Deutsche Forschungsgemeinschaft
Priority Programme 1962

Hierarchical Convex Multiobjective Optimization by the Euclidean Reference Point Method

Stefan Banholzer, Stefan Volkwein

SPP 1962
Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization

Preprint Number SPP1962-117

received on August 15, 2019

Edited by
SPP1962 at Weierstrass Institute for Applied Analysis and Stochastics (WIAS)
Leibniz Institute in the Forschungsverbund Berlin e.V.
Mohrenstraße 39, 10117 Berlin, Germany
E-Mail: spp1962@wias-berlin.de

World Wide Web: `http://spp1962.wias-berlin.de/`

Hierarchical Convex Multiobjective Optimization by the Euclidean Reference Point Method

Stefan Banholzer, Stefan Volkwein

University of Konstanz, Department of Mathematics and Statistics, Universitätsstraße 10, D-78464 Konstanz, Germany

Stefan.Banholzer@uni-konstanz.de☆, Stefan.Volkwein@uni-konstanz.de

Abstract

In the present article convex multiobjective optimization problems with an arbitrary number of cost functions are considered. Since the weighted sum method has some deficiencies when it comes to approximating the Pareto front equidistantly, the Euclidean reference point method is investigated. However, for this method it is not clear how to choose reference points, i.e., the parameters in the scalarization function, guaranteeing a complete approximation of the Pareto front in the case of more than two cost functions. It is shown that by hierarchically solving subproblems of the original problem, it is possible to get a characterization of these reference points which is also numerically applicable independent of the number of cost functions. The resulting algorithm can thus be used for an arbitrary number of cost functions, which is shown in numerical tests for up to four cost functions.

Keywords: Convex multiobjective optimization, Pareto front, weighted sum method, Euclidean reference point method, hierarchical algorithm.

1. Introduction

Many optimization problems arising in applications can be formulated using several objective functions. This leads to the notion of multiobjective or multicriterial optimization problems. In most cases there is no single point which optimizes all functions simultaniously, i.e., the cost functions are conflicting with each other. Thus, the goal in multiobjective optimization is to compute the set of optimal compromises between the cost functions. Hereby, the definition of an optimal compromise depends on the chosen optimality concept, the most popular one being Pareto optimality or Pareto efficiency (c.f. [Par71]), which we will use in this article. In this case the set of optimal compromises is called the Pareto set and its image under all cost functions is called the Pareto front.

There are three popular approaches to tackle multiobjective optimization problems: Evolutionary algorithms (see e.g. [Deb01]), homotopy approaches as described in e.g. [Hil01], and last but not least scalarization methods; c.f. [Ehr05, Eic08, Mie99] amongst many others. The idea of scalarization methods is to transform the multiobjective optimization problem into a series of (parameter-dependent) scalar optimization problems, which can then be solved using well-known

☆S. Banholzer gratefully acknowledges support by the German DFG-Priority Program 1962 and by the Landesgraduiertenförderung of Baden-Württemberg.

techniques from scalar optimization. An approximation of the set of optimal compromises is computed by using different parameter values in the scalarization function.

Here we want to deal with the quite specific case of convex multiobjective optimization, which is still relevant for many applications, however. These kind of problems arise for example in the field of multiobjective optimal control of linear partial differential equations (see e.g. [BBV17, ITV15, IUV13]). For convex problems the weighted sum method (see e.g. [Ehr05, Mie99]), in which the weighted sum of all cost functions is minimized for different nonnegative weighting parameters, seems to be the natural choice. Firstly, it is very easy to understand and to implement, and secondly it can compute all Pareto optimal points (see e.g. [Ehr05, Theorem 3.5]). However, the method has some major drawbacks when it comes to computing an equidistant covering of the Pareto front, which is for example shown in [DD97]. Thus, in this article we investigate the so-called Euclidean reference point method (c.f. [Wie80, Wie86, Wie98]), in which the scalarization consists of minimizing the Euclidean distance between a given reference point and the image space of the cost functions. It has the advantage of a guaranteed fineness of the approximation of the Pareto front, which is lacking for the weighted sum method (c.f. Theorem 3.10).
One big issue of the Euclidean reference point method is the choice of reference points. Whereas it is a-priori clear in which range the weighting parameters for the weighted sum method have to be chosen, it is a-priori not known which reference points have to be chosen to

(i) guarantee solvability and Pareto optimality of the solution of the Euclidean reference point problem,

(ii) obtain a complete approximation of the Pareto front.

For the case of two cost functions this issue can be solved by simple geometrical arguments and iteratively defining new reference points using the information of the previously solved problems, see e.g. [Ban17, BBV16, BBV17, POBD15]. However, for more than two cost functions this iterative procedure cannot be used anymore, since it is not clear in which direction to move with the iteration and when to stop it.
To tackle this problem in the case of more than two cost functions the notion of Pareto admissible reference points is introduced in this article. By definition a Pareto admissible reference point guarantees the solvability and Pareto optimality of the solution of the Euclidean reference point problem. The connection of the Euclidean reference point method to the weighted sum method is used to obtain a characterization of Pareto admissible reference points. Using this it is then shown that the set of all Pareto admissible reference points is sufficient to obtain all Pareto optimal points.
However, its characterization relies on the solutions to all weighted sum problems and is therefore numerically not applicable. In a second step it is shown that Pareto admissible reference points restricted to a certain set are already completely characterized by the solutions to subproblems of the original problem, i.e., to multiobjective optimization problems, where one or several of the cost functions of the original problem are neglected. This gives rise to a hierarchical procedure of computing the Pareto front.

The idea of solving subproblems to build up the Pareto front hierarchically was already used in [MGGS09] for possibly nonconvex multiobjective optimization problems with a bounded and connected Pareto front using the Pascoletti-Serafini scalarization ([PS84]). Therefore, the concept of a 'boundary' of the Pareto front was introduced (which does not coincide with the topological

boundary) as the union of so-called trade-off limits. For these it was shown that they can be computed by solving subproblems neglecting single cost functions. The inner part of the Pareto front was then filled using the Pascoletti-Serafini scalarization, for which the reference points were chosen by using the 'boundary' information. In [KSd15] some weaknesses of this approach are shown and the notion of the boundary of the Pareto front is extended (though it still does not coincide with the topological boundary) to overcome these issues. The approach of filling the inner part of the Pareto front is the same as in [MGGS09]. There are more publications using this hierarchical method for computing the Pareto front, see e.g. [BKR17, DK19]. What is lacking in all of these papers is the proof that their method of filling the inner part of the Pareto front when the 'boundary' is given results in the complete Pareto front. In this article, we will prove that this is indeed true for our method, if all cost functions are convex.
Let us also mention that in [GPD19] it was shown that for unconstrained multiobjective optimization problems with twice continuously differentiable cost functions the solution of subproblems leads to a superset of the boundary of the so-called Pareto critical set. This was used to investigate the hierarchical structure of the Pareto critical set.

The structure of this article is as follows: We start by briefly introducing the basic concepts of multiobjective optimization in Section 2. In Section 3 we first study the weighted sum method before turning to the Euclidean reference point method. The notion of Pareto admissible reference points is introduced, for which a characterization is shown using the connection between the weighted sum method and the Euclidean reference point method. For a subset of all Pareto admissible reference point this characterization is investigated more deeply in Section 4. It is shown that this subset of Pareto admissible reference points is already characterized by solutions to subproblems of the original problem. This result is then used in Section 5 to develop a hierarchical method of computing reference points, which is translated into a numerical algorithm for computing the Pareto front. In Section 6 this algorithm is used to solve a multiobjective optimization problem involving four cost functions, and for an example with three cost functions the Euclidean reference point method and the weighted sum method are compared. Finally, we draw a conclusion and give an outlook on possible future work in Section 7.

2. Multiobjective Optimization

In this section we want to introduce convex multiobjective optimization problems and the optimality concept which we will use in the following.

Let $(U, \langle\cdot,\cdot\rangle_U)$ be a real Hilbert space, $U_{ad} \subset U$ non-empty, convex and closed, $k \geq 2$ arbitrary and denote by $J_1, \dots, J_k : U_{ad} \subset U \to \mathbb{R}$ the multiobjective cost functions. We write $J := (J_1, \dots, J_k)^T : U_{ad} \to \mathbb{R}^k$.
In the following we want to deal with the multiobjective optimization problem

$$\min_{u \in U_{ad}} J(u). \tag{MOP}$$

2.1 Definition. The set U_{ad} is called the **admissible set** and a vector $u \in U_{ad}$ is called **admissible**. The image set $J(U_{ad})$ is called the **objective space**. An element $y = J(u) \in J(U_{ad})$ is called **objective vector**.

For the cost functions we will mostly make use of the following assumptions. These assumptions assure, in particular, that all cost functions possess a global minimizer.

1 Assumption. *Let the cost functions $J_1, \dots, J_k : U_{ad} \subset U \to \mathbb{R}$ be convex, continuous and bounded from below. In the case that U_{ad} is unbounded, suppose that $\lim_{\|u\|_U \to \infty} J_i(u) = \infty$ for all $i \in \{1, \dots, k\}$.*

In contrast to a one-dimensional optimization problem it is a priori not clear how to define a solution of (MOP). In fact, there are several different concepts of solutions for a multiobjective optimization problem (see e.g. [Ehr05, Eic08] for very general optimality concepts). Here we will work with the notion of the so-called (weak) Pareto optimality, c.f. [Par71].

2.2 Notation. On $\mathbb{R}^k$ we use the partial ordering $\geq$ and $>$, respectively, which are given by

$$x \geq y \iff (\forall i \in \{1, \dots, k\} : x_i \geq y_i),$$
$$x > y \iff (\forall i \in \{1, \dots, k\} : x_i > y_i)$$

for all $x, y \in \mathbb{R}^k$. Analogously $\leq$ and $<$ are defined.
For convenience we write

$$x \gneqq y \iff (x \geq y \ \ \& \ \ x \neq y)$$

for all $x, y \in \mathbb{R}^k$ and define the set

$$\mathbb{R}^k_{\gneqq} := \{y \in \mathbb{R}^k \mid y \gneqq 0\}.$$

2.3 Definition. (i) An admissible vector $u \in U_{ad}$ and its corresponding objective vector $y := J(u) \in J(U_{ad})$ are called **weakly Pareto optimal** if there is no $\tilde{y} \in J(U_{ad})$ with $\tilde{y} < y$. The set

$$\mathcal{U}_{opt,\mathrm{w}} := \{u \in U_{ad} \mid u \text{ is weakly Pareto optimal}\} \subset U_{ad}$$

is called the **weak Pareto set**, and the set

$$\mathcal{J}_{opt,\mathrm{w}} := J(\mathcal{U}_{opt,w}) \subset \mathbb{R}^k$$

is called the **weak Pareto front**.

(ii) An admissible vector $u \in U_{ad}$ and its corresponding objective vector $y := J(u) \in J(U_{ad})$ are called **Pareto optimal** if there is no $\tilde{y} \in J(U_{ad})$ with $\tilde{y} \lneqq y$. The set

$$\mathcal{U}_{opt} := \{u \in U_{ad} \mid u \text{ is Pareto optimal}\} \subset U_{ad}$$

is called the **Pareto set**, and the set

$$\mathcal{J}_{opt} := J(\mathcal{U}_{opt}) \subset \mathbb{R}^k$$

is called the **Pareto front**.

(iii) We define the **ideal objective vector** $y^{id} \in \mathbb{R}^k$ by

$$y_i^{id} := \inf_{u \in U_{ad}} J_i(u) \quad \text{for } i \in \{1, \dots, k\}.$$

2.4 Lemma. *If all cost functions are strictly convex, it holds $\mathcal{U}_{opt} = \mathcal{U}_{opt,\mathrm{w}}$ and $\mathcal{J}_{opt} = \mathcal{J}_{opt,\mathrm{w}}$, i.e., the notion of weak Pareto optimality and Pareto optimality coincide.*

Proof. It is clear that $\mathcal{U}_{opt} \subset \mathcal{U}_{opt,\mathrm{w}}$ holds. The other inclusion is shown by contraposition. Let $u \in U_{ad}$ with $u \notin \mathcal{U}_{opt}$ be arbitrary. Then there is $\bar{u} \in U_{ad}$ with $J(\bar{u}) \lneqq J(u)$. Since U_{ad} is convex, the cost functions are strictly convex and $\bar{u} \neq u$, this implies that there is $\tilde{u} \in U_{ad}$ with $J(\tilde{u}) < J(u)$. Thus, it also holds $u \notin \mathcal{U}_{opt,\mathrm{w}}$, which completes the proof of $\mathcal{U}_{opt,\mathrm{w}} \subset \mathcal{U}_{opt}$. The equality $\mathcal{J}_{opt} = \mathcal{J}_{opt,\mathrm{w}}$ follows directly from $\mathcal{U}_{opt} = \mathcal{U}_{opt,\mathrm{w}}$. $\square$

3. Scalarization Techniques

The aim of multiobjective optimization is to compute the (weak) Pareto front or at least an approximation of it. One solution approach is the use of scalarization methods, in which the multiobjective optimization problem is transformed into a series of scalar optimization problems, see e.g. [Ehr05, Eic08, Mie99] for an (by far not complete) overview of scalarization methods. One possible way of scalarization consists of composing an arbitrary function $g : \mathbb{R}^k \to \mathbb{R}$ with the multiobjective function J. In this way, we obtain the scalar optimization problem

$$\min_{u \in U_{ad}} J_g(u) := (g \circ J)(u). \tag{SOP}$$

3.1. The Weighted Sum Method

One popular scalarization method is the weighted sum method (WSM) (see e.g. [Ehr05, Chapter 3], [Mie99, pp. 78–85]). The idea of the weighted sum method is to provide the different objective functions with nonnegative weights and then to minimize the sum of the weighted objective functions. Thus, for a weight $\alpha \in \mathbb{R}^k_{\gneq}$ we consider the parameter-dependent scalarization function

$$g_\alpha : \mathbb{R}^k \to \mathbb{R},\ x \mapsto \sum_{i=1}^k \alpha_i x_i.$$

This leads to the weighted sum problem

$$\min_{u \in U_{ad}} \sum_{i=1}^k \alpha_i J_i(u). \tag{WSP}$$

When studying a scalarization method for solving multiobjective optimization problems two questions are central:

(i) Under which assumptions and for which parameter values do we obtain (weakly) Pareto optimal points?

(ii) Can all (weakly) Pareto optimal points be computed by solving the scalarized problem for some parameter value?

The following two theorems answer these questions for the weighted sum method. These results are not new. In a similar form, they can be found in e.g. [Ehr05, Chapter 3].

3.1 Theorem. *Let Assumption 1 be satisfied.*

(i) For any weight $\alpha \in \mathbb{R}^k_{\gneq}$ the weighted sum problem (WSP) *has at least one global minimizer, and every local minimizer is already global. Every global minimizer is weakly Pareto optimal. If $\alpha > 0$ holds, it is even Pareto optimal.*

(ii) If there is $i \in \{1, \dots, k\}$ such that the cost function J_i is strictly convex and it holds $\alpha_i > 0$, then the global minimizer of (WSP) *is unique and Pareto optimal. Consequently, if all cost functions $J_1, \dots, J_k$ are strictly convex, then there is a unique global, Pareto optimal minimizer of* (WSP) *for all $\alpha \in \mathbb{R}^k_{\gneq}$.*

Proof. A proof of these statements can be found in [Bee19, Corollary 3.24]. □

3.2 Theorem. *Assume that the cost functions $J_1, \dots, J_k$ are convex. Let $u \in \mathcal{U}_{\text{opt,w}}$ be arbitrary. Then there is a weight $\alpha \in \mathbb{R}^k_{\gneq}$ such that u is a global minimizer of* (WSP) *with weight α.*

Proof. The proof can be found in [Bee19, Theorem 3.25]. For a finite-dimensional admissible set the statement was also shown in [Ehr05, Theorem 3.5]. The proof uses the convexity of the set $J(U_{ad}) + [0, \infty)^k$ and a separation theorem for convex sets. □

3.3 Remark. (i) Since all Pareto optimal points are in particular weakly Pareto optimal, the statement from Theorem 3.2 is trivially also true for all Pareto optimal points.

(ii) By scaling we can see that every weakly Pareto optimal point solves (WSP) to a weight

$$\alpha \in \Delta_k := \{\beta \in \mathbb{R}^k_{\gneq} \mid \sum_{i=1}^k \beta_i = 1\}.$$

Since the additional condition $\alpha > 0$ is also needed in some statements, we define

$$\Delta_k^> := \{\beta \in (0, \infty)^k \mid \sum_{i=1}^k \beta_i = 1\}.$$

Now we want to summarize the results about the weighted sum method by introducing (set-valued) solution mappings.

