SOLVING ILL-POSED BILEVEL PROGRAMS

ALAIN B. ZEMKOHO

ABSTRACT. This paper deals with ill-posed bilevel programs, i.e., problems admitting multiple lower-level solutions for some upper-level parameters. Many publications have been devoted to the standard optimistic case of this problem, where the difficulty is essentially moved from the objective function to the feasible set. This new problem is simpler but there is no guaranty to obtain local optimal solutions for the original optimistic problem by this process. Considering the intrinsic non-convexity of bilevel programs, computing local optimal solutions is the best one can hope to get in most cases. To achieve this goal, we start by establishing an equivalence between the original optimistic problem an a certain set-valued optimization problem. Next, we develop optimality conditions for the latter problem and show that they generalize all the results currently known in the literature on optimistic bilevel optimization. Our approach is then extended to multiobjective bilevel optimization, and completely new results are derived for problems with vector-valued upper- and lower-level objective functions. Numerical implementations of the results of this paper are provided on some examples, in order to demonstrate how the original optimistic problem can be solved in practice, by means of a special set-valued optimization problem.

1. INTRODUCTION

2014 marked 80 years after the publication of the landmark book by Heinrich Freiherr von Stackelberg [44] in the field of economic game theory, which introduced a mathematical analysis of a competition based market economy with a hierarchical structure. The ideas of von Stakelberg remain at the heart of discussions and extensive research in the field of economics; see, for example, the 5/6th issue of the 23rd volume of the Journal of Economic Studies, entirely dedicated to the views of many experts on the various aspects of his ground-breaking work. See also [45] for a recent translation into English of the aforementioned book. Applications of the Stackelberg game also continue to grow in the area of economics, see, e.g., [25] for a recent literature review on its implementation in the areas of supply and marketing channels. One should however recognize that applications of the topic literally exploded with its introduction to the field of optimization by Bracken and McGill [4], where the problem is labeled as "mathematical program with optimization problem in the constraints". The vocabulary related to the Stackelberg game has thus evolved and the problem is commonly referred to as bilevel programming or two-level optimization problem within the optimization community. This is actually the vocabulary

1

that is used in this paper. The model has also evolved within the optimization theory framework, where perturbations on the lower-level feasible set and sometimes on the upper-level feasible set, are incorporated in the development of most results. Applications of bilevel programming have now expanded to areas such as engineering, chemistry, physics, health technology, climate change, etc, see, e.g., [5, 7, 8] for further details and references.

Despite the huge developments that have occur in the field of bilevel programming, in the last four decades, most solution methods that currently exist are designed only for well-posed problems, i.e., for the simplistic case where the reaction of the follower to choices of the leader is predictable. In other words, this refers to the situation where the lower-level player is restricted to a unique optimal solution for each strategy of the upper-level player. Well-posedness can be explicitly expressed by considering the model

$$(\mathsf{P_i}) \qquad \min_x \mathsf{F}(x, y(x)) \ \ \text{s.t.} \ \ x \in \mathsf{X}, \tag{1.1}$$

where F and X denote the upper-level objective function and feasible set, respectively, while $y(x)$ stands for the optimal solution (or reaction) of the lower-level optimization (or player) problem, which can be formulated as

$$\min_y \mathsf{f}(x, y) \ \ \text{s.t.} \ \ y \in \mathsf{K}(x). \tag{1.2}$$

Here, f and $\mathsf{K}(x)$ represent the lower-level objective and feasible set, respectively. Well-posedness is also imposed or implicity assumed in a large number of solution methods designed to solve the standard optimistic bilevel optimization problem

$$(\mathsf{P}) \qquad \min_{x,y} \mathsf{F}(x, y) \ \ \text{s.t.} \ \ x \in \mathsf{X}, \ y \in \mathsf{S}(x), \tag{1.3}$$

which is actually designed to handle problems with multiple lower-level solutions. The set-valued mapping S in the latter problem denotes the set of optimal solutions of the lower-level problem (1.2), i.e., precisely,

$$\mathsf{S}(x) := \arg\min_y \{\mathsf{f}(x, y) : \ y \in \mathsf{K}(x)\}. \tag{1.4}$$

More details on solution methods using explicitly or implicitly imposed well-posedness can be found in [5, 8, 7, 24, 31, 32, 40] and references therein. It is important to note that well-posed problems are naturally advantageous for the leader, but do not reflect most real-world systems, where the reaction of the follower should be based on her/his best interest, whether it suits the upper-level or not. This is the fair principle of freedom of choice. If the follower decides to make choices in favor of the leader, we have the optimistic bilevel optimization problem. Otherwise, the resulting problem is known as the pessimistic bilevel optimization problem.

To put the latter point into context, let us recall that from the conceptual definition of the Stackelberg game, the bilevel program is a hierarchical optimization problem involving an upper-level player (leader) and lower-level player (follower). It is assumed that the upper-level player, who is in control of the variable x, plays first and the lower-level player reacts by choosing a value y (depending on x) under his/her control. This leads to the problem

$$\text{``}\min_x\text{''} \ \mathsf{F}(x, y) \ \ \text{s.t.} \ \ x \in \mathsf{X}, \ y \in \mathsf{S}(x). \tag{1.5}$$

Clearly, in this problem, also known as the upper-level problem, the minimization is only w.r.t. x, the variable controlled by the corresponding player. In the case where the problem is well-posed, $S(x) = \{y(x)\}$ for all $x \in X$. As mentioned above, the resulting problem is either investigated in the form (1.1) or kept in the form (1.3) but the assumption is exploited in solution procedures, see [5, 8, 7, 40] for details and related references.

Recall that as the minimization in (1.5) is only w.r.t. x, the quotation marks are used in order to symbolize the ambiguity that appears for problems with multiple lower-level optimal solutions. This corresponds to ill-posed bilevel programs. A number of authors have suggested regularization techniques to implement well-posed bilevel optimization techniques in the framework of ill-posed programs. This essentially consist to replace the lower-level problem (1.2) by

$$\min_{y} f(x, y) + \alpha\pi(x, y) \ \text{ s.t. } \ y \in K(x), \tag{1.6}$$

where α and π are the regularization parameter and function, respectively. Under suitable assumptions, this perturbation then forces the lower-level problem to approximately generate a unique optimal solution. π can be chosen as $\pi(x, y) := F(x, y)$ (upper-level objective function) [15] or $\pi(x, y) := \|y\|^2$ (Tikhonov regularization) [6, 9, 39]. It was shown in [39], and to some extend in the other related references, that (1.6) often leads to solutions that are far away from the real optimal solution. Thus, it is important to develop methods that directly deal with ill-posed problems without modifying their nature. This is the main goal of this paper.

The original optimistic approach to deal with ill-posed bilevel programs is the problem

$$(P_o) \qquad \min_{x \in X} \varphi_o(x) := \min_{y} \{F(x, y) : \ y \in S(x)\}. \tag{1.7}$$

From the economics viewpoint, this corresponds to a situation where the follower participates in the profit of the leader, i.e., cooperation is possible between both players on the upper-and lower-levels. Problem (P) can be seen as a regularization of the original optimistic problem given that the difficulty in the objective function of (P_o) is essentially moved to the constraint set in (P), where the leader is assumed to have full control over both the upper- and lower-level variables x and y, respectively. Moreover, problems (P) and (P_o) are globally equivalent and a local optimal solution of (P_o) implies a local optimal solution of (P). However, for the converse of the latter statement, a condition has to be satisfied, in particular, the set-valued mapping

$$S_o(x) := \arg\min_{y} \{F(x, y) : \ y \in S(x)\} \tag{1.8}$$

has to be inner semicontinuous [13]. The latter assumption is strong and fails for very simple examples, as shown in the latter reference. Because of the special structure of the above models of the bilevel programs, they are all non-convex optimization problems. Thus it would be important to develop solution methods that compute local optimal solutions for (P_o), as this is the best one can get in most instances. Before providing a taste of the method investigated in this paper, it is worth mentionning the pessimistic bilevel program

$$(P_p) \qquad \min_{x \in X} \varphi_p(x) := \max_{y} \{F(x, y) : \ y \in S(x)\}, \tag{1.9}$$

which reflects the fact that it would not always be possible for the leader to convince the follower to make choices that are favorable for him/her. Hence it is necessary for the upper-level player to bound damages resulting from undesirable selections at the lower-level. One can easily construct examples showing that (P_o) and (P_p) are completely different from each other; see the example below and the other examples in Section 4.

In this paper, we consider the alternative approach to deal with ill-posed bilevel programs which consist to insert the lower-level solution mapping (1.4) in the upper-level objective function. This leads to the problem

$$(P_s) \qquad \min_x \mathcal{F}(x) := F(x, S(x)) \text{ s.t. } x \in X, \qquad (1.10)$$

where the subscript s refers to the fact that this is a set-valued optimization problem. (P_s) can be seen as an extension of the implicit function approach (1.1) to the context of ill-posed bilevel programs. The natural example to motivate our discussion here, is the well-known
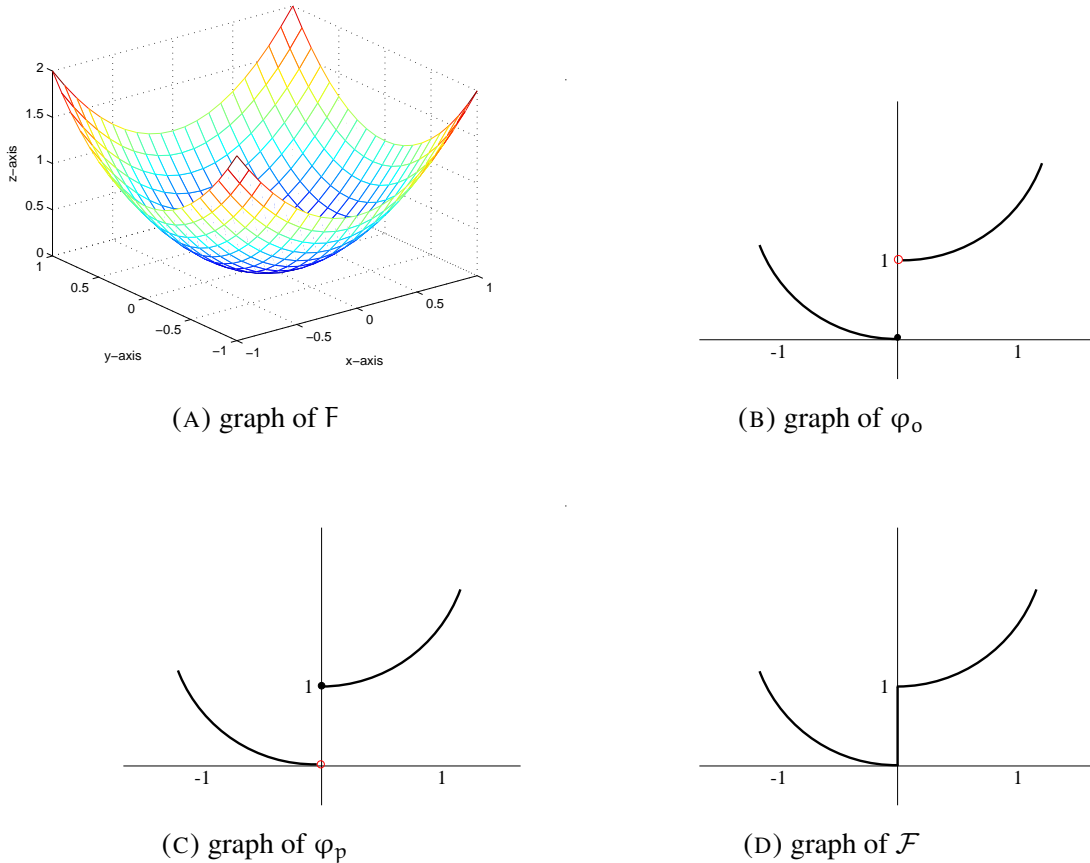

(A) graph of F

(B) graph of φ_o

(C) graph of φ_p

(D) graph of $\mathcal{F}$

FIGURE 1. Variation of the objective mapping with the problem reformulation

one by Lucchetti et al. [30], where

$$F(x, y) := x^2 + y^2, \ X := [-1, 1] \text{ and } S(x) := \arg\min_y \{-xy : y \in [0, 1]\}. \qquad (1.11)$$

One can check that S and $\mathcal{F}$ can be respectively obtained as

$$S(x) = \left\{ \begin{array}{ll} 0 & \text{if } x < 0, \\ 1 & \text{if } x > 0, \\ [0,1] & \text{if } x = 0, \end{array} \right. \text{ and } \mathcal{F}(x) = \left\{ \begin{array}{ll} x^2 & \text{if } x < 0, \\ x^2 + 1 & \text{if } x > 0, \\ \{y^2 : y \in [0,1]\} & \text{if } x = 0. \end{array} \right.$$

In Figure 1, we map the objective "function" of (P), (P_o), (P_p) and (P_s), respectively. This picture highlights the typical behavior of the corresponding graphs for ill-posed bilevel programs. Going from a smooth function F in (P), the best we can generally get for φ_o and φ_p is lower-semicontinuity and upper-semicontinuity, respectively, at points where S is multi-valued [13]. As for $\mathcal{F}$, it is set-valued at such points. To be precise, note that the graph of $\mathcal{F}$ coincides with that of (P_o) and (P_p) at points where S (1.4) is single-valued. At the point $\bar{x} = 0$ where S is multi-valued, $\mathcal{F}(\bar{x})$ is the interval connecting $\varphi_o(\bar{x})$ and $\varphi_p(\bar{x})$. In the next section, we show that this property remains valid for a number of problems.

To analyze the optimal solution of the problem, we consider the cone $K := \mathbb{R}_+$. Hence, following Jahn [27], a couple $(\bar{x}, \bar{z}) \in \text{gph } \mathcal{F} := \{(x, z) \in \mathbb{R}^n \times \mathbb{R} : z \in \mathcal{F}(x)\}$ will be said to be a local Pareto optimal solution of (P_s) if there exists a neighborhood U of $\bar{x}$ such that

$$(\mathcal{F}(X \cap U) - \bar{z}) \cap] - \infty, 0[= \emptyset. \tag{1.12}$$

It would be said to be global if the neighborhood of $\bar{x}$ can be taken as large as possible, that is, $U = \mathbb{R}^n$. Following this definition, we have the following three cases for (1.11):

(a) For $x < 0$ and $z = \mathcal{F}(x) = x^2$, we have $-x^2 \in (\mathcal{F}(X) - z) \cap] - \infty, 0[$. Hence (x, z) cannot be a Pareto optimal point.

