Deutsche Forschungsgemeinschaft
Priority Programme 1962

Semivectorial Bilevel Programming versus Scalar Bilevel Programming

Stephan Dempe, Patrick Mehlitz

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

Preprint Number SPP1962-082

received on October 9, 2018

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

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

Semivectorial bilevel programming versus scalar bilevel programming

With applications to existence theory for semivectorial bilevel optimal control problems

Stephan Dempe* Patrick Mehlitz†

October 9, 2018

We consider an optimistic semivectorial bilevel programming problem in Banach spaces. The associated lower level multicriterial optimization problem is assumed to be convex w.r.t. its decision variable. This property implies that all its weakly efficient points can be computed applying the weighted-sum-scalarization technique. Consequently, it is possible to replace the overall semivectorial bilevel programming problem by means of a standard bilevel programming problem whose upper level variables comprise the set of suitable scalarization parameters for the lower level problem. In this note, we consider the relationship between this surrogate bilevel programming problem and the original semivectorial bilevel programming problem. As it will be shown, this is a delicate issue as long as locally optimal solutions are investigated. The obtained theory is applied in order to derive existence results for semivectorial bilevel programming problems with not necessarily finite-dimensional lower level decision variables. Some regarding examples from bilevel optimal control are presented.

Keywords: Bilevel programming, Existence theory, Multiobjective optimization, Optimal control
MSC: 49J20, 49J27, 90C29, 90C48

1 Introduction

Semivectorial bilevel optimization problems possess a similar structure as standard bilevel programming problems, see [Bard, 1998, Dempe, 2002, Shimizu et al., 1997] for an intro-

*Technische Universität Bergakademie Freiberg, Faculty of Mathematics and Computer Science, 09596 Freiberg, Germany, dempe@math.tu-freiberg.de, http://www.mathe.tu-freiberg.de/~dempe/
†Brandenburgische Technische Universität Cottbus-Senftenberg, Institute of Mathematics, Chair of Optimal Control, 03046 Cottbus, Germany, mehlitz@b-tu.de, https://www.b-tu.de/fg-optimale-steuerung/team/dr-patrick-mehlitz

duction to hierarchical optimization, apart from the fact that the objective mapping of the underlying lower level (or follower's) problem is a vector function. This problem class has been introduced in [Bonnel, 2006, Bonnel and Morgan, 2006]. Noting that the common notion of a minimizer does not apply to multiobjective programs, the lower level decision maker has to compute e.g. the set of efficient or weakly efficient points for any fixed value of the upper level (or leader's) variable in general, see [Ehrgott, 2005, Jahn, 2004] for an introduction to multicriteria optimization. Apart from the general observation that the lower level decision maker may have multiple goals to optimize, semivectorial bilevel programming applies to real-world problems connected to electricity markets, see [Alves and Antunes, 2018] and the references therein, as well as to the modeling of inverse multicriterial optimization problems. Necessary optimality conditions for finite-dimensional semivectorial bilevel optimization problems were investigated in [Dempe et al., 2013, Liu et al., 2014]. Existence results and necessary optimality conditions for a semivectorial bilevel optimal control problem of ordinary differential equations have been derived in [Bonnel and Morgan, 2012, 2013]. Algorithms for the numerical solution of semivectorial bilevel programming problems with fully linear lower level problem can be found in [Ren and Wang, 2016, Zheng and Wan, 2011].

Supposing that the lower level decision maker has to solve a convex multicriterial optimization problem for any fixed value of the upper level variable, the overall weak Pareto front can be computed by means of a simple scalarization approach where a weighted sum of all lower level objective functionals is minimized, see [Ehrgott, 2005]. Interpreting the scalarization parameters as new upper level variables, the original semivectorial bilevel programming problem can be transformed into a standard bilevel programming problem. This idea was investigated in the papers [Bonnel and Morgan, 2012, 2013, Dempe et al., 2013] where the authors commented on the relation of this surrogate problem to the original semivectorial bilevel programming problem.

In this note, we will show by means of examples that the results obtained in [Bonnel and Morgan, 2012, Dempe et al., 2013] are not fully correct. Afterwards, we will amend these achievements. In order to guarantee applicability to the setting of semivectorial bilevel optimal control, all investigations will be carried out in the setting of Banach spaces. Finally, we apply our findings in order to study the existence of solutions to certain semivectorial bilevel programming problems where the lower level decision variable does not need to be chosen from a finite-dimensional space. This allows us to infer existence results for special classes of semivectorial bilevel optimal control problems.

This manuscript is organized as follows: In the remaining part of Section 1, we comment on the notation we exploit in this paper. We introduce the semivectorial bilevel programming problem of interest as well as its associated surrogate standard bilevel program in Section 2. Section 3 is dedicated to the study of the relationship between both hierarchical optimization problems. A theoretical existence result for semivectorial bilevel programming problems is derived in Section 4. Afterwards, the latter is applied to derive the existence of optimal solutions for two important classes of semivectorial bilevel optimal control problems. Finally, we briefly summarize the obtained results in Section 5.

Notation Let us briefly clarify some notation used in this note.

For some Banach space $\mathcal{X}$, $\|\cdot\|_{\mathcal{X}}$ denotes its norm. Furthermore, $\mathbb{U}^{\varepsilon}_{\mathcal{X}}(x)$ represents the open ε-ball around $x \in \mathcal{X}$ w.r.t. the norm $\|\cdot\|_{\mathcal{X}}$. Strong and weak convergence of a sequence $\{x_l\}_{l\in\mathbb{N}} \subset \mathcal{X}$ to some $\bar{x} \in \mathcal{X}$ will be denoted by $x_l \to \bar{x}$ and $x_l \rightharpoonup \bar{x}$, respectively. For a set $A \subset \mathcal{X}$, $\operatorname{cone} A$ and $\operatorname{conv} A$ denote the conic hull and the convex hull of A, respectively.

Let $\mathcal{Y}$ be another Banach space and assume that $A \subset \mathcal{X}$ is closed. A mapping Γ which assigns to any $x \in A$ a subset $\Gamma(x) \subset \mathcal{Y}$ is called a set-valued mapping or multifunction and will be denoted by $\Gamma\colon A \rightrightarrows \mathcal{Y}$. The sets defined by $\operatorname{dom}\Gamma := \{x \in A \,|\, \Gamma(x) \neq \varnothing\}$ and $\operatorname{gph}\Gamma := \{(x,y) \in A \times \mathcal{Y} \,|\, y \in \Gamma(x)\}$ are referred to as domain and graph of Γ, respectively. Fix some point $\bar{x} \in A$. The mapping Γ is said to be closed at $\bar{x}$ if for any sequences $\{x_l\}_{l\in\mathbb{N}} \subset A$ and $\{y_l\}_{l\in\mathbb{N}} \subset \mathcal{Y}$ satisfying $x_l \to \bar{x}$ and $y_l \to \bar{y}$ for some $\bar{y} \in \mathcal{Y}$ as well as $y_l \in \Gamma(x_l)$ for all $l \in \mathbb{N}$, we have $\bar{y} \in \Gamma(\bar{x})$. On the other hand, Γ is called lower semicontinuous at $\bar{x}$ if for any open set $V \subset \mathcal{Y}$ satisfying $\Gamma(\bar{x}) \cap V \neq \varnothing$, there exists $\varepsilon > 0$ such that $\Gamma(x) \cap V \neq \varnothing$ is satisfied for all $x \in A \cap \mathbb{U}^{\varepsilon}_{\mathcal{X}}(\bar{x})$. We call Γ closed (lower semicontinuous) if it is closed (lower semicontinuous) at all points of its domain.

For any two vectors $x, y \in \mathbb{R}^k$, $x \cdot y$ denotes their Euclidean inner product. If not stated otherwise, we equip $\mathbb{R}^k$ with the Euclidean norm $|\cdot|_2$. Furthermore, we will exploit $\Delta^k \subset \mathbb{R}^k$ to denote the standard simplex in $\mathbb{R}^k$, i.e. we have

$$\Delta^k := \left\{ z \in \mathbb{R}^k \,\middle|\, z \geq 0,\ \textstyle\sum_{j=1}^k z_j = 1 \right\}.$$

For a vector function $h\colon \mathcal{X} \to \mathbb{R}^k$ whose components are denoted by $h_1, \ldots, h_k\colon \mathcal{X} \to \mathbb{R}$ and a nonempty set $M \subset \mathcal{X}$, let us investigate the multicriterial optimization problem

$$\left.\begin{array}{c} h_1(x) \\ \vdots \\ h_k(x) \end{array}\right\} \to \text{“min”} \qquad x \in M. \tag{MOP}$$

Due to the presence of multiple objectives in case $k \geq 2$, the standard notions of local and global minimizers are not applicable to (MOP) which is why we used quotation marks around the min-operator. Instead, weaker notions of efficiency like optimality in Pareto's sense have been introduced which characterize whether some point in M is a *reasonable* solution of (MOP), see [Ehrgott, 2005, Jahn, 2004]. In this note, we will focus on weak efficiency. Recall that a point $\bar{x} \in M$ is weakly efficient for (MOP) if there does not exist $\hat{x} \in M$ which satisfies

$$h_j(\hat{x}) < h_j(\bar{x})$$

for all $j = 1, \ldots, k$. In order to find weakly efficient points of (MOP), a common approach is to consider the scalarized optimization problem

$$\begin{aligned} z \cdot h(x) &\to \min \\ x &\in M \end{aligned} \tag{SOP(z)}$$

where $z \in \Delta^k$ is a scalarization vector. This approach is known as the *weighted-sum-scalarization* or *linear scalarization* technique. It is well known that any solution of (SOP(z)) is weakly efficient for (MOP). Supposing that $h_1, \ldots, h_k$ are convex while M is convex, for any $\bar{x} \in \mathbb{R}^n$ which is weakly efficient for (MOP), there is some $z \in \Delta^k$ such that $\bar{x}$ solves (SOP(z)), see [Ehrgott, 2005, Section 3.1].

