Fakultät für Mathematik und Informatik

Preprint 2012-01

Stephan Dempe, Boris S. Mordukhovich,
Alain B. Zemkoho

Necessary optimality conditions
in pessimistic bilevel programming

ISSN 1433-9307

Stephan Dempe, Boris S. Mordukhovich,
Alain B. Zemkoho

Necessary optimality conditions in pessimistic bilevel programming

TU Bergakademie Freiberg

Fakultät für Mathematik und Informatik

Prüferstraße 9

09596 FREIBERG

http://www.mathe.tu-freiberg.de

ISSN 1433 – 9307

Herausgeber: Dekan der Fakultät für Mathematik und Informatik

Herstellung: Medienzentrum der TU Bergakademie Freiberg

NECESSARY OPTIMALITY CONDITIONS IN PESSIMISTIC BILEVEL PROGRAMMING

December 21, 2011

STEPHAN DEMPE[1], BORIS S. MORDUKHOVICH[2] and ALAIN B. ZEMKOHO[1]

Abstract. This paper is devoted to the so-called pessimistic version of bilevel programming programs. Minimization problems of this type are challenging to handle partly because the corresponding value functions are often merely upper (while not lower) semicontinuous. Employing advanced tools of variational analysis and generalized differentiation, we provide rather general frameworks ensuring the Lipschitz continuity of the corresponding value functions. Several types of *lower subdifferential* necessary optimality conditions are then derived by using the lower-level value function (LLVF) approach and the Karush-Kuhn-Tucker (KKT) representation of lower-level optimal solution maps. We also derive *upper subdifferential* necessary optimality conditions of a new type, which can be essentially stronger than the lower ones in some particular settings. Finally, certain links are established between the obtained necessary optimality conditions for the pessimistic and optimistic versions in bilevel programming.

Keywords: optimization and variational analysis; pessimistic bilevel programs; two-level value functions; sensitivity analysis; generalized differentiation; optimality conditions

Mathematical Subject Classification 2000: 48J53; 90C26; 90C30; 90C31; 90C46

1 Introduction

Bilevel programs belong to problems of hierarchical optimization written in the form

$$\text{“}\min_x\text{”}\{F(x,y)|\, x\in X,\ y\in S(x)\}, \tag{1.1}$$

where the upper-level player (leader) intends to minimize his/her cost function F with respect to the variable x while taking into account the reaction y of the lower-level player (follower). Here $S : X \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a set-valued mapping defined by

$$S(x) := \arg\min_y\{f(x,y)|\, y\in K(x)\}, \tag{1.2}$$

which describes sets of optimal solutions of the lower-level parametric optimization problem

$$\min_y\ f(x,y) \text{ s.t. } y\in K(x) \tag{1.3}$$

for any given choice $x \in X$ of the leader. The sets X and $K(x)$ are usually called upper-level and lower-level feasible sets, respectively. For simplicity we confine ourselves to the case where the upper and lower-level constraint sets are given explicitly as

$$X := \{x\in\mathbb{R}^n|\, G(x)\le 0\} \text{ and } K(x) := \{y\in\mathbb{R}^m|\, g(x,y)\le 0\}, \tag{1.4}$$

[1]Department of Mathematics and Computer Science, TU Bergakademie Freiberg, Freiberg, Germany (dempe@math.tu-freiberg.de, zemkoho@daad-alumni.de). Research of Alain B. Zemkoho was supported by the Deutscher Akademischer Austausch Dienst (DAAD).

[2]Department of Mathematics, Wayne State University, Detroit, USA (boris@math.wayne.edu). Research of this author was also partially supported by the Australian Research Council under grant DP-12092508, by the European Regional Development Fund (FEDER), and by the following Portuguese agencies: Foundation for Science and Technology, Operational Program for Competitiveness Factors, and Strategic Reference Framework under grant PTDC/MAT/111809/2009.

respectively, with $G : \mathbb{R}^n \to \mathbb{R}^k$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$. Furthermore, all the functions involved are assumed to be continuously differentiable. The reader may observe from our analysis that most of the results obtained can be extended to the case of equality and other types of constraints as well as to the case of nonsmooth functions via appropriate subdifferential constructions. Also recall that $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are (single-valued) upper-level and lower-level objective/cost functions, respectively.

One way to understand the general bilevel program (1.1) is to treat it as the following set-valued optimization problem:

$$\min_{x \in X} F(x, S(x)) := \bigcup_{y \in S(x)} \{F(x, y)\}, \tag{1.5}$$

where the minimization is considered with respect to some ordering cone. Such an approach is discussed in [9] while related developments are given in [1, 2, 28] to derive necessary optimality conditions for multiobjective bilevel programs. In order to investigate problem (1.1) from the viewpoint of scalar optimization, observe that (1.5) becomes a usual optimization problem

$$\min\ F(x, S(x)) \text{ with } x \in X \tag{1.6}$$

provided that $S(x)$ is single-valued for all $x \in X$; see [5] and the references therein for details on algorithms and necessary optimality conditions via this approach. It is worth mentioning that the main difficulty in both approaches (1.5) and (1.6) is the implicit nature of the objective.

If we can not ensure the uniqueness of optimal solutions for the lower-level problem (1.3), it has be noted earlier (see, e.g., [5]) that the formulation of problem (1.1) is ambiguous from the view point of scalar-objective optimization. This is the reason for the quotation marks in (1.1). To overcome this, two main approaches have been suggested in the literature. They are the optimistic and the pessimistic reformulations defined respectively as

$$(P_o) \quad \min_{x \in X} \varphi_o(x) := \mathop{\mathbf{min}}_{y} \{F(x, y) |\, y \in S(x)\}, \tag{1.7}$$

$$(P_p) \quad \min_{x \in X} \varphi_p(x) := \mathop{\mathbf{max}}_{y} \{F(x, y) |\, y \in S(x)\}. \tag{1.8}$$

The optimistic formulation (P_o) is much simpler to handle and has therefore been the most investigated, while mainly for its simplified version:

$$\min_{x,y} F(x, y) \text{ s.t. } x \in X,\ y \in S(x), \tag{1.9}$$

where the difficulty in (P_o) is shifted to the constraints. More details on the latter problem can be found in the books [5, 25] and the annotated bibliography [6]. For more recent results on the topic we refer the reader to the papers [8, 11, 12, 20, 29] and the bibliographies therein.

It is easy to construct examples of bilevel programs where (P_o) has an optimal solution while (P_p) does not have any. Even when both problems have optimal solutions, they may differ from each other; see, e.g., [5, 13] for more discussions. To the best of our knowledge, a detailed investigation on necessary optimality conditions for the original optimistic formulation (P_o) is conducted in [10] for the first time in the literature. It has been well recognized that the optimistic and pessimistic reformulations of the bilevel program (1.1) are optimization problems with objectives of the marginal/value function type. In fact, in both cases (P_o) and (P_p) the objectives are described by *two-level value function*, i.e., as the optimal value function of a parametric optimization problem partly constrained by another one.

As mentioned above, little attention has been given to the pessimistic bilevel program (1.8) also known as weak Stackelberg problem. For some algorithmic issues in particular classes of problem (P_p) we refer the reader to [3, 13, 14] and the bibliographies therein. We are not familiar with any work on optimality conditions, apart from [4], where several reformulations of (P_p) were suggested and necessary conditions are derived in terms of the coderivative of certain set-valued mappings in some special cases. In [5, Chapter 3] some optimality conditions are suggested for the linear pessimistic bilevel programming problem while assuming among other things that the functions φ_o and φ_p coincide at the reference point.

One major difficulty to handle the pessimistic problem relates to the fact that the objective function φ_p in (1.8) is usually only upper semicontinuous (u.s.c.), which makes it hard to detect optimal solutions by using the conventional "lower subdifferential" objects that work well for lower semicontinuous (l.s.c.) functions. Secondly, observe that (P_p) is a special minimax problem. Although many publications have been devoted to optimality conditions for the latter class of problem, most of them can not be applied to our pessimistic program because the corresponding inner problem $\max_y\{F(x,y)|\, y \in S(x)\}$ violates the imposed constraint qualifications (CQs).

In this paper we develop two approaches to establish necessary optimality conditions for pessimistic bilevel programs. Our first approach to derive *lower* (i.e., conventional) subdifferential optimality conditions for (P_p) is to get the lower semicontinuity of the two-level value function φ_p by constructing frameworks where it is actually locally Lipschitz continuous. This would be possible by applying the results obtained in [10] for the two-level value function approach of the optimistic problem (P_o). Thus we obtain various types of first-order necessary optimality conditions via the lower-level value function (LLVL) and Karush-Kuhn-Tucker (KKT) reformulations of φ_p. The descriptions of these terms are given in Sections 3 and 4, respectively.

The second approach considered in Section 5 seems to be more suitable for particular structures of the pessimistic program. Namely, we derive the so-called *upper* subdifferential optimality conditions for (P_p) in the sense of [17], which do not just require the function φ_p to be l.s.c., but eventually leads to much stronger necessary optimality conditions in certain important settings.

The rest of the paper is organized as follows. Section 2 presents some basic constructions and properties of variational analysis and generalized differentiation needed for formulations and proofs of the main results. In Sections 3 and 4 we derive necessary optimality conditions for (P_p), which essentially fell into the following two categories:

(i) Those with the usual structure of known conditions in minimax programming, where the partial gradients of the upper-level objective function F are included in the convex combinations of a certain family of gradients of the Lagrangians.

(ii) Those we label as the "KKT-type" conditions without any convex combination of the gradients of the functions involved in (P_p).

The last class of conditions of the upper subdifferential type is derived in Section 5. We finally establish clear and useful links between the necessary optimality conditions of (P_p) and (P_o) and then present the concluding remarks in Section 6.

Throughout the paper we assume that $S(x) \neq \emptyset$ whenever $x \in X$ for the lower-level solution sets (1.2) and further use the notation

$$\begin{aligned} S_o(x) &:= \arg\min_y\{F(x,y)|\, y \in S(x)\} \neq \emptyset, \\ S_p(x) &:= \arg\max_y\{F(x,y)|\, y \in S(x)\} \neq \emptyset. \end{aligned} \tag{1.10}$$

2 Tools of Variational Analysis

More details on the material briefly discussed in this section can be found in the books [18, 24] and their references. First recall that for a set-valued mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the expression

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

signifies the *Painlevé-Kuratowski outer/upper limit* of Ψ at $\bar{x}$. Given an extended-real-valued function $\psi:\mathbb{R}^n\to\overline{\mathbb{R}}:=(-\infty,\infty]$ and a point $\bar{x}\in\operatorname{dom}\psi:=\{x\in\mathbb{R}^n|\,\psi(x)<\infty\}$, the *regular/Frećhet subdifferential* of ψ at $\bar{x}$ is given by

$$\widehat{\partial}\psi(\bar{x}):=\Big\{v\in\mathbb{R}^n\Big|\ \liminf_{x\to\bar{x}}\frac{\psi(x)-\psi(\bar{x})-\langle v,x-\bar{x}\rangle}{\|x-\bar{x}\|}\ge 0\Big\},\tag{2.2}$$

while the *regular upper subdifferential* (or superdifferential) of ψ at $\bar{x}$ is

$$\widehat{\partial}^+\psi(\bar{x}):=-\widehat{\partial}(-\psi)(\bar{x}).\tag{2.3}$$

The *limiting/Mordukhovich subdifferential* of ψ at $\bar{x}$ (our basic construction here) is defined via the outer limit (2.1) of the regular subgradients (2.2) by

$$\partial\psi(\bar{x}):=\mathop{\rm Limsup}_{x\xrightarrow{\psi}\bar{x}}\widehat{\partial}\psi(x),\tag{2.4}$$

where the symbol $x\xrightarrow{\psi}\bar{x}$ means that $x\to\bar{x}$ with $\psi(x)\to\psi(\bar{x})$. The *upper* counterpart $\partial^+\psi(\bar{x})$ of (2.4) is defined similarly to (2.3). Provided further that ψ is Lipschitz continuous around $\bar{x}\in\operatorname{dom}\psi$, its *convexified/Clarke subdifferential* (or *generalized gradient*) at $\bar{x}$ can be defined by

$$\overline{\partial}\psi(\bar{x}):=\operatorname{co}\partial\psi(\bar{x}),\tag{2.5}$$

where "co" stands for the convex hull of the set in question. We have the following *convex hull property* for locally Lipschitzian functions that plays a significant role in this paper:

$$\operatorname{co}\partial(-\psi)(\bar{x})=-\operatorname{co}\partial\psi(\bar{x}).\tag{2.6}$$

The function ψ is said to be *lower* (resp. *upper*) *regular* at $\bar{x}$ if

$$\widehat{\partial}\psi(\bar{x})=\partial\psi(\bar{x})\quad\big(\text{resp. }\widehat{\partial}^+\psi(\bar{x})=\partial^+\psi(\bar{x})\big).\tag{2.7}$$

Obviously ψ is upper regular if and only if $-\psi$ is lower regular at the point in question. For convex functions ψ all the three (lower) subdifferential constructions $\widehat{\partial}\psi(\bar{x})$, $\partial\psi(\bar{x})$, and $\overline{\partial}\psi(\bar{x})$ reduce to the classical subdifferential of convex analysis. Furthermore, the upper constructions $\widehat{\partial}^+\psi(\bar{x})$ and $\partial^+\psi(\bar{x})$ agree with the superdifferential of concave functions, while the upper counterpart $\overline{\partial}\psi(\bar{x})$ of ψ is not different from its lower version (2.5) for any locally Lipschitzian functions.

