

Methods for multiobjective bilevel optimization

GABRIELE EICHFELDER¹

August 12, 2019

Abstract

This paper is on multiobjective bilevel optimization, i.e. on bilevel optimization problems with multiple objectives on the lower or on the upper level, or even on both levels. We give an overview on the major optimality notions used in multiobjective optimization. We provide characterization results for the set of optimal solutions of multiobjective optimization problems by means of scalarization functionals and optimality conditions. These can be used in theoretical and numerical approaches to multiobjective bilevel optimization. As multiple objectives arise in multiobjective optimization as well as in bilevel optimization problems, we also point out the results on the connection between these two classes of optimization problems. Finally, we give reference to numerical approaches which have been followed in the literature to solve these kind of problems. We concentrate in this paper on nonlinear problems, while the results and statements then naturally also hold for the linear case.

1 Introduction

In bilevel optimization one typically assumes that one has scalar-valued objective functions in the two levels of the bilevel program. This may be, for instance, costs or time which has to be minimized. However, many real world problems can only be modeled adequately by considering several, in most cases conflicting, objective function simultaneously, as weight and stability, or costs and time.

This occurs for instance in transportation planning [51]. We illustrate this with an example, cf. [17]: 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 motivate 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 several, maybe competing, objectives, too, as to minimize their transportation time and costs. Hence, the usage of

¹Institute for Mathematics, TU Ilmenau, Po 10 05 65, 98684 Ilmenau, Germany, email: gabriele.eichfelder@tu-ilmenau.de

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. Thus, such a problem can be modeled by multiobjective bilevel optimization.

In [40, 17] also a problem in medical engineering was discussed. There, the task was the configuration of coils. In the original version, this problem was 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.

Dempe and Mehlitz consider in [12] a bilevel control problem, where the three objective functions are approximating a given target state as good as possible, minimizing the overall control effort, and the sparsity of the chosen control.

In case we have more than one objective function than we speak of a multi-objective optimization problem. Collecting for instance m objective functions in an m -dimensional vector also leads to the name vector optimization, as now vector-valued maps have to be optimized. In case we have a bilevel program with a vector-valued objective on one or both of the levels we speak of a multiobjective bilevel optimization problem. In case one has multiple functions just on one level one also uses the name semivectorial bilevel problem.

For instance, Pilecka studies in [39] bilevel optimization problems with a single-objective lower level and a vector-valued function on the upper level. The same structure was assumed by Gadhi and Dempe in [23, 24] or Ye in [50]. Bonnel and co-authors [3] study semivectorial bilevel problems with a multiobjective optimization problem on the lower level. For optimality conditions for bilevel problems with such a structure (multiobjective on the lower level, single-objective on the upper level), see also Dempe and co-authors in [10].

In this paper we will give a short introduction to multiobjective optimization. First, an optimality concept for such optimization problems has to be formulated. By doing that one can define stricter or weaker concepts, which may have advantages from applications point of view or from the theoretical or numerical point of view, respectively.

We give scalarization results and optimality conditions which allow to characterize the set of optimal solutions of a multiobjective optimization problem. We also point out some of the developments on numerical solvers. Throughout the paper we emphasize more on nonlinear problems. For the linear case of a multiobjective bilevel optimization problem with multiple functions on each level see for instance [9] and the references therein. We will also study reformulations of the feasible set of the upper level problem, which can be used to formulate numerical solution approaches.

In bilevel optimization also multiple objectives arise: a function on the upper level and a function on the lower level. Hence, we will also discuss

the relationship between multiobjective and bilevel optimization from this perspective.

2 Basics of Multiobjective Optimization

As we will consider bilevel optimization problems with a multiobjective optimization problem on the lower or/and on the upper level, we start by giving an overview on the main optimality notions typically used in multiobjective optimization.

In multiobjective optimization one studies optimization problems formally defined by

$$\begin{aligned} \min f(x) &= (f_1(x), \dots, f_m(x))^\top \\ \text{subject to the constraint} & \\ x &\in \Omega, \end{aligned} \tag{MOP}$$

with a vector-valued objective function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($m, n \in \mathbb{N}$, $m \geq 2$) and a nonempty set of feasible points $\Omega \subseteq \mathbb{R}^n$.

For defining minimality for the multiobjective optimization problem (MOP), we need a partial ordering in the image space. Thus one considers partial orderings introduced by closed pointed convex cones $K \subseteq \mathbb{R}^m$. 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 $K \cap (-K) = \{0\}$. A partial ordering introduced by a pointed convex cone is antisymmetric. The partial ordering is then defined by

$$x \leq_K y \quad :\Leftrightarrow \quad y - x \in K .$$

Hence, in the following, let $K \subseteq \mathbb{R}^m$ be a closed pointed convex cone.

For $K = \mathbb{R}_+^m$ we have for $x, y \in \mathbb{R}^m$

$$x \leq_{\mathbb{R}_+^m} y \quad \Leftrightarrow \quad x_i \leq y_i, \quad i = 1, \dots, m,$$

i.e. the componentwise ordering. Throughout the paper, we denote by \leq without a subscript the componentwise ordering defined by $\leq := \leq_{\mathbb{R}_+^m}$. For more details on ordering structures in vector optimization we refer to [21].

By using the concept of a partial ordering, we are now able to define two types of optimal solutions of (MOP)

Definition 2.1. *A point $\bar{x} \in \Omega$ is called a K -minimal point, or efficient, for (MOP) if*

$$(\{f(\bar{x})\} - K) \cap f(\Omega) = \{f(\bar{x})\}.$$

Additionally, for $\text{int}(K) \neq \emptyset$, a point $\bar{x} \in \Omega$ is called a weakly K -minimal point, or weakly efficient, for (MOP) if

$$(\{f(\bar{x})\} - \text{int}(K)) \cap f(\Omega) = \emptyset.$$

Of course, any efficient point is also weakly efficient, in case $\text{int}(K) \neq \emptyset$. Whenever we speak of weakly efficient points we assume that the ordering

cone K has a nonempty interior. We denote the set of all K -minimal points, i. e. of all efficient points, as $\mathcal{E}(f(\Omega), K)$, and the set of all weakly efficient points as $\mathcal{E}^w(f(\Omega), K)$. The set $\mathcal{N}(f(\Omega), K) := \{f(x) \in \mathbb{R}^m \mid x \in \mathcal{E}(f(\Omega), K)\}$ is called *nondominated set* and the set $\mathcal{N}^w(f(\Omega), K) := \{f(x) \in \mathbb{R}^m \mid x \in \mathcal{E}^w(f(\Omega), K)\}$ *weakly nondominated set*.

For $K = \mathbb{R}_+^m$ the efficient points are also denoted as *Edgeworth-Pareto (EP)-minimal* points. Then the definition can also be formulated as follows: a point $\bar{x} \in \Omega$ is an EP-minimal point for (MOP) if there is no $x \in \Omega$ with

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

Similarly, a point is called weakly EP-minimal or weakly efficient for (MOP) w.r.t. $K = \mathbb{R}_+^m$ if there is no $x \in \Omega$ with

$$f_i(x) < f_i(\bar{x}), \quad i = 1, \dots, m.$$

The weakly efficient points are in most cases more of interest from a theoretically or numerically point of view, see, for instance, Theorem 3.3, or the results on optimality conditions for a semivectorial bilevel optimization problem as given in [23, 24, 10]. From an application point of view one would in general only be interested in a weakly efficient point which is also efficient, as otherwise one could improve the value of one objective function without deteriorating the others. Nevertheless, in many studies on multiobjective bilevel optimization problems it is assumed that the lower level decision maker would also be satisfied with a weakly efficient solution, see, for instance, [10]. Note that in case Ω is convex and the functions f_i , $i = 1, \dots, m$ are strictly convex, i.e., if for all $i = 1, \dots, m$

$$f_i(\lambda x + (1 - \lambda)y) < \lambda f_i(x) + (1 - \lambda)f_i(y) \quad \text{for all } x, y \in \Omega, \quad x \neq y, \quad \lambda \in]0, 1[,$$

then any weakly efficient point for (MOP) w.r.t. $K = \mathbb{R}_+^m$ is also an efficient point of (MOP) w.r.t. $K = \mathbb{R}_+^m$.

