Math. Program., Ser. A
DOI 10.1007/s10107-008-0259-0

FULL LENGTH PAPER

Multiobjective bilevel optimization

Gabriele Eichfelder

Received: 5 December 2006 / Accepted: 3 November 2008

Abstract In this work nonlinear non-convex multiobjective bilevel optimization problems are discussed using an optimistic approach. It is shown that the set of feasible points of the upper level function, the so-called induced set, can be expressed as the set of minimal solutions of a multiobjective optimization problem. This artificial problem is solved by using a scalarization approach by Pascoletti and Serafini combined with an adaptive parameter control based on sensitivity results for this problem. The bilevel optimization problem is then solved by an iterative process using again sensitivity theorems for exploring the induced set and the whole efficient set is approximated. For the case of bicriteria optimization problems on both levels and for a one dimensional upper level variable, an algorithm is presented for the first time and applied to two problems: a theoretical example and a problem arising in applications.

Keywords Multicriteria optimization · Vector optimization · Sensitivity · Bilevel optimization · Two-level optimization

Mathematics Subject Classification (2000) 90C29 · 90C31 · 90C59

1 Introduction

Bilevel optimization is an active research area in mathematical programming (see the monographs of Dempe [7] and Bard [3] and the bibliography reviews by Calamai and Vicente [41] and Dempe [9]). Many papers have been published in the last two

G. Eichfelder (✉)
Institute of Applied Mathematics, University of Erlangen-Nürnberg,
Martensstr. 3, 91058 Erlangen, Germany
e-mail: Gabriele.Eichfelder@am.uni-erlangen.de

Published online: 19 December 2008

decades about bilevel optimization but there are only very few of them dealing with multiobjective bilevel problems.

An application for these problems is given in [42] and procedures for solving linear multiobjective bilevel problems are presented, e.g., in [29]. Again even less papers are dealing with nonlinear multiobjective bilevel problems: Shi and Xia [36,37] present an interactive method, Abo-Sinna [1], Osman et al. [30] propose the usage of fuzzy set theory for convex problems, and Teng et al. [39] give an approach for a convex multiperson multiobjective bilevel problem. Besides, Bonnel and Morgan consider a so-called semivectorial bilevel optimization problem and propose a solution method based on a Penalty approach in [4] but no numerical results are yet given.

In bilevel optimization, also called two-level optimization, problems are considered where the set of feasible points of the so-called upper level problem is given by the solution set of a so-called lower level parametric optimization problem. The variables of the upper level are the parameters of the lower level problem and again the solutions of the optimization problem on the lower level influence the upper level objective function value. With a vector-valued objective on one or both of the levels we speak of a multiobjective bilevel optimization problem.

Despite multiobjective bilevel optimization problems have not yet received a broad attention in the literature, they are very interesting in the view of possible applications (see, e.g., [40]). We illustrate this with an example. Let us consider a city bus transportation system financed by the public authorities. They have as target the reduction of the money losses in this non-profitable business. As a second target they want to bring as many people as possible to use the buses instead of their own cars, as it is a public mission to reduce the overall traffic. The public authorities can decide about the bus ticket price but this will influence the customers in their usage of the buses. The public has maybe several competing objectives, too, as to minimize their transportation time and costs. Hence the usage of the public transportation system can be modeled on the lower level with the bus ticket price as parameter and with the solutions influencing the objective values of the public authorities on the upper level again and thus such a problem can be modeled by bilevel multiobjective optimization.

For solving such a problem we show that the set of feasible points of the upper level problem can be expressed completely as the solution set of a multiobjective optimization problem. Then we can determine an approximation of this solution set based on a scalarization approach by Pascoletti and Serafini using results from multiobjective optimization.

Here our aim is not a good approximation of the image set of the minimal solutions, the so-called efficient set, as it is the aim generally in multiobjective optimization (see, e.g., [6,14,16,35]), but of the solution set itself. A good approximation of this minimal solution set and thus of the set of feasible points of the upper level problem is meant in the sense of an approximation with almost equidistant points. This is important for getting a representative approximation of the set of feasible points with a given accuracy and for avoiding to neglect larger parts.

For generating such an approximation of the set of feasible points we use sensitivity results for controlling the parameters of the scalarization problem adaptively. We will use these sensitivity results again for solving the upper level problem in an iterative process: the approximation of the set of feasible points is refined adaptively around the approximated minimal solutions of the upper level. Thereby not only one minimal solution but an approximation of the whole efficient set of the multiobjective bilevel optimization problem is determined.

We continue our presentation in Sect. 2 with a short introduction to multiobjective optimization and the scalarization approach used and in Sect. 3 with the basic notations in bilevel optimization. In Sect. 4 we explain some important theoretical results which we use in Sect. 5 for developing a numerical method for solving nonlinear multiobjective bilevel problems without convexity assumptions but demanding twice continuously differentiable functions. Thereby we consider scalarizations of multiobjective optimization problems, demanding in the case of nonconvex problems appropriate solvers for determining global solutions of the scalar problems. For the case of a bicriteria lower and upper level problem and a one-dimensional upper level variable, an algorithm is given for the first time which is finally applied to two problems, an academic example and a topical problem arising in an application. The results are presented in Sect. 6.

2 Basic notations in multiobjective optimization

In multiobjective optimization one discusses problems formally defined by

$$\begin{aligned}&\min f(x) = (f_1(x), \ldots, f_m(x))^\top\\ &\text{subject to the constraint}\\ &x \in \Omega \subset \mathbb{R}^n,\end{aligned} \tag{2.1}$$

with a vector valued objective function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ $(m,\ n \in \mathbb{N},\ m \geq 2)$ and a set of feasible points Ω. For defining minimality for the multiobjective optimization problem (2.1) we need a partial ordering in the image space. Thus one considers partial orderings introduced by arbitrary closed pointed convex cones $K \subset \mathbb{R}^m$. A set K is a convex cone if $\lambda(x+y) \in K$ for all $\lambda \geq 0,\ x, y \in K$, and K is a pointed convex cone if additionally $K \cap (-K) = \{0_m\}$. An ordering introduced by a pointed convex cone is antisymmetric. The ordering is then given by

$$\leq_K := \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m \mid y - x \in K\}.$$

Throughout the paper we mean by $\leq$ without a subscript the componentwise ordering defined by $\leq := \leq_{\mathbb{R}^m}$.

Definition 2.1 Let K be a closed pointed convex cone. A point $\bar{x} \in \Omega$ is called a K-minimal point of (2.1) or non-dominated w.r.t. K if $(f(\bar{x}) - K) \cap f(\Omega) = \{f(\bar{x})\}$. Additionally for $\operatorname{int}(K) \neq \emptyset$ a point $\bar{x} \in \Omega$ is called a weakly K-minimal point of (2.1) if $(f(\bar{x}) - \operatorname{int}(K)) \cap f(\Omega) = \emptyset$.

We denote the set of all K-minimal points as $\mathcal{M}(f(\Omega), K)$ and the set of all weakly K-minimal points as $\mathcal{M}^w(f(\Omega), K)$. The set $\mathcal{E}(f(\Omega), K) := \{f(x) \in \mathbb{R}^m | x \in \mathcal{M}(f(\Omega), K)\}$ is called *efficient set* and the set $\mathcal{E}^w(f(\Omega), K) := \{f(x) \in \mathbb{R}^m | x \in$

$\mathcal{M}^w(f(\Omega), K)\}$ *weakly efficient set*. For $K = \mathbb{R}^m_+$ the K-minimal points are denoted as *Edgeworth-Pareto (EP)-minimal* points, too.

In connection with the iterative bilevel algorithm we want to present, the following result for EP-minimal points is useful:

Lemma 2.2 *For two sets $A^0, A^1 \subset \mathbb{R}^n$, and a vector valued function $f : \mathbb{R}^n \to \mathbb{R}^m$, consider the sets*

$$\begin{aligned} A &= A^0 \cup A^1 \quad \text{and} \\ \tilde{A} &= \mathcal{M}(f(A^0), \mathbb{R}^m_+) \cup A^1. \end{aligned}$$

Let $f(A^0)$ be compact. Then it is $\mathcal{M}(f(A), \mathbb{R}^m_+) = \mathcal{M}(f(\tilde{A}), \mathbb{R}^m_+)$.

Proof As the result is trivial for $A^0 = \emptyset$ or $A^1 = \emptyset$ we assume in the following $A^0, A^1 \neq \emptyset$. The compactness of $f(A^0)$ then implies $\mathcal{M}(f(A^0), \mathbb{R}^m_+) \neq \emptyset$ [34, Theorem 3.2.3].

First we show the inclusion $\mathcal{M}(f(A), \mathbb{R}^m_+) \subset \mathcal{M}(f(\tilde{A}), \mathbb{R}^m_+)$. For that we assume $\bar{x} \in \mathcal{M}(f(A), \mathbb{R}^m_+)$. Then there exists no point $x' \in A$ with

$$f(x') \leq f(\bar{x}) \text{ and } f(x') \neq f(\bar{x}). \tag{2.2}$$

We have $\bar{x} \in A = A^0 \cup A^1$. If $\bar{x} \in A^0$ then there is no point $x' \in A^0 \subset A$ with (2.2) and hence $\bar{x} \in \mathcal{M}(f(A^0), \mathbb{R}^m_+) \subset \tilde{A}$. For $\bar{x} \in A^1$ we have $\bar{x} \in \tilde{A}$, too. Because of $\tilde{A} \subset A$ there also exists no $x' \in \tilde{A}$ with (2.2) and hence $\bar{x} \in \mathcal{M}(f(\tilde{A}), \mathbb{R}^m_+)$.

It remains to show $\mathcal{M}(f(\tilde{A}), \mathbb{R}^m_+) \subset \mathcal{M}(f(A), \mathbb{R}^m_+)$. For that we assume $\bar{x} \in \mathcal{M}(f(\tilde{A}),\mathbb{R}^m_+)$, i.e., there is no $x' \in \tilde{A}$ with (2.2). We have $\bar{x} \in \mathcal{M}(f(A^0), \mathbb{R}^m_+) \cup A^1$. For $\bar{x} \in \mathcal{M}(f(A^0), \mathbb{R}^m_+)$ there exists no $x' \in A^0$ with (2.2), too. Because of $A = \tilde{A} \cup A^0$ we conclude $\bar{x} \in \mathcal{M}(f(A), \mathbb{R}^m_+)$. And for $\bar{x} \in A^1$ we assume $\bar{x} \notin \mathcal{M}(f(A), \mathbb{R}^m_+)$. Then there exists $x' \in A \backslash \tilde{A} = A^0 \backslash \mathcal{M}(f(A^0), \mathbb{R}^m_+)$ with (2.2). As the set $f(A^0)$ is compact it holds according to [34, Theorem 3.2.10], $f(A^0) \subset \mathcal{E}(f(A^0), \mathbb{R}^m_+) + \mathbb{R}^m_+$. Because of $x' \in A^0$ there exists $x^0 \in \mathcal{M}(f(A^0), \mathbb{R}^m_+)$ with $f(x^0) \leq f(x')$ resulting in $f(x^0) \leq f(\bar{x})$ and $f(x^0) \neq f(\bar{x})$. Due to $x^0 \in \tilde{A}$ this is a contradiction to $\bar{x} \in \mathcal{M}(f(\tilde{A}), \mathbb{R}^m_+)$. Hence $\bar{x} \in \mathcal{M}(f(A), \mathbb{R}^m_+)$. □

For our numerical calculations the notion of ε-EP-minimal solutions (see [21,26, 38], and others) can be useful:

Definition 2.3 Let $\varepsilon \in \mathbb{R}^m$ with $\varepsilon_i > 0,\ i = 1, \ldots, m$, be given. A point $\bar{x} \in \Omega$ is an ε-EP-minimal solution of the multiobjective optimization problem (2.1) if there is no $x \in \Omega$ with

$$\begin{aligned} & f_i(x) + \varepsilon_i \leq f_i(\bar{x}) \quad \text{for all } i \in \{1, \ldots, m\} \\ \text{and} \quad & f_j(x) + \varepsilon_j < f_j(\bar{x}) \quad \text{for at least one } j \in \{1, \ldots, m\}. \end{aligned}$$

For solving problem (2.1) several methods are discussed in the literature (for surveys see [11,19,20,28,33]). A general approach is to replace the vector optimization problem by a suitable scalar optimization problem. Our aim is an approximation of the whole efficient set and for that we use a scalarization by Pascoletti and Serafini [31]. We use this scalarization approach as it allows us to determine K-minimal points (for an arbitrary ordering cone K) while many well known scalarization problems as the ε-constraint problem (see [11,18,27,28]) are mainly developed for determining EP-minimal points, only. Further the Pascoletti–Serafini scalarization is also adequate for nonconvex multiobjective optimization problems. For example the weighted sum scalarization has the disadvantage that it is generally only for convex problems possible to determine all EP-minimal points by an appropriate parameter choice (see [5,24]). What is more, the Pascoletti–Serafini scalarization problem is a very general problem in the sense that many other scalarization approaches as the mentioned ε-constraint method, the normal boundary intersection method by Das and Dennis [6], the modified Polak method by Jahn and Merkel [22], the weighted Chebyshev norm [10,25], and others are included in this general formulation (see [13]). Thus, results gained for the Pascoletti–Serafini problem can be applied to these scalarization problems, too.

