mathbb{R}^n, \mathbb{R})$ that are spanned by a set of basis functions $\mathcal{B} \subseteq C^1(\mathbb{R}^n, \mathbb{R})$ to approximate it. This will turn (IMOPc) into a homogeneous linear system in the coefficients of the basis functions, which can be solved via singular value decomposition (SVD). In this way, we obtain an objective vector for which the data is either exactly or "almost" extended Pareto critical, depending on the properties of that linear system.

While the prescription of the KKT vectors in addition to the Pareto critical points seems necessary from a theoretical point of view, it makes the generation of

the data more demanding in practice. Recall that geometrically, if we attach the KKT vector of a Pareto optimal point $\bar{x}$ to $f(\bar{x})$, then it is orthogonal to the Pareto front (cf. Lemma 2.2.7). Thus, the assumption that the KKT vectors are given in the data means that the data contains significantly more information than just the location of Pareto critical points in the variable space. Nonetheless, there are applications where this data is available or obtainable. The first application is the generation of test problems for MOP solvers, where the KKT vectors can be used to control topological and geometrical properties of the Pareto critical set. Here, the results from Chapter 5 about the structure of Pareto critical sets can be used. For example, we can try to force certain points to lie on the boundary of the Pareto critical set by choosing KKT vectors with a zero component. The second application lies within the field of *stochastic multiobjective optimization* [FX11; GP13], where the objective vector contains a random variable and the task is to optimize its expected value. Here, the data can be generated by computing extended Pareto critical points of the stochastic objective vector using a number of samples of the random variable. Our goal is to infer the expected value of the objective vector from the data. The third application we will consider is the generation of surrogate models for expensive MOPs, where the idea is to compute a few Pareto critical points of the expensive problem and to then use these points as data for our inverse approach. To obtain the corresponding KKT vectors, most standard methods for solving MOPs provide them either explicitly or implicitly. For example, in the weighting method (cf. Section 2.3.1), the KKT vector corresponding to a solution is given by the weight vector that was used to compute it.

For the single-objective case (where $k = 1$), problems like (IMOP) are addressed in the field of *inverse optimization*. For the linear case, inverse problems of the form $\min_x c^\top x$ (with some linear constraints) were considered in [AO01; HL99], with the goal of finding the cost vector c such that a given feasible point is optimal. In [KWB11], convex parameter-dependent problems were considered with the goal of approximating the objective function based on observations of parameter values and associated optimal solutions. Similarly to our approach, this is done by considering linear combinations of a pre-selected set of basis functions and then minimizing the residuals of the first-order optimality condition in the given observations with respect to the coefficients. (It is worth mentioning that part of the literature in the single-objective case is concerned with finding a weighting vector for the objectives of an MOP such that a given feasible point is optimal for the weighted sum (cf. [Cha+14; CL18]). This is also referred to as *inverse multiobjective optimization*, but clearly differs from our context.) Recently, a first result for the multiobjective case was published. In [DZ18], a method was proposed for finding the parameters of a parameter-dependent, convex and constrained MOP such that its Pareto set contains a given set of noisy data points (modeled via probability distributions). Their strategy involved a parameter-dependent discretization of the Pareto set by a finite number of solutions of the weighting method and then minimizing the sum over the distances of the discretization to the given data points. Formally, this can be denoted as a mixed-integer linear problem, for which a heuristic solution method was proposed. If we interpret the coefficients in the linear combinations of the basis functions in our approch as parameters, then we are in a similar situation as in [DZ18]. Compared to their approach, we will not require convexity or rely on heuristics to solve the inverse problem, but require the KKT vectors to be contained

in the data.

The remainder of this chapter is organized as follows. In Section 6.1, we will transform (IMOPc) into a linear problem in the expansion coefficients of the basis functions and derive a solution method that is based on SVD. In Section 6.2, we will discuss the applications mentioned above. We will begin by showcasing the generation of MOPs where the Pareto critical set has certain prescribed properties (Section 6.2.1), before considering stochastic MOPs (Section 6.2.2) and the generation of surrogate models (Section 6.2.3). Finally, we will discuss open problems of our approach in Section 6.3.

Parts of this chapter have been previously published in [GP21b], to which the author of this thesis was the main contributor.

6.1 Linearity of the inverse problem and its solution via SVD

In this section, we will transform the inverse problem (IMOPc) into a linear problem and propose an algorithm to solve it. Our goal is to construct an objective vector for which the extended Pareto critical set $P_{\mathcal{M}}$ contains a finite set of data points $\mathcal{D}_x = \{\bar{x}^1, \dots, \bar{x}^N\} \subseteq \mathbb{R}^n$ with KKT vectors $\mathcal{D}_\alpha = \{\bar{\alpha}^1, \dots, \bar{\alpha}^N\} \subseteq \Delta_k$, i.e., for which $(\bar{x}^j, \bar{\alpha}^j) \in P_{\mathcal{M}}$ for all $j \in \{1, \dots, N\}$. The general idea of our inverse approach is to plug the data points $(\bar{x}^j, \bar{\alpha}^j)$ into the KKT conditions and solve them for the objective vector f. So instead of the classical task of searching for an $x \in \mathbb{R}^n$ for which an $\alpha \in \Delta_k$ exists such that (KKT) holds, we now search for an $f \in C^1(\mathbb{R}^n, \mathbb{R}^k)$ for which (KKT) holds for all $\bar{x}^j$ and $\bar{\alpha}^j$, $j \in \{1, \dots, N\}$. Since $C^1(\mathbb{R}^n, \mathbb{R}^k)$ is infinite-dimensional, we will identify it with $C^1(\mathbb{R}^n, \mathbb{R})^k$ and restrict the problem to a finite-dimensional subspace of $C^1(\mathbb{R}^n, \mathbb{R})$ that is spanned by a set of basis functions $\mathcal{B} = \{b_1, \dots, b_d\} \subseteq C^1(\mathbb{R}^n, \mathbb{R})$. An example for the choice of basis functions is the set of monomials in n variables such that $\operatorname{span}(\mathcal{B})$ is the set of multivariate polynomials (up to a certain degree). By replacing $C^1(\mathbb{R}^n, \mathbb{R}^k)$ with $\operatorname{span}(\mathcal{B})^k$, we can represent each element of our search space via its coefficient vector $c \in \mathbb{R}^{k \cdot d}$. This allows us to turn (IMOPc) into a homogeneous linear problem in c, which we can solve via SVD. We will show that the smallest singular value can be used as a measure for how well the given data set can be fitted into the extended Pareto critical set of an objective vector in $\operatorname{span}(\mathcal{B})^k$.

For the remainder of this section, we assume that we are given

- a data set $\mathcal{D} = \{(\bar{x}^1, \bar{\alpha}^1), \dots, (\bar{x}^N, \bar{\alpha}^N)\} \subseteq \mathbb{R}^n \times \Delta_k$ (which implicitly prescribes the number of objectives k),
- a set of basis functions $\mathcal{B} = \{b_1, \dots, b_d\} \subseteq C^1(\mathbb{R}^n, \mathbb{R})$ with linearly independent derivatives (as elements of $C^0(\mathbb{R}^n, \mathbb{R}^n)$).

The linear independence of the derivatives of the basis functions implies linear independence of the basis functions themselves, since

$$\sum_{j=1}^{d} c_j f_j = 0 \;\Rightarrow 0 = \nabla \left(\sum_{j=1}^{d} c_j f_j \right) = \sum_{j=1}^{d} c_j \nabla f_j \;\Leftrightarrow c = 0.$$

Thus, every element of $\operatorname{span}(\mathcal{B})$ is uniquely determined by its coefficient vector. Furthermore, the linear independence assures that the trivial solution of a constant objective vector (for which every point is Pareto critical for every KKT vector) corresponds to the coefficient vector $0 \in \mathbb{R}^{k \cdot d}$.

After replacing $C^1(\mathbb{R}^n, \mathbb{R}^k)$ in (IMOPc) with $\operatorname{span}(\mathcal{B})^k$, our goal is to find a function $f : \mathbb{R}^n \to \mathbb{R}^k$, $f = (f_i)_{i \in \{1,\dots,k\}}$, $f \neq 0$ with $f_i \in \operatorname{span}(\mathcal{B})$ for all $i \in \{1, \dots, k\}$ and

$$\sum_{i=1}^{k} \bar{\alpha}_i \nabla f_i(\bar{x}) = 0 \quad \forall (\bar{x}, \bar{\alpha}) \in \mathcal{D}. \tag{6.1}$$

Since $f_i \in \operatorname{span}(\mathcal{B})$, we can write

$$f_i = \sum_{j=1}^{d} c_{ij} b_j \tag{6.2}$$

for some $c_i \in \mathbb{R}^d$. Thus, for all $x \in \mathbb{R}^n$ we have

$$\begin{aligned} \sum_{i=1}^{k} \alpha_i \nabla f_i(x) = \sum_{i=1}^{k} \alpha_i \sum_{j=1}^{d} c_{ij} \nabla b_j(x) = \sum_{i=1}^{k} \sum_{j=1}^{d} \alpha_i c_{ij} \nabla b_j(x) \\ = L(x, \alpha) c \end{aligned} \tag{6.3}$$

where

$$c := (c_{11}, \dots, c_{1d}, c_{21}, \dots, c_{2d}, \dots, c_{k1}, \dots, c_{kd})^\top \in \mathbb{R}^{k \cdot d} \tag{6.4}$$

and

$$\begin{aligned} &L(x, \alpha) \\ &:= (\alpha_1 \nabla b_1(x), \dots, \alpha_1 \nabla b_d(x), \alpha_2 \nabla b_1(x), \dots, \alpha_2 \nabla b_d(x), \dots, \alpha_k \nabla b_1(x), \dots, \alpha_k \nabla b_d(x)) \\ &\in \mathbb{R}^{n \times (k \cdot d)}. \end{aligned}$$

Let

$$\mathcal{L} := \begin{pmatrix} L(\bar{x}^1, \bar{\alpha}^1) \\ \vdots \\ L(\bar{x}^N, \bar{\alpha}^N) \end{pmatrix} \in \mathbb{R}^{(n \cdot N) \times (k \cdot d)}. \tag{6.5}$$

Then by construction, (6.1) is equivalent to the homogeneous linear system

$$\mathcal{L} c = 0. \tag{6.6}$$

In particular, a (non-trivial) function satisfying (6.1) exists if and only if

$$\operatorname{rk}(\mathcal{L}) < k \cdot d. \tag{6.7}$$

We will consider two cases with respect to the size of $\mathcal{L}$:

1. The system (6.6) is underdetermined, i.e., $n \cdot N < k \cdot d$. In this case, (6.7) automatically holds, such that (6.6) always has a non-trivial solution. Here, our approach behaves like an interpolation method. In fact, for $n = 1$, $k = 1$ and monomials as basis functions, $\mathcal{L}$ is similar to the *Vandermonde matrix* [HJ12] from polynomial interpolation (without the constant column).

2. The system (6.6) is square or overdetermined, i.e., $n \cdot N \geq k \cdot d$. This means that generically, we do not expect (6.6) to have a non-trivial solution, i.e., we do not expect that $\mathcal{D}$ can be covered by an extended Pareto critical set of a (non-trivial) function from $\operatorname{span}(\mathcal{B})^k$. Thus we actually have to check the condition (6.7) in this case.

The following example shows a case where (6.6) is underdetermined.

Example 6.1.1. *Consider the data set $\mathcal{D} = (\mathcal{D}_x, \mathcal{D}_\alpha)$ with*

$$\mathcal{D}_x = \{(0,0)^\top, (1,0)^\top, (1,1)^\top, (0,1)^\top\},$$
$$\mathcal{D}_\alpha = \left\{(0,1)^\top, \left(\frac{1}{3}, \frac{2}{3}\right)^\top, \left(\frac{2}{3}, \frac{1}{3}\right)^\top, (1,0)^\top\right\},$$

i.e., we have $n = 2$ variables, $k = 2$ objectives and $N = 4$ data points. For the set of basis functions, we consider the (non-constant) monomials up to degree 2, i.e.,

$$\mathcal{B} = \{x_1, x_1^2, x_2, x_1x_2, x_2^2\},$$

such that we have $d = 5$ basis functions. Then $n \cdot N = 8 < 10 = k \cdot d$, so (6.6) is underdetermined. Based on the data set and the basis functions, we can assemble $\mathcal{L} \in \mathbb{R}^{8\times 10}$ via the formula in (6.5) to obtain

$$\mathcal{L} = \frac{1}{3}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 & 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 2 & 2 & 0 \\ 2 & 4 & 0 & 2 & 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 2 & 4 & 0 & 0 & 1 & 1 & 2 \\ 3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 6 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The solution of (6.6) is then given by the span of the two vectors

$$c^1 = (-4, 0, 0, 4, 0, 0, 1, 0, -2, -3)^\top,$$
$$c^2 = (0, 2, 12, 0, -6, 0, -1, 0, -6, 3)^\top.$$

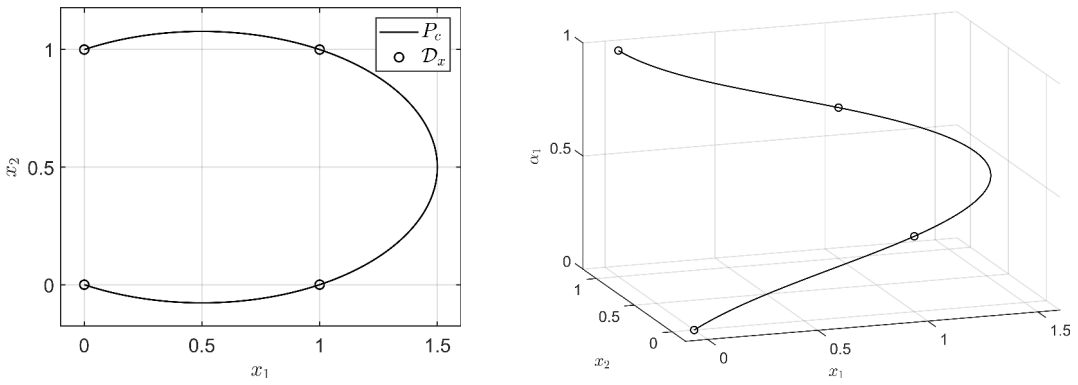
According to (6.2), these vectors correspond to the objective vectors

$$f^1(x) := \begin{pmatrix} 4x_1x_2 - 4x_1 \\ x_1^2 - 2x_1x_2 - 3x_2^2 \end{pmatrix} \quad \text{and} \quad f^2(x) := \begin{pmatrix} 2x_1^2 - 6x_2^2 + 12x_2 \\ -x_1^2 - 6x_1x_2 + 3x_2^2 \end{pmatrix},$$

respectively. It is possible to show that the extended Pareto critical set of f^1 is given by

$$P_{\mathcal{M}} = \left\{ \left(\left(\frac{-12\alpha_2^2 + 12\alpha_2}{8(3\alpha_2^2 - 3\alpha_2 + 1)}, \frac{12\alpha_2^2 - 20\alpha_2 + 8}{8(3\alpha_2^2 - 3\alpha_2 + 1)} \right)^\top, (1 - \alpha_2, \alpha_2)^\top \right) : \alpha_2 \in [0,1] \right\}.$$

The result is visualized in Figure 6.2. It is easy to verify that $\mathcal{D} \subseteq P_{\mathcal{M}}$, as expected. Furthermore, it is possible to show that the extended Pareto critical sets of f^1 and f^2 coincide in this example.



(a) **(b)**

Figure 6.2: **(a)** The Pareto critical set P_c of f^1 and the data $\mathcal{D}_x$ in Example 6.1.1. **(b)** A projection of the extended Pareto critical set $P_{\mathcal{M}}$ onto (x_1, x_2, α_1) and the corresponding projection of the data $\mathcal{D}$.