2 The problem and its surrogate

For any parameter $x \in X_{\text{ad}}$ where $X_{\text{ad}} \subset \mathcal{X}$ denotes the set of admissible parameters, we consider the parametric multicriterial optimization problem

$$\left.\begin{array}{c} f_1(x,y) \\ \vdots \\ f_k(x,y) \end{array}\right\} \to \text{“}\min_y\text{”} \qquad g(x,y) \in K. \tag{P(x)}$$

Here, the quotation marks emphasize that the minimization has to be interpreted in the sense of weak efficiency. The precise standing assumptions on (P(x)) are summarized below.

Assumption 2.1. *For Banach spaces $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$, let the mappings $f\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^k$ and $g\colon \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ be continuous. Here, $k \in \mathbb{N}$ satisfies $k \geq 2$. The components of f are denoted by $f_1, \ldots, f_k\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$. Furthermore, let $K \subset \mathcal{Z}$ be a nonempty, closed, convex cone. For each $x \in X_{\text{ad}}$, $f_1(x,\cdot), \ldots, f_k(x,\cdot)\colon \mathcal{Y} \to \mathbb{R}$ are assumed to be convex. Furthermore, for each $x \in X_{\text{ad}}$, $g(x,\cdot)\colon \mathcal{Y} \to \mathcal{Z}$ needs to satisfy*

$$g(x, \alpha y + (1-\alpha)y') - \alpha g(x,y) - (1-\alpha) g(x,y') \in K$$

for all $y, y' \in \mathcal{Y}$ and $\alpha \in [0,1]$, i.e. the mapping $g(x,\cdot)\colon \mathcal{Y} \to \mathcal{Z}$ is $-K$-convex in the sense of [Jahn, 2004, Definition 2.4].

By $\Psi_{\text{we}}\colon X_{\text{ad}} \rightrightarrows \mathcal{Y}$, we denote the set-valued mapping which assigns to any parameter $x \in X_{\text{ad}}$ the set $\Psi_{\text{we}}(x) \subset \mathcal{Y}$ of weakly efficient points of (P(x)). The superordinate upper level problem associated with (P(x)) is given by

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ x &\in X_{\text{ad}} \\ y &\in \Psi_{\text{we}}(x). \end{aligned} \tag{SVBPP}$$

Note that in contrast to a classical bilevel optimization problem where the upper level objective is only minimized w.r.t. x, we minimize w.r.t. all variables in (SVBPP). This formulation is related to the so-called optimistic approach of bilevel programming, see [Dempe, 2002, Zemkoho, 2012]. Below, we postulate the standing assumptions on our model problem (SVBPP).

Assumption 2.2. *The function $F\colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is continuous and $X_{\mathrm{ad}} \subset \mathcal{X}$ is a nonempty, closed set.*

Next, let us introduce a set-valued mapping $\Phi\colon X_{\mathrm{ad}} \times \Delta^k \rightrightarrows \mathcal{Y}$ by means of

$$\Phi(x, z) := \operatorname*{argmin}_y \{z \cdot f(x, y) \,|\, g(x, y) \in K\}$$

for all $x \in X_{\mathrm{ad}}$ and $z \in \Delta^k$. Recall that the set $\Delta^k \subset \mathbb{R}^k$ represents the standard simplex in $\mathbb{R}^k$. Exploiting the postulated convexity properties of (P(x)), we obtain

$$\Psi_{\mathrm{we}}(x) = \bigcup\nolimits_{z \in \Delta^k} \Phi(x, z)$$

for any $x \in X_{\mathrm{ad}}$. This shows that (SVBPP) is equivalent to

$$\begin{aligned} F(x, y) \; &\to \; \min_{x,y} \\ x \; &\in \; X_{\mathrm{ad}} \\ y \; &\in \; \bigcup\nolimits_{z \in \Delta^k} \Phi(x, z). \end{aligned}$$

Unfortunately, z does not play the role of a variable in the above hierarchical model which makes its theoretical handling very difficult. In order to overcome this shortcoming, one may take a look at the following bilevel programming model:

$$\begin{aligned} F(x, y) \; &\to \; \min_{x,y,z} \\ (x, z) \; &\in \; X_{\mathrm{ad}} \times \Delta^k \\ y \; &\in \; \Phi(x, z). \end{aligned} \tag{BPP}$$

Note that (BPP) is a standard bilevel programming problem since Φ is the solution set mapping of a scalar parametric optimization problem. Particularly, (BPP) can be dealt with (w.r.t. optimality conditions, solution algorithms, etc.) using standard techniques from bilevel programming. This has been done in the finite- and infinite-dimensional setting e.g. in [Bonnel and Morgan, 2012, 2013, Dempe et al., 2013]. For later use, let us define the feasible set mapping $\Gamma\colon X_{\mathrm{ad}} \rightrightarrows \mathcal{Y}$ of the parametric optimization problem (P(x)) by means of

$$\Gamma(x) := \{y \in \mathcal{Y} \,|\, g(x, y) \in K\} \tag{1}$$

for any $x \in X_{\mathrm{ad}}$. Noting that $g(x, \cdot)\colon \mathcal{Y} \to \mathcal{Z}$ is $-K$-convex and continuous, it is easily seen that Γ possesses closed and convex images. Furthermore, we will exploit the function $\varphi\colon X_{\mathrm{ad}} \times \Delta^k \to \overline{\mathbb{R}}$ given by

$$\varphi(x, z) := \inf_y \{z \cdot f(x, y) \,|\, y \in \Gamma(x)\} \tag{2}$$

for any choice of $x \in X_{\mathrm{ad}}$ and $z \in \Delta^k$. Obviously, φ is the so-called optimal value function of the parametric optimization problem

$$\begin{aligned} z \cdot f(x, y) \; &\to \; \min_y \\ g(x, y) \; &\in \; K. \end{aligned} \tag{3}$$

In [Dempe et al., 2013, Proposition 3.1], the authors postulate that the optimization problems (SVBPP) and (BPP) are (in a certain sense) equivalent w.r.t. globally and locally optimal solutions, respectively, provided that the set-valued mapping Φ is closed. In [Bonnel and Morgan, 2012, Proposition 2], it is claimed that this equivalence holds without any additional assumption. By means of a simple example, we show now that these results are not correct in general.

Example 2.3. *Let us consider the semivectorial bilevel programming problem*

$$\begin{aligned} y - x &\to \min_{x,y} \\ x &\in [0,1] \\ y &\in \Psi_{\mathrm{we}}(x) \end{aligned} \tag{4}$$

where $\Psi_{\mathrm{we}}\colon \mathbb{R} \rightrightarrows \mathbb{R}$ *denotes the set-valued mapping which assigns to any* $x \in \mathbb{R}$ *the set of weakly efficient points associated with*

$$\begin{aligned} \left.\begin{matrix} xy \\ 1-y \end{matrix}\right\} &\to \text{“}\min_{y}\text{”} \\ y &\in [0,1]. \end{aligned}$$

One can easily check that $\Psi_{\mathrm{we}}(x) = [0,1]$ *is valid for all* $x \in [0,1]$*. Thus,* $(\bar{x}, \bar{y}) := (1,0)$ *is the unique globally optimal solution of (4) and there do not exist any other locally optimal solutions than* $(\bar{x}, \bar{y})$*.*