(b) For $x > 0$ and $z = \mathcal{F}(x) = x^2 + 1$, it holds that $-x^2 - 1 \in (\mathcal{F}(X) - z) \cap] - \infty, 0[$. Similarly, this means that (x, z) is not a Pareto optimal point.

(c) For $x = 0$, $z \in \mathcal{F}(x)$ means that there exists $a \in [0, 1]$ such that $z = a^2$. Now, consider the case where $z = a^2$ for some $a \in]0, 1]$; then it follows that $-a^2 \in (\mathcal{F}(X) - z) \cap] - \infty, 0[$. This implies that no point of the form $(0, a)$ with $a \in]0, 1]$ can be a Pareto optimal solution of the problem. Finally, for $z = a = 0$, we have

$$\mathcal{F}(X) - z = \{x^2 : x \in [-1, 1[\} \cup \{x^2 + 1 : x \in]0, 1]\} \subseteq [0, \infty[.$$

In conclusion, $(0, 0)$ is the unique optimal solution of problem (1.11) in the sense of (1.12). On the other hand, one can easily see from Figure 1 (c) that $x = 0$ is the optimal solution for the original optimistic version (1.7) of the problem in (1.11). This observation motivates us to look at what happens in a general framework. We actually prove, see next section, that problems (P_o) and (P_s) are globally and locally equivalent, without any assumption. Note that (P_p) has no optimal solution.

It is worth to recall here that (1.10) and its link with the optimistic problem was mentioned in [7]. However, we are not aware of any thorough analysis of the problem. Recently, a branch and bound method based on a semi-infinite reformulation was proposed in [46] to compute approximate global solutions of problems (P_o) and (P_p). Using a KKT reformulation for S, (1.4) a combination of solvers is used in [43] to construct a procedure to compute global solutions for a special case of (P_p). For the first time, optimality conditions for problems (P_o) and (P_p) were derived in [13] and [14], respectively. Though a small number of publications have recently focused on the computation of stationary/local

optimal solutions (without imposing lower-level uniqueness) of (P), see [11, 28, 47], we are not aware of any method that locally solves (P_o). It is our opinion that the latter problem is the most important model in the context of ill-posed bilevel optimization. At first, as already mentioned above, a local optimal solution of (P_o) generates a local optimal solution of (P), but the converse in not true, cf. [13]. In the latter paper, it is also shown that the optimality conditions of (P_o) are equivalent to those of (P) under a mild assumption.

Secondly, note that for the bilevel program (1.5) where there is no perturbation in the lower-level problem, i.e., $K(x) := K$ for all $x \in X$, Molodtsov [33] proposed the following family of problems in the form of (P_o) to solve the corresponding pessimistic problem:

$$\min_{x \in X} \varphi_o^{\theta,\epsilon}(x) := \min_y \{F(x,y) : y \in S^{\theta,\epsilon}(x)\}$$
$$\text{with} \quad \begin{cases} S^{\theta,\epsilon}(x) := \{y \in K : f(x,y) - \theta F(x,y) \leq \varphi^\theta(x) + \epsilon\}, \\ \varphi^\theta(x) := \min_y \{f(x,y) - \theta F(x,y) : y \in K\}, \end{cases} \quad (1.13)$$

where $\theta \geq 0$ and $\epsilon \geq 0$. Observe that if θ and ϵ are set to zero, we get problem (P_o). Denote by $\varphi_{oo}^{\theta,\epsilon}$ the optimal value of a given problem from the family in (1.13) and let φ_{pp} be the optimal value of the corresponding pessimistic problem, (P_p). Then, it is shown in [33] that we have

$$\varphi_{pp} = \lim_{n \to \infty} \varphi_{oo}^{\theta_n,\epsilon_n} \text{ if } \theta_n \to 0^+, \epsilon_n \to 0^+ \text{ with } \frac{\theta_n}{\epsilon_n} \to 0^+$$

provided X and K are metric compact spaces and the functions F and f are continuous on $X \times K$. This result is derived in [29] under weaker assumptions. The result clearly indicates that solving problem (P_o) will provide a new direction to develop solution methods for the pessimistic bilevel program (P_p), which is the most difficult class of ill-posed bilevel programs. Further recall that it was already shown in [14] that the optimality conditions of (P_p) can be obtained from (P_o).

The latter point mainly followed from the stability analysis of the two-level value function φ_o (1.7) developed in the paper [13]. It can easily be checked that the stability/sensitivity analysis of this function leads to the stability/sensitivity analysis of the problems (P), (P_o) and (P_p). In summary, studying problem (P_o) provides a base to solve problems (P) and (P_p), and also to develop stability analysis results for the three problems.

The aim of this paper is not to provide a formal algorithm for (P_o), but rather a first step towards that goal, consisting of a theoretical framework with numerical implementations, showing how the problem can be solved in practice via the set-valued problem (P_s). The first motivation to consider this approach is that problem (P_o) is a three-level optimization problem while (P_s) is just two-level, thus simpler. Secondly, (P_s) is an extension of the implicit function model (P_i), thus solution schemes known for the latter problem could be implemented to the former one; cf. discussion in Section 5. The work in this paper can also be put in the context of works developing equivalence between bilevel programs and multiobjective/set-valued optimization problems. This aspect of research has drawn attention recently and a thorough review of the literature on related developments can be found in [42]. The main idea in these works is to establish an equivalence between (P) and a multiobjective program or an optimization problem over an efficient set. The results are based on specially crafted order relations, which are sometimes difficult to work with and may

not be induced by cones with nice properties [22]. Our approach is completely different as our focus is on problem (P_o) and our equivalent problem is set-valued. Moreover, we go a step further and provide a clear path to solve the new set-valued optimization problem. Also observe that the order relation that we work with (1.12) is the simplest Pareto concept induced by the most basic cone $\mathbb{R}_+$, which satisfies the properties relevant for tractable solution procedures.

The remainder of the rest of the paper is as follows. Our main focus in the next section is to formulate the observations made on the example above in the context of more general problems. Thus, we first establish the link between the objective maps of (P_o), (P_p) and (P_s). Next, we investigate the relationships between all the models above, including (P_i) and (P), while particularly paying attention to the link between the set-valued model and the original optimistic one. Optimality conditions are then derived from the perspective of (1.10) and shown to generalize all the results on optimistic bilevel optimization currently available in the literature. In Section 3, another powerful aspect of our approach is revealed, as we show that it helps to easily extend the results from the previous section to multiobjective bilevel optimization. In particular, we derive completely new results for parametric set-valued optimization, semivectorial bilevel optimization and multiobjective bilevel programs with vector-valued upper-and lower-level objective functions. Numerical implementations are provided in Section 4 in order to illustrate how the results of this paper can be used to solve (P_o) in practice by means of (P_s). Moreover, the examples considered contribute to a better understanding of the models discussed above.

2. OPTIMISTIC BILEVEL OPTIMIZATION

In this section we assume that the upper-and lower-level objective functions F and f are defined from $\mathbb{R}^n \times \mathbb{R}^m$ to $\mathbb{R}$ and are continuously differentiable, while the upper-and lower-level feasible sets X and $K(x)$, $x \in X$ are respectively closed.

2.1. **Relationships with the set-valued model.** The aim of this subsection is to confirm the observations made on the motivating example above. Before moving to our main result here, establishing the equivalence between (P_s) and (P_o), we quickly show that the objective "functions" of (P_s), (P_o) and (P_p) are closely related. To proceed, we denote by $|S(x)|$ the cardinality of $S(x)$ and by $X^u := \{x \in X : |S(x)| = 1\}$ and $X^m := \{x \in X : |S(x)| \geq 2\}$. Furthermore, we assume from here on that the two-level value functions φ_o (1.7) and φ_p (1.9) are well-defined, i.e., $|\varphi_o(x)| < \infty$ and $|\varphi_p(x)| < \infty$, for all $x \in X$.

Theorem 2.1. The following properties are satisfied:

(a) $(\mathrm{gph}\,\varphi_o \cup \mathrm{gph}\,\varphi_p) \cap (X \times \mathbb{R}) \subseteq \mathrm{gph}\,\mathcal{F}$.
(b) Let $F(x,y) := a(x)^\top y + b(x)$ with $a : \mathbb{R}^n \to \mathbb{R}^m$ and $b : \mathbb{R}^n \to \mathbb{R}$, and assume that S is convex-valued on X^m, i.e., $S(x)$ is convex, for all $x \in X^m$. Then,

$$
\mathrm{gph}\,\mathcal{F} \cap (X \times \mathbb{R}) = \begin{aligned} &\{(x, \varphi_o(x)) : x \in X^u\} \\ &\cup \{(x, \lambda\varphi_o(x) + (1-\lambda)\varphi_p(x)) : x \in X^m, \lambda \in [0,1]\}, \end{aligned} \tag{2.1}
$$

or equivalently, for all $x \in X$,

$$\mathcal{F}(x, S(x)) = \begin{cases} \varphi_o(x) = \varphi_p(x) & \text{if } x \in X^u, \\ [\varphi_o(x), \varphi_p(x)] & \text{if } x \in X^m. \end{cases}$$

Proof. For (a), let $(x, z) \in (\text{gph } \varphi_o \cup \text{gph } \varphi_p) \cap (X \times \mathbb{R})$. Then, $x \in X$ and $z = \varphi_o(x)$ or $z = \varphi_p(x)$. In either case, it follows from the definition of the corresponding function that we can find $y \in S(x)$ such that $z = F(x, y)$. Thus, $(x, z) \in \text{gph } \mathcal{F}$. As for (b), denote by C the right-hand-side of (2.1) and let $(x, z) \in \text{gph } \mathcal{F} \cap (X \times \mathbb{R})$. If $x \in X^u$, we get

$$(x, z) = (x, \mathcal{F}(x)) = (x, F(x, S(x)) = (x, \varphi_o(x)) = (x, \varphi_p(x)),$$

following the definitions of φ_o (1.7) and φ_p (1.9). Similarly, note that for any $x \in X^m$, it follows that we can find $y', y'' \in S(x)$ such that for all $\lambda \in [0, 1]$,

$$\lambda \varphi_o(x) + (1 - \lambda)\varphi_p(x) = a(x)^\top (\lambda y' + (1 - \lambda)y''). \tag{2.2}$$

On the other hand, with $x \in X^m$, we have $z \in F(x, y)$ for some $y \in S(x)$. Hence, $z = F(x, y) = F(x, \lambda y + (1 - \lambda)y)$ for any $\lambda \in [0, 1]$. It then follows from (2.2) that $(x, z) \in C$. Conversely, let $(x, z) \in C$. If $x \in X^u$, then we obviously have $(x, z) \in \text{gph } \mathcal{F}$. Finally, if $x \in X^m$, then by (2.2), it holds that $z = a(x)^\top (\lambda y' + (1 - \lambda)y'') \in \mathcal{F}(x)$, for some $y', y'' \in S(x)$ and $\lambda \in [0, 1]$, considering the fact that $S(x)$ is a convex set. $\square$

In the next theorem, we prove that the original optimistic bilevel program is equivalent to the set-valued model/implicit function-type problem in (1.10). For the latter problem, we use the concept of Pareto optimal solution in (1.12), while for the former problem, the standard notion from scalar objective optimization is applied.

Theorem 2.2. The point $\bar{x}$ is a local (resp. global) optimal solution of problem (P_o) if and only if there exists a vector $\bar{z} \in \mathcal{F}(\bar{x})$ such that $(\bar{x}, \bar{z})$ is a local (resp. global) Pareto optimal solution of problem (P_s).

Proof. We prove only the local case, as the global one can easily be shown in a similar manner. As first step, we claim that $(\bar{x}, \bar{z}) \in \text{gph } \mathcal{F}$ is a local Pareto optimal solution of (P_s) if and only if there exists a neighborhood U of $\bar{x}$ and a vector $\bar{y} \in S(\bar{x})$ such that we have $\bar{z} = F(\bar{x}, \bar{y})$ and

$$F(\bar{x}, \bar{y}) \leq F(x, y), \ \forall x \in X \cap U, \ \forall y \in S(x). \tag{2.3}$$

To check this, observe that $(\bar{x}, \bar{z}) \in \text{gph } \mathcal{F}$ is equivalent to saying that there exists $\bar{y} \in S(\bar{x})$ such that $\bar{z} = F(\bar{x}, \bar{y})$. Inserting the latter value in (1.12), while noting that

$$\mathcal{F}(X \cap U) = \{F(x, y) : \ x \in X \cap U, \ y \in S(x)\},$$

we can confirm that the statement is true.

Now, let $\bar{x}$ be a local optimal solution of (P_o); then, there exists a neighborhood U of $\bar{x}$ such that we have $\varphi_o(\bar{x}) \leq \varphi_o(x), \forall x \in X \cap U$. Considering the definition of φ_o, the latter implies that we can find some $\bar{y} \in S(\bar{x})$ such that

$$F(\bar{x}, \bar{y}) = \varphi_o(\bar{x}) \leq \varphi_o(x) \leq F(x, y), \ \forall x \in X \cap U, \ \forall y \in S(x).$$

Comparing this relation with (2.3), it is clear that $(\bar{x}, \bar{z}) \in \text{gph } \mathcal{F}$ (with $\bar{z} = F(\bar{x}, \bar{y})$) is a local Pareto optimal solution of (P_s). Conversely, suppose that $\bar{x}$ is not a local optimal

solution of problem (P_o), and let $\mathcal{V}(\bar{x})$ denote the collection of all the neighborhoods of this point in $\mathbb{R}^n$. Then it holds that

$$\forall U \in \mathcal{V}(\bar{x}), \ \exists x^* \in X \cap U \text{ s.t. } \varphi_o(x^*) < \varphi_o(\bar{x}).$$

This implies that we have

$$\forall U \in \mathcal{V}(\bar{x}), \ \exists x^* \in X \cap U, \ \exists y^* \in S(x^*) \text{ s.t. } F(x^*, y^*) = \varphi_o(x^*) < \varphi_o(\bar{x}).$$

Finally, returning again to the definition of $\varphi_o(\bar{x})$, it follows that

$$\forall y \in S(\bar{x}), \ \forall U \in \mathcal{V}(\bar{x}), \ \exists x^* \in X \cap U, \ \exists y^* \in S(x^*)$$
$$\text{s.t. } F(x^*, y^*) = \varphi_o(x^*) < \varphi_o(\bar{x}) \le F(\bar{x}, y).$$

Comparing this relation with (2.3), it clearly appears that the definition of the local Pareto optimal solution is contradicted. $\square$

This may seem surprising as solving set-valued optimization problems generally leads to a Pareto front. But one has to bear in mind that (P_s) is special problem which still takes its values on subsets of the real-line. The relationship, w.r.t. optimal solution, between the various formulation of bilevel optimization mentioned in the previous section can be summarized as follows:

$$(P_p) \overset{(**)}{\Longleftrightarrow} (P_i) \overset{(**)}{\Longleftrightarrow} (P_o) \Longleftrightarrow (P_s)$$
$$\Updownarrow (*)$$
$$(P)$$

Here, $(*)$ is used to symbolize the fact that the standard and original optimistic problems are equivalent only in the global case. As for condition $(**)$, it stands for the uniqueness of lower-level optimal solution for all upper-level parameter; i.e., $S(x) = \{y(x)\}$ for all $x \in X$.