Hence we consider the parameter dependent scalarization problems (according to Pascoletti and Serafini)

$$\begin{aligned} &\min\ t \\ &\text{subject to the constraints} \\ &a + t\,r - f(x) \in K, \\ &x \in \Omega, \\ &t \in \mathbb{R} \end{aligned} \tag{2.3}$$

for parameters $a, r \in \mathbb{R}^m$ to the multiobjective optimization problem (2.1). The main properties of this scalarization approach are the following:

Theorem 2.4 (a) *Let $\bar{x}$ be a K-minimal point of* (2.1)*, then* $(0, \bar{x})$ *is a minimal solution of* (2.3) *with $a = f(\bar{x})$, $r \in K\setminus\{0_m\}$.*
(b) *Let $(\bar{t}, \bar{x})$ be a minimal solution of* (2.3)*, then $\bar{x} \in \mathscr{M}^w(f(\Omega), K)$.*

Thus all K-minimal points of the multiobjective optimization problem can be found even for non-convex problems by choosing suitable parameters. We are even able to restrict the choice of the parameter a to a hyper plane in the image space according to the following theorem [14, Theorem 3.2].

Theorem 2.5 *Let $\bar{x} \in \mathscr{M}(f(\Omega), K)$ and $r \in K$ be given. We define a hyper plane H by $H = \{y \in \mathbb{R}^m \mid b^\top y = \beta\}$ with $b \in \mathbb{R}^m,\ b^\top r \neq 0,\ \beta \in \mathbb{R}$. Then there is a parameter $a \in H$ and some $\bar{t} \in \mathbb{R}$ such that $(\bar{t}, \bar{x})$ is a minimal solution of the problem* (2.3).

Note that the results of Theorems 2.4 and 2.5 are related to global minimal solutions of the scalar optimization problems. For the case of only local solutions similar results can be stated which give a connection to local K-minimal solutions of the problems (2.1) [12].

3 Basic notations in bilevel optimization

In this section we give a short introduction to bilevel optimization. For a detailed presentation, see, e.g., [7]. We formulate the general bilevel optimization problem (see (3.3)). Then we introduce the so-called optimistic approach (see (3.4)) which is a modification of the general bilevel optimization problem. It is a common procedure in bilevel optimization just to deal with this modification [7], [8, p. 5], [29, p. 166] and here we also consider the optimistic approach to bilevel optimization only.

As mentioned in the introduction, in bilevel optimization we have a parameter dependent optimization problem on the lower level depending on the variable $y \in \mathbb{R}^{n_2}$ of the upper level problem called upper level variable:

$$\begin{aligned} &\min_x f(x, y) \\ &\text{subject to the constraint} \\ &(x, y) \in G \end{aligned} \tag{3.1}$$

with a (vector valued) function $f : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{m_1}$ and a set $G \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ ($n_1, n_2, m_1 \in \mathbb{N}$). The constraint $(x, y) \in G$ can be replaced by

$$x \in G(y) := \{x \in \mathbb{R}^{n_1} \mid (x, y) \in G\}. \tag{3.2}$$

The parameters $x \in \mathbb{R}^{n_1}$ of this lower level problem are called lower level variables. For a constant $y \in \mathbb{R}^{n_2}$ let $x(y)$ be a minimal solution of (3.1), hence

$$x(y) \in \Psi(y) := \operatorname{argmin}_x\{f(x, y) \mid (x, y) \in G\} \subset \mathbb{R}^{n_1}.$$

The optimization problem of the upper level is then given by

$$\begin{aligned} &\text{“}\min_y\text{"} F(x(y), y) \\ &\text{subject to the constraints} \\ &x(y) \in \Psi(y), \\ &y \in \tilde{G} \end{aligned} \tag{3.3}$$

with a (vector valued) function $F : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{m_2}$, $m_2 \in \mathbb{N}$, and a compact set $\tilde{G} \subset \mathbb{R}^{n_2}$. Here the constraint $y \in \tilde{G}$ is uncoupled from the lower level variable as it is assumed, e.g., in [3, p. 303], [7, p. 123f], [9,17]. An even more general formulation of the bilevel problem is reached by allowing the constraint $y \in \tilde{G}(x)$ to depend on the lower level variable x (compare [3,7,8,41]), what we will not consider here.

We speak of a multiobjective bilevel optimization problem if $m_1 \geq 2$ or $m_2 \geq 2$ and in this work we even assume $m_1, m_2 \geq 2$. Bonnel and Morgan [4] denote a bilevel optimization problem with a scalar-valued optimization problem on the upper level and a vector-valued problem on the lower level also a semivectorial bilevel optimization problem.

If the minimal solution of the lower level problem (3.1) is not unique, i.e., the set $\Psi(y)$ consists of more than one point, the objective function $F(x(\cdot), \cdot)$ is not well-defined for $y \in \mathbb{R}^{n_2}$. That is the reason why the word "min" is written in quotes in (3.3). This difficulty is in some works avoided by just assuming that the solution of the lower level problem is unique. But in the case of a multiobjective optimization problem ($m_1 \geq 2$) on the lower level this cannot be done any more. In the case of non-uniqueness a common procedure is the optimistic approach. There it is assumed that the decision maker of the lower level chooses among all minimal solutions (for a fixed value of y) that minimal solution, which is best for the upper level, i.e., which is minimal for the objective function of the upper level. Thus, it is solved

$$\min_x \{F(x, y) \mid x \in \Psi(y)\} =: \varphi_0(y).$$

Using this function, the bilevel problem (3.3) is then, in the case of a scalar-valued map F, written as

$$\begin{aligned} &\min_y \varphi_0(y) \\ &\text{subject to the constraint} \\ &y \in \tilde{G} \end{aligned}$$

being (w.r.t. global minimal solutions) equivalent to

$$\begin{aligned} &\min_{x,y} F(x, y) \\ &\text{subject to the constraints} \\ &x \in \Psi(y), \\ &y \in \tilde{G}. \end{aligned} \tag{3.4}$$

Note that in the optimistic modification (3.4) the objective function of the upper level is minimized w.r.t. x and y, while in the general formulation (3.3) it is only minimized w.r.t. the upper level variable y. In this work we consider the optimistic approach only, i.e., the bilevel optimization problem as in (3.4), also for a vector-valued objective function F.

For the multiobjective optimization problem on the upper level we assume that the partial ordering is given by the closed pointed convex cone $K^2 \subset \mathbb{R}^{m_2}$ and for the lower level by the closed pointed convex cone $K^1 \subset \mathbb{R}^{m_1}$. In this work we assume that for any $y \in \tilde{G}$ of the upper level there exists a minimal solution of the lower level problem.

For the case $m_1 \geq 2$ the set $\Psi(y)$ is the solution set of a multiobjective optimization problem w.r.t. the ordering cone K^1. Thus, following the notation according to Sect. 2, p. 4 and using (3.2), we write instead of $x \in \Psi(y)$

$$x \in \mathcal{M}(f(G(y), y), K^1) =: \mathcal{M}_y(f(G), K^1), \tag{3.5}$$

with $\mathscr{M}(f(G(y), y), K^1)$ the set of K^1-minimal points of the multiobjective optimization problem (3.1) parameterized by y.

4 Theoretical results

In this section we show that the set of feasible points of the upper level problem of the considered multiobjective bilevel optimization problem can be expressed as the set of minimal solutions of a multiobjective optimization problem. That multiobjective optimization problem is solved by using the parameter dependent scalarization according to Pascoletti and Serafini introduced in Sect. 2. We further present a simplified scalarization and show the connection between the two scalarizations. Besides we examine the sensitivity on the parameters for the simplified scalarization. These results are used in Sect. 5 for developing an algorithm for solving the multiobjective bilevel optimization problem.

Using the optimistic approach as presented in Sect. 3 we consider the multiobjective bilevel optimization problem (3.4). For convenience we recall this problem using the notation in (3.5) instead of $x \in \Psi(y)$:

$$\begin{aligned} &\min_{x,y} F(x, y) \\ &\text{subject to the constraints} \\ &x \in \mathscr{M}_y(f(G), K^1), \\ &y \in \tilde{G}. \end{aligned} \tag{4.1}$$

The set of feasible points Ω of the upper level problem in (4.1), also called induced set, is then given by

$$\Omega = \{(x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mid x \in \mathscr{M}_y(f(G), K^1),\ \ y \in \tilde{G}\}.$$

In the following theorem we show that the set of feasible points Ω of problem (4.1) is equivalent to the set of $\hat{K}$-minimal points of the multiobjective optimization problem

$$\begin{aligned} &\min_{x,y} \hat{f}(x, y) := \begin{pmatrix} f(x, y) \\ y \end{pmatrix} \\ &\text{subject to the constraints} \\ &(x, y) \in G, \\ &y \in \tilde{G} \end{aligned} \tag{4.2}$$

w.r.t. the ordering cone $\hat{K} := K^1 \times \{0_{n_2}\} \subset \mathbb{R}^{m_1} \times \mathbb{R}^{n_2}$. A point $(\bar{x}, \bar{y})$ is thus an element of the set of feasible points Ω if and only if it is a minimal solution of the multiobjective optimization problem (4.2). This is shown in the following theorem.

Theorem 4.1 *Let $\hat{\mathscr{M}}$ be the set of $\hat{K}$-minimal points of the multiobjective optimization problem* (4.2) *with $\hat{K} = K^1 \times \{0_{n_2}\}$. Then it is $\Omega = \hat{\mathscr{M}}$.*

Proof

$$
\begin{aligned}
&\text{We have}\quad (\bar{x},\bar{y})\in\Omega\\
&\quad\Leftrightarrow \bar{x}\in\mathscr{M}_{\bar{y}}(f(G),K^1)\quad\wedge\quad \bar{y}\in\tilde{G}\\
&\quad\Leftrightarrow \Big(\nexists\ \ x\in G(\bar{y}) \text{ with } f(\bar{x},\bar{y})\in f(x,\bar{y})+K^1\backslash\{0_{m_1}\}\Big)\\
&\qquad\wedge\ \ \bar{y}\in\tilde{G}\quad\wedge\quad \bar{x}\in G(\bar{y})\\
&\quad\Leftrightarrow \Big(\nexists\ \ (x,y)\in G \text{ with } f(\bar{x},\bar{y})\in f(x,y)+K^1\backslash\{0_{m_1}\}\ \wedge\quad y=\bar{y}\Big)\\
&\qquad\wedge\ \ \bar{y}\in\tilde{G}\quad\wedge\quad (\bar{x},\bar{y})\in G\\
&\quad\Leftrightarrow (\nexists\ \ (x,y)\in G \text{ with}\\
&\qquad\begin{pmatrix} f(\bar{x},\bar{y})\\ \bar{y}\end{pmatrix}\in\begin{pmatrix} f(x,y)\\ y\end{pmatrix}+(K^1\times\{0_{n_2}\})\backslash\{0_{m_1+n_2}\}\Big)\\
&\qquad\wedge\ \ \bar{y}\in\tilde{G}\quad\wedge\quad (\bar{x},\bar{y})\in G\\
&\quad\Leftrightarrow \Big(\nexists(x,y)\in G \text{ with } \hat{f}(\bar{x},\bar{y})\in\hat{f}(x,y)+\hat{K}\backslash\{0_{m_1+n_2}\}\Big)\\
&\qquad\wedge\ \ \bar{y}\in\tilde{G}\quad\wedge\quad (\bar{x},\bar{y})\in G\\
&\quad\Leftrightarrow (\bar{x},\bar{y})\in\hat{\mathscr{M}}.
\end{aligned}
$$

□

Hence, if we are able to determine the solution set of the multiobjective optimization problem (4.2), we have already the set of feasible points of the upper level problem which we can solve then. Thus the upper level problem is reduced to $\min_{x,y}\{F(x,y) \mid (x,y)\in\hat{\mathscr{M}}\}$.

Generally, we cannot determine the whole solution set of problem (4.2) but we can calculate an approximation of this set and then, based on sensitivity information, we can refine this approximation depending on the behavior of the upper level function. For determining single solution points of problem (4.2) we use the scalarization according to Pascoletti and Serafini as discussed in Sect. 2. Thus we consider the scalarization problem (SP($\hat{a},\hat{r}$))

$$
\begin{aligned}
&\min_{t,x,y}\ t\\
&\text{subject to the constraints}\\
&\hat{a}+t\,\hat{r}-\hat{f}(x,y)\in\hat{K},\\
&(x,y)\in G,\\
&y\in\tilde{G},\\
&t\in\mathbb{R}
\end{aligned}
\qquad (\mathrm{SP}(\hat{a},\hat{r}))
$$

with $\hat{a}\in\mathbb{R}^{m_1+n_2}$, $\hat{r}\in\hat{K}=K^1\times\{0_{n_2}\}$. In the following we assume, for getting an easier notation, $e_{m_1}\in K^1$ with e_{m_1} the m_1-th unit vector in $\mathbb{R}^{m_1}$. This is, e.g., satisfied for $K^1=\mathbb{R}^{m_1}_+$, i.e., for the natural ordering. Then it is sufficient (see Theorem 2.5) to consider only parameters

$$\hat{a} := \begin{pmatrix} a \\ \tilde{a} \end{pmatrix} \in \hat{H} = \{x \in \mathbb{R}^{m_1+n_2} \mid x_{m_1} = 0\},\ \hat{r} = \begin{pmatrix} r \\ 0_{n_2} \end{pmatrix} \text{ with } r = e_{m_1} \quad (4.3)$$

and with $a \in \mathbb{R}^{m_1}$ and $\tilde{a} \in \mathbb{R}^{n_2}$.

