Ordering Structures in Vector Optimization and Applications in Medical Engineering

Gabriele Eichfelder[1]

January 2, 2013

Abstract

This manuscript is on the theory and numerical procedures of vector optimization w.r.t. various ordering structures, on recent developments in this area and, most important, on their application to medical engineering.

In vector optimization one considers optimization problems with a vector-valued objective map and thus one has to compare elements in a linear space. If the linear space is the finite dimensional space $\mathbb{R}^m$ this can be done componentwise. That corresponds to the notion of an Edgeworth-Pareto-optimal solution of a multiobjective optimization problem. Among the multitude of applications which can be modeled by such a multiobjective optimization problem, we present an application in intensity modulated radiation therapy and its solution by a numerical procedure.

In case the linear space is arbitrary, maybe infinite dimensional, one may introduce a partial ordering which defines how elements are compared. Such problems arise for instance in magnetic resonance tomography where the number of Hermitian matrices which have to be considered for a control of the maximum local specific absorption rate can be reduced by applying procedures from vector optimization. In addition to a short introduction and the application problem, we present a numerical solution method for solving such vector optimization problems.

A partial ordering can be represented by a convex cone which describes the set of directions in which one assumes that the current values are deteriorated. If one assumes that this set may vary dependently on the actually considered element in the linear space, one may replace the partial ordering by a variable ordering structure. This was for instance done in an application in medical image registration. We present a possibility of how to model such variable ordering structures mathematically and how optimality can be defined in such a case. We also give a numerical solution method for the case of a finite set of alternatives.

Key Words: Multiobjective optimization, vector optimization, partial ordering, variable ordering structure, intensity modulated radiation therapy, magnetic resonance tomography, local specific absorption rate.

Mathematics subject classifications (MSC 2000): 90C29, 90C30.

[1] Institute of Mathematics, Ilmenau University of Technology, Weimarer Str. 25, D-98693 Ilmenau, Germany, email: `Gabriele.Eichfelder@tu-ilmenau.de`

1 Introduction

In classical optimization one considers a scalar-valued objective function $f\colon X \to \mathbb{R}$ with X some real linear space and one searches for the minimal value of f over some nonempty set $S \subseteq X$, i.e. one aims on solving

$$\min_{x \in S} f(x). \tag{1}$$

An element $\bar{x} \in S$ is a minimal solution of (1) if $f(\bar{x}) \le f(x)$ for all $x \in S$. The unique minimal value of (1) is then $f(\bar{x})$.

However, many applications, plenty of those also in the area of medical engineering, require to minimize more than one objective function at the same time. If $m \geq 2$ objectives $f_i\colon X \to \mathbb{R}$, $i = 1, \ldots, m$ have to be minimized simultaneously one speaks of a *multiobjective* or *multicriteria optimization problem*

$$\min_{x \in S} \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \tag{MOP}$$

with a vector-valued objective function $f\colon X \to \mathbb{R}^m$ with $f(x) := (f_1(x), \ldots, f_m(x))$. Using the componentwise partial ordering, the values of this objective function f can be compared and it can be defined what an optimal solution is. We discuss such optimization problems together with an application to intensity modulated radiation therapy in Section 2.

Multiobjective optimization problems are a special case of *vector optimization problems*. There, one assumes to have a vector-valued objective map $f\colon X \to Y$ mapping in an arbitrary real linear space Y. The space Y may be partially ordered (for the definition of a partial ordering see Definition 3.1). Using this partial ordering, several optimality notions can be defined. In Section 3 we give an introduction to vector optimization in partially ordered spaces. The discussed concepts and a numerical solution procedure will be illustrated on an application in magnetic resonance tomography. There, a subset of a finite set of Hermitian matrices has to be determined for allowing a fast control of the maximum local specific absorption rate (SAR).

A partial ordering can be represented by a convex cone which is a set with a special structure (for the definition see p.12) and which describes the set of directions in which one assumes that the current values are deteriorated. If one assumes that this set may vary dependently on the actual element in the linear space, one may replace the partial ordering by a variable ordering structure. This was for instance necessary for being able to model an application problem in medical image registration. We present a possibility of how to mathematically formulate such variable ordering structures and how optimality can be defined in such a case together with a numerical solution procedure in Section 4.

This manuscript is not intended to be a comprehensive review of the theory of vector optimization and its application to problems in medical engineering, but to provide a survey on the different possibilities to model preferences in vector optimization and to present case studies in medical engineering which give examples for the different ordering structures.

2 Componentwise Ordering and an Application in Intensity Modulated Radiation Therapy

In the following, we assume that we have $m \geq 2$ objective functions $f_i \colon X \to \mathbb{R}$, $i = 1, \ldots, m$ which have to be minimized all at the same time over some nonempty subset $S \subseteq X$ of the linear space X. For instance, $X = \mathbb{R}^n$ and S may be given by inequality and equality constraints. Recall that maximizing some objective function f_i leads to the same optimal solutions (and the same absolute optimal value) as minimizing $-f_i$. Thus we can restrict ourselves to minimization in this section. We consider in the following the optimization problem as defined in (MOP) which is denoted a *multiobjective* or a *multicriteria optimization problem* with the vector-valued objective function $f \colon X \to \mathbb{R}^m$, $f(x) := (f_1(x), \ldots, f_m(x))$.

Applications which are modeled by such multiobjective optimization problems, next to the one discussed in Subsection 2.3, are for instance in chemotherapy control the maximization of the tumor cell killing while minimizing the toxicity and achieving a tolerable drug concentration [2] or in medical image registration the maximization of intensity similarity while minimizing the energy required to accomplish the transformation [1].

2.1 Multiobjective Optimization with the Componentwise Ordering

In the applied sciences F.Y. Edgeworth (1881, [13]) and V. Pareto (1906, [42]) were probably the first who introduced an optimality concept for multiobjective optimization problems. Therefore, optimal points are called Edgeworth-Pareto optimal points in the modern special literature, see [26].

Definition 2.1. A point $\bar{x} \in S$ is called an *Edgeworth-Pareto optimal (EP optimal) solution* of (MOP) if there exists no other $x \in S$ with

$$\begin{array}{rrcll} & f_i(x) & \leq & f_i(\bar{x}) & \text{for all } i = 1, \ldots, m, \\ \text{and} & f_j(x) & < & f_j(\bar{x}) & \text{for at least one } j \in \{1, \ldots, m\}\ . \end{array}$$

Hence, some point $\bar{x} \in S$ is EP optimal, if

$$\left(\{f(\bar{x})\} - \mathbb{R}^m_+\right) \cap f(S) = \{f(\bar{x})\}. \tag{2}$$

For Definition 2.1 the elements of the linear space $\mathbb{R}^m$ are compared componentwise. This ordering is also called the natural ordering: for all $a, b \in \mathbb{R}^m$

$$a \leq b \ :\Leftrightarrow \ a_i \leq b_i, \ i = 1, \ldots, m \ \Leftrightarrow \ b - a \in \mathbb{R}^m_+.$$

Then $\bar{x} \in S$ is an EP optimal solution of (MOP) if $\bar{y} := f(\bar{x})$ is a minimal element of the image set $f(S) := \{f(x) \in \mathbb{R}^m \mid x \in S\}$ in the sense of

$$y \leq \bar{y}, \quad y \in f(S) \ \Rightarrow \ y = \bar{y}.$$

For an illustration see Fig. 1.

Also weaker and stronger optimality notions are known in the literature, see for instance the books [14, 41] for a collection of notions.

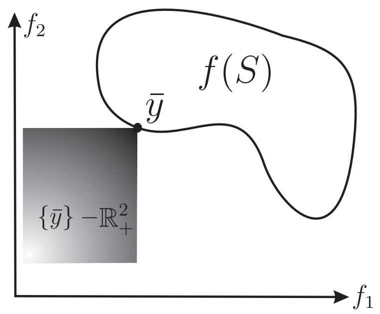

Figure 1: The point $\bar{y} := f(\bar{x})$ is a minimal element of $f(S) \subseteq \mathbb{R}^2$ and thus $\bar{x}$ is an EP-optimal solution of (MOP).

Definition 2.2. (i) A point $\bar{x} \in S$ is called a *weakly EP optimal solution* of (MOP), if there exists no other $x \in S$ with $f_i(x) < f_i(\bar{x})$ for all $i = 1, \ldots, m$.

(ii) A point $\bar{x} \in S$ is called a *strongly EP optimal solution* of (MOP), if $f_i(\bar{x}) \leq f_i(x)$ for all $i = 1, \ldots, m$ and all $x \in S$.

Thus $\bar{x}$ is weakly EP optimal if and only if

$$\big(\{f(\bar{x})\} - \mathrm{int}(\mathbb{R}^m_+)\big) \cap f(S) = \emptyset \ , \tag{3}$$

with $\mathrm{int}(\cdot)$ denoting the interior, and $\bar{x}$ is strongly EP optimal if and only if

$$f(S) \subseteq \{f(\bar{x})\} + \mathbb{R}^m_+ \ . \tag{4}$$

Of course, any strongly EP optimal solution is also EP optimal and any EP optimal solution is also weakly EP optimal. If there is a strongly EP optimal solution $\bar{x}$ then $\bar{x}$ simultaneously minimizes all objective functions f_i and hence the objectives are not concurrent. The weakly EP optimal solutions which are not also EP optimal are not desirable from the point of view of applications, as for these solutions still an improvement w.r.t. at least one objective function is possible without deteriorating the others. The notion of weakly EP optimal solutions is more of interest from a theoretical point of view, see for instance the comments in [35] and the following section on numerical procedures for solving (MOP).

Example 2.1. Let $f_i \colon \mathbb{R}^2 \to \mathbb{R}$ be defined by $f_i(x_1, x_2) = x_i$, $i = 1, 2$ and let $S = [1, 2] \times [1, 2]$. Then $f(S) = S$. The set $\{x \in S \mid x_1 = 1\} \cup \{x \in S \mid x_2 = 1\}$ equals the set of all weakly EP optimal solutions of $\min_{x \in S} f(x)$. The unique EP optimal solution is $x = (1, 1)$ which is at the same time a strongly EP optimal solution.

It is easy to proof that the image $f(\bar{x})$ of an EP optimal solution is always an element of the boundary $\partial f(S)$ of $f(S)$:

Lemma 2.1. *If $\bar{x} \in S$ is an EP optimal solution of (MOP), then $f(\bar{x}) \in \partial f(S)$.*

Hence, in numerical procedures, one can concentrate on determining elements of the boundary of $f(S)$.

2.2 Numerical Procedures

Many numerical procedures for solving multiobjective optimization problems presuming the componentwise ordering are proposed in the literature. Most of these approaches are based on the formulation of a parameter dependent scalar-valued optimization problem to the multiobjective optimization problem (MOP). Such a problem is for instance the weighted-sum scalarization

$$min_{x \in S}\ w_1 f_1(x) + \ldots + w_m f_m(x) \tag{5}$$

with nonnegative weights w_i, $i = 1, \ldots, m$. A survey over such scalarization approaches is for instance provided in [45] or [32]. Other approaches are based on evolutionary algorithm, for a survey see for instance [9, 10, 50, 56], or on the Newton's method [28].

In opposition to optimizing a scalar-valued optimization problem as the one in (1), the set of images of the optimal solutions of a multiobjective optimization problem, i.e.

$$\mathcal{E} := \{y = f(x) \in \mathbb{R}^m \mid x \text{ is an EP optimal solution of (MOP)}\},$$

is in general not a singleton, but infinitely many EP optimal solutions x and correspondent points $f(x) \in \mathbb{R}^m$ exist. Especially in engineering problems, one is in general not interested in one EP optimal solution only but in the complete set $\mathcal{E}$, which is also denoted the *efficient set*.

As this, in general infinite, set can in most cases not be determined, one aims on calculating an approximation of it. Such an approximation can be gained by solving a parameter dependent scalarization problem for a choice of parameters. Thereby it is in general the aim to approximate the complete set $\mathcal{E}$ concisely with almost equidistant approximation points and to avoid to neglect some parts of the set $\mathcal{E}$. Thus, an important question is how to choose the parameters appropriately in advance, as especially in application problems the solution of one scalarization problem may be very costly.

We present in the following a procedure for an adaptive parameter control using sensitivity information [18] and that is based on a scalarization known in the literature as ε-constrained method. This procedure was also used to solve the application problem in Section 2.3. The advantage of the ε-constraint scalarization compared with the weighted-sum scalarization mentioned above is that it is also applicable for non-convex problems. If the set $f(S) + \mathbb{R}^m_+$ is not convex, it might happen that even by varying the weights arbitrarily, not all EP optimal solutions might be found by solving the problem (5), see for instance [35] for a discussion.