2.2. **Optimality conditions.** To proceed in this subsection, we recall some concepts from the literature on variational analysis. For a closed set $C \subseteq \mathbb{R}^n$, the Mordukhovich normal cone [34] to C is the set

$$N_C(\bar{x}) := \{v \in \mathbb{R}^n : \ \exists v_k \to v, \ x_k \to \bar{x} \ (x_k \in C) \text{ with } v_k \in \widehat{N}_C(x_k)\}, \qquad (2.4)$$

for $\bar{x} \in C$, where $\widehat{N}_C$ denotes the Fréchet normal cone to C. Note that if $C := \psi^{-1}(\Xi)$, where $\Xi \subseteq \mathbb{R}^m$ is a closed set and $\psi : \mathbb{R}^n \to \mathbb{R}^m$ a continuously differentiable function at $\bar{x}$, then we have

$$N_C(\bar{x}) \subseteq \nabla\psi^\top N_\Xi(\psi(\bar{x})), \qquad (2.5)$$

provided $v = 0$ whenever $\nabla\psi(\bar{x})^\top v = 0$ with $v \in N_\Xi(\psi(\bar{x}))$, see, e.g. [35]. Equality holds in (2.5) if $N_\Xi(\psi(\bar{x})) = \widehat{N}_\Xi(\psi(\bar{x}))$. This is obviously the case if Ξ is a convex set. Finally, for a set-valued map $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, the coderivative [34] of Ψ at some point $(\bar{x}, \bar{y}) \in \text{gph } \Psi$ is a positively homogeneous set-valued mapping $D^*\Psi(\bar{x}|\bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, defined by

$$D^*\Psi(\bar{x}|\bar{y})(v) := \{u \in \mathbb{R}^n : \ (u, -v) \in N_{\text{gph }\Psi}(\bar{x}, \bar{y})\}, \qquad (2.6)$$

for all $v \in \mathbb{R}^m$. Here, $N_{\text{gph }\Psi}$ denotes the basic normal cone (2.4) to $\text{gph } \Psi$.

Our main goal here is to derive the necessary optimality conditions of (P_s) and subsequently confirm the equivalence in the previous subsection with (P_o) by showing how the

obtained result actually generalizes well-known ones on optimistic bilevel optimization. For $w \in \{x, y\}$ and $q \in \{n, m\}$, the following notations are used in the sequel:

$$\nabla_w F(x, y) := \left[\frac{\partial F}{\partial w_1}(x, y), \ldots, \frac{\partial F}{\partial w_q}(x, y) \right]^\top, \quad \nabla F(x, y) := \left[\nabla_x F(x, y)^\top, \nabla_y F(x, y)^\top \right]^\top.$$

Theorem 2.3. Let $(\bar{x}, \bar{z}) \in \mathrm{gph}\, \mathcal{F}$ be a local Pareto optimal solution of problem (P_s) and assume that S (1.4) is closed, locally bounded around $\bar{x}$ and

$$D^*S(\bar{x}|y)(0) \cap (-N_X(\bar{x})) = \{0\} \quad \text{for all } y \in S(\bar{x}) \text{ s.t. } \bar{z} = F(\bar{x}, y). \tag{2.7}$$

Then, there exists $\bar{y} \in S(\bar{x})$ with $\bar{z} = F(\bar{x}, \bar{y})$ such that we have

$$-\nabla_x F(\bar{x}, \bar{y}) \in D^*S(\bar{x}|\bar{y})(\nabla_y F(\bar{x}, \bar{y})) + N_X(\bar{x}). \tag{2.8}$$

Proof. To proceed, first note that $\mathcal{F} := F \circ E$ with $E(x) := (x, S(x))$ for all x. One can easily check that E is closed since S is assumed to be closed. It is also necessary to show that $\mathcal{F}$ is closed. Hence, let $(x^k, z^k) \in \mathcal{F}$ such that $(x^k, z^k) \to (\bar{x}, \bar{z})$. Then we can find a sequence $\{y^k\}$ such that $y^k \in S(x^k)$ and $z^k = F(x^k, y^k)$. Since S is locally bounded around $\bar{x}$, then the sequence $\{y^k\}$ is also bounded. Thus, by the Bolzano-Weierstrass theorem, we can extract a convergent subsequence of this sequence, that we also denote by $\{y^k\}$ provided there is no confusion. Taking into account the closedness of S, it follows that $y^k \to \bar{y} \in S(\bar{x})$. Furthermore, observe that the continuous differentiability of F guaranties that $\bar{z} = F(\bar{y})$. Thus $\bar{z} \in \mathcal{F}(\bar{x})$, implying that $\mathcal{F}$ is closed.

Let us now consider the set-valued mapping $M(x, z) := E(x) \cap F^{-1}(z)$ and show that it is locally bounded. Given that S is locally bounded around $\bar{x}$, denote by U a neighborhood of $\bar{x}$ such that $S(U) \subseteq V$, where V is a bounded set in $\mathbb{R}^m$. Now let W be a neighborhood of $\bar{z}$. Then we have

$$M(U \times W) = E(U) \cap F^{-1}(W) = (U \times S(U)) \cap F^{-1}(W) \subseteq U \times S(U) \subseteq U \times V.$$

It is then obvious that by choosing U as a bounded neighborhood of $\bar{x}$, which can be done without lost of generality, we can conclude that M is locally bounded around $(\bar{x}, \bar{z})$. Combining this with the closedness of E, we have from the coderivative chain rule by Mordukhovich [35, Corollary 5.3] that

$$D^*\mathcal{F}(\bar{x}|\bar{z})(z^*) \subseteq \left\{ D^*E(\bar{x}|\bar{w})(\nabla F(\bar{w})z^*) : \bar{w} \in E(\bar{x}) \cap F^{-1}(\bar{z}) \right\}. \tag{2.9}$$

To fully formulate this estimate in terms of our problem data, we have to evaluate the coderivative $D^*E(\bar{x}, \bar{w})(\nabla F(\bar{w})z^*)$. For this purpose, observe that the graph of E can be rewritten as

$$\mathrm{gph}\, E = \{(x, y, z) : \psi(x, y, z) \in \Lambda\}$$
$$\text{with } \psi(x, y, z) := (x - y, y, z) \text{ and } \Lambda := \{0_n\} \times \mathrm{gph}\, S.$$

Obviously, Λ is closed given that S is closed; and one can easily verify that $\alpha = 0$, $\beta = 0$ and $\gamma = 0$ for any triple (α, β, γ) satisfying

$$\psi(\bar{x}, \bar{y}, \bar{z})^\top(\alpha, \beta, \gamma) = 0 \text{ and } (\alpha, \beta, \gamma) \in N_\Lambda(\psi(\bar{x}, \bar{y}, \bar{z})).$$

Thus, it follows from inclusion (2.5) that we have

$$N_{\mathrm{gph}\, E}(\bar{x}, \bar{y}, \bar{z}) \subseteq \left\{ (\alpha, -\alpha + \beta, \gamma) : \alpha \in \mathbb{R}^n, (\gamma, \beta) \in N_{\mathrm{gph}\, S}(\bar{y}, \bar{z}) \right\}.$$

Considering the definition of the coderivative, it follows that we have the inclusion
$$D^*E(\bar{x}|(\bar{y}, \bar{z}))(y^*, z^*) \subseteq y^* + D^*S(\bar{y}|\bar{z})(z^*).$$
Inserting the latter in (2.9), we get
$$D^*\mathcal{F}(\bar{x}|\bar{z})(z^*) \subseteq \Big\{ \nabla_x F(\bar{x}, \bar{y})z^* + D^*S(\bar{x}|\bar{y})(\nabla_y F(\bar{x}, \bar{y})z^*) : \tag{2.10}$$
$$\bar{y} \in S(\bar{x}), F(\bar{x}, \bar{y}) = \bar{z} \Big\}$$
while considering the fact that $\bar{w} \in E(\bar{x})$ reduces to $\bar{w} = (\bar{x}, \bar{y})$ for some $\bar{y} \in S(\bar{x})$. Obviously, we can deduce that for $z^* = 0$, it holds that
$$D^*\mathcal{F}(\bar{x}|\bar{z})(0) \subseteq \Big\{ D^*S(\bar{x}|\bar{y})(0) : \bar{y} \in S(\bar{x}), F(\bar{x}, \bar{y}) = \bar{z} \Big\}.$$
This obviously guaranties that (2.7) is a sufficient condition for the following equality to be satisfied:
$$D^*\mathcal{F}(\bar{x}|\bar{z})(0) \cap (-N_X(\bar{x})) = \{0\}. \tag{2.11}$$
Hence, applying Theorem 4.5 from the paper by Ha [23], it follows that we can find some $z^* \in K^*$ with $|z^*| = 1$ such that
$$0 \in D^*\mathcal{F}(\bar{x}|\bar{z})(z^*) + N_X(\bar{x}) \tag{2.12}$$
while taking into account that the set X is closed (by assumption, cf. Introduction) and the set-valued mapping $\mathcal{F}$ is closed, as shown above. Inserting (2.10) in (2.12), we get (2.8) while considering the fact that the cone $K := \mathbb{R}_+$ is self dual; thus, the only element of K^* satisfying $|z^*| = 1$ is 1. This terminates the proof of the result. $\qquad \square$

We now analyze the link between this result and other known results on optimality conditions for optimistic bilevel optimization. Considering the equivalence established above between (P_s) and (P_o), we start with the latter problem. If $\bar{x}$ is a local optimal solution of (P_o) and S is closed, then we have the optimality condition (2.8) for some $\bar{y} \in S(\bar{x})$, provided that the mapping S_o (1.8) is inner semicompact at $\bar{x}$ and the following QC holds:
$$\partial^\infty \varphi_o(\bar{x}) \cap (-N_X(\bar{x})) = \{0\}, \tag{2.13}$$
cf. [13, 36]. Note that by definition, $\partial^\infty \varphi_o(\bar{x}) := \{v \in \mathbb{R}^n : (v, 0) \in N_{\mathrm{epi}\,\varphi_o}(\bar{x}, \varphi_o(\bar{x}))\}$ and considering the fact that φ_o (1.7) is an optimal value function, it holds that
$$\partial^\infty \varphi_o(\bar{x}) \subseteq \bigcup_{y \in S_o(\bar{x})} D^*S(\bar{x}|y)(0) \subseteq \bigcup_{y \in S(\bar{x})} D^*S(\bar{x}|y)(0), \tag{2.14}$$
see [36], where the second inclusion follows from $S_o(\bar{x}) \subseteq S(\bar{x})$, cf. (1.8). Obviously, inserting the second inclusion of (2.14) in (2.13), we get the QC in (2.7). As for the inner semicompactness of S_o necessary for problem (P_o), recall that it is automatically satisfied if S_o is locally bounded. With the inclusion $S_o(x) \subseteq S(x)$ for all x, this implies that S_o is locally bounded under the local boundedness of S imposed in Theorem 2.3. Details on the optimality conditions of (P_o) can be found in [13].

For the link with the optimality conditions result for the standard optimistic problem, one can easily check that for a local optimal solution $(\bar{x}, \bar{y})$ of (P), it holds that
$$0 \in \nabla F(\bar{x}, \bar{y}) + N_{\mathrm{gph}\,S}(\bar{x}, \bar{y}) + N_{X \times \mathbb{R}^m}(\bar{x}, \bar{y}), \tag{2.15}$$

provided the lower-level solution map S (1.4) is closed and the constraint qualification (CQ)

$$N_{\mathrm{gph}\,S}(\bar{x}, \bar{y}) \cap (-N_{X \times \mathbb{R}^m}(\bar{x}, \bar{y})) = \{(0, 0)\} \tag{2.16}$$

is satisfied. Conditions (2.15) and (2.16) can easily be recast as (2.8) and (2.7), respectively. The local boundedness of S and the quantifier "for all" (see QC (2.7)) are not needed for (P) because it is assumed here that the upper-level player is in control of both the upper-and lower-level variables. These assumptions are needed in (P_o) and (P_s) in order to mitigate the choices of the lower-level player $y \in S(\bar{x})$ as the leader does not have control over them. Note that special cases of this result on (P) are derived in [10, 21]. For further clarity on the fact that the result in Theorem 2.3 will effectively lead to standard results in the literature on KKT and value functions reformulations of (P_o) and (P), the interested reader is referred to the discussion in [52, pages 51-52]. See also [18, 16, 17, 48, 50, 51] for other approaches and special cases for optimality conditions of (P).

Finally, considering (P_i) where the lower-level solution function $y(.)$ is differentiable, all the assumptions in Theorem 2.3 are satisfied, and it is well known that a necessary condition of optimality for a point $\bar{x} \in X$ is that

$$0 \in \nabla_x F(\bar{x}, \bar{y}) + \nabla y(\bar{x})^\top \nabla_y F(\bar{x}, \bar{y}) + N_X(\bar{x});$$

see, e.g., [7, 40]. Obviously, the latter condition is the corresponding version of (2.8). The case where the lower-level solution function $y(.)$ is locally Lipschitz continuous and not necessarily differentiable can also be obtained in a similar way, cf. [16].