Now, we consider the associated bilevel programming problem (BPP). By definition, $\Delta^2 = \operatorname{conv}\{(1,0),(0,1)\}$ *is valid. Thus, for any* $z \in \Delta^2$*, there is a unique* $s \in [0,1]$ *satisfying* $z = (s, 1-s)$*. As a consequence, we may consider the scalarized lower level problem*

$$\begin{aligned} sxy + (1-s)(1-y) &\to \min_{y} \\ y &\in [0,1] \end{aligned}$$

for $s \in [0,1]$*. A simple calculation reveals*

$$\Phi(x, s, 1-s) = \begin{cases} \{0\} & \text{if } s > \frac{1}{x+1}, \\ [0,1] & \text{if } s = \frac{1}{x+1}, \\ \{1\} & \text{if } s < \frac{1}{x+1} \end{cases}$$

for all $x \in [0,1]$ *and* $s \in [0,1]$*. This implies that* Φ *is closed at all points* $(x, s, 1-s)$ *satisfying* $x \in [0,1]$ *and* $s \in [0,1]$*.*

Next, we show that $(\hat{x}, \hat{y}, \hat{z}_1, \hat{z}_2) := (1,1,0,1)$ *is a locally optimal solution of*

$$\begin{aligned} y - x &\to \min_{x,y,z} \\ (x, z_1, z_2) &\in [0,1] \times \Delta^2 \\ y &\in \Phi(x, z_1, z_2). \end{aligned} \tag{5}$$

Therefore, choose $\varepsilon := \frac{1}{4}$ *and fix an arbitrary feasible point* (x, y, z_1, z_2) *of* (5) *from an* ε*-neighborhood (here, w.l.o.g. we choose the supremum norm* $|\cdot|_\infty$*) of* $(\hat{x}, \hat{y}, \hat{z}_1, \hat{z}_2)$*. Supposing that* $y - x < \hat{y} - \hat{x} = 0$ *is valid,* $y \in (\frac{3}{4}, 1)$ *must hold true. This leads to* $z_1 = \frac{1}{1+x}$ *and* $z_2 = 1 - \frac{1}{1+x}$ *for* $x \in (\frac{3}{4}, 1]$*. Thus, we have*

$$|(\hat{z}_1, \hat{z}_2) - (z_1, z_2)|_\infty = \left|\frac{1}{1+x}\right| \geq \frac{1}{2} > \varepsilon$$

which contradicts the choice of (x, y, z_1, z_2)*.*

Hence, $(\hat{x}, \hat{y}, \hat{z}_1, \hat{z}_2)$ *is a locally optimal solution of* (5) *but* $(\hat{x}, \hat{y})$ *is not a locally optimal solution of* (4)*.*

3 On the relationship between semivectorial bilevel programming and scalar bilevel programming

In this section, we clarify the precise relationship between (SVBPP) and the associated bilevel programming problem (BPP). Thereby, we correct the results from [Bonnel and Morgan, 2012, Dempe et al., 2013]. For that purpose, we first introduce a set-valued mapping $\Theta\colon X_{\text{ad}} \times \mathcal{Y} \rightrightarrows \mathbb{R}^k$ given by

$$\Theta(x, y) := \left\{ z \in \Delta^k \,\middle|\, y \in \Phi(x, z) \right\}$$

for any choice of $x \in X_{\text{ad}}$ and $y \in \mathcal{Y}$. Observe that $\operatorname{dom} \Theta = \operatorname{gph} \Psi_{\text{we}}$ is valid. Furthermore, by linearity of the scalarization and continuity of the Euclidean inner product, the image sets of Θ are convex and compact, respectively.

First, we take a look at the relationship between globally optimal solutions of (SVBPP) and (BPP).

Theorem 3.1. 1. *Let* $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ *be a globally optimal solution of* (SVBPP)*. Then, for each* $\bar{z} \in \Theta(\bar{x}, \bar{y})$, $(\bar{x}, \bar{y}, \bar{z})$ *is a globally optimal solution of* (BPP)*.*

2. *Let* $(\bar{x}, \bar{y}, \bar{z}) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^k$ *be a globally optimal solution of* (BPP)*. Then,* $(\bar{x}, \bar{y})$ *is a globally optimal solution of* (SVBPP)*.*

Proof. 1. Since $(\bar{x}, \bar{y})$ is globally optimal solution of (SVBPP), we have $\bar{y} \in \Psi_{\text{we}}(\bar{x})$ by feasibility. Particularly, $\Theta(\bar{x}, \bar{y})$ is nonempty. Suppose on the contrary that there exists a scalarization vector $z \in \Theta(\bar{x}, \bar{y})$ such that $(\bar{x}, \bar{y}, z)$ is not globally optimal for (BPP). Then, there is a feasible point $(\hat{x}, \hat{y}, \hat{z}) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^k$ of (BPP) which satisfies $F(\hat{x}, \hat{y}) < F(\bar{x}, \bar{y})$. From $\hat{z} \in \Theta(\hat{x}, \hat{y})$, we deduce $\hat{y} \in \Psi_{\text{we}}(\hat{x})$, i.e. $(\hat{x}, \hat{y})$ is feasible to (SVBPP). This, however, is a contradiction since $(\bar{x}, \bar{y})$ solves (SVBPP) globally.

2. Suppose that $(\bar{x}, \bar{y})$ is not a globally optimal solution of (SVBPP). Then, there exists a feasible point $(\hat{x}, \hat{y}) \in \mathcal{X} \times \mathcal{Y}$ of (SVBPP) which satisfies $F(\hat{x}, \hat{y}) < F(\bar{x}, \bar{y})$. From $\hat{y} \in \Psi_{\text{we}}(\hat{x})$, we find some $\hat{z} \in \Theta(\hat{x}, \hat{y})$, i.e. $(\hat{x}, \hat{y}, \hat{z})$ is feasible to (BPP). This, however, contradicts the assumption that $(\bar{x}, \bar{y}, \bar{z})$ solves (BPP) globally.

□

Noting that the problems (SVBPP) and (BPP) are inherently nonconvex, the computation of their respective globally optimal solutions is often not possible in numerical practice. Instead, keeping available numerical methods in mind, locally optimal solutions are of essential interest. In the following theorem, we compare (SVBPP) and (BPP) w.r.t. local minimizers. Here, the situation is much more delicate than in Theorem 3.1 where globally optimal solutions were under consideration.

Theorem 3.2. *1. Let $(\bar{x}, \bar{y}) \in \mathcal{X} \times \mathcal{Y}$ be a locally optimal solution of (SVBPP). Then, for each $\bar{z} \in \Theta(\bar{x}, \bar{y})$, $(\bar{x}, \bar{y}, \bar{z})$ is a locally optimal solution of (BPP).*

2. Let $(\bar{x}, \bar{y}, z) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^k$ be a locally optimal solution of (BPP) for all $z \in \Theta(\bar{x}, \bar{y})$ and let Φ be closed at all points from $\{\bar{x}\} \times \Delta^k$. Then, $(\bar{x}, \bar{y})$ is a locally optimal solution of (SVBPP).

Proof. 1. The proof of the theorem's first assertion parallels the proof of the first statement of Theorem 3.1 and, thus, is omitted.

2. Suppose on the contrary that $(\bar{x}, \bar{y})$ is no locally optimal solution of (SVBPP). Then, we find a sequence $\{(x_l, y_l)\}_{l \in \mathbb{N}} \subset \operatorname{gph} \Psi_{\text{we}} \subset X_{\text{ad}} \times \mathcal{Y}$ which converges to $(\bar{x}, \bar{y})$ and satisfies

$$\forall l \in \mathbb{N}: \quad F(x_l, y_l) < F(\bar{x}, \bar{y}). \tag{6}$$

Since $y_l \in \Psi_{\text{we}}(x_l)$ is valid, we find a vector $z_l \in \Theta(x_l, y_l)$ for all $l \in \mathbb{N}$. Noting that Δ^k is compact, $\{z_l\}_{l \in \mathbb{N}} \subset \Delta^k$ converges w.l.o.g. to some $\hat{z} \in \Delta^k$ (otherwise, choose an appropriate subsequence). By definition of Θ and Φ, we obtain $y_l \in \Phi(x_l, z_l)$ for all $l \in \mathbb{N}$. Exploiting the closedness of Φ at $(\bar{x}, \bar{z})$, $\bar{y} \in \Phi(\bar{x}, \hat{z})$ is obtained. Hence, $(\bar{x}, \bar{y}, \hat{z})$ is feasible to (BPP). However, recalling (6), $(\bar{x}, \bar{y}, \hat{z})$ is no local minimizer of (BPP). This contradicts the theorem's assumptions.

□

Putting aside the closedness of Φ which can be ensured by standard assumptions, see Lemma 3.7, the second assertion of the above theorem means that, in order to check whether a given feasible point of (SVBPP) is a locally optimal solution of the latter, one has to verify that *all* the associated feasible points of (BPP) are local minimizers of the surrogate standard bilevel programming problem. By definition of Θ, infinitely many such points may exist in most of the practically relevant situations.