If the system (6.6) is underdetermined, then the solution of the inverse problem just boils down to the computation of the kernel of $\mathcal{L}$, as shown in the previous example. In the following, we will consider the case where the system is square or overdetermined. Although generically, we cannot expect to have exact solutions for this case, we will show that we can find approximate solutions of the inverse problem by finding approximate solutions of the linear system.

For ease of notation, we begin by making the following definition.

Definition 6.1.2. *Let*

$$\mathcal{F} : \mathbb{R}^{k\cdot d} \to C^1(\mathbb{R}^n, \mathbb{R}^k), \quad c \mapsto (f_i)_{i\in\{1,\dots,k\}} = \left(\sum_{j=1}^{d} c_{ij} b_j\right)_{i\in\{1,\dots,k\}}$$

be the function that maps a coefficient vector c onto the corresponding objective vector $f = (f_i)_{i\in\{1,\dots,k\}}$ (cf. (6.2) and (6.4)).

Consider the problem

$$\min_{\|c\|=1} \|\mathcal{L}c\|, \tag{6.8}$$

where the coefficient vector c is constrained to the unit sphere $\mathcal{S}^{(k\cdot d)-1}$ in $\mathbb{R}^{k\cdot d}$ to avoid the trivial solution $c^* = 0$. If c^* is a solution of (6.8) and $f = \mathcal{F}(c^*)$ is the corresponding objective vector, then the definition of $\mathcal{L}$ implies that

$$\|Df(\bar{x})^\top \bar{\alpha}\| = \|L(\bar{x}, \bar{\alpha})c^*\| \leq \|\mathcal{L}c^*\| \quad \forall (\bar{x}, \bar{\alpha}) \in \mathcal{D}. \tag{6.9}$$

In other words, the optimal value of (6.8) is an upper bound for the left-hand side of the KKT condition $\|Df(\bar{x})^\top \bar{\alpha}\|$ of f in each data point $(\bar{x}, \bar{\alpha}) \in \mathcal{D}$. The following lemma is a standard result for SVD and shows that the solution of (6.8) can be computed via an SVD of $\mathcal{L}$. As a reminder, an SVD of $\mathcal{L}$ is a decomposition

$$\mathcal{L} = USV^\top,$$

where $U \in \mathbb{R}^{(n\cdot N)\times(n\cdot N)}$, $V \in \mathbb{R}^{(k\cdot d)\times(k\cdot d)}$ are orthogonal matrices and $S \in \mathbb{R}^{(n\cdot N)\times(k\cdot d)}$ is a diagonal matrix. The diagonal entries of S are the singular values of $\mathcal{L}$ and the columns of V are the right-singular vectors of $\mathcal{L}$.

Lemma 6.1.3. *Let $p = \min(\{n \cdot N, k \cdot d\})$. Let $0 \leq s_1 \leq \ldots \leq s_p$ be the sorted singular values and let $v_1, \ldots, v_{k \cdot d} \in \mathbb{R}^{k \cdot d}$ be (corresponding) right-singular vectors of $\mathcal{L}$.*

(a) If $n \cdot N < k \cdot d$, then the optimal value of (6.8) is zero and the set of optimal solutions is given by $\operatorname{span}(\{v_i : s_i = 0\} \cup \{v_{(n \cdot N)+1}, \ldots, v_{k \cdot d}\}) \cap \mathcal{S}^{(k \cdot d)-1}$.

(b) If $n \cdot N \geq k \cdot d$, then we have

$$\min_{\|c\|=1} \|\mathcal{L}c\| = s_1 \quad \text{and} \quad \arg\min_{\|c\|=1} \|\mathcal{L}c\| = \operatorname{span}(\{v_i : s_i = s_1\}) \cap \mathcal{S}^{(k \cdot d)-1}.$$

Proof. (a) For $n \cdot N < k \cdot d$, the definition of the SVD implies that

$$\ker(\mathcal{L}) = \operatorname{span}(\{v_i : s_i = 0\} \cup \{v_{(n \cdot N)+1}, \ldots, v_{k \cdot d}\}).$$

The assertion then follows by enforcing the constraint $\|c\| = 1$.
(b) This follows from the Courant-Fischer Minimax Theorem for singular values (cf. Theorem 1 in [Dax13]). □

By the previous lemma, the smallest singular value s_1 can be seen as a measure for how well the data set $\mathcal{D}$ can be approximated by the extended Pareto critical set of an MOP where the objectives are linear combinations of basis functions in $\mathcal{B}$. Furthermore, if there are multiple singular values of the same (small) magnitude, then the number of those singular values corresponds to the dimension of the space of approximating objective vectors. The numerical procedure based on Lemma 6.1.3 for the solution of (IMOPc) is summarized in Algorithm 6.1.

Algorithm 6.1 Generate objective vector from data

Given: Data set $\mathcal{D} \subseteq \mathbb{R}^n \times \Delta_k$ of size $N \in \mathbb{N}$, set of basis functions $\mathcal{B} \subseteq C^1(\mathbb{R}^n, \mathbb{R})$ of size $d \in \mathbb{N}$, threshold $\bar{s} \geq 0$.

1: Assemble $\mathcal{L}$ as in (6.5). Let $p = \min(\{n \cdot N, k \cdot d\})$.
2: Calculate an SVD of $\mathcal{L}$ with singular values $0 \leq s_1 \leq \ldots \leq s_p$ and right-singular vectors $v_1, \ldots, v_{k \cdot d}$.
3: Identify the indices $I = \{1, \ldots, i^*\} \subseteq \{1, \ldots, p\}$ such that $s_i \leq \bar{s}$ for all $i \in I$.
4: Choose an element

$$c^* \in \operatorname{span}(\{v_i : i \in I\} \cup \{v_{p+1}, \ldots, v_{k \cdot d}\}) \subseteq \mathbb{R}^{k \cdot d}$$

with $\|c^*\| = 1$.
5: Assemble the objective vector $f = \mathcal{F}(c^*)$ as in (6.2).

In terms of the quality of the result of Algorithm 6.1, we have the following theorem.

Theorem 6.1.4. *Let f be the result of Algorithm 6.1 and let s_{i^*} be the largest singular value of $\mathcal{L}$ less or equal to $\bar{s}$. Then*

$$\|Df(\bar{x})^\top \bar{\alpha}\| \leq s_{i^*} \leq \bar{s} \quad \forall (\bar{x}, \bar{\alpha}) \in \mathcal{D}.$$

In particular, if $s_{i^*} = 0$, then all $(\bar{x}, \bar{\alpha}) \in \mathcal{D}$ are extended Pareto critical for the MOP with objective vector f.

Proof. Let $\mathcal{L} = USV^\top$ be an SVD of $\mathcal{L}$ where the columns of V are given by the right-singular vectors $v_1, \ldots, v_{k\cdot d}$ from Algorithm 6.1. Let c^* be the chosen coefficient vector in step 4 such that $f = \mathcal{F}(c^*)$. Let $p = \min(\{n \cdot N, k \cdot d\})$. Then there is some $\lambda \in \mathbb{R}^{k\cdot d}$ with $c^* = V\lambda$, $\lambda_{i^*+1} = \ldots = \lambda_p = 0$ and $1 = \|c^*\| = \|V\lambda\| = \|\lambda\|$. Thus,

$$\begin{aligned}\|\mathcal{L}c^*\| = \|\mathcal{L}V\lambda\| = \|US\lambda\| = \|S\lambda\| &= \sqrt{\sum_{i=1}^{p} s_i^2\lambda_i^2} = \sqrt{\sum_{i=1}^{i^*} s_i^2\lambda_i^2} \\ &\leq \sqrt{\sum_{i=1}^{i^*} s_{i^*}^2\lambda_i^2} = s_{i^*}\sqrt{\sum_{i=1}^{i^*}\lambda_i^2} \leq s_{i^*}\sqrt{\sum_{i=1}^{k\cdot d}\lambda_i^2} = s_{i^*}\|\lambda\| = s_{i^*}.\end{aligned}$$

The proof follows as in (6.9). □

Some properties of Algorithm 6.1 are discussed in the following remark.

Remark 6.1.5. *(a) In practice, it can make sense to first compute all singular values of $\mathcal{L}$ before choosing the threshold $\bar{s}$. In this way, gaps in the singular values can be taken into account.*

(b) In general, if $i^* > 1$, there is no obvious choice for c^* in step 4. For the interpretability of the result, it can make sense to choose a c^* that is as sparse as possible. (A similar approach was chosen in [BPK16] for the discovery of governing equations in the context of dynamical systems.)

(c) Recall that in (IMOPc)*, the data set is only required to be a subset of the extended Pareto critical set. Thus, it can happen that the Pareto critical set of the objective vector f from Algorithm 6.1 possesses additional components which were not prescribed in the data. Due to this, there are cases where the smallest singular value is close to zero, but the corresponding MOP is not desirable. (This will later be observed in practice in Section 6.2.1.)*

(d) By Lemma 6.1.3, if $s_{i^*} = 0$, then we do not have to normalize c^* in step 4.

We conclude this section with a discussion of the choice of basis functions. From the theory, the only requirements we have are differentiability and linear independence of the derivatives. While these are sufficient for our theoretical results to hold, there are additional properties that are desirable in practice:

(i) Since the derivatives of the basis functions have to be evaluated in every data point in $\mathcal{D}_x$ for the assembly of $\mathcal{L}$, the evaluation of the derivatives should be efficient.

(ii) In practice, it is difficult to say a priori how many basis functions are needed. Thus, it should be possible to generate new basis function (with linear independent derivatives) without much effort.

As mentioned earlier, an intuitive choice for the basis functions are the (non-constant) monomials in n variables up to a degree $l \in \mathbb{N}$, i.e.,

$$\mathcal{B} = \{x_1^{l_1}x_2^{l_2}\cdots x_n^{l_n} : l_i \in \mathbb{N}\cup\{0\}, i \in \{1,\ldots,n\}, 0 < l_1 + \ldots + l_n \leq l\},$$

such that span($\mathcal{B}$) is the space of (non-constant) polynomials up to degree l. Clearly, (derivatives of) monomials can be evaluated efficiently. Furthermore, the number of basis functions can be increased by increasing the maximum degree l. On top of that, for the space of polynomials, there are well-known results from approximation theory like the *Stone-Weierstrass theorem* [Rud91], which are also promising for our setting. Therefore, for the examples in the following section, we will always choose the monomials up to a certain degree as basis functions.

6.2 Applications

In this section, we will show how Algorithm 6.1 can be used to solve the inverse problems that arise in three different applications. We will begin by demonstrating how certain topological and geometrical properties can be enforced in the Pareto critical set, which can be useful for the generation of test problems for MOP solvers. Afterwards, we show how our results can be used to infer objectives of stochastic MOPs from stochastic data of their extended Pareto critical sets. Finally, we use our approach to generate surrogate models for (potentially expensive) objective vectors. The practical implementation of Algorithm 6.1 was carried out in MATLAB.

6.2.1 Generating MOPs with prescribed properties

Due to the many different topological and geometrical features Pareto sets can have, test problems and generators of test problems are important tools to analyze the behavior of MOP solvers and to compare them in practice [Deb99; Ker+16; Zha+08]. By Theorem 6.1.4, if the smallest singular value in Algorithm 6.1 is vanishing, then we found an MOP for which the data we prescribed is extended Pareto critical. In this way, our method can be used to create test problems with (partly) prescribed Pareto critical sets.

To control the properties of the Pareto critical set, the results from Section 2.2.1 and Chapter 5 can be used. In this context, they can be roughly summarized in the following two rules:

- By Theorem 2.2.10, the set of extended Pareto critical points with positive KKT vectors is a smooth manifold. Here, this means that data points with similar values in $\mathcal{D}_x$ should also have similar values in $\mathcal{D}_\alpha$.
- By our results in Section 5.1.3, points on the boundary of the Pareto critical set have a KKT vector with a zero component. Thus, we can (try to) force a point $\bar{x}^j$ to lie on the boundary of the Pareto critical set by choosing $\bar{\alpha}_i^j = 0$ for some $i \in \{1, \ldots, k\}$.

If we want to fully prescribe the Pareto critical set of the MOP resulting from our algorithm (and not just points from it), then the data in $\mathcal{D}_x$ should be a fine pointwise approximation of the prescribed set. This implies that the number of data points N is large and the system (6.6) is (typically) overdetermined. Thus, we can generally not expect that the MOP we are looking for actually exists (with respect to our choice of basis functions). Nonetheless, it will turn out that there are non-trivial cases where this approach does work, which we will demonstrate in the following.

In our our first example, we will generate an MOP where the Pareto critical set is the unit sphere $\mathcal{S}^1$ in $\mathbb{R}^2$.

Example 6.2.1. *For $N \in \mathbb{N}$ consider the points*

$$\bar{x}^j := \begin{pmatrix} \cos(2\pi \frac{j}{N}) \\ \sin(2\pi \frac{j}{N}) \end{pmatrix}, \quad j \in \{1, \dots, N\},$$

which are distributed equidistantly on $\mathcal{S}^1$. For the choice of the corresponding KKT vectors $\bar{\alpha}^j$, $j \in \{1, \dots, N\}$, note that $\mathcal{S}^1$ has an empty boundary (as defined in Definition 5.1.14), such that only the first one of the two rules above apply here. By construction, $\bar{x}^j$ is close to $\bar{x}^{j+1}$ for $j \in \{1, \dots, N-1\}$ and $\bar{x}^N$ is close to $\bar{x}^1$, so the same should hold for the $\bar{\alpha}^j$. One way of assuring this is to define the $\bar{\alpha}^j$ such that they periodically depend on j. For example, we can choose

$$\bar{\alpha}^j := \begin{pmatrix} \frac{1}{2}(\cos(4\pi \frac{j}{N}) + 1) \\ 1 - \frac{1}{2}(\cos(4\pi \frac{j}{N}) + 1) \end{pmatrix}, \quad j \in \{1, \dots, N\}. \tag{6.10}$$

This choice is just one possibility and by no means unique. (We chose a different "frequency" for $\bar{\alpha}^j$ than for $\bar{x}^j$ to avoid a linear relationship between $\bar{\alpha}^j$ and $\bar{x}^j_1$, which would be undesirable structure in the data in this example.) The resulting data set for Algorithm 6.1 is

$$\mathcal{D} := \{(\bar{x}^j, \bar{\alpha}^j) \in \mathbb{R}^2 \times \Delta_2 : j \in \{1, \dots, N\}\}$$

and we choose $N = 100$ as the number of data points for our computations. For the set of basis functions, we choose the monomials up to degree 3, i.e.,

$$\mathcal{B} := \{x_1, x_1^2, x_1^3, x_2, x_1x_2, x_1^2x_2, x_2^2, x_1x_2^2, x_2^3\}.$$

The singular values of the resulting matrix $\mathcal{L} \in \mathbb{R}^{200 \times 18}$ are shown in Figure 6.3. The two smallest singular values $s_1 = 1.2 \cdot 10^{-15}$ and $s_2 = 1.66 \cdot 10^{-15}$ are practically

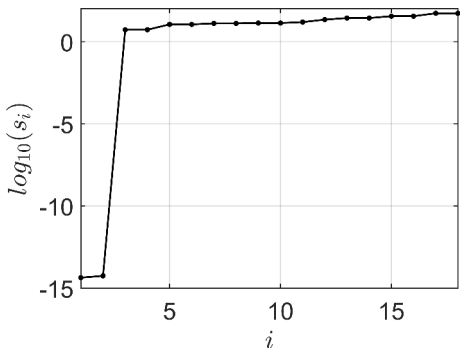


Figure 6.3: Singular values of $\mathcal{L}$ in Example 6.2.1.