For iterative numerical approaches, the following result is useful: in case the feasible set becomes larger and one has already determined the minimal solutions of the smaller set, then this can be used for determining the minimal solutions w.r.t. a larger superset.

Lemma 2.2. [17] *For two sets $A^0, A^1 \subseteq \mathbb{R}^n$, and a vector-valued function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$, consider the sets*

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

Let $f(A^0)$ be compact. Then it holds $\mathcal{E}(f(A), \mathbb{R}_+^m) = \mathcal{E}(f(\tilde{A}), \mathbb{R}_+^m)$.

For numerical calculations, also the notion of ε -minimal solutions can be useful. We state here a definition for the ordering cone $K = \mathbb{R}_+^m$, see also for instance [30, 31, 48]:

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

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

Example 2.4. Example to optimality notions Let $\Omega = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 \geq 1\}$, $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by $f(x) = x$ for all $x \in \mathbb{R}^2$ and $K = \mathbb{R}_+^2$. Then the set of efficient points is

$$\mathcal{E}(f(\Omega), K) = \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 = 1\}$$

and the set of weakly efficient points is

$$\begin{aligned} \mathcal{E}^w(f(\Omega), K) &= \{x \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1 + x_2 = 1\} \\ &\cup \{(0, x_2) \in \mathbb{R}^2 \mid x_2 \geq 1\} \cup \{(x_1, 0) \in \mathbb{R}^2 \mid x_1 \geq 1\}. \end{aligned}$$

As the objective map f is the identity we have

$$\mathcal{N}(f(\Omega), K) = \mathcal{E}(f(\Omega), K) \quad \text{and} \quad \mathcal{N}^w(f(\Omega), K) = \mathcal{E}^w(f(\Omega), K).$$

For $\varepsilon_i = 1/4$ for $i = 1, 2$, the point $\bar{x} = (1/2 + 1/8, 1/2 + 1/8)$ is an ε -EP-minimal solution of the multiobjective optimization problem $\min_{x \in \Omega} f(x)$ but not an efficient point.

For multiobjective optimization problems also more general optimality notions have been introduced, which use more general ordering structures. For instance, Ye discusses in [50] semivectorial bilevel problems where the preference relation is locally satiated and almost transitive. Moreover, the concept of using a fixed ordering cone can be replaced by using an ordering map which associates to each element y of the space an ordering cone $K(y)$. This was done by Dempe and Gadhi in [11]. They study semivectorial bilevel problems with a multiobjective optimization problem on the upper level and where the cones $K(y)$ are assumed to be Bishop-Phelps-cones. For such multiobjective optimization problems also proper nonlinear scalarizations exist, see [18], which have also been used by Dempe and Gadhi in their approach for formulating necessary optimality conditions. Variable ordering structures have been intensively studied, cf. [19] and, within the context of set-optimization, in [20]. There is a strong relation between bilevel optimization and set optimization, i.e. with optimization with a set-valued objective function. This arises whenever the optimal solution set on the lower level is not unique, see [39].

3 Characterization of optimal solutions of a multiobjective optimization problem

For solving problem (MOP) several methods are discussed in the literature. For surveys see [13, 29, 35, 42]. An often used approach is to replace the vector

optimization problem by a suitable scalar-valued optimization problem. This fact is also used in numerical methods for multiobjective bilevel problems to reformulate the lower level problem by a single-objective problem or by the optimality conditions of this reformulation.

Because of its simplicity, a widely used scalarization for multiobjective optimization is linear scalarization. For that one uses elements of the dual cone

$$K^* := \{w \in \mathbb{R}^m \mid w^\top y \geq 0 \text{ for all } y \in K\}$$

as well as of the quasi-interior of the dual cone

$$K^\# := \{w \in \mathbb{R}^m \mid w^\top y > 0 \text{ for all } y \in K \setminus \{0\}\}.$$

For $K = \mathbb{R}_+^m$ we have $K^* = \mathbb{R}_+^m$ and $K^\# = \text{int}(\mathbb{R}_+^m) = \{y \in \mathbb{R}^m \mid y_i > 0, i = 1, \dots, m\}$.

For convex multiobjective optimization problems one can use linear scalarization for a full characterization of the weakly efficient points.

Theorem 3.1. (a) Let $w \in K^* \setminus \{0\}$ and $\bar{x} \in \Omega$ with

$$w^\top f(\bar{x}) \leq w^\top f(x) \text{ for all } x \in \Omega.$$

Then \bar{x} is a weakly efficient point for (MOP). If, additionally, $w \in K^\#$, then \bar{x} is even efficient for (MOP).

(b) Let $f(\Omega) + K$ be convex. Then for any weakly efficient point $\bar{x} \in \Omega$ of (MOP) (and thus for any efficient point) there exists some $w \in K^* \setminus \{0\}$ with

$$w^\top f(\bar{x}) \leq w^\top f(x) \text{ for all } x \in \Omega.$$

Hence, in the convex case, we have a full characterization of weakly efficient points. A similar result as in (b) does not exist for efficient points and for $K^\#$. This is shown with the next example.

Example 3.2. Example on linear scalarization Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ with $f(x) = x$, $K = \mathbb{R}_+^2$, and $\Omega = \{x \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\}$. Then

$$\mathcal{E}(f(\Omega), \mathbb{R}_+^2) = \{x \in \mathbb{R}^2 \mid \|x\|_2 = 1, x_1 \leq 0, x_2 \leq 0\}.$$

To the point $\bar{x} = (0, -1) \in \mathcal{E}(f(\Omega), \mathbb{R}_+^2)$ there exists no $w \in K^\# = \text{int}(\mathbb{R}_+^m)$ with

$$w^\top f(\bar{x}) = -w_2 \leq w^\top f(x) = w_1 x_1 + w_2 x_2 \text{ for all } x \in \Omega.$$

Only for $w = (0, 1)^\top$ the point \bar{x} is a minimal solution of

$$\min_{x \in \Omega} w^\top f(x).$$

This characterization is often used for solving multiobjective bilevel optimization problems. If we have multiple objectives and if the problem is convex, then the set

$$\{\bar{x} \in \Omega \mid \exists w \in K^* \setminus \{0\} : w^\top f(\bar{x}) \leq w^\top f(x) \text{ for all } x \in \Omega\} \quad (1)$$

equals the set of weakly efficient points $\mathcal{E}^w(f(\Omega), K)$. Such linear scalarizations are also called weighted-sum approach, and the components w_i of the vector $w \in K^*$ are understood as weights of the objective function.

The approach in (1) is for instance used in [32, 33]. The authors reformulate a linear semivectorial bilevel programming problem (with a multiobjective optimization problem on the lower level) as a special bilevel programming problem, where the lower level is a parametric linear scalar-valued optimization problem. For characterizing the set in (1) optimality conditions from scalar-valued optimization can be used. This approach with using the weighted sum is also examined in [12] and it is pointed out that this is a delicate issue in case one studies locally optimal solutions of the parameterized problem only.

