Sensitivity analysis for two-level value functions with applications to bilevel programming

S. DEMPE[*], B. S. MORDUKHOVICH[†], AND A. B. ZEMKOHO[‡]

Abstract. This paper contributes to a deeper understanding of the link between a now conventional framework in hierarchical optimization spread under the name of the optimistic bilevel problem and its initial more difficult formulation that we call here the *original* optimistic bilevel optimization problem. It follows from this research that, although the process of deriving necessary optimality conditions for the latter problem is more involved, the conditions themselves do not–to a large extent–differ from those known for the conventional problem. It has been already well recognized in the literature that for optimality conditions of the usual optimistic bilevel program appropriate *coderivative* constructions for the set-valued solution map of the lower-level problem could be used, while it is shown in this paper that for the original optimistic formulation we have to go a step further to require and justify a certain Lipschitz-like property of this map. This occurs to be related to the local Lipschitz continuity of the optimal value function of an optimization problem constrained by solutions to another optimization problem; this function is labeled here as the *two-level value function*. More generally, we conduct a detailed sensitivity analysis for value functions of mathematical programs with extended complementarity constraints. The results obtained in this vein are applied to the two-level value function and then to the original optimistic formulation of the bilevel optimization problem, for which we derive verifiable stationarity conditions of various types entirely in terms of the initial data.

Keywords. Bilevel programming, Coderivative, Lipschitz-like property, Sensitivity analysis, Two-level value function, MPCC value functions, Optimality conditions

AMS subject classifications. 90C26, 90C30-31, 90C46, 91C12, 91A65

1. Introduction. This paper is mainly motivated by the study of a class of the so-called bilevel programming problems generally formalized as

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

where $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $X \subset \mathbb{R}^n$ stands for the *upper level/leader's* objective function and the feasible solution set, respectively, while the multifunction $S : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ denotes the set-valued *solution/argminimum map* for the *lower level/follower's* problem

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

with the lower-level objective function $f : \mathbb{R}^n \times \mathbb{R}^m$. For simplicity we confine ourselves to the case where the upper and lower level constraint sets are given explicitly as

$$X := \{x \in \mathbb{R}^n \,|\, G(x) \le 0\} \text{ and } K(x) := \{y \in \mathbb{R}^m \,|\, g(x,y) \le 0\}, \tag{1.3}$$

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 will be 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.

The quotation marks in (1.1) are used to express the uncertainty in the formalization of the bilevel optimization problem in the case of nonuniquely determined lower-level optimal solutions. In the latter case two major approaches have been suggested in the literature in order to easily handle the problem. On one hand, we have the *optimistic formulation*

$$(\mathrm{P}_o) \qquad \min\{\varphi_o(x)|\, x \in X\} \text{ with } \varphi_o(x) := \min_y\{F(x,y)|y \in S(x)\}.$$

¿From the economics viewpoint this corresponds to a situation where the follower participates in the profit of the leader, i.e., some cooperation is possible between both players on the upper and lower levels. However, it would not always be possible for the leader to convince the follower to make choices that are favorable for him or her. Hence it is necessary for the upper-level player to bound damages resulting from undesirable selections on the lower level. This gives the *pessimistic formulation* of the bilevel optimization problem as follows:

$$\min\{\varphi_p(x)|\, x \in X\} \text{ with } \varphi_p(x) := \max_y\{F(x,y)|y \in S(x)\}.$$

The latter problem is a special class of minimax problems. Static minimax problems, corresponding in our case to a situation where the feasible set of the inner problem $S(x)$ is independent of x, have been highly investigated in the literature; see, e.g., [15, 53, 54]. At the same time, it has been well recognized that when $S(x)$ stands for varying sets of solutions to another optimization problem, the pessimistic formulation above faces many challenges. Some of them are highlighted in [8] and the references therein. Recent developments on pessimistic bilevel programs can be found in [4, 7, 30].

Our main concern in this paper is the original optimistic formulation (P_o) in bilevel programming that has been eventually substituted in the literature, under the name of "optimistic bilevel program," by the following optimization problem:

$$(\mathrm{P}) \qquad \min_{x,y}\{F(x,y)|\, x \in X, \, y \in S(x)\}.$$

The latter problem, which we label here as the *auxiliary bilevel program*, has been well investigated. In the last two decades, problem (P) has attracted a lot of interest from both viewpoints of optimization theory and applications. The reader is referred to [2, 8, 54] and the bibliographies therein for detailed discussions. For more recent results on the topic we refer to [6, 16, 10, 11, 12, 27, 40, 61]. In addition, a vast literature on related mathematical problems with equilibrium constraints (MPECs) is widely available; see the books [32, 38, 46] with their commentaries and references. Note here that investigating problem (P) and related MPECs faces the issue of passing to an equivalent single-level reformulation, especially when the so-called Karush-Kuhn-Tucker (KKT) reformulation is in question [10]. As it will be clear in this paper, investigating the bilevel program (P_o) of our main interest in what follows does not lead to such a difficulty.

Unfortunately, very little is known about the initial bilevel program (P_o) that is the *original optimistic* model in the bilevel programming (1.1) and is labeled as

such. It has been well recognized that problems (P$_o$) and (P) are equivalent for global solutions while *not for local ones*: a local optimal solution to (P) may not be a local optimal solution to (P$_o$); see [8, 17]. It is clear that there is no distinction between both problems in the case where the optimal solutions of the lower-level problem are uniquely determined. Ruling out the latter possibility, a crucial question that arises is: *how are stationary points of* (P$_o$) *and* (P) *related to each other*? Among other things, we attempt to answer this question in the present paper.

To proceed in this direction, we aim to derive rather comprehensive first-order necessary optimality conditions, via various types of stationarity in bilevel programming, for the original problem (P$_o$) and compare them with known ones for (P). According to the general approach to "abstract" problems of this type developed in [38], sensitivity analysis and necessary optimality conditions for such problems are closely related to deriving appropriate subdifferential estimates for the optimal value function

$$\varphi_o(x) := \min_y \{F(x,y)|y \in S(x)\}. \tag{1.4}$$

We assume with no further mentioning that the minimum in (1.4) and similar settings below is realized. In the framework of this paper the value function (1.4) is not just defined via an abstract mapping S but it is associated with the two-level optimization problem (P$_o$) under consideration, where S is the solution map of the specifically given lower-level problem of parametric optimization. For this reason we call (1.4) the *two-level value function*.

A large literature exists for value functions (known also as *marginal functions*) in classical optimization problems constrained by functional inequalities and/or equalities; see, e.g., [3, 5, 19, 23, 49] to name just a few. Since marginal functions are *intrinsically nonsmooth*, generalized derivatives of various kinds are used to study their properties. More recently, significant progress in the study and applications of various classes of marginal/value functions has been made by using generalized differential constructions introduced by the second author [34]; see more details in [11, 16, 38, 39, 40, 41, 42]. Note that in problems of nonlinear and nondifferentiable programming the key conditions needed for subdifferential estimates and sensitivity analysis of the corresponding marginal functions are the classical constraint qualification by Mangasarian and Fromovitz [33] (MFCQ) and its nonsmooth extension introduced in [36]. It happens that these qualification conditions are not applicable to the two-level value function φ_o written in a marginal function form under parametric functional constraints; see Section 5. Thus adequate rules tailored for φ_o have to be developed.

In order to tackle this, we consider in this paper *three* possible approaches to sensitivity analysis for two-level value functions of type (1.4) involving certain representations/transformations of the solution map

$$S(x) := \operatorname*{argmin}_y \{f(x,y)|y \in K(x)\} \tag{1.5}$$

of the lower-level problem in the construction of φ_o. Here we label them conditionally as the *complementarity/MPCC* approach, as the *generalized equation/OPEC* approach, and as the *lower-level value function/LLVF* approach.

In the first two approaches, function (1.4) becomes the optimal value function of a mathematical program with complementarity constraints (MPCC) and an optimization problem with a generalized equation constraint (OPEC), respectively. To

4

the best of our knowledge, the initial results for value functions of this type were obtained by Lucet and Ye [31] and by Hu and Ralph [29]. Paper [29] is devoted to the study of the strict differentiability as well as the first-order and second-order directional derivatives of the value function for the corresponding MPCC under the MPCC/MPEC linear independence constraint qualification (MPEC-LICQ). Another approach is developed in [31], which employs the limiting subdifferential constructions by Mordukhovich to conduct a local sensitivity analysis of value functions of the aforementioned types.

The developments of this paper within the MPCC approach to sensitivity analysis of value functions are much closer to those by Lucet and Ye while we also try to bridge the gap between our work and that by Hu and Ralph; see Subsection 3.3. Note that the results of [31] for MPCC value functions focus only on the case where one of the functional components involved in the crucial complementarity condition is given by the simplest linear function. Some of our results obtained in Section 3 can be seen as extensions of those in [31] to the general function setting in the complementarity condition. On the other hand, our results in Section 4 within the OPEC approach cover the generalized equation description of (1.5) via the normal cone to *moving* convex sets (of the quasi-variational inequality type), which was not considered in [31]. In this way we derive more detailed upper estimates for the limiting subdifferential of the corresponding value function and establish their clear link with those obtained via the MPCC approach; see Subsection 5.1. Another important difference between our work on sensitivity analysis for value functions via the MPCC and OPEC approaches and the one by Lucet and Ye is that we do not use their growth hypothesis, which plays a significant role in the results of [31]. We replace it by the weaker inner semicompactness assumption imposed on the solution map of the upper-level problem

$$S_o(x) := \{y \in S(x) \,|\, F(x,y) \leq \varphi_o(x)\} \qquad (1.6)$$

and derive even tighter upper bounds for the limiting subdifferential of φ_o under the inner semicontinuity of the mapping S_o in (1.6); see Section 2 for the precise definitions.

In the third (LLVF) of the aforementioned approaches, originated by Outrata [44] for a special class of bilevel programs/Stackeklberg games, we represent the solution map (1.5) of the lower-level problem as an inequality system containing the lower-level value function of (1.2). In this way we provide verifiable conditions in terms of the initial data to evaluate the coderivative of S and establish the Lipschitz-like property of this mapping. This leads us in turn to new conditions ensuring the local Lipschitz continuity of the two-level value function φ_o; see Subsection 5.2 for all the details.

The rest of the paper is organized as follows. Section 2 presents basic notions and results of variational analysis and generalized differentiation widely used in the subsequent parts. Section 3 is mainly concerned with sensitivity analysis of MPCC value functions. Here we derive upper estimates of the limiting subdifferential for such functions from various perspectives, depending on the type of optimality/stationary conditions of interest for the the original bilevel model (P_o). It should be mentioned that the results in Section 3 can stand on their own. Indeed, they also provide efficient rules to obtain estimates of the coderivative and the fulfillment of the Lipschitz-like property for mappings of special structures (inequality and equality systems with complementarity constraints) important for other classes of optimization-related problems, not just for bilevel programming. Sensitivity analysis of OPEC value functions, which is of its own interest as well, is conducted in Section 4.

The first part of Section 5 mainly deals with the applications of the results from Sections 3 and 4 to sensitivity analysis of the two-level value function (1.4) via the MPCC and OPEC approaches. In the second part of this section (i.e., in Subsection 5.2) we develop lower-level value function approach to analyze the two-level value function φ_o. Here a detailed discussion is given on rules to derive subdifferential estimates and establish the local Lipschitz continuity of φ_o from a perspective completely different from the previous ones.

In the concluding Section 6 we employ the results obtained above to deriving necessary optimality conditions for the original optimistic formulation (P_o) in the various forms of stationarity conditions including the new types introduced in this paper.

2. Background material. More details on the material briefly discussed in this section can be found in the books [38, 50, 52] and the references therein. We start with the *Painlevé-Kuratowski outer/upper limit* of a set-valued mapping $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ as $x \to \bar{x}$ defined by

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

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

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

while our basic construction in this paper known as the *Mordukhovich/limiting subdifferential* of ψ at $\bar{x} \in \mathrm{dom}\, \psi$ is defined via the outer limit (2.1) by

$$\partial\psi(\overline{x}) := \mathrm{Limsup}_{x \to \overline{x}} \widehat{\partial}\psi(x). \quad (2.2)$$

If ψ is convex, then the subdifferential $\partial\psi(\overline{x})$ reduces to the classical subdifferential of convex analysis. If ψ is locally Lipschitzian around $\bar{x}$, then the set $\partial\psi(\overline{x})$ is nonempty and compact. Moreover, its convex hull agrees with the subdifferential/generalized gradient by Clarke. If ψ is strictly differentiable at $\bar{x}$, i.e.,

$$\lim_{v \to \overline{x}, \, x \to \overline{x}} \frac{\psi(v) - \psi(x) - \langle \nabla\psi(\overline{x}), v - x\rangle}{\|v - x\|} = 0 \quad (2.3)$$

(with $\nabla\psi(\overline{x})$ denoting the classical gradient of ψ at $\overline{x}$), then $\partial\psi(\overline{x}) = \{\nabla\psi(\overline{x})\}$. It should be mentioned that every function continuously differentiable around some point is strictly differentiable at this point and that every function locally Lipschitzian around $\bar{x}$ is strictly differentiable at $\bar{x}$ provided that its subdifferential (2.2) is a singleton.

Given a nonempty set $\Omega \subset \mathbb{R}^n$, our basic *normal cone* to it at $\overline{x} \in \Omega$ corresponding to the subdifferential construction (2.2) is defined by

$$N_\Omega(\overline{x}) := \mathrm{Limsup}_{x \to \overline{x} \, (x \in \Omega)} \widehat{N}_\Omega(x) \quad (2.4)$$

via the outer limit (2.1) of the regular counterpart

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

at points $x \in \Omega$ near $\bar{x}$. Note that for sets $\Omega \subset \mathbb{R}^n$ locally closed around $\bar{x}$ the given definition (2.4) reduces to the original one

$$N_\Omega(\overline{x}) = \underset{x \to \overline{x}}{\text{Limsup}} \ [\text{cone}(x - \Pi_\Omega(x))].$$

introduced in [34], 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. Using the normal cone (2.4), we can equivalently describe the basic subdifferential (2.2) by

$$\partial\psi(\overline{x}) = \{v \in \mathbb{R}^n \,|\, (v, -1) \in N_{\text{epi}\psi}(\overline{x}, \psi(\overline{x}))\}$$

for lower semicontinuous (l.s.c.) functions with the epigraph $\text{epi}\,\psi$ and define the *singular subdifferential* of ψ at $\bar{x} \in \text{dom}\,\psi$ by

$$\partial^\infty\psi(\overline{x}) := \{v \in \mathbb{R}^n \,|\, (v, 0) \in N_{\text{epi}\psi}(\overline{x}, \psi(\overline{x}))\}. \tag{2.5}$$

It is worth mentioning that for functions ψ l.s.c. around $\bar{x}$ we have $\partial^\infty\psi(\overline{x}) = \{0\}$ if and only if ψ is locally Lipschitzian around this point.