The ε-constraint problem to the multiobjective optimization problem (MOP) is defined by

$$\begin{array}{lll} \min & f_m(x) & \\ \text{s.t.} & f_i(x) \leq \varepsilon_i, & i = 1, \ldots, m-1, \\ & x \in S & \end{array} \tag{6}$$

with parameter $\varepsilon \in \mathbb{R}^{m-1}$. Thus we minimize only one of the m objectives and convert the other objective functions into constraints. For this scalarization approach we have the following results. For proofs see e. g. [41, Ch. 3.2].

Theorem 2.1. *(a) If $\bar{x}$ is a solution of (6), then $\bar{x}$ is a weakly EP optimal solution of (MOP).*

(b) If $\bar{x}$ is a unique solution of (6), then $\bar{x}$ is an EP optimal solution of (MOP).

(c) If $\bar{x}$ is an EP optimal solution of (MOP), then it is also a solution of (6) with $\varepsilon_i := f_i(\bar{x})$, $i = 1, \ldots, m-1$.

For the case of only two objective functions we have the following stronger result [18]:

Lemma 2.2. *Let $m = 2$ and let $\bar{x}^1$ be a minimal solution of $\min_{x \in S} f_1(x)$ and $\bar{x}^2$ be a minimal solution of $\min_{x \in S} f_2(x)$. If $\bar{x}$ is an EP optimal solution of (MOP) then there exists a parameter $\varepsilon \in \mathbb{R}$ such that $\bar{x}$ is a minimal solution of (6) and $f_1(\bar{x}^1) \leq \varepsilon \leq f_1(\bar{x}^2)$.*

In case of three or more objective functions we cannot give such strong boundaries for the parameter ε, but we may reduce the parameter space to a compact set (for $f(S)$ compact), for instance in the case $m = 3$ by the following: Solve $\min_{x \in S} f_i(x)$ for $i = 1, 2$ with minimal solutions $x^{\min,i}$ and minimal values $f_i(x^{\min,i}) =: \varepsilon_i^{\min}$ as well as $\max_{x \in S} f_i(x)$ for $i = 1, 2$ with maximal solutions $x^{\max,i}$ and maximal values $f_i(x^{\max,i}) =: \varepsilon_i^{\max}$. Then for every EP optimal solution $\bar{x}$ of (MOP) there exists an $\varepsilon \in \mathbb{R}^2$ such that $\bar{x}$ is a minimal solution of (6) and $\varepsilon_i^{\min} \leq \varepsilon_i \leq \varepsilon_i^{\max}$, $i = 1, 2$.

For the remaining of this section we concentrate on the case $m = 2$, i.e. on the biobjective case. It still remains to clarify how to choose the parameters ε from the interval $[f_1(\bar{x}^1), f_1(\bar{x}^2)]$. We assume that the objective functions f_1 and f_2 are twice continuously differentiable. In addition to that, suppose that we have already solved the problem (6) for some parameter $\varepsilon^0 \in \mathbb{R}$ with $x^0 := x(\varepsilon^0)$ a minimal solution with Lagrange-multiplier $\mu^0 \geq 0$ to the constraint $f_1(x) - \varepsilon^0 \leq 0$ and that the point x^0 satisfies some first- and second-order optimality conditions and nondegeneracy is given. For more details on the assumptions needed to be satisfied for the following algorithm, we refer to [17, 18, 19]. The point x^0 is by Theorem 2.1 a weakly EP optimal solution and $f(x^0)$ serves as an approximation point of the efficient set $\mathcal{E}$. Next we want to find a parameter ε^1 with

$$\|f(x(\varepsilon^1)) - f(x^0)\| = \alpha \tag{7}$$

for a given value $\alpha > 0$. Throughout, let $x(\varepsilon)$ denote an optimal solution of the problem (6) for some $\varepsilon \in \mathbb{R}^{m-1}$. We suppose that the constraint $f_1(x) \leq \varepsilon^0$ is active in x^0, i.e. it is fulfilled with $f_1(x^0) = \varepsilon^0$. Otherwise we can easily find a parameter $\tilde{\varepsilon}^0$ with $f_1(x^0) = \tilde{\varepsilon}^0$. Under the above assumptions, the local minimal value function $\tau^\delta \colon \mathbb{R} \to \overline{\mathbb{R}}$,

$$\tau^\delta(\varepsilon) := \inf\{f_2(x) \mid f_1(x) \leq \varepsilon,\ x \in S,\ x \in B_\delta(x^0)\}\ ,$$

with $B_\delta(x^0)$ a closed ball around x^0 with radius δ for some small $\delta > 0$, is differentiable on a neighborhood of ε^0 with

$$(\tau^\delta)'(\varepsilon^0) = -\mu^0\ ,$$

see [17, Theorem 6], [18]. Using the derivative of the local minimal value function for a Taylor approximation (assuming this is possible) and assuming that the constraint $f_1(x) \leq \varepsilon$ remains active, we obtain

$$f_2(x(\varepsilon^1)) \approx f_2(x^0) - \mu^0(\varepsilon^1 - \varepsilon^0) \ .$$

As a consequence, the equation (7) is approximately satisfied for

$$\varepsilon^1 = \varepsilon^0 \pm \alpha \left(\sqrt{1 + (\mu^0)^2}\right)^{-1} \ .$$

This leads to the procedure summarized as Algorithm 1.

Algorithm 1 Approximation of the efficient set for $m = 2$

Require: distance $\alpha > 0$, starting distance $\beta \in (0, \alpha)$

1: solve $\min_{x \in S} f_2(x)$ with minimal solution x^1

2: set $\varepsilon^2 := f_1(x^1) - \beta$ and $l := 2$

3: solve $\min_{x \in S} f_1(x)$ with minimal solution x^E

4: **while** $\varepsilon^l \geq f_1(x^E)$ **do**

5: solve (6) for the parameter ε^l with minimal solution x^l and Lagrange-multiplier μ^l

6: set

$$\varepsilon^{l+1} := \varepsilon^l - \frac{\alpha}{\sqrt{1 + (\mu^l)^2}}$$

and $l := l + 1$

7: **end while**

8: **return** the set $A := \{f(x^1), \ldots, f(x^{l-1}), f(x^E)\}$ is an approximation of $\mathcal{E}$

A generalization of these results to the case $m \geq 3$ for generating locally equidistant points can be done easily, but for an equidistant approximation of the complete efficient set problems occur as discussed in [18]: as we have seen we cannot give as strong boundaries for the parameter ε as given in Lemma 2.2 for the case $m = 2$. An additional difficulty is that if we want to use sensitivity information to determine a new approximation point, we have to know which points are neighbors of the new point. We give the idea of a procedure for generating locally equidistant approximation points in the next paragraph.

Assume we have solved problem (6) for a parameter $\varepsilon^0 \in \mathbb{R}^{m-1}$ with minimal solution x^0 and Lagrange multiplier $\mu^0 \in \mathbb{R}^{m-1}$ to the constraints $f_i(x) - \varepsilon_i^0 \leq 0$, $i = 1, \ldots, m-1$ (assuming the constraints to be active in x^0) and we now want to find a new parameter $\varepsilon^1 \in \mathbb{R}^{m-1}$ with $\varepsilon^1 := \varepsilon^0 + s \cdot v$ for $s \in \mathbb{R}$ in a given direction $v \in \mathbb{R}^{m-1}$. Similarly as discussed above we approximatively get $\|f(x(\varepsilon^1)) - f(x^0)\| = \alpha$ for

$$s = \pm \frac{\alpha}{\sqrt{\|v\|^2 + (\mu^{0\top} v)^2}} \ . \tag{8}$$

One can use as directions v one or all of the $(m-1)$-dimensional unit vectors $e_1, \ldots, e_{m-1}$. For more details we refer again to [17, 18, 19].

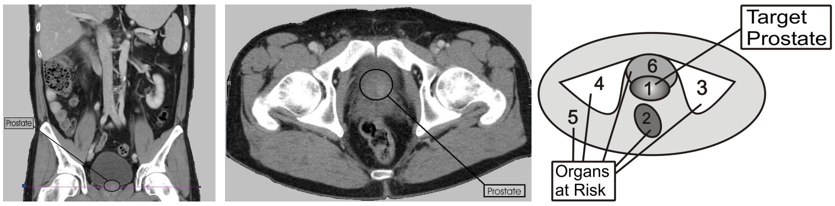

Figure 2: (a) and (b) Coronar and axial CT-cut (Figure courtesy of Dr. R. Janka, Institute of Diagnostic Radiology, Univ. Erlangen-Nuremberg). (c) Schematic axial body cut.

2.3 Application to Intensity Modulated Radiation Therapy

In this section, we present an optimization problem in intensity modulated radiation therapy and apply Algorithm 1 to it. The task is to find an irradiation plan for a patient with a cancer tumor to destroy the tumor while to spare the surrounding healthy organs. For a survey on the optimization problems arising in intensity modulated radiation therapy see [15]. The tasks which have to be considered in this context include the selection of beam angles (geometry problem), see for instance [16], the computation of an intensity map for each selected beam angle (intensity problem), and finding a sequence of configurations of a multileaf collimator to deliver the treatment (realization problem) [15].

We consider here the intensity problem which we model as a multiobjective optimization problem, see also for instance [37, 11, 12, 33]. The aim is to find an irradiation plan for a patient with prostate cancer. The tumor is irradiated with five equidistant beams which can be decompounded in 400 distinct controllable beamlets. We assume that the beam geometry is fixed. The relevant part of the patients body is mapped with the help of a computer tomography (CT) (see Fig. 2(a), (b)) and according to the thickness of the slices dissected into cubes, the so-called voxels.

With a clustering method [37], [46] where voxels with equal radiation exposure are collected, the very high number of 435 501 voxels can be reduced to 11 877 clusters $c_j,\ j = 1, \dots, 11\,877$. Then each cluster is allocated to one of the interesting volume structures $V_0, \dots, V_6$ by a physician. In our example these are the tumor (volumes $V_0,\ V_1$), the rectum (V_2), the left (V_3) and the right (V_4) hip-bone, the remaining surrounding tissue (V_5) and the bladder (V_6) (see Fig. 2(c)). Examinations have shown that the bladder and the rectum are opponents whereas the other critical organs follow these dominating organs in their stress caused by different irradiation plans. The emission by the beamlets B_i ($i \in \{1, \dots, 400\}$) in the clusters c_j ($j \in \{1, \dots, 11\,877\}$) at one radiation unit is described by the matrix $P = (P_{ji})_{j=1,\dots,11\,877, i=1,\dots,400}$. Let $x \in \mathbb{R}^{400}$ be the intensity profile. Then $P_j x$ with P_j the j-th row of the matrix P denotes the irradiation dose in the cluster c_j caused by the beamlets $B_i,\ i = 1, \dots, 400$.

For evaluating and comparing the radiation stress in the organs we use the concept of the equivalent uniform dose by Nimierko based on p-norms (here with

Table 1: Critical values for the organs at risk.

	number of organ (k)	p_k	U_k	Q_k	$N(V_k)$
rectum	2	3.0	30	36	6 459
left hip-bone	3	2.0	35	42	3 749
right hip-bone	4	2.0	35	42	4 177
remaining tissue	5	1.1	25	35	400 291
bladder	6	3.0	35	42	4 901

respect to the clustered voxels):

$$\mathrm{EUD}_k(x) = \frac{1}{U_k}\left(\frac{1}{N(V_k)}\sum_{\{j|c_j\in V_k\}} N(c_j)\cdot(P_jx)^{p_k}\right)^{\frac{1}{p_k}} - 1, \qquad k = 2,\ldots,6.$$

The scalar $p_k \in [1,\infty[$ is an organ depending constant reflecting the more parallel or more serial structure of the organ, $N(V_k)$ is the number of voxels in organ V_k and $N(c_j)$ is the number of voxels in cluster c_j, thus $\sum_{\{j|c_j\in V_k\}} N(c_j) = N(V_k)$. The value U_k is a dose limit for each organ which should not be exceeded and it is a statistical evaluated value which, in our example, can be taken from Table 1.

A feasible treatment plan has now to satisfy several constraints. First, a dangerous overdosing of the critical tissue should be avoided and thus, the maximal value Q_k must not be exceeded for all organs at risk V_k, $k = 2,\ldots,6$, i. e.

$$U_k(\mathrm{EUD}_k(x) + 1) \leq Q_k, \qquad k = 2,\ldots,6.$$

These restrictions can be restated as

$$\sum_{\{j|c_j\in V_k\}} N(c_j)(P_jx)^{p_k} \leq Q_k^{p_k}\cdot N(V_k) \qquad k = 2,\ldots,6.$$

It is also important that the dose in the tumor tissue remains below a maximal value to avoid injuries in the patients body and to achieve homogeneity of the irradiation. Besides, to have the desired effect of destroying all tumor cells, a certain curative dose has to be reached. Here, we differentiate between the so-called target-tissue V_0 and the boost-tissue V_1, which is tumor tissue that has to be irradiated especially high. Those conditions result in the following constraints for every cluster of the target and the boost volume:

$$\begin{array}{rcccll} & L_0(1-\varepsilon_0) & \leq & P_jx & \leq & L_0(1+\delta_0), & \forall j \text{ with } c_j \in V_0 \\ \text{and} & L_1(1-\varepsilon_1) & \leq & P_jx & \leq & L_1(1+\delta_1), & \forall j \text{ with } c_j \in V_1, \end{array} \qquad (9)$$

where L_0, L_1, ε_0, ε_1, δ_0 and δ_1 are constants given by the physician and tabulated in Table 2.

