Optimization problems with value function objectives

ALAIN B. ZEMKOHO*

Abstract. The family of optimization problems with value function objectives includes the minmax programming problem and the bilevel optimization problem. In this paper, we derive necessary optimality conditions for this class of problems. The main focus is on the case where the functions involved are nonsmooth and the constraints are the very general operator constraints.

Keywords. minmax programming, optimistic an pessimistic bilevel programming, value functions, variational analysis, optimality conditions

AMS subject classifications. 49K35, 49K40, 90C26, 90C30

1. Introduction. An optimization problem with value function objective is a problem of the form

$$\text{minimize } \psi(x) \text{ subject to } x \in X \tag{1.1}$$

where $X \subseteq \mathbb{R}^n$ denotes the feasible set and ψ is the optimal value function of the parametric optimization problem to

$$\underset{y}{\text{minimize}} \ F(x,y) \text{ subject to } y \in K(x) \tag{1.2}$$

with K denoting a certain set-valued mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ while F is a real-valued function on $\mathbb{R}^n \times \mathbb{R}^m$. If we replace the minimization in the latter problem by a maximization, problem (1.1) becomes the well-known minmax optimization problem, see e.g. [21]. If instead, one sets K to represent the solution/argminimum set-valued mapping of another optimization problem, then (1.1) results to the optimistic formulation of the bilevel optimization problem [2]:

$$\left\{\begin{array}{ll} \text{minimize} & \varphi_o(x) \text{ subject to } x \in X \\ \text{where} & \varphi_o(x) := \min\limits_y \{F(x,y)|\, y \in S(x)\} \\ & S(x) := \{y \in K(x)|\, f(x,y) \leq \varphi(x)\} \\ & \varphi(x) := \min\limits_y \{f(x,y)|\, y \in K(x)\}. \end{array}\right. \tag{1.3}$$

It should however be mentioned that originally, the bilevel optimization problem consist to

$$\text{"}\underset{x}{\text{minimize}}\text{"} \ F(x,y) \text{ subject to } x \in X,\, y \in S(x). \tag{1.4}$$

But with the ambiguity (marked by the brackets) that occur in handling this problem when S is not a single-valued mapping, the concepts of optimistic and pessimistic reformulation have been considered in the literature, see [2] for details on these reformulations. The optimistic reformulation is given in (1.3) whereas the pessimistic one

*Institut für Numerische Mathematik und Optimierung, Technische Universität Bergakademie Freiberg (zemkoho@daad-alumni.de). Financial support by the Deutscher Akademischer Austauscht Dienst (DAAD) is acknowledged.

is obtained by replacing the value function φ_o in the latter problem by the maximization value function

$$\varphi_p(x) := \max_y \{F(x,y)|\, y \in S(x)\}. \tag{1.5}$$

As far as the optimistic formulation of the bilevel optimization problem is concerned, the model mostly investigated is the following one:

$$\underset{x,y}{\text{minimize}}\ F(x,y) \text{ subject to } x \in X,\ y \in S(x). \tag{1.6}$$

This essentially consist of shifting the difficulty, which lies in the inclusion $y \in S(x)$, to the constraint, whereas it is situated in the objective function when dealing with the initial formulation (1.3).

Unfortunately, a local optimal solution of problem (1.6) may not correspond to a local optimal solution of (1.3), cf. [2, 4]. This is a major drawback when one intends to solve the original optimistic problem (1.3) by solving the auxiliary problem (1.6) given that these problems are both nonconvex.

In [9], the static minmax optimization problem

$$\min_{x\in X} \max_{y\in Y} F(x,y)$$

was first transformed into a semi-infinite programming problem

$$\begin{aligned} \text{minimize } z \quad & \text{subject to } x \in X \\ & \text{and } F(x,y) \le z,\ \forall y \in Y \end{aligned}$$

and then latter converted to an optimization problem with finitely many constraints, in order to investigate necessary optimality conditions. Such a process is also made to push the difficulty in a minmax program to the constraints.

Recently though, the minmax program, the optimistic and pessimistic bilevel programs have been considered respectively in [24], [8] and [7], as they are, that is optimization problems with value functions objectives. For the minmax program, this is not new as demonstrated for example by the approach in [21, Chapter 9]. More precisely, in [24, 8, 7], the generalized differentiation tools of Mordukhovich [13] were used to derive necessary optimality conditions for the aforementioned problems. A great literature exist on the bilevel optimization problem in its classical/auxiliary form (1.6), see e.g. [2, 4, 5, 6, 16, 22] and references therein for recent results. As for the minmax optimization problem, a number of results and references can be found in [21, 9, 24].

In the line of the works in [24] and [8, 7] respectively, we also treat the minmax, optimistic and pessimistic bilevel programs as optimization problems with value function objectives. Though acknowledging the fact that (1.3) is a special case of the simplest albeit general optimization problem with value function objective (1.1), it seems more reasonable to focus our attention to the former problem, considering the much more complicated structure of its objective function. Moreover, problem (1.3) appears to provide for a unified framework to derive necessary optimality conditions for the minmin, minmax, optimistic and pessimistic bilevel optimization problems. To see this, first note that if one sets the lower level objective function f in (1.3) to be a given constant everywhere, then the lower level value function φ would also be the same constant. Hence, (1.3) reduces to a problem of minimizing a classical/simple optimal value function. A similar observation can be made for the minmax problem as contained in the pessimistic bilevel program.

As for the pessimistic bilevel program itself, let us recall the following obvious interplay

$$\varphi_p(x) = -\varphi_p^o(x), \text{ for all } x \in X \tag{1.7}$$

between the maximization and minimization value functions, with φ_p given in (1.5) while φ_p^o is defined as

$$\varphi_p^o(x) := \min_y \{-F(x,y)|\, y \in S(x)\}. \tag{1.8}$$

Since we will be using the generalized differentiation tools of Mordukhovich [13] to analyze problem (1.3), equality (1.7) implies that the basic subdifferential of φ_p can be obtained from that of φ_p^o by means of the following convex hull property

$$\mathrm{co}\,\partial(-\psi)(\overline{x}) = -\mathrm{co}\,\partial\psi(\overline{x})$$

provided ψ is a given function which is Lipschitz continuous near $\overline{x}$. Here "co" stands for the convex hull of the set in question whereas $\partial\psi$ denotes the basic/Mordukhovich subdifferential that will be defined in the next section.

For these reasons, our first focus in this paper will be the following optimistic formulation of the bilevel optimization problem, where the upper- and lower-level problems are constrained by the very general operator-type constraints:

$$(P_o) \quad \left\{ \begin{array}{ll} \text{minimize} & \varphi_o(x) \text{ subject to } x \in X \\ \text{where} & X := \Omega_1 \cap \psi_1^{-1}(\Lambda_1) \\ & \varphi_o(x) := \min_y \{F(x,y)|\, y \in S(x)\} \\ & S(x) := \{y \in K(x)|\, f(x,y) \leq \varphi(x)\} \\ & \varphi(x) := \min_y \{f(x,y)|\, y \in K(x)\} \\ & K(x) := \{y|\, (x,y) \in \Omega_2 \cap \psi_2^{-1}(\Lambda_2)\} \end{array} \right.$$

where $\Omega_1 \subseteq \mathbb{R}^n$, $\Lambda_1 \subseteq \mathbb{R}^k$, $\Omega_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$ and $\Lambda_1 \subseteq \mathbb{R}^p$ are closed sets. The upper and lower level objective functions $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ whereas $\psi_2 : \mathbb{R}^n \to \mathbb{R}^p$ and $\psi_1 : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$. All these functions are assumed to be continuous throughout the paper. The upper- and lower-level constraints are determined by the inclusions $x \in \Omega_1 \cap \psi_1^{-1}(\Lambda_1)$ and $(x,y) \in \Omega_2 \cap \psi_2^{-1}(\Lambda_2)$, respectively. Such constraints structures, called operator constraints in [13], are in fact very general. One can easily check that they contain most of the constraint structures usually considered in optimization. The special value function

$$\varphi_o(x) := \min_y \{F(x,y)|\, y \in S(x)\}, \tag{1.9}$$

is called a two-level value function considering the nature of the constraint set $S(x)$ which is the solution set of a second optimization problem, namely the lower level problem

$$\min_y \{f(x,y)|\, y \in K(x)\}.$$

In the next section, we present the tools from variational analysis that will be used in this paper. We essentially introduce the Mordukhovich normal cone, subdifferential and coderivative and some of the related properties that are needed in the subsequent sections. In Section 3, we study the variational properties (that is, the sensitivity analysis) of the two-level value function φ_o (1.9) and deduce necessary optimality conditions of the optimistic and minmin optimization problems. The constraint qualification (CQ) mostly used here is the fairly weak calmness condition related to set-valued mappings since most of the well-known CQs can not be satisfied for the parametric problem associated to the two-level value function (1.9). Necessary optimality conditions for the minmax and pessimistic optimization problems are investigated in Section 4. It should be mentioned that our main focus is on the case where the functions involved in the aforementioned problems are nonsmooth. In [8, 7] the functions are assumed to be smooth, hence the generalized equation and Karush-Kuhn-Tucker reformulations are also investigated whereas here we only consider the (lower-level) optimal value reformulation marked by the presence of φ in (P_o). Furthermore, the constraints in [8, 7] are simple inequality constraints. Clearly, this implies that the work in this paper completes those of [8, 7, 24].