Given further a nonempty set $\Omega\subset\mathbb{R}^n$, our basic *normal cone* to Ω at $\bar{x}\in\Omega$ corresponding to the subdifferential (2.4) is defined by

$$N_\Omega(\bar{x}):=\mathop{\rm Limsup}_{x\xrightarrow{\Omega}\bar{x}}\widehat{N}_\Omega(x),\tag{2.8}$$

where $x\xrightarrow{\Omega}\bar{x}$ stands for $x\to\bar{x}$ with $x\in\Omega$, and where

$$\widehat{N}_\Omega(\bar{x}):=\Big\{v\in\mathbb{R}^n\Big|\ \limsup_{x\xrightarrow{\Omega}\bar{x}}\frac{\langle v,x-\bar{x}\rangle}{\|x-\bar{x}\|}\le 0\Big\}.\tag{2.9}$$

When Ω is locally closed around $\bar{x}$, the normal cone (2.8) reduces to

$$N_\Omega(\bar{x})=\mathop{\rm Limsup}_{x\to\bar{x}}\big[\operatorname{cone}(x-\Pi_\Omega(x))\big],$$

where the symbol "cone" stands for the conic hull of the corresponding set, and where Π denotes the Euclidean projection on the set in question. This was in fact the original definition of the normal cone in [15]. The convexified normal cone is defined via the convex closure of (2.8) by

$$\overline{N}_\Omega(\bar{x}):=\operatorname{clco}N_\Omega(\bar{x}).\tag{2.10}$$

Next let $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping with the graph

$$\operatorname{gph}\Psi := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m |\ y \in \Psi(x)\}.$$

Following [16], we define the *coderivative* for Ψ at $(\bar{x},\bar{y}) \in \operatorname{gph}\Psi$ by

$$D^*\Psi(\bar{x},\bar{y})(v) := \{u \in \mathbb{R}^n |\ (u,-v) \in N_{\operatorname{gph}\Psi}(\bar{x},\bar{y})\} \text{ for } v \in \mathbb{R}^m \tag{2.11}$$

via the normal cone (2.8) to the graph of Ψ. The regular coderivative $\widehat{D}^*\Psi(\bar{x},\bar{y})$ and convexified coderivative $\overline{D}^*\Psi(\bar{x},\bar{y})$ of Ψ at $(\bar{x},\bar{y})$ are defined similarly by replacing the basic normal cone $N_{\operatorname{gph}\Psi}(\bar{x},\bar{y})$ by its counterparts (2.9) and (2.10), respectively.

We conclude this section with some continuity properties of set-valued mappings that are of a particular interest in this paper. A mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *inner semicompact* at a point $\bar{x}$ with $\Psi(\bar{x}) \neq \emptyset$ if for every sequence $x_k \to \bar{x}$ with $\Psi(x_k) \neq \emptyset$ there is a sequence of $y_k \in \Psi(x_k)$ that contains a convergent subsequence. This property surely holds if the sets $\Psi(x)$ are uniformly bounded around $\bar{x}$. This mapping is *inner semicontinuous* at $(\bar{x},\bar{y}) \in \operatorname{gph}\Psi$ if for every sequence $x_k \to \bar{x}$ there is a sequence of $y_k \in \Psi(x_k)$ that converges to $\bar{y}$. Further, Ψ is *calm* at $(\bar{x},\bar{y})$ if there exist neighborhoods U of $\bar{x}$, V of $\bar{y}$ and a constant $\ell > 0$ such that

$$\Psi(x) \cap V \subset \Psi(\bar{x}) + \ell\|x-\bar{x}\|\mathbb{B} \text{ for all } x \in U \tag{2.12}$$

with $\mathbb{B}$ denoting the closed unit ball in $\mathbb{R}^m$. Finally, a set-valued mapping $\Psi : D \subset \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ admits a *local upper Lipschitzian selection* at $(\bar{x},\bar{y})$ if there is a single-valued mapping $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$, which is local upper Lipschitzian at $\bar{x}$ satisfying $f(\bar{x}) = \bar{y}$ and $f(x) \in \Psi(x)$ for all $x \in D$ in a neighborhood of $\bar{x}$. The upper Lipschitzian property of f is understood in the sense of Robinson [23]: there is a neighborhood U of $\bar{x}$ and a constant $\ell \geq 0$ such that

$$\|f(x) - f(\bar{x})\| \leq \ell\|x-\bar{x}\| \text{ for all } x \in U \cap D.$$

It is easy to see that the latter reduces to calmness (2.12) in the case of single-valued mappings.

3 LLVF Approach to Pessimistic Bilevel Programs

We start this section by noting that the lower-level solution map S in (1.2) can be described by

$$S(x) = \{y |\ g(x,y) \leq 0,\ f(x,y) - \varphi(x) \leq 0\} \text{ with } \varphi(x) := \min_y\{f(x,y) |\ g(x,y) \leq 0\}, \tag{3.1}$$

where φ denotes the *lower-level value function* (LLVF), i.e., the value function of the lower-level optimization problem (1.3). Thus we have the following *LLVF reformulation* for the maximization two-level value function φ_p (1.8):

$$\varphi_p(x) = \max_y\{F(x,y) |\ g(x,y) \leq 0,\ f(x,y) - \varphi(x) \leq 0\}. \tag{3.2}$$

In order to apply the results obtained in [10] for the optimistic problem (P_o), we consider the minimization two-level value function as follows:

$$\varphi_{op}(x) = \min_y\{-F(x,y) |\ g(x,y) \leq 0,\ f(x,y) - \varphi(x) \leq 0\}. \tag{3.3}$$

The classical interplay between the minimization and maximization of a given objective function induces the equality $\varphi_p(x) = -\varphi_{op}(x)$. The solution map associated with φ_{op} from (3.3) is

$$S_{op}(x) := \{x \in S(x) |\ F(x,y) + \varphi_{op}(x) \geq 0\}. \tag{3.4}$$

To proceed, recall the *constraint qualifications* introduced in [10] to derive necessary optimality conditions for the optimistic bilevel program via the LLVF approach:

(A_1^o) The mapping $\Xi(\vartheta) := \{(x,y)|\, g(x,y) \leq 0,\ f(x,y) - \varphi(x) \leq \vartheta\}$ is calm at $(0, \bar{x}, \bar{y})$.

This CQ is automatically satisfied if, e.g., the functions f and g are linear in (x, y) [10]. More details on (A_1^o) and other sufficient conditions ensuring its fulfilment can be found in [10] and the references therein. Next we provide another CQ, which is needed to ensure the local *Lipschitz continuity* of φ_{op} and hence of φ_p:

(A_2^o) $\big[(\lambda, \beta) \in \Lambda_y^o(\bar{x}, \bar{y}, 0),\ x^* \in \partial(-\varphi)(\bar{x})\big] \Longrightarrow \lambda x^* = -\lambda \nabla_x f(\bar{x}, \bar{y}) - \nabla_x g(\bar{x}, \bar{y})^\top \beta,$

where $\Lambda_y^o(\bar{x}, \bar{y}, v)$ for $v \in \mathbb{R}^m$ denotes the set of multipliers

$$\begin{aligned} \Lambda_y^o(\bar{x}, \bar{y}, v) := \ & \{(\lambda, \beta)|\, \lambda \geq 0,\ \beta \geq 0,\ \beta^\top g(\bar{x}, \bar{y}) = 0, \\ & \qquad v + \lambda \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^\top \beta = 0\}. \end{aligned} \tag{3.5}$$

Some frameworks where the latter condition holds are presented in [10].

To close this list of assumptions, we formulate the lower-level and upper-level regularity conditions that are used in the whole paper. The pair $(\bar{x}, \bar{y}) \in \operatorname{gph} K$ (resp. $\bar{x} \in X := \{x|\, G(x) \leq 0\}$) is *lower-level regular* (resp. *upper-level regular*) if there is no nonzero vector u (resp. α) such that

$$\nabla_y g(\bar{x}, \bar{y})^\top u = 0,\ u \geq 0, u^\top g(\bar{x}, \bar{y}) = 0\, (\text{resp. } \nabla G(\bar{x})^\top \alpha = 0, \alpha \geq 0, \alpha^\top G(\bar{x}) = 0).$$

The lower-level regularity (resp. upper-level regularity) conditions correspond to the well-known Mangasarian-Fromovitz CQ for the parametric lower-level problem (1.3) (resp. upper-level constraint system $G(x) \leq 0$). Let us mention that the terms "lower-level regularity" and "upper-level regularity" are borrowed from the paper [8] and no confusion should be made with the notion "lower regularity" and "upper regularity" (2.7) defined in the previous section.

In the spirit of [10] we use the prefixes "KM" and "KN" in this section to designate the optimality conditions for the pessimistic bilevel program (P_p) obtained via the LLVF reformulation (3.2) while assuming the inner semico**n**tinuity and inner semico**m**pactness of the mapping S_{op} from (3.4), respectively. The KM and/or KN prefixes are preceded by the "KKT-type" if no number is attached to the gradients of the upper-level objective function F. A similar terminology (i.e., analogously to [10]) is used in the next section to facilitate comparisons between the necessary optimality conditions for (P_o) and (P_p).

Theorem 3.1 (KM-stationarity conditions for pessimistic bilevel programs). *Let $\bar{x}$ be a local optimal solution to problem (P_p) that is assumed to be upper-level regular. Suppose also that the mapping S_{op} in* (3.4) *is inner semicompact at $\bar{x}$, that the lower-level regularity is satisfied at $(\bar{x}, y)$ for all $y \in S(\bar{x})$, and that the constraint qualifications (A_1^o) and (A_2^o) hold at $(\bar{x}, y)$ for all $y \in S_{op}(\bar{x})$. Then there exist $\alpha \in \mathbb{R}^k$, $(\beta_s, \gamma_t, \lambda_s, v_s, w_t) \in \mathbb{R}^{2p+3}$, $y_s \in S_{op}(\bar{x})$, and $y_t \in S(\bar{x})$ with*

$s = 1, \ldots, n+1$ *and* $t = 1, \ldots, n+1$ *such that we have the conditions:*

$$\sum_{s=1}^{n+1} v_s\big(\nabla_x F(\bar{x}, y_s) + \lambda_s \nabla_x f(\bar{x}, y_s) + \nabla_x g(\bar{x}, y_s)^\top \beta_s\big)$$

$$-\big(\sum_{s=1}^{n+1} v_s\lambda_s\big) \sum_{t=1}^{n+1} w_t\big(\nabla_x f(\bar{x}, y_t) + \nabla_x g(\bar{x}, y_t)^\top \gamma_t\big) + \nabla G(\bar{x})^\top \alpha = 0, \tag{3.6}$$

$$\forall s = 1, \ldots, n+1,\ \nabla_y F(\bar{x}, y_s) + \lambda_s \nabla_y f(\bar{x}, y_s) + \nabla_y g(\bar{x}, y_s)^\top \beta_s = 0, \tag{3.7}$$

$$\forall t = 1, \ldots, n+1,\ \nabla_y f(\bar{x}, y_t) + \nabla_y g(\bar{x}, y_t)^\top \gamma_t = 0, \tag{3.8}$$

$$\forall s = 1, \ldots, n+1,\ \lambda_s \leq 0,\ \beta_s \leq 0,\ \beta_s^\top g(\bar{x}, y_s) = 0, \tag{3.9}$$

$$\forall t = 1, \ldots, n+1,\ \gamma_t \geq 0,\ \gamma_t^\top g(\bar{x}, y_t) = 0, \tag{3.10}$$

$$\alpha \geq 0,\ \alpha^\top G(\bar{x}) = 0, \tag{3.11}$$

$$\sum_{s=1}^{n+1} v_s = 1,\ \forall s = 1, \ldots, n+1,\ v_s \geq 0, \tag{3.12}$$

$$\sum_{t=1}^{n+1} w_t = 1,\ \forall t = 1, \ldots, n+1,\ w_t \geq 0. \tag{3.13}$$