Also note that from an application point of view, the weakly efficient points are typically not so much of interest. A point $x \in \Omega$ which is a weakly EP-minimal point of (MOP) but not an EP-minimal point of (MOP) means, that there is still some \tilde{x} which satisfies all objective functions equally good and is better with respect to one of the objective functions. By considering instead the set

$$\{\tilde{x} \in \Omega \mid \exists w \in K^\# : w^\top f(\tilde{x}) \leq w^\top f(x) \text{ for all } x \in \Omega\}$$

one only has a subset of the set of efficient points of (MOP) as the Example 3.2 has shown. Note that there are numerical approaches for multiobjective bilevel problems which in fact aim on the efficient instead of the weakly efficient points, see for instance [25]

Theorem 3.1(b) requires the set $f(\Omega) + K$ to be convex. This holds in case the functions f_i are convex, $K = \mathbb{R}_+^m$, and Ω is a convex set. For nonconvex multiobjective optimization problems one has to use nonlinear scalarizations. A widely used one is a scalarization by Pascoletti and Serafini, [38], also known as Tammer-Weidner-functional [26, 27]. This scalarization approach allows to determine weakly efficient points for an arbitrary ordering cone K while many other well known nonlinear scalarizations as the ε -constraint problem (see [13, 28, 34, 35]) are mainly developed for determining EP-minimal points, only. Moreover, those can also be seen as a special case of the Pascoletti-Serafini-scalarization, see [15]. 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)

$$\begin{aligned} & \min t \\ & \text{subject to the constraints} \\ & a + tr - f(x) \in K, \\ & x \in \Omega, \\ & t \in \mathbb{R} \end{aligned} \tag{2}$$

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

Theorem 3.3. (a) Let \bar{x} be an efficient point of (MOP), then $(0, \bar{x})$ is a minimal solution of (2) with $a = f(\bar{x})$, $r \in K \setminus \{0\}$.

(b) Let (\bar{t}, \bar{x}) be a minimal solution of (2), then $\bar{x} \in \mathcal{E}^w(f(\Omega), K)$.

Thus all efficient 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 ([16], Theorem 3.2).

Theorem 3.4. Let $\bar{x} \in \mathcal{E}(f(\Omega), K)$ 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).

Also other scalarizations are used for deriving results on multiobjective bilevel optimization problems. For instance, Gadhi and Dempe use in [23] the scalarization known as signed-distance functional for deriving optimality conditions. This functional evaluates the distance of a point to the negative of the ordering cone as well as to the complement of the negative of the ordering cone.

We end this section by giving the KKT-type (i.e., Karush-Kuhn-Tucker-type) optimality conditions for constrained multiobjective optimization problems. KKT-conditions, and optimality conditions in general, are often used to characterize the optimal solutions of the lower level of a bilevel optimization problem. Such an approach was for instance named to be a classical approach by Deb and Sinha in [6] and it was used by Chuong in [4] for a nonsmooth multiobjective bilevel optimization problem. As it is known from single-objective optimization, KKT-conditions typically include complementarity conditions which are difficult to handle mathematically. Moreover, the necessary conditions require some regularity assumptions (cf. (A2) in [33]). For more details on the sufficiency and necessity of the optimality conditions we refer for instance to [13, 29].

We require the feasible set to be given by inequality and equality constraints (also basic convex sets could be handled). For shortness of representation we restrict ourselves here to inequality constraints. Thus, let $g_k: \mathbb{R}^n \rightarrow \mathbb{R}$, $k = 1, \dots, p$ be continuously differentiable functions such that

$$\Omega := \{x \in \mathbb{R}^n \mid g_k(x) \leq 0, k = 1, \dots, p\}. \quad (3)$$

We give in the following the necessary conditions which also hold for just locally efficient solutions, too. For a proof see for instance [13, Theorem 3.21].

Theorem 3.5 (Necessary optimality conditions). Let Ω be given as in (3) and let f_j , $j = 1, \dots, m$ and g_k , $k = 1, \dots, p$ be continuously differentiable. Let \bar{x} be a weakly efficient solution of (MOP) and assume that there exists some $d \in \mathbb{R}^n$ such that

$$\nabla g_k(\bar{x})^\top d < 0 \text{ for all } k \in I(\bar{x}) := \{k \in \{1, \dots, p\} \mid g_k(\bar{x}) = 0\},$$

i. e. a regularity assumption is satisfied. Then there exist Lagrange multipliers $\lambda \in \mathbb{R}_+^m \setminus \{0\}$ and $\mu \in \mathbb{R}_+^p$ such that

$$\begin{aligned} \sum_{j=1}^m \lambda_j \nabla f_j(\bar{x}) + \sum_{k=1}^p \mu_k \nabla g_k(\bar{x}) &= 0, \\ \mu_k g_k(\bar{x}) &= 0, \quad k = 1, \dots, p. \end{aligned} \quad (4)$$

With appropriate convexity assumptions we can state also sufficient conditions. We state here a simplified version. For weaker assumptions on the convexity of the functions, as quasiconvexity for the constraints, we refer to [29].

Theorem 3.6 (Sufficient optimality conditions). *Let Ω be given as in (3) and let f_j , $j = 1, \dots, m$ and g_k , $k = 1, \dots, p$ be convex and continuously differentiable. Let $\bar{x} \in \Omega$ and assume that there exist Lagrange multipliers $\lambda \in \mathbb{R}_+^m \setminus \{0\}$ and $\mu \in \mathbb{R}_+^p$ such that (4) is satisfied. Then, \bar{x} is a weakly efficient solution of (MOP).*

4 Connection between multiobjective optimization and bilevel optimization

In bilevel optimization one has (at least) two objective functions: one on the upper and one on the lower level. In multiobjective optimization one also studies problems with two or more objective functions. This naturally leads to the question on how these problem classes are related, and whether bilevel problems can be reformulated as multiobjective optimization problems. This question was answered for nonlinear problems by Fliege and Vicente in [22]: only by introducing a very special order relation the optimal solutions of the bilevel problem are exactly the efficient points of a multiobjective optimization problem. Only for continuously differentiable problems on the lower level this order relation can be formulated as a more tractable order relation. We give some of the basic ideas in the following, all based on [22]. For a more recent survey and unifying results as well as more general characterization of the connections between single-level and bilevel multiobjective optimization we refer to [41]. They also provide a thorough literature survey on this topic as well as the historic sources of this approach for bilevel linear problems.

We assume the bilevel problem to be of the form

$$\begin{aligned} \min_{x_u \in \mathbb{R}^{n_u}, x_\ell \in \mathbb{R}^{n_\ell}} f_u(x_u, x_\ell) \\ \text{s.t.} \quad x_\ell \in \operatorname{argmin}_{y \in \mathbb{R}^{n_\ell}} \{f_\ell(x_u, y)\}, \end{aligned} \quad (5)$$

where $f_u, f_\ell: \mathbb{R}^{n_u} \times \mathbb{R}^{n_\ell} \rightarrow \mathbb{R}$ are the upper-level and lower-level objective functions respectively. Hence, we assume that there are no constraints in each level and we assume the optimistic approach. The handling of constraints is also shortly described in [22].

As a related multiobjective optimization problem, it is not enough to study just a problem with the vector-valued function $(f_u(x_u, x_\ell), f_\ell(x_u, x_\ell))$. Instead, we need a map which includes the vector x_u as a component. The

authors in [22] present a complete reformulation with an ordering which is not numerically tractable. They also formulate a weaker relation which we present in the following but which gives a sufficient condition only.