In the following theorem we show that for these parameters and $\tilde{a} \in \tilde{G}$, a point $(\bar{t}, \bar{x}, \bar{y})$ is a minimal solution of (SP($\hat{a}, \hat{r}$)) with $\bar{y} = \tilde{a}$ if $(\bar{t}, \bar{x})$ is a minimal solution of the simplified problem (SP($a, r, \tilde{a}$)) defined by

$$\begin{aligned} &\min_{t,x}\ t \\ &\text{subject to the constraints} \\ &a + t\,r - f(x, \tilde{a}) \in K^1, \\ &(x, \tilde{a}) \in G, \\ &t \in \mathbb{R} \end{aligned} \qquad (\text{SP}(a, r, \tilde{a}))$$

with $a \in H := \{x \in \mathbb{R}^{m_1} \mid x_{m_1} = 0\}, \tilde{a} \in \tilde{G}$, and $r = e_{m_1}$. Problem (SP($\hat{a}, \hat{r}$)) has, for the parameter $\hat{r}$ as in (4.3), no minimal solution for $\tilde{a} \notin \tilde{G}$. Thus it is only interesting to consider problem (SP($\hat{a}, \hat{r}$)) for $\hat{a} = (a, \tilde{a})$ with $\tilde{a} \in \tilde{G}$. Then we can ignore the constraint $y \in \tilde{G}$ and we can consider the simplified problem (SP($a, r, \tilde{a}$)) instead of (SP($\hat{a}, \hat{r}$)). This is shown in the following theorem. This connection between the two scalarizations will be important for the application of the sensitivity results in the following section.

Note that the simplified problem (SP($a, r, \tilde{a}$)) is just the Pascoletti–Serafini scalarization applied to problem (3.1) for $y = \tilde{a}$. In the following, for an arbitrary cone K, we denote by K^* its dual cone.

Theorem 4.2 *We consider the optimization problems* (SP($\hat{a}, \hat{r}$)) *and* (SP($a, r, \tilde{a}$)) *with*

$$G := \{(x, y) \in \mathbb{R}^{n_1+n_2} \mid g(x, y) \in C\}$$

with continuously differentiable functions $f : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^{m_1}$ and $g : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^p$, $p \in \mathbb{N}$, and $C \subset \mathbb{R}^p$ a convex cone. Let the point $(\bar{t}, \bar{x})$ be a minimal solution of (SP($a, r, \tilde{a}$)) *with $a \in H$, $r = e_{m_1}$, and $\tilde{a} \in \tilde{G}$, with Lagrange multipliers $\mu \in (K^1)^*$ to the constraint $a + t\,r - f(x, \tilde{a}) \in K^1$ and $\nu \in C^*$ to the constraint $g(x, y) \in C$. Then $(\bar{t}, \bar{x}, \tilde{a})$ is a minimal solution of* (SP($\hat{a}, \hat{r}$)) *with $\hat{a}$ and $\hat{r}$ as in* (4.3) *and Lagrange multipliers $(\mu, \tilde{\mu}) \in \hat{K}^* = (K^1)^* \times \mathbb{R}^{n_2}$ with*

$$\tilde{\mu} = -\sum_{i=1}^{m_1} \mu_i \nabla_y f_i(\bar{x}, \tilde{a}) + \sum_{j=1}^{p} \nu_j \nabla_y g_j(\bar{x}, \tilde{a}) \qquad (4.4)$$

to the constraint $\hat{a} + t\,\hat{r} - \hat{f}(x, y) \in \hat{K}$ and $\nu \in C^*$ the Lagrange multiplier to the constraint $g(x, y) \in C$.

Proof Let $(\bar{t}, \bar{x})$ be a minimal solution of (SP($a, r, \tilde{a}$)). Then we have for the related Lagrange function $\mathscr{L}(\bar{t}, \bar{x}, \mu, \nu) = \bar{t} - \mu^\top (a + \bar{t}\,r - f(\bar{x}, \tilde{a})) - \nu^\top g(\bar{x}, \tilde{a})$ and

$$\nabla_{(t,x)}\mathscr{L}(\bar{t},\bar{x},\mu,\nu) = \begin{pmatrix} 1 \\ 0_{n_1} \end{pmatrix} + \sum_{i=1}^{m_1-1} \mu_i \begin{pmatrix} 0 \\ \nabla_x f_i(\bar{x},\tilde{a}) \end{pmatrix} + \mu_{m_1} \begin{pmatrix} -1 \\ \nabla_x f_{m_1}(\bar{x},\tilde{a}) \end{pmatrix} - \sum_{j=1}^{p} \nu_j \begin{pmatrix} 0 \\ \nabla_x g_j(\bar{x},\tilde{a}) \end{pmatrix} = 0_{n_1+1}. \tag{4.5}$$

We consider now the constraint $\hat{a} + t\,\hat{r} - \hat{f}(x,y) \in \hat{K}$ of problem (SP($\hat{a},\hat{r}$)) with $\hat{K} = K^1 \times \{0_{n_2}\}$. This constraint is equivalent to the constraints $a + t\,r - f(x,y) \in K^1$ and $\tilde{a} - y \in \{0_{n_2}\}$. Hence it follows $\tilde{a} = y$ and we conclude for $\tilde{a} \in \tilde{G}$ immediately that $(\bar{t},\bar{x},\bar{y})$ with $\bar{y} = \tilde{a}$ is a minimal solution of (SP($\hat{a},\hat{r}$)). The Lagrange function $\hat{\mathscr{L}}$ to problem (SP($\hat{a},\hat{r}$)) is defined by (with $\hat{\nu} \in \hat{K}^*$, $\hat{\mu} \in C^*$)

$$\hat{\mathscr{L}}(\bar{t},\bar{x},\bar{y},\hat{\mu},\hat{\nu}) = \bar{t} - \hat{\mu}^\top (\hat{a} + \bar{t}\,\hat{r} - \hat{f}(\bar{x},\bar{y})) - \hat{\nu}^\top g(\bar{x},\bar{y})$$

and thus together with $\bar{y} = \tilde{a}$ and $\hat{f}_{m_1+i}(x,y) = y_i,\ i = 1,\ldots,n_2$, we get

$$\nabla_{(t,x,y)}\hat{\mathscr{L}}(\bar{t},\bar{x},\bar{y},\hat{\mu},\hat{\nu}) = \begin{pmatrix} 1 \\ 0_{n_1} \\ 0_{n_2} \end{pmatrix} + \sum_{i=1}^{m_1-1} \hat{\mu}_i \begin{pmatrix} 0 \\ \nabla_x f_i(\bar{x},\tilde{a}) \\ \nabla_y f_i(\bar{x},\tilde{a}) \end{pmatrix} + \hat{\mu}_{m_1} \begin{pmatrix} -1 \\ \nabla_x f_{m_1}(\bar{x},\tilde{a}) \\ \nabla_y f_{m_1}(\bar{x},\tilde{a}) \end{pmatrix} + \sum_{i=1}^{n_2} \hat{\mu}_{m_1+i} \begin{pmatrix} 0 \\ 0_{n_1} \\ e_i \end{pmatrix} - \sum_{j=1}^{p} \hat{\nu}_j \begin{pmatrix} 0 \\ \nabla_x g_j(\bar{x},\tilde{a}) \\ \nabla_y g_j(\bar{x},\tilde{a}) \end{pmatrix}$$

with e_i the i-th unit vector in $\mathbb{R}^{n_2}$. With (4.5) we get for $\hat{\mu} = (\mu,\tilde{\mu})$ and $\hat{\nu} = \nu$

$$\nabla_{(t,x)}\hat{\mathscr{L}}(\bar{t},\bar{x},\bar{y},(\mu,\tilde{\mu}),\nu) = 0_{n_1+1}$$

and by setting $\tilde{\mu}$ as in (4.4) we conclude

$$\nabla_y\hat{\mathscr{L}}(\bar{t},\bar{x},(\mu,\tilde{\mu}),\nu) = \sum_{i=1}^{m_1} \mu_i \nabla_y f_i(\bar{x},\tilde{a}) + \sum_{i=1}^{n_2} \tilde{\mu}_i\, e_i - \sum_{j=1}^{p} \nu_j \nabla_y g_j(\bar{x},\tilde{a}) = 0_{n_2}.$$

Further, as μ and ν are Lagrange multipliers to problem (SP($a,r,\tilde{a}$)), it holds

$$\mu^\top(a + \bar{t}\,r - f(\bar{x},\tilde{a})) = 0 \text{ and } \nu^\top g(\bar{x},\tilde{a}) = 0.$$

For the problem (SP($\hat{a},\hat{r}$)) it is for $\hat{\mu} = (\mu,\tilde{\mu})$ because of $\tilde{a} = \bar{y}$

$$\hat{\mu}^\top(\hat{a} + \bar{t}\,\hat{r} - \hat{f}(\bar{x},\bar{y})) = \mu^\top(a + \bar{t}\,r - f(\bar{x},\bar{y})) + \tilde{\mu}^\top(\tilde{a} - \bar{y}) = 0$$

and thus $\hat{\mu} = (\mu, \tilde{\mu})$ and ν are Lagrange multipliers to the point $(\bar{t}, \bar{x}, \bar{y}) = (\bar{t}, \bar{x}, \tilde{a})$ for the problem (SP($\hat{a}, \hat{r}$)), too. □

Of course equality constraints for the set G can be included also. If we are searching for solutions of the problem (SP($\hat{a}, \hat{r}$)) with $\hat{a} = (a, \tilde{a}) \in \hat{H}$ and $\hat{r} = (r, 0_{n_2})$ it is therefore satisfactory to solve (SP(a, r, y)) for $a \in H$, $r = e_{m_1}$, and $y = \tilde{a} \in \tilde{G}$.

For determining an approximation of the solution set of (4.2) we proceed as follows: We consider problems (SP($\hat{a}, \hat{r}$)) for a choice of parameters $\hat{a}, \hat{r}$ as in (4.3) with $\tilde{a} \in \tilde{G}$. Instead of solving problem (SP($\hat{a}, \hat{r}$)) directly we switch to the simplified problem (SP($a, r, \tilde{a}$)) with parameters $a \in H$, $r = e_{m_1}$, and $\tilde{a} \in \tilde{G}$. The aim is to cover the whole solution set of problem (4.2). For achieving this we discretize the set $\tilde{G}$ with equal distances, e.g., for $\tilde{G} = [c, d] \subset \mathbb{R}$ by $y^1 := c \leq y^2 := y^1 + \beta \leq y^3 := y^1 + 2\beta \leq \ldots \leq y^{n^y} := y^1 + (n^y - 1)\beta \leq d$ ($\beta \in \mathbb{R}_+$, $n^y \in \mathbb{N}$), and solve for any $\tilde{a} = y^k$ of this discretization the problem (SP($a, r, \tilde{a}$)) for a variation of the remaining parameters a and r. This is equivalent to determine the solution set of the lower level problem (3.1) for the parameter $y = y^k$. For the remaining parameters a and r we have seen that it is sufficient to choose the parameter r constant and to choose the parameter a from a particular hyper plane (see Theorem 2.5).

Instead of choosing the parameter a, e.g., equidistantly from the hyper plane (and thereby having no control over the distribution of the found points in the set Ω) we get more evenly spread approximation points of Ω by using sensitivity information for an adaptive controlled choice. Based on these sensitivity results, we try to control the choice of the parameter a adaptively in such a way that for the minimal solutions $(t, x) = (t(a), x(a))$ found (dependent on the parameter a) we achieve a predefined (equal) distance between the points $x(a)$ for different values of a. As the points $(x(a), y^k)$ are points of the feasible set Ω of the upper level problem we achieve a controlled approximation of this set with almost equidistant points w.r.t. x. Such a representative approximation of Ω is important for avoiding to neglect parts of the feasible set and thereby for assuring that the approximation points of the solution set of the original bilevel problem found are in fact close to the real solution set.

In what follows, we describe such sensitivity information for the case $K^1 = \mathbb{R}^{m_1}_+$ and $G = \{(x, y) \in \mathbb{R}^{n_1+n_2} \mid g_j(x, y) \geq 0,\ j = 1, \ldots, p\}$ (i.e., $C = \mathbb{R}^p_+$). We define for $y = \tilde{a}$ the index sets $I := \{1, \ldots, m_1\}$ and $J := \{1, \ldots, p\}$ which can be split in the disjunct sets $I = I^+ \cup I^0 \cup I^-$ and $J = J^+ \cup J^0 \cup J^-$ by

$$\begin{aligned} I^+ &= \{i \in I \mid a_i + t\, r_i - f_i(x, \tilde{a}) = 0, \quad \mu_i > 0\}, \\ I^0 &= \{i \in I \mid a_i + t\, r_i - f_i(x, \tilde{a}) = 0, \quad \mu_i = 0\}, \\ I^- &= \{i \in I \mid a_i + t\, r_i - f_i(x, \tilde{a}) > 0, \quad \mu_i = 0\} \end{aligned}$$

and

$$\begin{aligned} J^+ &= \{j \in J \mid g_j(x, \tilde{a}) = 0, \quad \nu_j > 0\}, \\ J^0 &= \{j \in J \mid g_j(x, \tilde{a}) = 0, \quad \nu_j = 0\}, \\ J^- &= \{j \in J \mid g_j(x, \tilde{a}) > 0, \quad \nu_j = 0\} \end{aligned} \tag{4.6}$$

subject to x and μ and ν respectively, with μ and ν Lagrange multipliers to the particular constraint.