zero as they are close to machine precision ($= 2^{-52} \approx 2.22 \cdot 10^{-16}$). This indicates that there are objective vectors in the span of our chosen basis functions for which the data is extended Pareto critical. From the second to the third smallest singular value $s_3 = 1.71$ there is an obvious gap, which motivates us to use $\bar{s} = s_2$ as the threshold for Algorithm 6.1. Thus, in step 4, we have to consider the span of the right-singular vectors corresponding to s_1 and s_2, which are

$$v_1 = (-0.8390, 0, 0.2797, 0, 0, 0, 0, 0, -0.1127, 0, 0, 0.2797, 0.3380, 0, 0, 0, 0, -0.1127)^\top,$$
$$v_2 = (0.3380, 0, -0.1127, 0, 0, 0, 0, 0, -0.2797, 0, 0, -0.1127, 0.8390, 0, 0, 0, 0, -0.2797)^\top,$$

respectively. In this example, considering the structure of these two vectors leads to the observation that

$$\begin{aligned}&\operatorname{span}(\{v_1, v_2\})\\&= \{(-3\sigma_1, 0, \sigma_1, 0, 0, 0, 0, 0, \sigma_2, 0, 0, \sigma_1, -3\sigma_2, 0, 0, 0, 0, \sigma_2)^\top : \sigma_1, \sigma_2 \in \mathbb{R}\}. \quad (6.11)\end{aligned}$$

While all coefficient vectors in this span would lead to objective vectors for which the data is extended Pareto critical, not all of these vectors are desirable. For example, choosing $\sigma_1 = 0$ and $\sigma_2 = 1$ leads to the coefficient vector

$$c = (0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, -3, 0, 0, 0, 0, 1)^\top,$$

which corresponds to the objective vector

$$\mathcal{F}(c)(x) = \begin{pmatrix} x_2^3 \\ x_2^3 - 3x_2 \end{pmatrix}. \quad (6.12)$$

It is possible to show that the data set $\mathcal{D}$ is contained in the extended Pareto critical set of this objective vector, as expected. However, the full Pareto critical set of this vector is given by $\mathbb{R} \times [-1, 1]$, which clearly contains significantly more than just our data set. In this case, it is easy to see that this degeneracy is caused by the fact that the objective vector does not depend on x_1. A better choice for a coefficient vector would be, e.g., $\sigma_1 = 1$ and $\sigma_2 = 1$, resulting in

$$c = (-3, 0, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, -3, 0, 0, 0, 0, 1)^\top$$

with the corresponding objective vector

$$f(x) := \mathcal{F}(c)(x) = \begin{pmatrix} -3x_1 + x_1^3 + x_2^3 \\ -3x_2 + x_1^3 + x_2^3 \end{pmatrix}.$$

It is easy to show that for this objective vector, the KKT condition is equivalent to

$$\begin{aligned} x_1^2 + x_2^2 &= 1, \\ \alpha_1 &= x_1^2, \\ \alpha_2 &= x_2^2, \end{aligned}$$

such that the Pareto critical set is precisely $\mathcal{S}^1$. The extended Pareto critical set of f and the data $\mathcal{D}$ are shown in Figure 6.4(a). Figure 6.4(b) shows the image of the Pareto critical set and a discretization of the image of f (around $\mathcal{S}^1$). It indicates that roughly half of the Pareto critical points of f are at least locally Pareto optimal.

Remark 6.2.2. *a) Motivated by the structure of f in Example 6.2.1, it is easy to show that the Pareto critical set of*

$$f : \mathbb{R}^n \to \mathbb{R}^n, \quad x \mapsto \begin{pmatrix} -3x_1 + \sum_{i=1}^n x_i^3 \\ \vdots \\ -3x_n + \sum_{i=1}^n x_i^3 \end{pmatrix}$$

is the $(n-1)$-dimensional unit sphere $\mathcal{S}^{n-1}$ in $\mathbb{R}^n$.

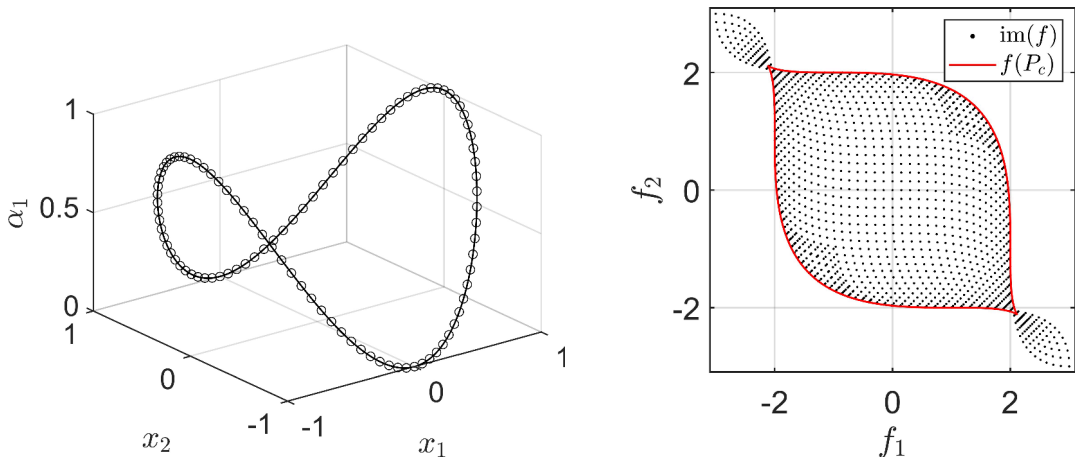


Figure 6.4: **(a)** Projection of the data $\mathcal{D}$ (circles) and the extended Pareto critical set of f (line) in Example 6.2.1 onto (x_1, x_2, α_1). **(b)** Image of the Pareto critical set of f and a pointwise approximation of $f([-1,1]^2)$.

b) With the same strategy as in Example 6.2.1, it is possible to show that arbitrary ellipses can be represented as Pareto critical sets of objective vectors consisting of polynomials of degree 3. *For* $a, b \in \mathbb{R}^{>0}$, *we merely have to replace* $\bar{x}^j$ *in Example 6.2.1 by*

$$\bar{x}^j := \begin{pmatrix} a \cdot \cos(2\pi \frac{j}{N}) \\ b \cdot \sin(2\pi \frac{j}{N}) \end{pmatrix}, \quad j \in \{1, \dots, N\}.$$

This leads to the same gap in the singular values and analogously to the expression (6.11), *we have*

$$\begin{aligned} &\operatorname{span}(\{v_1, v_2\}) \\ &= \{(-3a^2\sigma_1, 0, \sigma_1, 0, 0, 0, 0, 0, \sigma_2, 0, 0, \sigma_1, -3b^2\sigma_2, 0, 0, 0, 0, \sigma_2)^\top : \sigma_1, \sigma_2 \in \mathbb{R}\}. \end{aligned}$$

For $\sigma_1 = 1$ *and* $\sigma_2 = 1$, *we obtain the objective vector*

$$\mathcal{F}(c)(x) = \begin{pmatrix} -3a^2x_1 + x_1^3 + x_2^3 \\ -3b^2x_1 + x_1^3 + x_2^3 \end{pmatrix}.$$

It is easy to show that the corresponding Pareto critical set is given by

$$P_c = \left\{ x \in \mathbb{R}^2 : \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \right\},$$

which is the ellipse centered at $(0,0)^\top$ *with radii* (a, b).

In Example 6.2.1 we were able to derive an objective vector with a simple expression whose Pareto critical set was precisely what we prescribed. Clearly, this comes down to the simplicity of the data set and will not work in general. To this end, a more complicated case is presented in the following example, where the prescribed data set consists of three connected components.

Example 6.2.3. *In this example, we are looking for an MOP where the Pareto critical set contains three connected components $C_i \subseteq \mathbb{R}^2$, $i \in \{1,2,3\}$, given by the following three (non-intersecting) straight lines:*

$$C_i := p_i + [0,1]\frac{1}{4}\frac{q_i}{\|q_i\|}$$

with

$$p_1 = \begin{pmatrix} 0.15 \\ -0.20 \end{pmatrix},\ q_1 = \begin{pmatrix} 0.47 \\ 0.04 \end{pmatrix},$$
$$p_2 = \begin{pmatrix} 0.47 \\ -0.32 \end{pmatrix},\ q_2 = \begin{pmatrix} 0.40 \\ 0.14 \end{pmatrix},$$
$$p_3 = \begin{pmatrix} 0.37 \\ 0.18 \end{pmatrix},\ q_3 = \begin{pmatrix} 0.38 \\ 0.28 \end{pmatrix}.$$

To obtain $\mathcal{D}_x$, we discretize each C_i with 20 points as shown in Figure 6.6(a). The corresponding points in $\mathcal{D}_\alpha$ are chosen linearly from $(0,1)^\top$ to $(1,0)^\top$. By this choice, both boundary points of each C_i have a zero multiplier. For the basis functions we again use monomials. For more complicated data sets like this one, it makes sense to first analyze what degree of monomials we need for a satisfactory approximation before actually choosing an objective vector. To this end, we repeat step 2 of Algorithm 6.1 for different degrees and look at the smallest singular value for each degree. The result is shown in Figure 6.5(a). We see that the monomials up to a

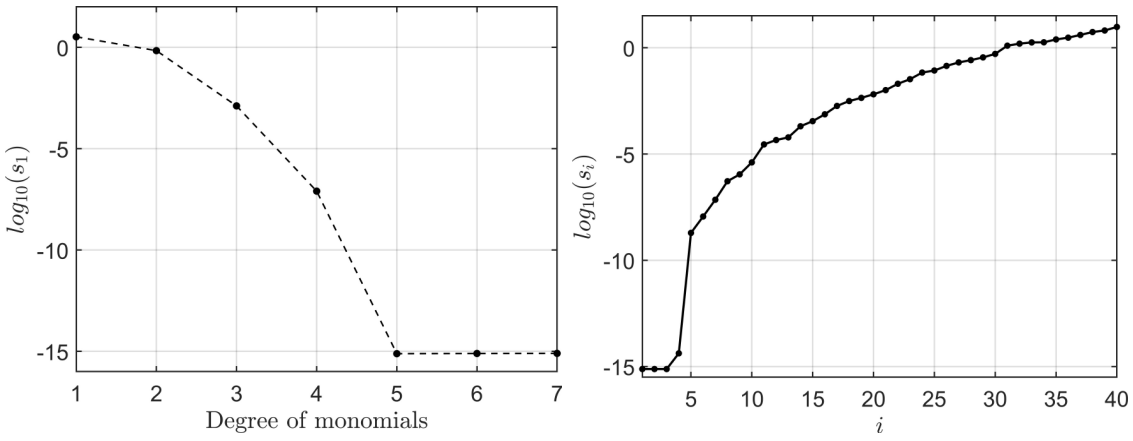
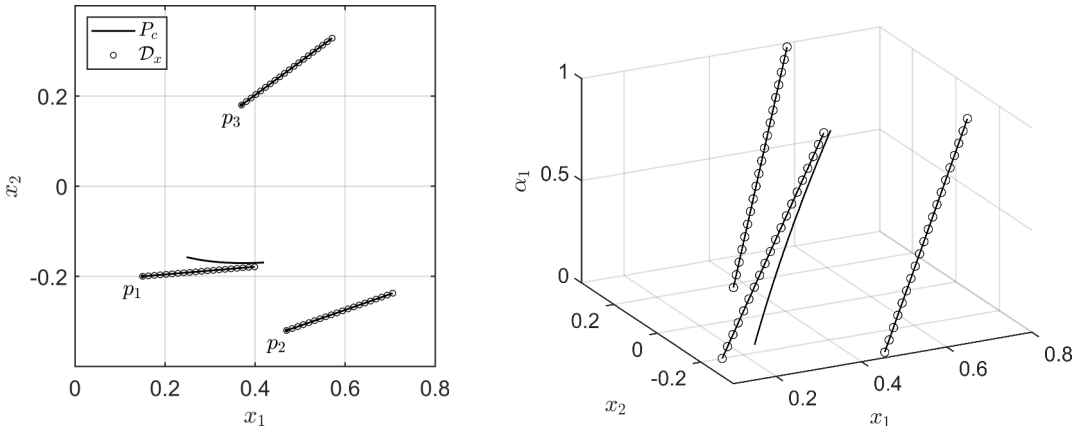


Figure 6.5: **(a)** Smallest singular value s_1 for different degrees of monomials in Example 6.2.3. **(b)** Singular values of $\mathcal{L}$ for polynomials of degree 5.

degree of 5 are a promising choice, since the smallest singular value is close to machine precision and therefore no meaningful further decrease is possible for higher degrees. All singular values of the matrix $\mathcal{L}$ corresponding to monomials of degree 5 are shown in Figure 6.5(b). We see that the first three singular values are identical with a value of $s_1 = s_2 = s_3 = 7.67 \cdot 10^{-16}$. Subsequently, there is a small gap from s_3 to $s_4 = 4.21 \cdot 10^{-15}$ and then a large gap from s_4 to $s_5 = 1.95 \cdot 10^{-9}$. Since the first three singular values are small and identical, we choose the threshold $\bar{s} = s_3$, such that $I = \{1,2,3\}$ in step 3 of Algorithm 6.1. As this example is more complex, there is no obvious way to obtain an expression like (6.11) in Example 6.2.1.

Thus, we simply choose $c = \frac{v_1+v_2+v_3}{\|v_1+v_2+v_3\|}$ *as the coefficient vector. The corresponding Pareto critical set and the extended Pareto critical set of the resulting objective vector* $f := \mathcal{F}(c)$ *can be computed via the continuation method from Chapter 3 and are shown in Figure 6.6. As expected from the small singular values, the prescribed*



(a) **(b)**

Figure 6.6: **(a)** Pareto critical set of f and the data $\mathcal{D}_x$ in Example 6.2.3. **(b)** Projection of the extended Pareto critical set (line) and the data (circles) onto (x_1, x_2, α_1).

data is (extended) Pareto critical for f*. But unfortunately, we observe an additional connected component that is not contained in the data. As mentioned in Remark 6.1.5(c), since our approach only assures that the data is contained in the extended Pareto critical set of our computed objective vector, additional Pareto critical points have to be expected in the general case. Finally, the image of the Pareto critical set and the image of* f *(around the data) are shown in Figure 6.7. It suggests that part*

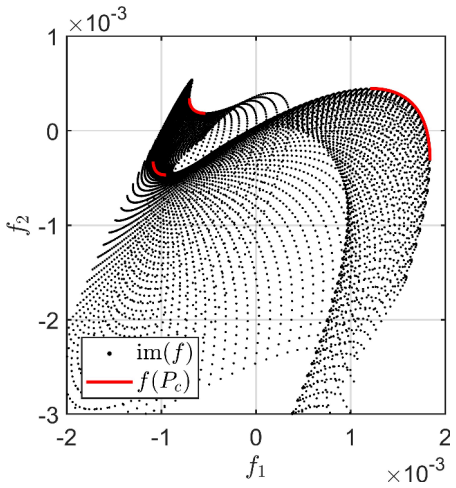


Figure 6.7: Image of the Pareto critical set of f and a pointwise approximation of $f([0, 0.8] \times [-0.4, 0.4])$ in Example 6.2.3.

of the Pareto critical set corresponds to (local) "Pareto maximal" points of f*.*

6.2.2 Inferring objectives of stochastic MOPs

In the previous application, we were looking for MOPs for which the prescribed data is exactly extended Pareto critical. We had the chance to find such MOPs since the data sets were precise pointwise discretizations of lower-dimensional sets with the theoretical properties of extended Pareto critical sets. Clearly, having exact data like that is a very strong assumption which cannot be made in other applications, where the data is not carefully constructed by hand. Thus, we generally have to expect that the data contains some amount of noise. To test the robustness of our method under noise, we will now consider data that comes from stochastic MOPs. Only a brief introduction to stochastic multiobjective optimization will be given here. For more details, see [FX11; GP13].