Proof. Apart from the upper-level regularity condition imposed on $\bar{x}$, all the other assumptions of the theorem imply that the value function φ_{op} in (3.3) is Lipschitz continuous around $\bar{x}$; cf. [10, Theorem 5.9(ii)]. Hence the function φ_p is also Lipschitz continuous around $\bar{x}$. It then follows from [18, Proposition 5.3] that

$$0 \in \partial\varphi_p(\bar{x}) + N_X(\bar{x}). \tag{3.14}$$

Using the upper-level regularity of $\bar{x}$, we get the equality

$$N_X(\bar{x}) = \{\nabla G(\bar{x})^\top \alpha |\, \alpha \geq 0,\ \alpha^\top G(\bar{x}) = 0\}. \tag{3.15}$$

The convex hull property (2.6) allows us to find a vector $\alpha \in \mathbb{R}^k$ satisfying the complementarity conditions (3.11) and the inclusion

$$\nabla G(\bar{x})^\top \alpha \in \operatorname{co} \partial\varphi_{op}(\bar{x}), \tag{3.16}$$

Applying now the classical Carathéodory's theorem to (3.16) gives us $v_s,\ s = 1, \ldots, n+1$, satisfying (3.12) and $x_s^* \in \partial\varphi_{op}(\bar{x})$, $s = 1, \ldots, n+1$, and such that

$$\nabla G(\bar{x})^\top \alpha = \sum_{s=1}^{n+1} v_s x_s^*. \tag{3.17}$$

On the other hand, we also have from [10, Theorem 5.9(ii)] that

$$\partial\varphi_{op}(\bar{x}) \subset \bigcup_{y \in S_{op}(\bar{x})} \bigcup_{(\lambda,\beta)\in\Lambda_y^o(\bar{x},y)} \big\{ -\nabla_x F(\bar{x}, y) + \lambda(\nabla_x f(\bar{x}, y) + \partial(-\varphi)(\bar{x})) + \nabla_x g(\bar{x}, y)^\top \beta \big\},$$

where $\Lambda_y^o(\bar{x}, y) := \Lambda_y^o(\bar{x}, y, -\nabla_y F(\bar{x}, y))$ with $\Lambda_y^o(\bar{x}, y, -\nabla_y F(\bar{x}, y))$ given by (3.5). The inclusion $x_s^* \in \partial\varphi_{op}(\bar{x})$ clearly implies that there exist $y_s \in S_{op}(\bar{x})$ and $(\lambda_s, \beta_s) \in \mathbb{R}^{1+p}$, with $s = 1, \ldots, n+1$ such that the following relationships hold:

$$\forall s = 1, \ldots, n+1,\ -\nabla_y F(\bar{x}, y_s) + \lambda_s \nabla_y f(\bar{x}, y_s) + \nabla_y g(\bar{x}, y_s)^\top \beta_s = 0, \tag{3.18}$$

$$\forall s = 1, \ldots, n+1,\ \lambda_s \geq 0,\ \beta_s \geq 0,\ \beta_s^\top g(\bar{x}, y_s) = 0, \tag{3.19}$$

$$\forall s = 1, \ldots, n+1,\ x_s^* + \nabla_x F(\bar{x}, y_s) - \lambda_s \nabla_x f(\bar{x}, y_s) - \nabla_x g(\bar{x}, y_s)^\top \beta_s \in \lambda_s \partial(-\varphi)(\bar{x}). \tag{3.20}$$

Remind at this point that the inner semicompactness of the mapping S in (3.1) (ensured by that of S_{op} in (3.4)) and the fulfilment of the lower-level regularity at $(\bar{x}, y)$ for all $y \in S(\bar{x})$ imply the local Lipschitz continuity of the function $-\varphi$ at $\bar{x}$; see [19]. Furthermore, we have an upper estimate of $\partial(-\varphi)(\bar{x})$ as in [10, Theorem 5.3]. This allows us to verify that inclusion (3.20) implies the existence of $y_t \in S(\bar{x})$, $\gamma_t \in \mathbb{R}^p$ and $w_t \in \mathbb{R}$ as $t = 1, \ldots, n+1$ such that relationships (3.8), (3.10). and (3.13) hold together with

$$\begin{aligned} x_s^* = \; & -\nabla_x F(\bar{x}, y_s) - \lambda_s \nabla_x f(\bar{x}, y_s) - \nabla_x g(\bar{x}, y_s)^\top \beta_s \\ & + \lambda_s \textstyle\sum_{s=1}^{n+1} w_t (\nabla_x f(\bar{x}, y_t) + \nabla_x g(\bar{x}, y_t)^\top \gamma_t). \end{aligned}$$

Substituting the values of x_s^* into (3.17) while multiplying the resulting equation together with (3.18) and (3.19) by -1, we arrive at the optimality conditions (3.6), (3.7), and (3.9). This completes the proof of the theorem. □

Under the additional convexity assumptions on the function involved in the lower-level problem (1.3) we get the following useful consequence.

Corollary 3.2 (KM-stationarity conditions for convex pessimistic programs). *Let $\bar{x}$ be local optimal solution to problem (P_p) under the assumptions of Theorem* 3.1. *Assume in addition that the functions f and g_i, $i = 1, \ldots, p$, are convex in (x, y). Then there exist $\alpha \in \mathbb{R}^k$, $(\beta_s, \gamma_s, \lambda_s, v_s) \in \mathbb{R}^{2p+2}$, and $y_s \in S_{op}(\bar{x})$ with $s = 1, \ldots, n+1$ such that the optimality conditions* (3.7), (3.9), (3.11), *and* (3.12) *are satisfied together with*

$$\sum_{s=1}^{n+1} v_s \Big(\nabla_x F(\bar{x}, y_s) + \nabla_x g(\bar{x}, y_s)^\top (\beta_s - \lambda_s \gamma_s) \Big) + \nabla G(\bar{x})^\top \alpha = 0, \tag{3.21}$$

$$\forall s = 1, \ldots, n+1, \; \nabla_y f(\bar{x}, y_s) + \nabla_y g(\bar{x}, y_s)^\top \gamma_s = 0, \tag{3.22}$$

$$\forall s = 1, \ldots, n+1, \; \gamma_s \geq 0, \; \gamma_s^\top g(\bar{x}, y_s) = 0. \tag{3.23}$$

Proof. It follows from Theorem 3.1 with the additional observation: whenever $y \in S(\bar{x})$ we have

$$\partial\varphi(\bar{x}) \subset \Big\{ \nabla_x f(\bar{x}, y) + \nabla_x g(\bar{x}, y)^\top \gamma \Big| \begin{array}{l} \nabla_x f(\bar{x}, y) + \nabla_x g(\bar{x}, y)^\top \gamma = 0, \\ \gamma \geq 0, \; \gamma^\top g(\bar{x}, y) = 0 \end{array} \Big\}$$

by taking into account the assumed lower-level regularity and the full convexity of f and g_i for all $i = 1, \ldots, p$. This allows us to complete the proof of the corollary while choosing $y := y_s \in S_{op}(\bar{x}) \subset S(\bar{x})$ as $s = 1, \ldots n+1$. □

It is important, in order to eliminate some difficulties in numerical algorithms to solve the problem, that the gradient of the objective function F not be involved in the convex combinations summations in (3.6) and (3.21). To get results of the latter type, we replace the inner semicompactness of the mapping S_{op} in the above Theorem 3.1 by the inner semicontinuity, which leads us to the following result.

Theorem 3.3 (KKT-type KN-stationarity conditions for pessimistic programs). *Let $\bar{x}$ be a local optimal solution to problem (P_p), where upper-level regularity is satisfied at $\bar{x}$ and the lower-level regularity is satisfied at $(\bar{x}, \bar{y})$. Assume that the mapping S_{op} in* (3.4) *is inner semicontinuous at $(\bar{x}, \bar{y})$ and that the constraint qualifications (A_1^o) and (A_2^o) hold at $(\bar{x}, \bar{y})$. Furthermore, suppose that the set $co\, N_{gph\, S}(\bar{x}, \bar{y})$ is closed. Then there exist $\alpha \in \mathbb{R}^k$ and $(\beta_s, \gamma_s, \lambda_s, v_s) \in \mathbb{R}^{2p+2}$,*

$s = 1, \dots, n+m+1$ *such that the optimality condition* (3.11) *holds together with*

$$\nabla_x F(\bar{x},\bar{y}) + \nabla G(\bar{x})^\top \alpha + \sum_{s=1}^{n+m+1} v_s \nabla_x g(\bar{x},\bar{y})^\top (\beta_s - \lambda_s \gamma_s) = 0, \tag{3.24}$$

$$\nabla_y F(\bar{x},\bar{y}) + \sum_{s=1}^{n+m+1} v_s (\lambda_s \nabla_y f(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top \beta_s) = 0, \tag{3.25}$$

$$\forall s = 1, \dots, n+m+1,\ \nabla_y f(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top \gamma_s = 0, \tag{3.26}$$

$$\forall s = 1, \dots, n+m+1,\ \lambda_s \le 0,\ \beta_s \le 0,\ \beta_s^\top g(\bar{x},\bar{y}) = 0, \tag{3.27}$$

$$\forall s = 1, \dots, n+m+1,\ \gamma_s \ge 0,\ \gamma_s^\top g(\bar{x},\bar{y}) = 0, \tag{3.28}$$

$$\sum_{i=1}^{n+m+1} v_s = 1,\ \forall s = 1, \dots, n+m+1,\ v_s \ge 0. \tag{3.29}$$

Proof. Similarly to the proof of the previous theorem we deduce from [10, Theorem 5.9(i)] that the function φ_{op} is Lipschitz continuous around $\bar{x}$. Also the same interplay between φ_{op} and φ_p implies that there exists a vector $\alpha \in \mathbb{R}^k$ such that (3.11) and the following inclusion hold with the convexified subdifferential of φ_{op}:

$$\nabla G(\bar{x})^\top \alpha \in \operatorname{co} \partial \varphi_{op}(\bar{x}) = \overline{\partial} \varphi_{op}(\bar{x}). \tag{3.30}$$

Since the upper-level objective function F is continuously differentiable and the solution map S_{op} is inner semicontinuous at $(\bar{x},\bar{y})$, we get from [20, Theorem 5.1] that

$$\overline{\partial} \varphi_{op}(\bar{x}) \subset -\nabla_x F(\bar{x},\bar{y}) + \overline{D}^* S(\bar{x},\bar{y})(-\nabla_y F(\bar{x},\bar{y})). \tag{3.31}$$

Combining (3.30)-(3.31) and using the coderivative definition give us

$$\begin{bmatrix} \nabla G(\bar{x})^\top \alpha + \nabla_x F(\bar{x},\bar{y}) \\ \nabla_y F(\bar{x},\bar{y}) \end{bmatrix} \in \overline{N}_{\text{gph}S}(\bar{x},\bar{y}) = \operatorname{co} N_{\text{gph}S}(\bar{x},\bar{y}), \tag{3.32}$$

where the last equality results from the closedness of the set $\operatorname{co} N_{\text{gph}}(\bar{x},\bar{y})$.