Below, we comment on the assumptions which appear in the second assertion of Theorem 3.2 and are, obviously, violated in the setting of Example 2.3. Our first example visualizes that there may exist situations, where precisely one scalarization parameter $\hat{z} \in \Theta(\bar{x}, \bar{y})$ exists such that $(\bar{x}, \bar{y}, \hat{z})$ is no local minimizer of (BPP).

Example 3.3. *We consider the same semivectorial bilevel programming problem as in Example 2.3 at $(\bar{x}, \bar{y}) := (1, 1)$. One can easily check*

$$\Theta(\bar{x}, \bar{y}) = \operatorname{conv}\left\{(0,1), (\tfrac{1}{2}, \tfrac{1}{2})\right\}.$$

Similar as in Example 2.3 it can be shown that for any $s \in [0, \frac{1}{2})$, $(\bar{x}, \bar{y}, s, 1-s)$ is a locally optimal solution of the associated scalar bilevel programming problem (BPP) *given in (5). However, the point $(\bar{x}, \bar{y}, \frac{1}{2}, \frac{1}{2})$ is no locally optimal solution of (5) since the sequence $\{(1, 1 - \frac{1}{l}, \frac{1}{2}, \frac{1}{2})\}_{l \in \mathbb{N}}$ converges to $(\bar{x}, \bar{y}, \frac{1}{2}, \frac{1}{2})$ but each point of it possesses a better objective value than $(\bar{x}, \bar{y}, \frac{1}{2}, \frac{1}{2})$.*

Particularly, there is only one point in the infinite set $\Theta(\bar{x}, \bar{y})$ where the assumptions of Theorem 3.2 are violated. However, $(\bar{x}, \bar{y})$ is no local minimizer of the semivectorial bilevel programming problem (4).

Supposing that Θ is at most singleton-valued, the technical assumptions associated with the second assertion of Theorem 3.2 reduce to the closedness of Φ which, as we will see in Lemma 3.7, follows from standard constraint qualifications. However, even in the case where $f_1(x, \cdot), \ldots, f_k(x, \cdot) \colon \mathcal{Y} \to \mathbb{R}$ are strictly convex functions for all $x \in X_{\mathrm{ad}}$, which implies the set $\Phi(x, z)$ to be either empty or a singleton for any $x \in X_{\mathrm{ad}}$ and $z \in \Delta^k$, Θ does not need to possess this property.

Example 3.4. *We investigate the semivectorial bilevel programming problem*

$$\begin{aligned} -x - y &\to \min_{x,y} \\ x &\in [-1, 1] \\ y &\in \Psi_{\mathrm{we}}(x) \end{aligned} \tag{7}$$

where $\Psi_{\mathrm{we}} \colon [-1, 1] \rightrightarrows \mathbb{R}$ is the multifunction which assigns to any $x \in \mathbb{R}$ the set of all weakly efficient points of the parametric bicriterial optimization problem

$$\begin{aligned} \left.\begin{matrix} (y+1)^2 \\ (y-1)^2 \end{matrix}\right\} &\to \text{“}\min_y\text{”} \\ y &\in [-1, 1] \\ xy &\geq 0. \end{aligned}$$

One easily calculates

$$\Psi_{\mathrm{we}}(x) = \begin{cases} [-1, 0] & \text{if } x \in [-1, 0), \\ [-1, 1] & \text{if } x = 0, \\ [0, 1] & \text{if } x \in (0, 1] \end{cases}$$

for all $x \in [-1, 1]$. Thus, $(\bar{x}, \bar{y}) := (1, 1)$ solves (7) while there do not exist any locally optimal solutions different from $(\bar{x}, \bar{y})$.

On the other hand, we obtain

$$\Phi(x, s, 1-s) = \begin{cases} \{\min\{1 - 2s; 0\}\} & \text{if } x \in [-1, 0), \\ \{1 - 2s\} & \text{if } x = 0, \\ \{\max\{1 - 2s; 0\}\} & \text{if } x \in (0, 1] \end{cases}$$

for any $x \in [-1, 1]$ and $s \in [0, 1]$. Obviously, Φ is closed at all points from $\{1\} \times \Delta^2$. Note that Φ is singleton-valued, but the same does not hold true for Θ: one easily obtains e.g.

$$\Theta(1, 0) = \operatorname{conv}\left\{(\tfrac{1}{2}, \tfrac{1}{2}), (1, 0)\right\}.$$

Furthermore, $(1, 0, s, 1-s)$ is a locally optimal solution of the associated scalar bilevel programming problem for any $s \in (\frac{1}{2}, 1]$ while $(1, 0)$ is no local minimizer of (7). This can be easily checked using similar arguments as in Examples 2.3 and 3.3.

Finally, we would like to comment on the postulated closedness assumption appearing in Theorem 3.2. In this regard, the following lemma might be of interest.

Lemma 3.5. *Fix $\bar{x} \in X_{\text{ad}}$ and assume that Φ is closed at all points from $\{\bar{x}\} \times \Delta^k$. Then, Ψ_{we} is closed at $\bar{x}$.*

Proof. Let $\{x_l\}_{l\in\mathbb{N}} \subset X_{\text{ad}}$ and $\{y_l\}_{l\in\mathbb{N}} \subset \mathcal{Y}$ be sequences such that $x_l \to \bar{x}$ holds true while we have $y_l \in \Psi_{\text{we}}(x_l)$ for all $l \in \mathbb{N}$. Furthermore, assume that there is $\bar{y} \in \mathcal{Y}$ such that $y_l \to \bar{y}$ holds true. Clearly, there is a sequence $\{z_l\}_{l\in\mathbb{N}} \subset \Delta^k$ such that $y_l \in \Phi(x_l, z_l)$ is satisfied for all $l \in \mathbb{N}$. Observing that Δ^k is compact, we may assume w.l.o.g. that $z_l \to \bar{z}$ holds true for some $\bar{z} \in \Delta^k$. Recalling that Φ is closed at $(\bar{x}, \bar{z})$, $\bar{y} \in \Phi(\bar{x}, \bar{z})$ follows. This yields $\bar{y} \in \Psi_{\text{we}}(\bar{x})$, i.e. Ψ_{we} is closed at $\bar{x}$. $\square$

One can easily check that the converse statement of Lemma 3.5 is not generally true.

Example 3.6. *For the set $X_{\text{ad}} := [0, 1]$, we consider the parametric bicriterial optimization problem*

$$\begin{aligned} \left.\begin{matrix} y \\ -y \end{matrix}\right\} &\to \text{“}\min_y\text{”} \\ y &\in [0,1] \\ xy &\leq 0. \end{aligned} \tag{8}$$

Clearly, we have

$$\Psi_{\text{we}}(x) = \begin{cases} [0,1] & \textit{if } x = 0, \\ \{0\} & \textit{if } x \in (0,1] \end{cases}$$

for all $x \in [0, 1]$ and, thus, Ψ_{we} is closed everywhere on its domain. On the other hand, a simple calculation shows

$$\Phi(x, s, 1-s) = \begin{cases} \{1\} & \textit{if } x = 0,\ s \in [0, \frac{1}{2}), \\ [0,1] & \textit{if } x = 0,\ s = \frac{1}{2}, \\ \{0\} & \textit{if } x = 0,\ s \in (\frac{1}{2}, 1], \\ \{0\} & \textit{if } x \in (0,1],\ s \in [0,1] \end{cases}$$

and this mapping is not closed at the point $(\bar{x}, \bar{z}_1, \bar{z}_2) := (0, 0, 1)$. In order to see this, consider the sequence $\{(\frac{1}{l}, 0, 1, 0)\}_{l\in\mathbb{N}} \subset \operatorname{gph}\Phi$ and observe that it converges to $(\bar{x}, \bar{z}_1, \bar{z}_2, 0) \notin \operatorname{gph}\Phi$.

Recall that $\operatorname{gph}\Psi_{\text{we}} \subset X_{\text{ad}} \times \mathcal{Y}$ holds by definition. Noting that the semivectorial bilevel programming problem (SVBPP) is equivalent to

$$\begin{aligned} F(x,y) &\to \min_{x,y} \\ (x,y) &\in \operatorname{gph}\Psi_{\text{we}}, \end{aligned}$$

the closedness of $\operatorname{gph}\Psi_{\mathrm{we}}$ (amongst overs) needs to be guaranteed in order to derive existence results for (SVBPP). Exploiting Lemma 3.5, the closedness of $\operatorname{gph}\Psi_{\mathrm{we}}$ follows from the closedness of $\operatorname{gph}\Phi$. This justifies the closedness assumption in Theorem 3.2. Observing that Φ is the solution set mapping of a scalar parametric optimization problem, one could use standard results, see e.g. [Bank et al., 1983], in order to infer its closedness. In the lemma below, a sufficient condition for the closedness of Φ is presented. In this regard, recall that Γ defined in (1) is the lower level feasible set mapping while the function φ is given in (2).