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

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

recall the notion of *coderivative* for Ψ at $(\bar{x}, \bar{y}) \in \text{gph}\,\Psi$ defined in [35] by

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

via the normal cone (2.4) to the graph of Ψ. If Ψ is single-valued and locally Lipschitzian around $\bar{x}$, its coderivative can be represented analytically as

$$D^*\Psi(\overline{x})(v) = \partial\langle v, \Psi\rangle(\overline{x}) \text{ for } v \in \mathbb{R}^m,$$

via the basic subdifferential (2.2) of the Lagrange scalarization $\langle v, \Psi\rangle(x) := \langle v, \Psi(x)\rangle$, where the component $\bar{y}(= \Psi(\overline{x}))$ is omitted in the coderivative notation for single-valued mappings. This implies the coderivative representation

$$D^*\Psi(\overline{x})(v) = \{\nabla\Psi(\overline{x})^\top v\} \text{ for } v \in \mathbb{R}^m,$$

when Ψ is strictly differentiable at $\bar{x}$ as in (2.3) with $\nabla\Psi(\overline{x})$ standing for its Jacobian matrix at $\bar{x}$ and with the symbol "$\top$" standing for transposition.

Some continuity properties of set-valued mappings are of a particular interest in this paper. We say that $\Psi : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is *inner semicompact* at $\bar{x}$ 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 in finite dimensions whenever Ψ is uniformly bounded around $\overline{x}$, i.e., there exists a neighborhood U of $\overline{x}$ and a bounded set $C \subset \mathbb{R}^m$ such that

$$\Psi(x) \subset C \text{ for all } x \in U. \tag{2.7}$$

The mapping Ψ is *inner semicontinuous* at $(\overline{x}, \overline{y}) \in \text{gph}\,\Psi$ 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 $k \to \infty$. The latter property reduces to the usual continuity for single-valued mappings while in the general set-valued case it is implied by the *Lipschitz-like/Aubin property* of Ψ around $(\overline{x}, \overline{y}) \in$

gph Ψ, which means that there are neighborhoods U of $\overline{x}$, V of $\overline{y}$, and a constant $\ell > 0$ such that

$$d(y, \Psi(x)) \le \ell\|u - x\| \quad \text{for all} \ \ u, x \in U \ \ \text{and} \ \ y \in \Psi(u) \cap V, \tag{2.8}$$

where d signifies the distance between a point and a set in $\mathbb{R}^m$. When $V = \mathbb{R}^m$ in (2.8), this property reads as to the classical local Lipschitz continuity of Ψ around $\overline{x}$. A complete characterization of the Lipschitz-like property (2.8), and hence a sufficient condition for the inner semicontinuity of Φ at $(\bar{x}, \bar{y})$, is given for closed-graph mappings by the following *coderivative/Mordukhovich criterion* (see [38, Theorem 5.7] and [50, Theorem 9.40]):

$$D^*\Psi(\overline{x}, \overline{y})(0) = \{0\}. \tag{2.9}$$

Furthermore, the infimum of all $\ell > 0$ for which (2.8) holds is equal to the coderivative norm $\|D^*\Psi(\bar{x}, \bar{y})\|$ as a positively homogeneous mapping $D^*\Psi(\overline{x}, \overline{y})\colon \mathbb{R}^m \rightrightarrows \mathbb{R}^n$.

If we fix $x = \overline{x}$ in (2.8), the resulting weaker property is known as *calmness* of Ψ at $(\bar{x}, \bar{y})$ [50]; for $V = \mathbb{R}^m$ it corresponds to the upper Lipschitz property of Robinson [48].

In order to analyze our two-level optimal value function φ_o in (1.4), we first consider a general "abstract" framework of the marginal functions

$$\mu(x) := \min_y \{\psi(x, y) \,|\, y \in \Psi(x)\} \tag{2.10}$$

with $\psi\colon \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ and $\Psi\colon \mathbb{R}^n \rightrightarrows \mathbb{R}^m$. Denoting the argminimum mapping in (2.10) by

$$\Psi_o(x) := \operatorname{argmin} \{\psi(x, y)|y \in \Psi(x)\} = \{y \in \Psi(x)| \ \psi(x, y) \le \mu(x)\},$$

we summarize in the next theorem some known results on general marginal functions needed in the paper; see [38, Corollary 1.109] and [39, Theorem 5.2].

THEOREM 2.1 (properties of general marginal functions). *Let the marginal function μ be given in* (2.10), *where the graph of Ψ is locally closed around $(\bar{x}, \bar{y}) \in \operatorname{gph}\Psi$, and where ψ is strictly differentiable at this point. The following assertions hold:*

(i) *Let Ψ_o be inner semicontinuous at $(\overline{x}, \overline{y})$. Then μ is lower semicontinuous at $\overline{x}$ and we have the following upper bound for its basic subdifferential*

$$\partial\mu(\overline{x}) \subset \nabla_x\psi(\overline{x}, \overline{y}) + D^*\Psi(\overline{x}, \overline{y})(\nabla_y\psi(\overline{x}, \overline{y})).$$

If in addition Ψ is Lipschitz-like around $(\overline{x}, \overline{y})$, then we also have the Lipschitz continuity of μ around $\overline{x}$.

(ii) *Let Ψ_o be inner semicompact at $\overline{x}$. Then μ is lower semicontinuous at $\overline{x}$ and*

$$\partial\mu(\overline{x}) \subset \bigcup_{\overline{y} \in \Psi_o(\overline{x})} \{\nabla_x\psi(\overline{x}, \overline{y}) + D^*\Psi(\overline{x}, \overline{y})(\nabla_y\psi(\overline{x}, \overline{y}))\}.$$

If in addition Ψ is Lipschitz-like around $(\overline{x}, \overline{y})$ for all vectors $\overline{y} \in \Psi_o(\overline{x})$, then μ is Lipschitz continuous around $\overline{x}$.

Depending on specific structures of the set-valued mapping Ψ, our aim in Sections 3-5 is to give detailed upper bounds for $D^*\Psi(\overline{x}, \overline{y})$ in terms of problem data. Verifiable rules for Ψ to be Lipschitz-like will also be provided. Thus, implying explicit upper

bounds for $\partial\mu(\overline{x})$ and the local Lipschitz continuity of μ. More discussions on the inner semicontinuity of argminimum mappings can be found in [11, Remark 3.2] and the references therein.

To conclude this section, we present constraint qualification and necessary optimality conditions for a general optimization problem with geometric constraints in terms of limiting normals and subgradients; see, e.g., [38, Proposition 5.3] and the commentaries to it.

THEOREM 2.2 (optimality conditions under geometric constraints). *Let $\overline{x}$ be a local optimal solution to the problem:*

$$minimize \ \ \psi(x) \ \ subject \ to \ \ x \in \Omega,$$

where $\psi : \mathbb{R}^n \to \overline{\mathbb{R}}$ is l.s.c. around $\bar{x} \in \Omega \cap \operatorname{dom}\psi$, and where $\Omega \subset \mathbb{R}^n$ is locally closed around this point. Then we have

$$0 \in \partial\psi(\overline{x}) + N_\Omega(\overline{x}) \tag{2.11}$$

provided the validity of the qualification condition

$$\partial^\infty\psi(\overline{x}) \cap (-N_\Omega(\overline{x})) = \{0\}, \tag{2.12}$$

which is the case, in particular, when ψ is locally Lipschitzian around $\bar{x}$.

3. Sensitivity analysis of MPCC value functions. In this section we consider the parametric optimization problem belonging to the class of *mathematical programs with complementarity constraints* (MPCCs):

$$\min_y\{F(x,y)| \quad g(x,y) \leq 0, \ h(x,y) = 0, \\ G(x,y) \geq 0, \ H(x,y) \geq 0, \ G(x,y)^\top H(x,y) = 0\}, \tag{3.1}$$

where $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^a$, $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^b$ and $G, H : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^d$ are all continuously differentiable functions. Denoting by

$$S^c(x) := \{y \in \mathbb{R}^m| \quad g(x,y) \leq 0, \ h(x,y) = 0, \\ G(x,y) \geq 0, \ H(x,y) \geq 0, \ G(x,y)^\top H(x,y) = 0\} \tag{3.2}$$

the sets of *feasible solutions* to (3.1), associate with (3.1) the optimal value function

$$\mu^c(x) := \min_y\{F(x,y)|y \in S^c(x)\}. \tag{3.3}$$

The main goal of this section is to conduct a local sensitivity analysis for the MPCC problem (3.1) around the given optimal solution. By this we understand deriving efficient subdifferential estimates for the optimal value function (3.3), verifiable conditions for its local Lipschitz continuity and for the Lipschitz-like property of the feasible solution map (3.2) entirely in terms of the initial data of (3.1). According to the variational analysis results discussed in Section 2, this relates to evaluating the *coderivative* (2.6) of the solution map (3.2). Adopting the terminology originated by Scheel and Scholtes [51], the sensitivity analysis results established below and the associated constraint qualifications are expressed via the sets of M-, C-, and S-type multipliers used in the corresponding *M(ordukhovich), C(larke)*, and *S(trong) stationarity* conditions for MPCCs; cf. Section 6.

Fix a pair $(\overline{x}, \overline{y}) \in \operatorname{gph} S^c$ and associate with it the following partition of the indices for the functions involved in the complementarity system of (3.2):

$$
\begin{aligned}
\eta := & \ \eta(\overline{x}, \overline{y}) := \{i = 1, \ldots, d| \ \ G_i(\overline{x}, \overline{y}) = 0, \ H_i(\overline{x}, \overline{y}) > 0\}, \\
\theta := & \ \theta(\overline{x}, \overline{y}) := \{i = 1, \ldots, d| \ \ G_i(\overline{x}, \overline{y}) = 0, \ H_i(\overline{x}, \overline{y}) = 0\}, \\
\nu := & \ \nu(\overline{x}, \overline{y}) := \{i = 1, \ldots, d| \ \ G_i(\overline{x}, \overline{y}) > 0, \ H_i(\overline{x}, \overline{y}) = 0\},
\end{aligned}
\tag{3.4}
$$

where the middle set θ in (3.4) is known as the *biactive* or *degenerate index set*. As it will be clear in what follows, the difference between the various types of multiplier sets depends on the structure of the components corresponding to the biactive set θ. We further consider a vector $z^* \in \mathbb{R}^{n+m}$ such that

$$
z^* + \nabla g(\overline{x}, \overline{y})^\top \alpha + \nabla h(\overline{x}, \overline{y})^\top \beta + \nabla G(\overline{x}, \overline{y})^\top \gamma + \nabla H(\overline{x}, \overline{y})^\top \zeta = 0, \tag{3.5}
$$

$$
\alpha \geq 0, \ \alpha^\top g(\overline{x}, \overline{y}) = 0, \tag{3.6}
$$

$$
\gamma_\nu = 0, \zeta_\eta = 0, \tag{3.7}
$$

$$
\forall i \in \theta, \ (\gamma_i < 0 \wedge \zeta_i < 0) \vee \gamma_i \zeta_i = 0. \tag{3.8}
$$

The set of *M-type multipliers* associated with problem (3.1) is defined by

$$
\Lambda^{cm}(\overline{x}, \overline{y}, z^*) := \{(\alpha, \beta, \gamma, \zeta) | (3.5) - (3.8) \text{ hold}\}.
$$

Similarly we define the set $\Lambda_y^{cm}(\overline{x}, \overline{y}, y^*)$, with $y^* \in \mathbb{R}^m$, obtained by replacing the gradients of g, h, G, and H in (3.5) by their partial derivatives with respect to y. In the case where $y^* := \nabla_y F(\overline{x}, \overline{y})$ we denote $\Lambda_y^{cm}(\overline{x}, \overline{y}) := \Lambda_y^{cm}(\overline{x}, \overline{y}, \nabla_y F(\overline{x}, \overline{y}))$.

The corresponding sets of *C-type multipliers* denoted by $\Lambda^{cc}(\overline{x}, \overline{y}, z^*)$, $\Lambda_y^{cc}(\overline{x}, \overline{y}, y^*)$, and $\Lambda_y^{cc}(\bar{x}, \bar{y})$ are defined similarly to (3.5)–(3.8) with the replacement of (3.8) by

$$
\gamma_i \zeta_i \geq 0 \ \text{ for all } \ i \in \theta.
$$

For the case of *S-type multipliers* we need to define only $\Lambda_y^{cs}(\overline{x}, \overline{y})$, which is a analog of $\Lambda_y^{cm}(\overline{x}, \overline{y})$ with the replacement of (3.8) therein by

$$
\gamma_i \leq 0 \wedge \zeta_i \leq 0 \ \text{ for all } \ i \in \theta.
$$

The following links between the sets $\Lambda_y^{cs}(\overline{x}, \overline{y})$, $\Lambda_y^{cm}(\overline{x}, \overline{y})$, and $\Lambda_y^{cc}(\overline{x}, \overline{y})$ is obvious:

$$
\Lambda_y^{cs}(\overline{x}, \overline{y}) \subset \Lambda_y^{cm}(\overline{x}, \overline{y}) \subset \Lambda_y^{cc}(\overline{x}, \overline{y}). \tag{3.9}
$$

To further simplify the presentation of this section, we introduce the following *Lagrange-type* and *singular Lagrange-type functions*, respectively, associated with problem (3.1):

$$
L(x, y, \alpha, \beta, \gamma, \zeta) := F(x, y) + g(x, y)^\top \alpha + h(x, y)^\top \beta + G(x, y)^\top \gamma + H(x, y)^\top \zeta,
$$

$$
L_o(x, y, \alpha, \beta, \gamma, \zeta) := g(x, y)^\top \alpha + h(x, y)^\top \beta + G(x, y)^\top \gamma + H(x, y)^\top \zeta.
$$

In the sequel the derivative of L_o with respect to (x, y) is often needed and is denoted by

$$
\begin{aligned}
\nabla L_o(x, y, \alpha, \beta, \gamma, \zeta) := & \ \nabla_{x,y} L_o(x, y, \alpha, \beta, \gamma, \zeta) \\
= & \ \nabla g(x, y)^\top \alpha + \nabla h(x, y)^\top \beta + \nabla G(x, y)^\top \gamma + \nabla H(x, y)^\top \zeta.
\end{aligned}
$$

The following *optimal solution/argminimum* map for the MPCC problem (3.1) given by

$$
S_o^c(x) := \{y \in S^c(x) | F(x, y) \leq \mu^c(x)\} \tag{3.10}
$$

plays a significant role in our subsequent sensitivity analysis in this section.