Theorem 4.3 *We consider problem* (SP$(a, r, \tilde{a})$) *for* $K^1 = \mathbb{R}^{m_1}_+$ *and* $G = \{(x, y) \in \mathbb{R}^{n_1+n_2} | g_j(x, y) \geq 0,\ j = 1, \ldots, p\}$. *Let* $\tilde{a} \in \tilde{G}$ *be given and let the functions* f *and* g *be twice continuously differentiable. Let the sets* $I^+, I^0, I^-, J^+, J^0, J^-$ *be defined as in (4.6) and assume non-degeneracy, i.e.,* $I^0 = J^0 = \emptyset$. *Assume* (t^0, x^0) *is a minimal solution of the so-called reference problem, i.e., of* (SP$(a, r, \tilde{a})$) *with parameters* $(a, r) = (a^0, r^0)$. *Further assume*:

(a) *Let the gradients w.r.t.* (t, x) *of the constraints being active in the point* (t^0, x^0) *be linearly independent, i.e., let*

$$\begin{pmatrix} r_i^0 \\ -\nabla_x f_i(x^0, \tilde{a}) \end{pmatrix}, \quad i \in I^+, \quad \begin{pmatrix} 0 \\ \nabla_x g_j(x^0, \tilde{a}) \end{pmatrix}, \quad j \in J^+$$

be linearly independent.

(b) *Assume there is some constant* $\xi > 0$, *such that for the Hessian of the Lagrange function* $\mathscr{L}$ *in the point* (t^0, x^0) *it is*

$$(t, x^\top)\, \nabla^2_{(t,x)} \mathscr{L}(t^0, x^0, \mu^0, \nu^0, a^0, r^0) \begin{pmatrix} t \\ x \end{pmatrix} \geq \xi \left\| \begin{pmatrix} t \\ x \end{pmatrix} \right\|^2$$

for all $(t, x) \in \{(t, x) \in \mathbb{R}^{n_1+1} \mid r_i^0 t = \nabla_x f_i(x^0, \tilde{a})^\top x,\ \forall i \in I^+, \nabla_x g_j(x^0, \tilde{a})^\top x = 0,\ \forall j \in J^+\}$.

Then (t^0, x^0) *is a local unique minimal solution of* (SP$(a^0, r^0, \tilde{a})$) *with unique Lagrange multipliers* (μ^0, ν^0) *and there is a* $\delta > 0$ *and a neighborhood* $N(a^0, r^0)$ *of* (a^0, r^0) *such that the function* $\phi\colon N(a^0, r^0) \to B_\delta(t^0, x^0) \times B_\delta(\mu^0, \nu^0)$ *with*

$$\phi(a, r) = (t(a, r), x(a, r), \mu(a, r), \nu(a, r))$$

is Lipschitzean on $N(a^0, r^0)$ (*with* $B_\delta(\cdot)$ *a closed ball around the particular point*). *Here* $(t(a, r), x(a, r))$ *denotes the local unique minimal solution of* (SP$(a, r, \tilde{a})$) *with* $\mu(a, r)$ *and* $\nu(a, r)$ *the related unique Lagrange multipliers for* $(a, r) \in N(a^0, r^0)$.

We have a first order approximation of $\phi(a, r)$ *with* (a, r) *from a neighborhood of* (a^0, r^0) *by*

$$\phi(a, r) = \phi(a^0, r^0) + M^{-1} N \begin{pmatrix} a - a^0 \\ r - r^0 \end{pmatrix} + o\left(\left\| \begin{pmatrix} a - a^0 \\ r - r^0 \end{pmatrix} \right\| \right)$$

with $M = [M_1 \mid M_2] \in \mathbb{R}^{(1+n_1+m_1+p)} \times \mathbb{R}^{(1+n_1+m_1+p)}$,

$$M_1 = \begin{bmatrix} 0 & 0_{n_1}^{\top} & -r_1 & \dots & -r_{m_1} \\ 0_{n_1} & \nabla_x^2 \mathscr{L}(\phi(a^0, r^0), a^0, r^0) & \nabla_x f_1(x^0, \tilde{a}) & \dots & \nabla_x f_{m_1}(x^0, \tilde{a}) \\ \mu_1^0 r_1 & -\mu_1^0 \nabla_x f_1(x^0, \tilde{a})^{\top} & k_1^0 & 0 \dots 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_{m_1}^0 r_{m_1} & -\mu_{m_1}^0 \nabla_x f_{m_1}(x^0, \tilde{a})^{\top} & 0 & 0 \dots 0 & k_{m_1}^0 \\ 0 & \nu_1^0 \nabla_x g_1(x^0, \tilde{a})^{\top} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \nu_p^0 \nabla_x g_p(x^0, \tilde{a})^{\top} & 0 & \dots & 0 \end{bmatrix}$$

and

$$M_2 = \begin{bmatrix} 0 & \dots & 0 \\ -\nabla_x g_1(x^0, \tilde{a}) & \dots & -\nabla_x g_p(x^0, \tilde{a}) \\ 0_{m_1} & \dots & 0_{m_1} \\ g_1(x^0, \tilde{a}) & 0 \dots 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 \dots 0 & g_p(x^0, \tilde{a}) \end{bmatrix}$$

with $k^0 = a^0 + t^0 r^0 - f(x^0, \tilde{a}) \in \mathbb{R}^{m_1}$ *and*

$$N = \left[\mathbf{0}_{2m_1 \times (n_1+1)}, -\mu_1^0 \begin{pmatrix} e_1 \\ t^0 e_1 \end{pmatrix}, \dots, -\mu_{m_1}^0 \begin{pmatrix} e_{m_1} \\ t^0 e_{m_1} \end{pmatrix}, \mathbf{0}_{2m_1 \times p} \right]^{\top}$$

($N \in \mathbb{R}^{(1+n_1+m_1+p)\times(2m_1)}$*) with* $\mathbf{0}_{2m_1\times(n_1+1)}$ *and* $\mathbf{0}_{2m_1\times p}$ *the matrices in* $\mathbb{R}^{2m_1\times(n_1+1)}$ *and in* $\mathbb{R}^{2m_1\times p}$ *respectively having only zero entries, and* $e_i,\ i = 1, \dots, m_1$*, the* i*-th unit vector in* $\mathbb{R}^{m_1}$*.*

Equality constraints can be included as well, see Theorem 3.2.5 in [12]. The result that the function ϕ is Lipschitzean follows from a theorem by Alt ([2], Theorems 5.3 and 6.1, see also Theorem 4.2 in [14]). The further conclusions are an application of a theorem by Fiacco [15, Cor. 3.2.4] applied to problem (SP($a, r, \tilde{a}$)). Here it is assumed that the constraints are non-degenerated, because otherwise the matrix M is not invertible. In the degenerate case we can apply a result by Jittorntrum [23, Theorems 3 and 4]. For doing this we have to define in which direction we want to vary the parameters. Let $v = (v^a, v^r) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_1}$ be such a direction, hence let

$$\begin{pmatrix} a \\ r \end{pmatrix} = \begin{pmatrix} a^0 \\ r^0 \end{pmatrix} + s\, v \quad \text{for } s \in \mathbb{R},\ s \geq 0. \tag{4.7}$$

Then, a directional derivative of the function ϕ in direction v is available, even in the degenerate case:

Theorem 4.4 *Let the assumptions of Theorem 4.3 be satisfied but without assuming $I^0 = J^0 = \emptyset$, i.e., allowing degenerate constraints. Let (a, r) be given as in (4.7). Then the directional derivatives are given by*

$$\lim_{s \to 0^+} \begin{pmatrix} \frac{t(s)-t(0)}{s} \\ \frac{x(s)-x(0)}{s} \\ \frac{\mu(s)-\mu(0)}{s} \\ \frac{\nu(s)-\nu(0)}{s} \end{pmatrix} = \begin{pmatrix} \bar{t} \\ \bar{x} \\ \bar{\mu} \\ \bar{\nu} \end{pmatrix}$$

with $(t(s), x(s), \mu(s), \nu(s)) = (t(a,r), x(a,r), \mu(a,r), \nu(a,r))$ the minimal solution and the Lagrange multipliers of problem (SP$(a, r, \tilde{a})$) *for the parameters $(a, r) = (a^0 + s\, v^a, r^0 + s\, v^r)$. Thereby $(\bar{t}, \bar{x}, \bar{\mu}, \bar{\nu})$ is the unique solution of the system of equations and inequalities given by*:

$$-\sum_{i=1}^{m_1} \bar{\mu}_i\, r_i^0 = \mu^{0\top} v^r,$$

$$\sum_{i=1}^{m_1} \mu_i^0 \nabla_x^2 f_i(x^0, \tilde{a})\bar{x} - \sum_{j=1}^{p} \nu_j^0 \nabla_x^2 g_j(x^0, \tilde{a})\bar{x} + \sum_{i=1}^{m_1} \bar{\mu}_i \nabla_x f_i(x^0, \tilde{a})$$
$$- \sum_{j=1}^{p} \bar{\nu}_j \nabla_x g_j(x^0, \tilde{a}) = 0_{n_1},$$

$$r_i^0\, \bar{t} - \nabla_x f_i(x^0, \tilde{a})^\top \bar{x} = -v_i^a - t^0\, v_i^r, \qquad \forall i \in I^+,$$

$$r_i^0\, \bar{t} - \nabla_x f_i(x^0, \tilde{a})^\top \bar{x} \geq -v_i^a - t^0\, v_i^r, \qquad \bar{\mu}_i \geq 0, \qquad \forall i \in I^0,$$

$$\bar{\mu}_i \left(r_i^0\, \bar{t} - \nabla_x f_i(x^0, \tilde{a})^\top \bar{x} + v_i^a + t^0\, v_i^r \right) = 0, \qquad \forall i \in I^0,$$

$$\bar{\mu}_i = 0, \qquad \forall i \in I^-,$$

$$\nabla_x g_j(x^0, \tilde{a})^\top \bar{x} = 0, \qquad \forall j \in J^+,$$

$$\nabla_x g_j(x^0, \tilde{a})^\top \bar{x} \geq 0, \qquad \bar{\nu}_j \geq 0, \qquad \forall j \in J^0,$$

$$\bar{\nu}_j \left(\nabla_x g_j(x^0, \tilde{a})^\top \bar{x} \right) = 0, \qquad \forall j \in J^0,$$

$$\bar{\nu}_j = 0, \qquad \forall j \in J^-.$$

This derivative information can be used for a first order approximation of the minimal solutions. Because the degenerate case is allowed here, equality constraints can be included as well.

We will use these results for problems of type (SP$(a, r, \tilde{a})$) with $\tilde{a} = y^k$ for determining an approximation of the minimal solution set of problem (3.1) with almost equal distances between the solutions x by choosing appropriate parameters $a \in H$. We discuss this procedure in the following section. In a second step, for refining our approximation of the set of feasible points in the course of an iterative solution process, we need again a sensitivity result for the scalar optimization problem (SP($\hat{a}, \hat{r}$)) similar to the one in Theorem 4.3, which will be stated in the following section.

5 Numerical methods

In this section we describe an algorithm for solving bilevel multiobjective optimization problems based on the results in the preceding sections. We first present an adaptive parameter control for the scalarization problem (SP($a, r, \tilde{a}$)). The aim is an equidistant approximation of the set of minimal solutions of the multiobjective optimization problems (3.1) for $\tilde{a} = y^k$. Based on this parameter control, for several discretization points $y^k \in \tilde{G}$ such approximations of the minimal solution sets are determined and are then unified to an approximation of the minimal solution set of the multiobjective optimization problem (4.2). This is the first step of our proposed method which will be discussed in the first part of this section. An algorithm for this first part is given in Algorithm 1.

As a result we have according to Theorem 4.1 an approximation of the set of feasible points of the upper level problem, too. This approximation is then mapped by the upper level function and the non-dominated points w.r.t. the ordering cone of the upper level are determined. These points are an approximation of the minimal solutions of the original bilevel optimization problem. For improving the accuracy of this approximation and thus for producing a better representation of the efficient set of the upper level, the approximation of the set of feasible points is refined around these points. Based on sensitivity results we can refine the set of feasible points with a given fineness and only in the interesting parts of the set of feasible points.

This is described in the second part of this section and the procedure is summarized in Algorithm 2.

Hence, we have the following three main steps:

(1) Determine equidistant approximations of the solution sets of the multiobjective problems (3.1) for various discretization points y in the set $\tilde{G}$.
(2) Select all non-dominated points of the lower level problems being non-dominated for the upper level problem.
(3) Refine the approximation of the efficient set of the original problem by refining the approximation of the set of feasible points of the upper level problem.