For $m \in \mathbb{N}$ let $\xi \in \mathbb{R}^m$ be a random vector and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$. Let

$$F : \mathbb{R}^n \to \mathbb{R}^k, \quad x \mapsto \mathbb{E}[f(x, \xi)],$$

where $\mathbb{E}[f(x, \xi)]$ is the (component-wise) expected value of $f(x, \xi)$ with respect to ξ. Then

$$\min_{x \in \mathbb{R}^n} F(x) \tag{SMOP}$$

is a *stochastic multiobjective optimization problem* (SMOP). We assume that the expected value is unknown, so that we cannot directly evaluate the objective vector F. To approximate it, we use the average

$$\tilde{f}^{N_s}(x) := \frac{1}{N_s} \sum_{j=1}^{N_s} f(x, \xi^j),$$

where $\xi^j \in \mathbb{R}^m$, $j \in \{1, \dots, N_s\}$, is an independently and identically distributed random sample of N_s realizations of ξ. The corresponding MOP

$$\min_{x \in \mathbb{R}^n} \tilde{f}^{N_s}(x) \tag{SAA}$$

is the so-called *sample average approximation problem* (SAA). Since the objective vector of (SAA) is an approximation of the objective vector of (SMOP), we expect the solution of (SAA) to be an approximation of the solution of (SMOP). In the following, we will check if the original objective vector of (SMOP) can (theoretically) be found with our inverse approach by using extended Pareto critical points of (SAA) as data. As an example, we consider the *Multiobjective Stochastic Location Problem* from [FX11].

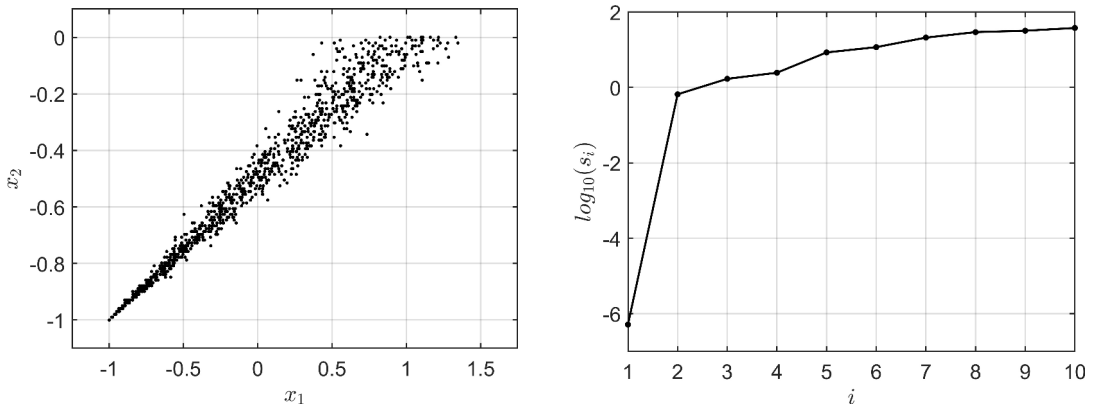
Example 6.2.4. *Let $a = (-1, -1)^\top$ and $\xi = (\xi_1, 0)^\top$ be a random vector with ξ_1 being uniformly distributed on $[0, 2]$. Let*

$$f(x, \xi) := \begin{pmatrix} \|x - a\|^2 \\ \|x - \xi\|^2 \end{pmatrix}.$$

The expected value of f can be computed by hand, resulting in

$$F(x) = \mathbb{E}[f(x, \xi)] = \begin{pmatrix} 2x_1 + x_1^2 + 2x_2 + x_2^2 + 2 \\ -2x_1 + x_1^2 + x_2^2 + \frac{4}{3} \end{pmatrix} = \begin{pmatrix} \|x - a\|^2 \\ \|x - (1, 0)^\top\|^2 + \frac{1}{3} \end{pmatrix}.$$

Thus, the Pareto critical (and in this case optimal) set of (SMOP) *is given by the line connecting* a *and* $(1,0)^\top$ *(cf. Example 2.1.6). For the sample average approximation of* F*, we use a sample size of* $N_s = 10$*. For the solution of* (SAA)*, we apply the weighting method (cf. Section 2.3.1) with* 100 *equidistant weights from* Δ_2 *and solve each of the resulting scalar problems* 10 *times (via the MATLAB function* `fminunc` *and with different realizations of* ξ *every time). The result is shown in Figure 6.8(a). Since the first objective is deterministic, the approximation is relatively*



(a) **(b)**

Figure 6.8: **(a)** Approximation of the solution of (SMOP) via the sample average approximation and weighted sum in Example 6.2.4. **(b)** Singular values of $\mathcal{L}$.

accurate close to $a = (-1,-1)^\top$ *and becomes worse when moving towards* $(1,0)^\top$*.*

For our inverse approach, we now use the points in Figure 6.8(a) as the data set $\mathcal{D}_x$*. For each* $\bar{x} \in \mathcal{D}_x$*, the point in* $\mathcal{D}_\alpha$ *corresponding to* $\bar{x}$ *is chosen as the weight in the weighting method that was used to compute* $\bar{x}$*. As basis functions we use the monomials up to degree* 2*, i.e.,*

$$\mathcal{B} := \{x_1, x_2^2, x_2, x_1x_2, x_2^2\}. \tag{6.13}$$

The singular values of $\mathcal{L}$ *when applying Algorithm 6.1 are shown in Figure 6.8(b). While* $s_1 = 5.19 \cdot 10^{-7}$ *is relatively small, the corresponding right-singular vector is given by*

$$v_1 = (0, 0, -0.8165, 0, -0.4082, 0, 0, 0, 0, -0.4082)^\top,$$

such that $\mathcal{F}(v_1)(x) = (-0.8165x_2 - 0.4082x_2^2, -0.4082x_2^2)^\top$*. Clearly, this objective vector is degenerate due to the missing dependency on* x_1*. Thus, we choose a larger threshold to try to obtain a more regular function. The next singular values are*

$$\begin{aligned} s_2 &= 0.6552, \\ s_3 &= 1.6864, \\ s_4 &= 2.4426, \\ s_5 &= 8.4318. \end{aligned}$$

Due to the slight gap from s_4 *to* s_5 *we choose* $\bar{s} = s_4$ *as the threshold, such that* $I = \{1,2,3,4\}$ *in step 3 of Algorithm 6.1. For better interpretability in step 4, we*

compute a sparse basis $\{w_1, w_2, w_3, w_4\}$ *of* $\operatorname{span}(\{v_1, v_2, v_3, v_4\})$ *(by computing linear independent elements with a small* ℓ_1*-norm), resulting in*

$$\begin{aligned}
w_1 &= (-0.4762, -0.2393, 0, 0, -0.0102, 0, 0.0048, 0.9626, -0.9282, 1)^\top,\\
w_2 &= (0, 0, -1, 0, -0.5, 0, 0, 0, 0, -0.5)^\top,\\
w_3 &= (0, 0.2444, -0.4969, -0.4880, 0.0002, 0, -0.0034, -0.4908, 0.4768, -1)^\top,\\
w_4 &= (-1, -0.5068, 0, -0.0040, -0.0001, 0.9563, -0.4468, -0.0026, 0, 0)^\top.
\end{aligned}$$

For our choice of basis functions, the objective vector F *of* (SMOP) *can be represented exactly (up to the constants in both components) by the coefficient vector*

$$\bar{c} = (2, 1, 2, 0, 1, -2, 1, 0, 0, 1)^\top. \tag{6.14}$$

In other words, $F - (2, \frac{4}{3})^\top = \mathcal{F}(\bar{c})$*. Considering our sparse basis, we see that its span contains*

$$\begin{aligned}
c^* &:= -2w_2 - 2w_4\\
&= (2, 1.0136, 2, 0.0080, 1.0001, -1.9126, 0.8936, 0.0051, 0, 1)^\top,
\end{aligned}$$

which is close to the original $\bar{c}$*. The Pareto critical sets of* F *and* $\mathcal{F}(c^*)$ *are shown in Figure 6.9. A numerical approximation of their Hausdorff distance (using a point-*

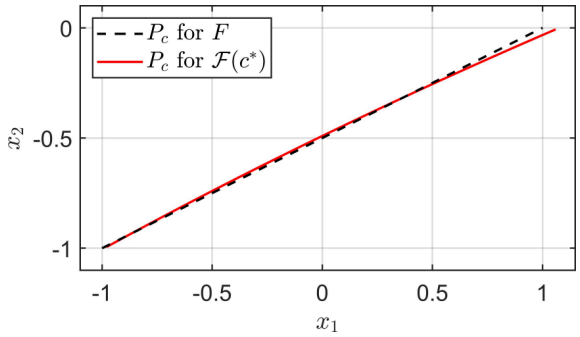


Figure 6.9: Pareto critical sets of F and $\mathcal{F}(c^*)$ in Example 6.2.4.

wise discretization) yields a distance of 0.06*. In terms of the functions themselves, a comparison of their values (up to constants) around the Pareto critical set yields*

$$\max_{x \in [-1.1, 1.1] \times [-1.1, 0.1]} \|(F(x) - (2, \tfrac{4}{3})^\top) - \mathcal{F}(c^*)(x)\|_\infty \approx 0.2308.$$

In our example, the objectives of (SMOP) were contained in the span of our chosen basis functions, which is clearly a special case. Furthermore, we were only able to reconstruct the corresponding coefficient vector via $\bar{c} \approx -2w_2 - 2w_4$ because we already knew $\bar{c}$. In the general case, there is no obvious way how F can be reconstructed without assuming further knowledge about the structure of the problem. (In Example 6.2.4, this further knowledge could be the observation that the randomness only occurs in a certain part of the second objective function.) Thus, we only showed the potential of our approach here and more work is needed to be able to actually infer objective vectors of stochastic MOPs in the general case. Nonetheless, our result demonstrates the ability of our approach to handle noisy data.

6.2.3 Generation of surrogate models

By Theorem 6.1.4, if there is no objective vector in the span of our basis functions for which the data is exactly extended Pareto critical, then we at least find an objective vector for which the data is "almost" (depending on the singular values) extended Pareto critical. In particular, if we use data which is extended Pareto critical for an objective vector f^e which cannot be expressed in our basis functions, then our approach generates an objective vector in our basis which has at least a similar extended Pareto critical set. In other words, our approach generates a surrogate model for the original objective vector f^e. If our basis functions are cheap to evaluate, then all objective vectors resulting from our method are cheap as well, as they are just linear combinations of those basis functions. In this way, our method can be used to generate cheap surrogate models for (potentially expensive) objective vectors.

Computationally expensive objectives frequently occur in practical applications. For example, in the optimization of physical systems, every evaluation of the objectives might involve the solution of a complex PDE [Lot+05; Trö09]. While it may be possible to compute a few single Pareto critical points of such a problem, the approximation of the full Pareto critical set with standard methods is often computationally infeasible. For the generation of surrogate models with our method, the idea is to only compute a few extended Pareto critical points of the expensive problem and then use those points as data in our inverse approach. It is important to note that in contrast to traditional surrogate modeling (see [BGW15; SVR08] for overviews), such a surrogate model is not an approximation of f^e in the sense of function values, but an approximation in the sense of their extended Pareto critical sets. In particular, the actual Pareto set of the surrogate model will generally differ from the Pareto set of the original problem.

For our inverse approach, the data must consist of Pareto critical points and corresponding KKT vectors of the original problem. Although the main goal of solution methods for MOPs is the computation of Pareto critical (or optimal) points, many standard methods also implicitly provide the corresponding KKT vectors:

- Weighting method: If $x^* \in \arg\min_{x\in\mathbb{R}^n} \sum_{i=1}^k \alpha_i^* f_i(x)$ for some weighting vector $\alpha^* \in \Delta_k$, then α^* is a KKT vector of x^* (by the first-order optimality condition).

- ε-constraint method: For $j \in \{1,\dots,k\}$ and $\varepsilon_i \in \mathbb{R}$, $i \in \{1,\dots,k\} \setminus \{j\}$, let x^* be a solution of

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f_j(x), \\ s.t. \quad & f_i(x) \leq \epsilon_i \ \forall i \in \{1,\dots,k\} \setminus \{j\}. \end{aligned}$$

  By the first-order optimality condition for this problem, there are $\mu_i^* \geq 0$, $i \in \{1,\dots,k\} \setminus \{j\}$, such that

$$\nabla f_j(x^*) + \sum_{i\neq j} \mu_i^* \nabla f_i(x^*) = 0.$$

  Let

$$\alpha^* = \frac{1}{\alpha_1 + \ldots + \alpha_k}\alpha \quad \text{with} \quad \alpha_i = \begin{cases} \mu_i^*, & \text{if } i \neq j, \\ 1, & \text{if } i = j. \end{cases}$$

Then α^* is a KKT vector of x^*.

- Reference point method (cf. Section 4.5 in [Ehr05]): Let $z \in \mathbb{R}^k$ and x^* be a solution of

$$\min_{x \in \mathbb{R}^n} \|f(x) - z\|^2$$

with $z_i \leq f_i(x^*)$ for all $i \in \{1, \ldots, k\}$ and $z \neq f(x^*)$. By the first-order optimality condition, we have

$$0 = \nabla(x \mapsto \|f(x) - z\|^2)(x^*) = \sum_{i=1}^{k} 2(f_i(x^*) - z_i)\nabla f_i(x^*).$$

By assumption, $f_i(x^*) - z_i \geq 0$ for all $i \in \{1, \ldots, k\}$ and $\sum_{i=1}^{k} f_i(x^*) - z_i > 0$, so a KKT vector of x^* is given by

$$\alpha^* = \frac{1}{\sum_{i=1}^{k} f_i(x^*) - z_i}(f(x^*) - z).$$

If a solution method does not implicitly provide KKT vectors during its application, like evolutionary algorithms, it is also possible to obtain them a posteriori by considering the image space. Recall that by Lemma 2.2.7, a KKT vector of $x^* \in P_c$ is given by the (positive) normal vector of $f(P_c)$ at $f(x^*)$. Thus, linear regression can be used to approximate the KKT vectors. But note that this requires a certain density of the approximation of the Pareto front, which can be difficult to obtain for expensive MOPs.

When searching for a surrogate model, it is important to balance out the number of basis functions and the number of data points to avoid *underfitting* and *overfitting*. These terms are common in statistics and machine learning and apply here in a similar fashion. In general, underfitting means that the chosen model is not able to capture all the features that are present in the data set. In our context, this means that we chose an inappropriate (e.g., too small) set of basis functions. When using monomials as basis functions, we can try to fix this by using a higher maximal degree. On the other hand, overfitting means that the model captures features in the data set that were caused by noise and are highly dependent on the specific data set that was used. In our context, this happens when the number of basis functions d is too large. A necessary condition to circumvent overfitting is that the linear system corresponding to the matrix $\mathcal{L} \in \mathbb{R}^{(n \cdot N) \times (k \cdot d)}$ from (6.5) is overdetermined, i.e., that

$$d \leq \frac{n \cdot N}{k}. \tag{6.15}$$

As discussed in Section 6.1, if this condition does not hold, then we always find an objective vector in the chosen basis for which all data points are extended Pareto critical. Thus, if (6.15) is violated, then overfitting is unavoidable.

As a first example for the generation of surrogate models, we consider the problem $L\&H_{2\times 2}$ from [HL14]. While its objective vector is cheap to evaluate, its Pareto critical set has interesting topological properties.