3.1. Sensitivity analysis via M-type multipliers. To proceed in this subsection, we define the *M-qualification conditions* at $(\bar{x}, \bar{y})$:

(A_1^1) $(\alpha, \beta, \gamma, \zeta) \in \Lambda^{cm}(\bar{x}, \bar{y}, 0) \Longrightarrow \alpha = 0, \beta = 0, \gamma = 0, \zeta = 0$;

(A_2^1) $(\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cm}(\bar{x}, \bar{y}, 0) \Longrightarrow \nabla_x L_o(\bar{x}, \bar{y}, \alpha, \beta, \gamma, \zeta) = 0$;

(A_3^1) $(\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cm}(\bar{x}, \bar{y}, 0) \Longrightarrow \alpha = 0, \beta = 0, \gamma = 0, \zeta = 0$

and observe the obvious links between them:

$$(A_2^1) \Longleftarrow (A_3^1) \Longrightarrow (A_1^1). \tag{3.11}$$

The next theorem provides a constructive upper estimate of the coderivative (2.6) of the MPCC feasible solution map (3.2) and gives a verifiable condition for its *robust Lipschitzian stability*, i.e., the validity of the Lipschitz-like property.

THEOREM 3.1 (coderivative estimate and Lipschitz-like property of MPCC feasible solutions via M-multipliers). *Let $(\bar{x}, \bar{y}) \in gph\, S^c$, and let (A_1^1) holds at $(\bar{x}, \bar{y})$. Then we have for all $y^* \in \mathbb{R}^m$*

$$D^* S^c(\bar{x}, \bar{y})(y^*) \subset \{\nabla_x L_o(\bar{x}, \bar{y}, \alpha, \beta, \gamma, \zeta) \,|\, (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cm}(\bar{x}, \bar{y}, y^*)\}. \tag{3.12}$$

If in addition (A_2^1) is satisfied at $(\bar{x}, \bar{y})$, then S^c is Lipschitz-like around this point.

Proof. We start by recalling that the complementarity system

$$G(x, y) \geq 0,\ H(x, y) \geq 0,\ G(x, y)^\top H(x, y) = 0 \tag{3.13}$$

is equivalent to the following inclusion:

$$(G_i(x, y), H_i(x, y)) \in \{(u, v) \in \mathbb{R}^2 \,|\, u \geq 0,\ v \geq 0,\ u^\top v = 0\} := \Lambda_i,\ \text{for } i = 1, \ldots, d, \tag{3.14}$$

and the graph of S^c can be rewritten in the form

$$gph\, S^c = \{(x, y) \,|\, \psi(x, y) \in \Lambda\}$$

via the vector-valued function ψ and the polyhedral set Λ defined by

$$\psi(x, y) := \big[g(x, y), h(x, y), (G_i(x, y), H_i(x, y))_{i=1}^d\big] \text{ and } \Lambda := \mathbb{R}_-^a \times \{0_b\} \times \prod_{i=1}^d \Lambda_i.$$

It follows from the calculus rules in [38, Theorem 3.8] and [50, Theorem 6.14] that

$$N_{gph\, S^c}(\bar{x}, \bar{y}) \subset \nabla \psi(\bar{x}, \bar{y})^\top N_\Lambda(\psi(\bar{x}, \bar{y})) \tag{3.15}$$

provided the validity of the qualification condition

$$\left. \begin{array}{l} \nabla \psi(\bar{x}, \bar{y})^\top (\alpha, \beta, \gamma, \zeta) = 0 \\ (\alpha, \beta, \gamma, \zeta) \in N_\Lambda(\psi(\bar{x}, \bar{y})) \end{array} \right\} \Longrightarrow \alpha = 0, \beta = 0, \gamma = 0, \zeta = 0. \tag{3.16}$$

It is easy to check the equality

$$\nabla \psi(\bar{x}, \bar{y})^\top (\alpha, \beta, \gamma, \zeta) = \nabla L_o(\bar{x}, \bar{y}, \alpha, \beta, \gamma, \zeta)$$

for any quadruple $(\alpha, \beta, \gamma, \zeta)$ and that, by the product formula for limiting normals,

$$N_\Lambda(\psi(\bar{x}, \bar{y})) = N_{\mathbb{R}_-^a}(g(\bar{x}, \bar{y})) \times N_{\{0_b\}}(h(\bar{x}, \bar{y})) \times \prod_{i=1}^d N_{\Lambda_i}(G_i(\bar{x}, \bar{y}), H_i(\bar{x}, \bar{y})).$$

Using the expression of the normal cone to the sets Λ_i, $i = 1, \ldots, d$, from [21] (cf. also [45, 56] for particular cases), we get

$$N_{\Lambda_i}(G_i(\overline{x}, \overline{y}), H_i(\overline{x}, \overline{y})) = \{(\gamma_i, \zeta_i)| \quad \gamma_i = 0 \text{ if } i \in \nu, \ \zeta_i = 0 \text{ if } i \in \eta \\ (\gamma_i < 0, \ \zeta_i < 0) \vee (\gamma_i \zeta_i = 0), \text{ if } i \in \theta\},$$

which implies that the qualification condition (3.16) reduces to (A_1^1) in this case and that inclusion (3.12) in the theorem results from (3.15) and the coderivative definition (2.6). Finally, the Lipschitz-like property of S^c around $(\overline{x}, \overline{y})$ under the additional M-qualification condition (A_2^1) follows from (3.12) due to the coderivative criterion (2.9). $\square$

Now we can readily get efficient estimates of the limiting subdifferential of the value function (3.3) and verifiable conditions for its local Lipschitz continuity.

THEOREM 3.2 (M-type sensitivity analysis for MPCC value functions). *The following assertions hold for the value function μ^c in* (3.3)*:*

(i) *Let the argminimum mapping S_o^c from* (3.10) *be inner semicontinuous at $(\overline{x}, \overline{y})$, and let (A_1^1) hold at $(\overline{x}, \overline{y})$. Then we have the subdifferential upper estimate*

$$\partial \mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta)| (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cm}(\overline{x}, \overline{y})\}.$$

If in addition (A_2^1) is satisfied at $(\overline{x}, \overline{y})$, then μ^c is Lipschitz continuous around $\overline{x}$.

(ii) *Assume that S_o^c is inner semicompact at $\overline{x}$ and that (A_1^1) holds at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^c(x)$. Then we have the subdifferential upper estimate*

$$\partial \mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta)| \overline{y} \in S_o^c(\overline{x}), \ (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cm}(\overline{x}, \overline{y})\}.$$

If in addition (A_2^1) is satisfied at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^c(\overline{x})$, then the value function μ^c is Lipschitz continuous around $\overline{x}$.

Proof. It follows from the results of Theorem 3.1 and Theorem 2.1. $\square$

Note that a subdifferential upper estimate similar to assertion (ii) Theorem 3.2 was obtained in [31] in the case of $G(x, y) := y$ under a certain growth hypothesis implying the inner semicompactness of the optimal solution map S_o^c from (3.10).

REMARK 3.3. *We do not pay any special attention to the lower semicontinuity of the value function* (3.3) *in Theorem 3.2 and subsequent results on value functions. By Theorem 2.1 this easily follows from the proof under the inner semicontinuity or the weaker inner semicompactness of the solution map S_o^c. There are various sufficient conditions for the validity of the qualification condition (A_1^1); see, e.g., [13]. Furthermore, (A_1^1) can be replaced by the weaker* calmness *assumption on the mapping*

$$\Phi(\vartheta) := \{(x, y)| \psi(x, y) + \vartheta \in \Lambda\}, \tag{3.17}$$

where ψ and Λ are defined in the proof of Theorem 3.1. Indeed, it is shown in [25, Theorem 4.1] that the calmness of (3.17) *is sufficient for inclusion* (3.15)*, which thus ensures the conclusions of Theorem 3.1 and 3.2 by the proofs above. Note that the latter calmness assumption automatically holds when the mappings g, h, G, and H are linear. Observe finally that due to the relationships* (3.11) *both assumptions (A_1^1) and (A_2^1) can be replaced by the fulfillment of the single condition (A_3^1).*

REMARK 3.4. *Following the pattern of Theorem 2.1, the basic difference between the upper estimate of $\partial \mu^c$ in assertions (i) and (ii) of Theorem 3.2 resides in the fact that in the first case we have to compute the gradient of the Lagrange-type function L associated with the MPCC* (3.1) *only at the point $(\overline{x}, \overline{y})$ where S_o^c is inner semicontinuous. In the second case though this should be done at all $(\overline{x}, y)$ with $y \in S_o^c(\overline{x})$.*

Thus the upper bound of $\partial\mu^c$ obtained under the inner semicontinuity is obviously much tighter since it is always a subset of the one in (ii). As already mentioned in Section 2, the inner semicontinuity of S_o^c is automatically satisfied if this mapping is Lipschitz-like around the point in question. Moreover, if S_o^c is inner semicompact at $\overline{x}$ and $S_o^c(\overline{x}) = \{\overline{y}\}$, then S_o^c is inner semicontinuous at $(\overline{x}, \overline{y})$.

REMARK 3.5. *The technique employed in Theorem 3.1 that transforms the complementarity system (3.13) into inclusion (3.14) is rather common in the field of MPCCs to study some issues different from those considered here. It is used, e.g., by Ye and Ye [59] and Ye [56] to derive necessary optimality conditions for the KKT reformulation of the classical optimistic bilevel problem (P), which is a special case of the unperturbed version of problem (3.1). Outrata [45] also uses a similar technique while studying constraint qualifications and optimality conditions for a unperturbed version of problem (3.1) with $G(x, y) = y$. Some differences occur in constructing the set Λ corresponding, in the proof of Theorem 3.1, to $\{(u, v) \in \mathbb{R}^d \,|\, u \geq 0, \, v \geq 0, \, u^\top v = 0\}$ while in the aforementioned papers $\Lambda = gph\,N_{\mathbb{R}_+^d}$. Note also in [28] a transformation in the vein of (3.14) is employed to derive an exact penalty result and then optimality conditions for the so-called mathematical programs with vanishing constraints; see [1]. Having in mind this transformation, the methods developed in our paper (cf., in particular, the proofs of Theorem 3.1 and Theorem 3.2) can readily be applied to conduct a local sensitivity analysis for the latter class of programs.*

3.2. Sensitivity analysis via C-type multipliers. Similarly to Subsection 3.1 we introduce the following *C-qualification conditions* at $(\overline{x}, \overline{y})$:

(A_1^2) $(\alpha, \beta, \gamma, \zeta) \in \Lambda^{cc}(\overline{x}, \overline{y}, 0) \Longrightarrow \alpha = 0, \, \beta = 0, \, \gamma = 0, \, \zeta = 0$;

(A_2^2) $(\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cc}(\overline{x}, \overline{y}, 0) \Longrightarrow \nabla_x L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) = 0$;

(A_3^2) $(\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cc}(\overline{x}, \overline{y}, 0) \Longrightarrow \alpha = 0, \, \beta = 0, \, \gamma = 0, \, \zeta = 0$

with the similar relationships between them:

$$(A_2^2) \Longleftarrow (A_3^2) \Longrightarrow (A_1^2).$$

To proceed, we use the well-known nonsmooth transformation of the feasible set to the MPCC introduced by Scheel and Scholtes [51]:

$$S^c(x) := \{y \in \mathbb{R}^m \,|\quad g(x, y) \leq 0, \, h(x, y) = 0, \\ \min\{G_i(x, y), H_i(x, y)\} = 0, \, i = 1, \ldots, d\}. \tag{3.18}$$

Employing this transformation, a *C*-counterpart of Theorem 3.2 can be derived with a different proof and a larger estimate for the coderivative of S^c under the *C*-qualification conditions.

THEOREM 3.6 (coderivative estimate and Lipschitz-like property of MPCC feasible solutions via C-multipliers). *Let $(\overline{x}, \overline{y}) \in gph\,S^c$, and let (A_1^2) hold at $(\overline{x}, \overline{y})$. Then we have for all $y^* \in \mathbb{R}^m$*

$$D^*S^c(\overline{x}, \overline{y})(y^*) \subset \{\nabla_x L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \,|\, (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cc}(\overline{x}, \overline{y}, y^*)\}.$$

If in addition (A_2^2) is satisfied at $(\overline{x}, \overline{y})$, then S^c is Lipschitz-like around this point.

Proof. From the expression of S^c in (3.18) we get

$$gph\,S^c = \{(x, y) \,|\, \psi(x, y) \in \Lambda\},$$

where ψ and Λ are defined by

$$\psi(x, y) := [g(x, y), \, h(x, y), \, V(x, y)] \text{ and } \Lambda := \mathbb{R}_-^a \times \{0_b\} \times \{0_d\} \tag{3.19}$$

with $V_i(x,y) := \min\{G_i(x,y), H_i(x,y)\} = 0$ for $i = 1, \ldots, d$. Since ψ is locally Lipschitzian around $(\bar{x}, \bar{y})$, it follows from [38, Theorem 3.8] that

$$N_{\text{gph }S^c}(\overline{x}, \overline{y}) \subset \{\partial\langle u^*, \psi\rangle(\overline{x}, \overline{y}) \,|\, u^* \in N_\Lambda(\psi(\overline{x}, \overline{y}))\} \tag{3.20}$$

provided that the qualification condition

$$\left. \begin{array}{l} 0 \in \partial\langle u^*, \psi\rangle(\overline{x}, \overline{y}) \\ u^* \in N_\Lambda(\psi(\overline{x}, \overline{y})) \end{array} \right\} \implies u^* = 0 \tag{3.21}$$

is satisfied. Furthermore, we have the normal cone representation

$$N_\Lambda(\psi(\overline{x}, \overline{y})) = N_{\mathbb{R}_-^a}(g(\overline{x}, \overline{y})) \times N_{\{0_b\}}(h(\overline{x}, \overline{y})) \times N_{\{0_d\}}(V(\overline{xy})) \tag{3.22}$$

and calculate the subdifferential of the scalarization in (3.20) by

$$\partial\langle(\alpha, \beta, \chi), \psi\rangle(\overline{x}, \overline{y}) = \nabla g(\overline{x}, \overline{y})^\top \alpha + \nabla h(\overline{x}, \overline{y})^\top \beta + \partial\langle\chi, V\rangle(\overline{x}, \overline{y}) \tag{3.23}$$

for $(\alpha, \beta, \chi) \in N_\Lambda(\psi(\overline{x}, \overline{y}))$. Since the function V is nondifferentiable and χ may contain negative components by (3.22), we apply the convex hull "co" to our basic subdifferential (2.2) in (3.23) in order to instate the plus/minus symmetry

$$\partial\langle\chi, V\rangle(\overline{x}, \overline{y}) \subset \text{co }\partial\langle\chi, V\rangle(\overline{x}, \overline{y}) \subset \sum_{i=1}^d \chi_i \overline{\partial} V_i(\overline{x}, \overline{y})$$

via Clarke's generalized gradient $\overline{\partial} V_i$. Considering the partition of the index set $\{1, \ldots, d\}$ in (3.4), we arrive by [5] at the following calculations:

$$\overline{\partial} V_i(\overline{x}, \overline{y}) = \begin{cases} \nabla G_i(\overline{x}, \overline{y}) & \text{if } i \in \eta, \\ \nabla H_i(\overline{x}, \overline{y}) & \text{if } i \in \nu, \\ \text{co }\{\nabla G_i(\overline{x}, \overline{y}), \nabla H_i(\overline{x}, \overline{y})\} & \text{if } i \in \theta. \end{cases}$$

Invoking the classical Carathéodory theorem gives us

$$\text{co }\{\nabla G_i(\overline{x}, \overline{y}), \nabla H_i(\overline{x}, \overline{y})\} = \{t_i \nabla G_i(\overline{x}, \overline{y}) + (1 - t_i) \nabla H_i(\overline{x}, \overline{y}) \,|\, t_i \in [0,1]\},$$

and hence we obtain from (3.23) the inclusions

$$\begin{aligned} \partial\langle(\alpha, \beta, \chi), \psi\rangle(\overline{x}, \overline{y}) \subset \quad & \{\nabla L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \,|\, \gamma_\eta = 0, \zeta_\nu = 0 \\ & \forall i \in \theta, \exists t_i \in [0,1], r_i \in \mathbb{R} \text{ s.t. } \gamma_i = r_i t_i, \zeta_i = r_i(1 - t_i)\} \\ \subset \quad & \{\nabla L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \,|\, \gamma_\eta = 0, \zeta_\nu = 0 \\ & \forall i \in \theta, \gamma_i \zeta_i \ge 0\}. \end{aligned} \tag{3.24}$$

Since the qualification condition (3.21) is equivalent to

$$\{(\alpha, \beta, \chi) \,|\, 0 \in \partial\langle(\alpha, \beta, \chi), \psi\rangle(\overline{x}, \overline{y}), \, (\alpha, \beta, \chi) \in N_\Lambda(\psi(\overline{x}, \overline{y}))\} = \{(0,0,0)\},$$

the second inclusion in (3.24) shows that (A_1^2) is sufficient for this to hold. Furthermore, by (3.20) the second inclusion of (3.24) leads to an upper estimate of $N_{\text{gph }S^c}$, which allows us via the coderivative definition (2.6) to recover the upper bound of $D^* S^c$ in the theorem. The latter implies the Lipschitz-like property of S^c under (A_2^2) as in Theorem 3.1. $\square$

As in the previous subsection, we arrive at the following sensitivity results for the MPCC value function (3.3) via C-multipliers.