The related vector optimization problem uses as objective map $F: \mathbb{R}^{n_u} \times \mathbb{R}^{n_\ell} \rightarrow \mathbb{R}^{n_u+3}$ defined by

$$F(x_u, x_\ell) = (x_u, f_u(x_u, x_\ell), f_\ell(x_u, x_\ell), \|\nabla_{x_\ell} f_\ell(x_u, x_\ell)\|_2). \quad (6)$$

In the image space, a cone K is used to define an ordering as before by $x \leq y$ if and only if $y - x \in K$:

$$K = \{(x, f_1, f_2, d) \in \mathbb{R}^{n_u+3} \mid (x = 0 \wedge f_2 > 0) \vee (f_1 > 0 \wedge d \geq 0)\} \cup \{0\}. \quad (7)$$

Theorem 4.1. [22, Corollary 4.2] *Let F and K be as in (6) and (7). Let a point $\bar{x} = (\bar{x}_u, \bar{x}_\ell) \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_\ell}$ be such that $F(\bar{x})$ is nondominated w.r.t. K , i.e.*

$$(\{F(\bar{x})\} - K \setminus \{0\}) \cap \{F(x) \mid x \in \mathbb{R}^{n_u} \times \mathbb{R}^{n_\ell}\} = \emptyset,$$

then \bar{x} is an optimal solution of the bilevel optimization problem (5).

The study about the relation between multiobjective and bilevel optimization problems is also continued and unified in [41] as already mentioned above. The authors also study thereby multiobjective bilevel optimization problems and their connection to multiobjective optimization.

5 A multiobjective bilevel program

In this section, we introduce the multiobjective bilevel optimization problem. We formulate the general bilevel optimization problem (see (10)). Then we introduce the so-called optimistic approach (see (11)) which is a special case of the general bilevel optimization problem.

As mentioned in the introduction, in bilevel optimization we have a parameter dependent optimization problem on the lower level, which depends on the variable $y \in \mathbb{R}^{n_2}$ of the upper level. In case of multiple objectives, this lower level problems reads as

$$\begin{aligned} & \min_{x \in \mathbb{R}^{n_1}} f(x, y) \\ & \text{subject to the constraint} \\ & (x, y) \in G \end{aligned} \quad (8)$$

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

$$x \in G(y) := \{x \in \mathbb{R}^{n_1} \mid (x, y) \in G\}. \quad (9)$$

The variables $x \in \mathbb{R}^{n_1}$ 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 (8), hence

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

For a multiobjective optimization problem on the lower level it has to be clarified what a minimal solution is, i.e. how the set $\Psi(y)$ is defined. An appropriate concept from a practical point of view might be the set of efficient elements w.r.t. a partial ordering, while from a theoretical point of view the set of weakly efficient points might be more suitable.

The optimization problem of the upper level is then given by

$$\begin{aligned} & \text{“ min ” } F(x(y), y) \\ & \text{subject to the constraints} \\ & \quad x(y) \in \Psi(y), \\ & \quad y \in \tilde{G} \end{aligned} \tag{10}$$

with a vector-valued function $F: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{m_2}$, $m_2 \in \mathbb{N}$, 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. 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 , what we will discuss shortly later. We speak of a multiobjective bilevel optimization problem if $m_1 \geq 2$ or $m_2 \geq 2$.

If the minimal solution of the lower level problem (8) 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 (10). 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 is in general not reasonable any more.

Hence, a possible approach is to study the set-valued map

$$y \mapsto \{F(x, y) \mid x \in \Psi(y)\}.$$

In the case of non-uniqueness, another common approach is the optimistic approach. There it is assumed that the decision maker of the lower level chooses among all minimal solutions that minimal solution, which is best for the upper level, i.e. which is minimal for the objective function of the upper level. Hence, it is solved

$$\begin{aligned} & \min_{x, y} F(x, y) \\ & \text{subject to the constraints} \\ & \quad x \in \Psi(y), \\ & \quad y \in \tilde{G}. \end{aligned} \tag{11}$$

Thus, in the optimistic modification (11), the objective function of the upper level is minimized w.r.t. x and y , while in the general formulation (10) it is only minimized w.r.t. the upper level variable y . In the following, we consider the optimistic approach only, i.e. the bilevel optimization problem as in (11).

When assuming this optimistic approach, one assumes at the same time that the decision maker on the lower level allows the decision maker on the

upper level to utilize any solution from the set of efficient solutions of the lower level, which might not be realistic. See also the discussion in [45, 47].

In [45], also the case where the lower level decision maker has sufficient power to make a deterministic decision from the set of efficient solutions of the lower level is studied. This is formulated by using a value function on the objective functions of the lower level, which needs to be known in this setting. By using a value function the lower level problem is somehow reduced to a single-objective problem. To overcome the difficulty of determining the lower level value function, Sinha and co-authors suggest in [46] to handle uncertainties in the lower level value function.

For the multiobjective optimization problem on the upper level we assume that the partial ordering is given by the closed pointed convex cone $K^2 \subseteq \mathbb{R}^{m_2}$ and for the lower level by the closed pointed convex cone $K^1 \subseteq \mathbb{R}^{m_1}$. 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, using (9), we write instead of $x \in \Psi(y)$

$$x \in \mathcal{E}(f(G(y), y), K^1) =: \mathcal{E}_y(f(G), K^1), \quad (12)$$

with $\mathcal{E}(f(G(y), y), K^1)$ the set of K^1 -minimal points of the multiobjective optimization problem (8) parameterized by y . Hence we study

$$\begin{aligned} & \min_{x,y} F(x, y) \\ & \text{subject to the constraints} \\ & x \in \mathcal{E}_y(f(G), K^1), \\ & y \in \tilde{G}. \end{aligned} \quad (13)$$

A helpful result, which can be used for numerical approaches, is 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. The set of feasible points Ω of the upper level problem in (13), also called induced set, is given by

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

We show that the set Ω is equivalent to the set of \hat{K} -minimal points of the multiobjective optimization problem

$$\begin{aligned} & \min_{x,y} \hat{f}(x, y) := \begin{pmatrix} f(x, y) \\ y \end{pmatrix} \\ & \text{subject to the constraints} \\ & (x, y) \in G, \\ & y \in \tilde{G} \end{aligned} \quad (14)$$

w. r. t. the ordering cone

$$\hat{K} := K^1 \times \{0\} \subset \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}.$$

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 (14). This is shown in the following theorem. There is also a close relation to the reformulation using the function F in (6), see [41].