2. Tools from variational analysis. The material presented here is essentially taken from [13]. Also see [20]. We start with the Kuratowski-Painlevée outer/upper limit of a set-valued mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, which is defined at a point $\overline{x}$ as

$$\operatorname*{Limsup}_{x \to \overline{x}} \Psi(x) := \{v \in \mathbb{R}^m | \exists x_k \to \overline{x},\ v_k \to v \text{ with } v_k \in \Psi(x_k) \text{ as } k \to \infty\}. \tag{2.1}$$

For an extended real-valued function $\psi : \mathbb{R}^n \to \overline{\mathbb{R}} := (-\infty, \infty]$, the Fréchet subdifferential of ψ at $\overline{x} \in \operatorname{dom} \psi := \{x \in \mathbb{R}^n |\ \psi(x) < \infty\}$ is given by

$$\widehat{\partial}\psi(\overline{x}) := \left\{ v \in \mathbb{R}^n | \liminf_{x \to \overline{x}} \frac{\psi(x) - \psi(\overline{x}) - \langle v, x - \overline{x}\rangle}{\|x - \overline{x}\|} \geq 0 \right\}$$

whereas the basic/limiting/Mordukhovich subdifferential of ψ is the Kuratowski-Painlevée upper limit of the set-valued mapping $\widehat{\partial}\psi$ at $\overline{x}$:

$$\partial\psi(\overline{x}) := \operatorname*{Limsup}_{x \to \overline{x}} \widehat{\partial}\psi(x).$$

If ψ is convex, then $\partial\psi(\overline{x})$ reduces to the subdifferential in the sense of convex analysis, that is

$$\partial\psi(\overline{x}) := \{v \in \mathbb{R}^n |\ \psi(x) - \psi(\overline{x}) \geq \langle v, x - \overline{x}\rangle, \forall x \in \mathbb{R}^n\}. \tag{2.2}$$

For a local Lipschitz continuous function, $\partial\psi(\overline{x})$ is nonempty and compact. Moreover, its convex hull is the subdifferential of Clarke, that is, one can define the Clarke subdifferential $\overline{\partial}\psi(\overline{x})$ of ψ at $\overline{x}$ by

$$\overline{\partial}\psi(\overline{x}) := \operatorname{co} \partial\psi(\overline{x}) \tag{2.3}$$

where "co" stands for the convex hull of the set in question. Thanks to this link between the Mordukhovich and Clarke subdifferentials, we have the following convex hull property which plays and important role in this paper:

$$\operatorname{co} \partial(-\psi)(\overline{x}) = -\operatorname{co} \partial\psi(\overline{x}). \tag{2.4}$$

For this equality to hold, ψ should be Lipschitz continuous near $\overline{x}$.

We now introduce the basic/limiting/Mordukhovich normal cone to a set $\Omega \subseteq \mathbb{R}^n$ at one of its points $\overline{x}$

$$N_\Omega(\overline{x}) := \operatorname*{Limsup}_{x \to \overline{x}\,(x \in \Omega)} \widehat{N}_\Omega(\overline{x}) \tag{2.5}$$

where $\widehat{N}_\Omega(\overline{x})$ denotes the prenormal/Fréchet normal cone to Ω at $\overline{x}$ defined by

$$\widehat{N}_\Omega(\overline{x}) := \Big\{ v \in \mathbb{R}^n | \limsup_{x \to \overline{x}(x \in \Omega)} \frac{\langle v, x - \overline{x} \rangle}{\|x - \overline{x}\|} \leq 0 \Big\}$$

and "Limsup" stands for the Kuratowski-Painlevée upper limit defined in (2.1).

A set-valued mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ will be said to be inner semicompact at a point $\overline{x}$, with $\Psi(\overline{x}) \neq \emptyset$, if for every sequence $x_k \to \overline{x}$ with $\Psi(x_k) \neq \emptyset$, there is a sequence of $y_k \in \Psi(x_k)$ that contains a convergent subsequence as $k \to \infty$. It follows that the inner semicompactness holds whenever Ψ is uniformly bounded around $\overline{x}$, i.e. there exists a neighborhood U of $\overline{x}$ and a bounded set $\Omega \subset \mathbb{R}^m$ such that

$$\Psi(x) \subseteq \Omega, \text{ for all } x \in U.$$

The mapping Ψ is inner semicontinuous at $(\overline{x}, \overline{y}) \in \operatorname{gph} \Psi$ if for every sequence $x_k \to \overline{x}$ there is a sequence of $y_k \in \Psi(x_k)$ that converges to $\overline{y}$ as $k \to \infty$. Obviously, if Ψ is inner semicompact at $\overline{x}$ with $\Psi(\overline{x}) = \{\overline{y}\}$, then Ψ is inner semicontinuous at $(\overline{x}, \overline{y})$. In general though, the inner semicontinuity is a property much stronger than the inner semicompactness and it is a necessary condition for the Lipschitz-like property to hold. If Ψ has a close graph, Ψ is Lipschitz-like around $(\overline{x}, \overline{y})$ if and only if the following coderivative/Mordukhovich criterion holds [13]:

$$D^*\Psi(\overline{x}, \overline{y})(0) = \{0\}. \tag{2.6}$$

For $(\overline{x}, \overline{y}) \in \operatorname{gph} \Psi := \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m | y \in \Psi(x)\}$, the coderivative of Ψ at $(\overline{x}, \overline{y})$ is a homogeneous mapping $D^*\Psi(\overline{x}, \overline{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$, defined at $v \in \mathbb{R}^m$ by

$$D^*\Psi(\overline{x}, \overline{y})(v) := \{u \in \mathbb{R}^n | (u, -v) \in N_{\operatorname{gph} \Psi}(\overline{x}, \overline{y})\}. \tag{2.7}$$

Here, $N_{\operatorname{gph} \Psi}$ denotes the basic normal cone (2.5) to $\operatorname{gph} \Psi$. Finally, let us mention the calmness property that will also be useful in this paper. The set-valued mapping Ψ will be said to be calm at some point $(\overline{x}, \overline{y}) \in \operatorname{gph} \Psi$, if there exist neighborhoods U of $\overline{x}$, V of $\overline{y}$, and a constant $\kappa > 0$ such that

$$\Psi(x) \cap V \subseteq \Psi(\overline{x}) + \kappa \|x - \overline{x}\| \mathbb{B}, \text{ for all } x \in U$$

with $\mathbb{B}$ denoting the unit ball in $\mathbb{R}^m$. Ψ is automatically calm at $(\overline{x}, \overline{y})$, if it Lipschitz-like around the same point. Further details on sufficient conditions ensuring the calmness property can be found in [12] and references therein.

3. The optimistic bilevel programming problem. Our basic aim in this section is to derive necessary optimality for the optimistic formulation (P_o) of the bilevel programming problem (1.4). To proceed, we first study the sensitivity analysis of the two-level optimal value function φ_o (1.9). The necessary optimality conditions would then be deduced by a well-known result of [13]. To conclude this section, we will show how to deduce necessary optimality conditions of a problem to minimize a classical value function by means of the results obtained for (P_o). As constraint qualifications to compute basic normal cones or subdifferentials, we will impose the calmness of the following set-valued mappings:

$$\begin{array}{ll} \Phi^X(u) & := \{x \in \Omega_1 |\, \psi_1(x) + u \in \Lambda_1\}, \\ \Phi^K(u) & := \{(x,y) \in \Omega_2 |\, \psi_2(x,y) + u \in \Lambda_2\}, \\ \Phi^S(u) & := \{(x,y) \in \operatorname{gph} K |\, f(x,y) - \varphi(x) + u \leq 0\}. \end{array} \tag{3.1}$$

The set-valued mapping Φ^X (resp. Φ^K) (resp. Φ^S) is automatically calm at a point of its graph provided Ω_1 and Λ_1 (resp. Ω_2 and Λ_2) (resp. $\operatorname{gph} K$) are polyhedral and ψ_1 (resp. ψ_2) (resp. f and φ) is affine linear in its variables [18]. The lower-level value function φ is affine linear if, for example, the set-valued mapping K is defined by a system of affine linear equalities and/or inequalities while f is affine linear.

In order to obtain the Lipschitz continuity of the lower and the two-level value functions, respectively, the following additional conditions could be needed:

$$\Big[(x^*, 0) \in \partial\langle u, \psi_2\rangle(\overline{x}, \overline{y}) + N_{\Omega_2}(\overline{x}, \overline{y}), \text{ with } u \in N_{\Lambda_2}(\psi_2(\overline{x}, \overline{y}))\Big] \Longrightarrow x^* = 0, \tag{3.2}$$

$$\left.\begin{array}{r} (x^*, 0) \in r\partial(f - \varphi)(\overline{x}, \overline{y}) + \partial\langle u, \psi_2\rangle(\overline{x}, \overline{y}) + N_{\Omega_2}(\overline{x}, \overline{y}) \\ r \geq 0,\, u \in N_{\Lambda_2}(\psi_2(\overline{x}, \overline{y})) \end{array}\right\} \Longrightarrow x^* = 0. \tag{3.3}$$

As it will be clear in the proof of Theorem 3.2, they are sufficient conditions for the coderivative criterion (2.6) to hold for the set-valued mappings $K(x) := \{y |\, (x, y) \in \Omega_2 \cap \psi_2^{-1}(\Lambda_2)\}$ and $S(x) := \{y \in K(x) |\, f(x, y) \leq \varphi(x)\}$, respectively, provided the corresponding calmness conditions above are satisfied. Further details on these types of conditions and more generally on the development of coderivatives can be found in [13, 20]. It is important to mention that the basic/Mangasarian-Fromovitz CQ is not satisfied for the constraint structure $y \in S(x)$ described by S when the above lower-level optimal value function reformulation is in consideration [5]. However, condition (3.3) can be satisfied, in particular, in the framework provided in the next proposition taken from [8].

Proposition 3.1 (validity of qualification condition (3.3)). *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}^p$ be two convex and continuously differentiable functions. Consider the value function*

$$\varphi(x) := \min_y \{f(x, y) |\, g(y) \leq 0\}.$$

Let $(\overline{x}, \overline{y}) \in gph\, S := \{(x, y) |\, g(y) \leq 0,\, f(x, y) \leq \varphi(x)\}$ and assume that $\varphi(\overline{x}) < \infty$ while $u \rightrightarrows \{y |\, g(y) + u \leq 0\}$ is calm at $(0, \overline{y})$. Then, CQ (3.3) holds at $(\overline{x}, \overline{y})$.