THEOREM 3.7 (C-type sensitivity analysis for MPCC value functions). *The following assertions hold for the value function μ^c in (3.3):*

(i) *Let the optimal solution map S_o^c is inner semicontinuous at $(\overline{x}, \overline{y})$, and let (A_1^2) holds at $(\overline{x}, \overline{y})$. Then we have the subdifferential upper estimate*

$$\partial \mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \mid (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cc}(\overline{x}, \overline{y})\}.$$

If in addition (A_2^2) holds at $(\overline{x}, \overline{y})$, then μ^c is Lipschitz continuous around $\overline{x}$.

(ii) *Assume that S_o^c is inner semicompact at $\overline{x}$ and that (A_1^2) holds at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^c(x)$. Then we have the subdifferential upper estimate*

$$\partial \mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \mid \overline{y} \in S_o^c(\overline{x}), \ (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cc}(\overline{x}, \overline{y})\}.$$

If in addition (A_2^2) also holds at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^c(\overline{x})$, then μ^c is Lipschitz continuous around $\overline{x}$.

Proof. It follows from the results of Theorem 3.7 and Theorem 2.1. □

Note that assertion (ii) of Theorem 3.7 can be found in [31] for $G(x, y) = y$ under the following assumption corresponding to the replacement of the set $\Lambda^{cc}(\overline{x}, \overline{y}, 0)$ in (A_1^2) by

$$\{(\alpha, \beta, \gamma, \zeta) \mid \quad \alpha \geq 0, \ \alpha^\top g(\overline{x}, \overline{y}) = 0,$$
$$\gamma_\eta = 0, \zeta_\nu = 0,$$
$$\forall i \in \theta, \exists t_i \in [0, 1], r_i \in \mathbb{R} \text{ s.t. } \gamma_i = r_i t_i, \ \zeta_i = r_i(1 - t_i),$$
$$\nabla g(\overline{x}, \overline{y})^\top \alpha + \nabla h(\overline{x}, \overline{y})^\top \beta + \nabla G(\overline{x}, \overline{y})^\top \gamma + \nabla H(\overline{x}, \overline{y})^\top \zeta = 0\}.$$

The latter assumption is weaker than (A_1^2), but in our assumption we simply need to check that the components of γ and ζ are of the same sign on θ rather than constructing them as in the above set. It is also important to mention that all the points made in Remark 3.3 can be restated here accordingly. In particular, (A_1^2) can be substituted by the weaker calmness of the set-valued mapping Φ from (3.17) with ψ and Λ given in (3.19). This is obviously satisfied if the functions g, h, G, and H are linear, because the one of $V_i(x, y) = \min\{G_i(x, y), H_i(x, y)\}$ is piecewise linear provided the linearity of G_i and H_i.

3.3. Sensitivity analysis via S-type multipliers. The need for S-type stationarity conditions in the context of MPCCs is the best one would want to have since these conditions are *equivalent* to the KKT type optimality conditions whenever the MPCC is treated as an ordinary nonlinear optimization problem.

Having this in mind, we attempt here to suggest a tighter upper bound for the basic subdifferential of the MPCC value function μ^c. In order to obtain an upper bound for $\partial \mu^c$ containing $\Lambda_y^{cs}(\overline{x}, \overline{y})$ rather than $\Lambda_y^{cm}(\overline{x}, \overline{y})$ or $\Lambda_y^{cc}(\overline{x}, \overline{y})$, we impose the following *S-qualification condition* with the index set I defined by $I := I(\overline{x}, \overline{y}) := \{i = 1, \ldots, a \mid g_i(\overline{x}, \overline{y}) < 0\}$:

(A_1^3) $\quad \left. \begin{array}{l} \nabla L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) = 0 \\ \alpha_I = 0, \ \gamma_\nu = 0, \ \zeta_\eta = 0 \end{array} \right\} \implies \gamma_\theta = 0, \ \zeta_\theta = 0$

introduced by Ye [56] and later named in [58] as *Partial MPEC-LICQ* (Linear Independence Constraint Qualification). This condition and another close while weaker one have also been used by Flegel, Kanzow and Outrata [22] to recover the S-stationarity conditions of a MPCC from the M-ones. In the next theorem we obtain

a new S-type upper bound for $\partial\mu^c$ by a similar methodology, i.e., going from the M-type bound provided above. Note that assumption (A_1^1) is the one introduced in Subsection 3.1.

THEOREM 3.8 (S-type sensitivity analysis for MPCC value functions). *The following assertions hold for the value function μ^c from (3.3):*

(i) *Let the optimal solution map S_o^c be inner semicontinuous at $(\overline{x}, \overline{y})$, and let assumptions (A_1^1) and (A_1^3) be satisfied at $(\overline{x}, \overline{y})$. Then we have*

$$\partial\mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \,|\, (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cs}(\overline{x}, \overline{y})\}.$$

(ii) *Let S_o^c be inner semicompact at $\overline{x}$ with (A_1^1) and (A_1^3) being satisfied at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^c(\overline{x})$. Then we have*

$$\partial\mu^c(\overline{x}) \subset \{\nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) \,|\, \overline{y} \in S_o^c(\overline{x}), \, (\alpha, \beta, \gamma, \zeta) \in \Lambda_y^{cs}(\overline{x}, \overline{y})\}.$$

Proof. We provide the proof only for assertion(i); the other case can be proved similarly.

Assuming (A_1^1) and the inner semicontinuity of S_o^c, we have the upper estimate of $\partial\mu^c$ from Theorem 3.2(i). Further, denote by $A(\overline{x}, \overline{y})$ (resp. $B(\overline{x}, \overline{y})$) the right-hand side of the inclusion in Theorem 3.2(i) (resp. Theorem 3.8(i)). It remains to show that $A(\overline{x}, \overline{y}) = B(\overline{x}, \overline{y})$, under the S-qualification condition (A_1^3).

We obviously have $A(\overline{x}, \overline{y}) \supset B(\overline{x}, \overline{y})$. To justify the opposite inclusion, pick any $a(\alpha, \beta, \gamma, \zeta) \in A(\overline{x}, \overline{y})$ and search for $b(\alpha^o, \beta^o, \gamma^o, \zeta^o) \in B(\overline{x}, \overline{y})$ such that $a(\alpha, \beta, \gamma, \zeta) = b(\alpha^o, \beta^o, \gamma^o, \zeta^o)$. If the latter equality were to hold, we would get

$$\begin{cases} \nabla L_o(\overline{x}, \overline{y}, \alpha - \alpha^o, \beta - \beta^o, \gamma - \gamma^o, \zeta - \zeta^o) = 0, \\ \alpha_I^o - \alpha_I^o = 0, \, \gamma_\nu - \gamma_\nu^o = 0, \, \zeta_\eta - \zeta_\eta^o = 0. \end{cases}$$

Thus it follows from (A_1^3) that $\gamma_\theta^o = \gamma_\theta$ and $\zeta_\theta^o = \zeta_\theta$. To conclude the proof, choose $\alpha^o := \alpha$, $\beta^o := \beta$, $\gamma_{\theta^c}^o := \gamma_{\theta^c}$ and $\zeta_{\theta^c}^o := \zeta_{\theta^c}$ with $\theta^c := \{i = 1, \ldots, d\} \setminus \theta$. $\square$

We can see from the proof that it can be repeated with using the C-type upper bound in Subsection 3.2 instead of the M-one. This shows that under the assumption (A_1^3) all the S-type, M-type, and C-type upper bounds for $\partial\mu^c$ are the same. It places us in the situation similar to that already recognized in the context of the various types of stationarity concepts known for MPCCs: they agree with each other an appropriate assumption.

We also mention two possibilities for the local Lipschitz continuity of μ^c in the framework of Theorem 3.8. The first one is either to replace (A_1^1) by (A_3^1) or to add (A_2^1) to the assumptions; cf. (3.11) and Theorem 3.2. The second possibility is to replace (A_1^3) by the following stronger qualification condition:

$$\left. \begin{array}{r} \nabla_y L_o(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta) = 0 \\ \alpha_I = 0, \, \gamma_\nu = 0, \, \zeta_\eta = 0 \end{array} \right\} \implies \alpha = 0, \, \beta = 0, \, \gamma = 0, \, \zeta = 0. \qquad (3.25)$$

The latter condition corresponding to the well-known MPEC-LICQ for the parametric MPCC (3.1) has the advantage, in the framework of Theorem 3.8(i), to ensure even more than the Lipschitz continuity of μ^c; namely, its *strict differentiability* as stated in the next corollary.

COROLLARY 3.9 (S-type sensitivity analysis for MPCC value functions under the MPEC-LICQ). *Assume that S_o^c is inner semicontinuous at the point $(\overline{x}, \overline{y})$, where the*

MPEC-LICQ (3.25) is also satisfied. Then the value function μ^c is strictly differentiable at $\overline{x}$ with

$$\nabla \mu^c(\overline{x}) = \nabla_x L(\overline{x}, \overline{y}, \alpha, \beta, \gamma, \zeta),$$

where $(\alpha, \beta, \gamma, \zeta)$ is the unique multiplier of the set $\Lambda_y^{cs}(\overline{x}, \overline{y})$.

Proof. We can see that the set on the right-hand side of the inclusion in Theorem 3.8(i) is a singleton; hence $\partial \mu^c(\bar{x})$ is a singleton as well. Since the value function μ^c is surely locally Lipschitzian around $\bar{x}$ under the MPEC-LICQ (3.25), the latter uniqueness ensures its strict differentiability at this point; see Section 2. $\square$

In case of (ii) we additionally need $S_o^c(\overline{x})$ to be a singleton to ensure the strict differentiability of μ^c at $\overline{x}$. The latter corresponds to the framework provided by Hu and Ralph [29], and hence it shows (see Remark 3.4) that the assumptions imposed in [29] imply the inner semicontinuity of the set-valued mapping S_o^c at the solution point. Note also that assertion (ii) of Theorem 3.8 closely relates to the corresponding result of [31] obtained in a particular case from a different perspective. Finally, we mention that the S-qualification condition (A_1^3) does *not imply* the equalities between the multiplier sets in (3.9); for this we need the stronger assumption consisting in replacing the gradients of g, h, G, and H in (A_1^3) by their partial gradients with respect to the y-variable.

4. Sensitivity analysis of OPEC value functions. This section is devoted to the study of the following parametric *optimization problem with generalized equation constraints* (OPEC):

$$\min_y \{F(x, y) \,|\, 0 \in h(x, y) + N_{K(x)}(y)\}, \tag{4.1}$$

where $F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ are continuously differentiable functions, and K denotes a set-valued mapping (moving set) defined by

$$K(x) := \{y \in \mathbb{R}^m \,|\, g(x, y) \leq 0\} \tag{4.2}$$

with g also continuously differentiable and $g(x, .)$ convex for all $x \in \mathbb{R}^n$. Note that model (4.1) is written in the form of *quasi-variational inequalities* described by the normal cone to moving sets; see, e.g., [43] and the references therein. On the other hand, problem (4.1) is closely related to the MPCC considered in the previous section. Indeed, it has been well recognized that the complementarity system (3.13) can equivalently be written as

$$0 \in -G(x, y) + N_{\mathbb{R}_-^d}(-H(x, y)),$$

which is in the form of OPEC constraints in (4.1) with the normal cone to the constant nonnegative orthant. In the other direction, by replacing the normal cone in (4.1) by its well-known expression

$$N_{K(x)}(y) = \{\nabla_y g(x, y)^\top u \,|\, u \geq 0, \, u^\top g(x, y) = 0\} \tag{4.3}$$

under a certain constraint qualification, we get a particular case of problem (3.1). Despite this equivalence, sensitivity analysis of the OPEC optimal value function

$$\mu^e(x) := \min_y \{F(x, y) \,|\, 0 \in h(x, y) + N_{K(x)}(y)\} \tag{4.4}$$

associated with problem (4.1) in its given form is of independent interest. Indeed, in this way we obtain different estimates for the limiting subdifferential of the two-level value function φ_o from (1.4), which is of our main attention in this paper. This issue will be comprehensively discussed in the next section.