3. MULTIOBJECTIVE BILEVEL OPTIMIZATION

In this section, we consider problem (1.5), where the upper- and lower-level objective functions are vector-valued and respectively defined by $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{l_1}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{l_2}$ with $l_1 \geq 1$ and $l_2 \geq 2$ being natural numbers. Though some of the results in this section will be valid for various solution concepts, we focus our attention on the case where the lower-level solution set-valued mapping

$$S_{wef}(x) := \arg\min_z \{f(x, z) : \ z \in K(x)\} \tag{3.1}$$

is made of weakly efficient points while the upper-level solution points will be understood in the sense of Pareto. Concretely, we would write $\bar{z} \in S_{wef}(\bar{x})$ if $f(\bar{x}, \bar{z}) - f(x, z) \in \mathrm{int}\,\mathbb{R}^{l_2}_+$ for all $(x, z) \in \mathrm{gph}\,K$, whereas $(\bar{x}, z^o) \in \mathrm{gph}\,\mathcal{F}$ would be a local optimal solution of the corresponding set-valued model

$$\min_x \mathcal{F}(x) := F(x, S_{wef}(x)) \ \text{s.t.} \ x \in X, \tag{3.2}$$

in the sense of Pareto, if there exists a neighborhood U of the point $\bar{x}$ such that the equality $(\mathcal{F}(X \cap U) - z^o) \cap (-\mathbb{R}^{l_1}_+ \setminus \{0_{l_1}\}) = \emptyset$ is satisfied.

First we consider the original optimistic version of the semivectorial bilevel program

$$\min_{x \in X} \varphi_o(x) := \min_y \{F(x, y) : \ y \in S_{wef}(x)\}, \tag{3.3}$$

which corresponds to the case where $l_1 = 1$ and $l_2 \geq 2$. In this case, the counterpart of Theorem 2.2 remains valid. We state the result in the next theorem without the proof, as it follows exactly on the path of the proof of Theorem 2.2. The case $l_1 \geq 2$ and $l_2 \geq 2$ will be

investigated in a separate work, as it involves the extension of the concept of optimal value function to the set-valued optimization framework, something which is out of the scope of this paper.

Theorem 3.1. If $l_1 = 1$ and $l_2 \geq 2$, then $\bar{x}$ is a local (resp. global) optimal solution of problem (3.3) if and only if there exists a vector $\bar{z} \in \mathcal{F}(\bar{x})$ such that $(\bar{x}, \bar{z})$ is a local (resp. global) Pareto optimal solution of problem (3.2).

Next, we derive the coderivative of the lower-level weakly efficient optimal solution set-valued mapping S_{wef} (3.1). Our approach is based on the sacalarization technique well-known in multiobjective optimization. Thus, we consider the solution map S_s of the scalarized problem, defined by

$$S_s(x, y) := \arg\min_z \{ \langle y, f(x, z) \rangle : \ z \in K(x) \} \tag{3.4}$$

and which has the same structure as the solution map of our lower-level problem in the previous section. Our aim is to write the coderivative of S_{wef} in terms of the coderivative of S_s, which can easily be obtained afterwards by using standard methods from stability analysis of scalar objective optimization problems. To proceed, consider the set

$$Y := \{ y \in \mathbb{R}^{l_2} : \ \geq 0, \ \|y\| = 1 \}$$

of scalarization parameters of (3.4) and let $\bar{y} \in Y$. We denote by $I(\bar{y}) := \{ i \in \{1, \ldots, l_2\} : \bar{y}_i = 0 \}$ and e_i a unit vector from $\mathbb{R}^{l_2}$ having 1 as component at position i and 0 everywhere else. Furthermore, define the vector

$$u_i^{\bar{y}}(\alpha, \beta) := \left\{ \begin{array}{ll} \alpha \bar{y}_i & \text{if } \ \bar{y}_i > 0, \\ \beta_i & \text{if } \ \bar{y}_i = 0, \end{array} \right. \tag{3.5}$$

for some $(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}_+^{|I(\bar{y})|}$ with $|I(\bar{y})|$ denoting the cardinality of the set $I(\bar{y})$. For any given vector $\bar{y} \in Y$, we collect all the corresponding vectors from (3.5) in the set $U(\bar{y})$. Then, we have the following result.

Proposition 3.2. Assume that S_s is closed and let $\bar{z} \in S_{wef}(\bar{x})$. Further suppose that for all $\bar{y} \in Y$ such that $\bar{z} \in S_s(\bar{x}, \bar{y})$, it holds that

$$\left[(0, -u) \in D^* S_s((\bar{x}, \bar{y})|\bar{z})(0), \ u \in U(\bar{y}) \right] \implies u = 0. \tag{3.6}$$

Then for all $z^* \in \mathbb{R}^m$, we have the following estimate for the coderivative of S_{wef}:

$$D^* S_{wef}(\bar{x}|\bar{z})(z^*) \subseteq \Big\{ x^* : \ (x^*, -\bar{u}) \in D^* S_s((\bar{x}, \bar{y})|\bar{z})(z^*), \tag{3.7}$$
$$\bar{y} \in Y, \ \bar{u} \in U(\bar{y}), \ \bar{z} \in S_s(\bar{x}, \bar{y}) \Big\}.$$

Proof. Note that $S_{wef}(x) = S_s \circ Y_s(x)$ for all $x \in X$, where $Y_s(x) := (x, Y)$. Hence,

$$\text{gph} \, Y_s = \psi^{-1}(\Lambda) \ \text{ with } \ \Lambda := \{0_n\} \times Y \ \text{ and } \ \psi(x, y, z) := (x - y, z).$$

Observe that Λ is closed and convex. Moreover, one can easily check that $\alpha = 0$ and $\beta = 0$ for any couple (α, β) satisfying

$$\psi(\bar{x}, \bar{y}, \bar{z})^\top (\alpha, \beta) = 0 \text{ and } (\alpha, \beta) \in N_\Lambda(\psi(\bar{x}, \bar{y}, \bar{z})).$$

Thus, it follows from (2.5) and the related discussion that we have the equality

$$N_{\mathrm{gph}\,Y_s}(\bar{x}, \bar{y}, \bar{z}) = \big\{(\alpha, \beta, \gamma): \ \alpha = -\beta, \ \gamma \in N_Y(\bar{z})\big\}.$$

Subsequently, the definition of the notion of coderivative leads us to the formula

$$D^*Y_s(\bar{x}|(\bar{y}, \bar{z}))(y^*, z^*) = \begin{cases} \{y^*\} & \text{if } y^* \in \mathbb{R}^n \text{ and } z^* \in -N_Y(\bar{z}), \\ \emptyset & \text{otherwise.} \end{cases} \tag{3.8}$$

As far as the normal cone to Y is concerned, note that the Mangasarian-Fromowitz constraint qualification (MFCQ) is satisfied at any point of the set. Thus, we can easily check that

$$N_Y(\bar{z}) = \Big\{\alpha\bar{z} + \sum_{i \in I(\bar{z})} \beta_i e^i: \ (\alpha, \beta) \in \mathbb{R} \times \mathbb{R}_+^{|I(\bar{v})|}\Big\} = U(\bar{z}). \tag{3.9}$$

Following the above expression of the coderivative of Y_s, the kernel of the set-valued mapping $D^*Y_s(\bar{x}|(\bar{y}, \bar{z}))$ can be obtained as

$$\begin{aligned} \ker D^*Y_s(\bar{x}|(\bar{y}, \bar{z})) &:= \{(y^*, z^*): \ 0 \in D^*Y_s(\bar{x}|(\bar{y}, \bar{z}))(y^*, z^*)\}, \\ &= \{0\} \times (-N_Y(\bar{z})). \end{aligned}$$

From this formula, we can observe that QC (3.7) is sufficient for the condition

$$D^*S_s((\bar{y}, \bar{v})|\bar{z})(0) \cap \ker D^*Y_s(\bar{x}|(\bar{y}, \bar{v})) = \{0\} \tag{3.10}$$

to hold, for all $(\bar{y}, \bar{v}) \in M(\bar{x}, \bar{z}) := Y_s(\bar{x}) \cap S_s^{-1}(\bar{z})$.

Finally, consider some neighborhoods V and W of $\bar{x}$ and $\bar{v}$, respectively, where V is bounded. Then we have

$$M(V \times W) = Y_s(V) \cap S_s^{-1}(W) = (V \times Y) \cap S_s^{-1}(W) \subseteq V \times Y.$$

Since Y is a bounded set, it follows that M is locally bounded around $(\bar{x}, \bar{v})$. Combining this with the fulfillment of (3.10), it follows from the chain rule by Mordukhovich [35, Theorem 5.1] that we have

$$\begin{aligned} D^*S_{\mathrm{wef}}(\bar{x}|\bar{z})(z^*) &\subseteq \bigcup_{(\bar{y}, \bar{v}) \in M(\bar{x}, \bar{z})} \big[D^*Y_s(\bar{x}|(\bar{y}, \bar{v})) \circ D^*S_s((\bar{y}, \bar{v})|\bar{z})(z^*)\big] \\ &= \bigcup_{(\bar{y}, \bar{v}) \in M(\bar{x}, \bar{z})} \big[D^*Y_s(\bar{x}|(\bar{y}, \bar{v}))(y^*, v^*): \ (y^*, v^*) \in D^*S_s((\bar{y}, \bar{v})|\bar{z})(z^*)\big] \\ &= \bigcup_{\bar{y} \in Y: \ \bar{z} \in S_s(\bar{x}, \bar{y})} \big\{x^*: \ u \in N_Y(\bar{y}), \ (x^*, -u) \in D^*S_s((\bar{x}, \bar{y})|\bar{z})(z^*)\big\} \end{aligned}$$

while taking into account that the set-valued mapping S_s is closed and $(\bar{y}, \bar{v}) \in M(\bar{x}, \bar{z})$ is equivalent to $\bar{x} = \bar{y}$, $\bar{v} \in Y$ and $\bar{z} \in S_s(\bar{y}, \bar{v})$. Hence, we get (3.7) by inserting equation (3.9) in the above estimate of $D^*S_{\mathrm{wef}}(\bar{x}|\bar{z})(z^*)$. $\qquad\square$

Obviously, QC (3.6) automatically holds if S_s is single-valued and Lipschitz continuous around $(\bar{x}, \bar{y})$. Secondly, assume that the lower-level feasible set is defined by

$$K(x) := \{z: \ g_i(x, z) \le 0, \ i = 1, \dots, p\} \tag{3.11}$$

and the MFCQ holds at the point $(\bar{x}, \bar{z})$ for the inequalities defining K, w.r.t. z. Also assume that the mapping S_s is inner semicontinuous at $(\bar{x}, \bar{y}, \bar{z})$, i.e., for any sequence

$(x^k, y^k) \to (\bar{x}, \bar{y})$, there exits a sequence $z^k \in S_s(x^k, y^k)$ such that $z^k \to \bar{z}$. Additionally, suppose that the set-valued mapping

$$\Psi(\vartheta) := \left\{ (x, y, z) : \begin{bmatrix} \langle y, f(x, z) \rangle - \varphi(x, y) \\ g(x, z) \end{bmatrix} + \vartheta \leq 0 \right\}$$

is calm at $(0, \bar{x}, \bar{y}, \bar{z})$, i.e., there exist neighborhoods V of 0, W of $(\bar{x}, \bar{y}, \bar{z})$, and a constant $\kappa > 0$ such that $\Psi(\vartheta) \cap W \subseteq \Psi(0) + \kappa \|\vartheta\| \mathbb{B}$ (unit ball centered at 0) for all $\vartheta \in V$. Then,

$$D^* S_s((\bar{x}, \bar{y})|\bar{z})(z^*)$$
$$\subseteq \left\{ \begin{bmatrix} \sum_{i=1}^p (\beta_i - \lambda \mu_i) \nabla_x g_i(\bar{x}, \bar{z}) \\ 0 \end{bmatrix} : (\beta, \mu, \lambda) \in \Lambda(\bar{x}, \bar{y}, \bar{z}, z^*) \right\}, \quad (3.12)$$

where $\Lambda(\bar{x}, \bar{y}, \bar{z}, z^*)$ denotes the set of all multipliers (β, μ, λ) such that

$$z^* + \lambda \langle \bar{y}, \nabla_z f(\bar{x}, \bar{z}) \rangle + \sum_{i=1}^p \beta_i \nabla_z g_i(\bar{x}, \bar{z}) = 0,$$
$$\lambda \geq 0, \ \beta \geq 0, \ \mu \geq 0, \ \beta^\top g(\bar{x}, \bar{z}) = \mu^\top g(\bar{x}, \bar{z}) = 0,$$

see [13, 37]. It is clear that for any $u \in U(\bar{y})$ such that $(0, u)$ is in the upper bound of $D^* S_s((\bar{x}, \bar{y})|\bar{z})(0)$ from (3.12), we have $u = 0$. This implies that QC (3.6) is automatically satisfied in this case. Note that the addition of equality constraints in (3.11) is possible without difficulties and does not change the aforementioned outcome on QC (3.6).

Recall that a different route is used in [26] to derive estimates of the coderivative of the solution mapping of a parametric multiobjective optimization problem. The results in the latter paper are written in terms of the frontier map, which is the extension of the notion of optimal value function to multiobjective programs. The main advantage of our result is that it is obtained in terms of the coderivative of the solution map of a scalar objective optimization problem, for which detailed estimates in terms of problem data are well-known in the literature, see, e.g., [13, 38]. Using Proposition 3.2, we now derive necessary optimality conditions for a local Pareto optimal solution of problem (3.2).