On the other hand, we derive from [10, Theorem 5.7] that, under the constraint qualification (A_1^o) and the assumptions on the lower-level regularity and the inner semicontinuity of S_{op}, the basic normal cone to the graph of S can be estimated as

$$N_{\text{gph}S}(\bar{x},\bar{y}) \subset \left\{ \begin{bmatrix} \nabla_x g(\bar{x},\bar{y})^\top (\beta - \lambda\gamma) \\ \lambda \nabla_y f(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top \beta \end{bmatrix} \middle| \begin{array}{r} \nabla_y f(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top \gamma = 0 \\ \lambda \ge 0,\ \beta \ge 0,\ \beta^\top g(\bar{x},\bar{y}) = 0 \\ \gamma \ge 0,\ \gamma^\top g(\bar{x},\bar{y}) = 0 \end{array} \right\}.$$

Substituting the inclusion obtained into (3.32) and applying Carathéodory's theorem complete the proof of the claimed optimality conditions. □

Observed that the derived KN-stationarity conditions of the KKT-type agree with the optimality conditions obtained in Theorem 3.1 provided that $S_{op}(\bar{x}) = \{\bar{y}\} = S(\bar{x})$ and $\Lambda(\bar{x},\bar{y}) = \{\gamma\}$ in the latter result while $\Lambda_y^o(\bar{x},\bar{y}) = \{(\lambda,\beta)\}$ in both theorems.

4 KKT Approach to Pessimistic Bilevel Programs

There are two other (closely related while somewhat different) approaches to transform the minimizing two-level value function

$$\varphi_{op}(x) := \min_y \{-F(x,y) |\, y \in S(x)\} \tag{4.1}$$

into the value function for a known class of optimization problem. They are: the primal Karush-Kuhn-Tucker (or generalized equation) reformulation

$$\varphi_{op}(x) = \min_y \{-F(x,y) | \, 0 \in \nabla_y f(x,y) + N_{K(x)}(y)\} \tag{4.2}$$

and the following classical KKT reformulation, which gives us the value function of a special mathematical program with complementarity constraints (MPCC):

$$\varphi_{op}(x) = \min_{y,u} \{-F(x,y) | u \in \Lambda(x,y), \, g(x,y) \leq 0\}, \tag{4.3}$$

$$\begin{array}{l} \text{where } \Lambda(x,y) := \{u | \, \mathcal{L}(x,y,u) = 0, \, u \geq 0, \, u^\top g(x,y) = 0\} \\ \text{with } \mathcal{L}(x,y,u) := \nabla_y f(x,y) + \nabla_y g(x,y)^\top u. \end{array} \tag{4.4}$$

For both transformations we need the functions $f(x,.)$ and $g(x,.)$ to be *convex* whenever $x \in X$. For the latter reformulation we additionally need an appropriate CQ, say the lower-level regularity defined in the previous section. Under such a CQ, the normal cone $N_{K(x)}(y)$ in (4.2) reduces to

$$N_{K(x)}(y) = \{\nabla_y g(x,y)^\top u | \, u \geq 0, \, u^\top g(x,y) = 0\}.$$

Thus we get the value function (4.3) from that of (4.2).

The underlying goal of this section is to derive necessary optimality/stationarity conditions for the pessimistic bilevel program (P_p) by using either the primal reformulation or the classical KKT reformulation of the two-level value function φ_{op} in (4.1). Considering the combinatorial nature of the complementarity system

$$u \geq 0, \, g(x,y) \leq 0, \, u^\top g(x,y) = 0, \tag{4.5}$$

the corresponding techniques from related fields of optimization theory suggest the following partition of the indices associated to the functions involved in the latter relationship:

$$\begin{array}{lll} \eta := & \eta(\bar{x},\bar{y},\bar{u}) := \{i = 1,\ldots,p | & \bar{u}_i = 0, \, g_i(\bar{x},\bar{y}) < 0\}, \\ \theta := & \theta(\bar{x},\bar{y},\bar{u}) := \{i = 1,\ldots,p | & \bar{u}_i = 0, \, g_i(\bar{x},\bar{y}) = 0\}, \\ \nu := & \nu(\bar{x},\bar{y},\bar{u}) := \{i = 1,\ldots,p | & \bar{u}_i > 0, \, g_i(\bar{x},\bar{y}) = 0\}. \end{array} \tag{4.6}$$

They lead to various types of stationarity concepts for the pessimistic bilevel program, which are completely different from those derived in the previous section. The functions f and g defining the lower-level problem (1.3) are assumed to be *twice continuously differentiable* throughout this section and also in the subsequent one.

To proceed, we recall that the *M(ordukhovich)-type* multipliers set associated with the above KKT reformulations of our two-level value function are defined by

$$\begin{array}{rl} \Lambda^{cm}(\bar{x},\bar{y},\bar{u},v) & := \{(\beta,\gamma) | v + \nabla g(\bar{x},\bar{y})^\top \beta + \nabla_{x,y} \mathcal{L}(\bar{x},\bar{y},\bar{u})^\top \gamma = 0, \\ & \qquad \nabla_y g_\nu(\bar{x},\bar{y})\gamma = 0, \beta_\eta = 0, \\ & \qquad (\nabla_y g_i(\bar{x},\bar{y})\gamma > 0 \, \wedge \, \beta_i > 0) \vee \beta_i(\nabla_y g_i(\bar{x},\bar{y})\gamma) = 0, \forall i \in \theta\}. \end{array} \tag{4.7}$$

The *C(larke)-type* multiplier set $\Lambda^{cc}(\bar{x},\bar{y},\bar{u},v)$ is defined analogously by replacing the relationship

$$(\nabla_y g_i(\bar{x},\bar{y})\gamma > 0 \, \wedge \, \beta_i > 0) \, \vee \, \beta_i(\nabla_y g_i(\bar{x},\bar{y})\gamma) = 0, \, \forall i \in \theta \tag{4.8}$$

in the right-hand-side of equality (4.7) by

$$\beta_i(\nabla_y g_i(\bar{x},\bar{y})\gamma) \geq 0, \, \forall i \in \theta. \tag{4.9}$$

Also of interest are the multipliers sets in the form $\Lambda_y^{c\ltimes}(\bar{x},\bar{y},\bar{u},v)$ (with $\ltimes = m, c$), where the gradients of the functions g and $\mathcal{L}$ in $\Lambda^{c\ltimes}(\bar{x},\bar{y},\bar{u},v)$ are considered only with respect to the lower-level variable y. The above multiplier sets will essentially help us to simplify the presentation of CQs needed in this section. However, M-, C- and *S(trong)-type* (see below) stationarity conditions are our main targets here. Most often the M-stationarity conditions are identified by

$$(\nabla_y g_i(\bar{x},\bar{y})\gamma < 0 \,\wedge\, \beta_i < 0) \,\vee\, \beta_i(\nabla_y g_i(\bar{x},\bar{y})\gamma) = 0, \,\forall i \in \theta.$$

Note the difference in the signs of $\nabla_y g_i(\bar{x},\bar{y})\gamma$ and β_i, which are the opposite to those in (4.8). As for the C-stationarity conditions, they are still distinguished by (4.9) while S-stationarity conditions are recognized by

$$\nabla_y g_i(\bar{x},\bar{y})\gamma \leq 0 \,\wedge\, \beta_i \leq 0, \,\forall i \in \theta. \tag{4.10}$$

4.1 M-Type Stationarity Conditions

Let us now state the following rules taken from [10] that allow us to derive the necessary optimality conditions for the pessimistic bilevel program (P_p) via the KKT reformulation of the two-level value function φ_{op} from (4.3):

(A_1^m) $\quad(\beta,\gamma) \in \Lambda^{cm}(\bar{x},\bar{y},\bar{u},0) \Longrightarrow \beta = 0,\, \gamma = 0,$
(A_2^m) $\quad(\beta,\gamma) \in \Lambda_y^{cm}(\bar{x},\bar{y},\bar{u},0) \Longrightarrow \nabla_x g(\bar{x},\bar{y})^\top\beta + \nabla_x\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0,$
(A_3^m) $\quad(\beta,\gamma) \in \Lambda_y^{cm}(\bar{x},\bar{y},\bar{u},0) \Longrightarrow \beta = 0,\, \gamma = 0.$

We obviously have the implications $(A_1^m) \Longleftarrow (A_3^m) \Longrightarrow (A_2^m)$.

To further simplify the presentation, consider the feasible set-valued mapping

$$S^h(x) := \{(y,u)|\, \mathcal{L}(x,y,u) = 0,\, u \geq 0,\, g(x,y) \leq 0,\, u^\top g(x,y) = 0\} \tag{4.11}$$

and the solution map related to the two-level value function (4.3):

$$S_{op}^h(x) := \{(y,u) \in S^h(x)|\, F(x,y) + \varphi_{op}(x) \geq 0\}. \tag{4.12}$$

Theorem 4.1 (M-type stationarity conditions for pessimistic programs). *Let $\bar{x}$ be a local optimal solution for (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x},y)$ for all $y \in S(\bar{x})$. Assume that $\bar{x}$ is upper-level regular, that the mapping S_{op}^h in* (4.12) *is inner semicompact at $\bar{x}$, and that CQs (A_1^m) and (A_2^m) hold at $(\bar{x},y,u)$ whenever $(y,u) \in S_{op}^h(\bar{x})$. Then there exist $\alpha \in \mathbb{R}^k$, $(\beta^s,\, \gamma^s,\, v_s) \in \mathbb{R}^{p+m+1}$, and $(y_s,u_s) \in S_{op}^h(\bar{x})$ with $s = 1,\ldots,n+1$ such that condition* (3.11) *holds together with the following relationships:*

$$\sum_{i=1}^{n+1} v_s\Big(\nabla_x F(\bar{x},y_s) + \nabla_x g(\bar{x},y_s)^\top\beta^s + \nabla_x\mathcal{L}(\bar{x},y_s,u_s)^\top\gamma^s\Big) + \nabla G(\bar{x})^\top\alpha = 0, \tag{4.13}$$

$$\forall s = 1,\ldots,n+1,\, \nabla_y F(\bar{x},y_s) + \nabla_y g(\bar{x},y_s)^\top\beta^s + \nabla_y\mathcal{L}(\bar{x},y_s,u_s)^\top\gamma^s = 0, \tag{4.14}$$

$$\forall s = 1,\ldots,n+1,\, \nabla_y g_{\nu^s}(\bar{x},y_s)\gamma^s = 0,\, \beta_{\eta^s}^s = 0, \tag{4.15}$$

$$\forall s = 1,\ldots,n+1, i \in \theta^s,\, (\beta_i^s < 0 \wedge \nabla_y g_i(\bar{x},y_s)\gamma^s < 0) \vee \beta_i^s(\nabla_y g_i(\bar{x},y_s)\gamma^s) = 0, \tag{4.16}$$

$$\sum_{i=1}^{n+1} v_s = 1,\, \forall s = 1,\ldots,n+1,\, v_s \geq 0, \tag{4.17}$$

$$\forall s = 1,\ldots,n+1,\, \eta^s := \eta(\bar{x},y_s,u_s), \theta^s := \theta(\bar{x},y_s,u_s), \nu^s := \nu(\bar{x},y_s,u_s). \tag{4.18}$$

Proof. By the counterpart of [10, Theorem 5.2], under the inner semicompactness of S^h_{op} in (4.12), the value function φ_{op} is Lipschitz continuous around $\bar{x}$ and we have

$$\partial\varphi_{op}(\bar{x}) \subset \bigcup_{(\bar{y},\bar{u})\in S^h_{op}(\bar{x})} \bigcup_{(\beta,\gamma)\in\Lambda^{cm}_y(\bar{x},\bar{y},\bar{u})} \Big\{ -\nabla_x F(\bar{x},\bar{y}) + \nabla_x g(\bar{x},\bar{y})^\top \beta + \nabla_x \mathcal{L}(\bar{x},\bar{y},\bar{u})^\top \gamma \Big\}.$$

Hence φ_p in (1.8) is also Lipschitz continuous around $\bar{x}$, and Carathédory's theorem yields

$$\begin{aligned}\partial\varphi_p(\bar{x}) \subset \ \Big\{ &\textstyle\sum_{s=1}^{n+1} v_s\big(\nabla_x F(\bar{x},y_s) - \nabla_x g(\bar{x},y_s)^\top \beta^s - \nabla_x \mathcal{L}(\bar{x},y_s,u_s)^\top \gamma^s\big) \big| \\ & (y_s,u_s)\in S^h_{op}(\bar{x}),\ (\beta^s,\gamma^s)\in\Lambda^{cm}_y(\bar{x},y_s,u_s),\ s=1,\ldots,n+1, \\ & \textstyle\sum_{i=1}^{n+1} v_s = 1,\ v_s\geq 0,\ s=1,\ldots,n+1\Big\},\end{aligned} \tag{4.19}$$

where the multiplier set $\Lambda^{cm}_y(\bar{x},\bar{y},\bar{u})$ is defined as $\Lambda^{cm}_y(\bar{x},\bar{y},\bar{u},-\nabla_y F(\bar{x},\bar{y}))$, i.e.,

$$\begin{aligned}\Lambda^{cm}_y(\bar{x},\bar{y},\bar{u}) \ := \{(\beta,\gamma)|\ & \nabla_y g_\nu(\bar{x},\bar{y})\gamma = 0, \beta_\eta = 0, \\ & (\nabla_y g_i(\bar{x},\bar{y})\gamma > 0 \wedge \beta_i > 0) \vee \beta_i(\nabla_y g_i(\bar{x},\bar{y})\gamma) = 0,\ \forall i\in\theta, \\ & -\nabla_y F(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top\beta + \nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0\}.\end{aligned}$$