To present our main result in this section on the generalized differentiation and Lipschitz continuity of the value function μ^e, we proceed similarly to Section 3 and consider first the *feasible solution map* of the parametric generalized equation in (4.1) defined by

$$S^e(x) := \{y \in \mathbb{R}^m \mid 0 \in h(x,y) + N_{K(x)}(y)\}. \tag{4.5}$$

A detailed study of the robust Lipschitzian stability of (4.5) based on the coderivative analysis has been carried out by Mordukhovich and Outrata [43]. Note that the work in [43] heavily relies on an estimate of the coderivative of normal cone mapping $(x,y) \rightrightarrows N_{K(x)}(y)$ given therein. Before introducing the rules to be used here (which emerged from [43]), some notation is necessary to simplify the presentation. Define $\mathcal{L}(x,y,u) := h(x,y) + \nabla_y g(x,y)^\top u$ and consider the set of *lower-level Lagrange multipliers*

$$\Lambda(\overline{x}, \overline{y}) := \{u \mid \mathcal{L}(\overline{x}, \overline{y}, u) = 0,\ u \geq 0,\ u^\top g(\overline{x}, \overline{y}) = 0\}. \tag{4.6}$$

Similarly to the previous section, we partition the indices of the functions involved in the complementarity system of (4.6) as follows:

$$
\begin{aligned}
\eta :=&\ \eta(\overline{x}, \overline{y}, \overline{u}) := \{i = 1, \ldots, p \mid & \overline{u}_i = 0,\ g_i(\overline{x}, \overline{y}) < 0\}, \\
\theta :=&\ \mu(\overline{x}, \overline{y}, \overline{u}) := \{i = 1, \ldots, p \mid & \overline{u}_i = 0,\ g_i(\overline{x}, \overline{y}) = 0\}, \\
\nu :=&\ \nu(\overline{x}, \overline{y}, \overline{u}) := \{i = 1, \ldots, p \mid & \overline{u}_i > 0,\ g_i(\overline{x}, \overline{y}) = 0\}.
\end{aligned}
\tag{4.7}
$$

Consider also the system of relationships that play an important role in the sequel:

$$z^* + \nabla g(\overline{x}, \overline{y})^\top \beta + \nabla_{x,y} \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0, \tag{4.8}$$

$$\nabla_y g_\nu(\overline{x}, \overline{y})\gamma = 0,\ \beta_\eta = 0, \tag{4.9}$$

$$\forall i \in \theta,\ (\beta_i > 0 \wedge \nabla_y g_i(\overline{x}, \overline{y})\gamma > 0) \vee \beta_i(\nabla_y g_i(\overline{x}, \overline{y})\gamma) = 0. \tag{4.10}$$

The corresponding set of multipliers, which are of a *special M-type*, are defined by:

$$\Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, z^*) := \{(\beta, \gamma) \mid (4.8) - (4.10) \ \text{hold}\}.$$

Similarly to Section 3, we further define $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, y^*)$, with $y^* \in \mathbb{R}^m$, by replacing (4.8) with $y^* + \nabla_y g(\overline{x}, \overline{y})^\top \beta + \nabla_y \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0$ and then set $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}) := \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, \nabla_y F(\overline{x}, \overline{y}))$. The following *EM-qualification conditions* deduced from [43] can be formulated as:

(A_1^4) $[\nabla_y g(\overline{x}, \overline{y})^\top \beta = 0,\ \beta \geq 0, \beta^\top g(\overline{x}, \overline{y}) = 0] \Longrightarrow \beta = 0$;

(A_2^4) $\forall \overline{u} \in \Lambda(\overline{x}, \overline{y}) : [\nabla g(\overline{x}, \overline{y})^\top \beta = 0,\ \beta_\eta = 0] \Longrightarrow \beta = 0$;

(A_3^4) $\forall \overline{u} \in \Lambda(\overline{x}, \overline{y}) : (\beta, \gamma) \in \Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0) \Longrightarrow \beta = 0,\ \gamma = 0$;

(A_4^4) $[\overline{u} \in \Lambda(\overline{x}, \overline{y}),\ (\beta, \gamma) \in \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0)] \Longrightarrow \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0$;

(A_5^4) $[\overline{u} \in \Lambda(\overline{x}, \overline{y}),\ (\beta, \gamma) \in \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0)] \Longrightarrow \beta = 0,\ \gamma = 0$.

It is easy to observe the relationships between these qualification conditions :

$$(A_4^4) \Longleftarrow (A_5^4) \Longrightarrow (A_3^4),\ (A_2^4).$$

We are now ready to establish the main result of this section, where S_o^e denotes the *optimal solution map* to the parametric optimization problem (4.1) given by

$$S_o^e(x) := \{y \in S^e(x)|\, F(x,y) - \mu^e(x) \le 0\}.$$

THEOREM 4.1 (M-type sensitivity analysis for OPEC value functions). *The following assertions hold for the value function μ^e from (4.5):*

(i) *Let the optimal solution map S_o^e be inner semicontinuous at the point $(\overline{x}, \overline{y})$, where the qualification conditions (A_1^4)–(A_3^4) are satisfied. Then we have the subdifferential estimate*

$$\partial\mu^e(\overline{x}) \subset \bigcup_{\overline{u} \in \Lambda(\overline{x},\overline{y})} \bigcup_{(\beta,\gamma) \in \Lambda_y^{em}(\overline{x},\overline{y},\overline{u})} \left\{ \nabla_x F(\overline{x},\overline{y}) + \nabla_x g(\overline{x},\overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x},\overline{y},\overline{u})^\top \gamma \right\}.$$

If in addition (A_4^4) holds at $(\overline{x}, \overline{y})$, then μ^e is Lipschitz continuous around $\overline{x}$.

(ii) *Let S_o^e be inner semicompact at $\overline{x}$, and let (A_1^4)–(A_3^4) be satisfied at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^e(\overline{x})$. Then we have the subdifferential estimate*

$$\partial\mu^e(\overline{x}) \subset \bigcup_{\overline{y} \in S_o(\overline{x})} \bigcup_{\overline{u} \in \Lambda(\overline{x},\overline{y})} \bigcup_{(\beta,\gamma) \in \Lambda_y^{em}(\overline{x},\overline{y},\overline{u})} \Big\{ \nabla_x F(\overline{x},\overline{y}) + \nabla_x g(\overline{x},\overline{y})^\top \beta$$
$$+ \nabla_x \mathcal{L}(\overline{x},\overline{y},\overline{u})^\top \gamma \Big\}.$$

If (A_4^4) is also satisfied at $(\overline{x}, \overline{y})$ for all $\overline{y} \in S_o^e(\overline{x})$, then μ^e is Lipschitz continuous around $\overline{x}$.

Proof. We justify only assertion (i); the one in (ii) can be proved similarly. Since F is continuously differentiable and S_o^e is inner semicontinuous, it follows from Theorem 2.1(i) that

$$\partial\mu^e(\overline{x}) \subset \nabla_x F(\overline{x},\overline{y}) + D^* S^e(\overline{x},\overline{y})(\nabla_y F(\overline{x},\overline{y})). \tag{4.11}$$

Applying further [43, Theorem 4.3] to the solution map S^e and taking into account that the EM-qualification conditions (A_1^4)–(A_3^4) are satisfied, we get the coderivative estimate

$$D^* S^e(\overline{x},\overline{y})(\nabla_y F(\overline{x},\overline{y})) \subset \bigcup_{\overline{u} \in \Lambda(\overline{x},\overline{y})} \bigcup_{(\beta,\gamma) \in \Lambda_y^{em}(\overline{x},\overline{y},\overline{u})} \left\{ \nabla_x g(\overline{x},\overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x},\overline{y},\overline{u})^\top \gamma \right\}.$$
$$\tag{4.12}$$

Then the upper estimate of the basic subdifferential of μ^e in the theorem follows by combining (4.11) and (4.12). The local Lipschitz continuity of μ^e around $\bar{x}$ also follows from Theorem 2.1 (i) by recalling [43] that S^e is Lipschitz-like around $(\overline{x}, \overline{y})$ under (A_1^4)–(A_4^4). $\square$

To the best of our knowledge, the first result in the direction of Theorem 4.1(ii) goes back to Lucet and Ye [31], where a similar subdifferential estimate was obtained under a growth hypothesis (implying the inner semicompactness of S_o^e) for a particular case of the problem under consideration. Note however that their result deals only with the case where K is *independent* of x. Assertion (i) of Theorem 4.1 clearly provides a tighter subdifferential upper bound under the inner semicontinuity assumption. We also mention the work by Mordukhovich, Nam and Yen [42] in the framework, where the regular and limiting subdifferentials of μ^e are estimated in the case of

$$S^e(x) := \{y|\, 0 \in h(x,y) + Q(x,y)\}$$

in (4.4) with a general set-valued mapping $Q(x,y)$ not specified to our setting $Q(x,y) := N_{K(x)}(y)$ in terms of the initial data of (4.3).

REMARK 4.2. *Following Mordukhovich and Outrata [43], the qualification condition (A_2^4) in Theorem 4.1 can be replaced by the weaker* calmness property *of the following set-valued mapping at $(0, \overline{x}, \overline{y}, \overline{u})$, for all $\overline{u} \in \Lambda(\overline{x}, \overline{y})$:*

$$M(\vartheta) := \left\{ (x,y,u) \,\middle|\, \left[\begin{array}{c} g(x,y) \\ u \end{array} \right] + \vartheta \in gph\, N_{\mathbb{R}_-^p} \right\}.$$

Similarly, condition (A_3^4) can be replaced by the calmness property of the mapping

$$P(z, \vartheta) := \left\{ (x,y,u) \,\middle|\, \left[\begin{array}{c} \mathcal{L}(x,y,u) \end{array} \right] + z = 0 \right\} \cap M(\vartheta)$$

at $(0, 0, \overline{x}, \overline{y}, \overline{u})$ for all $\overline{u} \in \Lambda(\overline{x}, \overline{y})$. Both calmness assumptions are automatic when the mappings g and $(x,y) \mapsto \nabla_y f(x,y)$ are linear.

5. Sensitivity analysis of two-level optimal value functions. Our main concern in this section is to conduct a local sensitivity analysis of the two-level optimal value function

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

defined in (1.4), where S is the optimal solution map of the lower-level problem (1.2) constrained by $y \in K(x)$ with $K(x)$ defined in (1.3). We explore all the three approaches to this issue discussed in Section 1.

5.1. MPCC and OPEC approaches. ¿From here and for the rest of this subsection we assume the lower-level problem (1.2) with $K(x)$ given by (1.3) is *convex*, i.e., the functions $f(x, .)$ and $g(x, .)$ are convex for all $x \in X$. Most of the notation below is either taken from Sections 4 or closely related to it. To be more precise, from here on the lower-level Lagrange multipliers set $\Lambda(\overline{x}, \overline{y})$ is considered as in (4.6) while the index sets η, θ and ν are given in (4.7). The lower-level Lagrangian $\mathcal{L}$ is considered now in the form $\mathcal{L}(x,y,u) := \nabla_y f(x,y) + \nabla_y g(x,y)^\top u$, i.e., with $h(x,y) := \nabla_y f(x,y)$. The next lemma involving $\mathcal{L}$ is useful in what follows.

LEMMA 5.1 (representation of the two-level value function). *Let $x \in X$ from (1.3), and let (A_1^4) be satisfied at all (x,y) with $y \in S(x)$. Then we have*

$$\varphi_o(x) = \min_{y,u} \{ F(x,y) | \quad \mathcal{L}(x,y,u) = 0 \tag{5.1}$$
$$u \geq 0,\, g(x,y) \leq 0,\, u^\top g(x,y) = 0 \}.$$

Proof. Fix $\overline{x} \in X$ and let $\overline{y}$ be a global optimal solution to the problem

$$\min_y \{ F(\overline{x}, y) | y \in S(\overline{x}) \}.$$

Then we have the relationships

$$\begin{aligned}
\varphi_o(\overline{x}) = \quad & F(\overline{x}, \overline{y}), \\
\leq \quad & F(\overline{x}, y) : \forall y \in S(\overline{x}), \\
\leq \quad & F(\overline{x}, y) : \forall y \text{ with } 0 \in \nabla_y f(\overline{x}, y) + N_{K(\overline{x})}(y) \\
& \qquad \text{(by convexity of } f(\overline{x}, .) \text{ and } g(\overline{x}, .)), \\
\leq \quad & F(\overline{x}, y) : \forall (y, u) \text{ with } \mathcal{L}(\overline{x}, y, u),\, u \geq 0, g(\overline{x}, y) \leq 0,\, u^\top g(\overline{x}, y) = 0,
\end{aligned}$$

where the last inequality is due to the normal cone representation (4.3) by taking into account that (A_1^4) holds at all $(\overline{x}, y)$ with $y \in S(\overline{x})$. $\square$

Having this transformation of the two-level value function φ_o, at least two observations can be made. First we note that for each $x \in X$ the value of $\varphi_o(x)$ is obtained from a *global solution* to the parametric problem

$$\min_{y,u}\{F(x,y)| \quad \mathcal{L}(x,y,u) = 0,$$
$$u \geq 0,\, g(x,y) \leq 0,\, u^\top g(x,y) = 0\}. \tag{5.2}$$

Thus the major difficulty arising when establishing the link between local solutions of the auxiliary problem (P) and its KKT reformulation (see [10] for details) does *not* appear here. Secondly, the presence of the complementarity constraints $u \geq 0,\, g(x,y) \leq 0,\, u^\top g(x,y) = 0$ in (5.2) leads to the violation of the MFCQ, while the results of Section 3 can be applied. To proceed, consider the *feasible solution map* associated with (5.2) by

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

and the *optimal solution map* of (5.2) given by

$$S_o^h(x) := \{(y,u) \in S^h(x)|\, F(x,y) \leq \varphi_o(x)\}. \tag{5.3}$$

Now we establish M-type sensitivity results for the two-level value function φ_o, which are *crucial* in the paper. The multiplier sets $\Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, z^*)$, $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$, and $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})$ used in the next theorem are exactly the ones defined in Section 4.