3.4 Definition. Let Assumption 1 be satisfied. Then we can define the (set-valued) solution mappings

$$\begin{aligned}
W_{opt,w} &: \Delta_k \rightrightarrows U_{ad} \subset U, && \alpha \mapsto \{u \in U_{ad} \mid u \text{ is a minimizer of (WSP) with weight } \alpha\},\\
W_{opt} &: \Delta_k \rightrightarrows U_{ad} \subset U, && \alpha \mapsto W_{opt,w}(\alpha) \cap \mathcal{U}_{opt},\\
F_{opt,w} &: \Delta_k \rightrightarrows \mathbb{R}^k, && \alpha \mapsto J(W_{opt,w}(\alpha)),\\
F_{opt} &: \Delta_k \rightrightarrows \mathbb{R}^k, && \alpha \mapsto J(W_{opt}(\alpha)).
\end{aligned}$$

For these mappings we can show the following properties.

3.5 Theorem. *Let Assumption 1 be satisfied. Then it holds*

$$W_{opt,w}(\Delta_k) = \mathcal{U}_{opt,w}, \quad W_{opt}(\Delta_k) = \mathcal{U}_{opt}, \quad W_{opt,w}\big|_{\Delta_k^>} = W_{opt}\big|_{\Delta_k^>},$$

and

$$F_{opt,w}(\Delta_k) = \mathcal{J}_{opt,w}, \quad F_{opt}(\Delta_k) = \mathcal{J}_{opt}, \quad F_{opt,w}\big|_{\Delta_k^>} = F_{opt}\big|_{\Delta_k^>}.$$

If all the cost functions $J_1, \dots, J_k$ are additionally strictly convex, the mappings $W_{opt,w}$, W_{opt}, $F_{opt,w}$ and F_{opt} are single-valued, and it holds $W_{opt,w} = W_{opt}$ and $F_{opt,w} = F_{opt}$.

Proof. All the statements are direct consequences of Theorems 3.1 and 3.2 and Definition 3.4. □

In Theorem 3.5 it was shown that the mappings $W_{opt,w}$, W_{opt}, $F_{opt,w}$ and F_{opt} are single-valued with $W_{opt,w} = W_{opt}$ and $F_{opt,w} = F_{opt}$, if the cost functions are strictly convex. In this case the mappings are treated as functions. Under even stronger assumptions we can then show the continuity of the functions W_{opt} and F_{opt}.

3.6 Theorem. *Let Assumption 1 be satisfied and assume additionally that the cost functions $J_1, \dots, J_k$ are strongly convex and twice continuously differentiable. Then the functions W_{opt} and F_{opt} are continuous.*

Proof. The continuity of W_{opt} under the given assumptions follows from [Ban17, Theorem 3.22]. The continuity of F_{opt} can then be concluded from $F_{opt} = J \circ W_{opt}$ and the continuity of W_{opt} and J. □

The problem of the weighted sum method is that a uniform distribution of the weights does not provide a uniform approximation of the Pareto front, even if all cost functions are convex. This issue has already been studied extensively in the literature, see e.g. [DD97].

3.2. The Euclidean Reference Point Method

As the weighted sum method has some disadvantages when it comes to the numerical approximation of the Pareto front, we investigate another scalarization method in the following, namely the so-called Euclidean reference point method (ERPM) (cf. [Wie80, Wie86, Wie98]).
For a given reference point $z \in \mathbb{R}^k$ we define the parameter-dependent scalarization function

$$g_z : \mathbb{R}^k \to \mathbb{R},\ x \mapsto \frac{1}{2}\|x - z\|_2^2 = \frac{1}{2}\sum_{i=1}^{k}(x_i - z_i)^2,$$

which measures the Euclidean distance of the point x to a given reference point z. The Euclidean reference point problem to the reference point z then reads

$$\min_{u \in U_{ad}} \frac{1}{2}\sum_{i=1}^{k}\left(J_i(u) - z_i\right)^2. \tag{ERPP}$$

Again the two main questions are:

(i) How does the reference point z have to be chosen such that solving (ERPP) has a global minimizer, which is (weakly) Pareto optimal?

(ii) Can all (weakly) Pareto optimal points be obtained by solving (ERPP) for a specific reference point z?

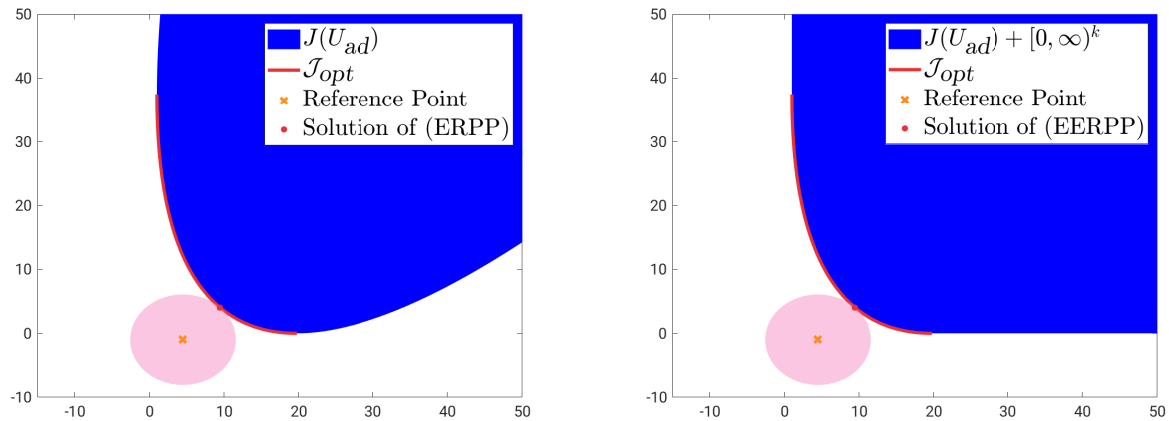
To investigate these questions, we first introduce a generalization of the (ERPP) as an auxiliary problem. For a given reference point $z \in \mathbb{R}^k$ we call

$$\min_{y \in J(U_{ad}) + [0,\infty)^k} \frac{1}{2}\|y - z\|_2^2 \tag{EERPP}$$

the extended Euclidean reference point problem (see also [Bee19, Definition 3.36]). For this, we can show the following statements.

3.7 Theorem. *Let Assumption 1 be satisfied. Then for all reference points $z \in \mathbb{R}^k$ the extended Euclidean reference point problem* (EERPP) *has a unique solution.*
Furthermore, a sufficient and necessary optimality condition for the unique solution $\bar{y}$ is given by

$$\langle \bar{y} - z, y - \bar{y} \rangle \geq 0 \quad \text{for all } y \in J(U_{ad}) + [0,\infty)^k. \tag{1}$$



(a) Solving (ERPP)

(b) Solving (EERPP)

Figure 1: The solutions for (ERPP) and (EERPP) coincide, so that $z \in \mathcal{Z}_{ad}$ holds.

Proof. Under the given assumptions one can show that the set $J(U_{ad}) + [0,\infty)^k$ is closed and convex, see [Ban17, Proof of Theorem 3.35]. By applying the Hilbert projection theorem ([Rud08, Theorem 4.10]) we get that (EERPP) has a unique solution $\bar{y} \in J(U_{ad}) + [0,\infty)^k$ for all reference points $z \in \mathbb{R}^k$.
The optimality condition (1) was shown in [Ban17, Definition and Theorem 2.14]. □

3.8 Remark. In the proof of the last theorem we used the proof of Theorem 3.35 in [Ban17]. Let us remark here that while the proof of the closure and convexity of the set $J(U_{ad}) + [0,\infty)^k$ is correct, the statement of Theorem 3.35 in [Ban17] only holds true for $k = 2$, i.e., for two cost functions. A counter example for three cost functions can be found in [Bee19, Figure 3.6].

3.9 Definition. Let Assumption 1 be satisfied. Then we can define the solution mapping

$$P : \mathbb{R}^k \to J(U_{ad}) + [0,\infty)^k,\ z \mapsto \bar{y},$$

which maps a reference point $z \in \mathbb{R}^k$ to the unique minimizer $\bar{y}$ of (EERPP).

Since the solution mapping P can be seen as a projection, we can show the following continuity result.

3.10 Theorem. *Let Assumption 1 be satisfied and $z^1, z^2 \in \mathbb{R}^k$ be two arbitrary reference points. Then it holds*

$$\left\|P(z^1) - P(z^2)\right\|_2 \leq \left\|z^1 - z^2\right\|_2,$$

i.e., the mapping P is Lipschitz-continuous with Lipschitz constant 1.

Proof. This was for example shown in [Ban17, Definition and Theorem 2.14] and follows from the optimality condition (1). □

3.11 Remark. The continuity result from Theorem 3.10 is the reason why the Euclidean reference point method performs better than the weighted sum method when it comes to approximating the Pareto front equidistantly. The advantage is that we can directly control the maximal distance between two Pareto optimal points by bounding the distance of the corresponding reference points. In this way it is possible to get a guaranteed approximation fineness of the Pareto front – in contrast to the weighted sum method.

3.12 Lemma. *Let Assumption 1 be satisfied and $z \in \mathbb{R}^k$ be arbitrary. Then it holds $P(z) \geq z$.*

Proof. Using the optimality condition (1) with $y = P(z) + d \in J(U_{\text{ad}}) + [0, \infty)^k$ for an arbitrary $d \in [0, \infty)^k$ yields $\langle P(z) - z, d \rangle \geq 0$. Hence, it must hold $P(z) \geq z$. □

In the following we want to investigate for which reference points the problems (ERPP) and (EERPP) are equivalent in the sense that the global minimizers of (ERPP) are given by the set $\{u \in U_{\text{ad}} \mid J(u) = P(z)\}$.

3.13 Definition. A reference point $z \in \mathbb{R}^k$ is called **admissible**, if the global minimizers of (ERPP) are given by the set $\{u \in U_{\text{ad}} \mid J(u) = P(z)\} \neq \emptyset$. The set of all admissible reference points is denoted by $\mathcal{Z}_{\text{ad}}$.

The following theorem gives a first (a-posteriori) characterization of admissible reference points.

3.14 Theorem. *Let Assumption 1 be satisfied and $z \in \mathbb{R}^k$ be arbitrary. Then*

$$z \in \mathcal{Z}_{\text{ad}} \iff P(z) \in J(U_{\text{ad}}).$$

Proof. '$\Rightarrow$' Assume that $\{u \in U_{\text{ad}} \mid J(u) = P(z)\} \neq \emptyset$ are the global minimizers of (ERPP). In particular, it holds $P(z) \in J(U_{\text{ad}})$.
'$\Leftarrow$' For the other implication suppose that $P(z) \in J(U_{\text{ad}})$ holds. Since $J(U_{\text{ad}}) \subset J(U_{\text{ad}}) + [0, \infty)^k$ it is clear that $\{u \in U_{\text{ad}} \mid J(u) = P(z)\} \neq \emptyset$ are global minimizers of (ERPP). To show that these are the only global minimizers, assume that there is a global minimizer $\bar{u} \in U_{\text{ad}}$ of (ERPP) with $J(\bar{u}) \neq P(z)$. But then we have $\|J(\bar{u}) - z\|_2^2 = \|P(z) - z\|_2^2$, which is a contradiction to the uniqueness of the minimizer of (EERPP). Hence, $\{u \in U_{\text{ad}} \mid J(u) = P(z)\}$ are the only global minimizers of (ERPP). □

Now we can show the following connection between the (extended) Euclidean reference point method and the weighted sum method, which was already stated in a similar way in [Wie80, Lemma 5].

3.15 Theorem. *Let Assumption 1 be satisfied.*

(i) If $z \in (J(U_{\text{ad}}) + [0, \infty)^k)^{\mathsf{c}}$ is a reference point with projection $P(z) = J(\bar{u}) + d \in J(U_{\text{ad}}) + [0, \infty)^k$ for some $\bar{u} \in U_{\text{ad}}$ and $d \in [0, \infty)^k$, then it holds $J(\bar{u}) \in F_{\text{opt,w}}(\alpha) \subset \mathcal{J}_{\text{opt,w}}$ with $\alpha := \frac{P(z)-z}{\|P(z)-z\|_1} \in \Delta_k$ and $\langle d, \alpha \rangle = 0$. In particular, if $d \neq 0$, there is $i \in \{1, \ldots, k\}$ with $\alpha_i = 0$, i.e., it holds $\alpha \in \partial \Delta_k$.

(ii) If, on the other hand, $\bar{u} \in W_{\text{opt,w}}(\alpha)$ for a weight $\alpha \in \Delta_k$, then it holds $P(J(\bar{u}) - t\alpha) = J(\bar{u})$ for any $t \geq 0$.

Proof. (i) Let $z \in (J(U_{\text{ad}}) + [0, \infty)^k)^{\mathsf{c}}$ be arbitrary. In particular, we have $P(z) = J(\bar{u}) + d \neq z$ and hence, Lemma 3.12 implies $P(z) \gneqq z$. Define $\alpha := \frac{P(z)-z}{\|P(z)-z\|_1} \in \Delta_k$. Choosing $y = J(u) + d$ for an arbitrary $u \in U_{\text{ad}}$ in the optimality condition (1) for the projection we get that it holds

$$\begin{aligned} 0 &\leq \langle P(z) - z, J(u) + d - P(z) \rangle = \langle P(z) - z, J(u) + d - J(\bar{u}) - d \rangle \\ &= \|P(z) - z\|_1 \langle \alpha, J(u) - J(\bar{u}) \rangle. \end{aligned}$$

This implies $J(\bar{u}) \in F_{opt,w}(\alpha) \subset \mathcal{J}_{opt,w}$ and $\bar{u} \in W_{opt,w}(\alpha) \subset \mathcal{U}_{opt,w}$.
To prove that d and α are orthogonal to each other, we plug $y = J(\bar{u})$ into the optimality condition (1). This yields

$$0 \leq \langle P(z) - z, J(\bar{u}) - P(z)\rangle = \langle P(z) - z, J(\bar{u}) - J(\bar{u}) - d\rangle = \|P(z) - z\|_1 \langle \alpha, -d\rangle.$$

Since $\alpha, d \geq 0$ and $\|P(z) - z\|_1 > 0$ this implies $\langle \alpha, d\rangle = 0$.

(ii) Now let $\alpha \in \Delta_k$ be arbitrary and $\bar{u} \in W_{opt,w}(\alpha)$. For any $t \geq 0$ define $z := J(\bar{u}) - t\alpha$. We show $P(z) = J(\bar{u})$ by plugging it into the optimality condition (1): For any $y \in J(U_{ad}) + [0, \infty)^k$ it holds

$$\langle J(\bar{u}) - z, y - J(\bar{u})\rangle = \langle J(\bar{u}) - (J(\bar{u}) - t\alpha), y - J(\bar{u})\rangle = t\langle \alpha, y - J(\bar{u})\rangle \geq 0,$$

since $\bar{u} \in W_{opt,w}(\alpha)$ solves (WSP) to the weight α and $y \in J(U_{ad}) + [0, \infty)^k$. Hence, $P(z) = J(\bar{u})$.

□

As a direct consequence of the Theorems 3.14 and 3.15 we get the next corollary. It tells us a-priori which elements of the set $\left(J(U_{ad}) + [0, \infty)^k\right)^{\mathsf{c}}$ are admissible – if the solutions to all weighted sum problems are known.

3.16 Corollary. *Let Assumption 1 be satisfied and $z \in \mathbb{R}^k$ with $z \notin J(U_{ad}) + [0, \infty)^k$ be arbitrary. Then the following statements are equivalent*

(i) $z \in \mathcal{Z}_{ad}$,

(ii) $P(z) \in J(U_{ad})$,

(iii) $P(z) \in \mathcal{J}_{opt,w}$,

(iv) $P(z) \in F_{opt,w}(\alpha)$ *with* $\alpha = \frac{P(z)-z}{\|P(z)-z\|_1}$,

(v) $\exists \alpha \in \Delta_k : \exists u \in W_{opt,w}(\alpha) : \exists t > 0 : z = J(u) - t\alpha$.

In particular, the equality

$$\mathcal{Z}_{ad} \cap \left(J(U_{ad}) + [0, \infty)^k\right)^{\mathsf{c}} = \left\{z \in \mathbb{R}^k \mid \exists \alpha \in \Delta_k : \exists u \in W_{opt,w}(\alpha) : \exists t > 0 : z = J(u) - t\alpha\right\}$$

holds.

Proof. (i) $\iff$ (ii) was shown in Theorem 3.14. (ii) $\iff$ (iii) $\iff$ (iv) follows from Theorem 3.15 (i). (iv) $\Rightarrow$ (v) is proved by rearranging the equality $\alpha = \frac{P(z)-z}{\|P(z)-z\|_1}$. Finally, Theorem 3.15 (ii) can be used to show (v) $\Rightarrow$ (iv). □

Having defined and characterized admissible reference points, the next step is to investigate for which admissible reference points the minimizers of (ERPP) are (weakly) Pareto optimal. Therefore, the notion of (weakly) Pareto admissible reference points is introduced.