The result of the theorem is then derived by the combination of (3.14), (3.15), and (4.19). □

As discussed in Section 3, it is important particularly for numerical reasons to obtain optimality conditions with no "multiplier" attached to the gradient of the upper-level objective function. To reach this goal in the case of M-type stationarity, our first approach is based on the KKT reformulation (4.3) of the two-level value function. However, assumption (A^m_1) in the previous theorem is replaced by the following one:

$$(A^m_1)' \quad \left.\begin{array}{r} \nabla g(\bar{x},\bar{y})^\top \sum_{s=1}^{2p+1}\mu_s v^s + \nabla_{x,y}\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0 \\ \sum_{s=1}^{2p+1}\mu_s u^s + \nabla_y g(\bar{x},\bar{y})\gamma = 0 \\ \forall s=1,\ldots,2p+1,\ u^s_\nu = 0,\ v^s_\eta = 0 \\ \forall s=1,\ldots,2p+1, i\in\theta,\ (u^s_i > 0\wedge v^s_i > 0)\vee u^s_i v^s_i = 0 \\ \sum_{s=1}^{2p+1}\mu_s = 1,\ \forall s=1,\ldots,2p+1,\ \mu_s\geq 0 \end{array}\right\} \Longrightarrow \left\{\begin{array}{l}\gamma = 0 \\ \sum_{s=1}^{2p+1}\mu_s v^s = 0\end{array}\right.$$

while imposing instead the inner semicontinuity of the solution map S^h_{op} in (4.12).

Theorem 4.2 (**M-stationarity conditions of the KKT-type for pessimistic bilevel programs, I).** *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is assumed to be convex and lower-level regular at $(\bar{x},y)$ for all $y\in S(\bar{x})$. Assume further that $\bar{x}$ is upper regular, that S^h_{op} in* (4.12) *is inner semicontinuous at $(\bar{x},\bar{y},\bar{u})$, and that the constraint qualifications $(A^m_1)'$ and (A^m_2) are satisfied. Furthermore we suppose that the set $\text{co}N_\Lambda(\bar{u},-g(\bar{x},\bar{y}))$ (with $\Lambda := \{(a,b)\in\mathbb{R}^{2p}|\, a,b\geq 0,\ a^\top b = 0\}$) is closed. Then there exist $(\alpha,\gamma)\in\mathbb{R}^{k+m}$ and $(u^s,v^s,\mu_s)\in\mathbb{R}^{2p+1}$ with $s=1,\ldots,2p+1$ such that condition* (3.11) *holds together with*

$$\nabla_x F(\bar{x},\bar{y}) + \nabla G(\bar{x})^\top\alpha + \nabla_x g(\bar{x},\bar{y})^\top\sum_{s=1}^{2p+1}\mu_s v^s + \nabla_x\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0, \tag{4.20}$$

$$\nabla_y F(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top\sum_{s=1}^{2p+1}\mu_s v^s + \nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0, \tag{4.21}$$

$$\sum_{s=1}^{2p+1}\mu_s u^s + \nabla_y g(\bar{x},\bar{y})\gamma = 0, \tag{4.22}$$

$$\forall s=1,\ldots,2p+1,\ u^s_\nu = 0,\ v^s_\eta = 0, \tag{4.23}$$

$$\forall s=1,\ldots,2p+1, i\in\theta,\ (u^s_i > 0\wedge v^s_i > 0)\vee u^s_i v^s_i = 0, \tag{4.24}$$

$$\sum_{s=1}^{2p+1}\mu_s = 1,\ \forall s=1,\ldots,2p+1,\ \mu_s\geq 0. \tag{4.25}$$

Proof. It follows the path of that of Theorem 3.3. Obviously the difference lies in the upper estimate of the normal cone to the graph of the mapping S^h. To proceed, first note that since the new variable u does not appear in the upper-level objective function F, after the transformation of φ_{op} obtained in (4.3), inclusion (3.32) should be replaced by

$$\begin{bmatrix} \nabla G(\bar{x})^\top \alpha + \nabla_x F(\bar{x}, \bar{y}) \\ \nabla_y F(\bar{x}, \bar{y}) \\ 0 \end{bmatrix} \in \overline{N}_{\mathrm{gph} S^h}(\bar{x}, \bar{y}, \bar{u}) \subset \nabla \psi(\bar{x}, \bar{y}, \bar{u})^\top \overline{N}_{\{0_m\} \times \Lambda}(\psi(\bar{x}, \bar{y}, \bar{u})), \tag{4.26}$$

where the last inclusion follows from [20, Theorem 5.2(ii)] provided that

$$\ker \nabla \psi(\bar{x}, \bar{y}, \bar{u})^\top \cap \overline{N}_{\{0_m\} \times \Lambda}(\psi(\bar{x}, \bar{y}, \bar{u})) = \{0\}. \tag{4.27}$$

Here the function ψ and the set Λ are defined respectively as

$$\psi(x, y, u) := [\mathcal{L}(x, y, u), (u, -g(x, y))] \ \text{ and } \ \Lambda := \{(a, b) \in \mathbb{R}^{2p} |\, a, b \geq 0, \, a^\top b = 0\}. \tag{4.28}$$

From the well-known expression of the basic normal cone to Λ (see, e.g., [27]) we get

$$N_\Lambda(\bar{u}, -g(\bar{x}, \bar{y})) = \left\{ (u, v) \,\middle|\, \begin{array}{l} u_\eta = 0, v_\nu = 0, \\ \forall i \in \theta, (u_i < 0 \wedge v_i < 0) \wedge u_i v_i = 0 \end{array} \right\}. \tag{4.29}$$

Employing Carathéodory's theorem then gives us

$$\operatorname{co} N_\Lambda(\bar{u}, -g(\bar{x}, \bar{y})) = \left\{ \begin{bmatrix} \sum_{s=1}^{2p+1} \mu_s u^s \\ \sum_{s=1}^{2p+1} \mu_s v^s \end{bmatrix} \,\middle|\, \begin{array}{l} \sum_{s=1}^{2p+1} \mu_s = 1, \, \forall s = 1, \ldots, 2p+1, \, \mu_s \geq 0, \\ u_\eta^s = 0, v_\nu^s = 0, \; s = 1, \ldots, 2p+1, \\ (u_i^s < 0 \wedge v_i^s < 0) \wedge u_i^s v_i^s = 0, \, \forall i \in \theta, \; s = 1, \ldots, 2p+1 \end{array} \right\}. \tag{4.30}$$

Since the set $\operatorname{co} N_\Lambda(\bar{u}, -g(\bar{x}, \bar{y}))$ is closed, we have

$$\overline{N}_\Lambda(\bar{u}, -g(\bar{x}, \bar{y})) = \operatorname{co} N_\Lambda(\bar{u}, -g(\bar{x}, \bar{y})). \tag{4.31}$$

Inserting thus the right-hand side of equality (4.30) into (4.27) yields the constraint qualification $(A_1^m)'$. Repeating this process with inclusion (4.26), we arrive at the necessary optimality conditions asserted in the theorem.

It should be mentioned that to obtain the Lipschitz continuity of φ_{op}, we need to observe that condition (4.27) implies the validity of

$$\ker \nabla \psi(\bar{x}, \bar{y}, \bar{u})^\top \cap N_{\{0_m\} \times \Lambda}(\psi(\bar{x}, \bar{y}, \bar{u})) = \{0\}. \tag{4.32}$$

However, it is a simple exercise to check that the latter condition is equivalent to (A_1^m). Then we can derive from [10] that the value function φ_{op} is Lipschitz continuous around $\bar{x}$ provided that the constraint qualification (A_2^m) and the inner semicontinuity of S_{op}^h are satisfied at $(\bar{x}, \bar{y}, \bar{u})$. This completes the proof of the theorem. □

In the next result we obtain M-stationarity conditions of the KKT-type from the viewpoint of the primal KKT reformulation (4.2). Instead of the inner semicontinuity of the mapping S_{op}^h in (4.12), we impose the inner semicontinuity of the solution map

$$S_{op}^e(x) := \{y \in S^e(x) |\, F(x, y) + \varphi_{op}(x) \geq 0\}, \tag{4.33}$$

where S^e denotes the feasible solution map for the parametric optimization problem related to the value function in (4.2), i.e., defined by

$$S^e(x) := \{y |\, 0 \in \nabla_y f(x, y) + N_{K(x)}(y)\}. \tag{4.34}$$

The following theorem requires more regularity assumptions while some of them are closely related to those used in the previous result:

(A_1^e) $\forall \bar{u} \in \Lambda(\bar{x}, \bar{y}):\ [\nabla g(\bar{x}, \bar{y})^\top \beta = 0,\ \beta_\eta = 0] \Longrightarrow \beta = 0,$
(A_2^e) $\forall \bar{u} \in \Lambda(\bar{x}, \bar{y}) : (\beta, \gamma) \in \Lambda^{cm}(\bar{x}, \bar{y}, \bar{u}, 0) \Longrightarrow \beta = 0,\ \gamma = 0,$
(A_3^e) $[\bar{u} \in \Lambda(\bar{x}, \bar{y}),\ (\beta, \gamma) \in \Lambda_y^{cm}(\bar{x}, \bar{y}, \bar{u}, 0)] \Longrightarrow \nabla_x g(\bar{x}, \bar{y})^\top \beta + \nabla_x \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma = 0,$
(A_4^e) $[\bar{u} \in \Lambda(\bar{x}, \bar{y}),\ (\beta, \gamma) \in \Lambda_y^{cm}(\bar{x}, \bar{y}, \bar{u}, 0)] \Longrightarrow \beta = 0,\ \gamma = 0.$

Recall that the multiplier set $\Lambda(\bar{x}, \bar{y})$ is given in (4.4). Also note the implications $(A_3^e) \Longleftarrow (A_4^e) \Longrightarrow (A_2^e)$ and $(A_4^e) \Longrightarrow (A_1^e)$.

Theorem 4.3 (M-stationarity conditions of the KKT-type for the pessimistic bilevel programs, II). *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is assumed to be convex. Assume that the lower-level regularity is satisfied at $(\bar{x}, \bar{y})$, that the upper-level regularity is satisfied at $\bar{x}$, and that the mapping S_{op}^e in* (4.33) *is inner semicontinuous at $(\bar{x}, \bar{y})$. Furthermore, let the constraint qualifications (A_1^e), (A_2^e), and (A_3^e) hold at $(\bar{x}, \bar{y})$, and let the set $\text{co}N_{gph\, S^e}(\bar{x}, \bar{y})$ be closed. Then there exist $\alpha \in \mathbb{R}^k$, $(\beta^s, u_s, \gamma^s, v_s) \in \mathbb{R}^{2p+m+1}$, with $s = 1, \ldots, m+n+1$ such that relationship* (3.11) *holds together with the following conditions:*

$$
\begin{aligned}
\nabla_x F(\bar{x}, \bar{y}) + \nabla G(\bar{x})^\top \alpha + \sum_{i=1}^{m+n+1} v_s\big(\nabla_x g(\bar{x}, \bar{y})^\top \beta^s + \nabla_x \mathcal{L}(\bar{x}, \bar{y}, u_s)^\top \gamma^s\big) = 0,\\
\nabla_y F(\bar{x}, \bar{y}) + \sum_{s=1}^{m+n+1} v_s\big(\nabla_y g(\bar{x}, \bar{y})^\top \beta^s + \nabla_y \mathcal{L}(\bar{x}, \bar{y}, u_s)^\top \gamma^s\big) = 0,\\
\forall s = 1, \ldots, m+n+1,\ \nabla_y g_{\nu^s}(\bar{x}, \bar{y})\gamma^s = 0,\ \beta_{\eta^s}^s = 0,\\
\forall s = 1, \ldots, m+n+1, i \in \theta^s,\ (\beta_i^s < 0 \wedge \nabla_y g_i(\bar{x}, \bar{y})\gamma^s < 0) \vee \beta_i^s(\nabla_y g_i(\bar{x}, \bar{y})\gamma^s) = 0,\\
\sum_{i=1}^{m+n+1} v_s = 1,\ \forall s = 1, \ldots, m+n+1,\ v_s \geq 0,\\
\forall s = 1, \ldots, m+n+1, \eta^s := \eta(\bar{x}, \bar{y}, u_s), \theta^s := \theta(\bar{x}, \bar{y}, u_s), \nu^s := \nu(\bar{x}, \bar{y}, u_s).
\end{aligned}
$$