THEOREM 5.2 (M-type sensitivity analysis for two-level value functions via the MPCC reformulation). *Assume that (A_1^4) is satisfied at all (x,y), $y \in S(x)$, that the optimal solution map S_o^h is inner semicontinuous at $(\overline{x}, \overline{y}, \overline{u})$, and the implication*

$$(\beta, \gamma) \in \Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0) \Longrightarrow \beta = 0,\, \gamma = 0 \tag{5.4}$$

holds at $(\overline{x}, \overline{y}, \overline{u})$. Then the limiting subdifferential of φ_o is estimated by

$$\partial \varphi_o(\overline{x}) \subset \bigcup_{(\beta,\gamma) \in \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})} \left\{\nabla_x F(\overline{x}, \overline{y}) + \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma\right\}. \tag{5.5}$$

Furthermore, φ_o is Lipschitz continuous around $\overline{x}$ provided that the following qualification condition is also satisfied at $(\overline{x}, \overline{y}, \overline{u})$:

$$(\beta, \gamma) \in \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0) \Longrightarrow \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0. \tag{5.6}$$

Proof. By setting $y := (y, u)$ in the framework of Theorem 3.2, we simply need to specify the various multiplier sets therein to our setting. It follows from Lemma 5.1 that

$$\varphi_o(x) = \min_{y,u}\{F(x,y)| \quad h(x,y,u) = 0,$$
$$G(x,y,u) \geq 0,\, H(x,y,u) \geq 0,\, G(x,y,u)^\top H(x,y,u) = 0\},$$

where $h(x,y,u) := \mathcal{L}(x,y,u)$, $G(x,y,u) := u$, and $H(x,y,u) := -g(x,y)$. Then using the notations of Section 3, we have

$$\Lambda^{cm}(\overline{x}, \overline{y}, \overline{u}, x^*, y^*, u^*) = \{(\beta, \gamma, \zeta)| \quad \zeta_\nu = 0, \beta_\eta = 0,$$
$$(\zeta_i < 0,\, \beta_i < 0) \vee (\beta_i \zeta_i = 0),\, \forall i \in \theta,$$
$$x^* + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma - \nabla_x g(\overline{x}, \overline{y})^\top \beta = 0,$$
$$y^* + \nabla_y \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma - \nabla_y g(\overline{x}, \overline{y})^\top \beta = 0,$$
$$u^* + \nabla_y g(\overline{x}, \overline{y})\gamma + \zeta = 0\}.$$

It implies that $\zeta = -\nabla_y g(\overline{x}, \overline{y})\gamma$ by setting $u^* := 0$ in the relationship $u^* + \nabla_y g(\overline{x}, \overline{y})\gamma + \zeta = 0$ above. Further, the multiplier ζ can be eliminated from the whole process since the multiplier set $\Lambda^{cm}(\overline{x}, \overline{y}, \overline{u}, 0, 0, 0)$ corresponds to

$$\begin{aligned}\Lambda^{cm}(\overline{x}, \overline{y}, \overline{u}, 0) \quad &= \{(\beta, \gamma)\,|\, \nabla_y g_\nu(\overline{x}, \overline{y})\gamma = 0,\, \beta_\eta = 0, \\ &\quad (\nabla_y g_i(\overline{x}, \overline{y})\gamma > 0,\, \beta_i < 0) \vee \beta_i(\nabla_y g_i(\overline{x}, \overline{y})\gamma) = 0,\, \forall i \in \theta, \\ &\quad \nabla_{x,y}\mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma - \nabla g(\overline{x}, \overline{y})^\top \beta = 0\} = \Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0).\end{aligned}$$

The relationships $\Lambda_y^{cm}(\overline{x}, \overline{y}, \overline{u}, 0) = \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$ and $\Lambda_y^{cm}(\overline{x}, \overline{y}, \overline{u}) = \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})$ can be derived in a similar way. Thus the results of this theorem follow from those in Theorem 3.2 by observing that conditions (5.4) and (5.6) correspond to (A_1^1) and (A_2^1), respectively. $\square$

If the inner semicontinuity of S_o^h is replaced by its inner semicompactness at $\overline{x}$ and if condition (5.4) holds at $(\overline{x}, \overline{y}, \overline{u})$ for all $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$, then

$$\partial \varphi_o(\overline{x}) \subset \bigcup_{(\overline{y}, \overline{u}) \in S_o^h(\overline{x})} \bigcup_{(\beta, \gamma) \in \Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})} \{\nabla_x F(\overline{x}, \overline{y}) + \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma\}. \tag{5.7}$$

In this case it follows from Theorem 3.2(ii) that φ_o is Lipschitz continuous around $\overline{x}$ when condition (5.6) is satisfied at $(\overline{x}, \overline{y}, \overline{u})$ for all $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$. It makes sense to recall here that if S_o^h is inner semicompact at $\overline{x}$ and $S_o^h(\overline{x}) = \{(\overline{y}, \overline{u})\}$, then S_o^h is inner semicontinuous at $(\overline{x}, \overline{y}, \overline{u})$. Note that condition $S_o^h(\overline{x}) = \{(\overline{y}, \overline{u})\}$ in the bilevel programming context is *far removed* from the local uniqueness of lower-level solutions that is usually required in numerical algorithms; see, e.g., [8]. Moreover, the set $S_o^h(\overline{x})$ in (5.3) can be a singleton while the lower-level problem (1.2) may not have a unique solution.

Following Remark 3.3, we conclude that condition (5.4) can be replaced by the weaker *calmness* property of the set-valued mapping

$$\Phi(z, \vartheta) := \{(x, y, u)\,|\, \mathcal{L}(x, y, u) + z = 0,\, (-g_i(x, y), u_i) + \vartheta_i \in \Lambda_i,\, i := 1, \ldots, p\}, \tag{5.8}$$

where $\Lambda_i := \{(a, b) \in \mathbb{R}^2 \,|\, a \geq 0,\, b \geq 0,\, ab = 0\}$. The latter assumption is automatically satisfied when the mappings $g\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ and $(x, y) \mapsto \nabla_y f(x, y)$ are linear.

We can similarly consider the *C-type* multiplier sets $\Lambda^{ec}(\overline{x}, \overline{y}, \overline{u}, 0)$, $\Lambda_y^{ec}(\overline{x}, \overline{y}, \overline{u}, 0)$, and $\Lambda_y^{ec}(\overline{x}, \overline{y}, \overline{u})$, which are obtained by replacing condition (4.10) in $\Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$, $\Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$, and $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})$ by that of

$$\beta_i(\nabla_y g_i(\overline{x}, \overline{y})\gamma) \geq 0 \quad \text{for all} \quad i \in \theta.$$

Then an upper bound of the limiting subdifferential via C-type multipliers and the local Lipschitz continuity of the two-level value function φ_o under the C-type conditions can be derived as in Theorem 5.2 with $\Lambda^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$, $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0)$, and $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})$ replaced by $\Lambda^{ec}(\overline{x}, \overline{y}, \overline{u}, 0)$, $\Lambda_y^{ec}(\overline{x}, \overline{y}, \overline{u}, 0)$, and $\Lambda_y^{ec}(\overline{x}, \overline{y}, \overline{u})$, respectively. The case where S_o^h is inner semicompact would also follow analogously as described above for M-type multipliers.

To consider *S-type* upper bound for the subdifferential of φ_o, define the set $\Lambda_y^{es}(\overline{x}, \overline{y}, \overline{u})$ similarly to $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u})$ with replacing condition (4.10) by

$$\beta_i \geq 0 \quad \text{and} \quad \nabla_y g_i(\overline{x}, \overline{y})\gamma \geq 0 \quad \text{for all} \quad i \in \theta.$$

and arrive at the following sensitivity result.

THEOREM 5.3 (S-type sensitivity analysis for the two-level value function φ_o via the MPCC reformulation). *Assume that (A_1^4) is satisfied at all (x, y) with $y \in S(x)$, that the optimal solution map S_o^h from (5.3) is inner semicontinuous at $(\overline{x}, \overline{y}, \overline{u})$, and that both qualification conditions (5.4) and*

$$\left.\begin{array}{r} \nabla_{x,y}\mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma + \nabla g(\overline{x}, \overline{y})^\top \beta = 0 \\ \nabla_y g_\nu(\overline{x}, \overline{y}) \gamma = 0, \ \beta_\eta = 0 \end{array}\right\} \implies \beta_\theta = 0, \ \nabla_y g_\theta(\overline{x}, \overline{y}) \gamma = 0 \qquad (5.9)$$

hold at $(\overline{x}, \overline{y}, \overline{u})$. Then we have the subdifferential upper estimate

$$\partial \varphi_o(\overline{x}) \subset \bigcup_{(\beta,\gamma) \in \Lambda_y^{es}(\overline{x}, \overline{y}, \overline{u})} \left\{ \nabla_x F(\overline{x}, \overline{y}) + \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma \right\}.$$

Proof. Follows the lines in the proof of Theorem 3.8 by using now Theorem 5.2. □

Similarly to the above, we get the upper bound of $\partial \varphi_o(\overline{x})$ containing additionally the union over $(y, u) \in S_o^h(\overline{x})$ if the inner semicontinuity of S_o^h is replaced by its inner semicompactness.

We conclude this subsection by the following remark summarizing what can be done by using the OPEC approach to the sensitivity analysis of φ_o.

REMARK 5.4. *Since the functions $f(x, .)$ and $g(x, .)$ are assumed to be convex, the following equivalent generalized equation transformation of the solution map to the lower-level problem (1.2) is well known in convex optimization:*

$$y \in S(x) \iff 0 \in \nabla_y f(x, y) + N_{K(x)}(y). \qquad (5.10)$$

This allows us to reformulate the two-level value function φ_o as an OPEC value function

$$\varphi_o(x) = \min_y \{ F(x, y) \,|\, 0 \in \nabla_y f(x, y) + N_{K(x)}(y) \}.$$

Then applying Theorem 4.1 with $h(x, y) := \nabla_y f(x, y)$ leads us to an upper bound of the basic subdifferential of φ_o with M-type multipliers and a conclusion on the local Lipschitz continuity of φ_o different from that of Theorem 5.2. In particular, when S_o is assumed to be inner semicontinuous at $(\overline{x}, \overline{y})$, the upper bound of $\partial \varphi_o$ derived from Theorem 4.1(i) contains the union over $\Lambda(\overline{x}, \overline{y})$, which makes it much larger than the one obtained in Theorem 5.2. This is in fact understandable by taking into account that the appearance of the lower-level multiplier set $\Lambda(\overline{x}, \overline{y})$ in the upper estimate of $\partial \varphi_o$ in Theorem 4.1 is a posteriori while it is a priori in Theorem 5.2. However, in the case of the inner semicompactness of S_o, the upper bounds of Theorem 4.1 (ii) and inclusion (5.7) happen to be the same. Nevertheless, the assumptions made in both cases are similar but not identical.

5.2. LLVF approach. In this subsection we develop the *lower-level value function* (LLVF) approach to sensitivity analysis of the two-level value function φ_o from (1.4). Let us start by recalling that the *argminimum/solution map* of the lower-level problem (1.2) can be written as

$$S(x) := \{ y \,|\, f(x, y) - \varphi(x) \leq 0, \ g(x, y) \leq 0 \} \qquad (5.11)$$

with φ denoting the optimal value function associated to the lower-level problem (1.2), i.e.,

$$\varphi(x) := \min_y \{ f(x, y) \,|\, g(x, y) \leq 0 \}. \qquad (5.12)$$

Hence we have the LLVF reformulation of the two-level value function φ_o written as

$$\varphi_o(x) := \min_y \{F(x,y) \mid g(x,y) \le 0, \ f(x,y) - \varphi(x) \le 0\}.$$

Since our basic subdifferential $\partial\varphi$ does not satisfy the *plus/minus* symmetry, an appropriate estimate of $\partial(-\varphi)$ is needed to proceed with this approach. It can be done by invoking the well-known *convex hull property*

$$\operatorname{co}\partial(-\varphi)(\bar{x}) = -\operatorname{co}\partial\varphi(\bar{x}) \tag{5.13}$$

for local Lipschitzian functions, which follows from [5] due to $\overline{\partial}\varphi(\bar{x}) = \operatorname{co}\partial\varphi(\bar{x})$. The next theorem collects the results in this direction needed in what follows.

THEOREM 5.5 (sensitivity analysis of the negative value function in the lower-level problem). *The following assertions hold for the negation of the value function φ in (5.12):*

(i) *If the solution map S in (5.11) is inner semicompact at $\bar{x}$ for all $(\bar{x},y) \in \operatorname{gph}S$ satisfying (A_1^4), then φ is Lipschitz continuous around $\bar{x}$ and we have the inclusion*

$$\partial(-\varphi)(\bar{x}) \subset \quad \Big\{ \textstyle\sum_{s=1}^{n+1} \eta_s \big(\nabla_x f(\bar{x},y_s) + \nabla_x g(\bar{x},y_s)^\top \beta_s\big) \Big| (y_s)_{s=1}^{n+1} \in \prod_{s=1}^{n+1} S(\bar{x})$$
$$(\beta_s)_{s=1}^{n+1} \in \textstyle\prod_{s=1}^{n+1} \Lambda(\bar{x},y_s)$$
$$(\eta_s)_{s=1}^{n+1} \in \mathbb{R}_-^{n+1}, \ \textstyle\sum_{s=1}^{n+1} \eta_s = -1 \Big\}.$$

(ii) *Assume that $(\bar{x},\bar{y}) \in \operatorname{gph}S$ with $\bar{x} \in \operatorname{dom}\varphi$ satisfies (A_1^4) and that* EITHER *S is inner semicontinuous at this point* OR *f and g are fully convex. Then φ is Lipschitz continuous around $\bar{x}$ and we have the inclusion*

$$\partial(-\varphi)(\bar{x}) \subset \bigcup_{\beta \in \Lambda(\bar{x},\bar{y})} \big\{ -\nabla_x f(\bar{x},\bar{y}) - \nabla_x g(\bar{x},\bar{y})^\top \beta \big\}.$$

Proof. The local Lipschitz continuity of φ is justified in [39] under the fulfillment of (A_1^4) in both inner semicontinuous and inner semicompactness cases. If the functions f and g are fully convex, then the value function φ is convex as well; in this case the Lipschitz continuity follows from [5]. To prove the subdifferential inclusion in (i), recall that

$$\partial\varphi(\bar{x}) \subset \bigcup_{y \in S(\bar{x})} \bigcup_{\beta \in \Lambda(\bar{x},y)} \big\{ \nabla_x f(\bar{x},y) + \nabla_x g(\bar{x},y)^\top \beta \big\},$$

by [42] under the assumptions made in (i). The claimed estimate of $\partial(-\varphi)$ follows from here by combining (5.13) and the classical Carathéodory's theorem; cf. [11].

When S is inner semicontinuous at $(\bar{x},\bar{y})$, we have by [40] that

$$\overline{\partial}\varphi(\bar{x}) \subset \bigcup_{\beta \in \Lambda(\bar{x},\bar{y})} \big\{ \nabla_x f(\bar{x},\bar{y}) + \nabla_x g(\bar{x},\bar{y})^\top \beta \big\}, \tag{5.14}$$

which implies the subdifferential inclusion in (ii) by (5.13). If both f and g are fully convex, inclusion (5.14) holds without the inner semicontinuity assumption; see [11, 16]. $\square$

Note that in the fully convex (even nonsmooth) case, assumption (A_1^4) in Theorem 5.5 can be replaced by a much weaker qualification condition [16] requiring that the set

$$\text{epi } f^* + \text{cone} \left(\bigcup_{i=1}^{p} \text{epi } g_i^* \right) \text{ is closed on } \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R},$$

where epiϑ^* denotes the conjugate function for an extended-real-valued convex function ϑ.