Theorem 5.1. *Let $\hat{\mathcal{E}}$ be the set of \hat{K} -minimal points of the multiobjective optimization problem (14) with $\hat{K} = K^1 \times \{0\}$. Then it holds $\Omega = \hat{\mathcal{E}}$.*

Proof.

$$\begin{aligned}
\text{We have } & (\bar{x}, \bar{y}) \in \Omega \\
\Leftrightarrow & \bar{x} \in \mathcal{E}_{\bar{y}}(f(G), K^1) \quad \wedge \quad \bar{y} \in \tilde{G} \\
\Leftrightarrow & (\nexists x \in G(\bar{y}) \text{ with } f(\bar{x}, \bar{y}) \in f(x, \bar{y}) + K^1 \setminus \{0\}) \\
& \wedge \quad \bar{y} \in \tilde{G} \quad \wedge \quad \bar{x} \in G(\bar{y}) \\
\Leftrightarrow & (\nexists (x, y) \in G \text{ with } f(\bar{x}, \bar{y}) \in f(x, y) + K^1 \setminus \{0\} \quad \wedge \quad y = \bar{y}) \\
& \wedge \quad \bar{y} \in \tilde{G} \quad \wedge \quad (\bar{x}, \bar{y}) \in G \\
\Leftrightarrow & (\nexists (x, y) \in G \text{ with} \\
& \quad \left(\begin{array}{c} f(\bar{x}, \bar{y}) \\ \bar{y} \end{array} \right) \in \left(\begin{array}{c} f(x, y) \\ y \end{array} \right) + (K^1 \times \{0\}) \setminus \{0\}) \\
& \wedge \quad \bar{y} \in \tilde{G} \quad \wedge \quad (\bar{x}, \bar{y}) \in G \\
\Leftrightarrow & (\nexists (x, y) \in G \text{ with } \hat{f}(\bar{x}, \bar{y}) \in \hat{f}(x, y) + \hat{K} \setminus \{0\}) \\
& \wedge \quad \bar{y} \in \tilde{G} \quad \wedge \quad (\bar{x}, \bar{y}) \in G \\
\Leftrightarrow & (\bar{x}, \bar{y}) \in \hat{\mathcal{E}}. \quad \square
\end{aligned}$$

□

A multiobjective bilevel optimization problem is a quite challenging problem in case there are multiple objectives on each level. Already in case there are only multiple objectives on the upper level, such semivectorial bilevel problems become quite difficult. This is even more the case when the lower level problem does not have unique optimal solutions. We illustrate this with an example taken from [39, Ex. 4.14]. This example also gives insight to the optimistic approach.

Example 5.2. *Example on semivectorial bilevel problem We study a bilevel optimization problem as in (10) with the lower level problem*

$$\min_{x \in \mathbb{R}} \{ \max\{0, -xy\} \mid x \leq 1 \}.$$

Then we have

$$\Psi(y) = \begin{cases}] - \infty, 1] & \text{if } y = 0, \\ [0, 1] & \text{if } y =]0, 1]. \end{cases}$$

For the lower level problem we are now interested in the problem

$$\begin{aligned}
& \text{“min”}_{y \in \mathbb{R}} F(x(y), y) \\
& \text{subject to the constraints} \\
& x(y) \in \Psi(y), \\
& y \in [0, 1]
\end{aligned}$$

with

$$F(x, y) = \left(\begin{array}{c} x - y - 1 \\ y - \max\{0, x\} \end{array} \right).$$

As the lower level solutions are not unique, to each $y \in [0, 1]$ we can define the sets

$$\tilde{F}(y) := \bigcup_{x \in \Psi(y)} F(x, y)$$

where

$$\tilde{F}(y) = \begin{cases} \text{conv}\{(-1, 0), (0, -1)\} \cup \{(u, 0) \mid u \in]-\infty, -1]\} & \text{if } y = 0 \\ \text{conv}\{(-y-1, y), (-y, y-1)\} & \text{if } y \in]0, 1] \end{cases}$$

Then $(\hat{x}, \hat{y}) = (\hat{x}, 1)$ with $\hat{x} \in \Psi(1) = [0, 1]$ and $F(\hat{x}, \hat{y}) = (\hat{x} - 2, 1 - \hat{x})$ is not an optimal solution of the optimistic formulation

$$\begin{aligned} & \min_{x, y} F(x(y), y) \\ & \text{subject to the constraints} \\ & x(y) \in \Psi(y) \\ & y \in [0, 1]. \end{aligned}$$

This can be seen by considering $(x, y) = (x, 0)$ with $x \in \Psi(0) \cap]-\infty, -1]$. These points have smaller objective function values: $F(x, 0) = (x - 1, 0)^\top$. By comparing just the nondominated elements of the sets $\tilde{F}(y)$, i.e. by solving a set optimization problem, as proposed in [39, p. 60], there would also be solutions with $\hat{y} = 1$.

For $K^1 = \mathbb{R}_+^{m_1}$ Lv and Wan use in [32, 33] a weighted-sum approach, as discussed in Section 3, to characterize the optimal solutions of the lower level in case the lower level problem is convex and some regularity assumptions are satisfied. For the theoretical verification see Theorem 3.1. Then, instead of (13), they study

$$\begin{aligned} & \min_{x, y, w} F(x, y) \\ & \text{subject to the constraints} \\ & y \in \tilde{G}, \\ & \sum_{i=1}^{m_1} w_i = 1, \\ & w \in \mathbb{R}_+^{m_1} \\ & x \in \text{argmin} \{w^\top f(x, y) \mid (x, y) \in G\}. \end{aligned}$$

Hence, the weighting vector w , which is here the scalarization parameter, is interpreted as a new upper level variable, and thus the original semivectorial bilevel programming problem is transformed into a standard bilevel programming problem (cf. [12, Theorem 3.1]). Thereby it is assumed that one is interested in weakly efficient solutions of the lower level—and not in efficient solutions only. While this approach can be done in view of globally optimal solutions, one has to take care in case one uses locally optimal solutions on the lower level: Dempe and Mehlitz show in [12] that the reformulation using the weighted-sum scalarization and the additional variable w might have locally optimal solutions which are not locally optimal for the original problem.

In a second step, in [32, 33], the constraint

$$x \in \operatorname{argmin} \left\{ w^\top f(x, y) \mid (x, y) \in G \right\}$$

is replaced by the KKT-conditions of this scalar-valued optimization problem.

As done in these reformulations, many authors take the weakly efficient elements instead of the efficient, i.e. they replace the condition $x \in \mathcal{E}(f(G(y), y), K^1)$ from the lower level problem by $x \in \mathcal{E}^w(f(G(y), y), K^1)$. The next example shows that this can have a strong impact. We use for this examples ideas from [12].

Example 5.3. *Example on weakly efficient and efficient solutions* We study the optimistic semivectorial bilevel optimization problem

$$\begin{aligned} & \min_{x,y} x + 2y \\ & \text{subject to the constraints} \\ & x \in \Psi(y) = \operatorname{argmin} \{ (xy, 1 - x) \mid x \in [0, 1] \}, \\ & y \in [0, 1]. \end{aligned}$$

Then we have

$$\Psi(y) = \begin{cases} \{1\} & \text{if } y = 0, \\ [0, 1] & \text{if } y \in (0, 1]. \end{cases}$$

The bilevel problem has no optimal solution: for $y = 0$ the objective function value is 1 and for $y > 0$ it is $2y$. In case one takes the set of weakly efficient solutions of the lower level, which we denote by $\Psi^w(y)$, one gets

$$\Psi^w(y) = [0, 1] \quad \text{for } y \in [0, 1]$$

and as optimal value of the bilevel problem we obtain 0. The minimal solution is then $(x, y) = (0, 0)$.

Instead of enlarging the feasible set by taking the weakly efficient solutions instead of the efficient solutions, one could also use the so-called properly efficient solutions, which are a subset of the efficient solutions. Those can be completely characterized by linear scalarizations, at least in the convex case, by using elements $w \in K^\#$. Such an approach was discussed by Gebhardt and Jahn in [25]. The advantage is that the gained optimal values of the bilevel problem are then upper bounds on the optimal values of the original problem, and, what is more, that the gained optimal solutions are in fact feasible for the original problem, as any properly efficient solution is also efficient.