Example 6.2.5. *Consider the problem*

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & f^e(x), \\ s.t. \quad & x \in [-0.75, 0.75] \times [-2.5, 0.12], \end{aligned} \tag{$L\&H_{2\times 2}$}$$

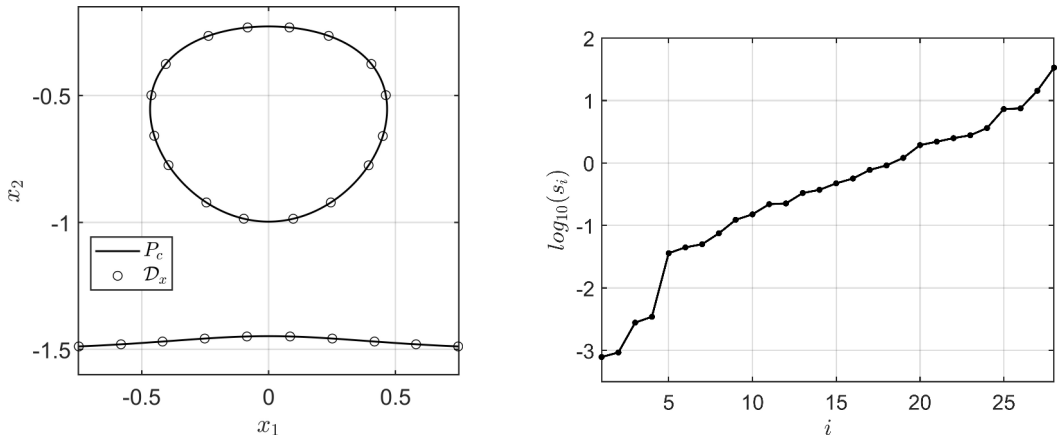
for

$$f^e(x) := -\begin{pmatrix} \frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}b(x) \\ -\frac{\sqrt{2}}{2}x_1 + \frac{\sqrt{2}}{2}b(x) \end{pmatrix}$$

with

$$\begin{aligned} b(x) &:= 0.2g(x, (0,0)^\top, 0.65) + 1.5g(x, (0,-1.5)^\top, 2.8), \\ g(x, p_0, \sigma) &:= \sqrt{\frac{2\pi}{\sigma}} \exp\left(-\frac{\|x - p_0\|_2^2}{\sigma^2}\right). \end{aligned}$$

The Pareto critical set of this problem (obtained via the continuation method from Section 3.1) is shown in Figure 6.10(a). Although this problem is constrained, all



(a) **(b)**

Figure 6.10: **(a)** Pareto critical set and data set $\mathcal{D}_x$ in Example 6.2.5. **(b)** Singular values of $\mathcal{L}$.

Pareto critical points on the boundary of the feasible set are also Pareto critical for the unconstrained problem. Thus, we will ignore the constraints for our purposes. For the construction of the surrogate model, we use the $N = 26$ data points in Figure 6.10(a). The corresponding KKT vectors were explicitly computed in the continuation method (cf. $X_{\mathcal{C}}$ in Algorithm 3.1). For the basis functions, we choose the monomials up to degree 4. By this choice we have

$$d = 14 \leq 26 = \frac{n \cdot N}{k},$$

such that the condition (6.15) is satisfied. The singular values of the resulting $\mathcal{L}$ in step 2 of Algorithm 6.1 are shown in Figure 6.10(b). In step 3 and 4, we simply choose the right-singular vector corresponding to the smallest singular value given by $s_1 = 7.84 \cdot 10^{-4}$, i.e., $c = v_1$. A comparison of the Pareto critical set of the resulting objective vector $f := \mathcal{F}(c)$ and the original f^e is shown in Figure 6.11. First of all,

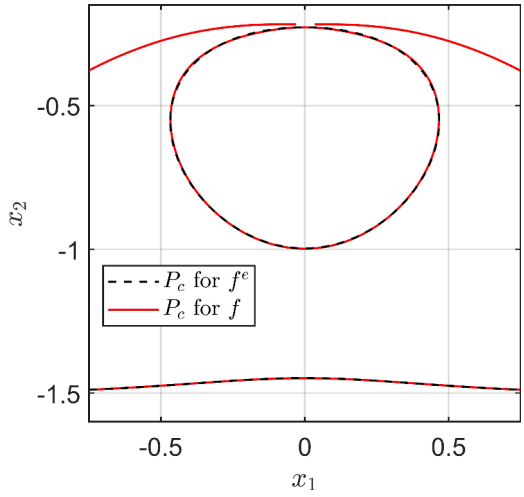


Figure 6.11: Pareto critical sets of the surrogate model $f = \mathcal{F}(c)$ and the original objective vector f^e in Example 6.2.5.

we see that all Pareto critical points of f^e are almost exactly Pareto critical for the surrogate model. On top of that, the topology of the two connected components of the original Pareto critical set was captured correctly, such that we obtained an almost perfect approximation with our surrogate model. Unfortunately, the surrogate model possesses additional connected components that were not contained in the data. As already seen in Example 6.2.3, such features can generally not be avoided.

In the next example, we will consider the MOP from [POBD18] that describes the optimal control of the Navier-Stokes equations. We will only give a short introduction of the problem here and refer to Section 2 in [POBD18] for the details.

Example 6.2.6. *In this example, we will consider the flow around a rotating cylinder governed by the 2D incompressible Navier-Stokes equation. The goal is to influence the flow field by controlling the rotation of the cylinder, as shown in Figure 6.12. The PDE that describes the physics behind this is given by*

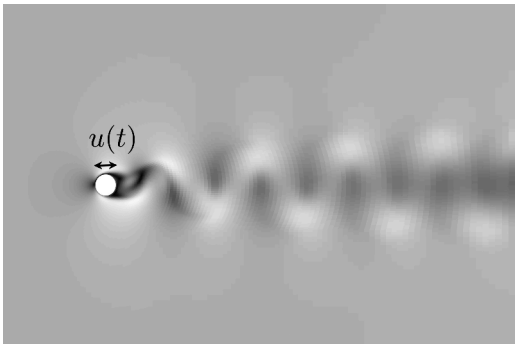


Figure 6.12: Flow around a cylinder and the resulting flow field in Example 6.2.6.

$$\begin{aligned} \frac{\partial y(x,t)}{\partial t} + (y(x,t)\cdot\nabla)y(x,t) &= -\nabla p(x,t) + \frac{1}{Re}\nabla^2 y(x,t), \\ \nabla \cdot y(x,t) &= 0, \\ (y(x,0), p(x,0)) &= (y_0(x), p_0(x)), \end{aligned} \tag{6.16}$$

where y is the flow velocity and p is the pressure. The control input u is introduced via a time-dependent Dirichlet boundary condition on the surface of the cylinder. If the cylinder does not rotate, then the so-called von Kármán vortex street *occurs, which is a periodic solution where vortices detach alternatingly from the upper and lower edge of the cylinder, as shown in Figure 6.12. The classical goal of the control problem is to stabilize the flow, i.e., to minimize the vertical velocity. This is associated with minimizing the vertical force on the cylinder, which is referred to as the* lift C_L*. The second objective is the minimization of the control effort. The resulting MOP is*

$$\min_{u \in L^2([t_0,t_e],\mathbb{R})} \begin{pmatrix} \int_{t_0}^{t_e} C_L(t)^2 dt \\ \int_{t_0}^{t_e} u(t)^2 dt \end{pmatrix}, \quad \text{s.t.} \quad (6.16) \text{ holds.} \tag{6.17}$$

By replacing the general control u with a sinusodial control $u(t) = x_1 \sin(2\pi x_2 t)$ and assuming injectivity of the control-to-state mapping, (6.17) *can be written as*

$$\min_{x \in \mathbb{R}^2} f^e(x) \quad \text{with} \quad f^e(x) := \begin{pmatrix} \int_{t_0}^{t_e} C_L(t)^2 dt \\ \int_{t_0}^{t_e} (x_1 \sin(2\pi x_2 t))^2 dt \end{pmatrix}, \tag{6.18}$$

which is now a finite-dimensional MOP. Since each evaluation of f^e requires the solution of the nonlinear PDE (6.16) *(e.g., via the finite element method), it is computationally infeasible to solve* (6.18) *with a standard solution method.*

To generate the data for our inverse approach, we apply the weighting method to the problem (6.18) *with the equidistant weights*

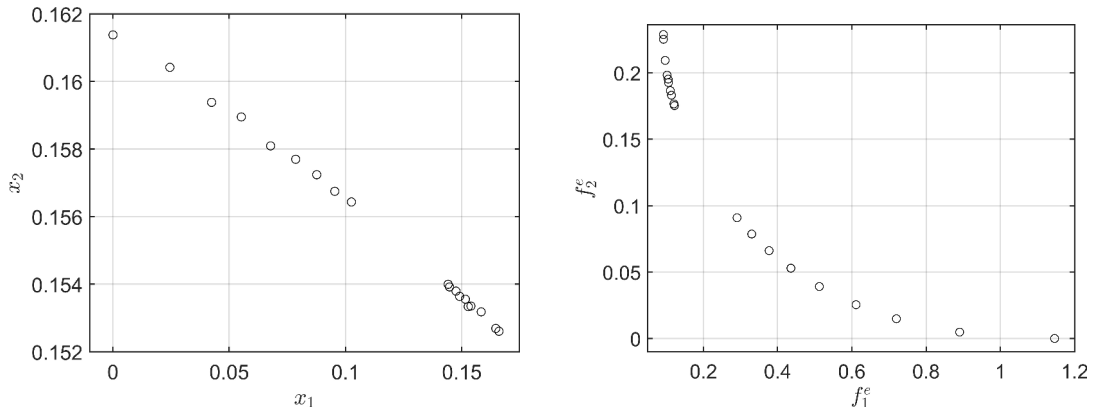
$$\alpha^j = \frac{1}{25}(j, 25-j)^\top \in \Delta_2, \quad j \in \{0, \dots, 25\}.$$

The resulting scalar problems are solved in MATLAB via the function `fminunc`*. As discussed at the beginning of this section, the KKT vectors corresponding to the solutions of the weighting method are given by the α^j. Since the solver for the scalarized problem did not converge for $j \in \{9, \dots, 15\}$, these weights were excluded from the data set. The remaining* 19 *points in $\mathcal{D}_x$ and their images under f^e are shown in Figure 6.13. For the basis functions, we choose the monomials up to degree* 2 *as in* (6.13)*. The singular values of the corresponding $\mathcal{L} \in \mathbb{R}^{38\times 10}$ are shown in Figure 6.14(a). The first two singular values $s_1 = 2.82 \cdot 10^{-4}$ and $s_2 = 5.94 \cdot 10^{-4}$ are of the same small magnitude. Unfortunately, for the corresponding right-singular vectors v_1 and v_2, all entries related to basis functions containing x_1 are relatively small, such that the resulting objective vectors are degenerate (similar to* (6.12) *in Example 6.2.1). The next singular value is $s_3 = 5.96 \cdot 10^{-3}$, and the objective vector corresponding to the third right-singular vector v_3 is given by*

$$\begin{aligned} f(x) &:= \mathcal{F}(v_3)(x) \\ &= \begin{pmatrix} -0.0519x_1^2 - 0.9285x_1x_2 + 0.1588x_1 + 0.1542x_2^2 + 0.1046x_2 \\ -0.0136x_1^2 - 0.2704x_1x_2 + 0.0437x_1 + 0.0054x_2^2 - 0.0008x_2 \end{pmatrix}. \end{aligned}$$

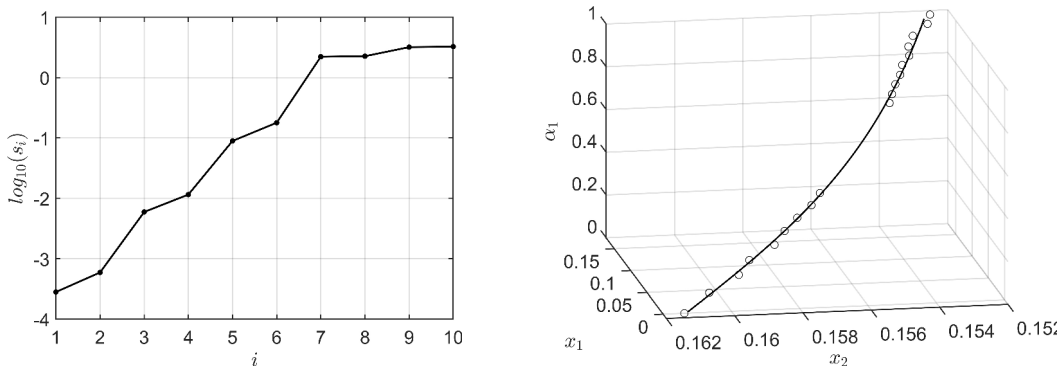
Figure 6.14(b) shows a comparison of the extended Pareto critical set of f and the data $\mathcal{D}$.

Since the actual Pareto critical set of the original problem (6.18) *is not known, we cannot definitively say if the Pareto critical set of our surrogate model is a good*



(a) **(b)**

Figure 6.13: **(a)** Data $\mathcal{D}_x$ resulting from the weighting method in Example 6.2.6. **(b)** Image of the data points under f^e.



(a) **(b)**

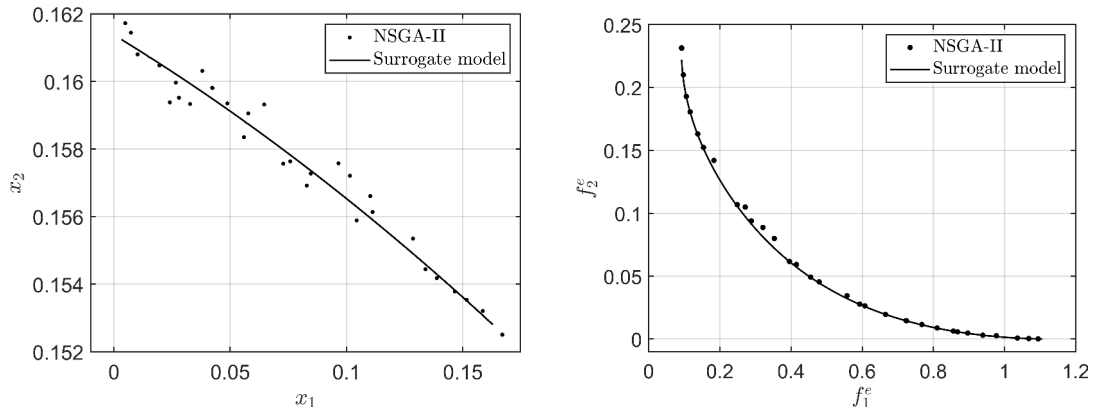
Figure 6.14: **(a)** Singular values of $\mathcal{L}$ in Example 6.2.6. **(b)** Projection of the extended Pareto critical set of the surrogate model f (line) and the data from the original problem (cricles) onto (x_1, x_2, α_1).

approximation. Nonetheless, we can at least try to compute a rough approximation of the Pareto set of (6.18) *using non-deterministic methods (cf. Section 2.3.4) and compare it to our result. To this end, we choose the NSGA-II algorithm (from MATLAB's Global Optimization Toolbox) and apply it directly to* (6.18)*. The result of NSGA-II and the Pareto critical set of our surrogate model are shown in Figure 6.15(a). To compare the results in the image space, we evaluated the original objective vector f^e in the Pareto critical points of our surrogate model f. The result is shown in Figure 6.15(b). It suggests that the Pareto critical set of our surrogate model is indeed a good approximation of the Pareto (critical) set of the original problem.*

6.3 Open problems of our approach

While several examples were presented where our method was able to compute satisfactory objective vectors from data, there are open problems that have to be addressed to make it applicable to general, non-academic examples:

- The requirement that the data does not only contain the Pareto critical points



(a) **(b)**

Figure 6.15: **(a)** Result of NSGA-II and the Pareto critical set of our surrogate model in Example 6.2.6. **(b)** The result of NSGA-II in the image space and the image of the Pareto critical set of our surrogate model under f^e.

but also the corresponding KKT vectors heavily restricts the range of application of our approach. But as seen in Example 6.0.1, this requirement is necessary to limit the number of different classes of solutions of (IMOPc). In [DZ18], this problem was avoided by restricting the search to convex objective vectors, where a pointwise approximation of the Pareto (critical) set can be obtained via the weighting method (for predetermined, fixed weighting vectors). In a way, their approach implicitly assigns a KKT vector to each of their data points. A similar idea could be introduced into our approach in (6.3), by considering α as a variable and not as given. But since this would break the linearity of (IMOPc) and likely require additional constraints to avoid trivial solutions, this is left for future work.