Proof. Since we work here with the primal KKT reformulation (4.2) of the two-level value function φ_{op}, it suffices to observe that the following upper estimate of the normal cone to the graph of S^e (4.34) holds:

$$
\begin{aligned}
N_{\text{gph}\, S^e}(\bar{x}, \bar{y}) \subset \Big\{ & \nabla g(\bar{x}, \bar{y})^\top \beta + \nabla_{x,y} \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma \Big| \\
& \bar{u} \geq 0,\ \bar{u}^\top g(\bar{x}, \bar{y}) = 0,\ \mathcal{L}(\bar{x}, \bar{y}, \bar{u}) = 0,\\
& \nabla_y g_\nu(\bar{x}, \bar{y})\gamma = 0, \beta_\eta = 0,\\
& (\nabla_y g_i(\bar{x}, \bar{y})\gamma > 0 \ \wedge\ \beta_i > 0) \ \vee\ \beta_i(\nabla_y g_i(\bar{x}, \bar{y})\gamma) = 0,\ \forall i \in \theta \Big\}.
\end{aligned}
$$

Indeed, it is implied by the convexity of problem (1.3) due to the assumptions (A_1^e) and (A_2^e); see [22]. The rest of the proof follows the lines in the proof of Theorem 3.3 by taking into account the sensitivity analysis of φ_{op} presented in [10, Theorem 4.1]. □

To close this subsection, note that the optimality conditions of Theorem 4.1 could also be obtained by using the primal KKT reformulation (4.2) of the two-level value function. In this case assumptions (A_1^e), (A_2^e) and (A_3^e) should replace those in (A_1^m) and (A_2^m), and instead the inner semicompactness of S_{op}^e in (4.33) is needed.

4.2 C-Type Stationarity Conditions

In this subsection the following C-qualification conditions are imposed:

(A_1^c) $(\beta, \gamma) \in \Lambda^{cc}(\bar{x}, \bar{y}, \bar{u}, 0) \Longrightarrow \beta = 0,\ \gamma = 0$
(A_2^c) $(\beta, \gamma) \in \Lambda_y^{cc}(\bar{x}, \bar{y}, \bar{u}, 0) \Longrightarrow \nabla_x g(\bar{x}, \bar{y})^\top \beta + \nabla_x \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma = 0$
(A_3^c) $(\beta, \gamma) \in \Lambda_y^{cc}(\bar{x}, \bar{y}, \bar{u}, 0) \Longrightarrow \beta = 0,\ \gamma = 0.$

As in the case of M-qualification conditions we have the implications $(A_1^c) \Longleftarrow (A_3^c) \Longrightarrow (A_2^c)$.

Theorem 4.4 (**C-stationarity conditions for pessimistic bilevel programs).** *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x}, y)$, for all $y \in S(\bar{x})$. Assume that $\bar{x}$ is upper-level regular, that the mapping S_{op}^h in* (4.12) *is inner semicompact at $\bar{x}$, and that the constraint qualifications (A_1^c) and (A_2^c) hold at $(\bar{x}, y, u)$ whenever $(y, u) \in S_{op}^h(\bar{x})$. Then there exist $\alpha \in \mathbb{R}^k$, $(\beta^s, \gamma^s, v_s) \in \mathbb{R}^{p+m+1}$, and $(y_s, u_s) \in S_{op}^h(\bar{x})$ with $s = 1, \ldots, n+1$ such that relationships* (4.13)–(4.15) *and* (4.17)–(4.18) *are satisfied together with*

$$\forall s = 1, \ldots, n+1, \, \forall i \in \theta^s, \, \beta_i^s(\nabla_y g_i(\bar{x}, y_s)\gamma^s) \geq 0.$$

Proof. It is similar to the proof of Theorem 4.1; see [10, Subsection 5.1] for the corresponding upper estimate of the subdifferential $\partial\varphi_{op}(\bar{x})$ in this case. □

The next theorem is a C-counterpart of Theorem 4.2.

Theorem 4.5 (**C-stationarity conditions of the KKT-type for the pessimistic bilevel programs).** *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x}, y)$ for all $y \in S(\bar{x})$. Assume that $\bar{x}$ is upper-level regular, the mapping S_{op}^h in* (4.12) *is inner semicontinuous at $(\bar{x}, \bar{y}, \bar{u})$, and the constraint qualifications (A_1^c) and (A_2^c) are satisfied. Then there exists $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that relationships* (3.11) *and* (4.9) *hold together with the following conditions:*

$$\nabla_x F(\bar{x}, \bar{y}) + \nabla G(\bar{x})^\top \alpha + \nabla_x g(\bar{x}, \bar{y})^\top \beta + \nabla_x \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma = 0, \tag{4.35}$$

$$\nabla_y F(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^\top \beta + \nabla_y \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma = 0, \tag{4.36}$$

$$\nabla_y g_\nu(\bar{x}, \bar{y})\gamma = 0, \, \beta_\eta = 0. \tag{4.37}$$

Proof. It follows the lines in the proof of Theorem 4.2. Observe to this end that the graph of the mapping S^h in (4.11) is represented as

$$\operatorname{gph} S^h = \{(x, y, u) | \, \mathcal{L}(x, y, u) = 0, \, \min\{u_i, -g_i(x, y)\} = 0, \, i = 1, \ldots, p\}.$$

Applying [20, Proposition 5.8] gives us the upper estimate

$$\overline{N}_{\operatorname{gph} S^h}(\bar{x}, \bar{y}, \bar{u}) \subset \left\{ \begin{bmatrix} \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma \\ \nabla_y g_\nu(\bar{x}, \bar{y})\gamma \end{bmatrix} + \sum_{i=1}^{p} \lambda_i \overline{\partial} V_i(\bar{x}, \bar{y}, \bar{u}) \right\} \tag{4.38}$$

provided that the following qualification condition

$$0 \in \begin{bmatrix} \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma \\ \nabla_y g_\nu(\bar{x}, \bar{y})\gamma \end{bmatrix} + \sum_{i=1}^{p} \lambda_i \overline{\partial} V_i(\bar{x}, \bar{y}, \bar{u}) \Longrightarrow \gamma = 0, \, \lambda_i = 0, \, i = 1, \ldots, p, \tag{4.39}$$

is satisfied, where $V_i(x, y, u) := \min\{u_i, -g_i(x, y)\}$ for $i = 1, \ldots, p$. Furthermore, we get from the proof of [10, Theorem 3.6] that

$$\sum_{i=1}^{p} \lambda_i \overline{\partial} V_i(\bar{x}, \bar{y}, \bar{u}) \subset \left\{ \begin{bmatrix} -\nabla_y g(\bar{x}, \bar{y})^\top \beta \\ \xi \end{bmatrix} \middle| \begin{array}{l} \beta_\eta = 0, \, \xi_\nu = 0, \\ \beta_i \xi_i \geq 0, \, \forall i \in \theta \end{array} \right\}. \tag{4.40}$$

Combining (4.39) and (4.40) shows that (A_1^c) is a sufficient condition for the validity of implication (4.39). Then the necessary optimality conditions of the theorem are derived by successively inserting the relationships in (4.38) and (4.40) into the inclusion

$$\begin{bmatrix} \nabla G(\bar{x})^\top \alpha + \nabla_x F(\bar{x}, \bar{y}) \\ \nabla_y F(\bar{x}, \bar{y}) \\ 0 \end{bmatrix} \in \overline{N}_{\operatorname{gph} S^h}(\bar{x}, \bar{y}, \bar{u})h$$

which is obtained similarly to the proof of Theorem 4.2. □

4.3 S-Type Stationarity Conditions

To establish stationary conditions of the S-type, we need the following additional constraint qualification, which is a version of the Partial MPEC Linear Independence CQ (Partial MPEC-LICQ) tailored for our problem:

$$(A_1^s) \qquad \left.\begin{array}{r}\nabla_{x,y}\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma+\nabla g(\bar{x},\bar{y})^\top\beta=0\\ \nabla_y g_\nu(\bar{x},\bar{y})\gamma=0,\ \beta_\eta=0\end{array}\right\}\Longrightarrow \beta_\theta=0,\ \nabla_y g_\theta(\bar{x},\bar{y})\gamma=0.$$

Recall that the Partial MPEC-LICQ was introduced in [26], where the reader can find more details on its original version and further developments.

Theorem 4.6 (**S-stationarity conditions for pessimistic bilevel programs).** *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x}, y)$ for all $y \in S(\bar{x})$. Assume $\bar{x}$ is upper-level regular, that the mapping S^h_{op} in* (4.12) *is inner semicompact at $\bar{x}$, and that either* $[(A_1^m)\wedge(A_2^m)]$ *or* $[(A_1^c)\wedge(A_2^c)]$ *holds at $(\bar{x}, y, u)$ whenever $(y,u) \in S^h_{op}(\bar{x})$. Suppose in addition that* (A_1^s) *also holds at $(\bar{x}, y, u)$ for all $(y, u) \in S^h_{op}(\bar{x})$. Then there exist $\alpha \in \mathbb{R}^k$, $(\beta^s, \gamma^s, v_s) \in \mathbb{R}^{p+m+1}$, and $(y_s, u_s) \in S^h_{op}(\bar{x})$ with $s = 1,\ldots,n+1$ such that relationships* (4.13)–(4.15) *and* (4.17)–(4.18) *are satisfied together with the condition*

$$\forall s = 1,\ldots,n+1,\ \forall i \in \theta^s,\ \beta_i^s \le 0 \wedge \nabla_y g_i(\bar{x}, y_s)\gamma^s \le 0.$$

Proof. We have from the counterpart of [10, Theorem 5.3], with inner semicompactness of S^h_{op} (4.12), that under the convexity of problem (1.3) and the lower-level regularity at $(\bar{x}, y)$ for all $y \in S(\bar{x})$ the following upper estimate of the limiting subdifferential of φ_{op} holds:

$$\partial\varphi_{op}(\bar{x}) \subset \bigcup_{(\bar{y},\bar{u})\in S^h_{op}(\bar{x})}\ \bigcup_{(\beta,\gamma)\in\Lambda^{cs}_y(\bar{x},\bar{y},\bar{u})} \big\{-\nabla_x F(\bar{x},\bar{y})+\nabla_x g(\bar{x},\bar{y})^\top\beta+\nabla_x\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma\big\}. \quad (4.41)$$

provided that either assumptions $[(A_1^m)\wedge(A_1^s)]$ or $[(A_1^c)\wedge(A_1^s)]$ are satisfied at $(\bar{x}, y, u)$, for all $(y, u) \in S^h_{op}(\bar{x})$. Here the multiplier set $\Lambda^{cs}(\bar{x},\bar{y},\bar{u})$ is defined by

$$\begin{array}{rl}\Lambda^{cs}(\bar{x},\bar{y},\bar{u}) & := \{(\beta,\gamma)|\ \nabla_y g_\nu(\bar{x},\bar{y})\gamma=0, \beta_\eta=0\\ & \qquad \nabla_y g_i(\bar{x},\bar{y})\gamma\le 0\ \wedge\ \beta_i\le 0,\ \forall i\in\theta\\ & \qquad -\nabla_y F(\bar{x},\bar{y})+\nabla_y g(\bar{x},\bar{y})^\top\beta+\nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma=0\}.\end{array} \quad (4.42)$$

Recall that assumption (A_2^m) or (A_2^c) is added to ensure the Lipschitz continuity of φ_{op}. The optimality conditions are now derived by proceeding as in the proof of Theorem 4.1 while applying Carathédory's theorem to the right-hand side of inclusion (4.41). □

In Theorem 4.6, the combination of assumptions (A_1^m) and (A_2^m) on one hand and of (A_1^c) and (A_2^c) on the other can be replaced by the single assumption (A_3^m) and (A_3^c), respectively. Next we provide a more general setting of S-stationarity conditions for the pessimistic problem that coincide with those of Theorem 4.6 provided the vector $(\beta, \gamma, \bar{y}, \bar{u})$ is unique in the latter result. The following occur however without requiring the aforementioned uniqueness. To proceed, we need the following well-know MPEC-LICQ condition tailored for our special constraint system described by the set-valued mapping S^h in (4.11):

$$(A_2^s) \qquad \left.\begin{array}{r}\nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma+\nabla_y g(\bar{x},\bar{y})^\top\beta=0\\ \nabla_y g_\nu(\bar{x},\bar{y})\gamma=0,\ \beta_\eta=0\end{array}\right\}\Longrightarrow \beta=0,\ \gamma=0.$$

The validity of this condition obviously implies that (A_1^s) is satisfied. This allows us to derive the next theorem on S-stationarity in pessimistic programs.