The target-tissue is pieced together by 8 593 clusters and the boost-tissue by 302 clusters which leads altogether to 17 790 additional constraints. Furthermore, it has to be assured that the intensity of the beams is nonnegative. Summarizing this we

Table 2: Critical values for the tumor tissues.

	number of organ (k)	L_k	δ_k	ε_k
target-tissue	0	67	0.11	0.11
boost-tissue	1	72	0.07	0.07

Table 3: Values of the approximation points and distances.

app.point	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$	$i=6$	$i=7$	$i=8$
$\text{EUD}_2(x^i)$	0.2000	0.1600	0.1203	0.0805	0.0421	0.0069	-0.0183	-0.0197
$\text{EUD}_6(x^i)$	0.0159	0.0164	0.0186	0.0283	0.0425	0.0782	0.1356	0.2000
δ^i	0.0400	0.0398	0.0410	0.0410	0.0501	0.0627	0.0644	-

have the following feasible set

$$\begin{aligned} S = \{x \in \mathbb{R}^{400}_+ \mid \quad & U_k(\text{EUD}_k(x)+1) \leq Q_k, & & k = 2,\ldots,6, \\ & L_0(1-\varepsilon_0) \leq P_j x \leq L_0(1+\delta_0), & & \forall j \text{ with } c_j \in V_0, \\ & L_1(1-\varepsilon_1) \leq P_j x \leq L_1(1+\delta_1), & & \forall j \text{ with } c_j \in V_1\} \end{aligned}$$

with 17 795 constraints and 400 variables. The aim is now to keep the dangerous overdosing of the organs at risk, the rectum (V_2) and the bladder (V_6), as low as possible, i. e. our two objectives are a minimization of the functions EUD_2 and EUD_6. Investigations [38] have shown, that these two organs are the dominating organs and that the other organs at risk follow in the level of their EUD-values these organs with a lower value.

Thus the bi-objective optimization problem can be written as

$$\min_{x \in S} \begin{pmatrix} f_1(x) \\ f_2(x) \end{pmatrix} = \min_{x \in S} \begin{pmatrix} \text{EUD}_2(x) \\ \text{EUD}_6(x) \end{pmatrix}.$$

As described in the preceding section we use the ε-constraint method as scalarization approach. This leads to the scalar-valued optimization problems

$$\begin{aligned} \min \quad & \text{EUD}_6(x) \\ \text{s. t.} \quad & \text{EUD}_2(x) \leq \varepsilon, \\ & x \in S \end{aligned}$$

with parameter $\varepsilon \in \mathbb{R}$. We first solve the problems $\min_{x \in S} f_i(x)$, $i = 1, 2$ and we get, according to Lemma 2.2, that it is sufficient to consider parameters $\varepsilon \in \mathbb{R}$ with $\varepsilon \in [-0.0197,\ 0.2000]$.

We apply Algorithm 1 with $\alpha = \beta = 0.04$. This results in the parameters $\varepsilon \in \{0.2000, 0.1600, 0.1203, 0.0805, 0.0421, 0.0069, -0.0183,\ -0.0197\}$ and the approximation shown in Fig. 3(a). The values of the approximation points and the distances $\delta^i := \|f(x^{i+1}) - f(x^i)\|_2$ between these points are tabulated in Table 3.

The physician can now choose one of the calculated, at least weakly, EP optimal solutions, can increase the fineness of the approximation by decreasing the value α and run the algorithm again, or can choose a point y determined by interpolation between existing approximation points and solve problem (6) again for $\varepsilon = y_1$, see also [49, p. 70].

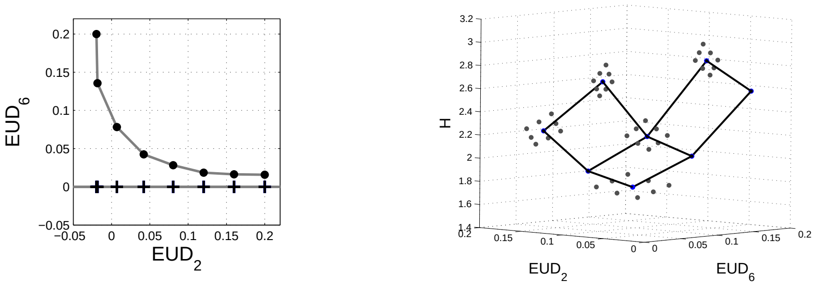

Figure 3: (a) Approximation of the efficient set of the biobjective problem for $\alpha = 0.04$, and the parameters $(\varepsilon, 0)$. (b) Approximation with locally equidistant approximation points of the efficient set of the multiobjective problem with three objective functions.

It is also of interest to include the additional target of homogeneity of the irradiation of the tumor in the problem formulation. This aim can be modeled by the objective function

$$H(x) := \sqrt{\frac{\sum\limits_{\{j|c_j \in V_0\}} N(c_j)\,(P_j x - L_0)^2 + \sum\limits_{\{j|c_j \in V_1\}} N(c_j)\,(P_j x - L_1)^2}{N(V_0) + N(V_1)}} \to \min!$$

Here, $N(V_0) = 13\,238$ and $N(V_1) = 2\,686$. The target of homogeneity competes with the previous objectives EUD_2 and EUD_6. Thus one may also investigate the following multiobjective optimization problem with three objective functions

$$\min_{x \in S} \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} = \min_{x \in S} \begin{pmatrix} \mathrm{EUD}_2(x) \\ \mathrm{EUD}_6(x) \\ H(x) \end{pmatrix} .$$

The corresponding scalarization approach is

$$\begin{aligned} \min \quad & H(x) \\ \text{s. t.} \quad & \mathrm{EUD}_2(x) \le \varepsilon_1, \\ & \mathrm{EUD}_6(x) \le \varepsilon_2, \\ & x \in S. \end{aligned}$$

The parameters $(\varepsilon_1, \varepsilon_2)$ can be determined according to (8). An approximation of the efficient set can be seen in Fig. 3(b). For more details we refer to [17].

3 Partial Orderings and an Application in Magnetic Resonance Tomography

As discussed in the introduction, multiobjective optimization problems, i.e. optimization problems with the objective function mapping in the real linear space $Y = \mathbb{R}^m$, can be seen as a special case of a vector optimization problem with an objective map mapping in an arbitrary real linear space Y. For comparing elements in

Y a binary relation has to be defined. Often, it is assumed that this binary relation is reflexive, transitive, and compatible with the linear structure of the space and thus a partial ordering. In this section we consider such vector optimization problems where the objective space is equipped with a partial ordering. Such optimization problems arise for instance in a data reduction problem in magnetic resonance tomography.

3.1 Vector Optimization with a Partial Ordering

First, we define the notion of a partial ordering.

Definition 3.1. Let Y be a real linear space.

(i) A nonempty subset R of the product space $Y \times Y$ is called a *binary relation* R on Y. We write yRz for $(y, z) \in R$.

(ii) A binary relation $\leq$ on Y is called a *partial ordering* on Y, if for arbitrary $w, x, y, z \in Y$
(reflexivity) $x \leq x$,
(transitivity) $x \leq y,\ y \leq z \ \Rightarrow\ x \leq z$,
$x \leq y,\ w \leq z \ \Rightarrow\ x + w \leq y + z$,
$x \leq y,\ \alpha \in \mathbb{R}_+ \ \Rightarrow\ \alpha x \leq \alpha y$.

(iii) A partial ordering $\leq$ on Y is called *antisymmetric*, if for arbitrary $y, z \in Y$

$$y \leq z,\ z \leq y \ \Rightarrow\ y = z.$$

A real linear space equipped with a partial ordering is called a *partially ordered linear space.* If $\leq$ is a partial ordering, then the set

$$K := \{y \in Y \mid y \geq 0_Y\}$$

is a convex cone. Recall that a set $K \subseteq Y$ is a *cone* if $\lambda y \in K$ for all $\lambda \geq 0$ and $y \in K$. And a cone is convex if $K + K \subseteq K$. Also, any convex cone $K \subseteq Y$ defines by

$$\leq_K := \{(y, z) \in Y \times Y \mid z - y \in K\}$$

a partial ordering on Y. Such a cone is then also called an *ordering cone.* A cone satisfying $K \cap (-K) = \{0_Y\}$ is called *pointed*, otherwise *non-pointed.* An ordering cone is pointed if and only if the associated partial ordering is antisymmetric.

For any element $y \in Y$, the set $(\{y\} + K) \setminus \{0_Y\} = \{z \in Y \mid z \geq y,\ z \neq y\}$ is the set of elements, which are considered to be worse than y, while the set $(\{y\} - K) \setminus \{0_Y\} = \{z \in Y \mid y \geq z,\ z \neq y\}$ describes the set of elements which are preferred to y. If $Y = \mathbb{R}^m$ and $K = \mathbb{R}^m_+$, then K defines the componentwise (natural) ordering in $\mathbb{R}^m$ as used in Section 2. But also other convex cones can be used to define a partial ordering in $\mathbb{R}^m$ as for instance the lexicographic ordering cone

$$K_{\text{lex}} := \{y \in \mathbb{R}^m \mid y_1 = \ldots = y_k = 0,\ y_{k+1} > 0 \text{ for some } k \in \{0, \ldots, m-1\}\} \cup \{0_{\mathbb{R}^m}\}$$

or the ice-cream cone, also known as Lorentz cone,

$$K_L := \{y \in \mathbb{R}^m \mid \|(y_1, \ldots, y_{m-1})\|_2 \leq y_m\}.$$

The vector optimization problems, which we are considering in the following, are

$$\begin{array}{c} \min_K \ f(x) \\ \text{such that} \\ x \in S \end{array} \tag{VOP}$$

with real linear spaces X, Y, a nonempty subset $S \subseteq X$, a vector-valued map $f\colon X \to Y$, and the linear space Y partially ordered by $\leq_K$ with $K \subseteq Y$ a pointed nontrivial (i.e. $K \neq \{0_Y\}$, $K \neq Y$) convex cone. The index K after min determines the convex cone and thus the partial ordering w.r.t. which the elements in Y are compared. In Section 2.1 we have seen that for deciding whether some $x \in S$ is an EP-optimal solution of (MOP) only the points $f(x)$ with $x \in S$, i.e. of the set $f(S)$, have to be compared. We thus first define what a (weakly, strongly) efficient element of a set w.r.t. a partial ordering is.

Definition 3.2. Let Y be partially ordered by some pointed convex cone $K \subseteq Y$ and let A be a nonempty subset of Y.

(i) An element $\bar{y} \in A$ is an *efficient element* of the set A if

$$(\{\bar{y}\} - K) \cap A = \{\bar{y}\}\ . \tag{10}$$

(ii) An element $\bar{y} \in A$ is a *strongly efficient element* of the set A if

$$A \subseteq \{\bar{y}\} + K\ . \tag{11}$$

(iii) Additionally, let Y be a topological space and the interior of the cone K, $\operatorname{int}(K)$, be nonempty. An element $\bar{y} \in A$ is a *weakly efficient element* of the set A if

$$(\{\bar{y}\} - \operatorname{int}(K)) \cap A = \emptyset\ . \tag{12}$$

In the following, if we speak of weak notions which are defined based on the interior of the cone K, we always assume the linear space Y to be a topological space and the cone K to have a nonempty interior. Note, that similar weak notions can also be defined in a linear space based on the algebraic interior of K, see for instance [35].

The notions "efficient" and "weakly efficient" are closely related. To see that, take an arbitrary weakly minimal element $\bar{y} \in A$ of the set A, i.e. $(\{\bar{y}\} - \operatorname{int}(K)) \cap A = \emptyset$. The set $\hat{K} := \operatorname{int}(K) \cup \{0_Y\}$ is a convex cone and it induces another partial ordering in Y. Consequently, $\bar{y}$ is also a minimal element of the set A with respect to the partial ordering induced by $\hat{K}$. In terms of lattice theory a strongly efficient element of a set A is also called *zero* element of A. It is a lower bound of the considered set. As this notion is very restrictive it is often not applicable in practice.