3.17 Definition. Let Assumption 1 be satisfied. A reference point $z \in \mathbb{R}^k$ is called **(weakly) Pareto admissible**, if $P(z) \in \mathcal{J}_{opt}$ ($P(z) \in \mathcal{J}_{opt,w}$) holds, i.e., if the set of global minimizers of (ERPP) fulfils

$$\{u \in U_{ad} \mid J(u) = P(z)\} \subset \mathcal{U}_{opt} \quad (\{u \in U_{ad} \mid J(u) = P(z)\} \subset \mathcal{U}_{opt,w}).$$

The set of all (weakly) Pareto admissible reference points is denoted by $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$, respectively.

3.18 Remark. Since it holds $\mathcal{J}_{opt} \subset \mathcal{J}_{opt,w} \subset J(U_{ad})$ we can conclude from Theorem 3.14 that $\mathcal{Z}_{opt} \subset \mathcal{Z}_{opt,w} \subset \mathcal{Z}_{ad}$ is satisfied.

A first characterization of the set of (weakly) Pareto admissible points is given by the next theorem.

3.19 Theorem. *Let Assumption 1 be satisfied. Then it holds*

$$\mathcal{Z}_{opt,w} = \left\{z \in \mathbb{R}^k \mid \exists \alpha \in \Delta_k : \exists u \in W_{opt,w}(\alpha) : \exists t \geq 0 : z = J(u) - t\alpha\right\}, \tag{2}$$

$$\mathcal{Z}_{opt} = \left\{z \in \mathbb{R}^k \mid \exists \alpha \in \Delta_k : \exists u \in W_{opt}(\alpha) : \exists t \geq 0 : z = J(u) - t\alpha\right\}. \tag{3}$$

Proof. We first show (2). Assume that there are $\alpha \in \Delta_k$, $u \in W_{opt,w}(\alpha)$ and $t \geq 0$ with $z = J(u) - t\alpha$. Then Theorem 3.15 implies that $P(z) = J(u) \in F_{opt,w}(\alpha) \subset \mathcal{J}_{opt,w}$ holds. Thus, $z \in \mathcal{Z}_{opt,w}$.
To show the other inclusion let $z \in \mathcal{Z}_{opt,w}$ be arbitrary, i.e., it holds $P(z) \in \mathcal{J}_{opt,w}$. If $z \in J(U_{ad}) + [0,\infty)^k$ we have $z = P(z) \in \mathcal{J}_{opt,w}$. Hence, there is $\alpha \in \Delta_k$ and $u \in W_{opt,w}(\alpha)$ with $z = P(z) = J(u) - 0 \cdot \alpha$. If $z \notin J(U_{ad}) + [0,\infty)^k$ we can directly apply (iii) $\Rightarrow$ (v) of Corollary 3.16 to get the statement. This concludes the proof of (2).
For the proof of (3) let first $\alpha \in \Delta_k$, $u \in W_{opt}(\alpha)$ and $t \geq 0$ with $z = J(u) - t\alpha$ be arbitrary. Again Theorem 3.15 yields that $P(z) = J(u) \in F_{opt}(\alpha) \subset \mathcal{J}_{opt}$ holds. Hence, $z \in \mathcal{Z}_{opt}$.
To prove the other inclusion let $z \in \mathcal{Z}_{opt}$ be arbitrary, i.e., it holds $P(z) \in \mathcal{J}_{opt}$. Again, $z \in J(U_{ad}) + [0,\infty)^k$ implies $z = P(z) \in \mathcal{J}_{opt}$, so that there is $\alpha \in \Delta_k$ and $u \in W_{opt}(\alpha)$ with $z = P(z) = J(u) - 0 \cdot \alpha$. If $z \notin J(U_{ad}) + [0,\infty)^k$ we can conclude from $P(z) \in J(U_{ad})$ and Theorem 3.15 (i) that $P(z) \in F_{opt,w}(\alpha)$ for $\alpha := \frac{P(z)-z}{\|P(z)-z\|_1}$. Since it even holds $P(z) \in \mathcal{J}_{opt}$ we have $P(z) \in F_{opt}(\alpha)$. This implies $z = P(z) - t\alpha$ for $t := \|P(z) - z\|_1$, which concludes the proof. □

Based on Theorem 3.15 we can also define the following function, which will become important later.

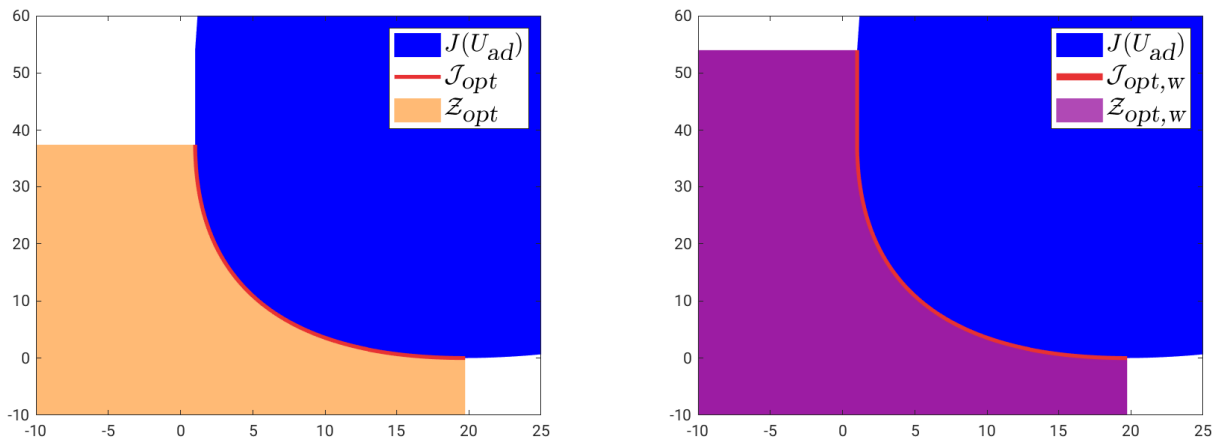
3.20 Definition. Let Assumption 1 be satisfied. Then we define the function H by

$$H : (J(U_{ad}) + [0,\infty)^k)^{\mathsf{c}} \to \Delta_k,\ z \mapsto \frac{P(z) - z}{\|P(z) - z\|_1}.$$

3.21 Lemma. *Let Assumption 1 be satisfied. Then the function H is continuous.*

Proof. For any $z \in (J(U_{ad}) + [0,\infty)^k)^{\mathsf{c}}$ it holds $P(z) \neq z$. Now the continuity of H follows from Theorem 3.10 as a concatenation of continuous functions. □

In the last theorem of this section it is shown that all (weakly) Pareto optimal points can be computed by solving (ERPP) for a (weakly) Pareto admissible reference point z.



(a) The set $\mathcal{Z}_{opt}$

(b) The set $\mathcal{Z}_{opt,w}$

Figure 2: Example for the sets $\mathcal{Z}_{opt} \neq \mathcal{Z}_{opt,w}$

3.22 Theorem. *Let Assumption 1 be satisfied. Then it holds*

$$P\left(\mathcal{Z}_{opt,w}\right) = \mathcal{J}_{opt,w}, \tag{4}$$

$$P\left(\mathcal{Z}_{opt}\right) = \mathcal{J}_{opt}. \tag{5}$$

Proof. For both statements (4) and (5) the inclusion '$\subset$' follows from the definition of $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$. To show $\mathcal{J}_{opt,w} \subset P\left(\mathcal{Z}_{opt,w}\right)$ let $y = J(u) \in \mathcal{J}_{opt,w}$ be arbitrary. By Theorem 3.5 there is a weight $\alpha \in \Delta_k$ such that $u \in W_{opt,w}(\alpha)$ holds. Define $z := y - t\alpha$ for an arbitrary $t \geq 0$. By (2) it holds $z \in \mathcal{Z}_{opt,w}$ and Theorem 3.15 (ii) implies $P(z) = y$, which concludes the proof. The inclusion $\mathcal{J}_{opt} \subset P\left(\mathcal{Z}_{opt}\right)$ is shown in the same way. $\square$

3.23 Definition. A set $Z \subset \mathbb{R}^k$ is called **(weakly) Pareto sufficient**, if $P(Z) = \mathcal{J}_{opt}$ ($P(Z) = \mathcal{J}_{opt,w}$) is satisfied.

3.24 Remark. By definition of the sets $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$, these are the biggest subsets of $\mathbb{R}^k$ fulfilling $P(\mathcal{Z}_{opt}) \subset \mathcal{J}_{opt}$ and $P(\mathcal{Z}_{opt,w}) \subset \mathcal{J}_{opt,w}$, respectively. Theorem 3.22 assures us that it also holds $\mathcal{J}_{opt} \subset P(\mathcal{Z}_{opt})$ and $\mathcal{J}_{opt,w} \subset P(\mathcal{Z}_{opt,w})$. In total, $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$ are Pareto sufficient and weakly Pareto sufficient, respectively, and there are no other reference points $z \notin \mathcal{Z}_{opt}$ with $P(z) \in \mathcal{J}_{opt}$ (or $z \notin \mathcal{Z}_{opt,w}$ with $P(z) \in \mathcal{J}_{opt,w}$). Moreover, for reference points in $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$ solving (ERPP) and (EERPP) is equivalent in the sense of Definition 3.13. Thus, the sets $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$ contain all relevant reference points.

However, thinking of a numerical implementation of the Euclidean reference point method, there are still two issues: First, given $\alpha \in \Delta_k$ and $y \in F_{opt}(\alpha)$ Theorem 3.15 tells us that $P(y - t\alpha) = y$ holds for all $t \geq 0$. Consequently, a set $Z \subset \mathcal{Z}_{opt}$ is already Pareto sufficient, if for each $\alpha \in \Delta_k$ and $y \in F_{opt}(\alpha)$ there is exactly one $t = t(y, \alpha) \geq 0$ such that $z = y - t(y, \alpha)\alpha \in Z$ (analogously for $Z \subset \mathcal{Z}_{opt,w}$). Hence, the sets $\mathcal{Z}_{opt}$ and $\mathcal{Z}_{opt,w}$ contain many unnecessary reference points, which would make a numerical algorithm inefficient. Second, the characterizations (2) and (3) given in Theorem 3.19 rely on the solutions of (WSP) to all weights in Δ_k, and are thus not applicable numerically.

So the goal is to find Pareto sufficient and weakly Pareto sufficient subsets $Z_{opt} \subset \mathcal{Z}_{opt}$ and $Z_{opt,w} \subset \mathcal{Z}_{opt,w}$, respectively, which can be characterized without using the solutions of (WSP) to all weights in Δ_k. This is the focus of the next section.

4. Characterization of Pareto Admissible Reference Points

So far the characterizations (2) and (3) of the sets $\mathcal{Z}_{opt,w}$ and $\mathcal{Z}_{opt}$ are not of great use, as was discussed in Remark 3.24. In the following we will show how solving (WSP) for all weights $\alpha \in \Delta_k$ can be avoided and instead only information of the solutions of (WSP) to $\alpha \in \partial\Delta_k = \{\beta \in \Delta_k \mid \exists i \in \{1,\dots,k\} : \beta_i = 0\}$ are used. Note that $\partial\Delta_k$ is the boundary of the set Δ_k in the subspace topology on $\{\beta \in \mathbb{R}^k \mid \sum_{i=1}^k \beta_i = 1\} \subset \mathbb{R}^k$.

4.1 Definition. For an arbitrary vector $\tilde{d} \in (0,\infty)^k$ define the **shifted ideal vector** $\tilde{y}^{id} := y^{id} - \tilde{d}$, where the ideal vector y^{id} was introduced in Definition 2.3. Let $D_i \subset \mathbb{R}^k$ be given by

$$D_i := \{y \in \mathbb{R}^k \mid y \geq \tilde{y}^{id},\ y_i = \tilde{y}_i^{id}\}$$

for all $i \in \{1,\dots,k\}$. Then the set $D \subset \mathbb{R}^k$ is defined by $D := \bigcup_{i=1}^k D_i$.

For the remaining part of this article let $\tilde{d} \in (0,\infty)^k$ be arbitrary but fixed, and $\tilde{y}^{id}$ and D be given as in Definition 4.1. In the following we want to investigate the sets $\mathcal{Z}_{opt} \cap D$ and $\mathcal{Z}_{opt,w} \cap D$. It turns out that these sets are (weakly) Pareto sufficient and that it is possible to characterize $\mathcal{Z}_{opt} \cap D$ by only looking at solutions of (WSP) for all weights $\alpha \in \partial\Delta_k$. For the set $\mathcal{Z}_{opt,w} \cap D$ we obtain a necessary, but not sufficient, condition for reference points to not be in $\mathcal{Z}_{opt,w} \cap D$.

4.2 Remark. The reason why we consider the shifted ideal vector $\tilde{y}^{id}$ instead of y^{id} is that we want to enforce $\big(J(U_{ad}) + [0,\infty)^k\big) \cap D = \emptyset$. Indeed, for an arbitrary $y \in D$ there is $i \in \{1,\dots,k\}$ with $y \in D_i$, i.e., $y_i = \tilde{y}_i^{id} = y_i^{id} - d_i < y_i^{id}$. This implies $y \notin J(U_{ad}) + [0,\infty)^k$.

4.3 Definition. Let Assumption 1 be satisfied. Given D as above we can define the (set-valued) mappings $G^D_{opt,w}$ and G^D_{opt} by

$$G^D_{opt,w} : \Delta_k \rightrightarrows D,\ \alpha \mapsto \left\{ y - t\alpha \mid y \in F_{opt,w}(\alpha),\ t = \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\alpha_i} \right\},$$
$$G^D_{opt} : \Delta_k \rightrightarrows D,\ \alpha \mapsto \left\{ y - t\alpha \mid y \in F_{opt}(\alpha),\ t = \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\alpha_i} \right\}.$$

4.4 Theorem. *Let Assumption 1 be satisfied. Then the mappings $G^D_{opt,w}$ and G^D_{opt} are well-defined.*

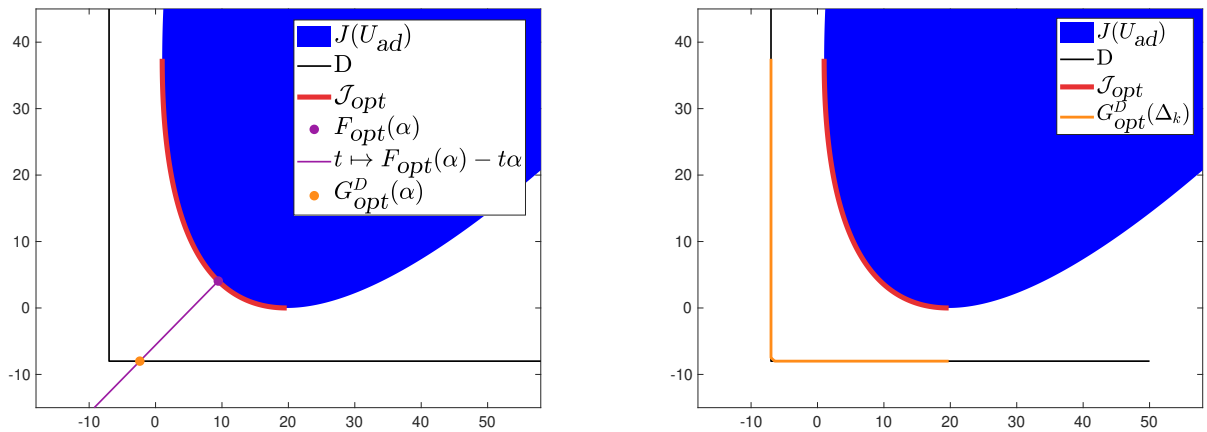
Proof. Let $\alpha \in \Delta_k$ be arbitrary. Because of $G^D_{opt}(\alpha) \subset G^D_{opt,w}(\alpha)$ we only need to show that $G^D_{opt,w}(\alpha) \subset D$ holds.
Since $\alpha \neq 0$ it is clear that $G^D_{opt,w}(\alpha) \subset \mathbb{R}^k$. For an arbitrary element $z = y - t\alpha \in G^D_{opt,w}(\alpha)$ choose $j \in \{1,\dots,k\}$ with

$$\frac{y_j - \tilde{y}_j^{id}}{\alpha_j} = \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\alpha_i} = t.$$

A straight-forward computation shows $z \in D_j$. Since $z \in G^D_{opt,w}(\alpha)$ was arbitrary this implies $G^D_{opt,w}(\alpha) \subset D$. Hence, the mappings $G^D_{opt,w}$ and G^D_{opt} are well-defined. □

At first we show some properties of the mappings $G^D_{opt,w}$ and G^D_{opt}, which will be needed in the algorithmic investigation in Section 5.