We start by describing the first of these three steps in detail. For our numerical method we assume that all necessary assumptions for applying the theoretical results of Sect. 4 are at least locally satisfied. As mentioned in the introduction of this section our aim is an approximation $x^1, x^2, \ldots, x^i, x^{i+1}, \ldots$ of the set of minimal solutions of problem (3.1) such that for a given distance $\alpha > 0$ we have

$$\|x^{i+1} - x^i\| \approx \alpha. \tag{5.1}$$

We restrict ourselves here to the bicriteria case on the lower level, i.e., to $m_1 = 2$. This simplifies and shortens the presentation because then, for instance, the hyper plane H is just a line allowing only one direction for a variation of the parameter a. Of course, the results can be generalized to higher dimensions.

For determining approximation points $x^1, x^2, \ldots, x^i, x^{i+1}, \ldots$ of problem (3.1) we use the scalarization (SP($a, r, \tilde{a}$)). We vary only the parameter a in a direction v and choose the parameter $r = r^0$ constant. Then, applying the results of Theorem 4.3 (assuming non-degeneracy, otherwise Theorem 4.4), we get with

$$a = \begin{pmatrix} a_1^0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in H$$

($\lambda > 0$) and $r = r^0$:

$$\phi(a, r) \approx \phi(a^0, r) + M^{-1} N \begin{pmatrix} a - a^0 \\ 0_2 \end{pmatrix}$$

and thus

$$\begin{pmatrix} t(a, r) \\ x(a, r) \\ \mu(a, r) \\ \nu(a, r) \end{pmatrix} \approx \begin{pmatrix} t^0 \\ x^0 \\ \mu^0 \\ \nu^0 \end{pmatrix} + \lambda\, M^{-1} N \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

with (t^0, x^0) the minimal solution of the reference problem with Lagrange multipliers μ^0, ν^0. For $\tilde{v} = (1, 0, 0, 0)^\top$ let $(M^{-1} N \tilde{v})\,|_x$ be the second to the $(n_1 + 1)$-th entry of the vector $M^{-1} N \cdot \tilde{v}$ (corresponding to the entries related to x). Then we get

$$\begin{aligned} \|x - x^0\| &= \|x(a, r) - x(a^0, r)\| \\ &= \|x(a^0 + \lambda\,(1, 0)^\top, r) - x(a^0, r)\| \\ &\approx \|x^0 + \lambda (M^{-1} N \tilde{v})\Big|_x - x^0\| \\ &= |\lambda|\, \|(M^{-1} N \tilde{v})\Big|_x\|. \end{aligned}$$

Thus we have (5.1) e. g. for $i = 0$ approximately fulfilled for $a^1 = a^0 + \bar{\lambda} \cdot (1, 0)^\top$ with

$$\bar{\lambda} = \frac{\alpha}{\|(M^{-1} N \tilde{v})\,|_x\|}. \tag{5.2}$$

Having the parameter a^1 we can repeat this procedure for calculating a^2 and so on till an approximation of the solution set of problems (3.1) for $y = y^k$ is generated by an adaptive parameter control. This has to be done for all discretization points y^k of $\tilde{G}$. Then, for $(\bar{t}, \bar{x})$ a minimal solution of (SP($a, r, \tilde{a}$)) with $\tilde{a} = y^k$ the point $(\bar{t}, \bar{x}, y^k)$ is according to Theorem 4.2 a minimal solution of the scalarization problem (SP($\hat{a}, \hat{r}$)) and thus $(\bar{x}, y^k)$ is a (weakly) $\hat{K}$-minimal solution of (4.2). Then $(\bar{x}, y^k)$ is an approximation point of the set $\hat{\mathscr{M}}$ and hence of Ω.

Now we can summarize the described method for determining a first approximation of the set of feasible points of the upper level problem. For that we assume

$K^1 = \mathbb{R}^2_+,\ m_1 = 2,\ C = \mathbb{R}^p_+,\ n_2 = 1,\ G = \{(x, y) \in \mathbb{R}^{n_1+1} \mid g_j(x, y) \geq 0,\ j = 1, \ldots, p\}$ and $\tilde{G} = [c, d] \subset \mathbb{R}$ for $c, d \in \mathbb{R}$ and we choose $r = (0, 1)^\top$ constant. Additionally we have to guarantee that the scalar optimization problems in Step 2a are solvable. Here we can assume for simplicity that the set G is nonempty and compact, but of course less restrictive assumptions are possible, too.

Algorithm 1 (Bilevel algorithm: Part 1)

Step 1: Choose $\beta > 0$ (or $n^y \in \mathbb{N}$) and discretize the interval $\tilde{G} = [c, d]$ by $y^1 := c,\ y^2 := y^1 + \beta,\ y^3 := y^1 + 2\beta, \ldots, y^{n^y}$ with $y^{n^y} \leq d,\ n^y \in \mathbb{N}$.

Step 2: Define an accuracy measure $\alpha > 0$. Repeat the following steps for all $y = y^k, k = 1, \ldots, n^y$ with $\tilde{a} := y^k$.

Step 2a: Solve $f_1(\bar{x}^1, y^k) := \min_x\{f_1(x, y^k) \mid (x, y^k) \in G\}$ and $f_2(\bar{x}^2, y^k) := \min\{f_2(x, y^k) \mid (x, y^k) \in G\}$ and determine $a^1 = (f_1(\bar{x}^1, y^k), 0)$ and $a^E = (f_1(\bar{x}^2, y^k), 0)$. Set $A^{0,k} := \{(\bar{x}^1, y^k)\}$, $a^2 = (f_1(\bar{x}^1, y^k) + \delta, 0)^\top$ for a small $\delta > 0$ and $l := 2$.

Step 2b: If $a_1^l \leq a_1^E$ solve problem $(\mathrm{SP}(a^l, r, \tilde{a}))$ for $\tilde{a} = y^k,\ r = (0, 1)^\top$ with minimal solution x^l and Lagrange multipliers (μ^l, ν^l) and set $A^{0,k} := A^{0,k} \cup \{(x^l, y^k)\}$. Calculate the matrices M, N according to Theorem 4.3, determine $\bar{\lambda}$ according to (5.2) with $\tilde{v} = (1, 0, 0, 0)^\top$ and set $a^{l+1} := a^l + \bar{\lambda} \cdot (1, 0)^\top,\ l := l + 1$ and repeat Step 2b. Else stop.

For Algorithm 1 we have assumed $n_2 = 1$, i.e., the compact set $\tilde{G} = [c, d]$ is an interval which is discretized in Step 1. Also the case $n_2 \geq 2$ could be treated. Then an (equidistant) discretization of the compact set $\tilde{G}$ has to be found. This can be done e.g. by finding a cuboid $I := [c_1, d_1] \times \cdots \times [c_{n_2}, d_{n_2}] \subset \mathbb{R}^{n_2}$ with $\tilde{G} \subset I$ and discretizing this cuboid with points

$$\{y \in \mathbb{R}^{n_2} \mid y_i = c_i + n_i\, \beta_i \text{ for a } n_i \in \mathbb{N} \cup \{0\},\ y_i \leq d_i,\ i = 1, \ldots, n_2\}$$

for distances $\beta_i > 0$ $(i = 1, \ldots, n_2)$ and then selecting only the discretization points that are elements of $\tilde{G}$.

In Step 2 we determine an approximation $A^{0,k}$ of problem (3.1) for all $y = y^k$, $k = 1, \ldots, n^y$, with the help of problem $(\mathrm{SP}(a, r, \tilde{a}))$. In Step 2a it is necessary to start with a^2, a slight mutation of the parameter a^1, to avoid numerical difficulties, because for a^1 the set of feasible points of $(\mathrm{SP}(a, r, \tilde{a}))$ might consist of one point only. With the union of the sets $A^{0,k}$, i.e., with $A^0(\alpha, \beta) := \bigcup_{k=1}^{n^y} A^{0,k}$, an approximation of the solution set of the tricriteria optimization problem (4.2) and hence of the set of feasible points Ω is calculated. For $\alpha \to 0$ and $\beta \to 0$ we even get $A^0(0, 0) = \Omega$. This result is included in the following convergence proof.

Theorem 5.1 *For $\alpha, \beta > 0$ denote by*

$$A^0(\alpha, \beta) := \bigcup_{y^k \in D(\beta)} A^{0,k} \text{ with } D(\beta) := \{c, c + \beta, c + 2\beta, \ldots, y^{n^y}\} \subset [c, d]$$

the output of Algorithm 1 for $\delta = 0$ *and by* $M^0(\alpha, \beta) := \mathscr{M}(F(A^0(\alpha, \beta), K^2) \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ *the set of* K^2*-minimal points in this set w.r.t. the upper level function* F*. Further denote by* $M^0(0, 0) := \mathscr{M}(F(\Omega), K^2)$ *the solution set of the bilevel optimization problem* (4.1)*. Let all assumptions of Theorem 4.4 hold. Moreover, assume that for all* $y \in \tilde{G}$ *the set of weakly* K^1*-minimal points of problem* (3.1) *coincides with the set of* K^1*-minimal points, i.e.,*

$$\mathscr{M}_y^w(f(G), K^1) = \mathscr{M}_y(f(G), K^1), \tag{5.3}$$

and the same for problem (4.1)*, i.e.,* $\mathscr{M}^w(F(\Omega), K^2) = \mathscr{M}(F(\Omega), K^2)$*. Further assume for all* $y \in \tilde{G}$ *that the set-valued map* $y \mapsto G(y)$ *is continuous and uniformly compact and assume that* F *is continuous on* G*, that* f *is injective on* G*, and that* Ω *is nonempty and compact.*

Then, the function M^0 *is upper semicontinuous at* $(\alpha, \beta) = (0, 0)$*. Furthermore, if the function* F *is injective on* Ω*, then* M^0 *is also lower semicontinuous at* $(\alpha, \beta) = (0, 0)$.

Proof For $\alpha > 0$ and $y = y^k \in \tilde{G}$ we consider the set-valued mapping

$$B^0 : (\alpha, y) \mapsto A^{0,k}.$$

Here we have $A^{0,k} \subset \Psi(y) \times \{y\}$ due to (5.3). Let us define $B^0(0, y) := \Psi(y) \times \{y\}$. Under the assumption of (5.3) and under the assumptions of Theorem 4.4 we immediately get that the map $B^0(\cdot, y)$ is continuous at $\alpha = 0$. (Moreover, for a fixed value of y, the set-valued map $\alpha \mapsto f(B^0(\alpha, y))$ is continuous at $\alpha = 0$ because f is continuous.)

Next, define for $\alpha \geq 0, \beta \geq 0$ the set-valued map $A^0 \colon (\alpha, \beta) \mapsto \bigcup_{y^k \in D(\beta)} A^{0,k} = \bigcup_{y^k \in D(\beta)} B^0(\alpha, y^k)$, with $D(\beta) := \{c, c+\beta, c+2\beta, \ldots, y^{n^y}\} \subset [c, d]$ for $\beta > 0$ as defined in Step 1, and with $D(0) := [c, d]$. Note that $A^0(0, 0) = \Omega$. As the set-valued map $y \mapsto G(y)$ is continuous and uniformly compact, as the function f is continuous and because of (5.3), we have [34, Theorem 4.4.1], that $y \mapsto \Psi(y)$ is upper semicontinuous for all $y \in \tilde{G}$, and as f is additionally injective on G we also get [34, Theorem 4.4.2] that $y \mapsto \Psi(y)$ is lower semicontinuous for all $y \in \tilde{G}$. Hence the map A^0 is continuous at $(\alpha, \beta) = (0, 0)$. (And the set-valued map $(\alpha, \beta) \mapsto F(A^0(\alpha, \beta))$ exhibits the same characteristics as F is continuous.)

Finally, consider the set-valued map $M^0\colon (\alpha, \beta) \mapsto \mathscr{M}(F(A^0(\alpha, \beta)), K^2)$ and note that $M^0(0, 0) = \mathscr{M}(F(\Omega), K^2)$. With the same arguments as before we conclude that M^0 is upper semicontinuous at $(\alpha, \beta) = (0, 0)$ [34, Theorem 4.4.1]. And because the function F is also injective on Ω and the sets $\mathscr{M}(F(A^0(\alpha, \beta)), K^2)$ are externally stable, i.e.,

$$F(A^0(\alpha, \beta)) \subset \mathscr{E}(F(A^0(\alpha, \beta)), K^2) + K^2,$$

for (α, β) near $(0, 0)$, as Ω is nonempty and compact [34, Theorem 3.2.9], the map M^0 is also lower semicontinuous at $(\alpha, \beta) = (0, 0)$ [34, Theorem 4.4.2]. □

We continue by discussing the second and the third step of the three main steps given in the introduction of this section. Now, we can evaluate the points approximating the induced set Ω with the objective function F of the upper level. Hence for all points $(x, y) \in A^0 := \bigcup_{k=1}^{n^y} A^{0,k}$ we determine the image points $F(x, y)$. We have restricted ourselves in the problems of Sect. 6 to $m_2 = 2$ and $K^2 = \mathbb{R}^2_+$ because then it is easily possible to visualize the results in the image space of the upper level problem. Therefore we make here also these assumptions, but a generalization to higher dimensions or other ordering cones goes straightforward because just the K^2-minimal points of a finite set of points in the space $\mathbb{R}^{m_2}$ have to be determined.