- The only result about the approximation quality of Algorithm 6.1 is Theorem 6.1.4, stating that the chosen threshold $\bar{s}$ is an upper bound for the norm of the KKT condition in the data points. Note that this bound is influenced by the scaling of the basis functions and cannot directly be used to obtain an estimate for the Hausdorff distance of the Pareto critical set to the data. Such an estimate would be helpful for the generation of surrogate models for expensive MOPs as a way to check if the Pareto critical set of the surrogate model is close to the actual (unknown) Pareto critical set. Thus, deeper analysis of the quality of our inferred objectives is needed.

- As discussed in Remark 6.1.5 (and seen in Examples 6.2.3 and 6.2.5), we can only expect the Pareto critical set P_c of the objective vector resulting from Algorithm 6.1 to be a superset of the data set $\mathcal{D}_x$. In particular, there can be many structures in P_c that were not contained in the data. Since this cannot be avoided in the general case, it might be helpful for some applications to develop a method for post-processing that filters out unwanted parts of P_c (e.g., using techniques from cluster analysis [ES00] to identify connected components of P_c).

- In step 4 of Algorithm 6.1, one has to choose an element in the span of certain right-singular vectors of $\mathcal{L}$. Depending on the chosen threshold, this span

can be large, and different elements can correspond to objective vectors with differing Pareto critical sets. If no additional information is available, then there is no favorable element in this span. This motivates the development of additional criteria in step 4 that may impose certain behavior or properties upon the resulting objective vector. For example, as already mentioned in Remark 6.1.5, sparsity may allow for better interpretability.

- By construction, the data points in $\mathcal{D}_x$ are (at most) Pareto critical for an objective vector resulting from our method, but not necessarily Pareto optimal. Thus, methods that can specifically compute (potentially non-optimal) Pareto critical sets have to be used when working with our inferred objective vectors. One example for such a method is our continuation method from Section 3.1.2. While it works well for the low-dimensional examples we considered here, it becomes more difficult in higher-dimensional cases. Thus, a way to assure that the data is actually Pareto optimal for our inferred objective vectors would make working with them a lot easier. A way to achieve this might be the consideration of sufficient optimality conditions involving second-order derivatives of the basis functions (cf. Theorem 3.2.17 in [Mie98]). But as the resulting conditions will likely cause the inverse problem to be nonlinear, this is left for future work.

7 Conclusion and outlook

In this thesis, numerical methods for the computation of Pareto critical sets for smooth and nonsmooth multiobjective optimization problems were presented. Furthermore, the hierarchical structure of Pareto critical sets was analyzed. Finally, the results about the structure were used to solve inverse multiobjective optimization problems. In this section, we will summarize the main results and discuss possible directions for future work.

Box-continuation methods for smooth problems

Chapter 3 was concerned with continuation methods for smooth problems based on box coverings of the Pareto critical set. For the case where exact gradients are available, we modified the continuation method from [SDD05] by introducing the problem (3.1), which checks whether a box contains Pareto critical points. We then proposed a novel extension to the case of equality and inequality constrained problems, which is based on the observation that active inequality constraints can (in some sense) be treated as additional equality constraints in the constrained KKT conditions. In the version of the method presented here (Algorithm 3.2), this required us to iterate over all possible combinations of active inequality constraints. As discussed in Remark 3.1.13, this is can be highly impractical. For future work, a method should be developed that filters out the combinations of active indices that are actually relevant before applying the continuation method. Alternatively, it may be possible to use the ideas from [BCS20], where the continuation method from [MS17] was generalized to inequality constrained problems.

For the case where only inexact gradient information is available in the sense of (3.9), we derived a tight superset P_ε^r of the Pareto critical set based on the inexact gradients and the upper bounds for the error. Since P_ε^r always has the same dimension as the variable space $\mathbb{R}^n$, it is more efficient to only compute a box covering of the boundary ∂P_ε^r of P_ε^r instead of P_ε^r itself. To this end, we showed that this boundary is contained in the zero level set of the function φ from Section 3.2.2 and then used a continuation method for its computation. While we analyzed some of the properties of φ to argue why we expect the zero level set to be piecewise smooth, we did not actually give a proof for this. For future work, an actual proof is required, which could possibly be achieved as outlined in Remark 3.2.10. Furthermore, the problem (3.19) that is used to check whether a box contains part of ∂P_ε^r can be difficult to solve, as the function $x \mapsto \varphi(x)^2$ that is minimized contains a minimization problem itself. It might be possible to make the solution more efficient by exploiting some of the structure of the problem, like the fact that the optimization problem (3.11) that is contained in φ is quadratic.

An efficient descent method for nonsmooth problems

In Chapter 4, we derived a descent method for multiobjective optimization problems where the objectives are assumed to be only locally Lipschitz continuous, i.e., potentially nonsmooth. Our method was based on [AGG15], where a descent direction is obtained as the element with the smallest norm in the negative convex hull of all Clarke subdifferentials of the objective functions. Due to the practical difficulties of working with the Clarke subdifferential, we modified this direction by instead using the Goldstein ε-subdifferential. For the approximation of the ε-subdifferential, we used the approach from [MAY12] and generalized it to the multiobjective case. Combined with an Armijo-like step length, this results in a descent method (Algorithm 4.3) that finds an (ε, δ)-critical point after a finite number of steps. By sequentially reducing ε and δ, we then obtained a method (Algorithm 4.4) that generates a sequence whose accumulation points are Pareto critical. A comparison to the multiobjective proximal bundle method from [MKW14] showed that the performance of our descent method is competitive. A combination with the subdivision method allowed us to compute the entire Pareto (critical) set. For future work, our descent method can likely be extended to the case of constrained nonsmooth MOPs by modifying the direction finding problem (4.7) to assure that the descent sequence remains feasible. Since our method basically reduces to the gradient descent method in the smooth and scalar case, we can only expect a linear convergence rate for our method. It might be possible to achieve a better rate by using variable norms in the direction finding problem, as was done in [CQ12] for the scalar case. Furthermore, it would be interesting to see if our method could be improved by using the generalized Jacobian of f (discussed in Remark 2.2.26) instead of the Clarke subdifferentials of the individual objectives. Finally, since the Clarke subdifferential can also be defined for locally Lipschitz functions on general Hilbert spaces, it might be possible to generalize our method to the infinite-dimensional case. This would make it possible to apply it directly to nonsmooth multiobjective optimal control problems.

Hierarchical structure of Pareto critical sets

In Chapter 5, we analyzed the structure of the Pareto critical set with respect to the Pareto critical sets of subproblems that only contain subsets of the set of objective functions. We began by classifying Pareto critical points in terms of the zero entries of their KKT vectors. Afterwards, we investigated the properties of tangent cones of the Pareto critical set to be able to define its boundary. This enabled us to show that the boundary of the Pareto critical set consists of Pareto critical sets of subproblems where at least one objective function is neglected. In other words, the boundary is covered by the union of all Pareto critical sets of subproblems of size at most $k-1$. Since this is not necessarily a tight covering, we then analyzed the smallest size of subproblems required to still obtain a covering of the boundary. It turned out that under certain regularity assumptions, the minimal required size is the maximal rank of the Jacobian of the objective vector f in the variable space. In particular, it is bounded by the number of variables n. For future work, it would be interesting to combine our results with the more abstract results in [Sma73; LP14] that were obtained using stratification theory. In particular, this might make it possible to relate our definition of the boundary (Definition 5.1.14) to a definition based on some manifold structure. Furthermore, since the Pareto critical set is just a superset of

the actual Pareto set, our results cannot directly be transferred to the Pareto set. Structural results about the Pareto set may be achievable by incorporating the sufficient optimality conditions based on the Hessians of the objective functions into our results.

After the unconstrained case, we discussed the structure of the boundary of the Pareto critical set in the presence of equality and inequality constraints. Here, we only presented the basic ideas without the actual proofs. Due to the structure of the constrained KKT conditions, our strategy was to interpret the constraints as additional objective functions and then apply our results from the unconstrained case. For the case where only equality constraints are present, we obtained the same result as in the unconstrained case. If we have inequality constraints, then the boundary may also contain points where the activity of the inequalities changes. Clearly, future work in the constrained case should start with rigorously proving our results. Furthermore, it might make sense to rethink the definition of the boundary via tangent cones in this case, as by this definition, "kinks" in the Pareto critical set related to the constraints also count as boundary points.

Following the smooth case with and without constraints, we considered the unconstrained, nonsmooth case. Due to the lack of smoothness, most of the tools we used before could not be applied here. Nonetheless, some basic results could be generalized. For the case $k > n+1$, a simple application of Carathéodory's theorem showed that the Pareto critical set can be covered by Pareto critical sets of subproblems of size $n+1$. For the case $k = n+1$, we showed that the topological boundary of the Pareto critical set can be covered by Pareto critical sets of subproblems of size n. Unfortunately, we have basically no results for $k < n+1$. For this case, future work may require new tools or a new approach. For Lemma 5.5.7, it might be possible to sharpen the bound $|I| \leq n+1$ by using a sharper version of Carathéodory's theorem (Theorem 3.1 in [Gal11]) which is based on *affine geometry*. In particular, the *affine dimension* of $\partial^{\cup} f(x)$ (cf. (2.16)) might be a natural generalization of the rank of the Jacobian of f in this context. Besides that, a definition for the boundary of the Pareto critical set in the nonsmooth case has to be found, since the topological boundary is not useful when $k < n+1$. Finally, it would be interesting to see if better results can be obtained for the set of points satisfying optimality conditions based on the generalized Jacobian from Remark 2.2.26.

In addition to the points mentioned so far, a promising direction for future research is the exploitation of the hierarchical structure in numerical methods. So far, this has been done in [Pei17] for the ε-constraint method and in [BV19] for the reference point method. This is especially promising for problems where $k \geq n+1$ since in this case, the Pareto critical set is typically not a null set. Solving all subproblems of size n creates a partition of the variable space (as in Example 5.3.1) and to determine the complete Pareto critical set, one merely has to test one point from each set in the partition for Pareto criticality.

Inferring objective vectors from Pareto critical data

Chapter 6 was concerned with the inverse problem of multiobjective optimization, where a data set is given and we are searching for a (smooth) objective vector for which the data is Pareto critical. We began by discussing which data is required for this and concluded that in addition to Pareto critical points, the data should also contain the corresponding KKT vectors to obtain a "reasonably posed" inverse prob-

lem. By replacing the space of continuously differentiable functions with the span of a finite number of basis functions, we showed that the inverse problem reduces to a homogeneous linear system in the coefficients of the basis functions. This allowed us to solve it efficiently via a singular value decomposition, where the smallest singular value was a measure for how well the given data set can be approximated by the (extended) Pareto critical set of an objective vector in the span of the chosen basis functions. After stating the algorithm for solving the inverse problem and proving a theorem about the quality of the result, we considered three applications. The first application was the creation of test problem for MOP solvers. Here, by recalling the structural results about Pareto critical sets from Chapter 5, we were able to construct data sets by hand such that our inverse approach creates MOPs with certain prescribed properties. The second application was the identification of objective vectors for stochastic MOPs, where we showed that the robustness of our approach with respect to (stochastic) noise in the data can potentially be used to reconstruct the actual objective vector. Finally, as a third application, we used our inverse approach to create cheap surrogate models for (potentially expensive) objective vectors. Here, the idea was to compute just a few Pareto critical points (with corresponding KKT vectors) for the original problem and then use the result as data for our inverse approach. For future work, the open problems discussed in Section 6.3 should be addressed. Furthermore, since we only considered monomials of varying maximal degree as basis functions here, it would be interesting to consider different choices for basis functions. For example, it might be possible to incorporate additional information into the inverse approach by choosing basis functions with certain properties. Finally, since the inverse approach is based on the KKT condition for the unconstrained case, this is currently the only case it can handle. A generalization to constrained problems might be possible by considering the KKT conditions for the constrained case and additionally requiring the data to contain information about the constraints.

List of symbols

$\mathbb{R}$	Real numbers
$\mathbb{R}^{\geq 0}$	Non-negative real numbers ($\mathbb{R}^{>0}$, $\mathbb{R}^{\leq 0}$ are defined analogously)
$\mathbb{N}$	Natural numbers
n	Number of variables in our MOP
k	Number of objective functions in our MOP
X	Set of feasible points (mostly $X = \mathbb{R}^n$)
f	Objective vector $f : X \to \mathbb{R}^k$
im	Image of a function
$\|\cdot\|$	Euclidean norm
$\|\cdot\|_1$	ℓ_1-norm
∂A	Topological boundary of a set A
$\overline{A}$	Topological closure of a set A
A°	Topological interior of a set A
$\operatorname{conv}(A)$	Convex hull of a set A
∇f_i	Gradient of f_i
Df	Jacobian matrix of f
$\nabla^2 f_i$	Hessian matrix of f_i
$\langle \cdot, \cdot \rangle$	Standard inner product in $\mathbb{R}^n$
Δ_k	The standard simplex in $\mathbb{R}^k$, i.e., the set of vectors in $\mathbb{R}^k$ with non-negative entries that sum up to 1
P	Pareto set
P_c	Pareto critical set
$\operatorname{rk}(B)$	Rank of a matrix B
$\ker(B)$	Kernel of a matrix B
$T_z M$	Tangent space of a manifold M in $z \in M$
$\mathcal{M}$	Set of Pareto critical points together with their positive KKT vectors (cf. Theorem 2.2.10)
pr_x	Projection onto the first n components
$B_\varepsilon(x)$	Open ball $\{y \in \mathbb{R}^n : \|x - y\| < \varepsilon\}$ with radius ε around x
$\overline{B_\varepsilon}(x)$	Closed ball $\{y \in \mathbb{R}^n : \|x - y\| \leq \varepsilon\}$ with radius ε around x
Ω_i	Set of nondifferentiable points of f_i
Ω	Set of nondifferentiable points of f
$\partial f_i(x)$	Clarke subdifferential of f_i in x
$\partial f(x)$	Generalized Jacobian of f in x

$\partial^{\cup} f(x)$	Convex hull of the union of all Clarke subdifferentials in x, i.e., $\operatorname{conv}(\bigcup_{i=1}^{k} \partial f_i(x))$
$\operatorname{sgn}(a)$	Sign of $a \in \mathbb{R}$ with $\operatorname{sgn}(0) = 0$
$\mathcal{B}$	Collection of boxes covering $\mathbb{R}^n$
$\mathcal{B}_c$	Collection of boxes containing Pareto critical points
$\mathcal{A}_g$	Set $\{i \in \{1, \ldots, k_g\} : g_i(x) = 0\}$ of active indices
f^r	Inexact objective vector $f^r : \mathbb{R}^n \to \mathbb{R}^k$ (cf. (3.9))
P_c^r	Pareto critical set of f^r
P_ε^r	Tight superset of P_c and P_c^r (cf. (3.10))
$\partial_\varepsilon f_i(x)$	Goldstein ε-subdifferential of f_i in x
$\partial_\varepsilon^{\cup} f(x)$	Convex hull of the union of all ε-subdifferentials in x, i.e., $\operatorname{conv}(\bigcup_{i=1}^{k} \partial_\varepsilon f_i(x))$
$A(x)$	Set of KKT vectors of x
P_{int}	Pareto critical points with a strictly positive KKT vector
P_0	Complement of P_{int} in P_c
$\operatorname{Tan}(Y, x)$	Tangent cone of Y at x
$\partial_T P_c$	Boundary of the Pareto critical set defined via tangent cones (Definition 5.1.14)
P_c^I	Pareto critical set of the subproblem corresponding to $I \subseteq \{1, \ldots, k\}$
$\mathcal{P}(A)$	Power set of a set A
$P_{\mathcal{M}}$	Extended Pareto critical set, i.e., the set of Pareto critical points paired up with their KKT vectors
$\mathcal{D}$	Data set $\mathcal{D} = (\mathcal{D}_x, \mathcal{D}_\alpha) \subseteq \mathbb{R}^n \times \Delta_k$
$\mathcal{L}$	Matrix of the linear system in the inverse approach (cf. (6.5))
$\mathcal{F}$	Function that maps a coefficient vector onto the corresponding objective vector (cf. Definition 6.1.2)

Bibliography

[AB06] C. Aliprantis and K. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. 3rd ed. June 2006. DOI: 10.1007/3-540-29587-9.