5.2.1. Employing calmness. The importance of various calmness properties has been well recognized in bilevel programming. In this subsection we discuss their roles in the LLVF approach to sensitivity analysis of the two-level optimal value function (1.4). Calculating the *coderivative* of the optimal solution map S in (5.11) is highly significant in our approach. This means computing the limiting normal cone to the graph of S:

$$\text{gph } S = \{(x, y) \in \Omega | f(x, y) - \varphi(x) \leq 0\} \text{ with } \Omega := \{(x, y) \mid g(x, y) \leq 0\} \quad (5.15)$$

in terms of the initial data. To proceed in this way by using the conventional results of the generalized differential calculus [38] requires the fulfillment of the *basic qualification condition*, which reads in this case as

$$\partial(f - \varphi)(\overline{x}, \overline{y}) \cap (-N_\Omega(\overline{x}, \overline{y})) = \emptyset. \quad (5.16)$$

However, it is shown in [12] that condition (5.16) fails in common situations; in particular, when φ is locally Lipschitzian around the point in question. The following weaker assumption helps circumventing this difficulty:
(A_1^5) The mapping $\Phi(\vartheta) := \{(x, y) \in \Omega | f(x, y) - \varphi(x) \leq \vartheta\}$ is *calm* at $(0, \overline{x}, \overline{y})$.
By applying the concept of stability regions known in linear programming (see, e.g., [8]), to the optimal value function φ it is possible to show, by means of Robinson's theorem [48] on the upper-Lipschitz continuity of a polyhedral set-valued mapping, that (A_1^5) is automatically satisfied if f and g are linear. Furthermore, condition (A_1^5) is satisfied at $(\overline{x}, \overline{y})$ for locally Lipschitzian functions φ if we pass to the boundary of the normal cone in (5.16), i.e., if the following qualification condition holds:

$$\partial(f - \varphi)(\overline{x}, \overline{y}) \cap (-\text{bd } N_\Omega(\overline{x}, \overline{y})) = \emptyset \quad (5.17)$$

with Ω being *semismooth*, in particular, convex; cf. [12, 26]. Condition (5.17) seems to be especially effective for the so-called *simple convex bilevel optimization problems*; see [9, 12] for more details. It is worth mentioning that for the latter class of problems condition (5.17) can be further weakened [12] by passing to the boundary of the subdifferential of f.

Another sufficient condition for the validity of (A_1^5) is provided by the notion of uniform weak sharp minima. The parametric optimization problem (1.2) is said to have a *uniform weak sharp minimum* around $(\overline{x}, \overline{y})$ if there exist positive numbers λ and δ such that

$$f(x, y) - \varphi(x) \geq \lambda d(y, S(x)) \text{ for all } (x, y) \in B((\overline{x}, \overline{y}), \delta) \cap \Omega. \quad (5.18)$$

The concept of uniform weak sharp minimum, which emerged from the notions of sharp minimum introduced by Polyak [47] and weak sharp minimum introduced by

Ferris [18], was developed by Ye and Zhu [60] while the above localized version (5.18) has been recently considered by Henrion and Surowiec [27] and by Mordukhovich, Nam and Phan [40]. It follows from [27, Proposition 3.8] that (A_1^5) holds at $(\overline{x}, \overline{y})$ if problem (1.2) has a uniform weak sharp minimum around $(\overline{x}, \overline{y})$. Furthermore, it is shown in [40] that if f and g are linear in y and (x, y), respectively, then the lower-level problem has a local uniform weak sharp minimum. Note that notion (5.18) is closely related to the *partial calmness* property introduced in [60]. A number of other efficient conditions insuring a uniform weak sharp minimum in bilevel programming can be found in [12, 16, 11, 40, 60, 55]. It is interesting to observe that the qualification condition (A_1^5) is a sufficient for the partial calmness of the bilevel program in question if we drop the upper-level constraint or include it in the constraint set Ω.

5.2.2. Sensitivity analysis for optimal solution maps. In this subsection we derive an upper estimate for the coderivative of the solution map S given in (5.11) and then establish its Lipschitz-like property. Our additional qualification condition is formulated as follows:

(A_2^5) $\left[(\lambda, \beta) \in \Lambda_y^o(\overline{x}, \overline{y}, 0), \; x^* \in \partial(-\varphi)(\overline{x}) \right] \Longrightarrow \lambda x^* = -\lambda \nabla_x f(\overline{x}, \overline{y}) - \nabla_x g(\overline{x}, \overline{y})^\top \beta$,

where $\Lambda_y^o(\overline{x}, \overline{y}, y^*)$ for $y^* \in \mathbb{R}^m$ denotes a *particular set of multipliers* that plays an important role in the rest of the section:

$$\Lambda_y^o(\overline{x}, \overline{y}, y^*) := \begin{array}{l} \{(\lambda, \beta) | \, \lambda \geq 0, \, \beta \geq 0, \, \beta^\top g(\overline{x}, \overline{y}) = 0, \\ \qquad y^* + \lambda \nabla_y f(\overline{x}, \overline{y}) + \nabla_y g(\overline{x}, \overline{y})^\top \beta = 0 \}. \end{array} \qquad (5.19)$$

The next proposition describes a setting where assumption (A_2^5) is automatically satisfied.

PROPOSITION 5.6 (validity of assumption (A_2^5)). *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*

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

and the corresponding solution map $S(x) = \min_y \{ f(x, y) | \, g(y) \leq 0 \}$. *Taking* $(\overline{x}, \overline{y}) \in$ gph S *with* $\varphi(\bar{x}) < \infty$, *we have* $(A_1^4) \Longrightarrow (A_2^5)$ *at* $(\overline{x}, \overline{y})$.

Proof. Under the setting of this proposition, it follows from the convex case of Theorem 5.5(ii) that the function $-\varphi$ is strictly differentiable at $\overline{x}$ and $\partial(-\varphi)(\overline{x}) = \{-\nabla_x f(\overline{x}, \overline{y})\}$, which therefore justifies our conclusion. $\square$

The main result of this subsection is as follows.

THEOREM 5.7 (coderivative estimate and Lipschitz-like property of lower-level solution maps). *Let the solution map* (5.11) *be inner semicontinuous at* $(\overline{x}, \overline{y}) \in$ gph S, *and let the qualification conditions* (A_1^4) *and* (A_1^5) *be satisfied at this point. Then we have for all* $y^* \in \mathbb{R}^m$

$$D^* S(\overline{x}, \overline{y})(y^*) \subset \bigcup_{(\lambda, \beta) \in \Lambda_y^o(\overline{x}, \overline{y}, y^*)} \left\{ \lambda(\nabla_x f(\overline{x}, \overline{y}) + \partial(-\varphi)(\overline{x})) + \nabla_x g(\overline{x}, \overline{y})^\top \beta \right\} \quad (5.20)$$

If in addition (A_2^5) *holds at* $(\overline{x}, \overline{y})$, *then* S *is Lipschitz-like around this point.*

Proof. It follows from Theorem 5.5(ii) that the lower-level value function φ is Lipschitz continuous around $\overline{x}$ under (A_1^4) and the inner semicontinuity assumptions. If we add the calmness property (A_1^5), then

$$N_{\mathrm{gph}\,S}(\overline{x}, \overline{y}) \subset \bigcup_{\lambda \geq 0} \{ \lambda(\nabla f(\overline{x}, \overline{y}) + \partial(-\varphi)(\overline{x}) \times \{0\}) + N_\Omega(\overline{x}, \overline{y}) \}$$

by [25, Theorem 4.1] while taking into account that the constraint $f(x,y) - \varphi(x) \le 0$ is active at $(\overline{x}, \overline{y})$. The coderivative estimate (5.20) of the theorem follows now from definition (2.6) and the well-known expression of the normal cone

$$N_\Omega(\overline{x}, \overline{y}) = \{\nabla g(\overline{x}, \overline{y})^\top \beta \mid \beta \ge 0,\ \beta^\top g(\overline{x}, \overline{y}) = 0\},$$

which holds under the validity of (A_1^4) at $(\overline{x}, \overline{y})$. Further, by (5.20) and the coderivative criterion (2.9) for the Lipschitz-like property we get that the latter holds provided that

$$\left.\begin{array}{r} x^* \in \lambda(\nabla_x f(\overline{x}, \overline{y}) + \partial(-\varphi)(\overline{x})) + \nabla_x g(\overline{x}, \overline{y})^\top \beta \\ (\lambda, \beta) \in \Lambda_y^o(\overline{x}, \overline{y}, 0) \end{array}\right\} \implies x^* = 0,$$

which is in fact equivalent to the assumed qualification condition (A_2^5). $\square$

REMARK 5.8. *It follows from the alternative statement in Theorem 5.5(ii) that the inner semicontinuity of S can be dropped in the assumptions of Theorem 5.7 if the functions f and g are fully convex. As usual, the inner semicontinuity can be replaced by inner semicompactness with a larger inclusion in* (5.20).

5.2.3. Sensitivity analysis for two-level value functions via the LLVF approach. To conduct a local sensitivity analysis of the two-level value function φ_o defined in (1.4), we associate with it the *optimal solution map S_o* of the *upper-level* problem defined in (1.6) with S given as in (5.11). Having in mind the definition of the multiplier set $\Lambda_y^o(\overline{x}, \overline{y}, y^*)$ in Subsection 5.2.2, we put $\Lambda_y^o(\overline{x}, \overline{y}) := \Lambda_y^o(\overline{x}, \overline{y}, \nabla_y F(\overline{x}, \overline{y}))$. Then sensitivity results for φ_o are given next.

THEOREM 5.9 (sensitivity analysis for the two-level value function φ_o). *In the settings of* (1.4) *and* (1.6) *the following assertions hold:*

(i) *Assume that S_o is inner semicontinuous at $(\overline{x}, \overline{y})$ and that conditions (A_1^4) and (A_1^5) hold at this point. Then we have*

$$\partial\varphi_o(\overline{x}) \subset \bigcup_{(\lambda,\beta)\in\Lambda_y^o(\overline{x},\overline{y})} \bigcup_{\gamma\in\Lambda(\overline{x},\overline{y})} \{\nabla_x F(\overline{x}, \overline{y}) + \nabla_x g(\overline{x}, \overline{y})^\top(\beta - \lambda\gamma)\}.$$

If in addition (A_2^5) is satisfied at $(\overline{x}, \overline{y})$, then φ_o is Lipschitz continuous around $\overline{x}$.

(ii) *Assume that S_o is inner semicompact at $\overline{x}$, that (A_1^4) holds at $(\overline{x}, y)$ for all $y \in S(\overline{x})$, while (A_1^5) holds at $(\overline{x}, y)$ for all $y \in S_o(\overline{x})$. Then we have*

$$\partial\varphi_o(\overline{x}) \subset \bigcup_{y\in S_o(\overline{x})} \bigcup_{(\lambda,\beta)\in\Lambda_y^o(\overline{x},y)} \{\nabla_x F(\overline{x}, y) + \lambda\nabla_x f(\overline{x}, y) + \lambda\partial(-\varphi)(\overline{x}) + \nabla_x g(\overline{x}, y)^\top\beta\},$$

where the subdifferential $\partial(-\varphi)(\overline{x})$ is estimated in Theorem 5.5(i). If in addition (A_2^5) is satisfied at $(\overline{x}, y)$ for all $y \in S_o(\overline{x})$, then φ_o is Lipschitz continuous around $\overline{x}$.

Proof. To justify (i), observe by Theorem 2.1(i) that

$$\partial\varphi_o(\overline{x}) \subset \nabla_x F(\overline{x}, \overline{y}) + D^* S(\overline{x}, \overline{y})(\nabla_y F(\overline{x}, \overline{y})).$$

under the inner semicontinuity assumption on S_o. Since we obviously have $S_o(x) \subset S(x)$ for all $x \in X$, the lower-level optimal solution map S in (1.5) is also inner semicontinuous at $(\overline{x}, \overline{y}) \in \text{gph}\, S_o$. Thus the upper estimate of $\partial\varphi_o(\overline{x})$ in this theorem follows from those for the coderivative of S in Theorem 5.7 and for the subdifferential of the lower-level value function φ in Theorem 5.5(ii). To justify the local Lipschitz continuity of φ_o in (i) under (A_2^5), recall that the latter condition implies the Lipschitz-like property of S around $(\overline{x}, \overline{y})$ by Theorem 5.7. Thus we have the claimed result from Theorem 2.1(i).

Assertion (ii) is proved similarly following the discussion in Remark 5.8. □

REMARK 5.10. *Observe that for the subdifferential estimate of φ_o in Theorem 5.9(i), the upper bound of the basic subdifferential does not contain the partial derivative of the lower-level cost function f with respect to the upper-level variable x. This will induce in the context of necessary optimality conditions for the original optimistic formulation (P_o) in the next section a remarkable phenomenon first discovered by Dempe, Dutta and Mordukhovich [11] in the framework concerning the auxiliary problem (P). Note that such a phenomenon is no longer true if the inner semicontinuity assumption on S_o is replaced by the inner semicompactness one in assertion (ii) of Theorem 5.9. Finally, we mention that the inner semicompactness of S_o in Theorem 5.9(ii) can be replaced by the easier while more restrictive uniform boundedness assumption imposed on S_o or even on S.*

6. Applications to necessary optimality in original optimistic model. The concluding section of the paper is devoted to applications of the above sensitivity results to deriving new necessary optimality conditions for the original optimistic formulation (P_o) in bilevel programming. In fact we establish certain *stationarity conditions* of various types among which are of those types known for more conventional auxiliary optimistic formulation (P) together with stationarity conditions of the novel types for (P_o).

Most of the notation and assumptions in this section were used above. For the reader's convenience, remind that S and S_o refer to the solution maps of the lower-level (1.5) and upper-level (1.6) problems, respectively. and that the lower-level Lagrange function $\mathcal{L}$ and Lagrange multipliers set $\Lambda(\overline{x}, \overline{y})$ are given by

$$\mathcal{L}(x, y, u) = \nabla_y f(x, y) + \nabla_y g(x, y)^\top u \ \text{ and }$$
$$\Lambda(\overline{x}, \overline{y}) = \{u|\ \mathcal{L}(\overline{x}, \overline{y}, u) = 0,\ u \geq 0, g(\overline{x}, \overline{y}) \leq 0,\ u^\top g(\overline{x}, \overline{y}) = 0\}.$$

Also the index sets η, θ, and ν of the major interest here are defined in (4.7).

We start with the notion of M-stationary and weak M-stationary points specified for the original optimistic bilevel program (P_o).

DEFINITION 6.1 (M-stationarity). *A point $\overline{x}$ is* M-STATIONARY (*resp.*WEAK M-STATIONARY) *for problem (P_o) if for every $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$ (resp. there exists $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$) we can find a triple $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that*

$$\nabla_x F(\overline{x}, \overline{y}) + \nabla G(\overline{x})^\top \alpha + \nabla_x g(\overline{x}, \overline{y})^\top \beta + \nabla_x \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0, \tag{6.1}$$
$$\nabla_y F(\overline{x}, \overline{y}) + \nabla_y g(\overline{x}, \overline{y})^\top \beta + \nabla_y \mathcal{L}(\overline{x}, \overline{y}, \overline{u})^\top \gamma = 0, \tag{6.2}$$
$$\alpha \geq 0,\ \alpha^\top G(\overline{x}) = 0, \tag{6.3}$$
$$\nabla_y g_\nu(\overline{x}, \overline{y})\gamma = 0,\ \beta_\eta = 0, \tag{6.4}$$
$$\forall i \in \theta,\ (\beta_i > 0 \land \nabla_y g_i(\overline{x}, \overline{y})\gamma > 0) \lor \beta_i(\nabla_y g_i(\overline{x}, \overline{y})\gamma) = 0. \tag{6.5}$$

Relationships (6.1)–(6.5) *are called the* M-STATIONARITY CONDITIONS.