4.5 Lemma. *Let Assumption 1 be satisfied. For $\alpha, \beta \in \Delta_k$ with $\alpha \neq \beta$ it holds $G^D_{opt,w}(\alpha) \cap G^D_{opt,w}(\beta) = \emptyset$ and $G^D_{opt}(\alpha) \cap G^D_{opt}(\beta) = \emptyset$.*



(a) $G^D_{opt}(\alpha)$

(b) $G^D_{opt}(\Delta_k)$

Figure 3: Examples for the mapping G^D_{opt}

Proof. Let $\alpha, \beta \in \Delta_k$ with $\alpha \neq \beta$ be arbitrary. Since we have $G^D_{opt}(\alpha) \cap G^D_{opt}(\beta) \subset G^D_{opt,w}(\alpha) \cap G^D_{opt,w}(\beta)$, we only need to show $G^D_{opt,w}(\alpha) \cap G^D_{opt,w}(\beta) = \emptyset$. To this end assume that there is $z \in G^D_{opt,w}(\alpha) \cap G^D_{opt,w}(\beta)$, i.e., there are $y_\alpha \in F_{opt,w}(\alpha)$ and $y_\beta \in F_{opt,w}(\beta)$ such that

$$y_\alpha - t_\alpha \alpha = z = y_\beta - t_\beta \beta,$$

where $t_\alpha, t_\beta > 0$ are given as in the definition of $G^D_{opt,w}$. From Theorem 3.15 (ii) we can conclude $y_\alpha = P(z) = y_\beta$, so that $t_\alpha \alpha = t_\beta \beta$. Due to $\alpha, \beta \in \Delta_k$, this implies $\alpha = \beta$, which is a contradiction. Thus, it must holds $G^D_{opt,w}(\alpha) \cap G^D_{opt,w}(\beta) = \emptyset$. □

4.6 Corollary. *Let Assumption 1 be satisfied. Then it holds* $G^D_{opt,w}\big|_{\Delta_k^>} = G^D_{opt}\big|_{\Delta_k^>}$. *In particular, this implies* $G^D_{opt,w}(\Delta_k^>) = G^D_{opt}(\Delta_k^>)$.

Proof. This is a direct consequence of Theorem 3.5, where $F_{opt,w}\big|_{\Delta_k^>} = F_{opt}\big|_{\Delta_k^>}$ was shown. □

Another direct consequence of Lemma 4.5 and Corollary 4.6 is the following statement.

4.7 Corollary. *Let Assumption 1 be satisfied. Then it holds*

$$G^D_{opt}(\Delta_k) \setminus G^D_{opt}(\partial\Delta_k) = G^D_{opt}(\Delta_k^>) = G^D_{opt,w}(\Delta_k^>) = G^D_{opt,w}(\Delta_k) \setminus G^D_{opt,w}(\partial\Delta_k).$$

Proof. Lemma 4.5 implies the equalities $G^D_{opt}(\Delta_k) \setminus G^D_{opt}(\partial\Delta_k) = G^D_{opt}(\Delta_k^>)$ and $G^D_{opt,w}(\Delta_k^>) = G^D_{opt,w}(\Delta_k) \setminus G^D_{opt,w}(\partial\Delta_k)$, since $\Delta_k = \Delta_k^> \dot\cup \partial\Delta_k$. The equality $G^D_{opt}(\Delta_k^>) = G^D_{opt,w}(\Delta_k^>)$ was shown in Corollary 4.6. □

As explained in the beginning of this section we want to investigate the sets $\mathcal{Z}_{opt,w} \cap D$ and $\mathcal{Z}_{opt} \cap D$. As it turns out these sets are the images of the mappings $G^D_{opt,w}$ and G^D_{opt}.

4.8 Theorem. *Let Assumption 1 be satisfied. Then it holds* $G^D_{opt,w}(\Delta_k) = \mathcal{Z}_{opt,w} \cap D$ *and* $G^D_{opt}(\Delta_k) = \mathcal{Z}_{opt} \cap D$.

Proof. We will only show the equality $G^D_{opt,w}(\Delta_k) = \mathcal{Z}_{opt,w} \cap D$, since the proof of $G^D_{opt}(\Delta_k) = \mathcal{Z}_{opt} \cap D$ can be obtained similarly.
The inclusion $G^D_{opt,w}(\Delta_k) \subset \mathcal{Z}_{opt,w} \cap D$ is easy to see by the definition of $G^D_{opt,w}$ and Theorem 3.19. To prove $\mathcal{Z}_{opt,w} \cap D \subset G^D_{opt,w}(\Delta_k)$ let $z \in \mathcal{Z}_{opt,w} \cap D$ be arbitrary. By Theorem 3.19 there is $\alpha \in \Delta_k$, $y \in F_{opt,w}(\alpha)$ and $t \geq 0$ with $z = y - t\alpha$. Additionally, from $z \in D$ we can conclude that there is $i \in \{1, \dots, k\}$ with $z \in D_i$. By definition of D_i this implies $z_i = \tilde{y}^{id}_i$ and $z \geq \tilde{y}^{id}$. On the one hand this implies $t\alpha_i = y_i - z_i = y_i - \tilde{y}^{id}_i > 0$, so that $\alpha_i > 0$ and thus, $t = \frac{y_i - \tilde{y}^{id}_i}{\alpha_i}$. On the other hand using $z \geq \tilde{y}^{id}$ implies $y_j - t\alpha_j \geq \tilde{y}^{id}_j$ for all $j \in \{1, \dots, k\}$. For any $j \in \{1, \dots, k\}$ we get that if $\alpha_j = 0$ we have $\frac{y_j - \tilde{y}^{id}_j}{\alpha_j} = \infty$ and if $\alpha_j > 0$ we get $t \leq \frac{y_j - \tilde{y}^{id}_j}{\alpha_j}$. So in total this yields

$$t = \frac{y_i - \tilde{y}^{id}_i}{\alpha_i} = \min_{j \in \{1,\dots,k\}} \frac{y_j - \tilde{y}^{id}_j}{\alpha_j}.$$

Hence, $z = y - t\alpha \in G^D_{opt,w}(\alpha) \subset G^D_{opt,w}(\Delta_k)$. □

Using Theorem 3.15 (ii), Theorem 3.22 and Theorem 4.8 we can prove that the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$ are (weakly) Pareto sufficient.

4.9 Theorem. *Let Assumption 1 be satisfied. Then it holds*

$$P(G^D_{opt,w}(\Delta_k)) = \mathcal{J}_{opt,w}, \tag{6}$$

$$P(G^D_{opt}(\Delta_k)) = \mathcal{J}_{opt}, \tag{7}$$

i.e., the set $G^D_{opt,w}(\Delta_k)$ is weakly Pareto sufficient and the set $G^D_{opt}(\Delta_k)$ is Pareto sufficient.

Proof. The inclusion '$\subset$' follows in both cases from Theorem 4.8 and Theorem 3.22.
To show $\mathcal{J}_{opt,w} \subset P(G^D_{opt,w}(\Delta_k))$ let $y \in \mathcal{J}_{opt,w}$ be arbitrary with $y \in F_{opt,w}(\alpha)$ for an $\alpha \in \Delta_k$. Then Theorem 3.15 (ii) implies that $P(y - t\alpha) = y$ for all $t \geq 0$. In particular, this is true for $z = y - \min_{i \in \{1,\dots,k\}} \frac{y_i - \tilde{y}^{id}_i}{\alpha_i}\alpha \in G^D_{opt,w}(\alpha)$. Thus, $y = P(z) \in P(G^D_{opt,w}(\Delta_k))$, which was the claim.
The proof of $\mathcal{J}_{opt} \subset P(G^D_{opt}(\Delta_k))$ is done in the same way. □

In the following we investigate the structure of the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$ more detailed. From Theorem 4.8 we have $G^D_{opt,w}(\Delta_k) = \mathcal{Z}_{opt,w} \cap D$ and $G^D_{opt}(\Delta_k) = \mathcal{Z}_{opt} \cap D$, and Theorem 4.9 tells us that these sets are weakly Pareto sufficient and Pareto sufficient, respectively. However, there is still the same problem as before: We need to solve (WSP) for all weights $\alpha \in \Delta_k$ to obtain the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$. One idea to circumvent this issue is to try to characterize these sets by their boundaries $\partial G^D_{opt,w}(\Delta_k)$ and $\partial G^D_{opt}(\Delta_k)$, respectively.

4.10 Definition. Let Assumption 1 be satisfied. We define the functions $H^D_{opt,w}$ and H^D_{opt} by

$$H^D_{opt,w} := H\big|_{G^D_{opt,w}(\Delta_k)}, \qquad H^D_{opt} := H\big|_{G^D_{opt}(\Delta_k)},$$

where the function H was defined in Definition 3.20.

4.11 Corollary. *Let Assumption 1 be satisfied. Then the functions $H^D_{opt,w}$ and H^D_{opt} are continuous.*

Proof. This follows directly from Lemma 3.21, where we showed that H is continuous. □

We have the following connection between the mapping G^D_{opt} and the function H^D_{opt} (and $G^D_{opt,w}$ and $H^D_{opt,w}$, respectively).

4.12 Theorem. *Let Assumption 1 be satisfied. For all $\alpha \in \Delta_k$ it holds*

$$(H^D_{opt,w})^{-1}(\{\alpha\}) = G^D_{opt,w}(\alpha),$$
$$(H^D_{opt})^{-1}(\{\alpha\}) = G^D_{opt}(\alpha),$$

i.e., $G^D_{opt,w}$ is the preimage mapping of $H^D_{opt,w}$ and G^D_{opt} is the preimage mapping of H^D_{opt}.

Proof. Let $\alpha \in \Delta_k$ be arbitrary. We only prove the equality $(H^D_{opt})^{-1}(\{\alpha\}) = G^D_{opt}(\alpha)$. The equality $(H^D_{opt,w})^{-1}(\{\alpha\}) = G^D_{opt,w}(\alpha)$ follows with the same arguments.
If $z \in (H^D_{opt})^{-1}(\{\alpha\})$, it holds $z \in G^D_{opt}(\Delta_k)$, since $H^D_{opt} := H|_{G^D_{opt}(\Delta_k)}$. Therefore, there are $\beta \in \Delta_k$ and $y \in F_{opt}(\beta)$ such that $z = y - t\beta$ for $t := \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\beta_i}$. From $H^D_{opt}(z) = \alpha$ we get

$$\frac{P(z) - z}{\|P(z) - z\|_1} = H^D_{opt}(z) = \alpha.$$

Theorem 3.15 (ii) implies that $P(z) = y$, i.e.,

$$\alpha = \frac{P(z) - z}{\|P(z) - z\|_1} = \frac{y - y + t\beta}{\|y - y + t\beta\|_1} = \beta.$$

Hence, $z \in G^D_{opt}(\alpha)$.
If, on the other hand, $z \in G^D_{opt}(\alpha)$ holds, i.e., $z = y - t\alpha$ for a $y \in F_{opt}(\alpha)$ and $t := \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\beta_i}$, Theorem 3.15 (ii) implies that $P(z) = y \in F_{opt}(\alpha)$. Thus, $P(z) - z = y - y + t\alpha = t\alpha$, i.e.,

$$H^D_{opt}(z) = \frac{P(z) - z}{\|P(z) - z\|_1} = \alpha,$$

which concludes the proof. □

4.13 Corollary. *Let Assumption 1 be satisfied. If the cost functions $J_1, \dots, J_k$ are additionally strictly convex, the function $G^D_{opt} = G^D_{opt,w}$ is single-valued and it holds $G^D_{opt} = (H^D_{opt})^{-1}$.*

With these tools we can start to characterize the boundary of the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$, respectively. It turns out that the sets $\partial G^D_{opt,w}(\Delta_k)$ and $G^D_{opt,w}(\partial\Delta_k)$ as well as $\partial G^D_{opt}(\Delta_k)$ and $G^D_{opt}(\partial\Delta_k)$ are closely connected. Note that by $\partial G^D_{opt,w}(\Delta_k)$ and $\partial G^D_{opt}(\Delta_k)$ we mean the boundary of the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$ with respect to the subspace topology on D. The proof can be found in the Appendix.

4.14 Theorem. *Let Assumption 1 be satisfied. Then it holds*

$$\partial G^D_{opt,w}(\Delta_k) \cap G^D_{opt,w}(\Delta_k) \subset G^D_{opt,w}(\partial\Delta_k),$$
$$\partial G^D_{opt}(\Delta_k) \cap G^D_{opt}(\Delta_k) = G^D_{opt}(\partial\Delta_k).$$

If we have stronger assumptions on the cost functions, it is possible to show the equality $\partial G^D_{opt}(\Delta_k) = G^D_{opt}(\partial\Delta_k)$.

4.15 Theorem. *Let Assumption 1 be satisfied and the cost functions be additionally strongly convex and twice continuously differentiable. Then the mapping G^D_{opt} is a homeomorphism between Δ_k and $G^D_{opt}(\Delta_k)$.*

Proof. We already know that under these assumptions the function H^D_{opt} is continuous and it holds $G^D_{opt} = \left(H^D_{opt}\right)^{-1}$. The continuity of G^D_{opt} can be shown by using the continuity of F_{opt}, see Theorem 3.6. Thus, G^D_{opt} is a homeomorphism between Δ_k and $G^D_{opt}(\Delta_k)$. □

4.16 Corollary. *Let Assumption 1 be satisfied and the cost functions be additionally strongly convex and twice continuously differentiable. Then it holds* $\partial G^D_{opt}(\Delta_k) = G^D_{opt}(\partial\Delta_k)$.

Proof. It is well known that homeomorphisms map the boundary of a set onto the boundary of the image of that set. □

The following two theorems are the main results of this section. As it turns out the sets $G^D_{opt,w}(\partial\Delta_k)$ and $G^D_{opt}(\partial\Delta_k)$ contain enough information to characterize the sets $G^D_{opt,w}(\Delta_k)$ and $G^D_{opt}(\Delta_k)$, even without the strong assumptions of Corollary 4.16.
For the set $G^D_{opt,w}(\Delta_k)$ we obtain a necessary condition for points $z \in D$ to not lie in $G^D_{opt,w}(\Delta_k)$ and a sufficient condition for points $z \in D$ to not lie in $G^D_{opt,w}(\Delta_k^>)$.

4.17 Theorem. *Let Assumption 1 be satisfied. For all* $z \in D \setminus G^D_{opt,w}(\Delta_k)$ *it holds*

$$\exists\, i \in \{1,\dots,k\} : \exists\, \tilde{z} \in G^D_{opt,w}(\partial\Delta_k) \cap D_i : \exists\, \beta \in \{H^D_{opt,w}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta. \tag{8}$$

On the other hand if a point $z \in D$ *satisfies* (8) *then it holds* $z \notin G^D_{opt,w}(\Delta_k^>)$.

Proof. The proof can be found in the appendix. □

An analogue, yet stronger, statement can be shown for the set $G^D_{opt}(\Delta_k)$, for which we can find a necessary and sufficient characterization.

4.18 Theorem. *Let Assumption 1 be satisfied and* $z \in D$. *Then the following two statements are equivalent*

(i) $z \notin G^D_{opt}(\Delta_k)$,

(ii) $\exists\, i \in \{1,\dots,k\} : \exists\, \tilde{z} \in G^D_{opt}(\partial\Delta_k) \cap D_i : \exists\, \beta \in \{H^D_{opt}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta$.

Proof. The proof of this statement can be found in the appendix. □

The following two corollaries summarize the results about characterizations of (weakly) Pareto admissible reference points.

4.19 Corollary. *Let Assumption 1 be satisfied and* $z \in D$ *be arbitrary. Look at the following statements*

(i) $z \in \mathcal{Z}_{opt,w}$,

(ii) $z \in G^D_{opt,w}(\Delta_k)$,

(iii) $z \in G^D_{opt,w}(\Delta_k^>)$,

(iv) $\exists\, i \in \{1,\dots,k\} : \exists\, \tilde{z} \in G^D_{opt,w}(\partial\Delta_k) \cap D_i : \exists\, \beta \in \{H^D_{opt,w}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta$,

(v) $\exists\, i \in \{1,\dots,k\} : \exists\, \alpha \in \partial\Delta_k : \exists \tilde{z} \in G^D_{opt,w}(\alpha) \cap D_i : \exists\, \beta \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta$.

Then it holds

$$(i) \iff (ii), \qquad \neg(ii) \Rightarrow (iv), \qquad (iii) \Rightarrow (ii), \qquad (iv) \Rightarrow \neg(iii), \qquad (iv) \iff (v).$$

Proof. This is a direct consequence of Theorem 4.8 and Theorem 4.17. □

4.20 Corollary. *Let Assumption 1 be satisfied and $z \in D$ be arbitrary. Then the following statements are equivalent*

(i) $z \in \mathcal{Z}_{opt}$,

(ii) $z \in G^D_{opt}(\Delta_k)$,

(iii) $\neg\Big(\exists\, i \in \{1,\dots,k\} : \exists\, \tilde{z} \in G^D_{opt}(\partial\Delta_k) \cap D_i : \exists\, \beta \in \{H^D_{opt}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta\Big)$,

(iv) $\neg\Big(\exists\, i \in \{1,\dots,k\} : \exists\, \alpha \in \partial\Delta_k : \exists \tilde{z} \in G^D_{opt}(\alpha) \cap D_i : \exists\, \beta \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq} : z = \tilde{z} + \beta\Big)$.