Lemma 3.7. *Let g be continuously Fréchet differentiable and fix $\bar{x} \in X_{\mathrm{ad}}$. If for each $y \in \Gamma(\bar{x})$, the constraint qualification*

$$g'_y(\bar{x}, y)[\mathcal{Y}] - \operatorname{cone}(K - \{g(\bar{x}, y)\}) = \mathcal{Z} \tag{9}$$

is valid, then the following assertions hold:

(i) Γ is lower semicontinuous at $\bar{x}$,

(ii) φ is upper semicontinuous at all points from $\{\bar{x}\} \times \Delta^k$, and

(iii) Φ is closed at all points from $\{\bar{x}\} \times \Delta^k$.

Proof. (i) Suppose on the contrary that Γ is not lower semicontinuous at $\bar{x}$. Then, there exist $\delta > 0$, a sequence $\{x_l\}_{l \in \mathbb{N}} \subset X_{\mathrm{ad}}$ converging to $\bar{x}$, and some $\bar{y} \in \Gamma(\bar{x})$ such that $\Gamma(x_l) \cap \mathbb{U}^{\delta}_{\mathcal{Y}}(\bar{y})$ is empty for all $l \in \mathbb{N}$. This shows

$$\inf_{y_l \in \Gamma(x_l)} \|\bar{y} - y_l\|_{\mathcal{Y}} \geq \delta \tag{10}$$

for all $l \in \mathbb{N}$. Due to the validity of the postulated constraint qualification, we can exploit [Robinson, 1976, Theorem 1] in order to find constants $\varepsilon > 0$, $\gamma > 0$, and $c > 0$ such that

$$\inf_{y' \in \Gamma(x)} \|y - y'\|_{\mathcal{Y}} \leq c \inf_{z \in K} \|g(x, y) - z\|_{\mathcal{Z}}$$

holds for all $x \in \mathbb{U}^{\varepsilon}_{\mathcal{X}}(\bar{x})$ and $y \in \mathbb{U}^{\gamma}_{\mathcal{Y}}(\bar{y})$. Particularly, $\Gamma(x)$ is nonempty for all $x \in X_{\mathrm{ad}} \cap \mathbb{U}^{\varepsilon}_{\mathcal{X}}(\bar{x})$. The above result implies that

$$\inf_{y_l \in \Gamma(x_l)} \|\bar{y} - y_l\|_{\mathcal{Y}} \leq c \inf_{z \in K} \|g(x_l, \bar{y}) - z\|_{\mathcal{Z}} \leq c \|g(x_l, \bar{y}) - g(\bar{x}, \bar{y})\|_{\mathcal{Z}}$$

is valid for sufficiently large $l \in \mathbb{N}$, where the last inequality follows from $\bar{y} \in \Gamma(\bar{x})$. Observing that g is continuous, this, however, contradicts (10). Thus, Γ is lower semicontinuous at $\bar{x}$.

(ii) Noting that the mapping $\mathcal{X} \times \mathcal{Y} \times \Delta^k \ni (x, y, z) \mapsto z \cdot f(x, y) \in \mathbb{R}$ is continuous, we can invoke [Bank et al., 1983, Theorem 4.2.2] in order to see that φ is upper semicontinuous at all points from $\{\bar{x}\} \times \Delta^k$ since Γ is lower semicontinuous at $\bar{x}$ due to the lemma's first assertion.

(iii) Since the continuity of g implies the closedness of Γ at $\bar{x}$, it is possible to apply [Bank et al., 1983, Theorem 4.2.1] in order to deduce the lemma's final assertion from the first two.

□

Below, we comment on the constraint qualification (9) which has been exploited in the above lemma.

Remark 3.8. *In the context of programming in Banach spaces, condition (9) is referred to as* Robinson's constraint qualification. *It has been utilized by Robinson in [Robinson, 1976] in order to characterize the stability properties of parameterized constraint systems. Furthermore, it has been used by Kurcyusz and Zowe to guarantee the existence of Lagrange multipliers at locally optimal solutions of optimization problems in Banach spaces, see [Zowe and Kurcyusz, 1979]. More information on Robinson's constraint qualification including equivalent formulations and its applications can be found in the monograph [Bonnans and Shapiro, 2000]. Particularly, (9) equals the* Mangasarian-Fromovitz constraint qualification *in the setting of standard nonlinear bilevel programming, i.e. when* $\mathcal{X} := \mathbb{R}^n$, $\mathcal{Y} := \mathbb{R}^m$, $\mathcal{Z} := \mathbb{R}^p$, *and* $K := \{z \in \mathbb{R}^p \,|\, z \leq 0\}$ *hold.*

4 Existence of solutions in semivectorial bilevel programming

In this section, we investigate the existence of optimal solutions to (SVBPP). Keeping Lemma 3.5, the associated remarks, and Weierstraß' classical theorem in mind, this task is not challenging whenever the spaces $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional. The infinite-dimensional situation, unfortunately, is far more difficult to handle due to two observations: First, in order to guarantee the existence of solutions to optimization problems in Banach spaces, one needs to ensure that their feasible sets are weakly sequentially closed which is often guaranteed via convexity assumptions. Recalling e.g. Examples 2.3, 3.4 and 3.6, it is, however, not realistic to assume the convexity of $\operatorname{gph}\Psi_{\text{we}}$. In bilevel optimal control, it is a standard trick to exploit the uniqueness of the lower level solution for fixed parameter and certain continuity properties of the associated solution operator in order to infer the existence of solutions, see [Dempe et al., 2018, Harder and Wachsmuth, 2018]. The mapping Ψ_{we}, however, assigns to any $x \in X_{\text{ad}}$ the set of all weakly efficient points of the multicriterial optimization problem (P(x)) which is not a singleton in all practically relevant situations. This is the second difficulty we have to face.

Recall that due to Theorem 3.1, one could infer the existence of optimal solutions to (SVBPP) from the existence of optimal solutions to the standard bilevel program (BPP). Thus, noting that the solution mapping Φ of the scalarized lower level problem might be single-valued under appropriate assumptions although Ψ_{we} is not, Theorem 3.1 opens a way to the investigation of the existence of optimal solutions to (SVBPP). This observation has been used in [Bonnel and Morgan, 2012] in order to study special semivectorial bilevel optimal control problems of ordinary differential equations. However, in order to carry out a similar analysis to address (SVBPP), one needs to assume that *all* the lower level objective functionals $f_1(x,\cdot),\ldots,f_k(x,\cdot)\colon \mathcal{Y}\to\mathbb{R}$ are strictly convex for each $x\in X_{\text{ad}}$. Such an assumption might by too restrictive in many real-world applications.

Here, we are going to provide conditions which guarantee that $\operatorname{gph}\Phi\subset X_{\text{ad}}\times\Delta^k\times\mathcal{Y}$ is weakly sequentially compact. Such an approach has been used in [Holler et al., 2018] to derive an existence result for a parameter learning model which is a bilevel optimal control problems with not necessarily unique lower level solution. The weak sequential compactness of $\operatorname{gph}\Phi$ allows us to infer the existence of an optimal solution to (BPP) and, keeping Theorem 3.1 in mind, to (SVBPP).

In this section, we first derive a theoretical existence result for (SVBPP) where the lower level decision space $\mathcal{Y}$ is not necessarily finite-dimensional. Afterwards, we apply our findings to two prominent classes of semivectorial bilevel optimal control problems, namely so-called *simple semivectorial bilevel optimal control problems* and *inverse multicriterial optimal control problems*.

4.1 Theoretical investigations

Besides Assumptions 2.1 and 2.2, let us postulate the following:

Assumption 4.1. *Let $\mathcal{X}$ be finite-dimensional and $\mathcal{Y}$ be reflexive. We assume that F as well as $f_1,\ldots,f_k$ are weakly sequentially lower semicontinuous. Let X_{ad} be compact. For any $x\in X_{\text{ad}}$, the set $\Gamma(x)$ is supposed to be nonempty. Finally, let Γ be lower semicontinuous on X_{ad} and let $\operatorname{gph}\Gamma\subset X_{\text{ad}}\times\mathcal{Y}$ be weakly sequentially compact.*

Clearly, the demanded weak sequential lower semicontinuity of F as well as $f_1,\ldots,f_k$ is inherent whenever these functionals are fully convex (since they are postulated to be continuous). The lower semicontinuity of Γ on X_{ad} can be guaranteed if Robinson's constraint qualification is satisfied at all points from $\operatorname{gph}\Gamma$, see Lemma 3.7 and Remark 3.8, or if Γ is constant. If the mapping g is fully $-K$-convex, then it can be shown that $\operatorname{gph}\Gamma$ is convex and, due to the continuity of g, closed as well. This would ensure the weak sequential closedness of $\operatorname{gph}\Gamma$.