To complete the list of assumptions mostly used in this section, we mention the inner semicompactness or semicontinuity of the following solution set-valued mapping:

$$S_o(x) := \arg\min_y \{F(x, y) |\, y \in S(x)\} = \{y \in S(x) | F(x, y) \leq \varphi_o(x)\}. \tag{3.4}$$

In the next result, we first derive the sensitivity analysis of the two-level value function φ_o when all the functions involved are locally Lipschitz continuous and the solution set-valued mapping S_o is inner semicompact.

Theorem 3.2 (sensitivity analysis of the two-level value function under inner semicompactness). *Consider the two-level value function φ_o (1.9), where the functions f, ψ_2 and F are Lipschitz continuous near $(\overline{x}, y)$, $y \in S(\overline{x})$ and $(\overline{x}, y)$, $y \in S_o(\overline{x})$,* *respectively. Assume that the sets Ω_2 and Λ_2 are closed and the set-valued mapping Φ^K (resp. Φ^S) is calm at $(0, \overline{x}, y)$, $y \in S(\overline{x})$ (resp. $(0, \overline{x}, y)$, $y \in S_o(\overline{x})$). Also let S_o be inner semicompact at $\overline{x}$ and CQs (3.2) and (3.3) be satisfied at $(\overline{x}, y)$, $y \in S(\overline{x})$ and $(\overline{x}, y)$, $y \in S_o(\overline{x})$, respectively. Then, φ_o is Lipschitz continuous around $\overline{x}$ and one has the following upper estimate of its basic subdifferential:*

$$\begin{aligned}\partial\varphi_o(\overline{x}) \subseteq & \bigcup_{y\in S_o(\overline{x})} \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \bigcup_{r\geq 0} \Big\{x^* \Big| \\ & \big(x^* + r\textstyle\sum_{s=1}^{n+1} v_s x_s^*, 0\big) \in \partial F(\overline{x}, y) + r\partial f(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y), \\ & \textstyle\sum_{s=1}^{n+1} v_s = 1 \text{ and for all } s = 1, \ldots, n+1 : \\ & v_s \geq 0,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)),\ y_s \in S(\overline{x}), \\ & (x_s^*, 0) \in \partial f(\overline{x}, y_s) + \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s)\Big\}.\end{aligned}$$

Proof. Since S_o is inner semicompact at $\overline{x}$, one has from [17, Theorem 7 (ii)] that

$$\partial\varphi_o(\overline{x}) \subseteq \bigcup_{y\in S_o(\overline{x})} \big\{x^* + D^*S(\overline{x}, y)(y^*) \big| \, (x^*, y^*) \in \partial F(\overline{x}, y)\big\} \tag{3.5}$$

taking into account the Lipschitz continuity of F near $(\overline{x}, y)$, $y \in S_o(\overline{x})$. Now, let us note that $\operatorname{gph} K = \Phi^K(0)$. Hence, applying [12, Theorem 4.1], one gets

$$N_{\operatorname{gph} K}(\overline{x}, y) = N_{\Phi^K(0)}(\overline{x}, y) \subseteq \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \big\{\partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y)\big\}, \tag{3.6}$$

for $y \in S(\overline{x})$, since Ω_2 and Λ_2 are closed, ψ_2 is Lipschitz continuous near $(\overline{x}, y)$ and the set-valued mapping Φ^K is calm at $(0, \overline{x}, y)$. An upper bound for the coderivative of K then results from the definition of the coderivative

$$D^*K(\overline{x}, y)(y^*) \subseteq \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \Big\{x^* \Big| \, (x^*, -y^*) \in \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y)\Big\}. \tag{3.7}$$

From this inclusion, it is clear, taking into account the homogeneity of the coderivative, that CQ (3.2) is a sufficient condition for the coderivative criterion to hold for the set-valued mapping K at $(\overline{x}, y)$, $y \in S(\overline{x})$. Combining the latter fact with the inner semicompactness of S (obtained from that of S_o since $S_o(x) \subseteq S(x)$ for all $x \in X$), one gets the Lipschitz continuity of the lower-level value function φ near $\overline{x}$, cf. [14, Theorem 5.2 (ii)].

Repeating the above process of estimating the coderivative of K on the lower-level solution set-valued mapping S, one can also observe that $\operatorname{gph} S = \Phi^S(0)$. Hence, for $y \in S_o(\overline{x})$, Ω_2 and Λ_2 being closed, we have

$$N_{\operatorname{gph} S}(\overline{x}, y) = N_{\Phi^S(0)}(\overline{x}, y) \subseteq \bigcup_{r\geq 0} \big\{r\partial(f - \varphi)(\overline{x}, y) + N_{\operatorname{gph} K}(\overline{x}, y)\big\}, \tag{3.8}$$

considering the calmness of Φ^S at $(\overline{x}, y)$, the Lipschitz continuity of f, ψ_2 and φ near $(\overline{x}, y)$, $y \in S(\overline{x})$ and $\overline{x}$, respectively, cf. [12, Theorem 4.1]. Combining (3.6) and (3.8), the definition of the coderivative yields the following upper estimate for the coderivative of S:

$$\begin{aligned}D^*S(\overline{x}, y)(y^*) \subseteq \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \bigcup_{r\geq 0} \Big\{x^* \Big| \, (x^*, -y^*) \in r\partial(f - \varphi)(\overline{x}, y) + \\ \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y)\Big\}.\end{aligned} \tag{3.9}$$

With this inclusion, it is now clear, also taking into account the homogeneity of the coderivative set-valued mapping, that in addition to the other assumptions of the theorem, the fulfilment of CQ (3.3) at $(\overline{x}, y)$ implies the satisfaction of the coderivative criterion for S at the same point. Hence, the two-level value function φ_o is also Lipschitz continuous near $\overline{x}$, cf. [14, Theorem 5.2 (ii)].

The presence of the lower-level value function φ in the right-hand-side of inclusion (3.9) gives room for a further description of the upper bound of $D^*S(\overline{x}, y)(y^*)$ in terms of problem data. Hence, recall that by a combination of inclusion (3.7) and the inner semicompactness of S (due to that of S_o), one has

$$\partial\varphi(\overline{x}) \subseteq \bigcup_{y\in S(\overline{x})} \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \Big\{x^* \big| (x^*, 0) \in \partial f(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y)\Big\}. \tag{3.10}$$

Since φ is Lipschitz continuous near $\overline{x}$, we have

$$\partial(-\varphi)(\overline{x}) \subseteq \mathrm{co}\,\partial(-\varphi)(\overline{x}) = -\mathrm{co}\,\partial\varphi(\overline{x}), \tag{3.11}$$

where the last equality follows from the convex hull property (2.4). Applying Carathéodory's well-known theorem on the evaluation of the convex hull of a set in finite dimension, one gets the following upper estimate of the basic subdifferential of $-\varphi$, while considering inclusions (3.10) and (3.11)

$$\begin{aligned}\partial(-\varphi)(\overline{x}) \subseteq \;\; & \Big\{ -\textstyle\sum_{s=1}^{n+1} v_s x_s^* \big| \sum_{s=1}^{n+1} v_s = 1 \text{ and for all } s = 1, \dots, n+1 : \\ & \quad v_s \geq 0,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)),\ y_s \in S(\overline{x}), \\ & \quad (x_s^*, 0) \in \partial f(\overline{x}, y_s) + \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s)\Big\}.\end{aligned} \tag{3.12}$$

Inserting this inclusion in (3.9), one gets the following fully detailed upper bound of the coderivative of the solution set-valued mapping S:

$$\begin{aligned}D^*S(\overline{x}, y)(y^*) \;\; & \subseteq \bigcup_{r\geq 0} \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \Big\{x^* \big| \\ & \big(x^* + r\textstyle\sum_{s=1}^{n+1} v_s x_s^*, -y^*\big) \in r\partial f(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y), \\ & \textstyle\sum_{s=1}^{n+1} v_s = 1 \text{ and for all } s = 1, \dots, n+1 : \\ & v_s \geq 0,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)),\ y_s \in S(\overline{x}), \\ & (x_s^*, 0) \in \partial f(\overline{x}, y_s) + \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s)\Big\}.\end{aligned}$$

Finally, we have the upper estimate of the basic subdifferential of the two-level value function φ_o in the theorem by a combination of the latter inclusion and the one in (3.5). □

The convex combination in the upper bound of $\partial\varphi_o$ obtained in this theorem can be avoided. One way, that we consider in the next result, is to assume that the functions involved in the lower-level problem are convex. The second possibility, studied in Theorem 3.5, is to replace the inner semicompactness of S_o (3.4) in the above theorem by the stronger inner semicontinuity. A third approach consist in allowing one to have the difference rule $\partial(\psi_1 - \psi_2)(\overline{x}) \subseteq \partial\psi_1(\overline{x}) - \partial\psi_2(\overline{x})$ for the basic subdifferential. According to the work in [15], this may require that the Fréchet subdifferential of ψ_2 be nonempty in a neighborhood of $\overline{x}$. The latter case is out of the scope of this paper. Further discussion on this issue can also be found in [4].

In the next result, we still impose the inner semicompactness of the set-valued mapping S_o (3.4), but we assume that all the functions involved in the two-level value function φ_o are instead fully convex, that is, convex in both variables (x, y).