We are only interested in the non-dominated points of the set $\{F(x, y) \mid (x, y) \in A^0\}$, hence in the points $(\bar{x}, \bar{y}) \in A^0$, such that there is no point $(x', y') \in A^0$ with $F(x', y') \neq F(\bar{x}, \bar{y})$ and with $F_i(x', y') \leq F_i(\bar{x}, \bar{y})$, $i = 1, 2$. We denote the set of non-dominated points as $\mathscr{M}(F(A^0), \mathbb{R}^2_+)$.

For a not too strict selection of points a better concept can be to select only the ε-EP-minimal points of the set A^0, because the ε-efficient points are also close to the efficient set, in which we are interested, and hence deliver useful information, too. Thus, a refinement around these points can deliver new points which are not only ε-EP-minimal but in fact EP-minimal. The determination of all non-dominated points can be very expensive if the set A^0 consists of many points. Then a so-called Pareto filter can be applied. The costs can be reduced by using the method of Graef and Younes, see [19, p. 14], [20, p. 337], [43], or, extended with a backward iteration, the Jahn–Graef–Younes method described in [21, p. 4f]. With the set $\mathscr{M}(F(A^0), \mathbb{R}^2_+)$ we have a first approximation of the solution set of the bilevel problem.

For improving this approximation of the solution set we refine the discretization of the set of feasible points in a neighborhood of these non-dominated points. We can do this for all points of the set $\mathscr{M}(F(A^0), \mathbb{R}^2_+)$, or, if the decision maker of the upper level problem is interested in some special chosen efficient points, only in their neighborhood. A third possibility is to consider only those points for which the approximation of the efficient set of the upper level problem is not accurate enough, i.e., only those points where the distance to the next neighbor point, mapped with the function F, is not small enough.

For obtaining a refinement of the set of feasible points around the chosen points we use again sensitivity information for determining the parameter a of the lower level scalarization in such a way that the refinement points have nearly a given equal distance. For any point $(x, y) \in \mathscr{M}(F(A^0), \mathbb{R}^2_+)$ there is some $t \in \mathbb{R}$ and some parameter a such that the point (t, x) is a minimal solution of (SP$(a, r, \tilde{a})$) for $\tilde{a} = y$ and according to Theorem 4.2 (t, x, y) is then a minimal solution of (SP$(\hat{a}, \hat{r})$) and the Lagrange multipliers are known, too, if the Lagrange multipliers of problem (SP$(a, r, \tilde{a})$) are given. Problem (SP$(\hat{a}, \hat{r})$) is the scalarization of problem (4.2). Any minimal solution (t, x, y) of (SP$(\hat{a}, \hat{r})$) is thus a (weakly) $\hat{K}$-minimal solution of (4.2). The point (x, y) is an approximation point of the set $\hat{\mathscr{M}}$ and according to Theorem 4.1 also of the set of feasible points Ω.

Therefore we are interested in the dependence of the minimal solutions of problem (SP$(\hat{a}, \hat{r})$) on the parameter $\hat{a}$. Doing this we include the variable y in our sensitivity and hence in our distance considerations. For $K^1 = \mathbb{R}^2_+$, $\hat{r} = (0, 1, 0)^\top \in K^1 \times \{0\}$ (i.e., $n_2 = 1$) and $\hat{a} = (a_1, a_2, a_3)^\top$ we can rewrite problem (SP($\hat{a}, \hat{r}$)) with the following equality and inequality constraints

$$\begin{aligned} &\min_{t,x,y}\ t \\ &\text{subject to the constraints} \\ &a_1 - f_1(x, y) \geq 0, \\ &a_2 + t - f_2(x, y) \geq 0, \\ &a_3 - y = 0, \\ &(x, y) \in G, \\ &y \in \tilde{G} = [c, d], \\ &t \in \mathbb{R} \end{aligned} \tag{5.4}$$

with $G = \{(x, y) \in \mathbb{R}^{n_1+1} | g_j(x, y) \geq 0,\ \ j = 1, \ldots, p\}$. Applying again Cor. 3.2.1 from [15] we get similar to Theorem 4.3 the following sensitivity results for the minimal solutions of problem (5.4).

Theorem 5.2 *We consider problem (5.4) with $f_1, f_2, g_j, j = 1, \ldots, p$, twice continuously differentiable. Let (t^0, x^0, y^0) be a local minimal solution of the reference problem (5.4) w.r.t. the parameter $\hat{a}^0 = (a_1^0, a_2^0, a_3^0)^\top$ with Lagrange multipliers $\mu_1^0, \mu_2^0 \in \mathbb{R}_+$, $\mu_3^0 \in \mathbb{R}$, $\nu^0 \in \mathbb{R}^p_+$. Assume that the constraints are non-degenerated and that the gradients w.r.t. (t, x, y) of the active constraints are linearly independent. Assume there exists some $\xi > 0$ such that for the Hessian of the Lagrange function $\mathscr{L}$ in the point (t^0, x^0, y^0) it is*

$$\left(t, x^\top, y\right) \nabla^2_{(t,x,y)} \mathscr{L}\left(t^0, x^0, y^0, \mu^0, \nu^0\right) \begin{pmatrix} t \\ x \\ y \end{pmatrix} \geq \xi \left\| \begin{pmatrix} t \\ x \\ y \end{pmatrix} \right\|^2$$

for all

$$\begin{aligned} &(t, x, y) \in \{(t, x, y) \in \mathbb{R}^{n_1+2} \mid \hat{r}_i t = \nabla_x f_i(x^0, y^0)^\top x \text{ if } \mu_i^0 > 0 \text{ for } i \in \{1, 2\}, \\ &\quad \nabla_x g_j(x^0, y^0)^\top x = 0 \text{ if } \nu_j^0 > 0 \text{ for } j \in \{1, \ldots, p\},\ \ y = 0\}. \end{aligned}$$

Then (t^0, x^0, y^0) is a local unique minimal solution of (5.4) for the parameter $\hat{a}^0$ and there is some $\delta > 0$ such that the function $\phi\colon N(\hat{a}^0) \to B_\delta(t^0, x^0, y^0) \times B_\delta(\mu^0, \nu^0)$ (for $N(\hat{a}^0)$ a neighborhood of $\hat{a}^0$)

$$\phi(\hat{a}) = (t(\hat{a}), x(\hat{a}), y(\hat{a}), \mu(\hat{a}), \nu(\hat{a}))$$

has the following first order approximation

$$\phi(\hat{a}) = \phi(\hat{a}^0) + \hat{M}^{-1}\hat{N}\left(\hat{a} - \hat{a}^0\right) + o\left(\|\hat{a} - \hat{a}^0\|\right)$$

with $\hat{M} =$

$$\begin{pmatrix} \nabla^2_{(t,x,y)}\mathscr{L} & \begin{matrix}0\\ \nabla_{(x,y)}f_1\end{matrix} & \begin{matrix}-1\\ \nabla_{(x,y)}f_2\end{matrix} & \begin{matrix}0\\ e_{(n_1+1)}\end{matrix} & \begin{matrix}0\\ -\nabla_{(x,y)}g_1\end{matrix} & \cdots & \begin{matrix}0\\ -\nabla_{(x,y)}g_p\end{matrix} \\ \mu_1^0(0,-\nabla_{(x,y)}f_1) & k_1 & 0 & 0 & 0 & \cdots & 0 \\ \mu_2^0(1,-\nabla_{(x,y)}f_2) & 0 & k_2 & 0 & 0 & \cdots & 0 \\ \mu_3^0(0,-e_{(n_1+1)}^\top) & 0 & 0 & 0 & 0 & \cdots & 0 \\ \nu_1^0(0,\nabla_{(x,y)}g_1) & 0 & 0 & 0 & g_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \nu_p^0(0,\nabla_{(x,y)}g_p) & 0 & 0 & 0 & 0 & \cdots & g_p \end{pmatrix}$$

in the point (t^0, x^0, y^0) *with* $k_1 = \hat{a}_1 - f_1(x^0, y^0)$, $k_2 = \hat{a}_2 + t^0 - f_2(x^0, y^0)$, $g_i = g_i(x^0, y^0)$, $i = 1, \ldots, p$, *and*

$$\hat{N} = \left[\mathbf{0}_{3\times(n_1+2)}, -\mu_1^0\begin{pmatrix}1\\0\\0\end{pmatrix}, -\mu_2^0\begin{pmatrix}0\\1\\0\end{pmatrix}, -\mu_3^0\begin{pmatrix}0\\0\\1\end{pmatrix}, \mathbf{0}_{3\times p}\right]^\top.$$

For the degenerate case we can adapt Theorem 4.4 also.

Let us assume that we want to refine now in the neighborhood of the point $(x^0, y^0) \in \mathscr{M}(F(A^0), \mathbb{R}^2_+)$. Then there is some $t^0 \in \mathbb{R}$ and some parameter $\hat{a}^0 = (a_1^0, a_2^0, a_3^0)^\top \in \hat{H} \subset \mathbb{R}^3$ such that (t^0, x^0, y^0) is a minimal solution of (5.4). Assume $\hat{\mu}^0 = (\mu_1^0, \mu_2^0, \mu_3^0) \in \mathbb{R}^3$, $\nu^0 \in \mathbb{R}^p$ are the related Lagrange multipliers using the results from Theorem 4.2. We want to find now new parameters $\hat{a} \in \hat{H}$, e.g., by

$$\hat{a} = \hat{a}^0 + s^1 \cdot v^1 + s^2 \cdot v^2, \qquad s^1, s^2 \in \mathbb{R},\ v^1 = (1,0,0)^\top,\ v^2 = (0,0,1)^\top,$$

such that we have for a given distance $\gamma > 0$

$$\|(x(\hat{a}^0 + s^i v^i), y(\hat{a}^0 + s^i v^i)) - (x^0, y^0)\| = \gamma, \qquad i = 1, 2,$$

with $(t(\hat{a}), x(\hat{a}), y(\hat{a}))$ the minimal solution of problem (5.4) dependent on the parameter $\hat{a}$. With the matrices $\hat{M}, \hat{N}$ as in Theorem 5.2 we conclude that this aim is approximately fulfilled for

$$|s^i| = \frac{\gamma}{\|\left(\hat{M}^{-1}\hat{N}\cdot v^i\right)\big|_{(x,y)}\|}, \quad i = 1, 2,$$

with $(\hat{M}^{-1}\hat{N}\cdot v^i)\big|_{(x,y)}$ the second to the (n_1+2)-th entry of the vector $\hat{M}^{-1}\hat{N}\cdot v^i$. According to the desired number of new discretization points we set for some $n^D \in \mathbb{N}$

$$\hat{a} = \hat{a}^0 + l^1 \cdot s^1 v^1 + l^2 \cdot s^2 v^2, \quad l^1, l^2 \in \{-n^D, \ldots, n^D\} \subset \mathbb{Z},\ (l^1, l^2) \neq (0,0). \tag{5.5}$$

Solving (5.4) for these new parameters $\hat{a}$ we get minimal solutions (t, x, y) and with that new points of the set of feasible points. We set

$$A^1_{(x^0,y^0)} := \{(x, y) \in \mathbb{R}^{n_1+1} \mid \exists t \in \mathbb{R} \text{ with } (t, x, y) \text{ a minimal solution of (5.4) for } \hat{a} \text{ as in (5.5)}\}.$$

By repeating this for all points $(x, y) \in \mathscr{M}(F(A^0), \mathbb{R}^2_+)$ we get the following new approximation of the induced set Ω

$$A^1 := A^0 \cup \bigcup_{(x,y)\in\mathscr{M}(F(A^0),\mathbb{R}^2_+)} A^1_{(x,y)}. \tag{5.6}$$

We map these points again under the upper level function F and select only the non-dominated (or ε-EP-minimal) points $\mathscr{M}(F(A^1), \mathbb{R}^2_+)$. Here we can use Lemma 2.2 and consider only the set

$$\tilde{A}^1 := \mathscr{M}(F(A^0), \mathbb{R}^2_+) \cup \bigcup_{(x,y)\in\mathscr{M}(F(A^0),\mathbb{R}^2_+)} A^1_{(x,y)}$$

because we have (according to Lemma 2.2)

$$\mathscr{M}(F(A^1), \mathbb{R}^2_+) = \mathscr{M}(F(\tilde{A}^1), \mathbb{R}^2_+). \tag{5.7}$$

The set $\mathscr{M}(F(A^1), \mathbb{R}^2_+)$ is now an improved approximation of the solution set of the multiobjective bilevel optimization problem but it can be refined further by repeating the described steps arbitrarily often. We summarize this in Part 2 of our algorithm.

Algorithm 2 (Bilevel algorithm: Part 2)

Step 3: Set $A^0 = \bigcup_{k=1}^{n^y} A^{0,k}$ and determine $\mathscr{M}(F(A^0), \mathbb{R}^2_+)$ with the help of a Pareto filter or the Jahn-Graef-Younes method. Set $i := 0$ and choose $\gamma^0 > 0$.

Step 4: For any point $(x, y) \in \mathscr{M}(F(A^i), \mathbb{R}^2_+)$ determine a refinement of the approximation of the set of feasible points of the upper level problem around this point by solving the scalarization (5.4) to the problem (4.2) and by choosing the parameters $\hat{a} \in \mathbb{R}^3$ as in (5.5). Determine A^{i+1} by (5.6).