Definition 3.3. Let the vector optimization problem (VOP) be given. An element $\bar{x} \in S$ is a (weakly/stronlg) *efficient solution* of (VOP) if $f(\bar{x})$ is a (weakly/stronlg) efficient element of the set $f(S)$.

Of course, for $Y = \mathbb{R}^m$ and $K = \mathbb{R}^m_+$ the notion of (weakly/strongly) EP optimal solutions and (weakly/stronlgy) efficient solutions coincide.

In the following we collect some basic results on efficient elements (and thus on efficient solutions) of a vector optimization problem. For proofs we refer to [35, 26]. The first result relates the different optimality notions.

Lemma 3.1. *(a) Every strongly efficient element of the set A is also an efficient element of A.*

(b) Every efficient element of the set A is also a weakly efficient element of the set A.

Efficient elements of some set A are an element of the boundary ∂A of the set A, compare for $Y = \mathbb{R}^m$ and $K = \mathbb{R}^m_+$ with Lemma 2.1:

Lemma 3.2. *If $\bar{y} \in A$ is a weakly efficient element of A, then $\bar{y} \in \partial A$.*

In (5) we have stated the weighted-sum scalarization for multiobjective optimization problems with the componentwise ordering. This scalarization can be generalized e.g. to real topological linear spaces with a partial ordering. For that we need the definition of the dual cone. The dual cone $K^* \subseteq Y^*$, with Y^* the linear space of all continuous linear functionals $y^* \colon Y \to \mathbb{R}$, to some convex cone K is defined by

$$K^* := \{y^* \in Y^* \mid y^*(y) \geq 0 \ \text{ for all } y \in K\} \ .$$

The set

$$K^\# := \{y^* \in Y^* \mid y^*(y) > 0 \ \text{ for all } y \in K \setminus \{0_Y\}\}$$

is denoted the quasi interior of the dual cone. For $Y = \mathbb{R}^m$ the definitions read as $K^* = \{w \in \mathbb{R}^m \mid w^\top y \geq 0 \ \text{ for all } y \in K\}$ and $K^\# = \{w \in \mathbb{R}^m \mid w^\top y > 0 \ \text{ for all } y \in K \setminus \{0_{\mathbb{R}^m}\}\}$. For $K = \mathbb{R}^m_+$ we obtain

$$(\mathbb{R}^m_+)^* = \mathbb{R}^m_+ \ \text{ and } \ (\mathbb{R}^m_+)^\# = \text{int}(\mathbb{R}^m_+) \ .$$

We first collect sufficient conditions [35].

Theorem 3.1. *(a) If there is some $l \in K^*$ such that $\bar{x} \in S$ is a unique minimal solution of*

$$\min_{x \in S} l(f(x)) \ , \tag{13}$$

then $\bar{x}$ is an efficient solution of (VOP).

(b) If there is some $l \in K^\#$ such that $\bar{x} \in S$ is a minimal solution of (13), then $\bar{x}$ is an efficient solution of (VOP).

(c) If there is some $l \in K^* \setminus \{0_{Y^*}\}$ such that $\bar{x} \in S$ is a minimal solution of (13), then $\bar{x}$ is a weakly efficient solution of (VOP).

For obtaining necessary results, we need convexity assumptions.

Theorem 3.2. *Let the set $S + K$ be convex. If $\bar{x} \in S$ is a weakly efficient solution of (VOP), then there is some $l \in K^* \setminus \{0_{Y^*}\}$ such that $\bar{x} \in S$ is a minimal solution of (13).*

As every efficient solution is also a weakly efficient solution, this theorem delivers also a necessary condition for efficient solutions of (VOP).

For avoiding the need of a convex set $f(S)+K$, nonlinear scalarization functionals can be used. Allowing two parameters $a \in Y$ and $r \in Y \setminus \{0_Y\}$, one can consider the following nonlinear scalarization function $\psi_{a,r} \colon Y \to \mathbb{R} \cup \{\pm\infty\}$,

$$\psi_{a,r}(y) := \inf\{t \in \mathbb{R} \mid a + t\,r - y \in K\} \text{ for all } y \in Y . \tag{14}$$

This function was used as separational functional by Gerstewitz (Tammer) [29], see also [30], and is denoted smallest monotone map in [40]. It was used in vector optimization by Pascoletti and Serafini [43] and was already studied by Rubinov [44]. Its properties are well studied, see for instance [31, Theorem 2.3.1, Corollary 2.3.5], [8, Prop. 2.1] and [47]. For an illustration see Fig. 4.

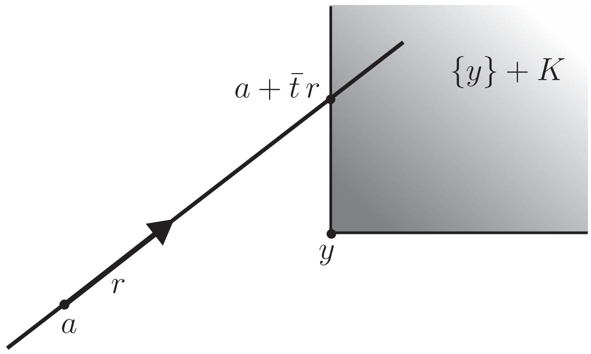

Figure 4: Illustration of the function $\psi_{a,r}$ with $\bar{t} := \psi_{a,r}(y)$.

Any minimal solution of

$$\min_{x \in S} \psi_{a,r}(f(x))$$

is an at least weakly efficient solution of (VOP). Any weakly efficient solution $\bar{x}$ of (VOP) is a minimal solution of this scalar problem if for instance $a = f(\bar{x})$ and $r \in \operatorname{int}(K)$. For more results on this scalarization we refer to [18]. Note that for $Y = \mathbb{R}^m$, $K = \mathbb{R}^m_+$, $a = (\varepsilon_1, \ldots, \varepsilon_{m-1}, 0)^\top$ and $r = (0, \ldots, 0, 1)^\top$, $\min_{x \in S} \psi_{a,r}(f(x))$ is equivalent to the ε-constraint problem discussed in (6).

3.2 Numerical Procedures

In case of a finite set $A = f(S)$ of the vector optimization problem the most simple approach for determining all optimal solutions is a pairwise comparison of all elements in A. This may be very time consuming, especially if the evaluation of the binary relation $\leq$ is costly. For that reason numerical methods as the Jahn-Graef-Younes method have been developed for reducing the numerical effort by reducing the number of necessary pairwise comparisons. For $\mathbb{R}^m$ partially ordered by the natural ordering, i.e. $K = \mathbb{R}^m_+$, this procedure was given by Jahn in [36], see also [35, Section 12.4], based on a procedure firstly presented by Younes in [53] and an algorithmic conception by Graef [35, p. 349]. In the following we present this algorithm for arbitrary linear spaces Y with some ordering cone $K \subseteq Y$.

Algorithm 2 Jahn-Graef-Younes method in partially ordered spaces

Require: $A = \{y^1, \ldots, y^k\}$, $K \subseteq Y$

put $U = \{y^1\}$ and $i = 1$
while $i < k$ **do**
    replace i by $i + 1$
    if $y^i \notin \{u\} + K$ for all $u \in U$ **then**
        replace U by $U \cup \{y^i\}$
    end if
end while
put $\{u^1, \ldots, u^p\} = U$
put $T = \{u^p\}$ and $i = p$
while $i > 1$ **do**
    replace i by $i - 1$
    if $u^i \notin \{t\} + K$ for all $t \in T$ **then**
        replace T by $\{u^i\} \cup T$
    end if
end while
return the set T is the set of efficient elements of A

Theorem 3.3. *Let A be a finite subset of Y and let U and T denote the sets gained by Algorithm 2.*

(a) If $\bar{y}$ is an efficient element of A, then $\bar{y} \in U$ and $\bar{y} \in T$.

(b) The set T is exactly the set of all efficient elements of A.

Proof. The proof for $Y = \mathbb{R}^m$ and $K = \mathbb{R}^m_+$ can be found in [35]. (a) Assume $\bar{y}$ is an efficient element of A but is not in U. Then there exists some $y \in U \subseteq A$, $y \neq \bar{y}$, with $\bar{y} \in \{y\} + K$ in contradiction to $\bar{y}$ an efficient element of the set A. Next, assume $\bar{y}$ is an efficient element of A but is not in T. According to the first part of the proof, $\bar{y} \in U$. Thus there exists some $y \in T \subseteq A$, $y \neq \bar{y}$, with $\bar{y} \in \{y\} + K$ in contradiction to $\bar{y}$ an efficient element of the set A.

(b) Let $T =: \{t^1, \ldots, t^q\}$ with $q \leq p \leq k$ and $t^j \in T$ be arbitrarily chosen with $1 \leq j \leq q$. We assume the elements of the sets to be ordered in the way they are generated in the algorithm. According to the first while-loop, $t^j \notin \{t^i\} + K$ for all i with $1 \leq i < j$ and according to the second while-loop, $t^j \notin \{t^i\} + K$ for all i with $j < i \leq q$. Hence, t^j is an efficient element of T.

According to (a), it remains to be shown that the elements of T are all also efficient elements of A. Let $y \in T$ and y be not an efficient element of A. Then there exists an efficient element $\bar{y}$ of A with $y \in \{\bar{y}\} + K \setminus \{0_Y\}$. According to (a), $\bar{y} \in T$ in contradiction to y an efficient element of T. □

In case of a non-finite set $f(S)$ for instance the scalarization (14) can be used for a numerical solution method using a procedure for the choice of the parameter a, while $r \in K \setminus \{0_Y\}$ can be chosen as constant [18].

3.3 Application to Magnetic Resonance Tomography

Parallel transmission (pTx) in magnetic resonance tomography uses multiple excitation coils driven by independent RF pulse waveforms. For the application of such systems, the management of local and global power deposition measured as specific absorption rate, SAR, in human subjects is a fundamental constraint [39]. So it is necessary to provide methods to enforces the satisfaction of local SAR constraints [34].

Based on a precalculated electrical field vector distribution from a unit voltage for each transmission channel, the local SAR in a region ν can be calculated by

$$\mathrm{SAR}(S_\nu) = \int_{\Delta t} U^H(t) \cdot S_\nu \cdot U(t)\mathrm{d}t$$

with $U_k(t)$ for all $t \in \Delta t$ the complex-valued waveform of the transmit channel k at the time t in some time interval Δt and S_ν an Hermitian $n \times n$ matrix. Typical values for n are $n = 2, 4, 8, 16$. In [23], $n = 8$ was assumed. The number m of subvolumes $\nu = 1, \ldots, m$ which have to be taken into account vary around 300 000 up to over 1 000 000.

The goal is to replace the precalculated SAR model (i.e. the set of matrices S_ν) by a smaller set of so-called virtual observation points A_j, $j = \{1, \ldots, N\}$ such that for all $\nu \in \{1, \ldots, m\}$ there exists some $j \in \{1, \ldots, N\}$ satisfying

$$\int_{\Delta t} U^H(t) \cdot S_\nu \cdot U(t)\mathrm{d}t \leq \int_{\Delta t} U^H(t) \cdot A_j \cdot U(t)\mathrm{d}t\ . \tag{15}$$

Then

$$\max_{\nu=1,\ldots,m} \int_{\Delta t} U^H(t) \cdot S_\nu \cdot U(t)\mathrm{d}t \leq \max_{j=1,\ldots,N} \int_{\Delta t} U^H(t) \cdot A_j \cdot U(t)\mathrm{d}t\ .$$

If $\{A_j \mid j = 1, \ldots, N\} \subseteq \{S_\nu \mid \nu = 1, \ldots, m\}$, then equality holds. Thereby, (15) is satisfied for arbitrary $U \colon \mathbb{R} \to \mathbb{C}^n$, if for all $\nu \in \{1, \ldots, m\}$ there exists some $j \in \{1, \ldots, N\}$ such that

$$A_j - S_\nu \text{ is positive semidefinite.}$$

Hence, among the set $M := \{S_1, \ldots, S_m\}$ one can determine the set of efficient elements $\mathcal{E}$ w.r.t. $K := -\mathcal{S}^n_+$ with $\mathcal{S}^n_+$ the cone of positive semidefinite Hermitian matrices. Then, if $S \in M \setminus \mathcal{E}$, there exists some matrix $A \in \mathcal{E}$ such that $S \in \{A\} + K = \{A\} - \mathcal{S}^n_+$, i.e. such that $A - S$ is positive semidefinite. For determining the efficient elements of the finite set M, in [23] Algorithm 2 was proposed.