In the lemmas below, we list some consequences of the postulated assumptions. For brevity, we introduce $\ell\colon\mathcal{X}\times\mathcal{Y}\times\Delta^k\to\mathbb{R}$ by means of

$$\ell(x,y,z):=z\cdot f(x,y)$$

for any $x\in\mathcal{X}$, $y\in\mathcal{Y}$, and $z\in\Delta^k$.

Lemma 4.2. *The functional ℓ is weakly sequentially lower semicontinuous.*

Proof. Let us choose sequences $\{x_l\}_{l\in\mathbb{N}}\subset\mathcal{X}$, $\{y_l\}_{l\in\mathbb{N}}\subset\mathcal{Y}$, and $\{z_l\}_{l\in\mathbb{N}}\subset\Delta^k$ such that $x_l\to\bar{x}$, $y_l\rightharpoonup\bar{y}$, as well as $z_l\to\bar{z}$ hold true for points $\bar{x}\in\mathcal{X}$, $\bar{y}\in\mathcal{Y}$, and $\bar{z}\in\Delta^k$ (note that $\mathcal{X}$ and $\mathbb{R}^k$ are finite-dimensional). We define a vector $\alpha\in\overline{\mathbb{R}}^k$ by means of

$$\alpha_j:=\liminf_{l\to\infty} f_j(x_l,y_l)$$

for all $j = 1, \ldots, k$. Noting that f possesses only finitely many components, there exists a subsequence $\{y_{l_s}\}_{s\in\mathbb{N}}$ of $\{y_l\}_{l\in\mathbb{N}}$ which satisfies

$$\alpha = \lim_{s\to\infty} f(x_{l_s}, y_{l_s}).$$

Recalling that the mappings $f_1, \ldots, f_k$ are weakly sequentially lower semicontinuous while the components of $\bar{z}$ are nonnegative, we obtain

$$\begin{aligned} \liminf_{l\to\infty} \ell(x_l, y_l, z_l) &= \lim_{s\to\infty} \ell(x_{l_s}, y_{l_s}, z_{l_s}) \\ &= \lim_{s\to\infty} z_{l_s} \cdot f(x_{l_s}, y_{l_s}) \\ &= \bar{z} \cdot \alpha \geq \bar{z} \cdot f(\bar{x}, \bar{y}) = \ell(\bar{x}, \bar{y}, \bar{z}) \end{aligned}$$

which shows the claim. □

Lemma 4.3. *The mapping Φ possesses nonempty images on its domain $X_{\text{ad}} \times \Delta^k$.*

Proof. First, we note that for fixed parameters $x \in X_{\text{ad}}$ and $z \in \Delta^k$, the functional $\mathcal{Y} \ni y \mapsto \ell(x, y, z) \in \mathbb{R}$ is weakly sequentially lower semicontinuous by means of Lemma 4.2. Furthermore, for fixed $x \in X_{\text{ad}}$, the lower level feasible set $\Gamma(x)$ is nonempty, convex, as well as bounded. The latter follows since $\operatorname{gph}\Gamma$ is assumed to be weakly sequentially compact and, thus, bounded. Now, the reflexivity of $\mathcal{Y}$ guarantees that $\Gamma(x)$ is weakly sequentially compact. As a consequence, standard arguments show that $\Phi(x, z)$ is nonempty. □

Lemma 4.4. *The mapping Φ possesses the following property: Let $\{x_l\}_{l\in\mathbb{N}} \subset X_{\text{ad}}$, $\{y_l\}_{l\in\mathbb{N}} \subset \mathcal{Y}$, as well as $\{z_l\}_{l\in\mathbb{N}} \subset \Delta^k$ be sequences which satisfy $x_l \to \bar{x}$ as well as $z_l \to \bar{z}$ for points $\bar{x} \in X_{\text{ad}}$ as well as $\bar{z} \in \Delta^k$, and $y_l \in \Phi(x_l, z_l)$ for all $l \in \mathbb{N}$. Then, there exists $\bar{y} \in \Phi(\bar{x}, \bar{z})$ such that $y_l \rightharpoonup \bar{y}$ holds at least along a subsequence.*
Particularly, the set $\operatorname{gph}\Phi \subset X_{\text{ad}} \times \Delta^k \times \mathcal{Y}$ is weakly sequentially compact.

Proof. Noting that Γ is lower semicontinuous on X_{ad} while ℓ is continuous, the function φ, which is the optimal value function of the parametric optimization problem (3), is upper semicontinuous by means of [Bank et al., 1983, Theorem 4.2.2], see Lemma 3.7 as well.

Choose sequences $\{x_l\}_{l\in\mathbb{N}} \subset X_{\text{ad}}$, $\{y_l\}_{l\in\mathbb{N}} \subset \mathcal{Y}$, and $\{z_l\}_{l\in\mathbb{N}} \subset \Delta^k$ satisfying $x_l \to \bar{x}$ and $z_l \to \bar{z}$ for some $\bar{x} \in X_{\text{ad}}$ and $\bar{z} \in \Delta^k$ as well as $y_l \in \Phi(x_l, z_l)$ for all $l \in \mathbb{N}$. Recalling that $\operatorname{gph}\Gamma$ is assumed to be bounded, $\bigcup_{l\in\mathbb{N}} \Gamma(x_l)$ is bounded as well. Thus, $\{y_l\}_{l\in\mathbb{N}}$ is bounded and, since $\mathcal{Y}$ is reflexive, possesses a weakly convergent subsequence (without relabeling) whose weak limit point will be denoted by $\bar{y} \in \mathcal{Y}$. It remains to show $\bar{y} \in \Phi(\bar{x}, \bar{z})$.

From $(x_l, y_l) \in \operatorname{gph}\Gamma$ for all $l \in \mathbb{N}$ and the postulated weak sequential compactness of the latter set, we particularly infer $(\bar{x}, \bar{y}) \in \operatorname{gph}\Gamma$, i.e. $\bar{y} \in \Gamma(\bar{x})$ holds. By definition of φ, the weak lower semicontinuity of ℓ (see Lemma 4.2), and upper semicontinuity of φ, we obtain

$$\varphi(\bar{x}, \bar{z}) \leq \ell(\bar{x}, \bar{y}, \bar{z}) \leq \liminf_{l\to\infty} \ell(x_l, y_l, z_l) = \liminf_{l\to\infty} \varphi(x_l, z_l) \leq \limsup_{l\to\infty} \varphi(x_l, z_l) \leq \varphi(\bar{x}, \bar{z}).$$

This yields $\varphi(\bar{x},\bar{z}) = \ell(\bar{x},\bar{y},\bar{z})$, i.e. $\bar{y} \in \Phi(\bar{x},\bar{z})$. Particularly, this shows that $\operatorname{gph}\Phi$ is weakly sequentially closed since X_{ad} and Δ^k are closed, respectively.

Clearly, we have

$$\operatorname{gph}\Phi \subset X_{\mathrm{ad}} \times \Delta^k \times \left(\bigcup_{x \in X_{\mathrm{ad}}} \Gamma(x) \right),$$

and the latter is bounded since $\operatorname{gph}\Gamma$ is bounded. Thus, $\operatorname{gph}\Phi$ is a weakly sequentially closed and bounded set. Consequently, it is weakly sequentially compact. This completes the proof. $\square$

Combining the above lemmas, the following theorem is obtained.

Theorem 4.5. *Under the postulated assumptions,* (SVBPP) *possesses an optimal solution.*

Proof. We note that (BPP) is equivalent to

$$\begin{aligned} F(x,y) &\to \min_{x,y,z} \\ (x,z,y) &\in \operatorname{gph}\Phi. \end{aligned}$$

The objective functional of this program is assumed to weakly sequentially lower semicontinuous. Furthermore, its feasible set is nonempty by Lemma 4.3 and weakly sequentially compact by Lemma 4.4. Particularly, (BPP) possesses an optimal solution. Applying Theorem 3.1, (SVBPP) possesses an optimal solution as well. This completes the proof. $\square$

4.2 Examples from bilevel optimal control

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and consider the associated space $L^2(\Omega)$ of all (equivalence classes) of (Lebesgue) measurable functions whose square is (Lebesgue) integrable. Furthermore, let $\mathtt{S}\colon L^2(\Omega) \to \mathcal{D}$ be a so-called *control-to-observation-operator* which assigns to any control in $L^2(\Omega)$ the (uniquely determined) solution of a given ordinary or partial differential equations and transfers it to the observation space $\mathcal{D}$ which is assumed to be a Hilbert space. In optimal control, one often aims to find a control function within a set of feasible controls such that e.g.