Theorem 3.3. Let $(\bar{x}, z^o) \in \text{gph} \mathcal{F}$ be a local Pareto optimal solution of problem (3.2), where $l_1 \geq 2$ and $l_2 \geq 2$, and assume that S_s (3.4) is closed and locally bounded around $\bar{x}$, i.e., for a neighborhood V of $\bar{x}$, $S_s(V \times Y) \subseteq W$ with W being a bounded set in $\mathbb{R}^m$. Furthermore, suppose that QC (3.6) holds for all $\bar{y} \in Y$ such that $\bar{z} \in S_s(\bar{x}, \bar{y})$ and for all (y, u, z) with $y \in Y$, $u \in U(y)$, $z \in S_s(\bar{x}, y)$ and such that $z^o = F(\bar{x}, z)$, it holds that

$$\left[(x^*, -u) \in D^* S_s((\bar{x}, y)|z)(0), \ x^* \in -N_X(\bar{x}) \right] \implies x^* = 0. \quad (3.13)$$

Then, there exist $z^* \in \mathbb{R}^{l_2}$ with $\|z^*\| = 1$ and a vector $(\bar{y}, \bar{u}, \bar{z})$ satisfying $\bar{y} \in Y$, $\bar{u} \in U(\bar{y})$, $\bar{z} \in S_s(\bar{x}, \bar{y})$ and $z^o = F(\bar{x}, \bar{z})$, such that we have

$$- (\nabla_x F(\bar{x}, \bar{z})^\top z^*, \bar{u}) \in D^* S_s((\bar{x}, \bar{y})|\bar{z})(\nabla_z F(\bar{x}, \bar{z})^\top z^*) + N_X(\bar{x}) \times \{0_{l_2}\}. \quad (3.14)$$

Proof. To follow the pattern of Theorem 2.3, let us first show that S_{wef} is closed and locally bounded under the assumptions made on S_s. In fact, let (x^k, z^k) be a sequence converging to $(\bar{x}, \bar{z})$ and such that $z^k \in S_{wef}(x^k)$. Then there exists another sequence y^k from Y and such that $z^k \in S_s(x^k, y^k)$. Given that Y is a compact set, then by Bolzano-Weierstrass' theorem, y^k admits a convergent subsequence that we denote similarly, provided there is confusion, and such that $y^k \to \bar{y} \in Y$. Moreover, considering the closedness of S_s, it

follows that $\bar{z} \in S_s(\bar{x}, \bar{y})$. Thus, $\bar{z} \in S_{wef}(\bar{x})$ given that $\bar{y} \in Y$. Obviously, with S_s locally bounded around $\bar{x}$, it follows that S_{wef} is locally bounded around this point.

Proceeding further as in Theorem 2.8, the next thing is show that the counterpart of (2.11) is satisfied. First, note that we have from Proposition 3.2 that

$$
\begin{aligned}
D^*\mathcal{F}(\bar{x}|z^o)(z^*) \quad &\subseteq \quad \Big\{ \nabla_x F(\bar{x}, \bar{z})^\top z^* + D^*S_{wef}(\bar{x}|\bar{z})(\nabla_z F(\bar{x}, \bar{z})^\top z^*) : \\
&\qquad\qquad \bar{z} \in S_{wef}(\bar{x}), \ F(\bar{x}, \bar{z}) = z^o \Big\} \\
&\subseteq \quad \Big\{ \nabla_x F(\bar{x}, \bar{z})^\top z^* + x^* : \ (x^*, -\bar{u}) \in D^*S_s((\bar{x}, \bar{y})|\bar{z})(z^*), \\
&\qquad\qquad \bar{z} \in S_s(\bar{x}, \bar{y}), \ F(\bar{x}, \bar{z}) = z^o, \ \bar{y} \in Y, \ \bar{u} \in U(\bar{y}) \Big\}
\end{aligned}
\tag{3.15}
$$

while taking into account that $S_s(\bar{x}, Y) \cap S_s(\bar{x}, \bar{y}) = S_s(\bar{x}, \bar{y})$. Since, $0 \in U(\bar{y})$, the origin of $\mathbb{R}^n$ is an element of the right-hand-side set that results from the corresponding upper estimate of $D^*\mathcal{F}(\bar{x}|z^o)(0)$ from the latter inclusion. Thus, the fulfillment of QC (3.13), for all (y, u, z) with $y \in Y$, $u \in U(y)$, $z \in S_s(\bar{x}, y)$ and such that $z^o = F(\bar{x}, z)$, is sufficient for the QC $D^*\mathcal{F}(\bar{x}|z^o)(0) \cap (-N_X(\bar{x})) = \{0\}$ to hold. This guaranties that the counterpart of inclusion (2.12) is satisfied for problem (3.2). Proceeding further with the counterpart of (2.12), we can find a vector $x^{*o} \in N_X(\bar{x})$ such that $-x^{*o}$ is an element of the right-hand-side set of the second inclusion in (3.15). Thus, we can subsequently find a vector $(\bar{y}, \bar{u}, \bar{z})$ satisfying $\bar{y} \in Y$, $\bar{u} \in U(y)$, $\bar{z} \in S_s(\bar{x}, y))$ and $z^o = F(\bar{x}, z)$, such that $x^* = -\nabla_x F(\bar{x}, \bar{z})^\top z^* - x^{*o}$ and $(x^*, -\bar{u}) \in D^*S_s((\bar{x}, \bar{y})|\bar{z})(z^*)$. This can easily be rearranged to get the condition in (3.14). $\qquad\qquad\qquad\square$

This theorem provides, for the first time, necessary optimality conditions for bilevel optimization problems involving vector-valued functions both at the upper-and lower-levels. To highlight the relationship with existing results tailored to some special cases of the problem, we apply Theorem 3.3 to the case where the lower-and upper-level feasible sets are defined as (3.11) and $X := \{x : G_j(x) \leq 0, \ j = 1, \dots, k\}$, respectively. The lower-level (resp. upper-level) regularity holds at $(\bar{x}, \bar{z})$ (resp. $\bar{x}$) if the MFCQ holds at this point, for the inequalities defining K w.r.t. z (resp. X). For this result, we assume that the lower-level problem is convex, i.e., the functions f_k for $k = 1, \dots l_2$ and g_i for $i = 1, \dots, p$ are all convex w.r.t. z. For the ease of presentation, we consider the following conditions:

$$
\sum_{i=1}^{p} \beta_i \nabla_z g_i(\bar{x}, \bar{z}) + \sum_{l=1}^{m} \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{zz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^{p} u_i \nabla_{zz_l}^2 g_i(\bar{x}, \bar{z}) \right] = 0, \tag{3.16}
$$

$$
-\sum_{l=1}^{m} \gamma_l \nabla_{z_l} f(\bar{x}, \bar{z}) \in U(\bar{y}), \tag{3.17}
$$

$$
\sum_{k=1}^{l_2} \bar{y}_k \nabla_z f_k(\bar{x}, \bar{z}) + \sum_{i=1}^{p} u_i \nabla_z g_i(\bar{x}, \bar{z}) = 0, \ u \geq 0, \ u^\top g(\bar{x}, \bar{z}) = 0, \tag{3.18}
$$

$$
\nabla_z g_v(\bar{x}, \bar{z})\gamma = 0, \ \beta_\eta = 0, \tag{3.19}
$$

$$
\forall i \in \theta : (\beta_i > 0 \wedge \nabla_z g_i(\bar{x}, \bar{z})\gamma > 0) \vee (\beta_i \nabla_z g_i(\bar{x}, \bar{z})\gamma = 0), \tag{3.20}
$$

$$
\alpha \geq 0, \ \alpha^\top G(\bar{x}) = 0, \tag{3.21}
$$

where the index sets η, θ and ν, respectively given below, define the partition of the complementarity system resulting from the optimality conditions of the lower-level problem:

$$\eta := \eta(\bar{x}, \bar{z}, \bar{u}) := \{i = 1, \dots, p \mid \bar{u}_i = 0, \ g_i(\bar{x}, \bar{z}) < 0\},$$
$$\theta := \theta(\bar{x}, \bar{z}, \bar{u}) := \{i = 1, \dots, p \mid \bar{u}_i = 0, \ g_i(\bar{x}, \bar{z}) = 0\},$$
$$\nu := \nu(\bar{x}, \bar{z}, \bar{u}) := \{i = 1, \dots, p \mid \bar{u}_i > 0, \ g_i(\bar{x}, \bar{z}) = 0\}.$$

Corollary 3.4. Let $(\bar{x}, z^o)$ be a local optimal solution of problem (3.2) and let the upper-level regularity be satisfied at $\bar{x}$ while the set-valued mapping K (3.11) is locally bounded around $\bar{x}$. Furthermore, suppose that the following conditions are satisfied:

(a) For all $(\bar{y}, \bar{z})$ such that $\bar{y} \in Y$ and $\bar{z} \in S_s(\bar{x}, \bar{y})$, the lower-level regularity condition is satisfied at $(\bar{x}, \bar{z})$ and we have $\beta = 0$ and $\gamma = 0$ whenever (3.16)–(3.20) hold.

(b) For all $(\bar{y}, \bar{u}, \bar{z})$ with $\bar{y} \in Y$, $\bar{u} \in U(\bar{y})$, $\bar{z} \in S_s(\bar{x}, \bar{y})$ and such that $z^o = F(\bar{x}, \bar{z})$, we have $\sum_{j=1}^k \alpha_j \nabla G_j(\bar{x}) = 0$ whenever (3.16)–(3.21) hold together with

$$\sum_{j=1}^k \alpha_j \nabla G_j(\bar{x}) + \sum_{i=1}^p \beta_i \nabla_x g_i(\bar{x}, \bar{z})$$
$$+ \sum_{l=1}^m \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{xz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^p u_i \nabla_{xz_l}^2 g_i(\bar{x}, \bar{z}) \right] = 0.$$

Then, there exist (α, β, γ) and $(\bar{u}, \bar{y}, \bar{z}, z^*)$ satisfying $z^* \in \mathbb{R}^{l_2}$, $\|z^*\| = 1$, $\bar{y} \in Y$, $\bar{u} \in U(\bar{y})$, $\bar{z} \in S_s(\bar{x}, \bar{y})$ and $z^o = F(\bar{x}, \bar{z})$, such that (3.16)–(3.21) hold together with the following conditions:

$$\nabla_x F(\bar{x}, \bar{z})^\top z^* + \sum_j^k \alpha_j \nabla G_j(\bar{x}) + \sum_{i=1}^p \beta_i \nabla_x g_i(\bar{x}, \bar{z})$$
$$+ \sum_{l=1}^m \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{xz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^p u_i \nabla_{xz_l}^2 g_i(\bar{x}, \bar{z}) \right] = 0,$$
$$\nabla_z F(\bar{x}, \bar{z})^\top z^* + \sum_{i=1}^p \beta_i \nabla_z g_i(\bar{x}, \bar{z})$$
$$+ \sum_{l=1}^m \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{zz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^p u_i \nabla_{zz_l}^2 g_i(\bar{x}, \bar{z}) \right] = 0.$$

Proof. Under the lower-level convexity, the function $z \to \sum_{k=1}^{l_2} y_k f_k(x, z)$ is convex for all $(x, y) \in \mathbb{R}^n \times \mathbb{R}_+^{l_2}$. Hence, given that $Y \subset \mathbb{R}_+^{l_2}$, it follows that S_s can be written as

$$S_s(x, y) := \left\{ z \in \mathbb{R}^{l_2} : \ 0 \in \sum_{k=1}^{l_2} y_k \nabla_z f_k(x, z) + N_K(x)(z) \right\}.$$

Let $(x^k, y^k, z^k) \in \mathrm{gph}\, S_s$ such that $(x^k, y^k, z^k) \to (\bar{x}, \bar{y}, \bar{z})$. Then under the lower-level regularity, we can find a sequence $u^k \in \Lambda(x^k, y^k, z^k)$, where the set-valued map Λ defines the set of Lagrange multipliers for the parametric problem (3.4):

$$\Lambda(\bar{x}, \bar{y}, \bar{z}) := \left\{ u \in \mathbb{R}^p : \ (3.18) \ \text{is satisfied} \right\}. \tag{3.22}$$

It is well-known that the set-valued map Λ is upper-semicontinuous under the lower-level regularity, see, e.g. [41]. Hence, there exists a convergent subsequence of u^k, for which we preserve the same notation, provided there is no confusion, such that $u^k \to \bar{u}$. Additionally, taking into account the closedness of the map Λ, we have $\bar{u} \in \Lambda(\bar{x}, \bar{y}, \bar{z})$. It then follows from the lower-level convexity that $(\bar{x}, \bar{y}, \bar{z}) \in \mathrm{gph}\, S_s$. This means that S_s is closed. As for the local boundedness of S_s, in the sense of Theorem 3.3, it follows from the local boundedness of the set-valued mapping K.

Under assumption (a), it follows from [38] (also see related discussion in [13]) that

$$D^*S_s((\bar{x}, \bar{y})|\bar{z})(z^*) \subseteq \bigcup_{\bar{u} \in \Lambda(\bar{x}, \bar{y}, \bar{z})} \bigcup_{(\beta, \gamma) \in \Lambda_{\bar{z}}^{em}(\bar{x}, \bar{y}, \bar{z}, \bar{u}, z^*)} \left\{ \begin{bmatrix} H^x(\bar{x}, \bar{y}, \bar{z}, \beta, \gamma) \\ \sum_{l=1}^{m} \gamma_l \nabla_{z_l} f(\bar{x}, \bar{z}) \end{bmatrix} \right\} \qquad (3.23)$$

for all $(\bar{y}, \bar{z})$ such that $\bar{y} \in Y$ and $\bar{z} \in S_s(\bar{x}, \bar{y})$. Recall that the map Λ is defined in (3.22), while $\Lambda_{\bar{z}}^{em}$ is given by

$$\Lambda_{\bar{z}}^{em}(\bar{x}, \bar{y}, \bar{z}, \bar{u}, z^*) := \left\{ (\beta, \gamma) : \nabla_z F(\bar{x}, \bar{z})^\top z^* + H^z(\bar{x}, \bar{y}, \bar{z}, \beta, \gamma) = 0, \right.$$
$$\left. \text{and } (3.19) - (3.20) \text{ are satisfied} \right\}$$

and for $w \in \{x, z\}$, $H^w(\bar{x}, \bar{y}, \bar{z}, \beta, \gamma)$ denotes the following expression:

$$\sum_{i=1}^{p} \beta_i \nabla_w g_i(\bar{x}, \bar{z}) + \sum_{l=1}^{m} \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{wz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^{p} u_i \nabla_{wz_l}^2 g_i(\bar{x}, \bar{z}) \right].$$

Denote by $\Xi(\bar{x}, \bar{y}, \bar{z})$ the upper bound of $D^*S_s((\bar{x}, \bar{y})|\bar{z})(0)$ resulting from (3.23). One can easily check that (3.6) holds if $u = 0$ whenever $(0, u) \in \Xi(\bar{x}, \bar{y}, \bar{z})$ and $u \in U(\bar{y})$. Next, observe that a vector $(0, u) \in \Xi(\bar{x}, \bar{y}, \bar{z})$ if and only if we can find a vector (β, γ) such that (3.16)–(3.20) holds together with $u = -\sum_{l=1}^{m} \gamma_l \nabla_{z_l} f(\bar{x}, \bar{z})$ and

$$\sum_{i=1}^{p} \beta_i \nabla_x g_i(\bar{x}, \bar{z}) + \sum_{l=1}^{m} \gamma_l \left[\sum_{k=1}^{l_2} y_k \nabla_{xz_l}^2 f_k(\bar{x}, \bar{z}) + \sum_{i=1}^{p} u_i \nabla_{xz_l}^2 g_i(\bar{x}, \bar{z}) \right] = 0.$$

In particular, the conditions (3.16)–(3.20) are satisfied. Thus, it follows from assumption (a) in the corollary that $\gamma = 0$. This implies that $u = -\sum_{l=1}^{m} \gamma_l \nabla_{z_l} f(\bar{x}, \bar{z}) = 0$. Hence, (3.6) holds. Proceeding similarly, it can be shown that assumption (b) is a sufficient condition for the fulfilment of QC (3.13). For the latter point, note that under the upper-level regularity at $\bar{x}$, $x^* \in N_X(\bar{x})$ if and only if $x^* = \sum_{j}^{k} \alpha_j \nabla G_j(\bar{x})$ with α satisfying (3.21). Finally, the optimality conditions in the result are obtained by applying Theorem 3.3.   □

To the best of our knowledge, the first result on optimality conditions for semivectorial bilevel programs (i.e., when $l_1 = 1$ and $l_2 \geq 2$) was derived [3]. These conditions, which were abstract were in nature, were obtained via the notion of contingent derivative using a problem reformulation closely related to (3.2). Detailed necessary optimality conditions for the semivectorial bilevel programs were first obtained in [12] for the standard optimistic model (P). The conditions in the latter paper can easily be recovered from Theorem 3.3 while implementing the optimal value function counterpart of Corollary 3.4 on the map S_s (3.4), see e.g., (3.12). It is however important to mention that Corollary 3.4 is the first result providing necessary optimality for the semivectorial bilevel program via the Karush-Kuhn-Tucker (KKT) approach.