However, it turned out that almost all matrices in M also belong to $\mathcal{E}$. Examinations with randomly generated 8×8 matrices show, that this happens also for such sets of matrices, but there are some few matrices which are not efficient. Therefore, in [23, 24] another approach was proposed to determine a smaller set of virtual observation points. New matrices $A_1, \ldots, A_N$, which are not an element of M, are determined such that these new matrices are exactly the efficient matrices of the enlarged set

$$\tilde{M} := \{S_1, \ldots, S_m, A_1, \ldots, A_N\}.$$

Evaluating then $\max_{j=1,\ldots,N} \int_{\Delta t} U^H(t) \cdot A_j \cdot U(t)\mathrm{d}t$ leads to an overestimation of the maximum local SAR. So the determination of the new matrices $A_1, \ldots, A_N$

has to be performed in such a way, that this overestimation is as small as possible using at the same time as few matrices (number N) as possible. This is again a multiobjective optimization problem.

For this biobjective optimization problem, an approach as given in (6) was chosen, known as ε-constraint method: in [23] an upper bound on the overestimation was chosen while to minimize the number N. The proposed procedure is given in Algorithm 3.

Algorithm 3 Data reduction by extending the efficient set

```
Require: M := {S_1, ..., S_m}, u > 0
put k = 1, N = 1
while M ≠ ∅ do
    choose S^{k*} ∈ argmax{‖S_j‖ | S_j ∈ M}
    sort all matrices S_j ∈ M w.r.t. λ_min(S^{k*} − S_j) in decreasing order, i.e.
    S^{k*} = S_1, ..., S_{nk} (nk ∈ ℕ, nk := |M|) with

        λ_min(S^{k*} − S_1) = 0 ≥ ... ≥ λ_min(S^{k*} − S_{nk})

    set l := 1, Z̄ := 0_{ℂ^{n×n}}, ε^k := 0 and C^k := ∅
    while ‖Z̄‖ ≤ u and l ≤ nk do
        set ε^k := −λ_min(S^{k*} − S_l), Z_k := Z̄, C^k := C^k ∪ {S_l} and l := l + 1
        if l ≤ nk then
            determine

                Z̄ ∈ argmin{‖Z‖ | S^{k*} + Z − S_j ∈ 𝒮^n_+ ∀S_j ∈ C^k ∪ {S_l}, Z ∈ 𝒮^n_+}

        end if
    end while
    set M := M \ C^k, N := k and k := k + 1
end while
return clusters C^k, matrices A_k := S^{k*} + Z_k, k = 1, ..., N
```

There, the matrices $S_1, \ldots, S_m$ are clustered based on the following similarity criteria: S is similar to a core matrix S^* for some given $\varepsilon \geq 0$ if

$$\lambda_{\min}(S^* - S) \geq -\varepsilon$$

with $\lambda_{\min}$ denoting the smallest eigenvalue. Then for arbitrary $U\colon \mathbb{R} \to \mathbb{C}^n$ the SAR-value of S is bounded from above by

$$\int_{\Delta t} U^H(t) \cdot S \cdot U(t)\mathrm{d}t \leq \int_{\Delta t} U^H(t) \cdot S^* \cdot U(t)\mathrm{d}t + \varepsilon \int_{\Delta t} \|U(t)\|^2\mathrm{d}t.$$

The clusters are generated iteratively: a core matrix S^* is chosen with largest norm among the remaining matrices. Then similar matrices are added to the cluster enlarging the similarity $\varepsilon \geq 0$ as long as the overestimation which results from a new-to-define matrix A for each cluster (see the next paragraphs) is bounded within a predefined bound u. For an illustration of the clusters see Fig. 5.

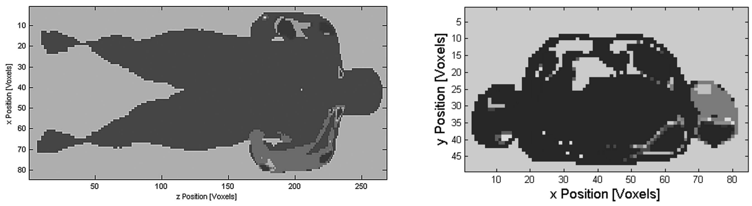

Figure 5: Clustering visualized in the human body with different gray scales. (a) coronal plane (b) transversal plane with $N = 62$.

For all matrices in a cluster around a core matrix S^*, a new matrix A is determined by solving

$$\begin{aligned} \min\{\|Z\| \mid \quad & S^* + Z - S_j \in \mathcal{S}^n_+ \ \forall S_j \in M \text{ with } \lambda_{\min}(S^* - S_j) \geq -\varepsilon, \\ & Z \text{ a Hermitian } n \times n \text{ matrix}\}\ , \end{aligned} \tag{16}$$

with $\|\cdot\|$ some matrix norm for instance consistent with the Euclidean norm, with minimal solution $\bar{Z}$. Then $A := S^* + \bar{Z}$ is the only efficient matrix in the set

$$\{S_j \in M \mid \lambda_{\min}(S^* - S_j) \geq -\varepsilon\} \cup \{A\}$$

and it holds

$$\int_{\Delta t} U^H(t) \cdot A \cdot U(t)\mathrm{d}t - \int_{\Delta t} U^H(t) \cdot S^* \cdot U(t)\mathrm{d}t \leq \|\bar{Z}\| \int_{\Delta t} \|U(t)\|^2 \mathrm{d}t.$$

Algorithm 3 determines $\varepsilon \geq 0$ for each cluster in such a way, that $\|\bar{Z}\| \leq u$ for some predefined $u > 0$ for all generated clusters, i.e. such that the overestimation is bounded.

Different choices of u result in different sizes of clusters and hence in different numbers N of clusters and thus of new matrices

$$A_k := S^{*k} + Z_k,\ k = 1, \ldots, N\ ,$$

with S^{*k} the core matrix and Z_k the minimal solution of (16), see Fig. 6. The matrices A_k, $k = 1, \ldots, N$ serve now as virtual observation points which satisfy (15).

For different human models and different landmark positions, resolution compression factors between 1 126 and 13 109 where reached. For instance 241 032 matrices where replaced by 214 virtual observation points and 891 418 matrices by 68.

4 Variable Ordering Structures and an Application in Medical Image Registration

In vector optimization one assumes in general, as we have seen in the previous subsections, that a partial ordering is given by some nontrivial convex cone K in the considered space Y. But already in 1974 in one of the first publications [54] related to the definition of optimal elements in vector optimization also the idea of variable

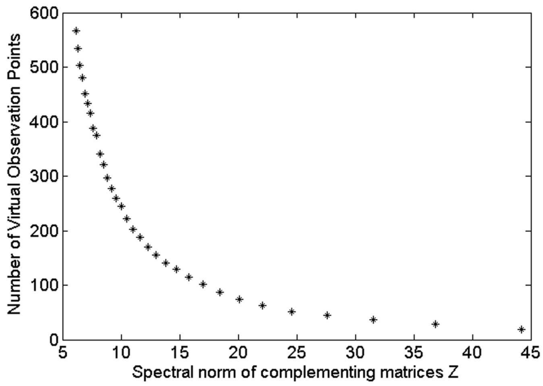

Figure 6: Number of virtual observation points N for different chosen upper limits u for the allowed maximum overestimation of the local SAR, i.e. the spectral norm of the complementing matrices $\|Z^k\|$.

ordering structures was given: to each element of the space a cone of dominated (or preferred) directions is defined and thus the ordering structure is given by a set-valued map. In [54] a candidate element was defined to be *nondominated* if it is not dominated by any other reference element w.r.t. the corresponding cone of this other element. Later, also another notion of optimal elements in the case of a variable ordering structure was introduced [5, 6, 7]: a candidate element is called a *minimal* (or *nondominated-like*) element if it is not dominated by any other reference element w.r.t. the cone of the candidate element.

Recently, there is an increasing interest in such variable ordering structures motivated by several applications for instance in medical image registration [20], see the next subsection, or in portfolio optimization [21, 3].

4.1 Application in Medical Image Registration

For modeling preferences of a totally rational decision maker in medical image registration, it turned out that a variable ordering structure better reflects the problem structure [51]. In medical image registration it is the aim to merge several medical images gained by different imaging methods as for instance computer tomography, magnetic resonance tomography, positron emission tomography, or ultrasound. For two data sets A and B a transformation map t, also called registration, has to be found (from a set T of allowed maps) such that some similarity measure comparing $t(A)$ and B is optimized. For some applications it is important that this transformation map is found automatically without a human decision maker. The quality of a transformation map, i.e. the similarity of the transformed data set to the target set, can be measured by a large variety of distance measures $f_i\colon (t, A, B) \to \mathbb{R}$, $i = 1, \ldots, m$ ($m \in \mathbb{N}$). They all evaluate distinct characteristics like the sum of square differences, mutual information or cross-correlation. Different measures may lead to different optimal transformation maps. Some measures fail on special data sets and can lead to mathematical correct but useless results. Thus it is important to combine several measures. Possible approaches are a weighted sum of different measures. But difficulties appear as badly scaled functions or non-convex functions.

Instead, the problem can be viewed as a multiobjective optimization problem [51, 52] by arranging the several distance measures in an objective vector $f := (f_1, \dots, f_m)^\top$. Then, for given data sets A and B, the vector optimization problem

$$\min_{t \in T} f(t, A, B)$$

has to be solved. For incorporating in the preference structure that some of the measures may fail on the given data sets, depending on the values $y \in \mathbb{R}^m$ in the objective space a weighting vector $w(y) \in \mathbb{R}^m_+$ is generated. This weight can be interpreted as some kind of voting between the different measures. Also a weight component equal to zero is allowed which corresponds to the negligence of the correspondent measure, because it seems for instance to fail on the data set. This weight may also depend on gradient information, conformity and continuity aspects and reflects therefore the preferences of a totally rational decision maker who puts a higher weight on promising measures dependent on the value $y = f(t, A, B)$.

To such a weight at a point $y \in \mathbb{R}^m$ a cone of more or equally preferred directions is defined by

$$\mathcal{P}(y) := \left\{ d \in \mathbb{R}^m \;\middle|\; \sum_{i=1}^{m} \operatorname{sgn}(d_i) w_i(y) \leq 0 \right\}$$

where

$$\operatorname{sgn}(d_i) := \begin{cases} 1 & \text{if } d_i > 0, \\ 0 & \text{if } d_i = 0, \\ -1 & \text{if } d_i < 0. \end{cases}$$

Then y is considered to be better than $\bar{y}$ if $y \in \{\bar{y}\} + \mathcal{P}(\bar{y})$. Note that for nonnegative weights $w(y) \in \mathbb{R}^m_+$ it holds $\mathbb{R}^m_+ \subseteq \mathcal{D}(y) := -\mathcal{P}(y)$ for all y.

For this special problem formulation Wacker proposed a solution procedure in [51].

4.2 Vector Optimization with Variable Ordering Structures

For a study of vector optimization problems with a variable ordering structure it is important to differentiate between the two optimality concepts mentioned in the introduction to this section as well as to examine the relation between the concepts. In view of applications it is important to formulate characterizations of optimal elements, for instance by scalarizations, for allowing numerical calculations.

In the following we assume Y to be a real topological linear space and A to be a nonempty subset of Y. Let $\mathcal{D}\colon Y \rightrightarrows Y$ be a set-valued map with $\mathcal{D}(y)$ a pointed convex cone for all $y \in Y$ and let $\mathcal{D}(A) := \bigcup_{y \in A} \mathcal{D}(y)$ denote the image of A under $\mathcal{D}$.

Based on the cone-valued map $\mathcal{D}$ one can define two different relations: for $y, \bar{y} \in Y$ we define

$$y \leq_1 \bar{y} \;\text{ if }\; \bar{y} \in \{y\} + \mathcal{D}(y) \tag{17}$$

and

$$y \leq_2 \bar{y} \;\text{ if }\; \bar{y} \in \{y\} + \mathcal{D}(\bar{y}) \;. \tag{18}$$

We speak here of a variable ordering (structure), given by the ordering map $\mathcal{D}$, despite the binary relations given above are in general not transitive nor compatible with the linear structure of the space, to express that the partial ordering given by a cone in the previous sections is replaced by a relation defined by $\mathcal{D}$.

Relation (17) implies the concept of nondominated elements originally defined in [54, 55]. We also state the definitions of weakly and strongly nondominated elements which can easily be derived from the original definition of nondominated elements.

Definition 4.1. (a) An element $\bar{y} \in A$ is a *nondominated element* of A w.r.t. the ordering map $\mathcal{D}$ if there is no $y \in A \setminus \{\bar{y}\}$ such that $\bar{y} \in \{y\} + \mathcal{D}(y)$, i.e., $y \not\leq_1 \bar{y}$ for all $y \in A \setminus \{\bar{y}\}$.

(b) An element $\bar{y} \in A$ is a *strongly nondominated* element of A w.r.t. the ordering map $\mathcal{D}$ if $\bar{y} \in \{y\} - \mathcal{D}(y)$ for all $y \in A$.