(i) the resulting state approximates a given target $y_{\mathrm{d}} \in \mathcal{D}$ as good as possible,

(ii) the overall control effort is minimal, or

(iii) the chosen control is *sparse*, i.e. it vanishes in most parts of Ω,

see [Hinze et al., 2009, Tröltzsch, 2009] for an introduction to optimal control of partial differential equations and Stadler [2007] where (iii) is promoted via an L^1-cost term w.r.t. the control. It is not difficult to imagine that the above goals are conflicting. However, in standard optimal control, these candidates for possible objectives are ranked by some (fixed) weights and their sum is minimized. In order to find a good *compromise*, it seems to be a nearby idea to consider a multicriterial optimal control problem instead. A satisfying overview of existing literature regarding multiobjective optimal control is given in Peitz and Dellnitz [2018].

In order to stay close to standard notation, the control function (which will be the lower level variable in our context) will be denoted by u. Furthermore, it is assumed that $\mathtt{S}$ is linear and continuous.

Simple semivectorial bilevel optimal control Let $U_{\mathrm{ad}} \subset L^2(\Omega)$ denote the set of feasible controls and assume that U_{ad} is nonempty, bounded, and convex. A typical choice is given by

$$U_{\mathrm{ad}} := \left\{ u \in L^2(\Omega) \,\middle|\, u_a \le u \le u_b \text{ a.e. on } \Omega \right\} \tag{11}$$

where $u_a, u_b \in L^2(\Omega)$ satisfying $u_a < u_b$ almost everywhere on Ω are fixed. Now, consider the multicriterial optimal control problem

$$\begin{aligned} \left.\begin{array}{r} \frac{1}{2}\left\|\mathtt{S}u - y_{\mathrm{d}}\right\|_{\mathcal{D}}^2 \\ \frac{1}{2}\left\|u\right\|_{L^2(\Omega)}^2 \\ \left\|u\right\|_{L^1(\Omega)} \end{array}\right\} &\rightarrow \text{“}\min\text{”} \\ u &\in U_{\mathrm{ad}}. \end{aligned} \tag{12}$$

Note that the objective functionals in (12) correspond to the goals (i), (ii), and (iii) listed above. The functional $L^2(\Omega) \ni u \mapsto \|u\|_{L^1(\Omega)} \in \mathbb{R}$ is continuous due to the postulated boundedness of Ω. Let $\Psi_{\mathrm{we}} \subset L^2(\Omega)$ denote the set of weakly efficient points associated with (12). Clearly, Ψ_{we} is not a singleton. Thus, in order to identify *reasonable* weakly efficient points of (12), one might consider the superordinate optimization problem

$$\begin{aligned} \mathcal{J}(u) &\rightarrow \min \\ u &\in \Psi_{\mathrm{we}} \end{aligned} \tag{SSVBOC}$$

where $\mathcal{J}\colon L^2(\Omega) \to \mathbb{R}$ is convex and continuous. A related problem is studied in [Bonnel and Kaya, 2010].

Noting that the data in (12) is convex while (SSVBOC) possesses a convex objective functional, (SSVBOC) seems to be related to the so-called *simple convex bilevel programming problem* (SCBPP for short), see e.g. [Dempe et al., 2010, Franke et al., 2018]. That is why we call (SSVBOC) the *simple semivectorial bilevel optimal control problem*. However, it has to be noted that in contrast to SCBPP, were a convex functional is minimized over the solution set of a scalar convex optimization program, (SSVBOC) is not a convex program since Ψ_{we} is not convex in general. Thus, one cannot apply standard results in order to infer the existence of an optimal solution.

It is possible to interpret (SSVBOC) as a particular instance of (SVBPP) where the upper level variable x simply vanishes (the set of feasible parameters X_{ad} may, therefore, be imagined as a singleton). The resulting scalarized bilevel optimal control problem is related to the model problems studied in [Harder and Wachsmuth, 2018, Holler et al., 2018]. One can easily check that the Assumptions 2.1, 2.2 and 4.1 hold for (SSVBOC). Thus, we can exploit Theorem 4.5 in order to obtain the following result.

Theorem 4.6. *Under the postulated assumptions,* (SSVBOC) *possesses an optimal solution.*

Inverse multiobjective optimal control Let us assume that $u_o \in L^2(\Omega)$ is a given weakly efficient point of the multiobjective optimal control problem (12) where the precise problem data, in particular the desired state y_d and the set of feasible controls U_{ad}, is unknown and shall be reconstructed. Thus, we have to consider (12) where y_d and U_{ad} depend from a parameter $p \in P$ where $P \subset \mathbb{R}^n$ is a nonempty and compact set. More precisely, for any fixed value of the parameter $p \in P$, the multicriterial optimal control problem

$$\left.\begin{array}{r} \frac{1}{2}\|\mathtt{S}u - \mathtt{T}p\|_{\mathcal{D}}^2 \\ \frac{1}{2}\|u\|_{L^2(\Omega)}^2 \\ \|u\|_{L^1(\Omega)} \end{array}\right\} \to \text{“}\min_u\text{”} \qquad u \in U_{ad}(p) \tag{MOC(p)}$$

is considered where $\mathtt{T}\colon \mathbb{R}^n \to \mathcal{D}$ is an affine and continuous operator while the set-valued mapping $U_{ad}\colon P \rightrightarrows L^2(\Omega)$ possesses nonempty, bounded, and convex images. A typical choice for the operator $\mathtt{T}$ would be given by

$$\mathtt{T}p := \sum_{i=1}^n p_i v^i$$

for all $p \in \mathbb{R}^n$ where $v^1, \ldots, v^n \in \mathcal{D}$ are given form functions. Let $\Psi_{we}\colon P \rightrightarrows L^2(\Omega)$ be the set-valued mapping which assigns to any $p \in P$ the set of weakly efficient points associated with (MOC(p)). We consider the superordinate *inverse multiobjective optimal control problem*

$$\begin{aligned} \frac{1}{2}\|u - u_o\|_{L^2(\Omega)}^2 &\to \min_{p,u} \\ p &\in P \\ u &\in \Psi_{we}(p). \end{aligned} \tag{IMOC}$$

Thus, we want to identify those parameters $p \in P$ for which u_o is (close to) a weakly efficient point of (MOC(p)). Once more, we apply Theorem 4.5 in order to infer the following existence result.

Theorem 4.7. *In addition to the above assumptions, assume that U_{ad} is lower semicontinuous. Furthermore, let $\operatorname{gph} U_{ad} \subset P \times L^2(\Omega)$ be weakly sequentially compact. Then,* (IMOC) *possesses an optimal solution.*

Clearly, all these assumptions on the set-valued mapping U_{ad} hold whenever it is constant. Below, a nontrivial example is presented where all the assumptions from Theorem 4.7 hold. It represents the situation where the upper bound in the definition of the set of feasible controls (11) is unknown. The example can be extended to the related cases where only the lower bound or both bounds are unknown doing some obvious changes.

Example 4.8. *Let $\{\Omega_i\}_{i=1}^n$ be a disjoint partition of Ω, let $u_a, u_b \in L^2(\Omega)$ be functions satisfying $u_a < u_b$ almost everywhere on Ω, let P be convex, and assume that we have $P \subset \{p \in \mathbb{R}^n \,|\, p \geq 0\}$. For each $p \in P$, assume that $U_{\mathrm{ad}}(p)$ is given by*

$$U_{\mathrm{ad}}(p) := \left\{ u \in L^2(\Omega) \,\middle|\, u_a \leq u \leq u_b + \sum_{i=1}^n p_i \chi_{\Omega_i} \text{ a.e. on } \Omega \right\}.$$

Here, for a measurable set $A \subset \Omega$, $\chi_A \colon \Omega \to \mathbb{R}$ denotes the characteristic function of A which equals 1 on A and vanishes otherwise.

Due to the convexity of P, one can easily check that the graph of $U_{\mathrm{ad}} \colon P \rightrightarrows L^2(\Omega)$ is convex. Moreover the compactness of P ensures that $\operatorname{gph} U_{\mathrm{ad}}$ is bounded and closed. Particularly, $\operatorname{gph} U_{\mathrm{ad}}$ is weakly sequentially compact.