If $l_1 \geq 2$ and $l_2 = 1$ in the standard optimistic bilevel program (P), we get the multiobjective bilevel pogram studied in [49]. The M-stationarity conditions obtained in this paper are a special case of the conditions in Corollary 3.4. The slight difference with our result is that the multiplier z^* attached to the upper-level objective function is an element of a special type of normal cone.

As already mentioned above, we are not aware of any result on optimality conditions for a bilevel program with multiobjective upper- and lower-level problems. Corollary 3.4 provides the M-type optimality conditions for this class of problems. Following the pattern of our work in [13], one can easily derive the other stationary concepts, well-known in standard optimistic bilevel program, cf. [19, 17, 18], while using appropriate transformations/reformulations for the set-valued map (3.4).

4. NUMERICAL IMPLEMENTATIONS

This section implements the results of Section 2 on some numerical examples. The main focus is to illustrate Theorem 2.3 while using the following steps to generate stationary points in the sense of (2.8): (1) estimate the Fréchet normal cone to the graph of S; (2) deduce an estimate of the limiting normal cone to the graph of S from (1); (3) deduce an estimate of the coderivative of S from (2); (4) estimate the limiting normal cone to X; and (5) insert the estimates of the coderivative of S and the normal cone to X in (2.8) and solve the resulting system. Note that Example 4.1 is the continuity of the problem discussed in Section 1, while the problems in Examples 4.2 and 4.3 are taken from [2].

Example 4.1 (one dimensional lower-level problem)**.** We continue here with the example in Section 1 (1.11). Clearly, S is closed and bounded given that its graph can be written as $\mathrm{gph}\, S = A \cup B \cup C$ with $A := [-1, 0] \times \{0\}$, $B := [0, 1] \times \{1\}$ and $C := \{0\} \times [0, 1]$. The Fréchet normal cone to $\mathrm{gph}\, S$ is obtained as

$$\widehat{N}_{\mathrm{gph}\, S}(\bar{x}, \bar{y}) = \begin{cases} \{0\} \times \mathbb{R} := \Omega_1 & \text{if } (-1 < \bar{x} < 0,\, \bar{y} = 0) \vee (0 < \bar{x} < 1,\, \bar{y} = 1), \\ \mathbb{R}_+ \times \mathbb{R}_- := \Omega_2 & \text{if } \bar{x} = 0,\, \bar{y} = 0, \\ \mathbb{R}_- \times \mathbb{R} := \Omega_3 & \text{if } \bar{x} = -1,\, \bar{y} = 0, \\ \mathbb{R}_- \times \mathbb{R}_+ := \Omega_4 & \text{if } \bar{x} = 0,\, \bar{y} = 1, \\ \mathbb{R}_+ \times \mathbb{R} := \Omega_5 & \text{if } \bar{x} = 1,\, \bar{y} = 1, \\ \mathbb{R} \times \{0\} := \Omega_6 & \text{if } \bar{x} = 0,\, 0 < \bar{y} < 1, \end{cases}$$

and we deduce the Mordukhovich normal cone while using (2.4):

$$N_{\mathrm{gph}\, S}(\bar{x}, \bar{y}) = \begin{cases} \Omega_1 & \text{if } (-1 < \bar{x} < 0,\, \bar{y} = 0) \vee (0 < \bar{x} < 1,\, \bar{y} = 1), \\ \Omega_2 \cup \Omega_1 \cup \Omega_6 & \text{if } \bar{x} = 0,\, \bar{y} = 0, \\ \Omega_3 \cup \Omega_1 & \text{if } \bar{x} = -1,\, \bar{y} = 0, \\ \Omega_4 \cup \Omega_1 \cup \Omega_6 & \text{if } \bar{x} = 0,\, \bar{y} = 1, \\ \Omega_5 \cup \Omega_1 & \text{if } \bar{x} = 1,\, \bar{y} = 1, \\ \Omega_6 & \text{if } 0 < \bar{x} < 1,\, \bar{y} = 1. \end{cases}$$

Thus, for each $(\bar{x}, \bar{y}) \in \mathbb{R}^2$ with $\bar{y} \in S(\bar{x})$ and each $y^* \in \mathbb{R}$ we get

$$D^*S(\bar{x}|\bar{y})(y^*) = \begin{cases} \{0\} & \text{if } (y^* \in \mathbb{R}, -1 < \bar{x} < 0,\, \bar{y} = 0) \vee (y^* \in \mathbb{R}, 0 < \bar{x} < 1,\, \bar{y} = 1) \\ & \quad \vee (y^* < 0,\, \bar{x} = 0,\, \bar{y} = 0) \vee (y^* > 0,\, \bar{x} = 0,\, \bar{y} = 1), \\ \mathbb{R} & \text{if } y^* = 0, (\bar{x} = 0, \bar{y} = 0) \vee (\bar{x} = 0, \bar{y} = 1) \vee (0 < \bar{x} < 1, \bar{y} = 1), \\ \mathbb{R}_+ & \text{if } (y^* > 0,\, \bar{x} = 0,\, \bar{y} = 0) \vee (y^* \in \mathbb{R},\, \bar{x} = 1,\, \bar{y} = 1), \\ \mathbb{R}_- & \text{if } (y^* \in \mathbb{R},\, \bar{x} = -1,\, \bar{y} = 0) \vee (y^* < 0,\, \bar{x} = 0,\, \bar{y} = 1), \\ \emptyset & \text{if } y^* \neq 0,\, 0 < \bar{x} < 1,\, \bar{y} = 1. \end{cases}$$

For the normal cone to X, we obviously have $N_X(\bar{x}) = \mathbb{R}_-$ for $\bar{x} = -1$, $N_X(\bar{x}) = \mathbb{R}_+$ for $\bar{x} = 1$ and $N_X(\bar{x}) = \{0\}$ otherwise. We can now proceed with the last step by first noting that $\nabla F(x, y) = (2x, 2y)^\top$. Inserting this in (2.8) and solving the resulting system, we get two stationary points $(\bar{x}, \bar{y}) = (0, 0)$ and $(\bar{x}, \bar{y}) = (0, 1)$. Clearly, the solution of the problem obtained in Section 1 is identified. Finally, note that since $N_X(\bar{x}) = \{0\}$ for $\bar{x} = 0$, QC (2.7) automatically holds. $\qquad\qquad\square$

Example 4.2 (two-dimensional lower-level problem with multiple values at a single point). Consider the bilevel program (1.5) with

$$F(x, y) := x + y_2, \quad X := [2, 4], \quad S(x) := \arg\min_y \{2y_1 + xy_2 : y \geq 0, x - y_1 - y_2 \leq -4\}.$$

The following values can be derived for the lower-level solution map S for x from 2 to 4:

$$S(x) = \begin{cases} \{(4 + x, 0)\} & \text{if } 2 < x \leq 4, \\ \{(y_1, 6 - y_1)| \, 0 \leq y_1 \leq 6\} & \text{if } x = 2, \end{cases}$$

and subsequently, we can observe that it is also closed and bounded, considering the fact that $\text{gph } S := A \cup B$ with the sets A and B respectively defined as follows

$$A := \{(x, y_1, y_2) : \, 2 \leq x \leq 4, \, x - y_1 + 4 = 0, \, y_2 = 0\},$$
$$B := \{(x, y_1, y_2) : \, x = 2, \, y_1 + y_2 - 6 = 0, \, 0 \leq y_1 \leq 6\}.$$

Based on this expression, we deduce the following partition of the set $\text{gph } S$:

$$\Delta_1 := \{(x, y_1, y_2) : \, (2 < x < 4) \wedge (y_1 = 4 + x) \wedge (y_2 = 0)\},$$
$$\Delta_2 := \{(x, y_1, y_2) : \, (x = 4) \wedge (y_1 = 4 + \bar{x} = 8) \wedge (y_2 = 0)\},$$
$$\Delta_3 := \{(x, y_1, y_2) : \, (x = 2) \wedge (0 < y_1 < 6) \wedge (y_2 = 6 - y_1)\},$$
$$\Delta_4 := \{(x, y_1, y_2) : \, (x = 2) \wedge (y_1 = 0) \wedge (y_2 = 6 - y_1 = 6)\},$$
$$\Delta_5 := \{(x, y_1, y_2) : \, (x = 2) \wedge (y_1 = 4 + x = 6) \wedge (y_2 = 0)\}.$$

Calculating the normal cones at elements of this partition, we get the following values for the Fréchet normal cone to $\text{gph } S$:

$$\widehat{N}_{\text{gph } S}(\bar{x}, \bar{y}_1, \bar{y}_2) = \begin{cases} \{(x, y_1, y_2) : \, x + y_1 = 0, \, y_2 \in \mathbb{R}\} =: \Omega_1 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1, \\ \{(x, y_1, y_2) : \, x + y_1 \geq 0, \, y_2 \in \mathbb{R}\} =: \Omega_2 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2, \\ \{(x, y_1, y_2) : \, y_1 - y_2 = 0, \, x \in \mathbb{R}\} =: \Omega_3 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3, \\ \{(x, y_1, y_2) : \, y_1 - y_2 \leq 0, \, x \in \mathbb{R}\} =: \Omega_4 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4, \\ \{(x, y_1, y_2) : \, x + y_1 \leq 0, \, y_1 - y_2 \geq 0\} =: \Omega_5 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5. \end{cases}$$

Subsequently, we use the Painlevé-Kuratowski upper limit in (2.4) to obtain the Mordukhovich normal cone

$$N_{\text{gph } S}(\bar{x}, \bar{y}_1, \bar{y}_2) = \begin{cases} \Omega_1 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1, \\ \Omega_2 \cup \Omega_1 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2, \\ \Omega_3 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3, \\ \Omega_4 \cup \Omega_3 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4, \\ \Omega_5 \cup \Omega_1 \cup \Omega_3 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5. \end{cases}$$
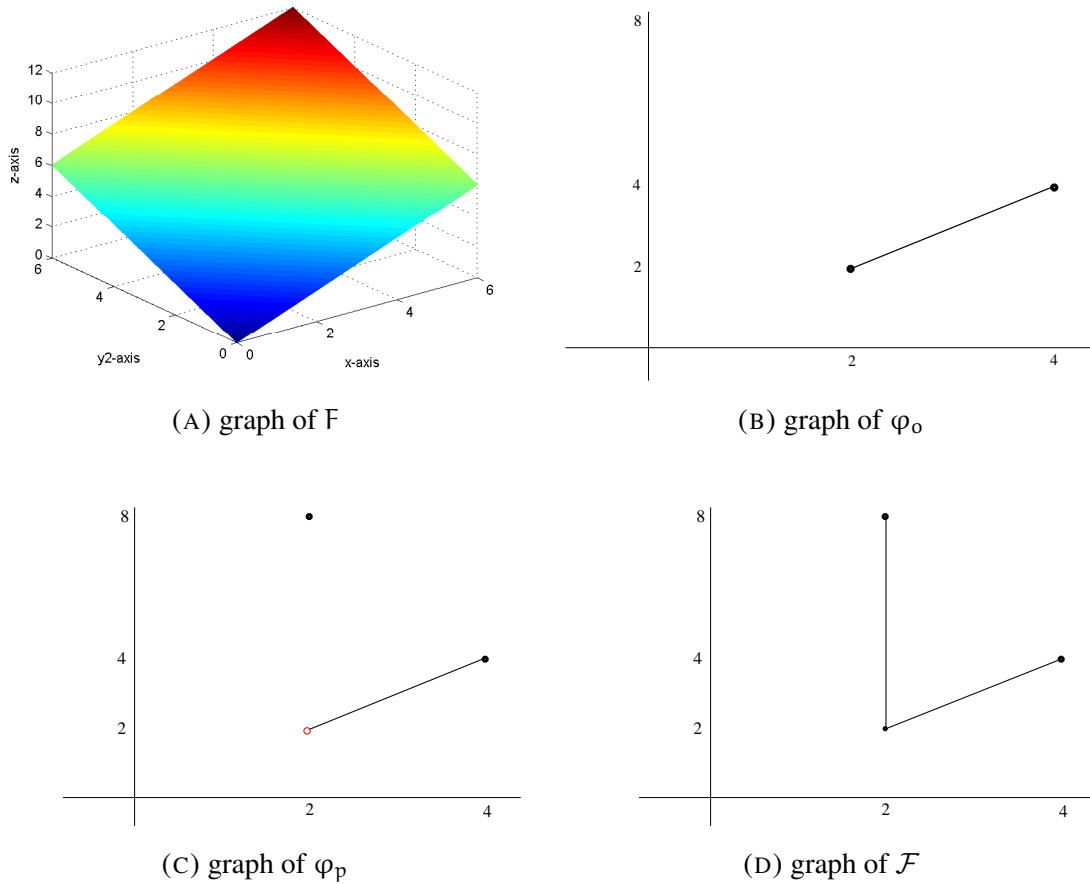
(A) graph of F

(B) graph of φ_o

(C) graph of φ_p

(D) graph of $\mathcal{F}$

FIGURE 2. The graphs of φ_o and φ_p coincide at all but the point $\bar{x} = 2$, where S is multi-valued. At this point, $\mathcal{F}$ is the segment joining $\varphi_o(2)$ and $\varphi_p(2)$. This illustrates Theorem 2.1. Note that the function F is represented only in terms of x and y_2, as it independent of y_1.