(c) Let $\mathcal{D}(y)$ have a nonempty interior, i.e. $\operatorname{int}(\mathcal{D}(y)) \neq \emptyset$, for all $y \in A$. An element $\bar{y} \in A$ is a *weakly nondominated* element of A w.r.t. the ordering map $\mathcal{D}$ if there is no $y \in A$ such that $\bar{y} \in \{y\} + \operatorname{int}(\mathcal{D}(y))$.

Example 4.1. Let $Y = \mathbb{R}^2$, the cone-valued map $\mathcal{D}\colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be defined by

$$\mathcal{D}(y_1, y_2) := \begin{cases} \text{cone conv}\{(y_1, y_2), (1, 0)\} & \text{if } (y_1, y_2) \in \mathbb{R}^2_+,\ y_2 \neq 0, \\ \mathbb{R}^2_+ & \text{otherwise,} \end{cases}$$

and

$$A := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \geq 0,\ y_2 \geq 0,\ y_2 \geq 1 - y_1\}\ .$$

Here cone and conv denote the conic hull and the convex hull, respectively. Then $\mathcal{D}(y_1, y_2) \subseteq \mathbb{R}^2_+$ for all $(y_1, y_2) \in \mathbb{R}^2$ and one can check that $\{(y_1, y_2) \in A \mid y_1 + y_2 = 1\}$ is the set of all nondominated elements of A w.r.t. $\mathcal{D}$ and that all elements of the set $\{(y_1, y_2) \in A \mid y_1 + y_2 = 1\ \vee y_1 = 0\ \vee\ y_2 = 0\}$ are weakly nondominated elements of A w.r.t. $\mathcal{D}$.

In Definition 4.1 the cone $\mathcal{D}(y) = \{d \in Y \mid y + d \text{ is dominated by } y\} \cup \{0_Y\}$ can be seen as the set of dominated directions for each element $y \in Y$. Note that when $\mathcal{D}(y) \equiv K$, where K is a pointed convex cone, and the space Y is partially ordered by K, the concepts of nondominated, strongly nondominated and weakly nondominated elements w.r.t. the ordering map $\mathcal{D}$ reduce to the classical concepts of efficient, strongly efficient and weakly efficient elements w.r.t. the cone K, compare Definition 3.2. Strongly nondominated is a stronger concept than nondominatedness, as it is not only demanded that $\bar{y} \in \{y\} + (Y \setminus \{\mathcal{D}(y)\})$ for all $y \in A \setminus \{\bar{y}\}$, but even $\bar{y} \in \{y\} - \mathcal{D}(y)$ for all $y \in A \setminus \{\bar{y}\}$ for $\bar{y}$ being strongly nondominated w.r.t. $\mathcal{D}$. This can be interpreted as the requirement of being far away from being dominated.

The second relation, relation (18), leads to the concept of minimal, also called nondominated-like, elements [5, 6, 7].

Definition 4.2. (a) An element $\bar{y} \in A$ is a *minimal element* of A w.r.t. the ordering map $\mathcal{D}$ if there is no $y \in A \setminus \{\bar{y}\}$ such that $\bar{y} \in \{y\} + \mathcal{D}(\bar{y})$, i.e., $y \not\leq_2 \bar{y}$ for all $y \in A \setminus \{\bar{y}\}$.

(b) An element $\bar{y} \in A$ is a *strongly minimal* element of A w.r.t. the ordering map $\mathcal{D}$ if $A \subseteq \{\bar{y}\} + \mathcal{D}(\bar{y})$.

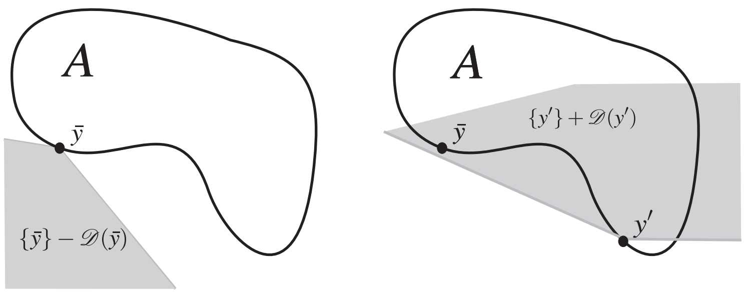

Figure 7: The element $\bar{y} \in A$ is a minimal element of A w.r.t. the ordering map $\mathcal{D}$ whereas $\bar{y}$ is not a nondominated element of A w.r.t. the ordering map $\mathcal{D}$ because of $\bar{y} \in \{y'\} + \mathcal{D}(y') \setminus \{0_Y\}$.

(c) An element $\bar{y} \in A$ with $\mathrm{int}(\mathcal{D}(\bar{y})) \neq \emptyset$ is a *weakly minimal* element of A w.r.t. the ordering map $\mathcal{D}$ if there is no $y \in A$ such that $\bar{y} \in \{y\} + \mathrm{int}(\mathcal{D}(\bar{y}))$.

Definition 4.2.(a) is equivalent to saying that $\bar{y}$ is a minimal element of A if and only if

$$(\{\bar{y}\} - \mathcal{D}(\bar{y})) \cap A = \{\bar{y}\}.$$

For an illustration of both optimality notions see Figure 7.

The concepts of strongly minimal and strongly nondominated elements w.r.t. an ordering map $\mathcal{D}$ are illustrated in the following example.

Example 4.2. Let $Y = \mathbb{R}^2$, the cone-valued map $\mathcal{D}\colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be defined by

$$\mathcal{D}(y_1, y_2) := \begin{cases} \mathbb{R}^2_+ & \text{if } y_2 = 0, \\ \text{cone conv}\{(|y_1|, |y_2|), (1,0)\} & \text{otherwise,} \end{cases}$$

and

$$A := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \leq y_2 \leq 2y_1\} \ .$$

One can check that $(0,0) \in A$ is a strongly minimal and also a strongly nondominated element of A w.r.t. $\mathcal{D}$.

The cone $\mathcal{D}(y)$ for some $y \in Y$ as used in the definition of the minimal elements is related to the set of preferred directions. First, one defines a set-values map $\mathcal{P} : Y \rightrightarrows Y$ with

$$\mathcal{P}(y) := \{d \in Y \mid y + d \ \text{ is preferred to } y\} \cup \{0_Y\}.$$

Then $\bar{y}$ is a minimal element if there is no preferred element, i.e. if

$$(\{\bar{y}\} + \mathcal{P}(\bar{y})) \cap A = \{\bar{y}\}.$$

For a unified representation (and as done in [6]) we set $\mathcal{D}(y) := -\mathcal{P}(y)$ which leads to Definition 4.2.(a). Note that the concepts of preference and of domination are two basically different approaches and that in general

$$\{d \in Y \mid y{+}d \ \text{ is dominated by } y\} \cup \{0_Y\} \neq -\{d \in Y \mid y{+}d \ \text{ is preferred to } y\} \cup \{0_Y\}.$$

Observe that $\bar{y}$ is a minimal element of some set $A \subseteq Y$ w.r.t. $\mathcal{D}$ if and only if it is an efficient element of the set A with Y partially ordered by $K := \mathcal{D}(\bar{y})$.

The following example illustrates that the concepts of nondominated and of minimal elements w.r.t. an ordering map $\mathcal{D}$ are not directly related.

Example 4.3. Let $Y = \mathbb{R}^2$, the cone-valued map $\mathcal{D}_1 \colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ be defined by

$$\mathcal{D}_1(y_1, y_2) := \begin{cases} \text{cone conv}\{(-1,1),(0,1)\} & \text{if } y_2 \geq 0, \\ \mathbb{R}^2_+ & \text{otherwise,} \end{cases}$$

and

$$A := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 \leq 1\} \ .$$

Then $(-1,0)$ is a nondominated but not a minimal element of A w.r.t. $\mathcal{D}_1$.

Considering instead the cone-valued map $\mathcal{D}_2 \colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ defined by

$$\mathcal{D}_2(y_1, y_2) := \begin{cases} \text{cone conv}\{(1,-1),(1,0)\} & \text{if } y_2 \geq 0, \\ \mathbb{R}^2_+ & \text{otherwise,} \end{cases}$$

then $(0,-1)$ is a minimal but not a nondominated element of A w.r.t. $\mathcal{D}_2$.

Considering instead the cone-valued map $\mathcal{D}_3 \colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ defined by

$$\mathcal{D}_3(y_1, y_2) := \begin{cases} \mathbb{R}^2_+ & \text{if } y \in \mathbb{R}^2 \setminus \{(0,-1),(-1,0)\}, \\ \{(z_1,z_2) \in \mathbb{R}^2 \mid z_1 \leq 0,\ z_2 \geq 0\} & \text{if } y = (0,-1), \\ \{(z_1,z_2) \in \mathbb{R}^2 \mid z_1 \geq 0,\ z_2 \leq 0\} & \text{if } y = (-1,0), \end{cases}$$

then all elements of the set $\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 = 1,\ y_1 \leq 0,\ y_2 \leq 0\}$ are minimal elements of A w.r.t. $\mathcal{D}$ but there is no nondominated element of the set A w.r.t. $\mathcal{D}$.

The two optimality concepts are only related under strong assumptions on $\mathcal{D}$:

Lemma 4.1. *(a) If $\mathcal{D}(y) \subseteq \mathcal{D}(\bar{y})$ for all $y \in A$ for some minimal element $\bar{y}$ of A w.r.t. $\mathcal{D}$, then $\bar{y}$ is also a nondominated element of A w.r.t. $\mathcal{D}$.*

(b) If $\mathcal{D}(\bar{y}) \subseteq \mathcal{D}(y)$ for all $y \in A$ for some nondominated element $\bar{y}$ of A w.r.t. $\mathcal{D}$, then $\bar{y}$ is also a minimal element of A w.r.t. $\mathcal{D}$.

These results are a direct consequence of the definitions.

Besides considering optimal elements of a set, all concepts apply also for a vector optimization problem with the linear space Y equipped with a variable ordering structure analogously to Definition 3.3.

For both optimality concepts, for minimal and for nondominated elements w.r.t. an ordering map $\mathcal{D}$, and for the related concepts of strongly and weakly optimal elements, we can easily derive the following properties.

Lemma 4.2. *(a) Any strongly nondominated element of A w.r.t. $\mathcal{D}$ is also a nondominated element of A w.r.t. $\mathcal{D}$. Any strongly minimal element of A w.r.t. $\mathcal{D}$ is also a minimal element of A w.r.t. $\mathcal{D}$.*

(b) If $\mathcal{D}(A)$ is pointed, then there is at most one strongly nondominated element of A w.r.t. $\mathcal{D}$.

(c) Let $int(\mathcal{D}(y)) \neq \emptyset$ for all $y \in A$. Any nondominated element of A w.r.t. $\mathcal{D}$ is also a weakly nondominated element of A w.r.t. $\mathcal{D}$. Any minimal element of A w.r.t. $\mathcal{D}$ is also a weakly minimal element of A w.r.t. $\mathcal{D}$.

(d) *If $\bar{y}$ is a strongly nondominated element of A w.r.t. $\mathcal{D}$, then the set of minimal elements of A w.r.t. $\mathcal{D}$ is empty or equals $\{\bar{y}\}$. If $\mathcal{D}(A)$ is additionally pointed, then $\bar{y}$ is the unique minimal element of A w.r.t. $\mathcal{D}$.*

(e) *If $\bar{y} \in A$ is a strongly minimal element of A w.r.t. $\mathcal{D}$ and if $\mathcal{D}(\bar{y}) \subseteq \mathcal{D}(y)$ for all $y \in A$, then $\bar{y}$ is also a strongly nondominated element of A w.r.t. $\mathcal{D}$.*

For proofs we refer to [20].

A common result is that the efficient elements of a set in a partially ordered space are a subset of the boundary of that set, see Lemma 3.2. The result remains true for variable ordering structures – at least under some assumptions.

Lemma 4.3. *(a) (i) Let $\mathrm{int}(\mathcal{D}(y)) \neq \emptyset$ for all $y \in Y$. If $\bar{y} \in A$ is a weakly minimal element of the set A w.r.t. the ordering map $\mathcal{D}$, then $\bar{y} \in \partial A$.*

(ii) If $\bar{y} \in A$ is a minimal element of the set A w.r.t. the ordering map $\mathcal{D}$ and $\mathcal{D}(\bar{y}) \neq \{0_Y\}$, then $\bar{y} \in \partial A$.

(b) (i) If $\bigcap_{y \in A} \mathrm{int}(\mathcal{D}(y)) \neq \emptyset$ and $\bar{y} \in A$ is a weakly nondominated element of the set A w.r.t. the ordering map $\mathcal{D}$, then $\bar{y} \in \partial A$.