Proof. This follows directly from Theorem 4.8 and Theorem 4.18. □

5. Hierarchical Computation of the Pareto Front

In this section we want to show how the theoretical results from the last section can be used to build up an algorithm for hierarchically computing the Pareto front.
Starting from the characterizations of the (weakly) Pareto sufficient sets $\mathcal{Z}_{opt} \cap D$ and $\mathcal{Z}_{opt,w} \cap D$ in Theorem 4.8, which used the solutions of (WSP) for all $\alpha \in \Delta_k$, we were able to show characterizations of $\mathcal{Z}_{opt} \cap D$ and $\mathcal{Z}_{opt,w} \cap D$ in Theorems 4.17 and 4.18, which only use the solutions of (WSP) for all $\alpha \in \partial\Delta_k$.
We can write the set $\partial\Delta_k$ in the following way

$$\partial\Delta_k = \bigcup_{i=1}^{k} \{\alpha \in \Delta_k \mid \alpha_i = 0\}.$$

The idea is that for a given $i \in \{1,\dots,k\}$ solving (WSP) for all $\alpha \in \Delta_k$ with $\alpha_i = 0$ corresponds to solving the multiobjective optimization problem $(J_j)_{j\in\{1,\dots,k\}\setminus\{i\}}$, which is a subproblem of the original problem. Hence, instead of solving the original problem for all $\alpha \in \partial\Delta_k$ we can also solve the subproblems $(J_j)_{j\in\{1,\dots,k\}\setminus\{i\}}$ for all $i \in \{1,\dots,k\}$. To construct the reference points for such a subproblem, we can again decompose it into all of its subproblems. This procedure can be repeated until the subproblems are of size one, i.e., the single cost functions $J_1,\dots,J_k$ have to be minimized. This gives rise to the following idea of hierarchically computing the Pareto front.

1. Compute the minimizers of the cost functions $J_1,\dots,J_k$
2. Use this information to compute $G^D_{opt}(\partial\Delta_2)$ ($G^D_{opt,w}(\partial\Delta_2)$) for all $\binom{k}{2}$ subproblems with two cost functions. Then generate the reference points for all of these subproblems using the characterization of Theorem 4.18 (Theorem 4.17) and solve all arising (ERPP).
3. Use the information of the $\binom{k}{2}$ subproblems with two cost functions to compute $G^D_{opt}(\partial\Delta_3)$ ($G^D_{opt,w}(\partial\Delta_3)$) for all $\binom{k}{3}$ subproblems with three cost functions. Then generate the reference points for all of these subproblems using the characterization of Theorem 4.18 (Theorem 4.17) and solve all arising (ERPP).

   ⋮

k. Use the information of the k subproblems with $k-1$ cost functions to compute $G^D_{opt}(\partial\Delta_k)$ ($G^D_{opt,w}(\partial\Delta_k)$). Then generate the reference points using the characterization of Theorem 4.18 (Theorem 4.17) and solve all arising (ERPP).

In what follows we will elaborate on this hierarchical procedure more precisely, translate it into a numerical algorithm and prove analytically that it yields the (weak) Pareto front.
To start with, we need to make some definitions concerning the solution of subproblems.

5.1 Definition. For $I \subset \{1,\dots,k\}$ we make the following definitions.

(i) For every vector $y \in \mathbb{R}^k$ we define the restriction to $\mathbb{R}^{|I|}$ by $y^I := (y_i)_{i\in I} \in \mathbb{R}^{|I|}$.

(ii) The sets $\mathcal{U}^I_{opt}, \mathcal{U}^I_{opt,w}$ denote the (weak) Pareto set and the sets $\mathcal{J}^I_{opt}, \mathcal{J}^I_{opt,w}$ denote the (weak) Pareto front for the multiobjective optimization problem $(J_i)_{i\in I}$.

(iii) We set $\Delta_I := \{\alpha \in \mathbb{R}^{|I|}_{\geq} \mid \sum_{i=1}^{|I|} \alpha_i = 1\}$, $\Delta^{>}_I := \{\alpha \in \Delta_I \mid \alpha > 0\}$. The boundary with respect to the subspace topology on $\{\beta \in \mathbb{R}^{|I|} \mid \sum_{i=1}^{|I|} \beta_i = 1\}$ is given by $\partial\Delta_I = \{\alpha \in \Delta_I \mid \exists i \in \{1,\dots,|I|\} : \alpha_i = 0\}$.

(iv) For $\alpha \in \Delta_I$ we define

$$\begin{aligned} W^I_{opt,w}(\alpha) &:= \{u \in U_{ad} \mid u \text{ minimizes (WSP) for the cost functions } (J_i)_{i\in I} \text{ with weight } \alpha\}, \\ W^I_{opt}(\alpha) &:= W^I_{opt,w}(\alpha) \cap \mathcal{U}^I_{opt}, \\ F^I_{opt,w}(\alpha) &:= (J_i(W^I_{opt,w}(\alpha)))_{i\in I}, \\ F^I_{opt}(\alpha) &:= (J_i(W_{opt}(\alpha)))_{i\in I}. \end{aligned}$$

(v) The mapping $P^I : \mathbb{R}^{|I|} \to (J_i(U_{ad}))_{i\in I} + [0,\infty)^{|I|}$ denotes the solution mapping of (EERPP) with respect to the cost functions $(J_i)_{i\in I}$.

(vi) For any $i \in I$ we set $D^I_i := \{y \in \mathbb{R}^{|I|} \mid y \geq (\tilde{y}^{id})^I,\ y_i = (\tilde{y}^{id})^I_i\}$ and $D^I := \cup_{i\in I} D^I_i$.

(vii) Given D^I as in (vi), the mappings $G^{D^I}_{opt}, G^{D^I}_{opt,w}, H^{D^I}_{opt}, H^{D^I}_{opt,w}$ are defined analogously to Section 4.

5.2 Remark. Let $I \subset \{1,\dots,k\}$ be arbitrary. Using the definitions from Definition 5.1 all the statements which were shown in Sections 3 and 4 for the original problem involving the cost functions $(J_1,\dots,J_k)$ can be directly transferred to the subproblem $(J_i)_{i\in I}$.

Moreover, we have the following result.

5.3 Lemma. *Let Assumption 1 be satisfied and let $I \subset \{1,\dots,k\}$ be arbitrary. Then it holds*

$$W^I_{opt,w}(\partial\Delta_I) = \bigcup_{K \subsetneq I} W^K_{opt}(\Delta_K).$$

Proof. '$\subset$' Let $\alpha \in \partial\Delta_I$ and $u \in W^I_{opt,w}(\alpha)$ be arbitrary. Define $K := \{i \in I \mid \alpha_i > 0\}$. Then it holds $K \subsetneq I$ and u solves (WSP) to the weight $\alpha^K > 0$ and the cost functions $(J_i)_{i\in K}$. Therefore, by Theorem 3.1 it holds $u \in W^K_{opt}(\alpha^K) \subset W^K_{opt}(\Delta_K)$.

'$\supset$' Let $K \subsetneq I$ and $u \in W^K_{opt}(\Delta_K)$ be arbitrary. Then there is $\alpha \in \Delta_K$ such that u solves (WSP) to the weight α and the cost functions $(J_i)_{i\in K}$. Defining $\tilde{\alpha} \in \mathbb{R}^{|I|}$ by

$$\tilde{\alpha} := \begin{cases} \alpha_i & i \in K, \\ 0 & i \in I \setminus K, \end{cases}$$

it holds $\tilde{\alpha} \in \partial\Delta_I$. Moreover, u also solves (WSP) to the weight $\tilde{\alpha}$ and the cost functions $(J_i)_{i\in I}$, i.e., it holds $u \in W^I_{opt,w}(\partial\Delta_I)$. □

Given $I \subset \{1,\dots,k\}$ and the solutions to all subproblems $(J_i)_{i\in K}$ for all $K \subset \{1,\dots,k\}$ with $|K| < |I|$, we only need to compute $W^I_{opt}(\Delta^>_I)$, since it holds

$$W^I_{opt,w}(\Delta_I) \setminus W^I_{opt,w}(\partial\Delta_I) = W^I_{opt}(\Delta_I) \setminus W^I_{opt}(\partial\Delta_I) \subset W^I_{opt}(\Delta^>_I) = W^I_{opt,w}(\Delta^>_I),$$

and $W^I_{opt,w}(\partial\Delta_I)$ – and thus also $W^I_{opt}(\partial\Delta_I)$ – were already obtained solving all the subproblems (see Lemma 5.3).

5.4 Lemma. *Let Assumption 1 be satisfied, $I \subset \{1,\dots,k\}$ and $\alpha \in \Delta_I$ be arbitrary. Then it holds*

$$P^I(G^{D^I}_{opt,w}(\alpha)) = F^I_{opt,w}(\alpha),$$
$$P^I(G^{D^I}_{opt}(\alpha)) = F^I_{opt}(\alpha).$$

In particular, this implies $P^I(G^{D^I}_{opt,w}(\Delta^>_I)) = P^I(G^{D^I}_{opt}(\Delta^>_I)) = F^I_{opt}(\Delta^>_I) = F^I_{opt,w}(\Delta^>_I)$.

Proof. Let $\alpha \in \Delta_I$ be arbitrary. We only show the equality $P^I(G^{D^I}_{opt,w}(\alpha)) = F^I_{opt,w}(\alpha)$, since the proof of $P^I(G^{D^I}_{opt}(\alpha)) = F^I_{opt}(\alpha)$ works with the same arguments.
'$\subset$' If $y \in P^I(G^{D^I}_{opt,w}(\alpha))$, there is $\tilde{y} \in F^I_{opt,w}(\alpha)$ such that $z := \tilde{y} - t\alpha \in G^{D^I}_{opt,w}(\alpha)$ with the usual $t > 0$ and $P^I(z) = y$. By Theorem 3.15 (ii) it also holds $P^I(z) = \tilde{y}$, i.e., $y = \tilde{y} \in F^I_{opt,w}(\alpha)$.
'$\supset$' Let $y \in F^I_{opt,w}(\alpha)$ be arbitrary. Define $z := y - t\alpha \in G^{D^I}_{opt,w}(\alpha)$ with the usual $t > 0$. Again by Theorem 3.15 (ii) it holds $P^I(z) = y \in F^I_{opt,w}(\alpha)$, which concludes the proof. □

Hence, by Lemma 5.4 reference points in

$$G^{D^I}_{opt,w}(\Delta^>_I) = G^{D^I}_{opt}(\Delta^>_I) = G^{D^I}_{opt}(\Delta_I) \setminus G^{D^I}_{opt}(\partial\Delta_I) = G^{D^I}_{opt,w}(\Delta_I) \setminus G^{D^I}_{opt,w}(\partial\Delta_I)$$

(see Corollary 4.7) are enough to compute $F^I_{opt}(\Delta^>_I)$ and therefore also $W^I_{opt}(\Delta^>_I)$.
Furthermore, modifying Corollary 4.20 yields the following characterization.

5.5 Corollary. *Let Assumption 1 be satisfied and let $I \subset \{1,\dots,k\}$ be arbitrary. Then the following two statements are equivalent*

(i) $z \in G^{D^I}_{opt}(\Delta^>_I) = G^{D^I}_{opt,w}(\Delta^>_I)$,

(ii) $\neg\Big(\exists i \in \{1,\dots,|I|\} : \exists \tilde{z} \in G^{D^I}_{opt}(\partial\Delta_I) \cap D^I_i : \exists \beta \in \{H^{D^I}_{opt}(\tilde{z}), e_i\}^\perp \cap [0,\infty)^{|I|} : z = \tilde{z} + \beta\Big)$.

Proof. This is a simple conclusion from Corollary 4.20 and the equality $G^{D^I}_{opt}(\Delta^>_I) = G^{D^I}_{opt}(\Delta_I) \setminus G^{D^I}_{opt}(\partial\Delta_I)$: It is clear that Corollary 4.20 also holds for all subsets $I \subset \{1,\dots,k\}$. Apart from that we only changed the condition $\beta \in \{H^{D^I}_{opt}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^{|I|}_{\gneq}$ to $\beta \in \{H^{D^I}_{opt}(\tilde{z}), e_i\}^\perp \cap [0,\infty)^{|I|}$ to obtain the slightly modified result. □

For checking the condition from Corollary 5.5, we need the set $G^{D^I}_{opt}(\partial\Delta_I)$, whose computation requires the set $W^I_{opt}(\partial\Delta_I)$. Thus, although we obtain $W^I_{opt,w}(\partial\Delta_I)$ by solving all of the subproblems (see Lemma 5.3), we only need $W^I_{opt}(\partial\Delta_I)$ to construct the new reference points for the subproblem $(J_i)_{i\in I}$.

Now we can state Algorithm 1 for solving a multiobjective optimization problem with k cost functions, and show that it computes the (weak) Pareto set (and therefore also the (weak) Pareto front) of the multiobjective optimization problem.

5.6 Theorem. *Let Assumption 1 be satisfied. Then Algorithm 1 computes the Pareto set $\mathcal{U}_{opt}$ (the weak Pareto set $\mathcal{U}_{opt,w}$) and therefore, also the Pareto front $\mathcal{J}_{opt}$ (the weak Pareto front $\mathcal{J}_{opt,w}$) can be obtained.*

Proof. We show by induction that after the i-th iteration step ($i \in \{1,\dots,k\}$) the sets $W^I_{opt,w}(\Delta_I)$ are obtained for all $I \subset \{1,\dots,k\}$ with $|I| \le i$. Then the statement follows from $\mathcal{U}_{opt} = W_{opt}(\Delta_k)$ (where we applied the Pareto filter in line 16 of Algorithm 1 to remove points which are not Pareto optimal) and $\mathcal{U}_{opt,w} = W_{opt,w}(\Delta_k)$.
<u>$i = 1$:</u> For $I = \{j\}$ for any $j \in \{1,\dots,k\}$, it holds

$$W^I_{opt,w}(\Delta_I) = W^I_{opt}(\Delta_I) = \{u \in U_{ad} \mid u \text{ minimizes } J_j\}.$$

This is computed in line 4 of Algorithm 1.
<u>$i \to i+1$:</u> ($i < k$) Let $I \subset \{1,\dots,k\}$ with $|I| = i+1$ be arbitrary. By induction assumption the sets $W^J_{opt,w}(\Delta_J)$ for all $J \subsetneq I$ are already computed. Thus, by Lemma 5.3 we have the set $W^I_{opt,w}(\partial\Delta_I)$, which can be used to compute $G^{D^I}_{opt}(\partial\Delta_I)$. Using the characterization from Corollary 5.5 we can obtain $G^{D^I}_{opt}(\Delta^>_I) = G^{D^I}_{opt,w}(\Delta^>_I)$. By Lemma 5.4 solving (ERPP) for all reference points in $G^{D^I}_{opt}(\Delta^>_I)$ results in the set $W^I_{opt,w}(\Delta^>_I) = W^I_{opt}(\Delta^>_I)$. In total, this gives us

$$W^I_{opt,w}(\Delta_I) = W^I_{opt,w}(\Delta^>_I) \cup W^I_{opt,w}(\partial\Delta_I),$$

which concludes the proof. □

6. Numerical Results

In this section we investigate a convex multiobjective optimization problem to explain how Algorithm 1 is realized numerically. In particular, we will show how the set $G^{D^I}_{opt}(\Delta^>_I)$ can be obtained given a discretization of the set $G^{D^I}_{opt}(\partial\Delta_I)$ (see line 11 of Algorithm 1).

6.1. Numerical Generation of Reference Points

Here we want to propose a way to generate reference points numerically by using the characterization from Corollary 5.5. To simplify the readability, we will use the following notation.

6.1 Notation. (i) Let $z \in \mathbb{R}^k$ and $i_1,\dots,i_l \in \{1,\dots,k\}$ be arbitrary, but pairwise different. Then $z_{-(i_1,\dots,i_l)} \in \mathbb{R}^{k-l}$ stands for the vector $(z_j)_{j\in\{1,\dots,k\}\setminus\{i_1,\dots,i_l\}}$.