THEOREM 3.3 (sensitivity analysis of the two-level value function with convex initial data). *Consider the two-level value function* φ_o (1.9)*, where the functions* f*,* $\psi_2 := g$ *and* F *are fully convex (that is, convex in* (x,y)*) and the sets* $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$ *and* $\Lambda_2 := \mathbb{R}^p_-$*. Assume the set-valued mapping* Φ^K *(resp.* Φ^S*) is calm at* $(0,\overline{x},y)$*,* $y \in S(\overline{x})$ *(resp.* $(0,\overline{x},y)$*,* $y \in S_o(\overline{x})$*). Also let* $\overline{x} \in dom\ \varphi$*, where* S_o *is inner semicompact and let CQ* (3.3) *be satisfied at* $(\overline{x},y)$*,* $y \in S_o(\overline{x})$*. Then,* φ_o *is Lipschitz continuous around* $\overline{x}$ *and one has the following upper estimate of its basic subdifferential:*

$$\partial\varphi_o(\overline{x}) \subseteq \bigcup_{y\in S_o(\overline{x})} \ \bigcup_{(r,\beta)\in\Lambda^o(\overline{x},y)} \ \bigcup_{\gamma\in\Lambda(\overline{x},y)} \Big\{\partial_x F(\overline{x},y) + r\big(\partial_x f(\overline{x},y) - \partial_x f(\overline{x},y)\big) + \sum_{i=1}^p \beta_i \partial_x g_i(\overline{x},y) - r\sum_{i=1}^p \gamma_i \partial_x g_i(\overline{x},y)\Big\}.$$

where the multiplier sets $\Lambda(\overline{x},y)$ *and* $\Lambda^o(\overline{x},y)$ *are given respectively as:*

$$\Lambda(\overline{x},y) := \{\gamma| \quad \begin{array}{l} 0 \in \partial_y f(\overline{x},y) + \sum_{i=1}^p \gamma_i \partial_y g_i(\overline{x},y) \\ \gamma_i \geq 0,\ \gamma_i g_i(\overline{x},y) = 0,\ i = 1,\dots,p\}, \end{array} \tag{3.13}$$

$$\Lambda^o(\overline{x},y) := \{(r,\beta)| \quad \begin{array}{l} 0 \in \partial_y F(\overline{x},y) + r\partial_y f(\overline{x},y) + \sum_{i=1}^p \beta_i \partial_y g_i(\overline{x},y) \\ r \geq 0,\ \beta_i \geq 0,\ \beta_i g_i(\overline{x},y) = 0,\ i = 1,\dots,p\}. \end{array} \tag{3.14}$$

Proof. The proof is essentially the same as that of Theorem 3.2. However, let us notice that with

$$\varphi(x) = \min_y \{f(x,y)|\, g(x,y) \leq 0\},$$

the functions f and g being fully convex, φ is convex and one has the following estimate of its subdifferential

$$\partial\varphi(\overline{x}) \subseteq \bigcup_{\gamma\in\Lambda(\overline{x},y)} \big\{\partial_x f(\overline{x},y) + \sum_{i=1}^p \gamma_i \partial_x g_i(\overline{x},y)\big\}, \tag{3.15}$$

for $y \in S(\overline{x})$, taking into account the calmness of the set-valued mapping Φ^K at $(0,\overline{x},y)$, cf. e.g. the proof of Theorem 4.1 in [24]. Moreover, φ is Lipschitz continuous near $\overline{x}$, as a convex function with $\overline{x} \in \operatorname{dom}\varphi$.

The second observation to make is that, φ being convex, inclusion (3.11) becomes

$$\partial(-\varphi)(\overline{x}) \subseteq -\partial\varphi(\overline{x}). \tag{3.16}$$

Combining (3.5), (3.15) and (3.16), one has the upper estimate of the basic subdifferential of φ_o in the corollary, while taking into account the decomposition formula

$$\partial\psi(x,y) \subseteq \partial_x\psi(x,y) \times \partial_y\psi(x,y)$$

which is valid for the functions F, f, and g since they are all fully convex. □

In [3], a simple bilevel programming problem is defined as the optimization problem to

$$\text{minimize } F(y) \text{ subject to } y \in S := \arg\min\{f(u)|\, g(u) \leq 0\}. \tag{3.17}$$

Obviously, perturbing this problem on the left- and-right-hand-sides yields our two-level value function φ_o (1.9). In the next result, we derive the sensitivity analysis of the simple bilevel programming problem in the case of left-hand-side perturbation.

PROPOSITION 3.4 (variational analysis of the value function of a simple bilevel programming problem). *Consider the two-level value function* φ_o (1.9)*, where for all* $x \in X$*,* $S(x) := S := \arg\min\{f(y)|\, g(y) \leq 0\}$ *with the functions* f*,* g *and* F *convex in* y *and* (x, y)*, respectively. Let* $\overline{x} \in dom\ \varphi_o$*, then* φ_o *is Lipschitz continuous around this point and for* $y \in S_o(\overline{x})$*, we have:*

$$\partial\varphi_o(\overline{x}) \subseteq \partial_x F(\overline{x}, y).$$

Proof. Under the assumptions of the proposition, φ_o is Lipschitz continuous near $\overline{x}$. Moreover, proceeding as in the proof of Theorem 4.1 in [24], take $\overline{y} \in S_o(\overline{x})$, then for $u \in \partial\varphi_o(\overline{x})$, $(\overline{x}, \overline{y})$ is a solution of the problem

$$\min_{x,y}\{F(x, y) - \langle u, x\rangle|\, (x, y) \in \mathbb{R}^n \times S\}.$$

Applying [13, Proposition 5.3], the optimality condition of the latter problem is obtained as

$$(u, 0) \in \partial F(\overline{x}, \overline{y}) + N_{\mathbb{R}^n \times S}(\overline{x}, \overline{y}).$$

Since F is fully convex, this implies the following inclusions

$$u \in \partial_x F(\overline{x}, \overline{y}) \text{ and } 0 \in N_S(\overline{y}).$$

Note that $0 \in N_S(\overline{y})$ is always true. Hence, the result. □

Reconsidering the two-level value function φ_o (1.9) in the case of right- and left-hand-side perturbations, we now derive its sensitivity analysis when the functions involved are locally Lipschitz continuous while S_o (3.4) is inner semicontinuous. For this, the calmness of the set-valued mapping K (3.1) will be replaced by the stronger dual form of the Mangasarian-Fromovitz constraint qualification in terms of Clarke's subdifferential:

$$\Big[0 \in \sum_{i=1}^{p} \gamma_i \overline{\partial} g_i(\overline{x}, \overline{y}),\ \gamma_i \geq 0, \gamma_i g_i(\overline{x}, \overline{y}) = 0, i = 1, \ldots, p\Big] \Longrightarrow \gamma_i = 0,\ i = 1, \ldots, p. \tag{3.18}$$

THEOREM 3.5 (sensitivity analysis of the two-level value function under the inner semicontinuity). *Consider the two-level value function* φ_o (1.9)*, where the functions* f*,* $\psi_2 := g$ *and* F *are Lipschitz continuous near* $(\overline{x}, \overline{y})$*. Assume that the sets* $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$ *and* $\Lambda_2 := \mathbb{R}^p_-$ *and the set-valued mapping* Φ^S *is calm at* $(0, \overline{x}, \overline{y})$*. Also let* S_o *be inner semicontinuous at* $(\overline{x}, \overline{y})$ *and CQs* (3.3) *and* (3.18) *be satisfied at* $(\overline{x}, \overline{y})$*. Then,* φ_o *is Lipschitz continuous around* $\overline{x}$ *and one has the following upper estimate of its basic subdifferential:*

$$\begin{aligned}\partial\varphi_o(\overline{x}) \subseteq \bigcup_{r\geq 0}\Big\{x^*\Big|\quad & (x^* + r x^*_\varphi, 0) \in \partial F(\overline{x}, \overline{y}) + r\partial f(\overline{x}, \overline{y}) + \textstyle\sum_{i=1}^p \beta_i \partial g_i(\overline{x}, \overline{y}),\\ & (x^*_\varphi, 0) \in \overline{\partial} f(\overline{x}, \overline{y}) + \textstyle\sum_{i=1}^p \gamma_i \overline{\partial} g_i(\overline{x}, \overline{y}),\\ & \beta_i \geq 0,\ \beta_i g_i(\overline{x}, \overline{y}) = 0,\ i = 1, \ldots, p,\\ & \gamma_i \geq 0,\ \gamma_i g_i(\overline{x}, \overline{y}) = 0,\ i = 1, \ldots, p\Big\}.\end{aligned}$$

Proof. The first thing to note is that since $S_o(x) \subseteq S(x)$, for all $x \in X$, S is inner semicontinous at $(\overline{x}, \overline{y})$, considering the fact that the latter property is assumed to also hold for S_o. In addition to CQ (3.18), it follows from [16, Theorem 5.9] that

$$\overline{\partial}\varphi(\overline{x}) \subseteq \big\{u| \quad (u,0) \in \overline{\partial} f(\overline{x},\overline{y}) + \textstyle\sum_{i=1}^{p} \gamma_i \overline{\partial} g_i(\overline{x},\overline{y}), \\ \text{for } i = 1,\dots,p,\ \gamma_i \geq 0,\ \gamma_i g_i(\overline{x},\overline{y}) = 0\big\}. \tag{3.19}$$

Furthermore, considering CQ (3.2), the lower-level value function φ is Lipschitz continuous near $\overline{x}$. On the other hand, applying [17, Theorem 7 (i)], one has

$$\partial\varphi_o(\overline{x}) \subseteq \big\{x^* + D^*S(\overline{x},\overline{y})(y^*)|\, (x^*,y^*) \in \partial F(\overline{x},\overline{y})\big\} \tag{3.20}$$

taking into account the Lipschitz continuity of F near $(\overline{x}, \overline{y})$ and the inner semicontinuity of S_o at the same point. Considering the Lipschitz continuity of φ near $\overline{x}$, one has from inclusion (3.9) that

$$D^*S(\overline{x},\overline{y})(y^*) \subseteq \bigcup_{r\geq 0} \Big\{x^*| \quad (x^* + rx^*_\varphi, -y^*) \in r\partial f(\overline{x},y) + \textstyle\sum_{i=1}^{p} \beta_i \partial g_i(\overline{x},y), \\ (x^*_\varphi, 0) \in \overline{\partial} f(\overline{x},\overline{y}) + \textstyle\sum_{i=1}^{p} \gamma_i \overline{\partial} g_i(\overline{x},y), \\ \beta_i \geq 0,\ \beta_i g_i(\overline{x},\overline{y}) = 0,\ i = 1,\dots,p, \\ \gamma_i \geq 0,\ \gamma_i g_i(\overline{x},y) = 0,\ i = 1,\dots,p\Big\} \tag{3.21}$$

while considering the convex hull property (3.11). Combining (3.20) and (3.21), one has the upper bound of $\partial\varphi_o(\overline{x})$ in the result. The Lipschitz continuity of the two-level value function φ_o is obtained as in the proof of Theorem 3.2, cf. [14, Theorem 5.2 (i)]. □