(ii) If $\bigcap_{y \in A} \mathcal{D}(y) \neq \{0_Y\}$ and $\bar{y} \in A$ is a nondominated element of the set A w.r.t. the ordering map $\mathcal{D}$, then $\bar{y} \in \partial A$.

The following example demonstrates that we need for instance in (b)(i) in Lemma 4.3 an assumption like

$$\bigcap_{y \in A} \mathrm{int}(\mathcal{D}(y)) \neq \emptyset \ . \tag{19}$$

Example 4.4. For the set $A = [1,3] \times [1,3] \subseteq \mathbb{R}^2$ and the ordering map $\mathcal{D} \colon \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$,

$$\mathcal{D}(y) := \begin{cases} \mathbb{R}^2_+ & \text{for all } y \in \mathbb{R}^2 \text{ with } y_1 \geq 2, \\ \{(z_1, z_2) \in \mathbb{R}^2 \mid z_1 \leq 0, \ z_2 \geq 0\} & \text{else,} \end{cases}$$

the point $\bar{y} = (2,2)$ is a weakly nondominated element of A w.r.t. $\mathcal{D}$ but $\bar{y} \notin \partial A$.

Next, we give some scalarization results for (weakly) nondominated and minimal elements w.r.t. a variable ordering structure. A basic scalarization technique in vector optimization is based on continuous linear functionals l from the topological dual space Y^*, see Theorem 3.1. Then one examines the scalar-valued optimization problems

$$\min_{y \in A} \ l(y) \ .$$

We get the following sufficient conditions for (weakly) optimal elements w.r.t. a variable ordering [27, 20]:

Theorem 4.1. *Let $\bar{y} \in A$.*

(a) (i) If for some $l \in (\mathcal{D}(\bar{y}))^{\#}$

$$l(\bar{y}) \leq l(y) \text{ for all } y \in A \ ,$$

then $\bar{y}$ is a minimal element of A w.r.t. the ordering map $\mathcal{D}$.

(ii) Let int$(\mathcal{D}(y)) \neq \emptyset$ *for all* $y \in A$*. If for some* $l \in (\mathcal{D}(\bar{y}))^* \setminus \{0_{Y^*}\}$

$$l(\bar{y}) \leq l(y) \text{ for all } y \in A\ ,$$

then $\bar{y}$ *is a weakly minimal element of* A *w.r.t. the ordering map* $\mathcal{D}$*.*

(b) (i) If for some $l \in (\mathcal{D}(A))^{\#}$

$$l(\bar{y}) \leq l(y) \text{ for all } y \in A\ ,$$

then $\bar{y}$ *is a nondominated element of* A *w.r.t. the ordering map* $\mathcal{D}$*.*

(ii) Let int$(\mathcal{D}(y)) \neq \emptyset$ *for all* $y \in A$ *and let* $\mathcal{D}(A)$ *be convex. If for some* $l \in (\mathcal{D}(A))^* \setminus \{0_{Y^*}\}$

$$l(\bar{y}) \leq l(y) \text{ for all } y \in A\ ,$$

then $\bar{y}$ *is a weakly nondominated element of* A *w.r.t. the ordering map* $\mathcal{D}$*.*

Because of $(\mathcal{D}(A))^* \subseteq (\mathcal{D}(\bar{y}))^*$ and $(\mathcal{D}(A))^{\#} \subseteq (\mathcal{D}(\bar{y}))^{\#}$ for any $\bar{y} \in A$ it it is also possible for simplicity to consider functionals l in $(\mathcal{D}(A))^*$ and in $(\mathcal{D}(A))^{\#}$ in (a), respectively. A necessary condition for the quasi interior of the dual cone of a convex cone to be nonempty is the pointedness of the cone [35, Lemma 1.27]. This shows the limitation of the above results if the variable ordering structure varies too much, i.e., if $\mathcal{D}(A)$ is no longer a pointed cone. Then the quasi-interior of the dual cone $(\mathcal{D}(A))^{\#}$ is empty and the above characterizations can no longer be applied. For that reason also nonlinear scalarization have to be considered, compare [25, 21].

Under the additional assumption that A is a convex set also necessary conditions for weakly optimal elements and hence also for optimal elements w.r.t. a variable ordering can be formulated with the help of linear functionals [20].

Theorem 4.2. *Let* A *be convex and let int*$(\mathcal{D}(y)) \neq \emptyset$ *for all* $y \in A$*.*

(a) For any weakly minimal element $\bar{y} \in A$ *of* A *w.r.t. the ordering map* $\mathcal{D}$ *there exists some* $l \in (\mathcal{D}(\bar{y}))^* \setminus \{0_{Y^*}\}$ *with*

$$l(\bar{y}) \leq l(y) \text{ for all } y \in A\ .$$

(b) Set

$$\hat{D} := \bigcap_{y \in A} \mathcal{D}(y)$$

and let int$(\hat{D})$ *be nonempty. For any weakly nondominated element* $\bar{y} \in A$ *of* A *w.r.t. the ordering map* $\mathcal{D}$ *there exists some* $l \in \hat{D}^* \setminus \{0_{Y^*}\}$ *with*

$$l(\bar{y}) \leq l(y) \quad \text{for all} \quad y \in A\ .$$

The necessary condition for weakly nondominated elements w.r.t. the ordering map $\mathcal{D}$ is very weak if the cones $\mathcal{D}(y)$ for $y \in A$ vary too much, because then the cone $\hat{D}$ is very small (or even trivial) and the dual cone is very large.

Example 4.5. Let $Y = \mathbb{R}^2$ and let $\mathcal{D}$ and A be defined as in Example 4.4. The unique nondominated element w.r.t. $\mathcal{D}$ is $(2,1)$ and all the elements of the set

$$\{(2,t) \in \mathbb{R}^2 \mid t \in [1,3]\} \cup \{(t,1) \in \mathbb{R}^2 \mid t \in [1,3]\}$$

are weakly nondominated w.r.t. $\mathcal{D}$. Further, $\mathcal{D}(A) = \{(z_1,z_2) \in \mathbb{R}^2 \mid z_2 \geq 0\}$ and thus $(\mathcal{D}(A))^* = \{(z_1,z_2) \in \mathbb{R}^2 \mid z_1 = 0,\ z_2 \geq 0\}$, i.e. $(\mathcal{D}(A))^\# = \emptyset$.

Let $l \in (\mathcal{D}(A))^* \setminus \{0_{\mathbb{R}^2}\}$ be arbitrarily chosen, i.e. $l_1 = 0$, $l_2 > 0$, and consider the scalar-valued optimization problem

$$\min_{y \in A} l^\top y \ .$$

Then all elements of the set $\{(t,1) \in \mathbb{R}^2 \mid t \in [1,3]\}$ are minimal solutions and hence are weakly nondominated elements of A w.r.t. $\mathcal{D}$ according to Theorem 4.1.(b)(ii). All the other weakly nondominated elements w.r.t. $\mathcal{D}$ cannot be found by the sufficient condition. Because of $\operatorname{int}(\hat{D}) = \emptyset$, the necessary condition of Theorem 4.2.(b) cannot be applied.

4.3 Numerical Procedures

Algorithm 2 can also be applied to try to determine the optimal elements w.r.t. a variable ordering structure: of course in lines 4 and 12 the binary relation $\leq_1$ or $\leq_2$, respectively, has to be used. However, Algorithm 2 may even then determine only a superset of the set of optimal elements. By adding a third while-loop, the exact set of optimal elements can be determined. We present in the following such an algorithm for the notion of nondominatedness. For a more general discussion as well as an example that a third while loop is necessary we refer to [22].

Theorem 4.3. *Let A be a finite subset of Y and let U, T and V denote the sets gained by Algorithm 4.*

(a) If $\bar{y}$ is a nondominated element of A w.r.t. $\mathcal{D}$, then $\bar{y} \in U$ and $\bar{y} \in T$.

(b) The elements of the set $T \subseteq A$ are all nondominated elements of T w.r.t. $\mathcal{D}$.

(c) If $\leq_1$ is a transitive and antisymmetric binary relation, then the set T is exactly the set of all nondominated elements of A w.r.t. $\mathcal{D}$.

(d) The set V is exactly the set of all nondominated elements of A w.r.t. $\mathcal{D}$.

Proof. The proof of (a) and (b) is similar to the proof of part (a) and the first part of (b) of Theorem 3.3.

(c) We first show that for all $y \in A$ there exists a nondominated element $\bar{y}$ of A w.r.t. $\mathcal{D}$ with $y \in \{\bar{y}\} + \mathcal{D}(\bar{y})$. For that, let $y \in A$ be arbitrarily given. If y is a nondominated element of A w.r.t. $\mathcal{D}$ then the assertion is proven. Now, let y be not a nondominated element of A w.r.t. $\mathcal{D}$, i.e. there exists some $y^1 \in A$ with $y^1 \leq_1 y$, $y^1 \neq y$. If y^1 is nondominated we are done. Otherwise there is some $y^2 \neq y^1$ with $y^2 \leq_1 y^1$ and by the transitivity also $y^2 \leq_1 y$, $y^2 \neq y$. If y^2 is not a nondominated element we can find $y^3 \in A \setminus \{y, y^1, y^2\}$ with $y^3 \leq_1 y^2 \leq_1 y^1 \leq_1 y$ and so on. As A is

Algorithm 4 Jahn-Graef-Younes method for nondominated elements

Require: $A = \{y^1, \dots, y^k\}$, $\mathcal{D}(y)$ for all $y \in A$

put $U = \{y^1\}$ and $i = 1$
while $i < k$ **do**
    replace i by $i + 1$
    if $y^i \notin \{y\} + \mathcal{D}(y)$ for all $y \in U$ **then**
        replace U by $U \cup \{y^i\}$
    end if
end while
put $\{u^1, \dots, u^p\} = U$
put $T = \{u^p\}$ and $i = p$
while $i > 1$ **do**
    replace i by $i - 1$
    if $u^i \notin \{u\} + \mathcal{D}(u)$ for all $u \in T$ **then**
        replace T by $\{u^i\} \cup T$
    end if
end while
put $\{t^1, \dots, t^q\} = T$
put $V = \emptyset$ and $i = 0$
while $i < q$ **do**
    replace i by $i + 1$
    if $t^i \notin \{y\} + \mathcal{D}(y)$ for all $y \in A \setminus T$ **then**
        replace V by $V \cup \{y^i\}$
    end if
end while
return the set V of nondominated elements of A w.r.t. $\mathcal{D}$.

finite and $\leq_1$ is antisymmetric, this procedure stops with a nondominated element $\bar{y} \in A$ of A w.r.t. $\mathcal{D}$ with $\bar{y} \leq_1 y$.

According to (a) and (b) all nondominated elements of the set A w.r.t. $\mathcal{D}$ are an element of T and all the elements of T are nondominated elements of T w.r.t. $\mathcal{D}$. It remains to be shown that the elements of T are also nondominated elements of A w.r.t. $\mathcal{D}$. Let $y \in T$ and y be not a nondominated element of A w.r.t. $\mathcal{D}$. As we have shown above, there exists a nondominated element $\bar{y}$ of A w.r.t. $\mathcal{D}$ with $y \in \{\bar{y}\}+\mathcal{D}(\bar{y})\setminus\{0_Y\}$. According to (a), $\bar{y} \in T$ in contradiction to y a nondominated element of T w.r.t. $\mathcal{D}$.

(d) This is a direct consequence of (a), (b) and the definition of nondominated elements. □

Conditions ensuring the transitivity and the antisymmetry of $\leq_1$ are given in [22].

Example 4.6. Let the set

$$A = \{(x_1, x_2) \in [0, \pi] \times (0, \pi] \mid x_1^2 + x_2^2 - 1 - 0.1 \cos\left(16\text{arctan}(\tfrac{x_1}{x_2})\right) \geq 0, \; (x_1 - 0.5)^2 + (x_2 - 0.5)^2 \leq 0.5\} \subset \mathbb{R}^2,$$

which was originally defined by Tanaka [48], be given. It holds $\inf_{y\in A} y_i > 0$, $i = 1, 2$. For the variable ordering structure we define the ordering map by

$$\mathcal{D}(y) = \{u \in \mathbb{R}^2 \mid \|u\|_2 \leq \ell(y)^\top u\}$$

with

$$\ell(y) := \frac{2}{\min_{i=1,2} y_i} y \quad \text{for all } y \in A.$$

Next, we generate a discrete approximation D of the set A with 5 014 points by

$$D := A \cap \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in \{0, 0.01, 0.02, \ldots, \pi\},\; x_2 \in \{0.01, 0.02, \ldots, \pi\}\},$$

compare the set of dots in Figure 8.