(ii) Let $K, I \subset \{1,\dots,k\}$ with $K \subset I$ be arbitrary. Then we write

$$\Delta^K_I := \{\alpha \in \Delta_I \mid \alpha_i = 0 \text{ for all } i \notin K\}.$$

Algorithm 1: Algorithm for solving a multiobjective optimization problem with k cost functions

Data: $J_1, \dots, J_k$: Cost functions;
computeParetoFront: Boolean variable to decide whether the Pareto front (true) or the weak Pareto front (false) shall be computed;

```
begin
    Sol := ∅;
    for j = 1 : k do
        Set I := {j};
        Compute {u ∈ U_ad | J_j(u) = min_{ũ∈U_ad} J_j(ũ)};
        Sol ← Sol ∪ {(u, e_j, I) | u minimizes J_j} with e_j ∈ ℝ^k;
    end
    Compute y^id and ỹ^id;
    for i = 2 : k do
        for all I ⊂ {1, ..., k} with |I| = i do
            Define D_l^I := {y ∈ ℝ^i | y ≥ (ỹ^id)^I, y_l = (ỹ^id)_l^I} and set D^I := ∪_{l=1}^i D_l^I;
            Compute G_opt^{D^I}(∂Δ_I) by using the solutions of all subproblems, i.e., all tupels
              (u, α, K) ∈ Sol with K ⊂ I (use the weight α^I ∈ ∂Δ_I to compute G_opt^{D^I}(α^I));
            Use the characterization from Corollary 5.5 to obtain the set G_opt^{D^I}(Δ_I^>);
            for all z ∈ G_opt^{D^I}(Δ_I^>) do
                Compute {u ∈ U_ad | u minimizes (ERPP)_z};
                Set
                  Sol ← Sol ∪ {(u, α, I) | u minimizes (ERPP)_z, α^I = H_opt^{D^I}(z), α^{{1,...,k}\I} = 0};
            end
        end
    end
    if computeParetoFront == true then
        Remove all tupels (u, α, I) ∈ Sol, which are not Pareto optimal (Pareto filter);
    end
end
```

Let $I \subset \{1, \dots, k\}$ be arbitrary and assume that all subproblems of $(J_i)_{i \in I}$ have already been solved, so that $G_{opt}^{D^I}(\partial \Delta_I)$ can be computed. Given $z \in D_i$ for an arbitrary $i \in \{1, \dots, |I|\}$ it holds $z \notin G_{opt}^{D^I}(\Delta_I^>)$ if and only if

$$\exists \tilde{z} \in G_{opt}^{D^I}(\partial \Delta_I) \cap D_i : \exists \beta \in \{H_{opt}^{D^I}(\tilde{z}), e_i\}^\perp \cap [0, \infty)^{|I|} : z = \tilde{z} + \beta. \tag{9}$$

This is equivalent to

$$\exists j \in I \setminus \{i\} : \exists\, \tilde{z} \in G_{opt}^{D^I}(\Delta_I^{I \setminus \{j\}}) \cap D_i : \exists\, \beta \in \{H_{opt}^{D^I}(\tilde{z}), e_i\}^\perp \cap [0, \infty)^{|I|} : z = \tilde{z} + \beta, \tag{10}$$

since $\partial \Delta_I = \cup_{j=1}^{|I|} \Delta_I^{I \setminus \{j\}}$. As we know that for $\tilde{z} \in G_{opt}^{D^I}(\Delta_I^{I \setminus \{j\}})$ it holds $(H_{opt}^{D^I}(\tilde{z}))_j = 0$, the characterization (10) is equivalent to

$$\begin{aligned} \exists j \in I \setminus \{i\} : &\exists \tilde{z} \in G_{opt}^{D^I}(\Delta_I^{I \setminus \{j\}}) \cap D_i : \\ &\exists \gamma \in \{H_{opt}^{D^I}(\tilde{z}), e_i, e_j\}^\perp \cap [0, \infty)^{|I|} : \exists t \geq 0 : z = \tilde{z} + \gamma + t e_j. \end{aligned} \tag{11}$$

Note that $\{H^{D^I}_{opt}(\tilde{z}), e_i, e_j\}^\perp \cap [0,\infty)^{|I|} \neq \{0\}$ if and only if $(H^{D^I}_{opt}(\tilde{z}))_l = 0$ for at least one $l \in I \setminus \{j\}$, i.e., if $\tilde{z} \in G^{D^I}_{opt}(\Delta_I^{I\setminus\{j,l\}})$ holds.
By defining a function

$$F^i_j : \{\tilde{z} + \gamma \mid \tilde{z} \in G^{D^I}_{opt}(\Delta_k^{K\setminus\{j\}}) \cap D_i,\ \gamma \in \{H^{D^I}_{opt}(\tilde{z}), e_i, e_j\}^\perp \cap [0,\infty)^k\} \to \mathbb{R}$$

by

$$F^i_j((z+\gamma)_{-(i,j)}) := z_j,$$

we can write this condition equivalently as

$$\begin{aligned}\exists j \in I \setminus \{i\} :& \exists \tilde{z} \in G^{D^I}_{opt}(\Delta_I^{I\setminus\{j\}}) \cap D_i : \\ & \exists \gamma \in \{H^{D^I}_{opt}(\tilde{z}), e_i, e_j\}^\perp \cap [0,\infty)^{|I|} : \Big(z_{-j} = (\tilde{z}+\gamma)_{-j} \ \&\ z_j \geq F^i_j(z_{-(i,j)})\Big). \end{aligned} \tag{12}$$

Numerically, we only have a discretization of the set $G^{D^I}_{opt}(\Delta_k^{K\setminus\{j\}}) \cap D_i$. Therefore, we propose the following Algorithm 2, which will be tested in the next section.

Algorithm 2: Computing reference points given $G^{D^I}_{opt}(\partial\Delta_I)$

Data: $G^{D^I}_{opt}(\partial\Delta_I)$: Computed by using the solutions to all subproblems of $(J_i)_{i\in I}$.

begin
- $Z = \emptyset$;
- **for** $i = 1 : |I|$ **do**
  - Define a grid $Gr_i \subset D_i^I$;
  - **for** $j = 1 : |I|$, $j \neq i$ **do**
    - Determine all points $z^1, \ldots, z^{N_j^i} \in G^{D^I}_{opt}(\partial\Delta_I) \cap D_i$ with $(H^{D^I}_{opt}(z^i))_j = 0$;
    - Discretize the sets $z^l + \{\gamma \mid \gamma \in \{H^{D^I}_{opt}(z^l), e_i, e_j\}^\perp \cap [0,\infty)^{|I|}\}$ for all $l \in \{1, \ldots, N_j^i\}$;
    - Inter-/ extrapolate the function F^i_j given the data $(z^l + \gamma^p)_{-(i,j)} \mapsto z^l_j$ with $z^l + \gamma^p \in z^l + \{\gamma \mid \gamma \in \{H^{D^I}_{opt}(z^l), e_i, e_j\}^\perp \cap [0,\infty)^k\}$ (If there are $(z^{l_1} + \gamma^{p_1})_{-(i,j)} = (z^{l_2} + \gamma^{p_2})_{-(i,j)}$ choose the smaller value of $z^{l_1}_j, z^{l_2}_j$.);
    - Evaluate the interpolated function F^i_j at $\{g_{-(i,j)} \mid g \in Gr_i\}$;
    - Set $Gr_i \leftarrow Gr_i \cap \{g \in Gr_i \mid g_j \leq \tilde{F}^i_j(g_{-(i,j)})\}$;
  - Set $Z \leftarrow Z \cup Gr_i$;

6.2. Numerical Computation of the Pareto Front

Note that in this section we leave away the index I in the captions and legends of figures, in order to preserve readability. However, from the context it is always clear if we refer to an object related with the original problem or a subproblem.
For our example we choose $U := \mathbb{R}^5$, the admissible set $U_{ad} := [-1,1]^5$ and define the four cost

functions $J_1, \dots, J_4 : U_{ad} \to \mathbb{R}$ by

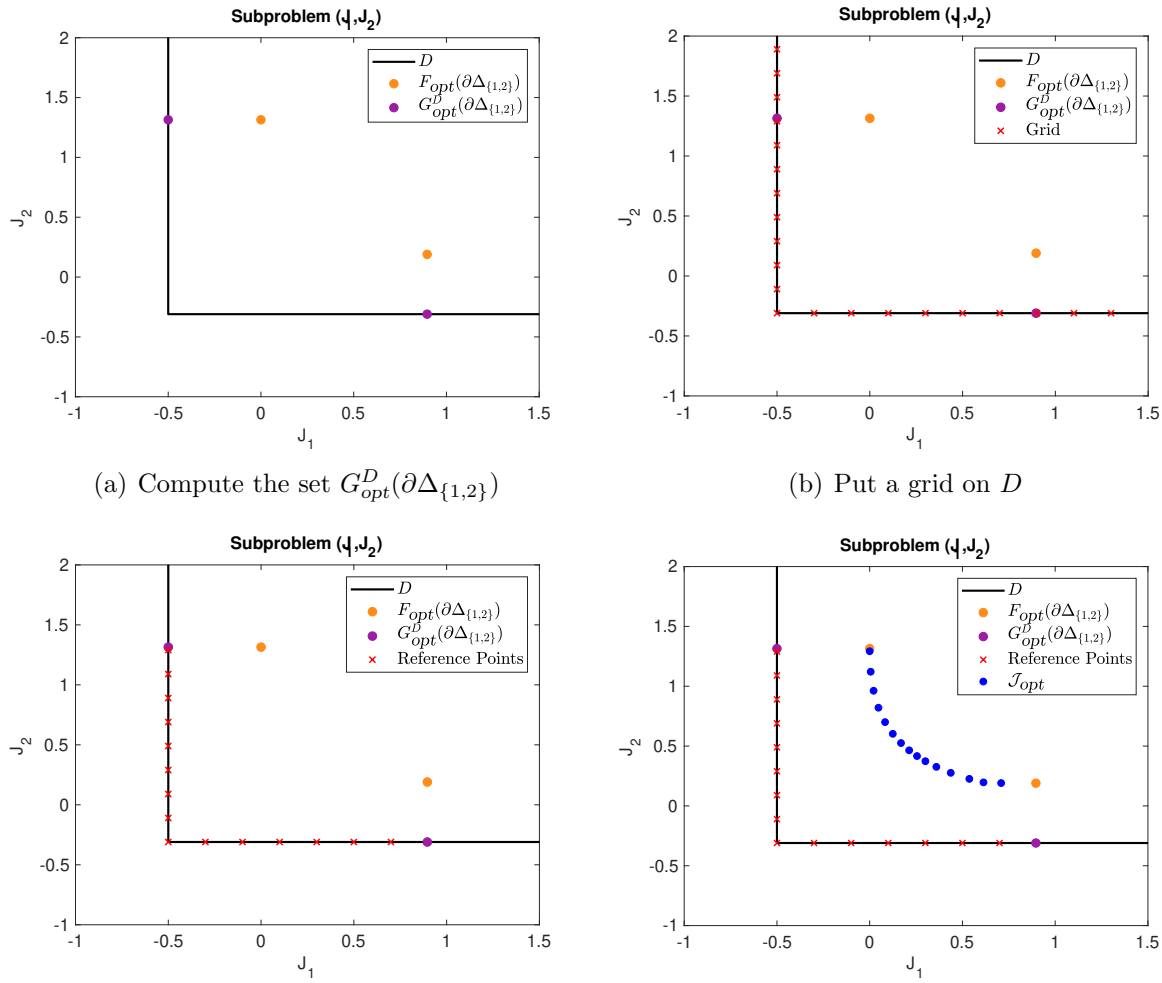
$$J_1(u) := \frac{1}{5}\left[u_1^2 + 5u_2^4 + 2u_3^2 + (u_4 - 1)^4 + 2(u_5 + 0.2)^4\right],$$
$$J_2(u) := \frac{1}{100}\left[3(u_1 - 2)^4 + 2u_2^2 + (u_3 + 3)^4 + u_4^2 + 3(u_5 - 0.5)^2\right],$$
$$J_3(u) := \frac{1}{40}\left[(u_1 + 5)^2 + 4(u_2 - 2)^4 + u_3^2 + u_4^4 + u_5^4\right],$$
$$J_4(u) := \frac{1}{10}\left[2(u_1 - 3)^2 + 3(u_2 + 1)^4 + 3(u_3 - 1)^4 + (u_4 + 1)^2 + 2(u_5 - 0.5)^2\right].$$

As can be easily verified the cost functions are even strongly convex and twice continuously differentiable, so that, in particular, Theorem 4.15 and Corollary 4.16 hold true in this case.
The first step of solving the multiobjective optimization problem is to compute the minimizers of the cost functions $J_1, \dots, J_4$. These minimizers are then used to compute the sets $G_{opt}^{D^I}(\partial\Delta_I)$ for any $I \subset \{1, \dots, 4\}$ with $|I| = 2$. For such sets I it holds $\partial\Delta_I = \{(1,0)^T, (0,1)^T\}$. For the set $I = \{1, 2\}$ the construction of reference points is shown in Figure 4. After putting a grid on D^I it is straightforward to obtain the set $G_{opt}^{D^I}(\Delta_I^>)$ by using the characterization from Corollary 5.5 directly, so that the inter-/extrapolation of the function F_j^i in Algorithm 2 is not needed in this case. The procedure is depicted in Figure 4 (c). Given the new reference points we can then solve all arising (ERPP) to obtain the solution to the subproblem (J_1, J_2) (see Figure 4 (d)).

After having solved all subproblems with two cost functions, we continue by looking at subproblems with three cost functions. Thereby, we proceed in the same way as for two cost functions. Given $I \subset \{1, \dots, 4\}$ with $|I| = 3$ the set $G_{opt}^{D^I}(\partial\Delta_I)$ is computed by using information of the subproblems of size 2. In Figure 5 (a) the sets $F_{opt}^I(\partial\Delta_I)$ and $G_{opt}^{D^I}(\partial\Delta_I)$ can be seen for $I = \{1, 2, 3\}$. In Figures 5 (b)-(e) the procedure of generating new reference points on D_1^I is shown. First, a grid on D_1^I is constructed which covers the set $G_{opt}^{D^I}(\partial\Delta_I) \cap D_1^I$. Then we follow Algorithm 2: Choose $j = 2$. All points $z^1, \dots, z^{N_2^1} \in G_{opt}^{D^I}(\Delta_{\{1,2,3\}}^{\{1,3\}})$ are determined, the sets $z^l + \{\gamma \mid \gamma \in \{H_{opt}^{D^I}(z^l), e_1, e_2\}^\perp \cap [0, \infty)^3\}$ are discretized and the function F_2^1 is inter-/extrapolated. (Note that $\{\gamma \mid \gamma \in \{H_{opt}^{D^I}(z^l), e_1, e_2\}^\perp \cap [0, \infty)^3\} \neq \{0\}$ only holds for $z^l = G_{opt}^{D^I}((1, 0, 0))$). This can be seen in Figure 5 (c). Then all grid points g fulfilling $g_2 \geq \tilde{F}_2^1(g_3)$ are removed (c.f. Figure 5 (d)). The same procedure is repeated for $j = 3$, as is shown in Figure 5 (e). The remaining grid points are the new reference points on D_1^I. Repeating the procedure for D_2^I and D_3^I results in a discretization of the set $G_{opt}^{D^I}(\Delta_I^>)$, i.e., the new reference points, see Figure 5 (f). Solving (ERPP) for these reference points results in the Pareto front for this subproblem (see Figure 7 (a)).

Finally, in the last step reference points for the original multiobjective optimization problem can be constructed using information about all previously solved subproblems. Since the sets $D_1, \dots, D_4$ are 3-dimensional planes in $\mathbb{R}^4$ in this case, we cannot show the complete set $G_{opt}^D(\partial\Delta_4)$. However, in Figure 6 (a) the set $G_{opt}^D(\partial\Delta_4) \cap D_1$ is depicted. In Figure 6 (b) the resulting reference points can be seen after Algorithm 2 has been applied. Doing the same for the other three sets D_2, D_3, D_4, we obtain a discretization of the set $G_{opt}^D(\Delta_4)$. Solving (ERPP) for all of these reference points yields a covering of the Pareto front of the original problem. Again, we cannot show this result since the Pareto front is a three dimensional object in $\mathbb{R}^4$.