It remains to show that the set-valued mapping U_{ad} is lower semicontinuous. Therefore, choose a pair $(\bar p, \bar u) \in \operatorname{gph} U_{\mathrm{ad}}$ as well as $\varepsilon > 0$ arbitrarily and fix $\hat p \in P \cap \mathbb{U}^{\delta}_{\mathbb{R}^n}(\bar p)$ for some $\delta > 0$ which will be specified below. Now, we define a function $\hat u \in U_{\mathrm{ad}}(\hat p)$ as follows: Fix $i \in \{1, \ldots, n\}$. If $\hat p_i \geq \bar p_i$ holds, set $\hat u(\omega) := \bar u(\omega)$ for all $\omega \in \Omega_i$. Supposing that $\hat p_i < \bar p_i$ holds true and considering $\omega \in \Omega_i$, we distinguish two cases. If $\bar u(\omega) \leq u_b(\omega) + \hat p_i$ is valid, we set $\hat u(\omega) := \bar u(\omega)$. Otherwise, $\hat u(\omega) := u_b(\omega) + \hat p_i$ is fixed. By definition, we obtain

$$\begin{aligned}
\|\hat u - \bar u\|^2_{L^2(\Omega)} &= \sum_{i=1}^n \int_{\Omega_i} (\hat u(\omega) - \bar u(\omega))^2 \mathrm{d}\omega \\
&\leq \sum_{i \colon \hat p_i < \bar p_i} \int_{\Omega_i} (u_b(\omega) + \hat p_i - \bar u(\omega))^2 \mathrm{d}\omega \\
&\leq \sum_{i \colon \hat p_i < \bar p_i} |\Omega_i| (\hat p_i - \bar p_i)^2 \\
&\leq C \, |\hat p - \bar p|_2^2 < C\delta^2
\end{aligned}$$

for the constant $C := \max\{|\Omega_1|, \ldots, |\Omega_n|\}$. Here, $|A|$ denotes the (Lebesgue) measure of a (Lebesgue) measurable set $A \subset \Omega$. Thus, choosing $\delta := \varepsilon/\sqrt{C}$, $\hat u \in U_{\mathrm{ad}}(\hat p) \cap \mathbb{U}^{\varepsilon}_{L^2(\Omega)}(\bar u)$ follows. This shows the lower semicontinuity of U_{ad}.

5 Concluding remarks

In this note, we investigated the relationship between a semivectorial bilevel programming (SVBPP) and an associated scalar bilevel programming problem (BPP) which is constructed by applying the *weighted-sum-scalarization* technique to the multiobjective lower level problem of (SVBPP). It has been shown that this relation is nonhazardous when globally optimal solutions are under consideration, while the investigation of locally optimal solutions is somehow delicate. By means of examples, we illustrated the necessity of the assumptions which are needed in order to guarantee that locally optimal solutions of (BPP) correspond to locally optimal solutions of (SVBPP). Thereby, we pointed out and clarified some inconsistencies in the literature. As in [Dempe and Dutta, 2010], where the Karush-Kuhn-Tucker reformulation of standard bilevel programming problems is considered, our results depict that surrogates of bilevel programming problems which are constructed via additional variables have to be investigated with extreme care when local minimizers are under consideration.

We exploited the global equivalence of (SVBPP) and (BPP) in order to derive an existence result for (SVBPP) which is applicable even in the setting where the lower level decision maker's variable comes from an infinite-dimensional Banach space. The associated theory has been used in order to obtain existence results for two classes of semivectorial bilevel optimal control problems, namely *simple semivectorial bilevel optimal control problems* and *inverse multicriterial optimal control problems*.

Grants and acknowledgments

This work is partially supported by the DFG grant *Analysis and Solution Methods for Bilevel Optimal Control Problems* (grant number DE 650/10-1) within the Priority Program SPP 1962 (Non-smooth and Complementarity-based Distributed Parameter Systems: Simulation and Hierarchical Optimization).

References

M. J. Alves and C. H. Antunes. A semivectorial bilevel programming approach to optimize electricity dynamic time-of-use retail pricing. *Computers & Operations Research*, 92:130 – 144, 2018. doi: 10.1016/j.cor.2017.12.014.

B. Bank, J. Guddat, D. Klatte, B. Kummer, and K. Tammer. *Non-Linear Parametric Optimization*. Birkhäuser, Basel, 1983.

J. F. Bard. *Practical Bilevel Optimization: Algorithms and Applications*. Kluwer Academic, Dordrecht, 1998.

J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.

H. Bonnel. Optimality conditions for the semivectorial bilevel optimization problem. *Pacific Journal of Optimization*, 2(3):447–468, 2006.

H. Bonnel and C. Y. Kaya. Optimization Over the Efficient Set of Multi-objective Convex Optimal Control Problems. *Journal of Optimization Theory and Applications*, 147(1): 93–112, 2010. doi: 10.1007/s10957-010-9709-y.

H. Bonnel and J. Morgan. Semivectorial Bilevel Optimization Problem: Penalty Approach. *Journal of Optimization Theory and Applications*, 131(3):365–382, 2006. doi: 10.1007/s10957-006-9150-4.

H. Bonnel and J. Morgan. Optimality Conditions for Semivectorial Bilevel Convex Optimal Control Problems. In D. H. Bailey, H. H. Bauschke, P. Borwein, F. Garvan, M. Théra, J. D. Vanderwerff, and H. Wolkowicz, editors, *Computational and Analytical Mathematics*, pages 45–78, New York, 2013. Springer.

H. Bonnel and M. Morgan. Semivectorial Bilevel Convex Optimal Control Problems: Existence Results. *SIAM Journal on Control and Optimization*, 50(6):3224–3241, 2012. doi: 10.1137/100795450.

S. Dempe. *Foundations of Bilevel Programming.* Kluwer Academic, Dordrecht, 2002.

S. Dempe and J. Dutta. Is bilevel programming a special case of a mathematical program with complementarity constraints? *Mathematical Programming, Series A*, 131:37–48, 2010. doi: 10.1007/s10107-010-0342-1.

S. Dempe, N. Dinh, and J. Dutta. Optimality Conditions for a Simple Convex Bilevel Programming Problem. In R. S. Burachik and J.-C. Yao, editors, *Variational Analysis and Generalized Differentiation in Optimization and Control*, volume 47 of *Springer Optimization and Its Applications*, chapter 7, pages 149–161. Springer, New York, 2010.

S. Dempe, N. Gadhi, and A. B. Zemkoho. New Optimality Conditions for the Semivectorial Bilevel Optimization Problem. *Journal of Optimization Theory and Applications*, 157(1):54–74, 2013. doi: 10.1007/s10957-012-0161-z.

S. Dempe, F. Harder, P. Mehlitz, and G. Wachsmuth. Solving Inverse Optimal Control Problems via Value Functions to Global Optimality. *Tech. Rep. SPP-1962-066, DFG Priority Programme 1962,*, pages 1–35, 2018. URL `https://spp1962.wias-berlin.de/preprints/066.pdf`.

M. Ehrgott. *Multicriteria Optimization.* Springer, Berlin, 2005.

S. Franke, P. Mehlitz, and M. Pilecka. Optimality conditions for the simple convex bilevel programming problem in Banach spaces. *Optimization*, 67(2):237–268, 2018. doi: 10.1080/02331934.2017.1394296.

F. Harder and G. Wachsmuth. Optimality conditions for a class of inverse optimal control problems with partial differential equations. *Optimization*, pages 1–29, 2018. doi: 10.1080/02331934.2018.1495205.

M. Hinze, R. Pinnau, M. Ulbich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer, Heidelberg, 2009.

G. Holler, K. Kunisch, and R. C. Barnard. A Bilevel Approach for Parameter Learning in Inverse Problems. *Preprint arXiv*, pages 1–26, 2018. URL https://arxiv.org/abs/1802.01365v1.

J. Jahn. *Vector Optimization*. Springer, Berlin, 2004.

B. Liu, Z. Wan, J. Chen, and G. Wang. Optimality conditions for pessimistic semivectorial bilevel programming problems. *Journal of Inequalities and Applications*, 41(1), 2014. doi: 10.1186/1029-242X-2014-41.

S. Peitz and M. Dellnitz. A Survey of Recent Trends in Multiobjective Optimal Control—Surrogate Models, Feedback Control and Objective Reduction. *Mathematical and Computational Applications*, 23(2), 2018. doi: 10.3390/mca23020030.

A. Ren and Y. Wang. A novel penalty function method for semivectorial bilevel programming problem. *Applied Mathematical Modelling*, 40(1):135 – 149, 2016. doi: 10.1016/j.apm.2015.04.041.

S. M. Robinson. Stability Theory for Systems of Inequalities, Part II: Differentiable Nonlinear Systems. *SIAM Journal on Numerical Analysis*, 13(4):497–513, 1976. doi: 10.1137/0713043.

K. Shimizu, Y. Ishizuka, and J. F. Bard. *Nondifferentiable and two-level mathematical programming*. Kluwer Academic, Dordrecht, 1997.

G. Stadler. Elliptic optimal control problems with L1-control cost and applications for the placement of control devices. *Computational Optimization and Applications*, 44: 159–181, 2007. doi: 10.1007/s10589-007-9150-9.

F. Tröltzsch. *Optimal Control of Partial Differential Equations*. Vieweg, Wiesbaden, 2009.

A. Zemkoho. *Bilevel Programming: Reformulations, Regularity, and Stationarity*. PhD thesis, Technische Universität Bergakademie Freiberg, 2012.

Y. Zheng and Z. Wan. A solution method for semivectorial bilevel programming problem via penalty method. *Journal of Applied Mathematics and Computing*, 37(1):207–219, 2011. doi: 10.1007/s12190-010-0430-7.

J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979. doi: 10.1007/BF01442543.