The first while-loop of Algorithm 4 selects 27 points (the set U) of the set D as candidates for being nondominated. For that, 61 128 evaluations of the binary relation defined by $\mathcal{D}$ have been necessary. The second while-loop reduces these 27 points to 12 points, the set T, compare Figure 8, by only 222 additional evaluations of the binary relation. By comparing these remaining points with all other 5 014 points of the discretization in the third while-loop (additionally, 60 156 evaluations of the binary relation) verifies that these 12 points are exactly the nondominated elements of the discretization set D w.r.t. $\mathcal{D}$. A total of 121 506 evaluations of the binary relation are thus needed.

A pairwise comparison of all 5 014 points with all other points (till it is shown that an element is dominated by another point or nondominated w.r.t. all) needs 4 472 290 evaluations of the binary relation, i.e. a reduction of around 97% is reached.

In addition to the above algorithm, also other numerical procedures have been developed to determine the optimal elements w.r.t. a variable ordering structure based on different ideas from vector optimization in partially ordered spaces. The

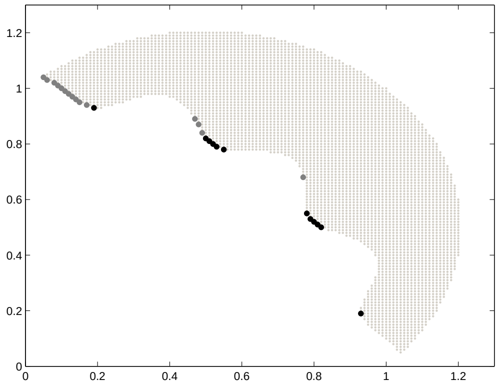

Figure 8: The sets D, U and T in light gray, dark gray and black, respectively.

first numerical procedure designed especially for the application problem presented in Subsection 4.1 was given in [51]. For a method for determining the minimal elements using a steepest descent method we refer to [4]. Also based on Algorithm 1 an algorithm for differentiable problems was developed assuming $\mathbb{R}^2_+ \subseteq \mathcal{D}(y)$ for all $y \in \mathbb{R}^2$. Not all points of the generated set A will be weakly nondominated, so an additional while-loop which selects only the weakly optimal elements has to be added. For a discussion of the method also in a more general setting we refer to [22].

References

[1] Alderliesten, T., Sonke, J.-J., and Bosman, P.A.N.:. Multi-objective optimization for deformable image registration: Proof of concept. Paper presented at the Progress in Biomedical Optics and Imaging - Proceedings of SPIE, **8314**, Article number 831420 (2012).

[2] Algoul, S., Alam, M.S., Hossain, M.A., and Majumder, M.A.A.: Multi-objective optimal chemotherapy control model for cancer treatment. Medical and Biological Engineering and Computing, **49**(1), 51–65 (2011).

[3] Baatar, D. and Wiecek, M.M.: Advancing equitability in multiobjective programming. Comput. Math. Appl. **52**, 225–234 (2006).

[4] Bello Cruz, J.Y., Bouza Allende, G.: A steepest descent-like method for variable order vector optimization problems. Submitted (2012)

[5] Chen, G.Y.: Existence of solutions for a vector variational inequality: an extension of the Hartmann-Stampacchia Theorem. J. Optim. Theory Appl. **74**, 445–456 (1992)

[6] Chen, G.Y., and Yang, X.Q.: Characterizations of variable domination structures via nonlinear scalarization. J. Optim. Theory Appl. **112** 97–110 (2002).

[7] G.Y. Chen, X. Huang and X. Yang, *Vector Optimization, Set-Valued and Variational Analysis* (Springer, Berlin, 2005).

[8] Chen, G.Y., Yang, X.Q., and Yu, H.: A nonlinear scalarization function and generalized quasi-vector equilibrium problems. J. Global Optim. **32** 451–466 (2005).

[9] Coello Coello, C.A.: An updated survey of evolutionary multiobjective optimization techniques: state of the art and future trends. Evolutionary Computation (1999)

[10] C.A. Coello Coello, D.A. Van Veldhuizen, and G.B. Lamont, *Evolutionary algorithms for solving multiobjective problems* (Springer, New York, 2007).

[11] Craft, D.L., Halabi, T.F., Shih, H.A., and Bortfeld, T.R.: Approximating convex Pareto surfaces in multiobjective radiotherapy planning. Medical Physics, **33**(9), 3399–3407 (2006).

[12] Craft, D., Halabi, T., Shih, H.A., and Bortfeld, T.: An approach for practical multiobjective IMRT treatment planning. International Journal of Radiation Oncology Biology Physics, **69**(5), 1600–1607 (2007).

[13] Edgeworth, F.Y.: *Mathematical psychics.* (Kegan Paul, London, 1881)

[14] M. Ehrgott, *Multicriteria Optimization* (Springer, Heidelberg, 2005).

[15] Ehrgott, M., Gler, Ç., Hamacher, H. W., and Shao, L.: Mathematical optimization in intensity modulated radiation therapy. Annals of Operations Research, **175**(1), 309–365 (2010).

[16] Ehrgott, M., and Johnston, R.: Optimisation of beam directions in intensity modulated radiation therapy planning. OR Spectrum, **25**(2), 251–264 (2003).

[17] Eichfelder, G.: ε-constraint method with adaptive parameter control and an application to intensity-modulated radiotherapy, in: *Multicriteria Decision Making and Fuzzy Systems, Theory, Methods and Applications*, eds.: K.-H. Küfer, H. Rommelfanger, C. Tammer and K. Winkler, 25 - 42 (Shaker, Aachen, 2006).

[18] G. Eichfelder, *Adaptive Scalarization Methods in Multiobjective Optimization.* (Springer, Berlin, 2008).

[19] Eichfelder, G.: An adaptive scalarization method in multi-objective optimization. SIAM Journal on Optimization, **19**(4) 1694–1718 (2009).

[20] Eichfelder, G.: Optimal elements in vector optimization with a variable ordering structure. J. Optim. Theory Appl. Vol. **151**(2), 217–240 (2011).

[21] Eichfelder, G.: Variable ordering structures in vector optimization, Chapter 4 (p.95-126) in: *Recent Developments in Vector Optimization*, Ansari, Q.H., Yao, J.-C. (Eds.) (Springer, Heidelberg, 2012).

[22] Eichfelder, G.: Numerical procedures in multiobjective optimization with variable ordering structures. Preprint-Series of the Institute of Mathematics, Ilmenau University of Technology, Germany (2012).

[23] Eichfelder, G., Gebhardt, M.: Local specific absorption rate control for parallel transmission by virtual observation points. Magnetic Resonance in Medicine **66**(5), 1468-1476 (2011).

[24] Eichfelder, G., Gebhardt, M.: Method for determining sensitivity matrices for hotspots. US Patent Application US13/045,832 RSC, Publication Number US 2011/0224924 A1, Filed 11.3.2011, published 15.9.2011.

[25] Eichfelder, G., Ha, T.X.D.: Optimality conditions for vector optimization problems with variable ordering structures. Optimization (2011) DOI 10.1080/02331934.2011.575939

[26] Eichfelder, G., Jahn, J.: Vector optimization problems and their solution concepts, Chapter 1 in: Recent Developments in Vector Optimization, Ansari, Q.H.,Yao, J.-C. (Eds.), p. 1-27, Springer, 2012.

[27] Engau, A.: Variable preference modeling with ideal-symmetric convex cones. J. Global Optim. **42**, 295–311 (2008)

[28] Fliege, J., Grana Drummond, L.M., and Svaiter, B.F.: Newton's method for multiobjective optimization. SIAM J. on Optimization **20**(2), 602–626 (2009).

[29] Gerstewitz (Tammer), C.: Nichtkonvexe Dualität in der Vektoroptimierung. Wissensch. Zeitschr. TH Leuna-Merseburg **25** 357–364 (1983).

[30] Gerstewitz (Tammer), C., and Iwanow, E.: Dualität für nichtkonvexe Vektoroptimierungsprobleme. Wissensch. Zeitschr. der Techn. Hochschule Ilmenau **31** 61–81 (1985).

[31] A. Göpfert, H. Riahi, C. Tammer, and C. Zălinescu, *Variational Methods in Partially Ordered Spaces* (Springer, New York, 2003).

[32] Hillermeier, C., and Jahn, J.: Multiobjective optimization: survey of methods and industrial applications. Surveys on Mathematics for Industry **11** 1–42 (2005).

[33] Holdsworth, C., Kim, M., Liao, J., and Phillips, M.: The use of a multiobjective evolutionary algorithm to increase flexibility in the search for better IMRT plans. Medical Physics **39**(4), 2261–2274 (2012).

[34] Homann, H., Graesslin, I., Eggers, H., Nehrke, K., Vernickel, P., Katscher, U., Dössel, O., Börnert, P.: Local SAR management by RF Shimming: A simulation study with multiple human body models. Magnetic Resonance Materials in Physics, Biology and Medicine **25**(3) 193-204 (2012).

[35] J. Jahn, *Vector Optimization - Theory, Applications, and Extensions*, 2nd edition (Springer, Heidelberg, 2011).

[36] Jahn, J. and Rathje, U.: Graef-Younes method with backward iteration, in: *Multicriteria decision making and fuzzy systems - theory, methods and applications*, K.-H. Küfer et al. (Eds.) (Shaker, Aachen, 2006) 75–81.

[37] Küfer, K.-H., Scherrer, A., Monz, M., Alonso, F., Trinkaus, H., Bortfeld, T., and Thieke, C.: Intensity-modulated radiotherapy - a large scale multi-criteria programming problem. OR Spectrum **25**, 223–249 (2003).

[38] Küfer, K.-H., and Monz, M., Private Communication (2004).

[39] Lee, J., Gebhardt, M., Wald, L.L., Adalsteinsson, E.: Local SAR in parallel transmission pulse design Magnetic Resonance in Medicine **67** (6) 1566–1578 (2012).

[40] D.T. Luc, *Theory of Vector Optimization* (Springer, Berlin, 1989).

[41] K.M. Miettinen, *Nonlinear Multiobjective Optimization* (Kluwer Academic Publishers, Boston, 1999).

[42] V. Pareto, Manuale di economia politica. (Societa Editrice Libraria, Milano, Italy, 1906), English translation: V. Pareto, *Manual of political economy*, translated by Schwier, A.S.. Augustus M. (Kelley Publishers, New York, 1971).

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

[44] Rubinov, A.M.: Sublinear operators and their applications. Uspekhi Mat. Nauk **32** 113–174 (1977).

[45] Ruzika S., and Wiecek, M.M.: Approximation methods in multiobjective programming. J. Optim. Theory Appl. **126** 473-501 (2005).

[46] Scherrer, A., Küfer, K.-H., Monz, M., Alonso, F., and Bortfeld, T.: IMRT planning on adaptive volume structures — a significant advance of computational complexity. Technical Report **60**. Fraunhofer ITWM, Kaiserslautern, Germany (2004).

[47] Tammer, C., and Zălinescu, C.: Vector variational principles for set-valued functions. Chapter 11 in: Recent Developments in Vector Optimization, Ansari, Q.H., Yao, J.-C. (Eds.) (Springer, Heidelberg, 2011).

[48] Tanaka, M.: GA-based decision support system for multi-criteria optimization. Proceedings of the international Conference on Systems, Man and Cybernetics 2, 1556–1561 (1995)

[49] C. Thieke, *Multicriteria Optimization in Inverse Radiotherapy Planning* (PhD thesis, University of Heidelberg, Germany, 2003).

[50] Van Veldhuizen, D.A., and Lamont, G.B: Multiobjective evolutionary algorithms: analyzing the state-of-the-art. Evolutionary Computation **8**(2): 125–147 (2000).

[51] Wacker, M.: *Multikriterielle Optimierung bei Registrierung medizinischer Daten.* (Diplomarbeit, Univ. Erlangen-Nürnberg, Germany, 2008).

[52] Wacker, M. and Deinzer, F.: Automatic robust medical image registration using a new democratic vector optimization approach with multiple measures. In: *Medical Image Computing and Computer-Assisted Intervention MICCAI 2009*, Eds. G.-Z. Yang et al., 590–597 (2009).

[53] Y.M. Younes, *Studies on discrete vector optimization* (Dissertation, University of Demiatta, 1993).

[54] Yu, P.L.: Cone convexity, cone extreme points, and nondominated solutions in decision problems with multiobjectives. J. Optim. Theory Appl. **14**, 319–377 (1974).

[55] P.L. Yu, *Multiple-criteria Decision Making: Concepts, Techniques and Extensions* (Plenum Press, New York, 1985).

[56] Zhou, A., Qu, B.-Y., Li, H., Zhao, S.-Z., Suganthan, P.N., and Zhang, Q.: Multiobjective evolutionary algorithms: A survey of the state-of-the-art. Swarm and Evolutionary Computation **1**(1) 32-49 (2011).