To conclude this section, Algorithm 1 using the Euclidean reference point method is compared to the weighted sum method. For this purpose we consider the multiobjective optimization problem



(a) Compute the set $G^D_{opt}(\partial\Delta_{\{1,2\}})$

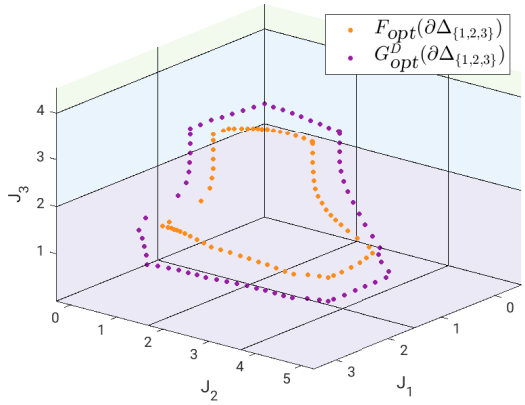
(b) Put a grid on D

(c) Remove all z with $z = G^D_{opt}((1,0)) + (0,t)^T$ or $z = G^D_{opt}((0,1)) + (t,0)^T$ for a $t \geq 0$

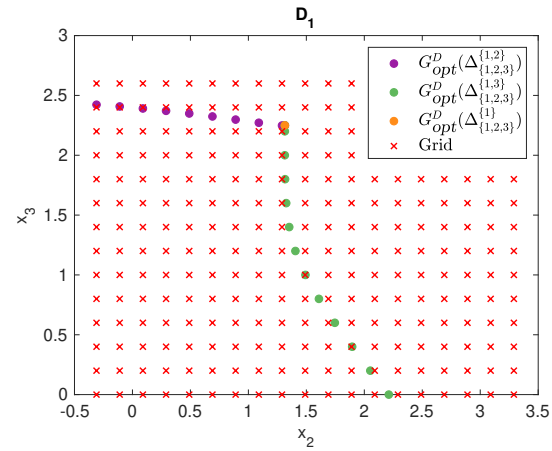
(d) Solve (ERPP) for all reference points to obtain an approximation of $\mathcal{J}_{opt}$

Figure 4: Generating new reference points from information about $G^D_{opt}(\partial\Delta_{\{1,2\}})$ and solving the subproblem (J_1, J_2)

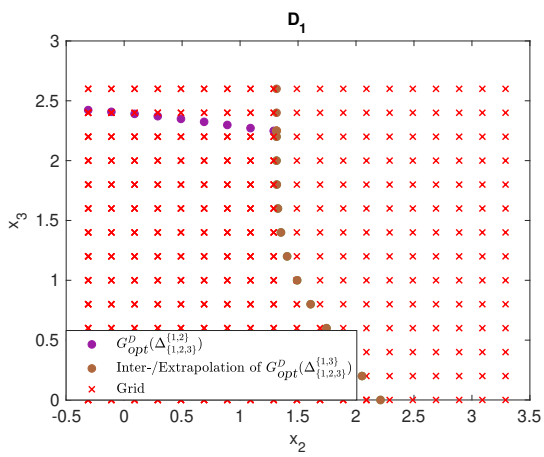
with the three cost functions (J_1, J_2, J_3) from the previous example.
As we can see in Figure 7 (a), Algorithm 1 using the Euclidean reference point method results in an almost equidistant covering of the Pareto front, also in regions which are close to the boundary. Only in the middle part of the Pareto front we can see a clustering of Pareto optimal points. This is due to the fact that we choose the reference points in $D = D_1 \cup D_2 \cup D_3$. Close to the intersection of the three coordinate planes D_1, D_2, D_3 there is a clustering of reference points which transfers to the Pareto front. In total 506 (ERPP) are solved to obtain this result.
For the weighted sum method we choose 819 equidistantly distributed weights. The resulting covering of the Pareto front is shown in Figure 7 (b). It is obvious that the upper part and some of the lower parts of the Pareto front are not approximated well by the weighted sum method. This behaviour was also observed in [DD97], and it leads to the conclusion that, in general, the Euclidean reference point method performs better than the weighted sum method.



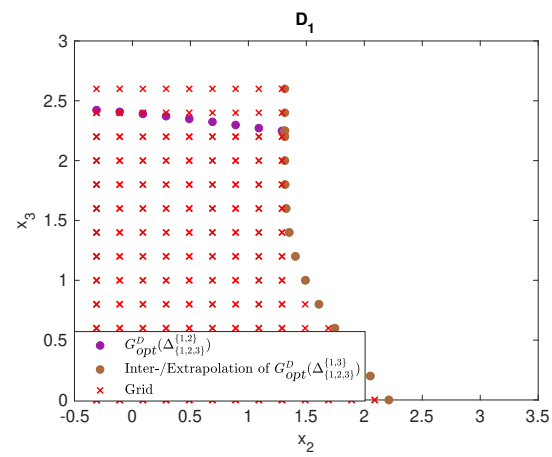
(a) The set $G^D_{opt}(\partial\Delta_{\{1,2,3\}})$



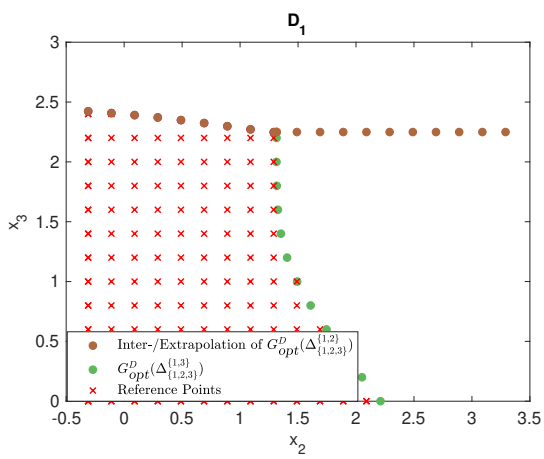
(b) Create a grid covering the set $G^D_{opt}(\partial\Delta_{\{1,2,3\}}) \cap D_1$



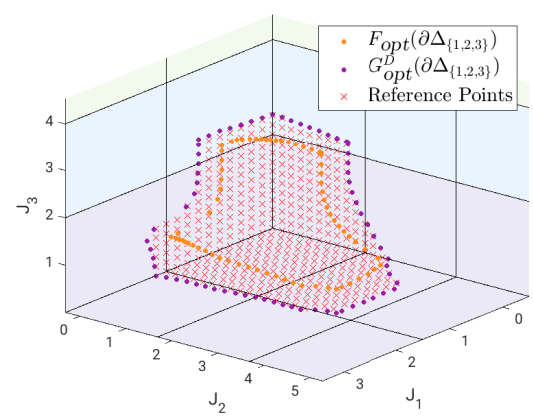
(c) Interpolate the points $G^D_{opt}(\Delta^{\{1,3\}}_{\{1,2,3\}}) \cap D_1$ in x_3-direction



(d) Remove all z with $z = y + (0, t, 0)^T$ with y an interpolation point of $G^D_{opt}(\Delta^{\{1,3\}}_{\{1,2,3\}}) \cap D_1$

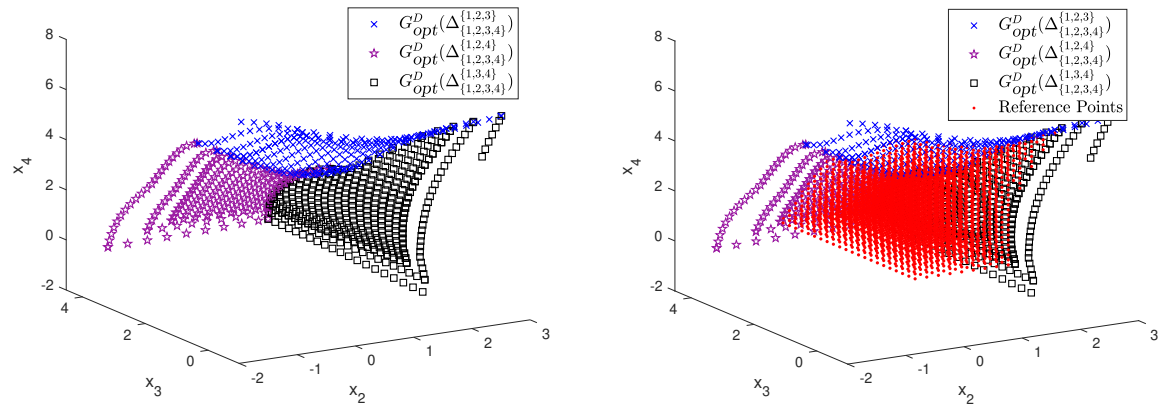


(e) Interpolate the points $G^D_{opt}(\Delta^{\{1,2\}}_{\{1,2,3\}}) \cap D_1$ in x_2-direction and remove the respective grid points



(f) Discretization of the set $G^D_{opt}(\Delta_{\{1,2,3\}})$

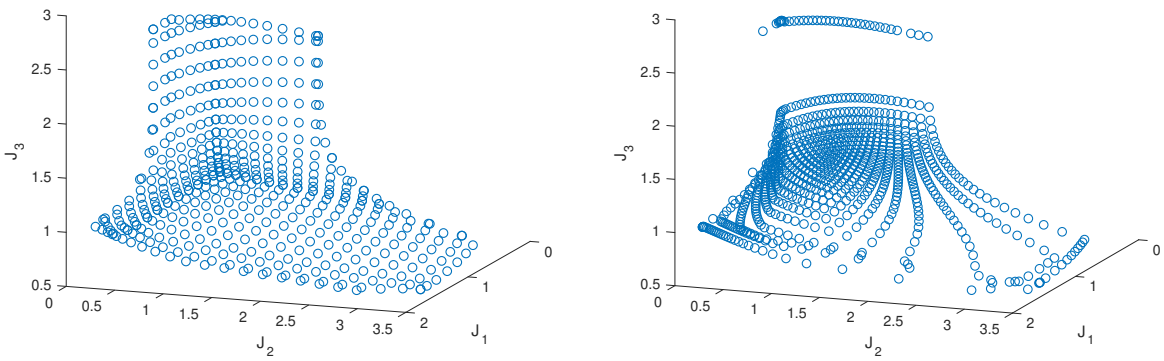
Figure 5: Generating new reference points on D_1 from information about $G^D_{opt}(\partial\Delta_{\{1,2,3\}})$



(a) The set $G^D_{opt}(\partial\Delta_4) \cap D_1$

(b) Generated reference points

Figure 6: Generation of reference points for four cost functions



(a) Approximation of the Pareto front using the ERPM

(b) Approximation of the Pareto front using the WSM

Figure 7: Comparison of the ERPM and the WSM for the problem (J_1, J_2, J_3)

7. Conclusion and Outlook

In this article we investigate the use of the Euclidean reference point method for solving convex multiobjective optimization problems with an arbitrary number of cost functions. With the help of an auxiliary problem – the extended Euclidean reference point method – the notion of (weakly) Pareto admissible reference points is introduced. For these reference points it is shown that

(i) (ERPP) is solvable and the minimizers are (weakly) Pareto optimal.

(ii) every weakly Pareto optimal point can be computed by solving (ERPP) to a (weakly) Pareto admissible reference point.

However, the set of all (weakly) Pareto admissible reference points is not suitable for a numerical implementation for two reasons:

(i) the set is too large in the sense that the solutions of (ERPP) to many reference points coincide (see Remark 3.24).

(ii) the characterization of (weakly) Pareto admissible reference points depends on the solution to all weighted sum problems, i.e., the (weak) Pareto front has to be computed to obtain (weakly) Pareto admissible reference points.

Therefore, the restriction of these sets to a union of shifted coordinate planes D is investigated. It turns out that the (weakly) Pareto admissible reference points on D can be characterized by the solutions to subproblems of the original problem, i.e., problems where one or several cost functions are neglected. This gives rise to a hierarchical method for computing both the reference points and in consequence the (weak) Pareto front.
In the numerical tests this method is demonstrated for an example with four cost functions. However, it can also be used for an arbitrary number of cost functions, since there are no restrictions on the dimension of the subproblems in Algorithm 2 for choosing new reference points. Consequently, this method overcomes the issue of not being able to deal with a large number of cost functions when using the Euclidean reference point method. In comparison with the weighted sum method the Euclidean reference point method yields a much more equidistant covering of the Pareto front, which is both shown theoretically (c.f. Theorem 3.10) and practically (see Figure 7 (a)).
For future work there are some theoretical and practical aspects that could be investigated further:

- Due to the structure of the set D, there is always a clustering of reference points in regions close to the intersection of two of the coordinate planes D_i and D_j, since the coordinate planes are orthogonal to each other. Thus, the reference points are also closer to each other in these regions. This transfers to the Pareto front, so that we can observe a clustering of reference points in the 'middle' of the Pareto front. A more intelligent construction of the grid for choosing the reference points might resolve this problem.

- Algorithm 2 for computing the new reference points becomes very costly for an increasing number of cost functions. This is due to the fact that the sets $z^l + \{\gamma \mid \gamma \in \{H_{opt}^{D^I}(z^l), e_i, e_j\}^\perp \cap [0,\infty)^{|I|}\}$ are up to $k-2$ dimensional. As they are discretized and the obtained data need to be inter-/extrapolated the procedure becomes computationally very expensive.

- The presented method can be easily parallelized in two ways. First, the solution of subproblems of the same size can be done in parallel, since only information of subproblems with fewer cost functions are needed. Second, having computed the reference points for a problem, solving all (ERPP) can also be parallelized.

- Here we restricted ourselves to convex problems. In [KSd15, MGGS09] a similar approach was studied for nonconvex problems using the Pascoletti-Serafini scalarization. However, both papers do not prove that their methods produce approximations of the complete Pareto front. We are optimistic that the approach from this article can be extended to the nonconvex case and the Pascoletti-Serafini scalarization in order to prove this in a similar manner as in Theorem 5.6.

Bibliography

[Ban17] S. Banholzer. POD-Based Bicriterial Optimal Control of Convection-Diffusion Equations. Master's thesis, University of Konstanz, 2017.

[BBV16] S. Banholzer, D. Beermann, and S. Volkwein. POD-Based Bicriterial Optimal Control by the Reference Point Method. *IFAC-PapersOnLine*, 49(8):210–215, 2016.

[BBV17] S. Banholzer, D. Beermann, and S. Volkwein. POD-Based Error Control for Reduced-Order Bicriterial PDE-Constrained Optimization. *Annual Reviews in Control*, 44:226–237, 2017.

[Bee19] D. Beermann. *POD-based scalar and multiobjective optimal control of evolution equations including operator-valued variables*. PhD thesis, University of Konstanz, 2019.

[BKR17] R. Burachik, C. Kaya, and M. Rizvi. A new scalarization technique and new algorithms to generate pareto fronts. *SIAM Journal on Optimization*, 27(2):1010–1034, 2017.

[DD97] I. Das and J.E. Dennis. A closer look at drawbacks of minimizing weighted sums of objectives for pareto set generation in multicriteria optimization problems. *Structural Optimization*, 14:63–69, 1997.

[Deb01] K. Deb. *Multi-Objective Optimization Using Evolutionary Algorithms*. John Wiley & Sons, first edition edition, 2001.

[DK19] A. Dolatnezhadsomarin and E. Khorram. Two efficient algorithms for constructing almost even approximations of the pareto front in multi-objective optimization problems. *Engineering Optimization*, 51(4):567–589, 2019.

[Ehr05] M. Ehrgott. *Multicriteria Optimization*. Springer, Berlin, Heidelberg, second edition edition, 2005.

[Eic08] G. Eichfelder. *Adaptive Scalarization Methods in Multiobjective Optimization*. Springer, Berlin Heidelberg, 2008.

[GPD19] B. Gebken, S. Peitz, and M. Dellnitz. On the hierarchical structure of pareto critical sets. *Journal of Global Optimization*, 73(4):891–913, 2019.

[Hil01] C. Hillermeier. *Nonlinear Multiobjective Optimization: A Generalized Homotopy Approach*. Birkhäuser, Basel Boston Berlin, 2001.

[ITV15] L. Iapichino, S. Trenz, and S. Volkwein. Reduced-order multiobjective optimal control of semilinear parabolic problems, 2015.

[IUV13] L. Iapichino, S. Ulbrich, and S. Volkwein. Multiobjective pde-constrained optimization using the reduced-basis method, 2013.

[KSd15] K. Khaledian and M. Soleimani-damaneh. A new approach to approximate the bounded pareto front. *Mathematical Methods of Operations Research*, 82(2):211–228, Oct 2015.

[MGGS09] D. Mueller-Gritschneder, H. Graeb, and U. Schlichtmann. A successive approach to compute the bounded pareto front of practical multiobjective optimization problems. *SIAM J. Optim.*, 20(2):915–934, 2009.

[Mie99] K.M. Miettinen. *Nonlinear Multiobjective Optimization*. Kluwer Academic Publishers, AH Dordrecht, first edition edition, 1999.

[Par71] V. Pareto. *Manual of Political Economy*. Augustus M. Kelly Publishers, New York, 1971.

[POBD15] S. Peitz, S. Ober-Blöbaum, and M. Dellnitz. Multiobjective optimal control methods for fluid flow using reduced order modeling, 2015.

[PS84] A. Pascoletti and P. Serafini. Scalarizing vector optimization problems. *Journal of Optimization Theory and Applications*, 42(4):499–524, 1984.

[Rud08] W. Rudin. *Real and Complex Analysis*. McGraw-Hill Higher Education, 3. edition edition, 2008.

[Wie80] A. P. Wierzbicki. The use of reference objectives in multiobjective optimization. In *Multiple Criteria Decision Making Theory and Application*, pages 468–486, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg.

[Wie86] A. P. Wierzbicki. On the completeness and constructiveness of parametric characterizations to vector optimization problems. *Operations-Research-Spektrum*, 8(2):73–87, Jun 1986.

[Wie98] A.P. Wierzbicki. Reference point methods in vector optimization and decision support. Iiasa interim report, IIASA, Laxenburg, Austria, April 1998.