Theorem 4.7 (S-stationarity conditions of the KKT-type for pessimistic bilevel programs). *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x}, y)$ for all $y \in S(\bar{x})$. Assume that $\bar{x}$ is upper-level regular, that the mapping S^h_{op} in* (4.12) *is inner semicontinuous at $(\bar{x}, \bar{y}, \bar{u})$, and that either $[(A^c_1) \wedge (A^c_2) \wedge (A^s_1)]$, $[(A^c_3) \wedge (A^s_1)]$ or $[(A^s_2)]$ hold at the same point. Then there is a triple $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that the optimality conditions* (3.11)*,* (4.10)*, and* (4.35)*–*(4.37) *are satisfied.*

Proof. First note that if, in addition to the convexity of (1.3) and the lower-level regularity at $(\bar{x}, y)$ as $y \in S(\bar{x})$, either the assumptions $[(A^c_1)\wedge(A^c_2)]$ or $[(A^c_3)]$ are satisfied, we have the necessary optimality conditions of Theorem 4.5. Furthermore, the fulfillment of (A^1_s) implies the validity of the asserted S-stationarity conditions.

Secondly, if instead (A^s_2) holds, we get from [10, Corollary 3.1] that the two-level value function φ_{op} (4.3) is strictly differentiable at $\bar{x}$ and

$$\nabla\varphi_{op}(\bar{x}) = -\nabla_x F(\bar{x}, \bar{y}) + \nabla G(\bar{x})^\top \alpha + \nabla_x g(\bar{x}, \bar{y})^\top \beta + \nabla_x \mathcal{L}(\bar{x}, \bar{y}, \bar{u})^\top \gamma,$$

where (β, γ) is a unique element of the set $\Lambda^{cs}(\bar{x}, \bar{y}, \bar{u})$ defined in (4.42). □

Based on this simple form of the optimality conditions for the pessimistic program and the observations made above for the other types of stationarity concepts, we highlight in the next subsection some links between the available necessary optimality conditions for pessimistic and optimistic versions in bilevel programming.

4.4 Pessimistic Versus Optimistic Bilevel Programs

As mentioned in Section 1, the optimistic and pessimistic bilevel programs (P_o) and (P_p) are different from each other only when we can not guaranty the uniqueness in the lower-level problem (1.3). It is shown in this subsection that the necessary optimality conditions for both problems coincide if the pessimistic upper-level solution set S^h_{op} in (4.12) agrees with its optimistic counterpart defined below. For simplicity in our comparison we consider only the C-stationarity and S-type stationarity conditions. Moreover, we focus here only on the KKT-type C-stationarity and S-stationarity conditions for the pessimistic problem obtained in Theorem 4.5 and Theorem 4.7, denoted by P_pC-stationarity and P_pS-stationarity conditions, respectively, in order to avoid confusion with those for (P_o) denoted similarly while replacing P_p by P_o.

Definition 4.8 (C-stationarity and S-stationarity concepts for pessimistic and optimistic bilevel programs). *We say that:*

(i) *A point $\bar{x}$ P_p*C-STATIONARY *(resp. P_p*S-STATIONARY*) if there exist $(\bar{y}, \bar{u}) \in S^h_{op}(\bar{x})$ and $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that relationships* (3.11) *and* (4.35)*–*(4.37) *are satisfied together with condition* (4.9) *(resp.* (4.10)*).*

(ii) *A point $\bar{x}$ is P_oC-STATIONARY (resp. P_oS-STATIONARY) if there exist $(\bar{y}, \bar{u}) \in S^h_o(\bar{x})$ and $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that relationships* (3.11) *and* (4.35)*–*(4.37) *hold together with the condition* (4.9) *(resp. $\beta_i \geq 0 \wedge \nabla_y g_i(\bar{x}, \bar{y})\gamma \geq 0$, $\forall i \in \theta$).*

Recall for these definitions that the pessimistic upper-level solution map S^h_{op} is defined in (4.12) while for the optimistic problem we have

$$S^h_o(x) := \{(y, u) \in S^h(x) | \, F(x, y) \leq \varphi_o(x)\} \tag{4.43}$$

with φ_o given in (1.7). Let us mention that the P_oC-stationarity and P_oS-stationarity conditions were labeled in [10] as weak C-stationarity and weak S-stationarity conditions, respectively, given that stronger versions were also provided in the aforementioned paper.

The next consequences of the previous results is to show that the necessary optimality conditions for pessimistic programs are closely related to those for optimistic ones.

Corollary 4.9 (relationships between the C-stationarity and S-stationarity conditions for the optimistic and pessimistic programs). *Let $\bar{x}$ be a local optimal solution to the bilevel problem (P_p). The following assertions hold:*

(i) *Let all the assumptions of Theorem 4.5 be satisfied with the inner semicontinuity of S^h_{op} replaced by the condition $S^h_{op}(\bar{x}) = \{(\bar{y},\bar{u})\} = S^h_o(\bar{x})$. Then $\bar{x}$ is both P_pC-stationary and P_oC-stationary.*

(ii) *Let all the assumptions of Theorem 4.7 be satisfied with the inner semicontinuity of S^h_{op} replaced by the condition $S^h_{op}(\bar{x}) = \{(\bar{y},\bar{u})\} = S^h_o(\bar{x})$. Then $\bar{x}$ is P_pS-stationary. Furthermore, it is also P_oS-stationary provided that either $\theta = \{i|\, \beta_i = 0 \wedge \nabla_y g_i(\bar{x},\bar{y})\gamma = 0\}$ or $\theta = \emptyset$ (i.e., strict complementarity) holds.*

Proof. Observe that assertion (i) follows from Theorem 4.5 by noting that condition $S^h_{op}(\bar{x}) = \{(\bar{y},\bar{u})\}$ ensures the inner semicontinuity of S^h_{op} in (4.12). It is also not hard to observe that the condition $S^h_{op}(\bar{x}) = \{(\bar{y},\bar{u})\} = S^h_o(\bar{x})$ ensures the fulfillment of the inclusions $(\bar{y},\bar{u}) \in S^h_{op}(\bar{x})$ and $(\bar{y},\bar{u}) \in S^h_o(\bar{x})$ in Definition 4.8 (i) and (ii), respectively. Similarly assertion (ii) is obtained from Theorem 4.7. □

Using the C- and S-counterparts of [10, Theorem 6.6], we can derive analogs of Corollary 4.9(i) and (ii), respectively, while considering local optimal solutions to program (P_o). Note that all the assumptions of Corollary 4.9 are satisfied for the problem considered in [10, Example 6.5].

To close this section, we focus on the bilevel optimization problem (1.1) when the upper-level feasible set defined by nonegativity inequality constraints

$$\text{“}\min_x\text{”}\{F(x,y)|\, G(x) \geq 0,\, y \in S(x)\}. \tag{4.44}$$

The corresponding pessimistic reformulation is as follows:

$$\min_x\{\varphi_p(x)|\, G(x) \geq 0\} \text{ with } \varphi_p(x) := \min_y\{F(x,y)|\, y \in S(x)\}. \tag{4.45}$$

Let us further introduce the optimistic program related to problem (4.44) with the objective function $-F$:

$$\min_x\{\varphi_{op}(x)|\, G(x) \geq 0\} \text{ with } \varphi_{op}(x) := \min_y\{-F(x,y)|\, y \in S(x)\}. \tag{4.46}$$

The next result shows that necessary optimality conditions for the pessimistic problem (4.45) can be obtained from those of problem (4.46), and vice versa.

Proposition 4.10 (optimality conditions or pessimistic programs via optimistic ones). *Let $\bar{x}$ be a local optimal solution to program* (4.46)*, where problem* (1.3) *is convex and lower-level regular at $(\bar{x},y)$ for all $y \in S(\bar{x})$. Assume that $\bar{x}$ is upper-level regular and that the mapping S^h_{op} in* (4.12) *is inner semicompact at $\bar{x}$. The following assertions hold:*

(i) *If (A^c_1) and (A^c_2) hold at $(\bar{x},y,u)$ for all $(y,u) \in S^h_{op}(\bar{x})$. Then $\bar{x}$ is P_pC-stationary.*

(ii) *If (A^c_1), (A^c_2), and (A^s_1) hold at $(\bar{x},y,u)$ for all $(y,u) \in S^h_{op}(\bar{x})$. Then $\bar{x}$ is P_pS-stationary.*

Proof. To prove (i), note that under the assumptions of the theorem it follows from the C-counterpart of [10, Theorem 6.6] the existence of $(\bar{y},\bar{u}) \in S^h_{op}(\bar{x})$ and $(\alpha,\beta,\gamma) \in \mathbb{R}^{k+p+m}$ such that relationships (3.11), (4.9), and (4.37) holds together with

$$\begin{aligned} -\nabla_x F(\bar{x},\bar{y}) - \nabla G(\bar{x})^\top\alpha + \nabla_x g(\bar{x},\bar{y})^\top\beta + \nabla_x \mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0,\\ -\nabla_y F(\bar{x},\bar{y}) + \nabla_y g(\bar{x},\bar{y})^\top\beta + \nabla_y \mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma = 0. \end{aligned}$$

Multiplying both equalities in the latter system by -1 and thus changing the signs of the multipliers β and γ, we arrive at the P_pC-stationarity conditions in the sense of Definition 4.8, which justifies (i). Assertion (ii) is proved similarly. □

The power of this result is that it allows us to relax some assumptions for the validity of the C-stationarity and S-stationarity conditions of the pessimistic programs obtained in Theorem 4.5 and Theorem 4.7, respectively. To this end, observe first that only the inner semicompactness of S^h_{op} is needed here, while the second assumption (A^c_2) can be dropped. More clarifications on the latter point are given in the next section.

5 Upper Subdifferential Conditions for Pessimistic Programs

The Lipschitz continuity of the two-level value function φ_p from (1.8) has been crucial in deriving the necessary optimality conditions in the previous sections. Most of the conditions obtained above required the combination of assumptions $(A^{\ltimes}_1)$ and $(A^{\ltimes}_2)$, with $\ltimes = o, m, c$. If we drop $(A^{\ltimes}_2)$, the Lipschitz continuity of the two-level value function may be lost. In this case only the upper semicontinuity of φ_p can be readily achieved under the inner semicompactness of either S_{op} in (3.4) or S^h_{op} in (4.11). Hence we would be out of the scope of [18, Proposition 5.3] used in the previous sections while we absolutely need the function φ_p to be lower semicontinuous. To deal with the absence of lower semicontinuity, we develop in this section the so-called *upper subdifferential* necessary optimality conditions for pessimistic bilevel programs in the line suggested in [17] (see also [18, Chapter 5]) for general minimization problems. To proceed in this way, we only need to require that φ_p is finite at the local optimal solution $\bar{x}$.

Observe that we do not face the above difficulty while investigating necessary optimality conditions for the optimistic bilevel program (P_o) since its value function φ_o is automatically l.s.c. under the inner semicompactness of either solution map $S_o(x) := \{y \in S(x)|\, F(x,y) \le \varphi_o(x)\}$ or $S^h_o(x) := \{(y,u) \in S^h(x)|\, F(x,y) \le \varphi_o(x)\}$. Hence assumption $(A^{\ltimes}_1)$ is sufficient to derive the optimistic counterparts of the results in Sections 3 and 4; see [10] for more details.