Remark 3.6 (Comments on the sensitivity analysis results of the two-level value function). *We provide some possible links between the results above and also relationships with previous work.*

(i) *If we assume that in the right-hand-side of the inclusion in Theorem 3.2 one has $S_o(\overline{x}) = \{y\}$, $S(\overline{x}) = \{y\}$ and the multiplier u in the upper estimate* (3.10) *of $\partial\varphi(\overline{x})$ is unique, then all the upper bounds of $\partial\varphi_o(\overline{x})$ obtained in Theorems 3.2, 3.3 and 3.5 coincide, provided that the functions are subdifferentially regular and the sets Ω_2 and Λ_2 are adjusted accordingly.*

(ii) *If we assume that all the functions involved in the two-level value function φ_o* (1.9) *are continuously differentiable, while setting $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$ and $\Lambda_2 := \mathbb{R}^p_-$, the upper bounds that we get for $\partial\varphi_o$ in Theorems 3.2 and 3.5 are exactly the same as those obtained in [8, Theorem 5.9]. This will induce the same observation when the necessary optimality conditions of the optimistic bilevel programming problem (P_o) are considered in Corollary 3.7.*

(iii) *Note that for the upper bounds of the basic subdifferential of φ_o* (1.9) *in Theorems 3.2, 3.3 and 3.5, CQ* (3.3) *is in fact not necessary. The latter condition comes into play only to obtain the Lipschitz continuity of the aforementioned two-level value function.*

For the rest of this section, we focus our attention on deriving necessary optimality conditions for the original optimistic formulation of the bilevel programming problem (P_o). Necessary optimality for an optimization problem with a classical value function objective will be discussed as well.

Corollary 3.7 (optimality conditions for the original optimistic bilevel programming problem). *Let $\overline{x}$ be a local optimal solution of (P_o), where ψ_1 is Lipschitz* *continuous near $\overline{x}$ and the sets Ω_1 and Λ_1 are closed. Furthermore, let the set-valued mapping Φ^X be calm at $(0, \overline{x})$. Then, the following assertions hold:*

(i) *Let all the assumptions of Theorem 3.2 be satisfied. Then, there exist $y \in S_o(\overline{x})$, $(\alpha, u, r) \in \mathbb{R}^{k+p+1}$, $(u_s, v_s) \in \mathbb{R}^{p+1}$ and $y_s \in S(\overline{x})$, $x_s^* \in \mathbb{R}^n$ with $s = 1, \ldots, n+1$ such that:*

$$\big(r\sum_{s=1}^{n+1} v_s x_s^*, 0\big) \in \partial F(\overline{x}, y) + r\partial f(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y) + \big[\partial\langle \alpha, \psi_1\rangle(\overline{x}) + N_{\Omega_1}(\overline{x})\big] \times \{0\}, \quad (3.22)$$

$$\text{for } s = 1, \ldots, n+1,\ (x_s^*, 0) \in \partial f(\overline{x}, y_s) + \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s), \quad (3.23)$$

$$\text{for } s = 1, \ldots, n+1,\ v_s \geq 0,\ \sum_{s=1}^{n+1} v_s = 1, \quad (3.24)$$

$$\text{for } s = 1, \ldots, n+1,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)), \quad (3.25)$$

$$u \in N_{\Lambda_2}(\psi_2(\overline{x}, y)), \quad (3.26)$$

$$r \geq 0,\ \alpha \in N_{\Lambda_1}(\psi_1(\overline{x})). \quad (3.27)$$

(ii) *Let all the assumptions of Theorem 3.3 be satisfied. Then, there exist $y \in S_o(\overline{x})$, α, β, γ and r such that we have condition* (3.27) *together with the following relationships to be satisfied:*

$$0 \in \partial_x F(\overline{x}, y) + r\big(\partial_x f(\overline{x}, y) - \partial_x f(\overline{x}, y)\big) + \sum_{i=1}^{p} \beta_i \partial_x g_i(\overline{x}, y) - r\sum_{i=1}^{p} \gamma_i \partial_x g_i(\overline{x}, y) + \partial\langle \alpha, \psi_1\rangle(\overline{x}) + N_{\Omega_1}(\overline{x}), \quad (3.28)$$

$$0 \in \partial_y F(\overline{x}, y) + r\partial_y f(\overline{x}, y) + \sum_{i=1}^{p} \beta_i \partial_y g_i(\overline{x}, y), \quad (3.29)$$

$$0 \in \partial_y f(\overline{x}, y) + \sum_{i=1}^{p} \gamma_i \partial_y g_i(\overline{x}, y), \quad (3.30)$$

$$\text{for } i = 1, \ldots, p,\ \beta_i \geq 0,\ \beta_i g_i(\overline{x}, y) = 0, \quad (3.31)$$

$$\text{for } i = 1, \ldots, p,\ \gamma_i \geq 0,\ \gamma_i g_i(\overline{x}, y) = 0. \quad (3.32)$$

(iii) *Let all the assumptions of Theorem 3.5 be satisfied. Then, there exist x^*, α, β, γ and r such that we have relationships* (3.27) *and* (3.31)–(3.32) (*where $y := \overline{y}$*) *together with the following conditions to hold:*

$$(rx^*, 0) \in \partial F(\overline{x}, \overline{y}) + r\partial f(\overline{x}, \overline{y}) + \sum_{i=1}^{p} \beta_i \partial g_i(\overline{x}, \overline{y}) + \big[\partial\langle \alpha, \psi_1\rangle(\overline{x}) + N_{\Omega_1}(\overline{x})\big] \times \{0\}, \quad (3.33)$$

$$(x^*, 0) \in \overline{\partial} f(\overline{x}, \overline{y}) + \sum_{i=1}^{p} \gamma_i \overline{\partial} g_i(\overline{x}, \overline{y}). \quad (3.34)$$

Proof. Under the assumptions of either (i), (ii) or (iii), the two-level value function φ_o (1.9) is Lipschitz continuous near $\overline{x}$. Hence, since $X := \Omega_1 \cap \psi^{-1}(\Lambda_1)$ is closed, one has from [13, Proposition 5.3] that

$$0 \in \varphi_o(\overline{x}) + N_X(\overline{x}). \tag{3.35}$$

Under the calmness of the set-valued mapping Φ^X at $(0, \overline{x})$, we have from [12, Theorem 4.1] that

$$N_X(\overline{x}) \subseteq \bigcup \left\{ \partial\langle \alpha, \psi_1 \rangle(\overline{x}) + N_{\Omega_1}(\overline{x}) |\, \alpha \in N_{\Lambda_1}(\psi_1(\overline{x})) \right\}. \tag{3.36}$$

Combining (3.35) and (3.36) on the one hand and Theorem 3.2, Theorem 3.3 and Theorem 3.5, respectively, on the other, one gets the result. □

If the assumptions of Remark 3.6 (i) are satisfied, the optimality conditions in Corollary 3.7 (i), (ii) and (iii) all coincide. Also note that since the optimality conditions of Corollary 3.7 (ii) depend on the couple $(\overline{x}, \overline{y})$, where S_o is inner semicontinuous, one should obtain further stronger optimality conditions in this case, provided the number of points $(\overline{x}, \overline{y}) \in \operatorname{gph} S_o$, where this property holds grows while making sure the other assumptions are also satisfied at these points [8]. It is also worth mentioning that the Lipschitz continuity of the two-level value function (1.9) is needed in the above result mostly to make sure that $\partial\varphi_o(\overline{x}) \neq \emptyset$. If we are sure that the latter is satisfied for a given problem, CQ (3.3) can be dropped, provided for instance that there is no upper level constraint, that is, $X := \mathbb{R}^n$.

Remark 3.8 (comparison with previous works on the optimality conditions of the optimistic bilevel program). *Consider the original optimistic bilevel optimization problem (P_o) where $\Omega_1 := \mathbb{R}^n$, $\Lambda := \mathbb{R}^k_-$, $\Omega := \mathbb{R}^n \times \mathbb{R}^m$ and $\Lambda := \mathbb{R}^p_-$, then the optimality conditions obtained in Corollary 3.7 (i), (ii) and (iii) are exactly those derived in [4, Theorem 5.1], [4, Theorem 4.1] and [16, Theorem 6.4], respectively, while using the auxiliary/classical optimistic bilevel programming problem* (1.6). *The same observation can be made in the case of smooth functions, cf. Remark 3.6 (ii). In the latter framework, a result similar to Corollary 3.7 (i) can also be found in [5, 23]. The only thing missing in the optimality conditions of [4, 5, 16, 23] is the condition that $y \in S_o(\overline{x})$. On the CQs, the assumptions used in Corollary 3.7 are parallel to those of the aforementioned works. In [4, 16, 23] the partial calmness is used as CQ. We replace it here by the calmness of the set-valued mapping Φ^S, which implies the satisfaction of the partial calmness, cf. Proof of Theorem 4.10 in [5]. The upper- and lower-level regularity conditions also used in [4, 16] imply the fulfilment of the calmness of the set-valued mappings Φ^X and Φ^K that we impose in Corollary 3.7. Furthermore, CQ* (3.2) *is satisfied provided the lower-level regularity holds. The reader is referred to [4, 5, 16, 23] and references therein for the definitions and discussions on the partial calmness, the upper- and lower-level regularity conditions. The only assumption in Corollary 3.7 that could be seen as additional as compared to those used in the aforementioned papers is CQ* (3.3)*, which is needed in our result to guaranty the Lipshitz continuity of the two-level value function φ_o* (1.9) *and hence ensuring that $\partial\varphi_o(\overline{x}) \neq \emptyset$. If we are sure that the latter condition is satisfied, as it happens in many non-Lipschitzian cases [13, 17, 20], then CQ* (3.3) *becomes superfluous.*