Appendix A. Proofs of Section 4

Proof of Theorem 4.14:

(i) Proof of $\partial G^D_{opt,w}(\Delta_k) \cap G^D_{opt,w}(\Delta_k) \subset G^D_{opt,w}(\partial\Delta_k)$:
Let $z \in D \setminus G^D_{opt,w}(\Delta_k)$ be arbitrary. Then there is $i \in \{1,\dots,k\}$ with $z \in D_i$ and $z \notin G^D_{opt,w}(\Delta_k) \cap D_i$. The projection $P(z)$ of z onto $J(U_{ad}) + [0,\infty)^k$ is given by $P(z) = J(u) + d$ for a $u \in U_{ad}$ and a $d \in [0,\infty)^k$. By Theorem 4.8 it holds $\mathcal{Z}_{opt,w} \cap D = G^D_{opt,w}(\Delta_k)$, so that we can conclude $z \notin \mathcal{Z}_{opt,w}$, i.e., $P(z) \notin \mathcal{J}_{opt,w}$. Corollary 3.16 directly implies $P(z) \notin J(U_{ad})$, i.e., it holds $d \gneqq 0$. Now Theorem 3.15 (i) yields $H(z) \in \partial\Delta_k$. As $z \in D \setminus G^D_{opt,w}(\Delta_k)$ was arbitrary this implies $H(D \setminus G^D_{opt,w}(\Delta_k)) \subset \partial\Delta_k$. Since H is continuous and $\partial\Delta_k$ is closed, we get

$$H(\partial G^D_{opt,w}(\Delta_k)) \subset H(\overline{D \setminus G^D_{opt,w}(\Delta_k)}) \subset \partial\Delta_k.$$

Finally, this implies

$$\partial G^D_{opt,w}(\Delta_k) \cap G^D_{opt,w}(\Delta_k) \subset H^{-1}(\partial\Delta_k) \cap G^D_{opt,w}(\Delta_k) = (H^D_{opt,w})^{-1}(\partial\Delta_k) = G^D_{opt,w}(\partial\Delta_k),$$

where we used Theorem 4.12 for the last equality. This concludes the proof.

(ii) Proof of $\partial G^D_{opt}(\Delta_k) \cap G^D_{opt}(\Delta_k) = G^D_{opt}(\partial\Delta_k)$:
We first show $\partial G^D_{opt}(\Delta_k) \cap G^D_{opt}(\Delta_k) \subset G^D_{opt}(\partial\Delta_k)$. Let $z \in D \setminus G^D_{opt}(\Delta_k)$ be arbitrary.
If even $z \notin G^D_{opt,w}(\Delta_k)$ we can conclude as in (i) that $H(z) \in \partial\Delta_k$ holds.
On the other hand, if $z \in G^D_{opt,w}(\Delta_k) \setminus G^D_{opt}(\Delta_k)$ holds, we can use the equality $G^D_{opt,w}(\Delta_k^{>}) = G^D_{opt}(\Delta_k^{>})$, which was shown in Corollary 4.6, and Theorem 4.12 to conclude that $z \in G^D_{opt,w}(\partial\Delta_k) = \left(H^D_{opt,w}\right)^{-1}(\partial\Delta_k)$, i.e., also in this case it holds $H(z) = H^D_{opt,w}(z) \in \partial\Delta_k$.

As $z \in D \setminus G^D_{opt}(\Delta_k)$ was arbitrary this implies $H(D \setminus G^D_{opt}(\Delta_k)) \subset \partial\Delta_k$. With the same arguments as in the proof of the first statement, we can show

$$\partial G^D_{opt}(\Delta_k) \cap G^D_{opt}(\Delta_k) \subset H^{-1}(\partial\Delta_k) \cap G^D_{opt}(\Delta_k) = (H^D_{opt})^{-1}(\partial\Delta_k) = G^D_{opt}(\partial\Delta_k).$$

This finishes the proof of $\partial G^D_{opt}(\Delta_k) \cap G^D_{opt}(\Delta_k) \subset G^D_{opt}(\partial\Delta_k)$.
To show the other implication let $z \in G^D_{opt}(\partial\Delta_k)$ be arbitrary. Since $G^D_{opt}(\partial\Delta_k) \subset G^D_{opt}(\Delta_k)$ is trivially satisfied, we only need to show that $z \in \partial G^D_{opt}(\Delta_k)$ holds. For this it suffices to show that there is a sequence $(z_n)_{n\in\mathbb{N}} \subset D$ with $z_n \to z$ as $n \to \infty$, but $z_n \notin G^D_{opt}(\Delta_k)$ for all $n \in \mathbb{N}$. In the following we will construct such a sequence.
Let $i \in \{1, \dots, k\}$ with $z \in D_i$. Since $z \in G^D_{opt}(\partial\Delta_k)$ there is $\alpha \in \partial\Delta_k$ such that $z = y - t\alpha$ for a $y \in F_{opt}(\alpha)$ with $t = \min_{j\in\{1,\dots,k\}} \frac{y_j - \tilde{y}^{id}_j}{\alpha_j} = \frac{y_i - \tilde{y}^{id}_i}{\alpha_i} > 0$. Choose a $j \in \{1, \dots, k\}$ with $\alpha_j = 0$. In particular, it holds $i \neq j$. Define $\tilde{z}(s) := z + se_j = y - t\alpha + se_j$ for every $s > 0$. Then it holds $\tilde{z}(s) \in D_i \subset D$ and furthermore, $P(\tilde{z}(s)) = P(z) + se_j = y + se_j \in J(U_{ad}) + [0, \infty)^k$ for all $s > 0$. To see the latter, we plug this expression into the sufficient optimality condition (1) and get for any $\tilde{y} \in J(U_{ad}) + [0, \infty)^k$

$$\begin{aligned}\langle y + se_j - \tilde{z}(s), \tilde{y} - y - se_j\rangle &= \langle y + se_j - y + t\alpha - se_j, \tilde{y} - y - se_j\rangle \\ &= \langle t\alpha, \tilde{y} - y - se_j\rangle = \langle t\alpha, \tilde{y} - y\rangle \geq 0,\end{aligned}$$

where we used $\langle \alpha, e_j\rangle = 0$ and the fact that $y \in F_{opt}(\alpha)$ is the solution of (WSP) to the weight α. Furthermore, $P(\tilde{z}(s)) \notin F_{opt}(\Delta_k) = \mathcal{J}_{opt}$, since we have $y \in \mathcal{J}_{opt}$ and $P(\tilde{z}(s)) \gneqq y$. By definition this yields $\tilde{z}(s) \notin \mathcal{Z}_{opt}$ and thus, Theorem 4.8 tells us that $\tilde{z}(s) \notin G^D_{opt}(\Delta_k)$ for all $s > 0$, but $\tilde{z}(s) \to z$ as $s \to 0$. Therefore, it holds $z \in \partial G^D_{opt}(\Delta_k)$ which concludes the proof. □

Proof of Theorem 4.17:

Let $z \in D \setminus G^D_{opt,w}(\Delta_k)$ be arbitrary and choose $i \in \{1, \dots, k\}$ with $z \in D_i$. By Theorem 4.8 it holds $\mathcal{Z}_{opt,w} \cap D = G^D_{opt,w}(\Delta_k)$, so that we can conclude $z \notin \mathcal{Z}_{opt,w}$, i.e., $P(z) \notin \mathcal{J}_{opt,w}$. Corollary 3.16 directly implies $P(z) \notin J(U_{ad})$, which means that $P(z) = J(u) + d$ for a $u \in U_{ad}$ and a $d \in \mathbb{R}^k_{\gneqq}$. Theorem 3.15 (i) tells us that $J(u) \in F_{opt,w}(\alpha)$ for $\alpha := H(z) = \frac{P(z)-z}{\|P(z)-z\|_1} \in \Delta_k$, and $\langle d, \alpha\rangle = 0$. Since $d \gneqq 0$, this implies $\alpha \in \partial\Delta_k$. Moreover, since $z \in D_i$ we know that $(P(z))_i > z_i$, and thus $\alpha_i > 0$. Hence, $d_i = 0$. Altogether, this yields $d \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$. Note that we have

$$z = P(z) - (P(z) - z) = P(z) - \|P(z) - z\|_1 \alpha = P(z) - \frac{(P(z) - z)_i}{\alpha_i}\alpha = P(z) - \frac{(P(z))_i - \tilde{y}^{id}_i}{\alpha_i}\alpha,$$

since $\alpha = \frac{P(z)-z}{\|P(z)-z\|_1}$ and $z_i = \tilde{y}^{id}_i$ as $z \in D_i$. Next, we define $\tilde{z}$ by

$$\begin{aligned}\tilde{z} &:= J(u) - \frac{J_i(u) - \tilde{y}^{id}_i}{\alpha_i}\alpha = P(z) - d - \frac{(P(z))_i - d_i - \tilde{y}^{id}_i}{\alpha_i}\alpha \\ &= P(z) - \frac{(P(z))_i - \tilde{y}^{id}_i}{\alpha_i}\alpha - d + \frac{d_i}{\alpha_i}\alpha = z - d,\end{aligned}$$

since $d \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$. Therefore, $z = \tilde{z} + d$ is satisfied. Moreover, it holds $\tilde{z}_i = \tilde{y}^{id}_i$, and for every $j \in \{1, \dots, k\}$ we have that if $\alpha_j = 0$, it holds $\tilde{z}_j = J_j(u) > \tilde{y}^{id}_j$, and if $\alpha_j > 0$ it holds $d_j = 0$

and thus $\tilde{z}_j = z_j \geq \tilde{y}_j^{id}$, since $z \in D_i$. Together this implies $\tilde{z} \in D_i$ as well, and by definition $\tilde{z} \in G^D_{opt,w}(\alpha) \subset G^D_{opt,w}(\partial\Delta_k)$. Finally, Theorem 4.12 yields $H^D_{opt,w}(\tilde{z}) = \alpha$, which finishes the proof of the first statement.

To prove the second statement assume that (8) is fulfilled. If $z \notin G^D_{opt,w}(\Delta_k)$ we are done. If $z \in G^D_{opt,w}(\Delta_k)$ holds, then on the one hand there is $i \in \{1,\dots,k\}$, $\alpha \in \Delta_k$, $y \in F_{opt,w}(\Delta_k)$ and $t := \min_{i\in\{1,\dots,k\}} \frac{y_i - \tilde{y}_i^{id}}{\alpha_i}$ with $z = y - t\alpha \in D_i \subset D$. On the other hand we can write $z = \tilde{z} + \beta$ for some $\tilde{z} \in G^D_{opt,w}(\partial\Delta_k) \cap D_i$ and $\beta \in \{H^D_{opt,w}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$. Using these equalities it holds for all $y \in J(U_{ad}) + [0,\infty)^k$

$$\begin{aligned}\langle P(\tilde{z}) + \beta - z, y - P(\tilde{z}) - \beta\rangle &= \langle P(\tilde{z}) + \beta - \tilde{z} - \beta, y - P(\tilde{z}) - \beta\rangle \\ &= \langle P(\tilde{z}) - \tilde{z}, y - P(\tilde{z}) - \beta\rangle = \langle P(\tilde{z}) - \tilde{z}, y - P(\tilde{z})\rangle \geq 0,\end{aligned}$$

since $H^D_{opt,w}(\tilde{z}) \parallel P(\tilde{z}) - \tilde{z}$ and thus, $\langle P(\tilde{z}) - \tilde{z}, \beta\rangle = 0$. This computation shows $P(z) = P(\tilde{z}) + \beta$, so that we get

$$\alpha = H^D_{opt,w}(z) = \frac{P(z) - z}{\|P(z) - z\|_1} = \frac{P(\tilde{z}) + \beta - \tilde{z} - \beta}{\|P(\tilde{z}) + \beta - \tilde{z} - \beta\|_1} = \frac{P(\tilde{z}) - \tilde{z}}{\|P(\tilde{z}) - \tilde{z}\|_1} = H^D_{opt,w}(\tilde{z}) \in \partial\Delta_k,$$

which implies $z \in G^D_{opt,w}(\partial\Delta_k)$. Since we have $G^D_{opt,w}(\partial\Delta_k) \cap G^D_{opt,w}(\Delta_k^>) = \emptyset$ by Lemma 4.5 this concludes the proof. □

<u>Proof of Theorem 4.18:</u>

To prove (i) ⇒ (ii) assume that $z \notin G^D_{opt}(\Delta_k)$ holds. Theorem 4.8 yields that $z \notin \mathcal{Z}_{opt}$, i.e., it holds $P(z) \notin \mathcal{J}_{opt}$. We have $P(z) = J(u) + d$ for a $u \in U_{ad}$ and a $d \in [0,\infty)^k$. Theorem 3.15 (i) tells us that $J(u) \in F_{opt,w}(\alpha)$ for $\alpha := \frac{P(z)-z}{\|P(z)-z\|_1} \in \Delta_k$. Since $J(u) \in F_{opt,w}(\alpha)$ there is $y \in F_{opt}(\alpha)$ and $\tilde{d} \in [0,\infty)^k$ with $J(u) = y + \tilde{d}$. If $d = 0$ holds, we have $\tilde{d} \gneqq 0$, since otherwise $P(z) = J(u) = y \in F_{opt}(\alpha) \subset \mathcal{J}_{opt}$. So in total we can write $P(z) = y + d$ with $y \in F_{opt}(\alpha)$ and $d \gneqq 0$.
From Theorem 3.15 (i) we conclude $\langle d, \alpha\rangle = 0$ and because of $d \gneqq 0$ also $\alpha \in \partial\Delta_k$. Since $z \in D$ there is $i \in \{1,\dots,k\}$ with $z \in D_i$, i.e., $z_i = \tilde{y}_i^{id}$. In particular, this yields $(P(z))_i > z_i$ and thus, $\alpha_i > 0$. Hence, $d_i = 0$. In total, this yields $d \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$. Note that we have

$$z = P(z) - (P(z) - z) = P(z) - \|P(z) - z\|_1 \alpha = P(z) - \frac{(P(z) - z)_i}{\alpha_i}\alpha = P(z) - \frac{(P(z))_i - \tilde{y}_i^{id}}{\alpha_i}\alpha,$$

since $\alpha = \frac{P(z)-z}{\|P(z)-z\|_1}$ and $z_i = \tilde{y}_i^{id}$ as $z \in D_i$. Next, we define $\tilde{z}$ by

$$\begin{aligned}\tilde{z} := y - \frac{y_i - \tilde{y}_i^{id}}{\alpha_i}\alpha &= P(z) - d - \frac{(P(z))_i - d_i - \tilde{y}_i^{id}}{\alpha_i}\alpha \\ &= P(z) - \frac{(P(z))_i - \tilde{y}_i^{id}}{\alpha_i}\alpha - d + \frac{d_i}{\alpha_i}\alpha = z - d,\end{aligned}$$

since $d \in \{\alpha, e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$. Therefore, $z = \tilde{z} + d$ is satisfied. Moreover, it holds $\tilde{z}_i = \tilde{y}_i^{id}$, and for every $j \in \{1,\dots,k\}$ we have that if $\alpha_j = 0$, it holds $\tilde{z}_j = J_j(u) > \tilde{y}_j^{id}$, and if $\alpha_j > 0$ it holds $d_j = 0$ and thus $\tilde{z}_j = z_j \geq \tilde{y}_j^{id}$, since $z \in D_i$. Together this implies $\tilde{z} \in D_i$ as well, and by definition $\tilde{z} \in G^D_{opt}(\alpha) \subset G^D_{opt}(\partial\Delta_k)$. Finally, Theorem 4.12 yields $H^D_{opt}(\tilde{z}) = \alpha$, which finishes the proof.

For the implication (ii) $\Rightarrow$ (i) assume that there is $i \in \{1, \dots, k\}$, $\tilde{z} \in G^D_{opt}(\partial\Delta_k) \cap D_i$ and $\beta \in \{H^D_{opt}(\tilde{z}), e_i\}^\perp \cap \mathbb{R}^k_{\gneqq}$ such that $z = \tilde{z} + \beta$ holds. We need to show that $z \notin G^D_{opt}(\Delta_k)$. According to Theorem 4.8 it suffices to show that $z \notin \mathcal{Z}_{opt}$, i.e., that $P(z) \notin \mathcal{J}_{opt}$.
To this end we claim that $P(z) = P(\tilde{z}) + \beta$. Indeed, we have $P(\tilde{z}) + \beta \in J(U_{ad}) + [0, \infty)^k$ and plugging this into the optimality condition (1) for $P(z)$ we get

$$\langle P(\tilde{z}) + \beta - \tilde{z} - \beta, \tilde{y} - P(\tilde{z}) - \beta \rangle = \langle P(\tilde{z}) - \tilde{z}, \tilde{y} - P(\tilde{z}) - \beta \rangle = \langle P(\tilde{z}) - \tilde{z}, \tilde{y} - P(\tilde{z}) \rangle \geq 0,$$

for all $\tilde{y} \in J(U_{ad}) + [0, \infty)^k$, since $0 = \|P(\tilde{z}) - \tilde{z}\| \langle \beta, H^D_{opt}(\tilde{z}) \rangle = \langle \beta, P(\tilde{z}) - \tilde{z} \rangle$. As $P(z) \gneqq P(\tilde{z})$ we can conclude $P(z) \notin \mathcal{J}_{opt}$, so that $z \notin \mathcal{Z}_{opt}$. This completes the proof. $\square$