Step 5: Calculate $\mathscr{M}(F(A^{i+1}), \mathbb{R}^2_+)$ by using (5.7) and a Pareto filter or the Jahn–Graef–Younes method. If this approximation of the solution set of the multiobjective bilevel problem is sufficient then stop. Else set $i := i + 1$, choose $\gamma^i > 0$ and go to Step 4.

6 Numerical results

In this section we apply the proposed algorithm to an academic test problem and to a problem which arose in a technical application.

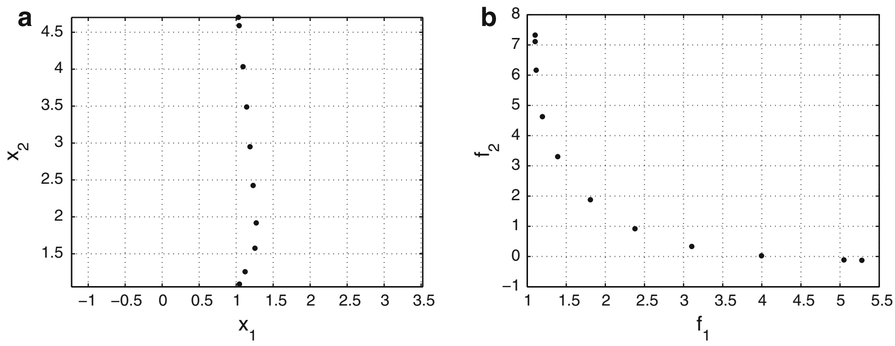

Fig. 1 Approximation **a** of the solution set and **b** the efficient set of problem (3.1) to the test problem for $y = 1.8$

6.1 Test Problem

We consider the following multiobjective bilevel problem (assuming the optimistic approach) with $n_1 = 2,\ n_2 = 1,\ m_1 = m_2 = 2$ and $K^1 = K^2 = \mathbb{R}^2_+$.

$$\min_{x,y} \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} = \begin{pmatrix} x_1 + x_2^2 + y + \sin^2(x_1 + y) \\ \cos(x_2) \cdot (0.1 + y) \cdot \left(\exp\left(-\frac{x_1}{0.1 + x_2}\right)\right) \end{pmatrix}$$

subject to the constraints

$$x \in \operatorname{argmin}_x \left\{ \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} \mid (x,y) \in G \right\},$$

$$y \in [0, 10]$$

with $f_1,\ f_2 \colon \mathbb{R}^3 \to \mathbb{R}$,

$$f_1(x,y) = \frac{(x_1 - 2)^2 + (x_2 - 1)^2}{4} + \frac{x_2 y + (5 - y)^2}{16} + \sin\left(\frac{x_2}{10}\right),$$

$$f_2(x,y) = \frac{x_1^2 + (x_2 - 6)^4 - 2x_1 y - (5 - y)^2}{80}$$

and

$$G = \{(x_1, x_2, y) \in \mathbb{R}^3 \mid x_1^2 - x_2 \le 0,\ 5x_1^2 + x_2 \le 10,\ x_2 - (5 - y/6) \le 0,\ x_1 \ge 0\}.$$

We apply Algorithms 1 and 2 with $\beta = 0.6$ for discretizing the interval [0, 10]. For example for $y = 1.8$ and the distance $\alpha = 0.6$ the Steps 2a and 2b lead to the approximation of the solution set of problem (3.1) shown in Fig. 1a and with that to an approximation of the efficient set of problem (3.1) shown in Fig. 1b. We used a Matlab implementation of the SQP-algorithm. Repeating the Steps 2a and 2b for all discretization points y of the interval [0, 10] we get the set A^0 shown in Fig. 2a which

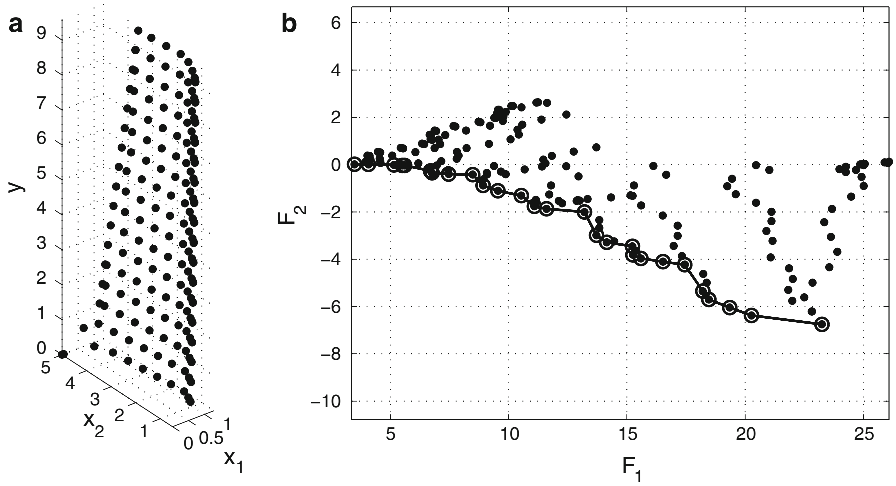

Fig. 2 **a** Approximation A^0 of the set of feasible points Ω and **b** the image $F(A^0)$ of this set under F

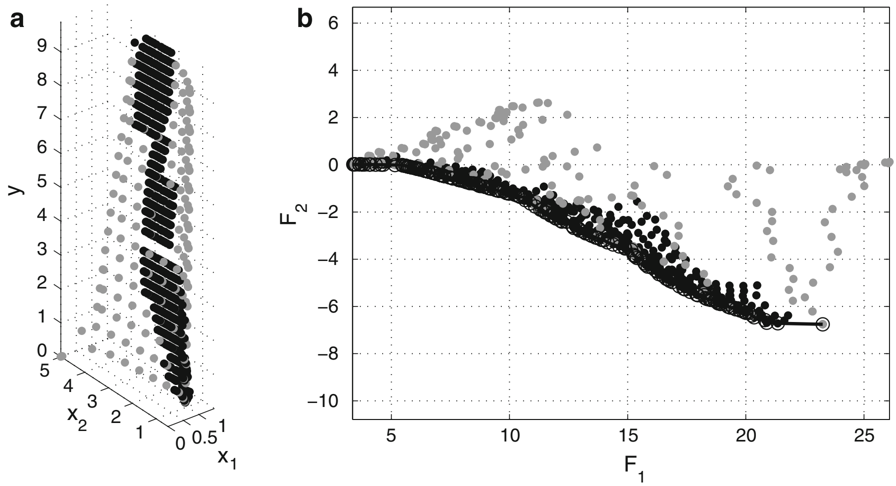

Fig. 3 **a** Refinement A^1 of the set of feasible points Ω and **b** its image under F.

is an approximation of the solution set of problem (4.2) and hence of the set of feasible points of the upper level problem. In Fig. 2b the set $F(A^0)$ is drawn. Here the image points under F of the set $\mathscr{M}(F(A^0), \mathbb{R}^2_+)$ are marked with circles and are connected with lines. We continue the algorithm by choosing $\gamma^0 = 0.3$ and $n^D = 2$. The set A^1 is given in Fig. 3a, whereby the points $A^1 \backslash A^0$ are drawn in black and the points A^0 in gray. The set $F(A^1)$ and the non-dominated points of this set can be seen in Fig. 3b. Repeating this with $\gamma^1 = 0.21$ and $\gamma^3 = 0.12$ and doing the refinement in Step 4 only for those points of the set $\mathscr{M}(F(A^i), \mathbb{R}^2_+)$ which have no neighbors in this set with a distance less than 0.3 we get the results shown in Figs. 4a,b and 5a, b.

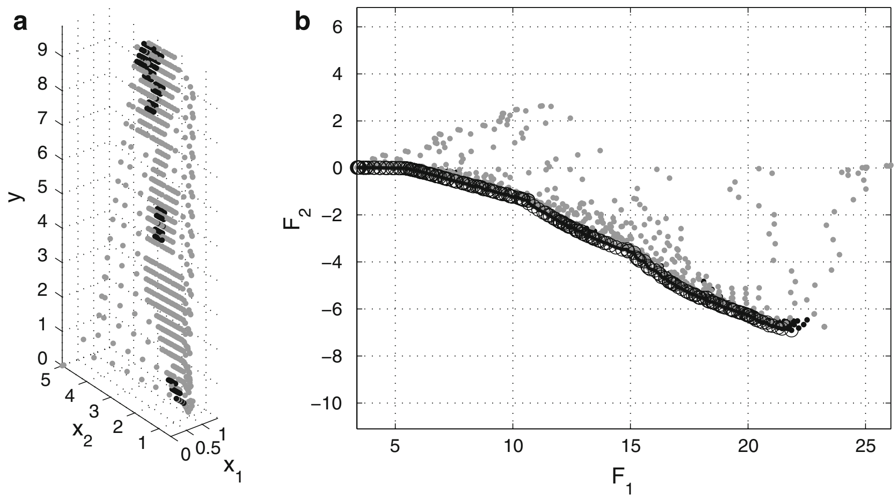

Fig. 4 **a** The set A^2 and **b** the image set $F(A^2)$

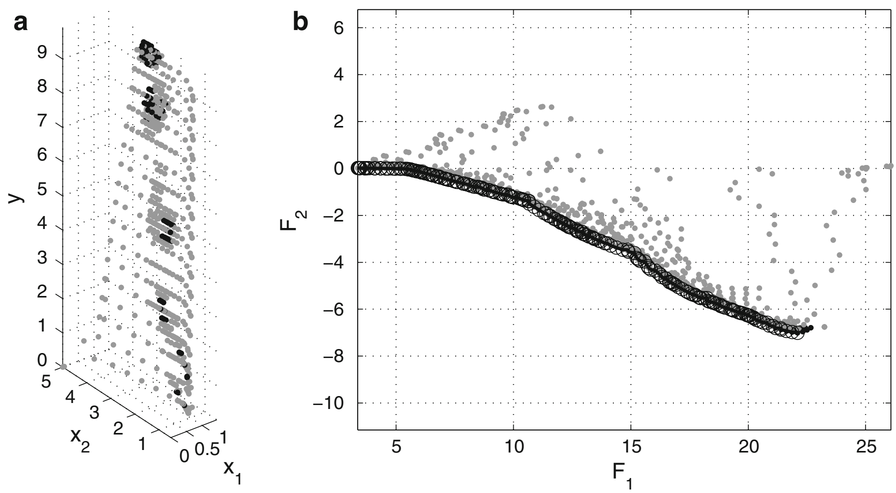

Fig. 5 **a** The set A^3 and **b** the image set $F(A^3)$

Then the algorithm is stopped as only small improvements for the approximation of the efficient set of the multiobjective bilevel optimization problem were gained by the last iteration.

6.2 Application problem

This optimization problem arose in a recent problem in medical engineering [32, p. 62f] dealing with the configuration of coils. In its original version this problem is a usual scalar valued standard optimization problem which had to be reformulated as a bilevel optimization problem due to the need of real-time solutions and because of its structure. It turned out that the engineers would accept or even prefer solutions which do not satisfy a previous equality constraint strictly in favor of incorporating some additional objectives. This led to the examination of two objective functions on both levels (see [12, p. 150f]). The resulting multiobjective bilevel optimization problem is the following

$$\begin{aligned} &\min_{x,y} \begin{pmatrix} \|x(y)\|_2^2 \\ \|x(y) - x_{\text{old}}\|_2^2 \end{pmatrix} \\ &\text{subject to the constraints} \\ &x = x(y) \in \operatorname{argmin}_x \left\{ \begin{pmatrix} \|A(y) \cdot Vx - b(y)\|_2^2 \\ \|x\|_2^2 \end{pmatrix} \;\middle|\; \right. \\ &\qquad \left. \|A(y) \cdot Vx - b(y)\|_2^2 \leq \Delta_{\max},\, x \in \mathbb{R}^{14} \right\}, \\ &y \in [0, \pi] \end{aligned} \tag{6.1}$$

with $n_1 = 14,\ n_2 = 1,\ K^1 = K^2 = \mathbb{R}^2_+$, $b(y) = (0, \cos y, \sin y, 1, 0, 0)^\top$, $\Delta_{\max} = 0.3$ and

$$A(y) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos y & \sin y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sin y & \cos y \\ 0 & 0 & 0 & -\sin y & 0 & 0 & \cos y & -\sin y \end{pmatrix}.$$

Here we choose $x_{\text{old}} \in \mathbb{R}^{14}$ as

$$\begin{aligned} x_{\text{old}} = (&0.1247, 0.1335, -0.0762, -0.1690, 0.2118, -0.0534, -0.1473, \\ &0.3170, -0.0185, -0.1800, 0.1700, -0.0718, 0.0058, 0.0985)^\top. \end{aligned}$$

The matrix $V \in \mathbb{R}^{8\times 14}$ is a non-sparse matrix with $\operatorname{rank}(V) = 8$ which depends on the considered medical system. For the calculation here a randomly chosen matrix is taken.