Definition 5.1 (upper subdifferential optimality conditions). *Let $\bar{x}$ be a local optimal solution in the minimization problem (P_p). Following [17], we say that an upper subdifferential necessary condition for $\bar{x}$ holds if*

$$-\widehat{\partial}^+\varphi_p(\bar{x}) \subset \widehat{N}_X(\bar{x}) \tag{5.1}$$

via the regular upper subdifferential (2.3)*.*

A number of results and discussions on this type of necessary conditions and their advantages in comparison with the lower subdifferential ones for some classes of constrained minimization problems can be found in [17] and [18, Chapter 5]. This equally applies to the next theorem that provides a detailed version of upper subdifferential necessary optimality conditions of the S-type for the pessimistic bilevel program (P_p) under consideration.

Theorem 5.2 (S-type upper subdifferential optimality conditions for pessimistic bilevel programs). *Let $\bar{x}$ be a local optimal solution to (P_p), where problem* (1.3) *is convex and the lower-level regularity holds at $(\bar{x}, y)$ for all $y \in S(\bar{x})$. Assume that the mapping $S^h_{op} : \operatorname{dom} S^h \rightrightarrows \mathbb{R}^{m+p}$ admits a local upper Lipschitzian selection at $(\bar{x}, \bar{y}, \bar{u})$. Then the following conditions hold:*

$$\begin{array}{l} \forall(\beta,\gamma) \in \mathbb{R}^{m+p} \;\; \textit{satisfying} \;\; (4.10), (4.36), \;\; \textit{and} \;\; (4.37); \\ \exists\alpha \in \mathbb{R}^k \;\; \textit{verifying} \;\; (3.7) \;\; \textit{such that} \;\; (4.35) \;\; \textit{is satisfied.} \end{array} \tag{5.2}$$

Proof. Since $|\varphi_p(\bar{x})| < \infty$ due to $S^h_{op}(\bar{x}) \neq \emptyset$, it follows from [18, Proposition 5.2] that inclusion (5.1) holds. By definition (2.3) of the upper subdifferential, we have

$$\widehat{\partial}^+\varphi_p(\bar{x}) = -\widehat{\partial}(-\varphi_p)(\bar{x}) = -\widehat{\partial}\varphi_{op}(\bar{x}). \tag{5.3}$$

Inserting (5.3) into (5.1) gives us the inclusion

$$\widehat{\partial}\varphi_{op}(\bar{x}) \subset N_X(\bar{x}). \tag{5.4}$$

Since the function $-F$ is Fréchet differentiable at the reference solution and since the mapping S_{op}^h admits a local upper Lipschitzian selection at $(\bar{x},\bar{y},\bar{u})$, it follows from [21, Theorem 2] that

$$\widehat{\partial}\varphi_{op}(\bar{x}) = -\nabla_x F(\bar{x},\bar{y}) + \widehat{D}^* S^h(\bar{x},\bar{y},\bar{u})(-\nabla_y F(\bar{x},\bar{y}),0). \tag{5.5}$$

Now recall that the mapping S^h can be equivalently represented as

$$S^h(x) := \{(y,u)|\, \psi(x,y,u) \in \{0\}\times\Lambda\}, \text{ with } \psi \text{ and } \Lambda \text{ given in (4.28)}.$$

Employing [18, Corollary 1.15], we get the following lower estimate of the regular normal cone to the graph of the solution map S^h:

$$\nabla\psi(\bar{x},\bar{y},\bar{u})^\top \widehat{N}_{\{0\}\times\Lambda}(\psi(\bar{x},\bar{y},\bar{u})) \subset \widehat{N}_{\mathrm{gph}\, S^h}(\bar{x},\bar{y},\bar{u}) \tag{5.6}$$

by taking into account that the function ψ is Fréchet differentiable at $(\bar{x},\bar{y},\bar{u})$. It clearly follows from the construction of the coderivative $\widehat{D}^*S^h$ that

$$\begin{aligned}\widehat{D}^*S^h(\bar{x},\bar{y},\bar{u})(-\nabla_y F(\bar{x},\bar{y}),0) \supset \;& \Big\{\nabla_x\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma - \nabla_x g(\bar{x},\bar{y})^\top\beta\big|\\ & \nabla_y\mathcal{L}(\bar{x},\bar{y},\bar{u})^\top\gamma - \nabla_y g(\bar{x},\bar{y})^\top\beta - \nabla_y F(\bar{x},\bar{y}) = 0\\ & \gamma\in\mathbb{R}^m,\, (-\nabla_y g(\bar{x},\bar{y})\gamma,\beta) \in \widehat{N}_\Lambda(\bar{u},-g(\bar{x},\bar{y}))\Big\}.\end{aligned} \tag{5.7}$$

Observe further from the expression of the regular normal cone to Λ (see, e.g., [26]) that

$$\widehat{N}_\Lambda(\bar{u},-g(\bar{x},\bar{y})) = \left\{(u,v)\;\middle|\; \begin{array}{l} u_\eta = 0, v_\nu = 0,\\ \forall i\in\theta, u_i\le 0 \wedge v_i \le 0\end{array}\right\}. \tag{5.8}$$

Combining finally (5.4), (5.5), (5.7), and (5.8) completes the proof of the theorem. □

Consider now a special case of the so-called *simple bilevel programming problem* from [7] and derive upper subdifferential optimality conditions for its pessimistic version

$$\min_{x\in X}\; \varphi_p(x) := \max_{y\in S}\; F(x,y) \text{ with } S := \arg\min\{f(y)|\, g(y)\le 0\}. \tag{5.9}$$

To simplify the presentation of the result, we set $S^{n+1} := \underbrace{S\times\ldots\times S}_{n+1-\text{times}}$.

Theorem 5.3 (upper subdifferential optimality conditions for the simple pessimistic bilevel program). *Let $\bar{x}$ be a local optimal solution of problem* (5.9)*, where the function $-F$ is convex in (x,y) and the solution set S is convex and compact. Then we have the conditions:*

$$\begin{array}{l}\forall y := (y_s)_{s=1}^{n+1}\in S^{n+1}, \forall\mu := (\mu_s)_{s=1}^{n+1}\in\mathbb{R}_+^{n+1} \;\; \textit{with} \;\; \sum_{s=1}^{n+1}\mu_s = 1,\\ \exists\alpha\in\mathbb{R}^k \;\; \textit{satisfying} \;\; (3.11) \;\; \textit{such that}\end{array}$$

$$\textstyle\sum_{s=1}^{n+1}\mu_s\nabla_x F(\bar{x},y_s) + \nabla G(\bar{x})^\top\alpha = 0.$$

Proof. Since the set S is compact and the function $-F$ continuously differentiable, it follows from Danskin's theorem (see, e.g., [29, Proposition 2.1]) that the value function

$$\varphi_{op}(x) := \min\{-F(x,y)|\, y\in S\}$$

is locally Lipschitz continuous and its generalized gradient is given by

$$\begin{aligned}\overline{\partial}\varphi_{op}(\bar{x}) = \;& \mathrm{co}\,\{-\nabla_x F(\bar{x},y)|\, y\in S\}\\ := \;& \Big\{-\textstyle\sum_{s=1}^{n+1}\mu_s\nabla_x F(\bar{x},y_s)|\, \sum_{s=1}^{n+1}\mu_s = 1, \mu_s\ge 0,\, y_s\in S, s=1,\ldots,n+1\Big\}.\end{aligned} \tag{5.10}$$

On the other hand, the convexity of the set S and the function $-F$ in (x,y) implies that φ_{op} is lower regular, and hence we have

$$\widehat{\partial}\varphi_{op}(\bar{x}) = \partial\varphi_{op}(\bar{x}) = \overline{\partial}\varphi_{op}(\bar{x}). \tag{5.11}$$

Combining (5.4), (5.10), and (5.11), completes the proof of the theorem. □

6 Concluding Remarks

In this paper we derived lower (i.e., conventional) subdifferential optimality conditions for pessimistic bilevel programs via the LLVL (Section 3) and KKT (Section 4) approaches. In these approaches the Lipschitz continuity of the two-level value function occurred to be a crucial requirement to achieve the required lower semicontinuity of φ_p in (1.8), the nonemptiness of its basic subdifferential $\partial\varphi_p$, and thus the application of the convex hull property (2.6). In Section 5 we discussed upper subdifferential optimality conditions. Despite the violation of the Lipschitz continuity of φ_p in this framework, the conditions of the latter type could be stronger than the lower subdifferential conditions for particular classes of problems. It should be mentioned that assumptions $(A_1^{\ltimes})$ (with $\ltimes = m, c$) and (A_1^e), (A_2^e) used in many results obtained above can be replaced by the weaker calmness property of certain set-valued mappings, which hold automatically when, e.g., the mappings g and $(x, y) \mapsto \nabla_y f(x, y)$ are linear. Finally, note also that the combination of assumptions $(A_1^{\ltimes})$ and $(A_2^{\ltimes})$ (resp. (A_1^e), (A_2^e) and (A_3^e)) used in many results (resp. Theorem 4.3) of the paper can be replaced by the single condition $(A_3^{\ltimes})$ (with $\ltimes = m, c$) (resp. (A_4^e)). More details on these assumptions can be found in [10].

References

[1] T.Q. Bao, P. Gupta and B.S. Mordukhovich, *Necessary conditions in multiobjective optimization with equilibrium constraints*, J. Optim. Theory Appl. **135** (2007), 179–203.

[2] H. Bonnel, *Optimality conditions for the semivectorial bilevel optimization problem*, Pacif. J. Optim. **2** (2006), no. 3, 447–467.

[3] M. Cervinka, M. Matonoha and J.V. Outrata, *On the computation of relaxed pessimistic solutions to MPECs*, Optim. Math. Soft., to appear.

[4] S. Dassanayaka, *Methods of Variational Analysis in Pessimistic Bilevel Programming*, Ph.D. Thesis (advisor: B.S. Mordukhovich), Department of Mathematics, Wayne State University, Detroit, 2010.

[5] S. Dempe, *Foundations of Bilevel Programming*, Kluwer Academic Publishers, Dordrecht, 2002.

[6] S. Dempe, *Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints*, Optimization **52** (2003), no. 3, 333–359.

[7] 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, pp. 149–161, Springer, Berlin, 2010,

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

[9] S. Dempe, V.V. Kalashnikov and N. Kalashnykova, *Optimality conditions for bilevel programming problems*, S. Dempe (ed.) et al., Optimization with Multivalued Mappings, Springer, New York, 2006.

[10] S. Dempe, B.S. Mordukhovich and A.B. Zemkoho, *Sensitivity analysis for two-level value functions with applications to bilevel programming*, Optimization-online Nr. 3227, 2011.

[11] 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.

[12] R. Henrion and T. Surowiec, *On calmness conditions in convex bilevel programming*, Appl. Analysis **90** (2011), 951–970.

[13] S. Lohse, *Eine Spezielle Klasse von Zwei-Ebenen-Optimierungsaufgaben*, Ph.D. Thesis (advisor: S. Dempe), Fakultät für Mathematik und Informatik, Technische Universität Bergakademie , Freiberg, 2010.

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

[15] B.S. Mordukhovich, *Maximum principle in problems of time optimal control with nonsmooth constraints*, J. Appl. Math. Mech. **40** (1976), 960–969.

[16] B.S. Mordukhovich, *Metric approximations and necessary optimality conditions for general classes of extremal problems*, Soviet Math. Dokl. **22** (1980), 526–530.

[17] B.S. Mordukhovich, *Necessary conditions in nonsmooth minimization via lower and upper subgradients*, Set-Valued Anal. **12** (2004), 163–193.

[18] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation. I: Basic Theory. II: Applications*, Springer, Berlin, 2006.

[19] B.S. Mordukhovich and N.M. Nam, *Variational stability and marginal functions via generalized differentiation*, Math. Oper. Res. **30** (2005), 800–816.

[20] B.S. Mordukhovich, M.N. Nam and H.M. Phan, *Variational analysis of marginal functions with applications to bilevel programming problems*, J. Optim. Theory Appl. **152** (2012).

[21] B.S. Mordukhovich, M.N. Nam, and N.D. Yen, *Subgradients of marginal functions in parametric mathematical programming*, Math. Program. **116** (2009), 369–396.

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

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

[24] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.

[25] K. Shimizu, Y. Ishizuka and J.F. Bard. *Nondifferentiable and Two-Level Mathematical Programming.* Kluwer Academic Publishers, Dordrecht, 1997.

[26] J.J. Ye, *Optimality conditions for optimization problems with complementarity constraints*, SIAM J. Optim. **9** (1999), 374–387.

[27] J.J. Ye, *Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints*, SIAM J. Optim. **10** (2000), 943–962.

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

[29] 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), 1885–1905.