[AGG15] H. Attouch, G. Garrigos, and X. Goudou. "A dynamic gradient approach to Pareto optimization with nonsmooth convex objective functions". In: *Journal of Mathematical Analysis and Applications* 422.1 (Feb. 2015), pp. 741–771. DOI: 10.1016/j.jmaa.2014.09.001.

[AO01] R. K. Ahuja and J. B. Orlin. "Inverse Optimization". In: *Operations Research* 49.5 (Oct. 2001), pp. 771–783. DOI: 10.1287/opre.49.5.771.10607.

[AO19] A. Asl and M. L. Overton. "Analysis of the gradient method with an Armijo–Wolfe line search on a class of non-smooth convex functions". In: *Optimization Methods and Software* 35.2 (Oct. 2019), pp. 223–242. DOI: 10.1080/10556788.2019.1673388.

[ASG01] A. C. Antoulas, D. C. Sorensen, and S. Gugercin. "A survey of model reduction methods for large-scale systems". In: *Contemporary Mathematics* 280 (2001), pp. 193–219. DOI: 10.1090/conm/280/04630.

[Ban+19] S. Banholzer, B. Gebken, M. Dellnitz, S. Peitz, and S. Volkwein. *ROM-based multiobjective optimization of elliptic PDEs via numerical continuation*. 2019. arXiv: 1906.09075. Accepted for publication.

[BBV17] S. Banholzer, D. Beermann, and S. Volkwein. "POD-based error control for reduced-order bicriterial PDE-constrained optimization". In: *Annual Reviews in Control* 44 (2017), pp. 226–237. DOI: 10.1016/j.arcontrol.2017.09.004.

[BCS20] F. Beltrán, O. Cuate, and O. Schütze. "The Pareto Tracer for General Inequality Constrained Multi-Objective Optimization Problems". In: *Mathematical and Computational Applications* 25.4 (Dec. 2020), p. 80. DOI: 10.3390/mca25040080.

[BGP21] Katharina Bieker, Bennet Gebken, and Sebastian Peitz. "On the Treatment of Optimization Problems with L1 Penalty Terms via Multiobjective Continuation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021). DOI: 10.1109/tpami.2021.3114962.

[BGW15] P. Benner, S. Gugercin, and K. Willcox. "A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems". In: *SIAM Review* 57.4 (Jan. 2015), pp. 483–531. DOI: 10.1137/130932715.

[BKM14] A. Bagirov, N. Karmitsa, and M. M. Mäkelä. *Introduction to Nonsmooth Optimization*. Springer International Publishing, 2014. DOI: 10.1007/978-3-319-08114-4.

[BLO05] J. V. Burke, A. S. Lewis, and M. L. Overton. "A Robust Gradient Sampling Algorithm for Nonsmooth, Nonconvex Optimization". In: *SIAM Journal on Optimization* 15.3 (Jan. 2005), pp. 751–779. DOI: 10.1137/030601296.

[BPK16] S. L. Brunton, J. L. Proctor, and J. N. Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems". In: *Proceedings of the National Academy of Sciences* 113.15 (Mar. 2016), pp. 3932–3937. DOI: 10.1073/pnas.1517384113.

[Bra+08] J. Branke, K. Deb, K. Miettinen, and R. Słowiński, eds. *Multiobjective Optimization. Interactive and Evolutionary Approaches*. Springer Berlin Heidelberg, 2008. DOI: 10.1007/978-3-540-88908-3.

[Bur+20] J. V. Burke, F. E. Curtis, A. S. Lewis, M. L. Overton, and L. E. A. Simões. "Gradient Sampling Methods for Nonsmooth Optimization". In: *Numerical Nonsmooth Optimization*. Springer International Publishing, 2020, pp. 201–225. DOI: 10.1007/978-3-030-34910-3_6.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Mar. 2004. DOI: 10.1017/cbo9780511804441.

[BV10] J. M. Borwein and J. D. Vanderwerff. *Convex Functions: Constructions, Characterizations and Counterexamples*. Cambridge University Press, 2010. DOI: 10.1017/cbo9781139087322.

[BV19] S. Banholzer and S. Volkwein. *Hierarchical convex multiobjective optimization by the Euclidean reference point method*. 2019. Preprint SPP1962-11.

[BW02] J. Brimberg and G. Wesolowsky. "Minisum Location with Closest Euclidean Distances". In: *Annals of Operations Research* 111 (Mar. 2002), pp. 151–165. DOI: 10.1023/A:1020901719463.

[CG59] W. Cheney and A. A. Goldstein. "Proximity maps for convex sets". In: *Proceedings of the American Mathematical Society* 10.3 (Mar. 1959), pp. 448–448. DOI: 10.1090/s0002-9939-1959-0105008-8.

[Cha04] A. Chambolle. "An algorithm for total variation minimization and applications". In: *Journal of Mathematical Imaging and Vision* 20 (Jan. 2004), pp. 89–97. DOI: 10.1023/B:JMIV.0000011325.36760.1e.

[Cha+14] T. C. Y. Chan, T. Craig, T. Lee, and M. B. Sharpe. "Generalized Inverse Multiobjective Optimization with Application to Cancer Therapy". In: *Operations Research* 62.3 (June 2014), pp. 680–695. DOI: 10.1287/opre.2014.1267.

[CL18] T. C. Y. Chan and T. Lee. "Trade-off preservation in inverse multiobjective convex optimization". In: *European Journal of Operational Research* 270.1 (Oct. 2018), pp. 25–39. DOI: 10.1016/j.ejor.2018.02.045.

[Cla90] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, Jan. 1990. DOI: 10.1137/1.9781611971309.

[CQ12] F. E. Curtis and X. Que. "An adaptive gradient sampling algorithm for non-smooth optimization". In: *Optimization Methods and Software* 28.6 (Sept. 2012), pp. 1302–1324. DOI: 10.1080/10556788.2012.714781.

[Cru13] J. Y. Bello Cruz. "A Subgradient Method for Vector Optimization Problems". In: *SIAM Journal on Optimization* 23.4 (Jan. 2013), pp. 2169–2182. DOI: 10.1137/120866415.

[Dax13] A. Dax. "From Eigenvalues to Singular Values: A Review". In: *Advances in Pure Mathematics* 03.09 (2013), pp. 8–24. DOI: 10.4236/apm.2013.39a2002.

[Deb01] K. Deb. *Multi-objective Optimization using Evolutionary Algorithms*. John Wiley & Sons, 2001.

[Deb+02] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. "A fast and elitist multiobjective genetic algorithm: NSGA-II". In: *IEEE Transactions on Evolutionary Computation* 6.2 (Apr. 2002), pp. 182–197. DOI: 10.1109/4235.996017.

[Deb99] K. Deb. "Multi-objective Genetic Algorithms: Problem Difficulties and Construction of Test Problems". In: *Evolutionary Computation* 7.3 (Sept. 1999), pp. 205–230. DOI: 10.1162/evco.1999.7.3.205.

[Del+09] M. Dellnitz, S. Ober-Blöbaum, M. Post, O. Schütze, and B. Thiere. "A multi-objective approach to the design of low thrust space trajectories using optimal control". In: *Celestial Mechanics and Dynamical Astronomy* 105.1-3 (Sept. 2009), pp. 33–59. DOI: 10.1007/s10569-009-9229-y.

[DGK63] L. Danzer, B. Grünbaum, and V. Klee. "Helly's Theorem and Its Relatives". In: Proceedings of Symposia in Pure Mathematics: Convexity (1963).

[DH97] M. Dellnitz and A. Hohmann. "A subdivision algorithm for the computation of unstable manifolds and global attractors". In: *Numerische Mathematik* 75.3 (Jan. 1997), pp. 293–317. DOI: 10.1007/s002110050240.

[DSH05] M. Dellnitz, O. Schütze, and T. Hestermeyer. "Covering Pareto Sets by Multilevel Subdivision Techniques". In: *Journal of Optimization Theory and Applications* 124.1 (Jan. 2005), pp. 113–136. DOI: 10.1007/s10957-004-6468-7.

[DZ18] C. Dong and B. Zeng. *Inferring Parameters Through Inverse Multiobjective Optimization*. 2018. arXiv: 1808.00935.

[EG15] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions, Revised Edition*. Chapman and Hall/CRC, Apr. 2015. DOI: 10.1201/b18333.

[Ehr05] M. Ehrgott. *Multicriteria Optimization*. Springer-Verlag, 2005. DOI: 10.1007/3-540-27659-9.

[ES00] M. Ester and J. Sander. *Knowledge Discovery in Databases*. Springer Berlin Heidelberg, 2000. DOI: 10.1007/978-3-642-58331-5.

[FDS09] J. Fliege, L. M. Graña Drummond, and B. F. Svaiter. "Newton's Method for Multiobjective Optimization". In: *SIAM Journal on Optimization* 20.2 (Jan. 2009), pp. 602–626. DOI: 10.1137/08071692x.

[FF93] C. M. Fonseca and P. J. Fleming. "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization". In: *Proceedings of the Fifth International Conference on Genetic Algorithms*. 1993, pp. 416–426.

[FS00] J. Fliege and B. F. Svaiter. "Steepest descent methods for multicriteria optimization". In: *Mathematical Methods of Operations Research (ZOR)* 51.3 (Aug. 2000), pp. 479–494. DOI: 10.1007/s001860000043.

[FX11] J. Fliege and H. Xu. "Stochastic Multiobjective Optimization: Sample Average Approximation and Applications". In: *Journal of Optimization Theory and Applications* 151.1 (May 2011), pp. 135–162. DOI: 10.1007/s10957-011-9859-6.

[Gal11] J. Gallier. *Geometric Methods and Applications*. Springer New York, 2011. DOI: 10.1007/978-1-4419-9961-0.

[GG92] G. Giorgi and A. Guerraggio. "On the notion of tangent cone in mathematical programming". In: *Optimization* 25.1 (Jan. 1992), pp. 11–23. DOI: 10.1080/02331939208843804.

[Gol77] A. A. Goldstein. "Optimization of lipschitz continuous functions". In: *Mathematical Programming* 13.1 (Dec. 1977), pp. 14–22. DOI: 10.1007/bf01584320.

[GP13] W. J. Gutjahr and A. Pichler. "Stochastic multi-objective optimization: a survey on non-scalarizing methods". In: *Annals of Operations Research* 236.2 (Apr. 2013), pp. 475–499. DOI: 10.1007/s10479-013-1369-5.

[GP21a] B. Gebken and S. Peitz. "An Efficient Descent Method for Locally Lipschitz Multiobjective Optimization Problems". In: *Journal of Optimization Theory and Applications* 80 (Mar. 2021), pp. 3–29. DOI: 10.1007/s10957-020-01803-w.

[GP21b] B. Gebken and S. Peitz. "Inverse multiobjective optimization: Inferring decision criteria from data". In: *Journal of Global Optimization* 80 (May 2021), pp. 696–723. DOI: 10.1007/s10898-020-00983-z.

[GPD19] B. Gebken, S. Peitz, and M. Dellnitz. "On the hierarchical structure of Pareto critical sets". In: *Journal of Global Optimization* 73.4 (Jan. 2019), pp. 891–913. DOI: 10.1007/s10898-019-00737-6.

[Gül10] O. Güler. *Foundations of Optimization*. Springer New York, 2010. DOI: 10.1007/978-0-387-68407-9.

[Gup+97] P. Gupta, R. Janardan, M. Smid, and B. Dasgupta. "The Rectangle Enclosure and Point-Dominance Problems Revisited". In: *International Journal of Computational Geometry & Applications* 07.05 (Oct. 1997), pp. 437–455. DOI: 10.1142/s0218195997000260.

[Gut+16] C. Gutiérrez, B. Jiménez, V. Novo, and G. Ruiz-Garzón. "Vector critical points and efficiency in vector optimization with Lipschitz functions". In: *Optimization Letters* 10.1 (Jan. 2016), pp. 47–62. DOI: 10.1007/s11590-015-0850-2.

[Har16] G. H. Hardy. "Weierstrass's Non-Differentiable Function". In: *Transactions of the American Mathematical Society* 17.3 (July 1916), p. 301. DOI: 10.2307/1989005.

[Hil01] C. Hillermeier. *Nonlinear Multiobjective Optimization*. Birkhäuser Basel, 2001. DOI: 10.1007/978-3-0348-8280-4.

[HJ12] R. A. Horn and C. R. Johnson. *Matrix Analysis*. 2nd edn. USA: Cambridge University Press, 2012. ISBN: 0521548233.

[HL14] M. E. Hartikainen and A. Lovison. "PAINT–SiCon: constructing consistent parametric representations of Pareto sets in nonconvex multiobjective optimization". In: *Journal of Global Optimization* 62.2 (Aug. 2014), pp. 243–261. DOI: 10.1007/s10898-014-0232-9.

[HL99] S. Huang and Z. Liu. "On the inverse problem of linear programming and its application to minimum weight perfect k-matching". In: *European Journal of Operational Research* 112.2 (Jan. 1999), pp. 421–426. DOI: 10.1016/s0377-2217(97)00444-x.

[HSS16] E. S. Helou, S. A. Santos, and L. E. A. Simões. "On the differentiability check in gradient sampling methods". In: *Optimization Methods and Software* 31.5 (May 2016), pp. 983–1007. DOI: 10.1080/10556788.2016.1178262.

[IUV17] L. Iapichino, S. Ulbrich, and S. Volkwein. "Multiobjective PDE-constrained optimization using the reduced-basis method". In: *Advances in Computational Mathematics* 43.5 (Jan. 2017), pp. 945–972. DOI: 10.1007/s10444-016-9512-x.

[Jah11] J. Jahn. *Vector Optimization*. Springer, 2011. DOI: 10.1007/978-3-642-17005-8.

[Ker+16] P. Kerschke, H. Wang, M. Preuss, C. Grimme, A. Deutz, H. Trautmann, and M. Emmerich. "Towards Analyzing Multimodality of Continuous Multiobjective Landscapes". In: *Parallel Problem Solving from Nature – PPSN XIV*. Springer International Publishing, 2016, pp. 962–972. DOI: 10.1007/978-3-319-45823-6_90.

[Kiw10] K. C. Kiwiel. "A Nonderivative Version of the Gradient Sampling Algorithm for Nonsmooth Nonconvex Optimization". In: *SIAM Journal on Optimization* 20.4 (Jan. 2010), pp. 1983–1994. DOI: 10.1137/090748408.