By definition of the coderivative (2.6), this leads to the following formula:

$$D^*S(\bar{x}|(\bar{y}_1, \bar{y}_2))(y_1^*, y_2^*) =$$
$$\begin{cases} \{y_1^*\} & \text{if } [(y_1^*, y_2^*) \in \mathbb{R}^2, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1] \vee [(y_1^* > y_2^*), (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5], \\ [y_1^*, \infty[& \text{if } (y_1^*, y_2^*) \in \mathbb{R}^2, \ (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2, \\ \mathbb{R} & \text{if } [y_1^* = y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3 \cup \Delta_5] \vee [(y_1^* \geq y_2^*), (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4], \\ \emptyset & \text{if } [y_1^* \neq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3] \vee [y_1^* < y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4], \\ ]-\infty, y_1^*] & \text{if } y_1^* < y_2^*, \ (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5. \end{cases}$$

Since $\nabla_y F(x, y) = (0, 1)^\top$, we derive from the above expression that

$$D^*S(\bar{x}|(\bar{y}_1, \bar{y}_2))(0, 1) = \begin{cases} \{0\} & \text{if } (2 < \bar{x} < 4) \wedge (\bar{y}_1 = 4 + \bar{x}) \wedge (\bar{y}_2 = 0), \\ \mathbb{R}_+ & \text{if } (\bar{x} = 4) \wedge (\bar{y}_1 = 4 + \bar{x} = 8) \wedge (\bar{y}_2 = 0), \\ \emptyset & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3 \cup \Delta_4, \\ \mathbb{R}_- & \text{if } (\bar{x} = 2) \wedge (\bar{y}_1 = 4 + \bar{x} = 6) \wedge (\bar{y}_2 = 0). \end{cases}$$

Obviously, $N_X(\bar{x}) = \mathbb{R}_-$ for $\bar{x} = 2$, $N_X(\bar{x}) = \mathbb{R}_+$ for $\bar{x} = 4$ and $N_X(\bar{x}) = \{0\}$ otherwise. For the final step, note that $\nabla_x F(x, y) = 1$; hence condition (2.8) leads to the the following family of stationary points:

$$(\bar{x}, \bar{y}_1, \bar{y}_2, \alpha, \beta) \text{ with } (\bar{x}, \bar{y}_1, \bar{y}_2) = (2, 6, 0), \ \alpha + \beta + 1 = 0, \ \alpha, \beta \in \mathbb{R}_-.$$

This suggest that the point $(\bar{x}, \bar{z})$, with $\bar{x} = 2$, $\bar{z} = F(\bar{x}, \bar{y})$ and $\bar{y} = (6, 0)$, is an optimal solution of (P_s). To confirm this, note that $\mathcal{F}(2) = \mathcal{F}(X) = [2, 8]$ and

$$\mathcal{F}(X) - \bar{z} = \begin{cases} [0, 6] & \text{if } \bar{z} = 2 = F(2, \bar{y}) \text{ with } \bar{y} = (6, 0), \\ [2 - \bar{z}, 8 - \bar{z}] & \text{if } \bar{z} \in \,]2, 8]. \end{cases}$$

Clearly, $2 - \bar{z} < 0$ for all $\bar{z} \in \,]2, 8]$. Hence, $(2, 2)$ is the only optimal solution of (P_s) in the sense of (1.12). Thus, from Theorem 2.2, $\bar{x} = 2$ is an optimal solution of (P_o). It is also obvious from Figure 2b that $\bar{x} = 2$ is the unique optimal solution of the latter problem.

Finally, let us check what happens with QC (2.7). First note that

$$D^*S(2|(\bar{y}_1, \bar{y}_2))(0, 0) = \mathbb{R} \text{ for all } (y_1, y_2) \in S(2).$$

It follows that $D^*S(2|(y_1, y_2))(0, 0) \cap (-N_X(2)) = \mathbb{R}_+$ for all $(y_1, y_2) \in S(2)$. This means that condition (2.7) fails at the optimal solution. $\square$

Example 4.3 (two-dimensional lower-level problem with multiple values at all but a single point)**.** Consider the problem (1.5) with $F(x, y) := -x + 10y_1 - y_2$, $X := [0, \infty[$ and

$$S(x) := \arg\min_y \{-y_1 - y_2 : \ y \geq 0, \ x - y_1 \leq 1, \ x + y_2 \leq 1, \ y_1 + y_2 \leq 1\}.$$

Simple calculations lead to the following values of S for $x \geq 0$:

$$S(x) = \begin{cases} \{(y_1, 1 - y_1) : \ x \leq y_1 \leq 1\} & \text{if } 0 \leq x < 1, \\ \{(1, 0)\} & \text{if } x = 1, \\ \emptyset & \text{if } x > 1. \end{cases}$$

Obviously, $\text{gph } S = \{(x, y_1, y_2) : \ 0 \leq x \leq 1, \ x \leq y_1 \leq 1, \ y_2 = 1 - y_1\}$. To proceed with the expression of the Fréchet normal cone to the graph of S, we partition the set as follows:

$$\begin{aligned} \Delta_1 &:= \{(x, y_1, y_2) : (\bar{x} = 0) \wedge (\bar{x} < \bar{y}_1 < 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_2 &:= \{(x, y_1, y_2) : (\bar{x} = 0) \wedge (\bar{x} = \bar{y}_1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_3 &:= \{(x, y_1, y_2) : (\bar{x} = 0) \wedge (\bar{y}_1 = 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_4 &:= \{(x, y_1, y_2) : (\bar{x} = 1) \wedge (\bar{y}_1 = 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_5 &:= \{(x, y_1, y_2) : (0 < \bar{x} < 1) \wedge (\bar{x} < \bar{y}_1 < 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_6 &:= \{(x, y_1, y_2) : (0 < \bar{x} < 1) \wedge (\bar{x} = \bar{y}_1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}, \\ \Delta_7 &:= \{(x, y_1, y_2) : (0 < \bar{x} < 1) \wedge (\bar{y}_1 = 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1)\}. \end{aligned}$$

Proceeding as in the previous example, the Fréchet normal cone to the graph of S is given
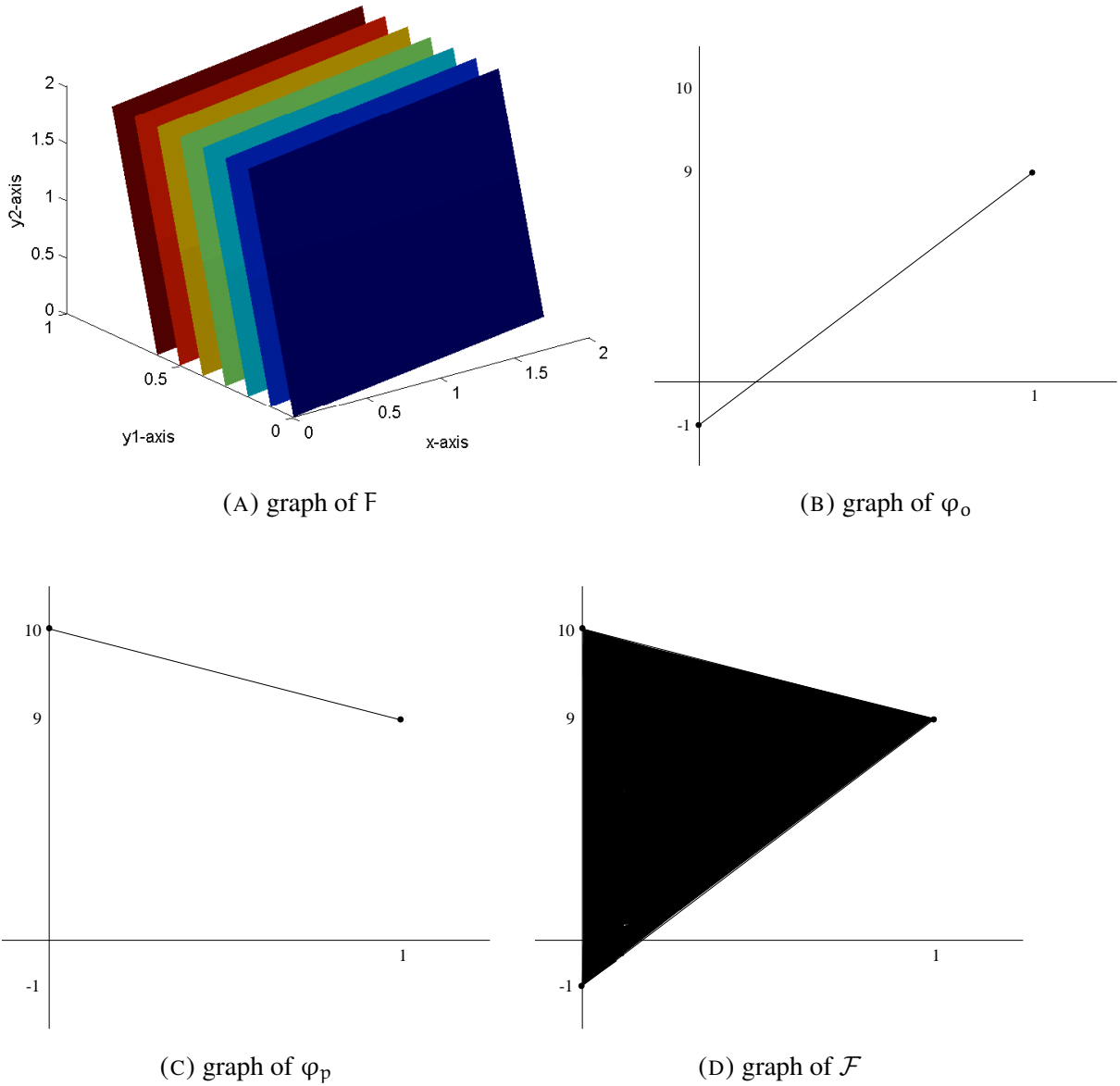
(A) graph of F

(B) graph of φ_o

(C) graph of φ_p

(D) graph of $\mathcal{F}$

FIGURE 3. (A) gives the isosurfaces of the upper-level objective function for the values $z_i = i = F(x, y), i = 0, 1, \ldots, 6$. For (D), note that as S is single-valued at a single point, $\bar{x} = 1$, φ_o and φ_p are distinct from each other at all the other points. Thus the graph of $\mathcal{F}$ is a full triangle given that $\mathcal{F}(x) = [\varphi_o(x), \varphi_p(x)]$ for all $x \in [0, 1[$, cf. Theorem 2.1.

by the following sequence of sets Ω_i, $i = 1, \ldots, 7$, at the corresponding points:

$$\Omega_1 := \{(x, y_1, y_2) : x \leq 0, \ y_1 - y_2 = 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1,$$
$$\Omega_2 := \{(x, y_1, y_2) : y_1 - y_2 \leq 0, \ x + y_1 - y_2 \leq 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2,$$
$$\Omega_3 := \{(x, y_1, y_2) : x \leq 0, \ y_1 - y_2 \geq 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3,$$
$$\Omega_4 := \bigcup_{t \geq 0} \{(x, y_1, y_2) : t - x \leq 0, \ t - x - y_1 + y_2 \leq 0\} \quad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4,$$
$$\Omega_5 := \{(x, y_1, y_2) : \ x = 0, \ y_1 - y_2 = 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5,$$
$$\Omega_6 := \{(x, y_1, y_2) : \ x \geq 0, \ x + y_1 - y_2 = 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_6,$$
$$\Omega_7 := \{(x, y_1, y_2) : \ x = 0, \ y_1 - y_2 \geq 0\} \qquad \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_7.$$

Subsequently, by also applying the Painlevé-Kuratowski upper limit, we arrive at

$$
N_{gph\,S}(\bar{x}, \bar{y}_1, \bar{y}_2) = \begin{cases}
\Omega_1 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1, \\
\Omega_2 \cup \Omega_1 \cup \Omega_5 \cup \Omega_6 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2, \\
\Omega_3 \cup \Omega_1 \cup \Omega_7 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3, \\
\Omega_4 \cup \Omega_5 \cup \Omega_6 \cup \Omega_7 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4, \\
\Omega_5 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5, \\
\Omega_6 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_6, \\
\Omega_7 & \text{if } (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_7.
\end{cases}
$$

By definition of the coderivative (2.6), this leads to the following equality

$$
D^*S(\bar{x}|(\bar{y}_1, \bar{y}_2))(y_1^*, y_2^*) =
\begin{cases}
]-\infty, y_1^* - y_2^*] & \text{if } y_1^* \geq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_2, \\
[y_1^* - y_2^*, \infty[& \text{if } y_1^* > y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4, \\
\{y_1^* - y_2^*\} & \text{if } y_1^* \geq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_6, \\
\mathbb{R}_- & \text{if } [y_1^* = y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_1] \vee [y_1^* \leq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_3], \\
\mathbb{R}_+ & \text{if } y_1^* \leq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_4, \\
\{0\} & \text{if } [y_1^* = y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_5] \vee [y_1^* \leq y_2^*, (\bar{x}, \bar{y}_1, \bar{y}_2) \in \Delta_7], \\
\emptyset & \text{otherwise.}
\end{cases}
$$