Remark 3.9 (what do we gain using the original optimistic model (P_o) to derive necessary optimality conditions for the bilevel program?). *As observed in the previous remark, the necessary optimality conditions obtained in Corollary 3.7 are the same as those derived in [4, 5, 16, 23] while using the auxiliary/classical model* (1.6)*, provided it is imposed in the latter case that $y \in S_o(\overline{x})$. As mentioned in the Introduction, the first motivation to investigate the original optimistic problem (P_o) separately is* *that it is not locally equivalent to the auxiliary problem* (1.6)*. This may well pose a problem while solving the bilevel problem especially when considering sufficient optimality conditions. Furthermore, as it will be clear in the next section, the another major advantage considering (*P_o*) is that the results in Theorem 3.2, Theorem 3.3 and Theorem 3.5 can readily be applied to derive necessary optimality conditions for the pessimistic bilevel optimization problem (*P_p*). Last but not least, it is worth mentioning that the results on the sensitivity analysis of the two-level value function* φ_o (1.9) *obtained above can directly be applied to derive the sensitivity analysis of the value function of classical optimistic problem* (1.6)*.*

In problem (P_o), if we set the lower-level objective function to be a constant, that is $f(x,y) := c$ for all $(x,y) \in \mathbb{R}^n \times \mathbb{R}^m$, then the lower-level optimal value function is also a constant, namely $\varphi(x) = c$ for all $x \in X$. Hence, (P_o) becomes a simple problem to minimize a classical value function:

$$(P) \quad \begin{cases} \text{minimize} & \varphi_o(x) \text{ subject to } x \in X \\ \text{where} & X := \Omega_1 \cap \psi_1^{-1}(\Lambda_1) \\ & \varphi_o(x) := \min\limits_y \{F(x,y) |\, y \in K(x)\} \\ & K(x) := \{y |\, (x,y) \in \Omega_2 \cap \psi_2^{-1}(\Lambda_2)\}. \end{cases}$$

Clearly, this means that (P) is contained in (P_o). No special work needs to be done for the sensitivity analysis of the value function φ_o in its current form

$$\varphi_o(x) := \min_y \{F(x,y) |\, y \in K(x)\}. \tag{3.37}$$

A great number of publications can be found in the literature on this topic. See e.g. [1, 10, 11, 19], when $y \in K(x)$ is replaced by the usual functional inequality and/or equality structure. For more general constraint structures, the notion of coderivative was used in [14, 17] to investigate the sensitivity analysis of φ_o (3.37). The only point we would like to make here is that the sensitivity analysis of a classical optimal value function φ_o (3.37) can be deduced from that of a two-level value function φ_o (1.9). Thus necessary optimality conditions for an optimization problem with minimization value function can be deduced from those of the optimistic bilevel optimization problem (P_o). The thing to do is deleting all the terms referring to the lower-level objective function f and value function φ from the basic subdifferential of the two-level value function (1.9) in Theorem 3.2, Theorem 3.3 and Theorem 3.5 (resp. the optimality conditions in Corollary 3.7). The assumptions insuring the Lipschitz continuity of φ_o (3.37) also have to be adjusted accordingly. Practically, for the upper bound of $\partial\varphi_o$ in Theorem 3.2, this would mean setting f to a constant and $x_s^* = 0$, for $s = 1, \ldots, n+1$ and deleting all the set of conditions:

$$\begin{aligned} &\sum_{s=1}^{n+1} v_s = 1 \text{ and for all } s = 1, \ldots, n+1 : \\ &v_s \geq 0,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)),\ y_s \in S(\overline{x}), \\ (x_s^*, 0) \in \partial f(\overline{x}, y_s) &+ \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s). \end{aligned}$$

The basic subdifferential of the classical value function φ_o (3.37) then follows as

$$\begin{aligned} \partial\varphi_o(\overline{x}) \quad \subseteq & \bigcup_{y \in S_o(\overline{x})} \bigcup_{u \in N_{\Lambda_2}(\psi_2(\overline{x},y))} \Big\{x^* \big| \\ & \big(x^*, 0\big) \in \partial F(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y)\Big\}, \end{aligned}$$

provided ψ_2 and F are Lipschitz continuous near $(\overline{x}, y)$, $y \in S_o(\overline{x})$, the sets Ω_2, Λ_2 are closed and the set-valued mapping Φ^K is calm at $(0, \overline{x}, y)$, $y \in S_o(\overline{x})$, while S_o is inner semicompact at $\overline{x}$. It should be reminded here that the value of the set-valued mapping S_o corresponding to φ_o (3.37) is

$$S_o(x) := \arg\min_y \{F(x,y) |\, y \in K(x)\}. \tag{3.38}$$

If in addition CQ (3.2) is satisfied at $(\overline{x}, y)$, $y \in S_o(\overline{x})$, then φ_o (3.37) is Lipschitz continuous near $\overline{x}$. This result was obtained in [14, 17], where instead of the calmness of the set-valued mapping K, the stronger basic/Mangasarian-Fromovitz CQ is assumed. This seems to suggest that the tow-level value function (1.9) can be seen as a kind of extension of the usual notion of value function.

From the above observations, it results that if $\overline{x}$ is a local optimal solution of (P) and the aforementioned assumptions are satisfied, then there exist $y \in S_o(\overline{x})$, $(\alpha, u) \in \mathbb{R}^{k+p}$ such that:

$$\begin{aligned}\big(0,0\big) \in \partial F(\overline{x}, y) + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y) + \big[\partial\langle \alpha, \psi_1\rangle(\overline{x}) + N_{\Omega_1}(\overline{x})\big] \times \{0\},\\ \alpha \in N_{\Lambda_1}(\psi_1(\overline{x})),\; u \in N_{\Lambda_2}(\psi_2(\overline{x}, \overline{y})),\end{aligned}$$

provided the set-valued mapping Φ^X is calm at $(0, \overline{x})$. Similar observations can be made for φ_o (3.37) in the frameworks of Theorem 3.3 and Theorem 3.5 (resp. Corollary 3.7 (ii) and (iii)).

We now conclude this section by considering the case where φ_o (3.37) is convex. Such a framework can not be thought of while dealing with the original optimistic bilevel problem (P_o) since the two-level value function φ_o (1.9) cannot be convex, except the lower-level problem is a nonperturbed convex optimization problem, see Proposition 3.4. This leads to necessary and sufficient optimality conditions for problem (P).

THEOREM 3.10 (necessary and sufficient optimality conditions for the minmin problem in the convex case). *Consider problem (P), where F and $\psi_2 := g$ are all continuously differentiable and fully convex (that is, convex in (x, y)) while $\psi_1 := G$ is a convex function. Assume that $\Lambda_1 := \mathbb{R}^p_-$, $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$, $\Lambda_2 := \mathbb{R}^p_-$ and there exists $\widetilde{x}$ such that $G_j(\widetilde{x}) < 0$, $j = 1, \ldots, k$. Furthermore consider $\overline{x} \in dom\,\varphi_o$, where S_o is inner semicompact and there exists $\widetilde{y}$ with $g_i(\overline{x}, \widetilde{y}) < 0$, $i = 1, \ldots, p$. Then, $\overline{x}$ is an optimal solution of (P) if and only if for any $y \in S_o(\overline{x})$, there exists $(\alpha, \beta) \in \mathbb{R}^{k+p}$ such that:*

$$0 \in \nabla_x F(\overline{x}, y) + \sum_{i=1}^p \beta_i \nabla_x g_i(\overline{x}, y) + \sum_{j=1}^k \alpha_j \partial G_j(\overline{x}), \tag{3.39}$$

$$\nabla_y F(\overline{x}, y) + \sum_{i=1}^p \beta_i \nabla_y g_i(\overline{x}, y) = 0, \tag{3.40}$$

$$\textit{for } i = 1, \ldots, p,\; \beta_i \geq 0,\; \beta_i g_i(\overline{x}, y) = 0, \tag{3.41}$$

$$\textit{for } j = 1, \ldots, k,\; \alpha_j \geq 0,\; \alpha_j G_j(\overline{x}) = 0. \tag{3.42}$$

Proof. Under the conditions on the sets Ω_2 and Λ_2, the value function φ_o (3.37) takes the form

$$\varphi_o(x) = \min_y \{F(x,y) |\, g_i(x,y) \leq 0,\; i = 1, \ldots, p\}. \tag{3.43}$$

This function is convex since F and g are fully convex. Moreover, it is Lipschitz continuous near $\overline{x}$, since $\overline{x} \in \operatorname{dom} \varphi_o$. Taking into account the fact that the functions G_j, $j = 1, \ldots, k$ are convex, problem (P) is a convex optimization problem. Hence, from a well-known fact in convex optimization, $\overline{x}$ is an optimal solution of (P) if and only if we have inclusion (3.35). In this case φ_o is given in (3.43) while $X := \{x |\, G_j(x) \leq 0,\, j = 1, \ldots, k\}$. Thus we have

$$N_X(x) = \bigcup \Big\{ \sum_{j=1}^{k} \alpha_j \partial G_j(\overline{x}) \,\big|\, \alpha_j \geq 0,\, \alpha_j G_j(\overline{x}) = 0,\, j = 1, \ldots, k \Big\} \tag{3.44}$$

given that there exists $\widetilde{x}$ with $G_j(\widetilde{x}) < 0$, $j = 1, \ldots, k$. It remains to note that since F and g are all continuously differentiable and fully convex, while S is inner semicompact at $\overline{x}$ and there exists $\widetilde{y}$ such that $g_i(\overline{x}, \widetilde{y}) < 0$, $i = 1, \ldots, p$, one has

$$\partial \varphi_o(\overline{x}) = \bigcup_{\beta \in \Lambda(\overline{x}, y)} \Big\{ \nabla_x F(\overline{x}, y) + \sum_{i=1}^{p} \beta_i \nabla_x g_i(\overline{x}, y) \Big\}, \tag{3.45}$$

for all $y \in S_o(\overline{x})$, cf. [21, Theorem 6.6.7]. Here, $\Lambda(\overline{x}, y)$ denotes the set of multipliers β such that conditions (3.40) and (3.41) are satisfied. The result follows from a combination of equations (3.35), (3.44) and (3.45). □