Finally, we would like to draw the attention to additional difficulties which can arise in case of constraints on the upper level which couple the upper and the lower level variables. Thus, in the remaining of this section, we study instead of problem (13) the more general formulation

$$\begin{aligned} & \min_{x,y} F(x, y) \\ & \text{subject to the constraints} \\ & x \in \mathcal{E}_y(f(G), K^1), \\ & (x, y) \in \tilde{G} \end{aligned} \tag{15}$$

with the constraint set $\tilde{G} \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$. This results in a coupling of the upper level variable y and the lower level variable x . Then

$$\Omega' := \{(x, y) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \mid x \in \mathcal{E}_y(f(G), K^1), (x, y) \in \tilde{G}\}$$

denotes the induced set of the problem (15).

First notice that the constraint $(x, y) \in \tilde{G}$ from the upper level cannot just be moved to the lower level as simple examples, see for instance [14, Example 7.4], show. This has also a practical interpretation. The same constraint on the lower level restricting the constraint set on the lower level has a different meaning as on the upper level. There, the feasibility is restricted after the determination of the minimal solution of the lower level and is thus an implicit constraint. For more details see [8, pp.25f].

A generalization of the results from Theorem 5.1 is not possible as discussed in [14, Chapter 7]. We also illustrate this with the following example.

Example 5.4. *Example on coupled constraints* We consider the biobjective bilevel optimization problem

$$\begin{aligned} \min_{y \in \mathbb{R}} F(x, y) &= \begin{pmatrix} x_1 - y \\ x_2 \end{pmatrix} \\ &\text{subject to the constraints} \\ x = x(y) &\in \operatorname{argmin}_{x \in \mathbb{R}^2} \left\{ f(x, y) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mid (x, y) \in G \right\}, \\ &(x, y) \in \tilde{G} \end{aligned}$$

with $n_1 = 2$, $n_2 = 1$, $K^1 = K^2 = \mathbb{R}_+^2$, $m_1 = m_2 = 2$,

$$G := \{(x, y) \in \mathbb{R}^3 \mid \|x\|_2^2 \leq y^2\}$$

and

$$\tilde{G} := \{(x, y) \in \mathbb{R}^3 \mid 0 \leq y \leq 1, x_1 + x_2 \geq -1\}.$$

Then

$$\mathcal{E}_y(f(G), K^1) = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid \|x\|_2^2 = y^2, x_1 \leq 0, x_2 \leq 0\}$$

and thus we obtain

$$\Omega' = \{(x, y) \in \mathbb{R}^3 \mid \|x\|_2^2 = y^2, x_1 \leq 0, x_2 \leq 0, 0 \leq y \leq 1, x_1 + x_2 \geq -1\}.$$

Let $\hat{\mathcal{E}}'$ be the set of $(\mathbb{R}_+^2 \times \{0\})$ -minimal points of the tricriteria optimization problem

$$\begin{aligned} \min_{x, y} &\begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} \\ &\text{subject to the constraints} \end{aligned}$$

$$(x, y) \in G \cap \tilde{G} = \{(x, y) \in \mathbb{R}^3 \mid 0 \leq y \leq 1, x_1 + x_2 \geq -1, \|x\|_2^2 \leq y^2\}$$

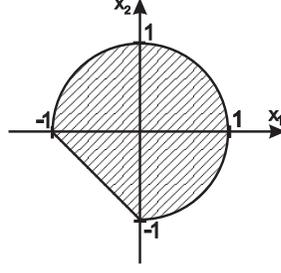


Figure 1: The set $\{x \in \mathbb{R}^2 \mid (x, y) \in G \cap \tilde{G}, y = 1\}$ of Example 5.4, cf. [14].

similar to (14). While it holds $\Omega' \subset \hat{\mathcal{E}}'$ (cf. [14, Theorem 7.5]) we show that it does not hold $\hat{\mathcal{E}}' \subset \Omega'$.

For $(\bar{x}, \bar{y}) = (-\frac{1}{2}, -\frac{1}{2}, 1)^\top$ we have $(\bar{x}, \bar{y}) \in \hat{\mathcal{E}}'$ because $(\bar{x}, \bar{y}) \in G \cap \tilde{G}$ and there is no $(x, y) \in G \cap \tilde{G}$ with

$$\begin{aligned} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{y} \end{pmatrix} &\in \begin{pmatrix} x_1 \\ x_2 \\ y \end{pmatrix} + (\mathbb{R}_+^2 \times \{0\}) \setminus \{0\} \\ \Leftrightarrow \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} &\in \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbb{R}_+^2 \setminus \{0\} \wedge y = \bar{y}. \end{aligned}$$

The set $\{x \in \mathbb{R}^2 \mid (x, y) \in G \cap \tilde{G}, y = 1\}$ is illustrated in Fig. 1. However, $(\bar{x}, \bar{y}) \notin \Omega'$ because $\|\bar{x}\|_2^2 = \frac{1}{2} \neq \bar{y}^2$.

In this example, as the induced set can be determined explicitly, the minimal solution set can be calculated by solving the biobjective optimization problem $\min_{x,y} \{F(x, y) \mid (x, y) \in \Omega'\}$. We get as solution set of the multiobjective bilevel optimization problem:

$$\mathcal{S}^{\min} = \left\{ (x_1, x_2, y) \mid x_1 = -1 - x_2, x_2 = -\frac{1}{2} \pm \frac{1}{4} \sqrt{8y^2 - 4}, y \in \left[\frac{\sqrt{2}}{2}, 1 \right] \right\}.$$

The image $\{F(x_1, x_2, y) \mid (x_1, x_2, y) \in \mathcal{S}^{\min}\}$ of this set is

$$\left\{ (z_1, z_2) \in \mathbb{R}^2 \mid z_1 = -1 - z_2 - y, z_2 = -\frac{1}{2} \pm \frac{1}{4} \sqrt{8y^2 - 4}, y \in \left[\frac{\sqrt{2}}{2}, 1 \right] \right\}.$$

These sets are shown in Fig. 2.

6 Numerical Methods

As stated in the introduction, we concentrate in this paper on nonlinear problems. Procedures for solving linear multiobjective bilevel problems are presented e. g. in [36].

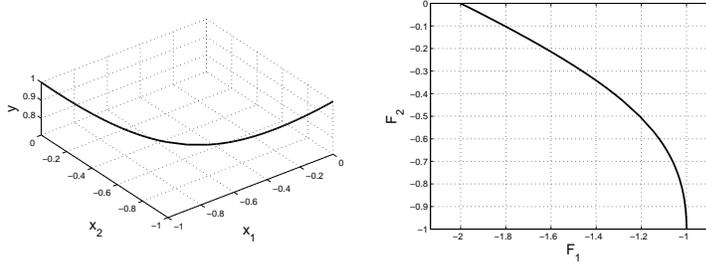


Figure 2: Solution set \mathcal{S}^{\min} of the biobjective bilevel optimization problem of Example 5.4 and the image $F(\mathcal{S}^{\min})$.

Much less papers are dealing with nonlinear multiobjective bilevel problems: Shi and Xia, [43, 44], present an interactive method, also using a scalarization (the ε -constraint scalarization, cf. [47]). Osman, Abo-Sinna et al., [1, 37], propose the usage of fuzzy set theory for convex problems. Teng et al., [49], give an approach for a convex multiperson multiobjective bilevel problem.

Bonnell and Morgan consider in [2] a semivectorial bilevel optimization problem and propose a solution method based on a penalty approach. However, no numerical results are given.