[Kiw85] K. C. Kiwiel. *Methods of Descent for Nondifferentiable Optimization*. Springer Berlin Heidelberg, 1985. DOI: 10.1007/bfb0074500.

[Kiw90] K. C. Kiwiel. "Proximity control in bundle methods for convex nondifferentiable minimization". In: *Mathematical Programming* 46.1-3 (Jan. 1990), pp. 105–122. DOI: 10.1007/bf01585731.

[KJ06] K. Klamroth and T. Jørgen. "Constrained optimization using multiple objective programming". In: *Journal of Global Optimization* 37.3 (July 2006), pp. 325–355. DOI: 10.1007/s10898-006-9052-x.

[Kön04] K. Königsberger. *Analysis 2*. Springer Berlin Heidelberg, 2004. DOI: 10.1007/3-540-35077-2.

[KP13] S. G. Krantz and H. R. Parks. *The Implicit Function Theorem*. Springer New York, 2013. DOI: 10.1007/978-1-4614-5981-1.

[KT51] H. W. Kuhn and A. W. Tucker. "Nonlinear Programming". In: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, Calif.: University of California Press, 1951, pp. 481–492.

[Kuh82] H. W. Kuhn. "Nonlinear Programming: A Historical Overview". In: *ACM SIGMAP Bulletin* 31 (June 1982), pp. 6–18. DOI: 10.1145/1111278.1111279.

[KV01] K. Kunisch and S. Volkwein. "Galerkin proper orthogonal decomposition methods for parabolic problems". In: *Numerische Mathematik* 90.1 (Nov. 2001), pp. 117–148. DOI: 10.1007/s002110100282.

[KV02] K. Kunisch and S. Volkwein. "Galerkin Proper Orthogonal Decomposition Methods for a General Equation in Fluid Dynamics". In: *SIAM Journal on Numerical Analysis* 40.2 (Jan. 2002), pp. 492–515. DOI: 10.1137/s0036142900382612.

[KWB11] A. Keshavarz, Y. Wang, and S. Boyd. "Imputing a convex objective function". In: *2011 IEEE International Symposium on Intelligent Control*. IEEE, Sept. 2011. DOI: 10.1109/isic.2011.6045410.

[Lee12] J. Lee. *Introduction to smooth manifolds*. 2nd ed. Springer, 2012. DOI: 10.1007/978-1-4419-9982-5.

[Lem89] C. Lemaréchal. "Chapter VII. Nondifferentiable optimization". In: *Handbooks in Operations Research and Management Science*. Elsevier, 1989, pp. 529–572. DOI: 10.1016/s0927-0507(89)01008-x.

[Lot+05] A. V. Lotov, G. K. Kamenev, V. E. Berezkin, and K. Miettinen. "Optimal control of cooling process in continuous casting of steel using a visualization-based multi-criteria approach". In: *Applied Mathematical Modelling* 29.7 (July 2005), pp. 653–672. DOI: 10.1016/j.apm.2004.10.009.

[Lov11] A. Lovison. "Singular Continuation: Generating Piecewise Linear Approximations to Pareto Sets via Global Analysis". In: *SIAM Journal on Optimization* 21.2 (Apr. 2011), pp. 463–490. DOI: 10.1137/100784746.

[Low+84] T. J. Lowe, J.-F. Thisse, J. E. Ward, and R. E. Wendell. "On Efficient Solutions to Multiple Objective Mathematical Programs". In: *Management Science* 30.11 (Nov. 1984), pp. 1346–1349. DOI: 10.1287/mnsc.30.11.1346.

[LP14] A. Lovison and F. Pecci. *Hierarchical stratification of Pareto sets*. 2014. arXiv: 1407.1755.

[LSB03] M. Lahanas, E. Schreibmann, and D. Baltas. "Multiobjective inverse planning for intensity modulated radiotherapy with constraint-free gradient-based optimization algorithms". In: *Physics in Medicine and Biology* 48.17 (Aug. 2003), pp. 2843–2871. DOI: 10.1088/0031-9155/48/17/308.

[Mäk03] M. M. Mäkelä. "Multiobjective proximal bundle method for nonconvex nonsmooth optimization: Fortran subroutine MPBNGC 2.0". In: *Reports of the Department of Mathematical Information Technology, Series B. Scientific Computing, B* 13 (2003), p. 2003.

[Mat12] J. Mather. "Notes on Topological Stability". In: *Bulletin of the American Mathematical Society* 49.4 (2012), pp. 475–506. DOI: 10.1090/s0273-0979-2012-01383-6.

[MAY12] N. Mahdavi-Amiri and R. Yousefpour. "An Effective Nonsmooth Optimization Algorithm for Locally Lipschitz Functions". In: *Journal of Optimization Theory and Applications* 155.1 (Apr. 2012), pp. 180–195. DOI: 10.1007/s10957-012-0024-7.

[MB94] C. Malivert and N. Boissard. "Structure of Efficient Sets for Strictly Quasi Convex Objectives". In: *Journal of Convex Analysis* 1.2 (1994), pp. 143–150.

[Meg83] N. Megiddo. "The Weighted Euclidean 1-Center Problem". In: *Mathematics of Operations Research* 8.4 (1983), pp. 498–504. DOI: 10.1287/moor.8.4.498.

[MEK14] M. M. Mäkelä, V.-P. Eronen, and N. Karmitsa. "On Nonsmooth Multiobjective Optimality Conditions with Generalized Convexities". In: *Optimization in Science and Engineering*. Springer New York, 2014, pp. 333–357. DOI: 10.1007/978-1-4939-0808-0_17.

[Mel76] W. de Melo. "On the Structure of the Pareto Set of Generic Mappings". In: *Boletim da Sociedade Brasileira de Matemática* 7.2 (Sept. 1976), pp. 121–126. DOI: 10.1007/bf02584786.

[MGGS09] D. Mueller-Gritschneder, H. Graeb, and U. Schlichtmann. "A Successive Approach to Compute the Bounded Pareto Front of Practical Multiobjective Optimization Problems". In: *SIAM Journal on Optimization* 20.2 (Jan. 2009), pp. 915–934. DOI: 10.1137/080729013.

[Mie98] K. Miettinen. *Nonlinear Multiobjective Optimization*. Springer US, 1998. DOI: 10.1007/978-1-4615-5563-6.

[MKM18] O. Montonen, N. Karmitsa, and M. M. Mäkelä. "Multiple subgradient descent bundle method for convex nonsmooth multiobjective optimization". In: *Optimization* 67.1 (2018), pp. 139–158. DOI: 10.1080/02331934.2017.1387259.

[MKW14] M. M. Mäkelä, N. Karmitsa, and O. Wilppu. "Multiobjective Proximal Bundle Method for Nonsmooth Optimization". In: *TUCS technical report No 1120, Turku Centre for Computer Science, Turku* (2014).

[MM93] K. Miettinen and M. M. Mäkelä. "An interactive method for nonsmooth multiobjective optimization with an application to optimal control". In: *Optimization Methods and Software* 2.1 (Jan. 1993), pp. 31–44. DOI: 10.1080/10556789308805533.

[MN92] M. M. Mäkelä and P. Neittaanmäki. *Nonsmooth Optimization*. World Scientific, May 1992. DOI: 10.1142/1493.

[MS17] A. Martín and O. Schütze. "Pareto Tracer: a predictor–corrector method for multi-objective optimization problems". In: *Engineering Optimization* 50.3 (June 2017), pp. 516–536. DOI: 10.1080/0305215x.2017.1327579.

[Naj+14] B. Najafi, A. Shirazi, M. Aminyavari, F. Rinaldi, and R. A. Taylor. "Exergetic, economic and environmental analyses and multi-objective optimization of an SOFC-gas turbine hybrid cycle coupled with an MSF desalination system". In: *Desalination* 334.1 (Feb. 2014), pp. 46–59. DOI: 10.1016/j.desal.2013.11.039.

[NE19] J. Niebling and G. Eichfelder. "A Branch–and–Bound-Based Algorithm for Nonconvex Multiobjective Optimization". In: *SIAM Journal on Optimization* 29.1 (Jan. 2019), pp. 794–821. DOI: 10.1137/18m1169680.

[Net+13] J. X. Da Cruz Neto, G. J. P. Da Silva, O. P. Ferreira, and J. O. Lopes. "A subgradient method for multiobjective optimization". In: *Computational Optimization and Applications* 54.3 (2013), pp. 461–472. DOI: 10.1007/s10589-012-9494-7.

[NW06] J. Nocedal and S. Wright. *Numerical Optimization*. Springer New York, 2006. DOI: 10.1007/978-0-387-40065-5.

[Oba+00] S. Obayashi, D. Sasaki, Y. Takeguchi, and N. Hirose. "Multiobjective Evolutionary Computation for Supersonic Wing-Shape Optimization". In: *IEEE Transactions on Evolutionary Computation* 4.2 (July 2000), pp. 182–187. DOI: 10.1109/4235.850658.

[OS14] W. de Oliveira and C. Sagastizábal. "Bundle methods in the XXIst century: A bird's-eye view". In: *Pesquisa Operacional* 34.3 (Dec. 2014), pp. 647–670. DOI: 10.1590/0101-7438.2014.034.03.0647.

[Par06] V. Pareto. *Manuale di Economia Politica*. Piccola biblioteca scientifica. Societa Editrice, 1906.

[PD17] S. Peitz and M. Dellnitz. "Gradient-Based Multiobjective Optimization with Uncertainties". In: *NEO 2016*. Springer International Publishing, Sept. 2017, pp. 159–182. DOI: 10.1007/978-3-319-64063-1_7.

[Pei17] S. Peitz. "Exploiting Structure in Multiobjective Optimization and Optimal Control". PhD thesis. Paderborn University, 2017.

[POBD18] S. Peitz, S. Ober-Blöbaum, and M. Dellnitz. "Multiobjective Optimal Control Methods for the Navier-Stokes Equations Using Reduced Order Modeling". In: *Acta Applicandae Mathematicae* 161.1 (Aug. 2018), pp. 171–199. DOI: 10.1007/s10440-018-0209-7.

[Pop05] N. Popovici. "Pareto reducible multicriteria optimization problems". In: *Optimization* 54.3 (June 2005), pp. 253–263. DOI: 10.1080/02331930500096213.

[PŽŽ17] P. M. Pardalos, A. Žilinskas, and J. Žilinskas. *Non-Convex Multi-Objective Optimization*. Springer International Publishing, 2017.

[QMN16] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced Basis Methods for Partial Differential Equations*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-15431-2.

[Que+05] N. V. Queipo, R. T. Haftka, W. Shyy, T. Goel, R. Vaidyanathan, and P. K. Tucker. "Surrogate-based analysis and optimization". In: *Progress in Aerospace Sciences* 41.1 (Jan. 2005), pp. 1–28. DOI: 10.1016/j.paerosci.2005.02.001.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[Rud91] W. Rudin. *Functional Analysis*. 2nd edn. McGraw-Hill, 1991. ISBN: 0070619883.

[SC16] D. G. Schaeffer and J. W. Cain. *Ordinary Differential Equations: Basics and Beyond*. Springer New York, 2016. DOI: 10.1007/978-1-4939-6389-8.

[Sch03] O. Schütze. "A New Data Structure for the Nondominance Problem in Multi-objective Optimization". In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2003, pp. 509–518. DOI: 10.1007/3-540-36970-8_36.

[Sch04] O. Schütze. "Set Oriented Methods for Global Optimization". PhD thesis. Paderborn University, 2004.

[Sch+09] O. Schütze, M. Vasile, O. Junge, M. Dellnitz, and D. Izzo. "Designing optimal low-thrust gravity-assist trajectories using space pruning and a multi-objective approach". In: *Engineering Optimization* 41.2 (Feb. 2009), pp. 155–181. DOI: 10.1080/03052150802391734.

[Sch12] S. Scholtes. *Introduction to Piecewise Differentiable Equations*. Springer New York, 2012. DOI: 10.1007/978-1-4614-4340-7.

[Sch16] B. Schröder. *Ordered Sets*. 2nd ed. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-29788-0.

[Sch+19] O. Schütze, O. Cuate, A. Martín, S. Peitz, and M. Dellnitz. "Pareto Explorer: a global/local exploration tool for many-objective optimization problems". In: *Engineering Optimization* 52.5 (May 2019), pp. 832–855. DOI: 10.1080/0305215x.2019.1617286.

[SDD05] O. Schütze, A. Dell'Aere, and M. Dellnitz. "On Continuation Methods for the Numerical Treatment of Multi-Objective Optimization Problems". In: *Practical Approaches to Multi-Objective Optimization*. Dagstuhl Seminar Proceedings 04461. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany, 2005.

[Sho85] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions*. Springer Berlin Heidelberg, 1985. DOI: 10.1007/978-3-642-82118-9.

[Sma73] S. Smale. "Global Analysis and Economics I: Pareto Optimum and a Generalization of Morse Theory". In: *Dynamical Systems*. Academic Press, 1973, pp. 531–544.

[SVR08] W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds. *Model Order Reduction: Theory, Research Aspects and Applications*. Springer Berlin Heidelberg, 2008. DOI: 10.1007/978-3-540-78841-6.

[SZ92] H. Schramm and J. Zowe. "A Version of the Bundle Idea for Minimizing a Nonsmooth Function: Conceptual Idea, Convergence Analysis, Numerical Results". In: *SIAM Journal on Optimization* 2.1 (Feb. 1992), pp. 121–152. DOI: 10.1137/0802008.

[Tib96] R. Tibshirani. "Regression Shrinkage and Selection via the Lasso". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 58.1 (Jan. 1996), pp. 267–288. DOI: 10.1111/j.2517-6161.1996.tb02080.x.

[Trö09] F. Tröltzsch. *Optimale Steuerung partieller Differentialgleichungen*. Vieweg+Teubner, 2009. DOI: 10.1007/978-3-8348-9357-4.

[Trö10] F. Tröltzsch. *Optimal Control of Partial Differential Equations*. American Mathematical Society, Apr. 2010. DOI: 10.1090/gsm/112.

[TŽ89] A. Törn and A. Žilinskas. *Global Optimization*. Springer Berlin Heidelberg, 1989. DOI: 10.1007/3-540-50871-6.

[Wal+05] R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban. "An interior algorithm for nonlinear optimization that combines line search and trust region steps". In: *Mathematical Programming* 107.3 (Nov. 2005), pp. 391–408. DOI: 10.1007/s10107-004-0560-5.

[Wit12] K. Witting. "Numerical algorithms for the treatment of parametric multiobjective optimization problems and applications". PhD thesis. Paderborn University, 2012.

[Zha+08] Q. Zhang, A. Zhou, S. Zhao, P. Suganthan, W. Liu, and S. Tiwari. "Multiobjective optimization Test Instances for the CEC 2009 Special Session and Competition". In: *Mechanical Engineering* (Jan. 2008).

[Zho+11] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, and Q. Zhang. "Multiobjective evolutionary algorithms: A survey of the state of the art". In: *Swarm and Evolutionary Computation* 1.1 (Mar. 2011), pp. 32–49. DOI: 10.1016/j.swevo.2011.03.001.

[ZLB04] E. Zitzler, M. Laumanns, and S. Bleuler. "A Tutorial on Evolutionary Multiobjective Optimization". In: *Lecture Notes in Economics and Mathematical Systems*. Springer Berlin Heidelberg, 2004, pp. 3–37. DOI: 10.1007/978-3-642-17144-4_1.

[ZLT01] E. Zitzler, M. Laumanns, and L. Thiele. "SPEA2: Improving the Strength Pareto Evolutionary Algorithm". In: *TIK-Report 103, ETH Zurich* (2001). DOI: 10.3929/ETHZ-A-004284029.