4. The pessimistic bilevel programming problem. In this section, we consider the following pessimistic formulation of the bilevel optimization problem (1.4):

$$(P_p) \quad \begin{cases} \text{minimize} & \varphi_p(x) \text{ subject to } x \in X \\ \text{where} & X := \Omega_1 \cap \psi_1^{-1}(\Lambda_1) \\ & \varphi_p(x) := \max\limits_y \{F(x, y) |\, y \in S(x)\} \\ & S(x) := \{y \in K(x) |\, f(x, y) \leq \varphi(x)\} \\ & \varphi(x) := \min\limits_y \{f(x, y) |\, y \in K(x)\} \\ & K(x) := \{y |\, (x, y) \in \Omega_2 \cap \psi_2^{-1}(\Lambda_2)\} \end{cases}$$

which is obtained by replacing the minimization two-level value function (1.9) in (P_o) by the maximization two-level value function φ_p (1.5). Considering the relationship (1.7) between the maximization and minimization two-level value functions, the Lipschitz continuity of φ_p (1.5) can be obtained from that of φ_p^o (1.8) by using the results of the previous section. The basic subdifferential of φ_p can also be derived from that of φ_p^o by means of the convex hull property (2.4). Proceeding with the latter operation, one should however have in mind that in the convexification process of the upper bound of $\partial \varphi_p^o$, the subdifferential of the lower-level value function φ can be kept unchanged. To make this point more clear, we provide a constructive way to deduce the sensitivity analysis of the maximization two-level value function φ_p (1.5) in the next result. The assumptions that we make are those of Theorem 3.2, Theorem 3.3 and Theorem 3.5, respectively, applied to the minimization two-level value function φ_p^o (1.8), while replacing the solution set-valued mapping S_o (3.4) by its analog S_o^p:

$$S_o^p(x) := \arg\min_y \{-F(x, y) |\, y \in S(x)\} = \{y \in S(x) | F(x, y) + \varphi_p^o(x) \geq 0\}. \tag{4.1}$$

THEOREM 4.1 (sensitivity analysis of the maximization two-level value function). *Consider the maximization two-level value function φ_p (1.5). Then, we have the following assertions:*

(i) *Let all the assumptions of Theorem 3.2 be satisfied for the two-level value function* φ_p^o (1.8) *with the set-valued mapping* S_o (3.4) *replaced by* S_o^p (4.1). *Then, the maximization two-level value function* φ_p (1.5) *is Lipschitz continuous near* $\overline{x}$ *and we have:*

$$\begin{aligned}\partial\varphi_p(\overline{x}) \quad \subseteq \Big\{ & -\textstyle\sum_{t=1}^{n+1}\eta_t x_t^* \big| \sum_{t=1}^{n+1}\eta_t = 1 \text{ and for all } t = 1,\dots,n+1 : \\ & \eta_t \geq 0,\ r_t \geq 0,\ u_t \in N_{\Lambda_2}(\psi_2(\overline{x}, y_t)),\ y_t \in S_o^p(\overline{x}), \\ & (x_t^* + r_t \textstyle\sum_{s=1}^{n+1} v_s x_s^*, 0) \in \partial(-F)(\overline{x}, y_t) + r_t\partial f(\overline{x}, y_t) + \\ & \qquad\qquad \partial\langle u_t, \psi_2\rangle(\overline{x}, y_t) + N_{\Omega_2}(\overline{x}, y_t), \\ & \textstyle\sum_{s=1}^{n+1} v_s = 1 \text{ and for all } s = 1,\dots,n+1 : \\ & v_s \geq 0,\ u_s \in N_{\Lambda_2}(\psi_2(\overline{x}, y_s)),\ y_s \in S(\overline{x}), \\ & (x_s^*, 0) \in \partial f(\overline{x}, y_s) + \partial\langle u_s, \psi_2\rangle(\overline{x}, y_s) + N_{\Omega_2}(\overline{x}, y_s) \Big\}.\end{aligned}$$

(ii) *Let all the assumptions of Theorem 3.3 be satisfied for the two-level value function* φ_p^o (1.8) *with the set-valued mapping* S_o (3.4) *replaced by* S_o^p (4.1). *Then, the maximization two-level value function* φ_p (1.5) *is Lipschitz continuous near* $\overline{x}$ *and we have:*

$$\begin{aligned}\partial\varphi_p(\overline{x}) \quad \subseteq \Big\{ & -\textstyle\sum_{t=1}^{n+1}\eta_t x_t^* \big| \sum_{t=1}^{n+1}\eta_t = 1 \text{ and for all } t = 1,\dots,n+1 : \\ & \eta_t \geq 0,\ (r_t, \beta^t) \in \Lambda^o(\overline{x}, y_t),\ \gamma \in \Lambda(\overline{x}, y_t),\ y_t \in S_o^p(\overline{x}), \\ & x_t^* \in \partial_x(-F)(\overline{x}, y_t) + r_t\big(\partial_x f(\overline{x}, y_t) - \partial_x f(\overline{x}, y_t)\big) + \\ & \qquad\qquad \textstyle\sum_{i=1}^p \beta_i^t \partial_x g_i(\overline{x}, y_t) - r_t \sum_{i=1}^p \gamma_i \partial_x g_i(\overline{x}, y_t) \Big\}.\end{aligned}$$

where $\Lambda(\overline{x}, y_t)$ *and* $\Lambda^o(\overline{x}, y_t)$ *are given in* (3.13) *and* (3.14)*, respectively.*

(iii) *Let all the assumptions of Theorem 3.5 be satisfied for the two-level value function* φ_p^o (1.8) *with the set-valued mapping* S_o (3.4) *replaced by* S_o^p (4.1). *Then, the maximization two-level value function* φ_p (1.5) *is Lipschitz continuous near* $\overline{x}$ *and we have:*

$$\begin{aligned}\partial\varphi_p(\overline{x}) \quad \subseteq \Big\{ & -\textstyle\sum_{t=1}^{n+1}\eta_t x_t^* \big| \sum_{t=1}^{n+1}\eta_t = 1 \text{ and for all } t = 1,\dots,n+1 : \\ & (x_t^* + r_t x_\varphi^*, 0) \in \partial(-F)(\overline{x}, \overline{y}) + r_t\partial f(\overline{x}, \overline{y}) + \textstyle\sum_{i=1}^p \beta_i^t \partial g_i(\overline{x}, \overline{y}), \\ & \eta_t \geq 0,\ r_t \geq 0,\ \beta_i^t \geq 0,\ \beta_i^t g_i(\overline{x}, \overline{y}) = 0,\ i = 1,\dots,p, \\ & (x_\varphi^*, 0) \in \overline{\partial} f(\overline{x}, \overline{y}) + \textstyle\sum_{i=1}^p \gamma_i \overline{\partial} g_i(\overline{x}, \overline{y}), \\ & \gamma_i \geq 0,\ \gamma_i g_i(\overline{x}, \overline{y}) = 0,\ i = 1,\dots,p \Big\}.\end{aligned}$$

Proof. The Lipschitz continuity of φ_p follows obviously from Theorem 3.2, Theorem 3.5 and Theorem 3.3, respectively, while considering equality (1.7). For the upper bound of $\partial\varphi_p$ in (i), note from the proof of Theorem 3.2 that

$$\begin{aligned}\partial\varphi_p^o(\overline{x}) \subseteq \bigcup_{y\in S_o^p(\overline{x})} \ \bigcup_{u\in N_{\Lambda_2}(\psi_2(\overline{x},y))} \ \bigcup_{r\geq 0} \Big\{ x^* \big| (x^*, 0) \in \partial(-F)(\overline{x}, y) + \\ r\partial f(\overline{x}, y) + r\partial(-\varphi)(\overline{x}) \times \{0\} + \partial\langle u, \psi_2\rangle(\overline{x}, y) + N_{\Omega_2}(\overline{x}, y) \Big\}.\end{aligned}$$

Applying the convex hull property (2.4), we get

$$\partial\varphi_p(\overline{x}) \subseteq -\text{co}\,\partial\varphi_p^o(\overline{x}).$$

For $x^* \in \text{co}\,\partial\varphi_p^o(\overline{x})$, one has from Carathédory's theorem that there exist x_t^*, $t = 1,\dots,n+1$ such that $x^* = \sum_{t=1}^{n+1}\eta_t x_t^*$, with $\sum_{t=1}^{n+1}\eta_t = 1$ and for $t = 1,\dots,n+1$, we have $y_t \in S_o^p(\overline{x})$, $\eta_t \geq 0$, $r_t \geq 0$, $u_t \in N_{\Lambda_2}(\psi_2(\overline{x}, y_t))$ and

$$(x_t^*, 0) \in \partial(-F)(\overline{x}, y_t) + r_t \partial f(\overline{x}, y_t) + r_t \partial(-\varphi)(\overline{x}) \times \{0\} + \partial\langle u_t, \psi_2\rangle(\overline{x}, y_t) + N_{\Omega_2}(\overline{x}, y_t).$$

The latter inclusion implies the existence of $x_\varphi^* \in \operatorname{co} \partial\varphi(\overline{x})$ such that

$$(x_t^* + r_t x_\varphi^*, 0) \in \partial(-F)(\overline{x}, y_t) + r_t \partial f(\overline{x}, y_t) + \partial\langle u_t, \psi_2\rangle(\overline{x}, y_t) + N_{\Omega_2}(\overline{x}, y_t).$$

Considering the upper estimate of $\partial\varphi(\overline{x})$ in (3.10), the result of (i) follows. The proofs of (ii) and (iii) follow in a similar way. However, in the latter cases, the upper bounds of $\operatorname{co} \partial\varphi(\overline{x})$ should instead be deduced from (3.15) and (3.19), respectively. □

As first consequence of this result, we derive necessary optimality conditions for the pessimistic bilevel programming problem (P_p).