According to Step 1 of Algorithm 1, the set $\tilde{G} = [0, \pi]$ is discretized with $\beta = \frac{\pi}{8}$ and for avoiding numerical difficulties we choose

$$0.0001,\ 0.0001 + \frac{\pi}{8},\ 0.0001 + 2 \cdot \frac{\pi}{8},\ 0.0001 + 3 \cdot \frac{\pi}{8}, \ldots, \pi.$$

It is sufficient to consider in Steps 2a and 2b the problem (SP$(a, r, \tilde{a})$) for the parameters $a \in H = \{y \in \mathbb{R}^2 \mid y_2 = 0\}$ with $a_1 \in [0, \Delta_{\max}]$. That is the case because the

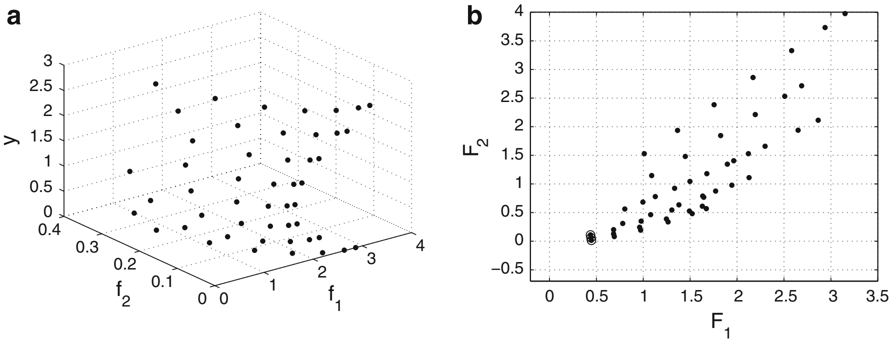

Fig. 6 **a** Approximation of the efficient set of problem (4.2) and **b** the set $F(A^0)$ of the application problem

constraint $a + t\,r - f(x, y^k) \in K^1$ for this problem is equivalent to

$$a_1 - \|A(y^k) \cdot Vx - b(y^k)\|_2^2 \geq 0,$$
$$t - \|x\|_2^2 \geq 0$$

and together with the constraint $(x, y) \in G$, i.e., $\|A(y) \cdot Vx - b(y)\|_2^2 \leq \Delta_{\max}$, problem (SP$(a, r, \tilde{a})$) has no feasible point for $a_1 < 0$. And for $a_1 > \Delta_{\max}$ the constraint $a_1 - \|A(y^k) \cdot Vx - b(y^k)\|_2^2 \geq 0$ is redundant and we get the same solutions as for $a_1 = \Delta_{\max}$.

Further we choose $\alpha = 0.2$ in Step 2 to compute the approximation $A^0 \subset \mathbb{R}^{15}$ of the set of feasible points and thus of the solution set of the tricriteria problem (4.2). We cannot draw the set A^0 and thus we show in Fig. 6a the image of the set A^0 under the objective functions of the problem (4.2). The set $F(A^0)$ and the nondominated points of $F(A^0)$ are given in Fig. 6b.

For refinements we choose $\gamma^0 = \frac{1}{35}$, $n^D = 3$ and $\gamma^1 = \frac{1}{50}$, $\gamma^2 = \frac{1}{70}$. The related approximations of the efficient set of the multiobjective bilevel optimization problem are given in Figs. 7a, b and 8a (note that it is only shown a very small section of the whole image set compared to Fig. 8 b).

The decision maker of the upper level can now choose his preferred solution from the set $\{y \in \mathbb{R} \mid (x, y) \in \mathscr{M}(F(A^2), \mathbb{R}^2_+)\}$.

7 Conclusion and outlook

We have presented several new theoretical results for general multiobjective bilevel optimization problems using the optimistic approach and we have given for the first time an algorithm for solving these types of problems, assuming some restrictions as bicriteria problems on both levels and a one-dimensional upper level variable. Yet based on the theoretical results generalizations of this algorithm can be done. The proposed method works on non-convex problems (assuming an appropriate solver for finding global solutions of scalar problems is available), too, because no reformulation of the lower level problem using the Karush–Kuhn–Tucker conditions is necessary.

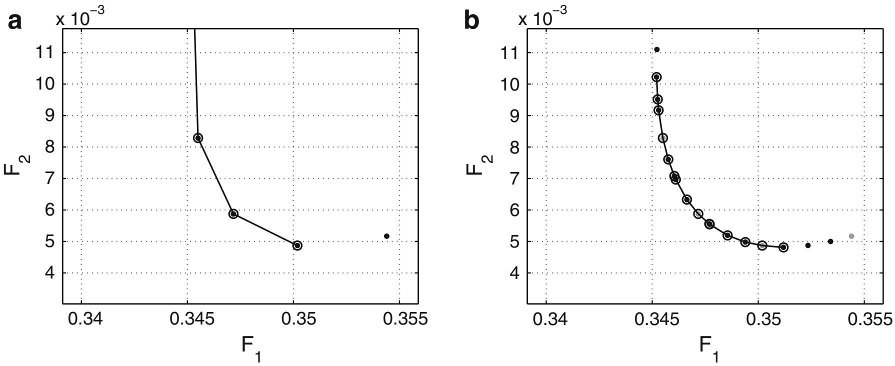

Fig. 7 Section of the approximation of the efficient set of the bilevel optimization problem after **a** the first and **b** the second refinement

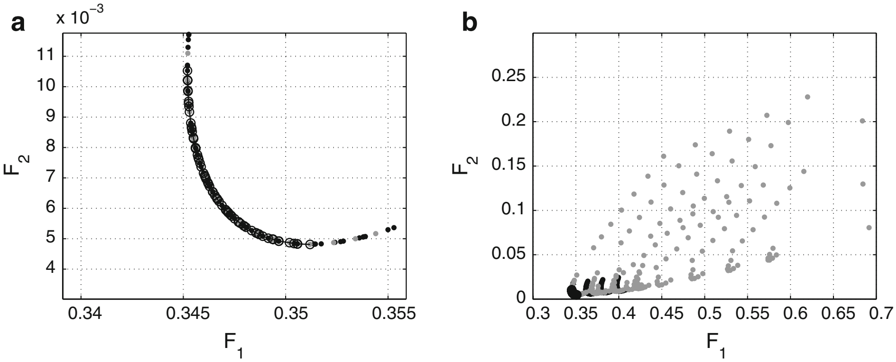

Fig. 8 **a** Section of the approximation of the efficient set of the bilevel optimization problem after the third refinement. **b** Complete image set $F(A^3)$

A more generalized algorithm and the consideration of further and more complex test problems and applications are subject to further research.

Acknowledgments The author wishes to thank Prof. Dr. J. Jahn for his supervision and his inspiring comments and suggestions. Further the author acknowledges Prof. Dr. J. Fliege for useful remarks as well as two anonymous referees for helpful comments and for sharing their ideas.

References

1. Abo-Sinna, M.: A bi-level non-linear multi-objective decision making under fuzziness. Opsearch **38**(5), 484–495 (2001)
2. Alt, W.: Parametric optimization with applications to optimal control and sequential quadratic programming. Bayreuther Math. Schr. **35**, 1–37 (1991)
3. Bard, J.F.: Practical Bilevel Optimization. Algorithms and Applications. Nonconvex Optimization and Its Applications, vol. **30**. Kluwer, Dordrecht (1998)
4. Bonnel, H., Morgan, J.: Semivectorial bilevel optimization problem: Penalty approach. J. Optim. Theory Appl. **131**(3), 365–382 (2006)

5. Das, I., Dennis, J.: A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. Struct. Optim. **14**, 63–69 (1997)
6. Das, I., Dennis, J.: Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM J. Optim. **8**(3), 631–657 (1998)
7. Dempe, S.: Foundations of Bilevel Programming. Nonconvex Optimization and Its Applications, vol. **61**. Kluwer, Dordrecht (2002)
8. Dempe, S.: Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization **52**(3), 333–359 (2003)
9. Dempe, S.: Bilevel programming—a survey. Preprint 2003-11, Fakultät für Mathematik und Informatik, TU Bergakademie Freiberg, Germany (2003)
10. Dinkelbach, W., Dürr, W.: Effizienzaussagen bei Ersatzprogrammen zum Vektormaximumproblem, iv. Oberwolfach-Tag. Operations Res 1971. Oper. Res. Verf. 12:69–77 (1972)
11. Ehrgott, M.: Multicriteria optimisation. Lect. Notes Econ. math. Syst, vol. **491**. Springer, Berlin (2000)
12. Eichfelder, G.: Parametergesteuerte Lösung nichtlinearer multikriterieller Optimierungsprobleme. PhD Thesis, Univ. Erlangen-Nürnberg, Germany (2006)
13. Eichfelder, G.: Scalarizations for adaptively solving multi-objective optimization problems. Comput. Optim. Appl. (2007) doi:10.1007/s10589-007-9155-4
14. Eichfelder, G.: An adaptive scalarization method in multi-objective optimization. SIAM J. Optim. (to appear)
15. Fiacco, A.V.: Introduction to Sensitivity and Stability Analysis in Nonlinear Programming. Mathematics in Science and Engineering, vol. **165**. Academic Press, London (1983)
16. Fliege, J.: Gap-free computation of Pareto-points by quadratic scalarizations. Math. Methods Oper. Res. **59**(1), 69–89 (2004)
17. Fliege, J., Vicente, LN: Multicriteria approach to bilevel optimization. J. Optim. Theory Appl. **131**(2), 209–225 (2006)
18. Haimes, Y., Lasdon, L., Wismer, D.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE Trans. Syst. Man Cybern. **1**, 296–297 (1971)
19. Hillermeier, C., Jahn, J.: Multiobjective optimization: survey of methods and industrial applications. Surv. Math. Ind. **11**, 1–42 (2005)
20. Jahn, J.: Vector Optimization: Theory, Applications and Extensions. Springer, Berlin (2004)
21. Jahn, J.: Multiobjective search algorithm with subdivision technique. Comput. Optim. Appl. **35**, 161–175 (2006)
22. Jahn, J., Merkel, A.: Reference point approximation method for the solution of bicriterial nonlinear optimization problems. J. Optim. Theory Appl. **74**(1), 87–103 (1992)
23. Jittorntrum, K.: Solution point differentiability without strict complementarity in nonlinear programming. Math. Program. Study **21**, 127–138 (1984)
24. Kim, I., de Weck, O.: Adaptive weighted sum method for bi-objective optimization. Struct. Multidisciplinary Optim. **29**, 149–158 (2005)
25. Lin, J.G.: On min-norm and min-max methods of multi-objective optimization. Math. Program. **103**(1), 1–33 (2005)
26. Loridan, P.: ε-solutions in vector minimization problems. J. Optim. Theory Appl. **43**, 265–276 (1984)
27. Marglin, S.: Public Investment Criteria. MIT Press, Cambridge (1967)
28. Miettinen, K.M.: Nonlinear Multiobjective Optimization. Kluwer, Boston (1999)
29. Nishizaki, I., Sakawa, M.: Stackelberg solutions to multiobjective two-level linear programming problems. J. Optim. Theory Appl. **103**(1), 161–182 (1999)
30. Osman, M., Abo-Sinna, M., Amer, A., Emam, O.: A multi-level nonlinear multi-objective decision-making under fuzziness. Appl. Math. Comput. **153**(1), 239–252 (2004)
31. Pascoletti, A., Serafini, P.: Scalarizing vector optimization problems. J. Optim. Theory Appl. **42**(4), 499–524 (1984)
32. Prohaska, J.: Optimierung von Spulenkonfigurationen zur Bewegung magnetischer Sonden. Diplomarbeit, Univ. Erlangen-Nürnberg, Germany (2005)
33. Ruzika, S., Wiecek, M.: Approximation methods in multiobjective programming. J. Optim. Theory Appl. **126**(3), 473–501 (2005)
34. Sawaragi, Y., Nakayama, H., Tanino, T.: Theory of Multiobjective Optimization. Number 176 in Mathematics in science and engineering. Academic Press, London (1985)
35. Schandl, B., Klamroth, K., Wiecek, M.M.: Norm-based approximation in bicriteria programming. Comput. Optim. Appl. **20**(1), 23–42 (2001)

36. Shi, X., Xia, H.: Interactive bilevel multi-objective decision making. J. Oper. Res. Soc. **48**(9), 943–949 (1997)
37. Shi, X., Xia, H.: Model and interactive algorithm of bi-level multi-objective decision-making with multiple interconnected decision makers. J. Multi-Criteria Decis. Anal. **10**, 27–34 (2001)
38. Staib, T.: On two generalizations of Pareto minimality. J. Optim. Theory Appl. **59**(2), 289–306 (1988)
39. Teng, C.-X., Li, L., Li, H.-B.: A class of genetic algorithms on bilevel multi-objective decision making problem. J. Syst. Sci. Syst. Eng. **9**(3), 290–296 (2000)
40. Tuy, H., Migdalas, A., Hoai-Phuong, N.: A novel approach to bilevel nonlinear programming. J. Global Optim. **38**(4), 527–554 (2007)
41. Vicente, L.N., Calamai, P.H.: Bilevel and multilevel programming: A bibliography review. J. Glob. Optim. **5**(3), 291–306 (1994)
42. Yin, Y.: Multiobjective bilevel optimization for transportation planning and management problems. J. Adv. Transp. **36**(1), 93–105 (2000)
43. Younes, Y.: Studies on discrete vector optimization. Dissertation, University of Demiatta, Egypt (1993)