More recently, in [17, 14] Theorem 5.1 was used for developing a numerical algorithm. Based on that result, the efficient set of the multiobjective optimization problem (14) was calculated. This gives the set of feasible points of the upper level problem, which then can be solved. In general, it is not possible to determine the whole solution set of problem (14). We can only calculate a representation or approximation of this set. In [17, 14] it is taken care that this approximation is generated with a high quality in the pre-image space. And then, in a second step, based on sensitivity information, this approximation is refined depending on the behavior of the upper level function. For determining single efficient points of problem (14) the scalarization according to Pascoletti and Serafini, see Theorem 3.3, is used. Thereby, one can also make use of the result given in Lemma 2.2. The approach is suitable for low dimensions of the upper level variable only. Summing up, the algorithm consists of the following steps:

- 1) Determine equidistant representations of the efficient sets of the multiobjective problems (8) 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 give the numerical results of this algorithm for one example, cf. [14,

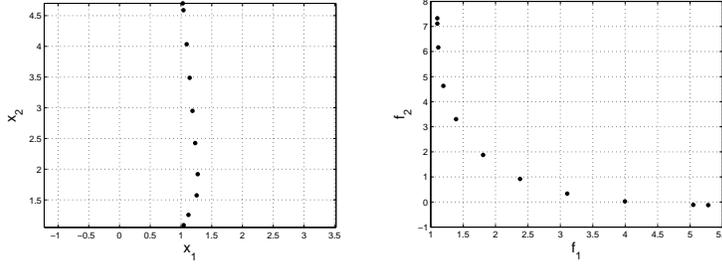


Figure 3: Representation (a) of the efficient set and (b) the nondominated set of problem (8) for $y = 1.8$.

Chapter 7]

Example 6.1. *Numerical example* 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

$$x \in \operatorname{argmin}_{x \in \mathbb{R}^2} \left\{ \begin{pmatrix} f_1(x,y) \\ f_2(x,y) \end{pmatrix} \mid (x,y) \in G \right\},$$

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

with $f_1, f_2: \mathbb{R}^3 \rightarrow \mathbb{R}$,

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

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

and

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

We apply the algorithm from [17, 14] with distance $\beta = 0.6$ for discretizing the interval $[0, 10]$. For example, for $y = 1.8$ one obtains the representation of the efficient set of problem (8) as shown in Figure 3(a). In the image space of the lower level problem this representation is shown in Figure 3(b).

The representation A^0 of the efficient sets of (8) for all discretization points y of the interval $[0, 10]$ is shown in Figure 4(a), which is an approximation of the solution set of problem (14) and hence of the set of feasible points of the upper level problem. In Figure 4(b) the set $F(A^0)$ is given. Here the image points under F of the set $\mathcal{E}(F(A^0), \mathbb{R}_+^2)$ are marked with circles and are connected with lines.

In several iterations the discretization of the feasible set of the upper level problem is refined around the efficient points which results in the representation A^f , see Figure 5. Then the algorithm is stopped as only small improve-

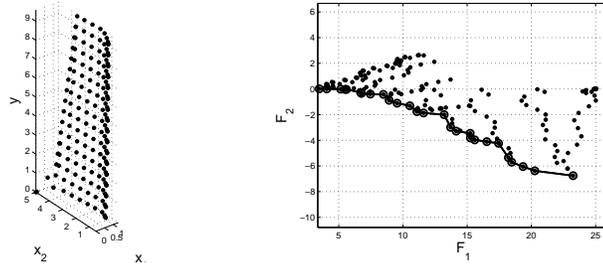


Figure 4: (a) Approximation A^0 of the set of feasible points Ω and (b) the image $F(A^0)$ of this set under F .

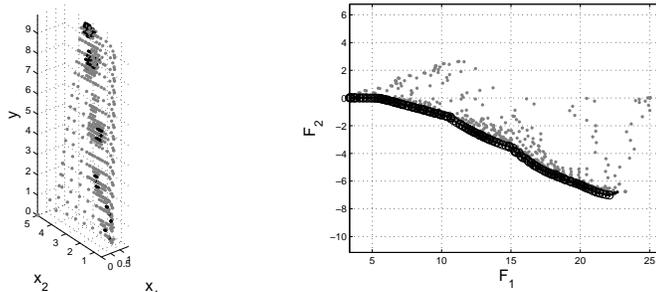


Figure 5: (a) Final approximation A^f of the set of feasible points Ω and (b) the image $F(A^f)$ of this set under F .

ments for the approximation of the efficient set of the multiobjective bilevel optimization problem were gained by the last iteration.

Later, another numerical deterministic approach for nonlinear problems was proposed by Gebhardt and Jahn in [25]. They propose to use a multiobjective search algorithm with a subdivision technique. The studied problems have multiple objectives on both levels. The authors also discuss optimality conditions for replacing the multiobjective problem on the lower level. Instead of taking the weakly efficient solutions of the lower level, they suggest to take the properly efficient solutions. By doing this the feasible set of the upper level gets smaller (instead of larger by using weakly efficient elements) and thus one gets upper bounds, and minimal solutions which are in fact feasible for the original problem. However, the authors do not use this for their numerical approach. Instead, the main ingredient of their algorithm is a subdivision technique for multiobjective problems which uses randomly generated points in boxes, where the size is decreased when the algorithm proceeds. It is assumed that the feasible set of the upper level variable \tilde{G} is such that one can easily work with a discretization of it, for instance an interval. While the procedure is appropriate only for small number of variables, it is capable of detecting globally optimal solutions also for highly nonlinear problems.

Also, the construction of numerical test instances is an important aspect. Test instances which are scalable in the number of variables and the number

of objectives and where the set of optimal solutions are known are important to evaluate and compare numerical approaches. Deb and Sinha propose five of such test problems in [5]. For the test instances the optimistic approach is chosen.

There have also been some numerical approaches for solving bilevel multiobjective optimization problems by using evolutionary algorithms. An attempt is for instance given by Deb and Sinha in [6]. A hybrid solver combined with a local solver was proposed by the same authors in [7].

References

- [1] M. Abo-Sinna. A bi-level non-linear multi-objective decision making under fuzziness. *Opsearch*, 38(5):484–495, 2001.
- [2] H. Bonnel and J. Morgan. Semivectorial bilevel optimization problem: Penalty approach. *J. Optim. Theory Appl.*, 131(3):365–382, 2006.
- [3] H. Bonnel, I. Tadjihoundé, C. Udriște. Semivectorial Bilevel Optimization on Riemannian Manifolds. *J. Optim. Theory Appl.*, 167(2):464–486, 2015.
- [4] T.D. Chuong. Optimality conditions for nonsmooth multiobjective bilevel optimization problems. *Annals of Operations Research*, doi=10.1007/s10479-017-2734-6, 2017.
- [5] K. Deb and A. Sinha. Constructing test problems for bilevel evolutionary multi-objective optimization. *IEEE Congress on Evolutionary Computation, Trondheim*, 1153–1160, 2009.
- [6] K. Deb and A. Sinha. Solving Bilevel Multi-Objective Optimization Problems Using Evolutionary Algorithms. *EMO 2009: Evolutionary Multi-Criterion Optimization* 110–124, 2009.
- [7] K. Deb and A. Sinha. An Efficient and Accurate Solution Methodology for Bilevel Multi-Objective Programming Problems Using a Hybrid Evolutionary-Local-Search Algorithm. *Evolutionary Computation*, 18(3):403–449, 2010.
- [8] S. Dempe. *Foundations of bilevel programming* Kluwer, 2012.
- [9] S. Dempe and S. Franke. Bilevel optimization problems with vectorvalued objective functions in both levels Working Paper Department of Mathematics and Computer Science, TU Bergakademie Freiberg, 2012.
- [10] S. Dempe, N. Gadhi and A.B. Zemkoho. New Optimality Conditions for the Semivectorial Bilevel Optimization Problem. *J. Optim. Theory Appl.*, 157(1):54–74, 2013.
- [11] S. Dempe and N. Gadhi. Optimality Conditions for Bilevel Vector Optimization Problems with a Variable Ordering Structure. *Numerical Functional Analysis and Optimization*, 38:8:988–1007, 2017.