Corollary 4.2 (optimality conditions for the pessimistic bilevel programming problem). *Let $\overline{x}$ be a local optimal solution of (P_p), where ψ_1 is Lipschitz continuous near $\overline{x}$ and the sets Ω_1 and Λ_1 are closed. Furthermore, let the set-valued mapping Φ^X be calm at $(0, \overline{x})$. Then, the following assertions hold:*

(i) *Let all the assumptions of Theorem 4.1 (i) be satisfied. Then, there exist $y_t \in S_o^p(\overline{x})$, η_t, r_t, u_t with $t = 1, \ldots, n+1$ and $y_s \in S(\overline{x})$, v_s, u_s with $s = 1, \ldots, n+1$ such that relationships* (3.23)-(3.25) *and* (3.27) (*where $r := r_t$*) *together with the following conditions are satisfied:*

$$\sum_{t=1}^{n+1} \eta_t x_t^* \in \partial\langle \alpha, \psi_1\rangle(\overline{x}) + N_{\Omega_1}(\overline{x}), \quad (4.2)$$

$$\text{for } t = 1, \ldots, n+1,\ (x_t^* + r_t \sum_{s=1}^{n+1} v_s x_s^*, 0) \in \partial(-F)(\overline{x}, y_t) + r_t \partial f(\overline{x}, y_t) + \partial\langle u_t, \psi_2\rangle(\overline{x}, y_t) + N_{\Omega_2}(\overline{x}, y_t), \quad (4.3)$$

$$\text{for } t = 1, \ldots, n+1,\ \eta_t \geq 0,\ \sum_{t=1}^{n+1} \eta_t = 1, \quad (4.4)$$

$$\text{for } t = 1, \ldots, n+1,\ u_t \in N_{\Lambda_2}(\psi_2(\overline{x}, y_t)). \quad (4.5)$$

(ii) *Let all the assumptions of Theorem 4.1 (ii) be satisfied. Then, for $y \in S(\overline{x})$, there exist γ and $y_t \in S_o^p(\overline{x})$, η_t, r_t, β^t with $t = 1, \ldots, n+1$ such that relationships* (3.27) (*with $r := r_t$*), (3.30) (*with $y := y_t$*), (3.32) (*with $y := y_t$*), (4.2) *and* (4.4), *together with the following conditions are satisfied:*

$$\text{for } t = 1, \ldots, n+1,\ x_t^* \in \partial_x(-F)(\overline{x}, y_t) + r_t\big(\partial_x f(\overline{x}, y_t) - \partial_x f(\overline{x}, y)\big) + \sum_{i=1}^{p} \beta_i^t \partial_x g_i(\overline{x}, y_t) - r_t \sum_{i=1}^{p} \gamma_i \partial_x g_i(\overline{x}, y), \quad (4.6)$$

$$\text{for } t = 1, \ldots, n+1,\ 0 \in \partial_y F(\overline{x}, y_t) + r_t \partial_y f(\overline{x}, y_t) + \sum_{i=1}^{p} \beta_i^t \partial_y g_i(\overline{x}, y_t), \quad (4.7)$$

$$\text{for } t = 1, \ldots, n+1;\ i = 1, \ldots, p,\ \beta_i^t \geq 0,\ \beta_i^t g_i(\overline{x}, y_t) = 0. \quad (4.8)$$

(iii) *Let all the assumptions of Theorem 4.1 (iii) be satisfied. Then, there exist η_t, r_t, u_t with $t = 1, \ldots, n+1$ such that that relationships* (3.27) (*with $r := r_t$*), (3.32) (*with* $y := \overline{y}$), (3.34), (4.2) *and* (4.4), *together with the following conditions are satisfied:*

$$\text{for } t = 1, \ldots, n+1,\ (x_t^* + r_t x_\varphi^*, 0) \in \partial(-F)(\overline{x}, \overline{y}) + r_t \partial f(\overline{x}, \overline{y}) + \sum_{i=1}^{p} \beta_i^t \partial g_i(\overline{x}, \overline{y}),$$

$$\text{for } t = 1, \ldots, n+1;\ i = 1, \ldots, p,\ \beta_i^t \geq 0,\ \beta_i^t g_i(\overline{x}, \overline{y}) = 0,\ i = 1, \ldots, p.$$

Proof. Combination of (3.35) (where φ_o is replaced by φ_p), (3.36) and the inclusions in Theorem (4.1) (i), (ii) and (iii), respectively. □

REMARK 4.3 (the pessimistic bilevel programming problem in the smooth case). *If we assume that all the functions involved in the pessimistic bilevel programming problem* (P_p) *are continuously differentiable while setting* $\Omega_1 := \mathbb{R}^n$, $\Lambda_1 := \mathbb{R}^k_-$, $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$ *and* $\Lambda_2 := \mathbb{R}^p_-$, *the optimality conditions obtained in [7] are recovered by applying Corollary 4.2 . In the latter paper, where only smooth functions are considered, the generalized equation and Karush-Kuhn-Tucker reformulations of the pessimistic bilevel programming problem are also investigated. These two approach lead to optimality conditions which are completely different from those obtained here.*

REMARK 4.4 (the minmax programming problem). *Another consequence of Theorem 4.1 is that, analogously to the discussion in the previous section, the sensitivity analysis of the classical maximization value function* $\varphi_p(x) := \max_y \{F(x,y) |\, y \in K(x)\}$ *can also be obtained from that of the maximization two-level value* φ_p (4.1), *by simply deleting the terms referring to the lower-level objective function and value function while adjusting the assumptions. Proceeding like this in Corollary 4.2 while setting* $\Omega_1 := \mathbb{R}^n$, $\Lambda_1 := \mathbb{R}^k_-$, $\Omega_2 := \mathbb{R}^n \times \mathbb{R}^m$ *and* $\Lambda_2 := \mathbb{R}^p_-$, *we recover the optimality conditions of the minmax programming problem derived in [24].*

REFERENCES

[1] J.F. Bonnans and A. Shapiro, *Perturbation analysis of optimization problems*, New York: Springer, 2000.

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

[3] S. Dempe, N. Dinh and J. Dutta, *Optimality conditions for a simple convex bilevel programming problem*, R.S. Burachik and J.-C. Yao (eds), Variational analysis and generalized differentiation in optimization and control. In honor of B.S. Mordukhovich. Berlin: Springer, 2010, pp. 149–161.

[4] S. Dempe, J. Dutta and B.S. Mordukhovich, *New necessary optimality conditions in optimistic bilevel programming*, Optimization **56** (2007), no. 5-6, 577–604.

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

[6] S. Dempe and A.B. Zemkoho, *On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem*, Nonlinear Analysis (2011), In press, doi:10.1016/j.na.2011.05.097.

[7] S. Dempe, B.S. Mordukhovich and A.B. Zemkoho, *Necessary optimality conditions in pessimistic bilevel programming*, in preparation.

[8] S. Dempe, B.S. Mordukhovich and A.B. Zemkoho, *Sensitivity analysis for two-level value functions with applications to bilevel programming*, submitted.

[9] A. Dhara and A. Mehra, *Approximate optimality conditions for minimax programming problems*, Optimization **59** (2010), no. 7-8, 1013–1032.

[10] A.V. Fiacco, *Introduction to sensitivity and stability analysis in nonlinear programming*, New York - London: Academic Press, 1983.

[11] J. Gauvin and F. Dubeau, *Differential properties of the marginal function in mathematical programming*, Math. Program. Study **18** (1982), 101–119.

[12] R. Henrion, A. Jourani and J.V. Outrata, *On the calmness of a class of multifunctions*, SIAM J. Optim. **13** (2002), no. 2, 603–618.

[13] B.S. Mordukhovich, *Variational analysis and generalized differentiation. I: Basic theory. II: Applications*, Berlin: Springer, 2006.
[14] B.S. Mordukhovich and N.M. Nam, *Variational stability and marginal functions via generalized differentiation*, Math. Oper. Res. **30** (2005), no. 4, 800–816.
[15] B.S. Mordukhovich, N.M. Nam and N.D. Yen, *Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming*, Optimization **55** (2006), no. 5-6, 685–708.
[16] B.S. Morukhovich, M.N. Nam and H.M. Phan, *Variational analysis of marginal function with applications to bilevel programming problems*, J. Optim. Theory Appl. (2011), to appear.
[17] B.S. Morukhovich, M.N. Nam and N.D. Yen, *Subgradients of marginal functions in parametric mathematical programming*, Math. Program. **116** (2009), no. 1-2, 369–396.
[18] S.M. Robinson, *Some continuity properties of polyhedral multifunctions*, Math. Program. Study **14** (1981), 206–214.
[19] R.T. Rockafellar, *Marginal values and second-order necessary conditions for optimality*, Math. Program. **26** (1983), 245–286.
[20] R.T. Rockafellar and R.J.-B. Wets, *Variational analysis*, Berlin: Springer, 1998.
[21] K. Shimizu, Y. Ishizuka and J.F. Bard, *Nondifferentiable and two-level mathematical programming*, Dordrecht: Kluwer Academic Publishers, 1997.
[22] 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** (2010), no. 4, 1885–1905.
[23] J.J. Ye and D.L. Zhu. *Optimality conditions for bilevel programming problems*, Optimization **33** (1995), no. 1, 9–27 (with Erratum in Optimization **39** (1995) no. 4, 361–366).
[24] A.B. Zemkoho, *A simple approach to optimality conditions in minmax programming*, submitted.