We are now ready to solve the optimality condition in (2.8). To proceed, note that we have $\nabla_x F(x, y) = -1$ and $\nabla_y F(x, y) = (10, -1)^\top$ for all (x, y). Hence, we get

$$
D^*S(\bar{x}|(\bar{y}_1, \bar{y}_2))(10, -1) = \begin{cases}
]-\infty, 11] & \text{if } (\bar{x} = 0) \wedge (\bar{x} = \bar{y}_1) \wedge (\bar{y}_2 = 1 - \bar{y}_1), \\
[11, \infty[& \text{if } (\bar{x} = 1) \wedge (\bar{y}_1 = 1) \wedge (\bar{y}_2 = 1 - \bar{y}_1), \\
\{11\} & \text{if } (0 < \bar{x} < 1) \wedge (\bar{x} = \bar{y}_1) \wedge (\bar{y}_2 = 1 - \bar{y}_1), \\
\emptyset & \text{otherwise.}
\end{cases}
$$

For the upper-level feasible set, we have $N_X(\bar{x}) = \mathbb{R}_-$ if $\bar{x} = 0$ and $N_X(\bar{x}) = \{0\}$ otherwise. Inserting these values in (2.8), we get the family of stationary points

$$
(\bar{x}, \bar{y}_1, \bar{y}_2, \alpha, \beta) \text{ with } \bar{x} = \bar{y}_1 = 0, \ \bar{y}_2 = 1, \ \alpha + \beta = 1, \ \alpha \leq 11, \ \beta \leq 0.
$$

To confirm that $\bar{x} = 0$ is an optimal solution of (P_o) by means of Theorem 2.2, note that $\mathcal{F}(0) = [-1, 10]$, $\mathcal{F}(X) = \{-x + 11y_1 - 1 : x \in [0, 1], y_1 \in [x, 1]\}$ and

$$
\mathcal{F}(X) - \bar{z} = \begin{cases}
\{-x + 11y_1 : x \in [0, 1], y_1 \in [x, 1]\} & \text{if } \bar{z} = -1 = F(0, \bar{y}), \ \bar{y} = (0, 1), \\
\{-x + 11y_1 - 1 - \bar{z} : x \in [0, 1], y_1 \in [x, 1]\} & \text{if } \bar{z} \in]-1, 10].
\end{cases}
$$

It then follows that $\mathcal{F}(X) - \bar{z} \subseteq [0, \infty[$ for $\bar{z} = -1$ and $-1 - \bar{z} \in (\mathcal{F}(X) - \bar{z}) \cap]-\infty, 0[$ for $\bar{z} \in]-1, 10]$. Hence, $(0, -1)$ is the only optimal solution of (P_s). Thus, $\bar{x} = 0$ is an optimal solution of (P_o) by Theorem 2.2. It is also clear from Figure 3b that $\bar{x} = 0$ is the unique optimal solution for (P_o). Furthermore, Figure 3c shows that (P_p) also has a solution in this case, $\bar{x} = 1$, which is completely different from the one of (P_o).

Finally, note that $D^*S(0|(\bar{y}_1, \bar{y}_2))(0, 0) = \mathbb{R}_-$ for all $(y_1, y_2) \in S(0)$ and $-N_X(0) = \mathbb{R}_+$; thus condition (2.7) is satisfied. $\qquad\square$

5. Discussion

The aim of the paper was to provide a first step towards solving the original optimistic bilevel program (P_o) using a set-valued optimization technique, which is actually an extension of the implicit function approach (1.1). As such, ideas tailored to (P_i) could be extended to (P_s). To proceed, the first step would be to further investigate the structure of the objective map $\mathcal{F}$ (1.10). It is important to note that this mapping is Lipschitz-like, in the sense of Aubin [1], if S is Lipschitz-like, cf. Proof of Theorem 2.3. Also note that in the multiobjective setting, the Lipschitz-like property of S_{wef} (3.1) can easily be deduced from that of S_s (3.4), see Proposition 3.2 and the discussion that follows it. Most methods for (P_i) are based on estimates of the subdifferentials of $y(.)$ which are sequentially inserted in the upper-level objective function F and the resulting subproblem is solved by an appropriate optimization technique. For instance, bundle methods are used in [40] for such a purpose. See also [7] for the descent method and other closely related approaches. In the context of (P_s), it is our opinion that such a process can be developed while replacing the estimates of the subdifferential of $y(.)$ by the estimates of the coderivative of the lower-level solution set-valued mapping S (1.4); cf. the discussion in Subsection 2.2 and related references on coderivatives of the solution maps of parametric optimization problems.

An alternative approach towards a formal algorithm to solve (P_s) is the one implemented in the examples of the previous section. This consist of solving the stationary conditions developed in Sections 2 and 3. It is however important to mention here that methods to solve set-valued mapping inclusions (see e.g., [20] and references therein) cannot be easily implemented directly on (2.8), considering the nature of S, as one would need to compute the second order coderivative of S. This implies evaluating the normal cone to the graph of the normal cone map to the graph of S, i.e. $N_{gph\,N_{gph\,S}}$. Since S is not normally regular in the sense of [36], computing an estimate for the latter normal cone would be quite complicate. Clearly, using this direction would imply inserting an estimate of the coderivative of S in (2.8), as we did for the examples above, and then solving the resulting conditions by a standard-type optimization technique. An example of method for the subproblems here would be the Newton method. Finally, it is worth mentioning that following the pattern of developments in Section 3, the possible methods mentioned above for the scalar objective bilevel optimization problem could easily be extended to the multiobjective framework.

References

[1] J.-P. Aubin. Lipschitz behavior of solutions to convex minimization problems. *Math. Oper. Res.*, 9(1):87–111, 1984.

[2] J.F. Bard. *Practical bilevel optimization: Algorithms and applications*. Dordrecht: Kluwer Academic Publishers, 1998.

[3] H. Bonnel. Optimality conditions for the semivectorial bilevel optimization problem. *Pac. J. Optim.*, 2(3):447–467, 2006.

[4] J. Bracken and J.T. McGill. Mathematical programs with optimization problems in the constraints. *Oper. Res.*, 21:37–44, 1973.

[5] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Ann. Oper. Res.*, 153:235–256, 2007.

[6] S. Dempe. A bundle algorithm applied to bilevel programming problems with non-unique lower level solutions. *Comput. Optim. Appl.*, 15(2):145–166, 2000.

[7] S. Dempe. *Foundations of bilevel programming*. Dordrecht: Kluwer Academic Publishers, 2002.

[8] S. Dempe. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization*, 52(3):333–359, 2003.

[9] S. Dempe and J.F. Bard. Bundle trust-region algorithm for bilinear bilevel programming. *J. Optim. Theory Appl.*, 110(2):265–288, 2001.

[10] S. Dempe, J. Dutta, and B.S. Mordukhovich. Variational analysis in bilevel programming. In S.K Neogy, R.B. Bapat, A.K Das, and T. Parthasarathy, editors, *Mathematical programming and game theory for decision making*, pages 257–277. World Scientific, 2008.

[11] S. Dempe and S. Franke. Solution algorithm for an optimistic linear stackelberg problem. *Comput. Oper. Res.*, 41:277 – 281, 2014.

[12] 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.

[13] S. Dempe, B. Mordukhovich, and A.B. Zemkoho. Sensitivity analysis for two-level value functions with applications to bilevel programming. *SIAM J. Optim.*, 22(4):1309–1343, 2012.

[14] S. Dempe, B.S. Mordukhovich, and A.B. Zemkoho. Necessary optimality conditions in pessimistic bilevel programming. *Optimization*, 63(4):505–533, 2014.

[15] S. Dempe and H. Schmidt. On an algorithm solving two-level programming problems with nonunique lower level solutions. *Comput. Optim. Appl.*, 6(3):227–249, 1996.

[16] S. Dempe and A.B. Zemkoho. The generalized Mangasarian-Fromowitz constraint qualification and optimality conditions for bilevel programs. *J. Optim. Theory Appl.*, 148:433–441, 2011.

[17] S. Dempe and A.B. Zemkoho. On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem. *Nonlinear Anal.*, 75(3):1202–1218, 2012.

[18] S. Dempe and A.B. Zemkoho. The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. *Math. Program.*, 138(1-2):447–473, 2013.

[19] S. Dempe and A.B. Zemkoho. KKT reformulation and necessary conditions for optimality in nonsmooth bilevel optimization. *SIAM J. Optim.*, 24(4):1639–1669, 2014.

[20] S. Dias and G. Smirnov. On the newton method for set-valued maps. *Nonlinear Anal.*, 75(3):1219 – 1230, 2012.

[21] J. Dutta and S. Dempe. Bilevel programming with convex lower level problems. In S. Dempe and V. Kalashnikov, editors, *Optimization with multivalued mappings*, pages 51–71. New York: Springer, 2006.

[22] J. Fliege and L.N. Vicente. Multicriteria approach to bilevel optimization. *J. Optim. Theory Appl.*, 131(2):209–225, 2006.

[23] T.X.D. Ha. Optimality conditions for various efficient solutions involving coderivatives: From set-valued optimization problems to set-valued equilibrium problems. *Nonlinear Anal.*, 75:1305–1323, 2012.

[24] K. Hatz, S. Leyffer, J.P. Schlöder, and H.G. Bock. Regularizing bilevel nonlinear programs by lifting. Preprint ANL/MCS-P4076-0613, 2013.

[25] X. He, A. Prasad, S.P. Sethi, and G.J. Gutierrez. A survey of stackelberg differential game models in supply and marketing channels. *J. Sys. Sci. Sys. Eng.*, 16(4):385–413, 2007.

[26] N.Q. Huy, B.S. Mordukhovich, and J.C. Yao. Coderivatives of frontier and solution maps in parametric multiobjective optimization. *Taiwanese J. Math.*, 12(8):2083–2111, 2008.

[27] J. Jahn. *Vector optimization. Theory, applications, and extensions. 2nd ed.* Berlin: Springer, 2011.

[28] G.-H. Lin, M. Xu, and J.J. Ye. On solving simple bilevel programs with a nonconvex lower level program. *Math. Program.*, 144(1-2):277–305, 2014.

[29] P. Loridan and J. Morgan. Weak via strong Stackelberg problem: New results. *J. Glob. Optim.*, 8(3):263–287, 1996.

[30] R. Lucchetti, F. Mignanego, and G. Pieri. Existence theorems of equilibrium points in Stackelberg games with contraints. *Optimization*, 18(6):857–866, 1987.

[31] A.G. Mersha and S. Dempe. Direct search algorithm for bilevel programming problems. *Comput. Optim. Appl.*, 49(1):1–15, 2011.

[32] A.G. Mersha and S. Dempe. Feasible direction method for bilevel programming problem. *Optimization*, 61(5):597–616, 2012.

[33] D.A. Molodtsov. The solution of a class of non antagonistic games. *USSR Compt. Maths. Math. Phys.*, 16:1451–1456, 1976.

[34] B.S. Mordukhovich. Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems. *Sov. Math., Dokl.*, 22:526–530, 1980.

[35] B.S. Mordukhovich. Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.*, 183(1):250–288, 1994.

[36] B.S. Mordukhovich. *Variational analysis and generalized differentiation. I: Basic theory. II: Applications*. Berlin: Springer, 2006.

[37] B.S. Mordukhovich, M.N. Nam, and H.M. Phan. Variational analysis of marginal function with applications to bilevel programming problems. *J. Optim. Theory Appl.*, 152(3):557–586, 2012.

[38] B.S. Mordukhovich and J.V. Outrata. Coderivative analysis of quasi-variational inequalities with applications to stability and optimization. *SIAM J. Optim.*, 18(2):389–412, 2007.

[39] J. Morgan and F. Patrone. Stackelberg problems: Subgame perfect equilibria via tikhonov regularization. In A. Haurie, S. Muto, L.A. Petrosjan, and T.E.S. Raghavan, editors, *Advances in Dynamic Games*, pages 209–221. Boston: Birkhäuser, 2006.

[40] J.V. Outrata, M. Kocvara, and J. Zowe. *Nonsmooth approach to optimization problems with equilibrium constraints*. Dordrecht: Kluwer Academic Publishers, 1998.

[41] S.M. Robinson. Generalized equations and their solutions. I: Basic theory. *Math. Program. Study*, 10:128–141, 1979.

[42] S. Ruuska, K. Miettinen, and M.M. Wiecek. Connections between single-level and bilevel multiobjective optimization. *J. Optim. Theory Appl.*, 153(1):60–74, 2012.

[43] M. Červinka, C. Matonoha, and J. V. Outrata. On the computation of relaxed pessimistic solutions to mpecs. *Optim. Methods Softw.*, 28(1):186–206, 2013.

[44] H.F. von Stackelberg. *Marktform und Gleichgewicht*. Berlin: Springer-Verlag, 1934.

[45] H.F. von Stackelberg. *Market Structure and Equilibrium: 1st Edition Translation into English by D. Bazin, R. Hill and L. Urch*. Heidelberg and New York: Springer, 2011.

[46] W. Wiesemann, A. Tsoukalas, P. Kleniati, and B. Rustem. Pessimistic bilevel optimization. *SIAM J. Optim.*, 23(1):353–380, 2013.

[47] M. Xu and J.J. Ye. A smoothing augmented lagrangian method for solving simple bilevel programs. *Comput. Optim. Appl.*, 59(1-2):353–377, 2014.

[48] J.J. Ye. Constraint qualifications and KKT conditions for bilevel programming problems. *Math. Oper. Res.*, 31(4):811–824, 2006.

[49] J.J. Ye. Necessary optimality conditions for multiobjective bilevel programs. *Math. Oper. Res.*, 36(1):165–184, 2011.

[50] J.J. Ye and D.L. Zhu. Optimality conditions for bilevel programming problems. *Optimization*, 33(1):9–27, 1995. (with Erratum in *Optimization*, 39(4):361-366, 1997).

[51] J.J. Ye and D.L. Zhu. New necessary optimality conditions for bilevel programs by combining MPEC and the value function approach. *SIAM J. Optim.*, 20(4):1885–1905, 2010.

[52] A.B. Zemkoho. *Bilevel programming: reformulations, regularity, and stationarity*. PhD thesis, Department of Mathematics and Computer Science, TU Bergakademie Freiberg, Freiberg, Germany, 2012.

SCHOOL OF MATHEMATICS, UNIVERSITY OF SOUTHAMPTON, SO17 1BJ SOUTHAMPTON, UK

E-mail address: `a.b.zemkoho@soton.ac.uk`