- [12] S. Dempe and P. Mehlitz. Semivectorial bilevel programming versus scalar bilevel programming. *Optimization*, 2019.
- [13] M. Ehrgott. *Multicriteria optimisation*, volume 491 of *Lect. Notes Econ. math. Syst.* Springer, Berlin, 2000.
- [14] G. Eichfelder. *Adaptive Scalarization Methods in Multiobjective Optimization*, Springer, Heidelberg, 2008.
- [15] G. Eichfelder. Scalarizations for adaptively solving multi-objective optimization problems. *Comput. Optim. Appl.*, 44(2): 249–273, 2009.
- [16] G. Eichfelder. An adaptive scalarization method in multi-objective optimization. *SIAM J. Optim.*, 19(4): 1694–1718, 2009.
- [17] G. Eichfelder. Multiobjective bilevel optimization. *Mathematical Programming*, Ser. A, 123(2):419–449, 2010.
- [18] G. Eichfelder and T.X.D. Ha. Optimality conditions for vector optimization problems with variable ordering structures. *Optimization* 62:597–627, 2013.
- [19] G. Eichfelder. Numerical Procedures in Multiobjective Optimization with Variable Ordering Structures. *J. Optim. Theory Appl.*, 162(2):489–514, 2014.
- [20] G. Eichfelder and M. Pilecka. Set Approach for Set Optimization with Variable Ordering Structures Part II: Scalarization Approaches *J. Optim. Theory Appl.*, 171(3):947–963, 2016.
- [21] G. Eichfelder and M. Pilecka. Ordering Structures and Their Applications. In: *Applications of Nonlinear Analysis*, Rassias, T.M. (Edt.), Springer, 256–304, 2018.
- [22] J. Fliege and L. N. Vicente. Multicriteria approach to bilevel optimization. *J. Optimization Theory Appl.*, 131(2):209–225, 2006.
- [23] N. Gadhi and S. Dempe. Necessary Optimality Conditions and a New Approach to Multiobjective Bilevel Optimization Problems. *J. Optim. Theory Appl.*, 155(1) 100–114, 2012.
- [24] N. Gadhi and S. Dempe. Sufficient Optimality Conditions for a Bilevel Semivectorial D.C. Problem. *Numerical Functional Analysis and Optimization*, 39(15):1622–1634, 2018.
- [25] E. Gebhardt and J. Jahn. Global solver for nonlinear bilevel vector optimization problems. *Pac. J. Optim.* 5(3):387–401, 2009.
- [26] C. Gerstewitz (Tammer). Nichtkonvexe Dualität in der Vektoroptimierung. *Wissensch. Zeitschr. TH Leuna-Merseburg* 25:357–364, 1983.
- [27] C. Gerstewitz (Tammer) and E. Iwanow. Dualität für nichtkonvexe Vektoroptimierungsprobleme. *Wissensch. Zeitschr. der Techn. Hochschule Ilmenau* 31:61–81, 1985.
- [28] Y. Haimes, L. Lasdon, and D. Wismer. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Trans. Syst. Man Cybern.*, 1:296–297, 1971.

- [29] J. Jahn. *Vector Optimization: Theory, Applications and Extensions*. Springer, Berlin, 2004.
- [30] J. Jahn. Multiobjective search algorithm with subdivision technique. *Comput. Optim. Appl.*, 35:161–175, 2006.
- [31] P. Loridan. ε -solutions in vector minimization problems. *J. Optimization Theory Appl.*, 43:265–276, 1984.
- [32] Y. Lv, Z. Wan. A solution method for the optimistic linear semivectorial bilevel optimization problem. *Journal of Inequalities and Applications* Lv and Wan *Journal of Inequalities and Applications* 164, 2014.
- [33] Y. Lv, Z. Wan. A Smoothing Method for Solving Bilevel Multiobjective Programming Problems. *Journal of the Operations Research Society of China*, 2(4):511–525, 2014.
- [34] S. Marglin. *Public investment criteria*. MIT Press, Cambridge, 1967.
- [35] K. Miettinen. *Nonlinear multiobjective optimization*. Kluwer Academic Publishers, Boston, 1999.
- [36] I. Nishizaki and M. Sakawa. Stackelberg solutions to multiobjective two-level linear programming problems. *J. Optimization Theory Appl.*, 103(1):161–182, 1999.
- [37] M. Osman, M. Abo-Sinna, A. Amer, and O. Emam. A multi-level nonlinear multi-objective decision-making under fuzziness. *Appl. Math. Comput.*, 153(1):239–252, 2004.
- [38] A. Pascoletti and P. Serafini. Scalarizing vector optimization problems. *J. Optimization Theory Appl.*, 42(4):499–524, 1984.
- [39] M. Pilecka. Set-valued optimization and its application to bilevel optimization. Dissertation, Technische Universität Bergakademie Freiberg, 2016.
- [40] J. Prohaska. Optimierung von Spulenkonfigurationen zur Bewegung magnetischer Sonden. Diplomarbeit, Univ. Erlangen-Nürnberg, Germany, 2005.
- [41] S. Ruuska, K. Miettinen and M.M. Wiecek. Connections Between Single-Level and Bilevel Multiobjective Optimization. *J. Optimization Theory Appl.*, 153(1): 60–74, 2012.
- [42] S. Ruzika and M. Wiecek. Approximation methods in multiobjective programming. *J. Optimization Theory Appl.*, 126(3):473–501, 2005.
- [43] X. Shi and H. Xia. Interactive bilevel multi-objective decision making. *J. Oper. Res. Soc.*, 48(9):943–949, 1997.
- [44] X. Shi and H. Xia. 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.

- [45] A. Sinha, P. Malo and K. Deb. Towards Understanding Bilevel Multi-objective Optimization with Deterministic Lower Level Decisions. in *Evolutionary Multi-Criterion Optimization*, A. Gaspar-Cunha, C. Henggeler Antunes and C.C. Coello (Eds.), 26–443, 2015.
- [46] A. Sinha, P. Malo, K. Deb, P. Korhonen and J. Wallenius. Solving Bilevel Multicriterion Optimization Problems With Lower Level Decision Uncertainty *IEEE Transactions on Evolutionary Computation* 20(2):199–217, 2015.
- [47] A. Sinha, P. Malo and K. Deb. A Review on Bilevel Optimization: From Classical to Evolutionary Approaches and Applications *IEEE Transactions on Evolutionary Computation*, 22(2):276–295, 2018.
- [48] T. Staib. On two generalizations of Pareto minimality. *J. Optimization Theory Appl.*, 59(2):289–306, 1988.
- [49] C.-X. Teng, L. Li, and H.-B. Li. A class of genetic algorithms on bilevel multi-objective decision making problem. *J. of Systems Science and Systems Engineering*, 9(3):290–296, 2000.
- [50] J.J. Ye. Necessary Optimality Conditions for Multiobjective Bilevel Programs. *Mathematics of Operations Research*, 36(1): 165–184, 2011.
- [51] Y. Yin. Multiobjective bilevel optimization for transportation planning and management problems. *J. Advanced Transportation*, 36(1):93–105, 2000.