Similarly we define the *C-stationarity* (resp. *S-stationarity*) by replacing condition (6.5) with

$$\forall i \in \theta,\ \beta_i(\nabla_y g_i(\overline{x}, \overline{y})\gamma) \geq 0 \ (\text{resp. } \ \forall i \in \theta,\ \beta_i \geq 0,\ \nabla_y g_i(\overline{x}, \overline{y})\gamma \geq 0).$$

We obviously have the implications: S-stationarity$\Longrightarrow$M-stationarity$\Longrightarrow$C-stationarity.

The following stationarity conditions of the new "KM" and "KN" types for the original optimistic bilevel program (P_o) reflect the difference between the **KKT**-type

optimality conditions obtained via the inner semicontinuity and inner semicompactness, respectively, of the optimal solution map S_o for the upper-level problem.

DEFINITION 6.2 (KN-stationarity). *A point $\overline{x}$ is* KN-STATIONARY (*resp.* WEAK KN-STATIONARY) *for problem (P_o) if for all $\overline{y} \in S_o(\overline{x})$ (resp. there exists $\overline{y} \in S_o(\overline{x})$) we can find a triple $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+2p}$ and a number $\lambda \in \mathbb{R}_+$ such that*

$$\nabla_x F(\overline{x}, \overline{y}) + \nabla G(\overline{x})^\top \alpha + \nabla_x g(\overline{x}, \overline{y})^\top (\beta - \lambda \gamma) = 0, \tag{6.6}$$

$$\nabla_y F(\overline{x}, \overline{y}) + \lambda \nabla_y f(\overline{x}, \overline{y}) + \nabla_y g(\overline{x}, \overline{y})^\top \beta = 0, \tag{6.7}$$

$$\nabla_y f(\overline{x}, \overline{y}) + \nabla_y g(\overline{x}, \overline{y})^\top \gamma = 0, \tag{6.8}$$

$$\alpha \geq 0, \ \alpha^\top G(\overline{x}) = 0, \tag{6.9}$$

$$\beta \geq 0, \ \beta^\top g(\overline{x}, \overline{y}) = 0, \tag{6.10}$$

$$\gamma \geq 0, \ \gamma^\top g(\overline{x}, \overline{y}) = 0. \tag{6.11}$$

Relationships (6.6)–(6.11) *are called the* KN-STATIONARITY CONDITIONS.

DEFINITION 6.3 (KM-stationarity). *A point $\overline{x}$ is* KM-STATIONARY *for problem (P_o) if there exist elements $\overline{y} \in S_o(\overline{x})$, $(y_s)_{s=1}^{n+1} \in \prod_{s=1}^{n+1} S(\overline{x})$, $(\alpha, \beta) \in \mathbb{R}^{k+p}$, $\lambda \in \mathbb{R}_+$, $(\gamma_s)_{s=1}^{n+1} \in (\mathbb{R}^k)^{n+1}$, and $(\eta_s)_{s=1}^{n+1} \in \mathbb{R}_-^{n+1}$ with $\sum_{s=1}^{n+1} \eta_s = -1$ such that we have relationships* (6.7), (6.9)–(6.10) *to be satisfied together with the following conditions:*

$$\begin{aligned} \nabla_x F(\overline{x}, \overline{y}) + \nabla G(\overline{x})^\top \alpha + \nabla_x g(\overline{x}, \overline{y})^\top \beta + \lambda \nabla_x f(\overline{x}, \overline{y}) \\ + \lambda \sum_{s=1}^{n+1} \eta_s \big(\nabla_x f(\overline{x}, y_s) + \nabla_x g(\overline{x}, y_s)^\top \gamma_s \big) = 0, \end{aligned} \tag{6.12}$$

$$\nabla_y f(\overline{x}, y_s) + \nabla_y g(\overline{x}, y_s)^\top \gamma_s = 0, \tag{6.13}$$

$$\gamma_s \geq 0, \ \gamma_s^\top g(\overline{x}, y_s) = 0. \tag{6.14}$$

All the relationships (6.7), (6.9)–(6.10), *and* (6.12)–(6.14) *considered together are called the* KM-STATIONARITY CONDITIONS.

To the best of our knowledge, necessary optimality condition of the KN-type were first obtained by Dempe, Dutta and Mordukhovich [11] for the standard/auxiliary version (P) in optimistic bilevel programming while those of the KM-type originated by Ye and Zhu [60] for (P) under additional assumptions involving *partial calmness*, which is not impose here. It is easy to see that the KM-stationarity agrees with the weak KN-stationarity provided that $S(\overline{x}) = \{\overline{y}\}$ and $\Lambda(\overline{x}, \overline{y}) = \{\gamma\}$. Moreover, if $\nabla_{x,y} \mathcal{L}(\overline{x}, \overline{y}, \overline{u}) = 0$, which is the case when f and g are linear in (x, y), then the S-stationarity conditions for a fixed $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ imply the KN-stationarity ones; cf. [12, 13]. In general there is no relationship between the M-, C- and S-stationarity conditions on the one hand and the KN- and KM-ones on the other.

We are now ready to establish one of the major results of the paper proving M-type necessary optimality conditions for the original optimistic bilevel program (P_o). Thus the result is derived from the sensitivity analysis of Theorem 5.2 and basic facts in variational analysis. To proceed, recall that a point $\overline{x} \in X := \{x \in \mathbb{R}^n \,|\, G(x) \leq 0\}$ is *upper-level regular* if there exists no nonzero vector $\alpha \geq 0$ such that $\alpha^\top G(\overline{x}) = 0$ and $\nabla G(\overline{x})^\top \alpha = 0$. This is nothing but the dual form of the classical MFCQ for the inequality system $G(x) \leq 0$. Finally, the *convexity* of the lower-level problem required in this theorem is understood in the sense that the functions $f(x, .)$ and $g(x, .)$ in the latter problem (1.2) are convex for all $x \in X$.

THEOREM 6.4 (M-type necessary optimality conditions for (P_o)). *Let $\overline{x}$ be an upper-level regular local optimal solution to (P_o), where the lower-level problem is convex. Assume that the qualification condition (A_1^4) holds at all (x, y) as $y \in S(x)$,*

that the solution map S_o^h from (5.3) is inner semicontinuous at $(\overline{x}, \overline{y}, \overline{u})$ for all $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$, and that relationships (5.4) and (5.6) are satisfied at $(\overline{x}, \overline{y}, \overline{u})$ for all $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$. Then $\overline{x}$ is M-stationary for problem (P_o).

Proof. Under the assumptions made, it follows from Theorem 5.2 that the two-level value function φ_o is Lipschitz continuous around $\overline{x}$. Thus $\partial \varphi_o(\overline{x}) \neq \emptyset$ while $\partial^\infty \varphi_o(\overline{x}) = \{0\}$, and the qualification condition (2.12) in Theorem 2.2 holds at $\overline{x}$. Employing now the optimality condition (2.11) of the latter theorem with the well-known formula

$$N_X(\overline{x}) = \{\nabla G(\overline{x})^\top \alpha \,|\, \alpha \geq 0, \, \alpha^\top G(\overline{x}) = 0\}$$

valid under the assumed upper-level regularity of $\bar{x}$ and then taking into account that the set on the right-hand side of inclusion (5.5) is nonempty, we arrive at the M-stationarity conditions of the theorem. $\square$

It is worth mentioning that the upper-level regularity in the assumptions of Theorem 6.4 can be replaced by the weaker *calmness* property of the mapping $v \rightrightarrows \{x \in \mathbb{R}^n \,|\, G(x) + v \leq 0\}$, which is automatically satisfied in if G is a linear function. Furthermore, as mentioned previously in Subsection 5.1, the qualification condition (5.4) can also be replaced by the weaker calmness property of the mapping Φ in (5.8), which holds if both functions g and $(x, y) \mapsto \nabla_y f(x, y)$ are linear. Next we provide a simple example illustrating Theorem 6.4.

EXAMPLE 6.5. *Consider the original optimistic bilevel program as in [13, Example 4.1]:*

$$\min_{x \in \mathbb{R}_+} \left\{ \min_{y \in \mathbb{R}} \{x^2 + y^2 \,|\, y \in S(x) := \arg\min\{xy + y \,|\, y \geq 0\}\} \right\}.$$

The KKT/complementarity reformulation of the corresponding two-level value function is

$$
\begin{aligned}
\varphi_o(x) \quad &:= \min_{y,u}\{x^2 + y^2 \,|\, x - u + 1 = 0, \\
&\qquad\qquad u \geq 0, \, y \geq 0, \, uy = 0\} \\
&= \begin{cases} x^2 & \text{if } x \geq -1, \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}
$$

It is obvious that $\overline{x} = 0$ is the (unique) optimistic optimal solution of this program and that φ_o is continuously differentiable near $\bar{x}$. On the other hand, we have $S_o^h(x) = \{(0, x+1)\}$ if $x \geq -1$ and $S_o^h(x) = \emptyset$ otherwise, and hence S_o^h reduces to a single-valued and continuous on its graph. Furthermore, $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0) = \{0\} \times \mathbb{R}$ if $(\overline{x}, \overline{y}, \overline{u}) = (-1, 0, 0)$ and $\Lambda_y^{em}(\overline{x}, \overline{y}, \overline{u}, 0) = \{(0, 0)\}$ for all the other points of $\mathrm{gph}\, S_o^h$. From the observations made in Sections 3-4, this implies that the qualification conditions (5.4) and (5.6) are satisfied at all points of the graph of S_o^h except $(-1, 0, 0)$, which is not optimal.

We can see from the proof of Theorem 6.4 that the local Lipschitz continuity of φ_o was used twice: to ensure the nonemptiness of $\partial\varphi_o(\overline{x})$ and the application of the optimality condition (2.11) of Theorem 2.2. Observe to this end that the Lipschitz property of φ_o is not needed for bilevel programs without upper-level constraints (i.e., if $X := \mathbb{R}^n$); in this case the qualification condition (2.12) holds automatically. The latter also allows us to drop assumption (5.6) in Theorem 6.4. However, we still have to make sure that $\partial\varphi_o(\bar{x}) \neq \emptyset$, which happens in many non-Lipschitzian situations; see, e.g., [38, 42, 50].

We can similarly derive *weak M-stationarity* conditions for the original optimistic bilevel formulation (P_o) under consideration.

THEOREM 6.6 (weak M-type necessary optimality conditions for (P_o)). *Let $\overline{x}$ be an upper-level regular local optimal solution to (P_o), where the lower-level problem is convex. Assume that (A_1^4) is satisfied at all (x, y) with $y \in S(x)$, that the solution map S_o^h in (5.3) is inner semicompact at $\overline{x}$, and that the qualification conditions (5.4) and (5.6) are satisfied at $(\overline{x}, \overline{y}, \overline{u})$ for all $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$. Then $\overline{x}$ is weak M-stationary for problem (P_o).*

Proof. Follows the lines in the proof of Theorem 6.4 by taking into account the discussion after the proof of Theorem 5.2. □

Another possibility to derive the above weak M-type necessary optimality conditions for (P_o) is by using the upper estimate of $\partial \varphi_o(\overline{x})$ obtained via the generalized equation transformation in Theorem 4.1. Note also that if the inner semicontinuity and qualification conditions (5.4) and (5.6) are satisfied only at one point $(\overline{x}, \overline{y}, \overline{u})$ in Theorem 6.4, we still can derive the weak M-type necessary optimality conditions for (P_o) at the difference that the reference couple $(\overline{y}, \overline{u}) \in S_o^h(\overline{x})$ is known a priori.

Proceeding similarly to the proof of Theorem 6.4 (resp. Theorem 6.6), the C-stationarity for a local optimal solution to problem (P_o) can be derived by a combination of Theorem 2.2 and the C-type counterpart of Theorem 5.2 (resp. C-type counterpart of inclusion (5.7)). The S-stationarity can be derived in this way by combining Theorem 2.2 and Theorem 5.3. Furthermore, based on Theorem 5.9(i,ii), we can respectively derive the following KN- and KM-stationarity conditions for the original optimistic bilevel program (P_o). We leave the proofs of these theorems to the reader.

THEOREM 6.7 (KN-type necessary optimality conditions for (P_o)). *Let $\overline{x}$ be an upper-level regular local optimal solution to the bilevel program (P_o). Assume that S_o (1.6) is inner semicontinuous at $(\overline{x}, \overline{y})$ (resp. for all $\overline{y} \in S_o(\overline{x})$) and that the conditions (A_1^1), (A_1^5) and (A_2^5) are satisfied at $(\overline{x}, \overline{y})$ (resp. for all $\overline{y} \in S_o(\overline{x})$). Then the point $\overline{x}$ is weak KN-stationary (resp. KN-stationary) for problem (P_o).*

THEOREM 6.8 (KM-type necessary optimality conditions for (P_o)). *Let $\overline{x}$ be an upper-level regular local optimal solution of (P_o), and let S_o (1.6) be inner semicompact at $\overline{x}$. Assume furthermore that (A_1^1) holds at all $(\overline{x}, \overline{y})$ with $\overline{y} \in S(\overline{x})$ and that (A_1^5), (A_2^5) are satisfied at all $(\overline{x}, \overline{y})$ with $\overline{y} \in S_o(\overline{x})$. Then the point $\overline{x}$ is KM-stationary for problem (P_o).*

Finally in this section, we establish the link between the above stationarity conditions for the original optimistic formulation (P_o) and those known for the *conventional/auxiliary optimistic* problem (P). Recall that a point $(\overline{x}, \overline{y})$ is M-, C- and S-stationary (resp. weak M-, C- and S-stationary) for problem (P) if for all $\overline{u} \in \Lambda(\overline{x}, \overline{y})$ (resp. there exists $\overline{u} \in \Lambda(\overline{x}, \overline{y})$) we can find a triple $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+p+m}$ such that the M-, C- and S-stationarity conditions are satisfied, respectively. Similarly we say that $(\overline{x}, \overline{y})$ is KM-stationary (resp. KN-stationary) for problem (P) if there exist $y \in S_o(\overline{x})$ (resp. for all $y \in S_o(\overline{x})$) with $(y_s)_{s=1}^{n+1} \in \prod_{s=1}^{n+1} S(\overline{x})$ and multipliers $(\alpha, \beta, \lambda) \in \mathbb{R}^k \times \mathbb{R}^p \times \mathbb{R}_+$, $(\gamma_s)_{s=1}^{n+1} \in (\mathbb{R}^k)^{n+1}$, and $(\eta_s)_{s=1}^{n+1} \in \mathbb{R}_-^{n+1}$ with $\sum_{s=1}^{n+1} \eta_s = -1$ (resp. $(\alpha, \beta, \gamma) \in \mathbb{R}^{k+2p}$ and $\lambda \in \mathbb{R}_+$) such that the KM- (resp. KN)-stationarity conditions are satisfied. The weak KN-stationarity for (P) can be defined analogously. As already mentioned in the introduction, problem (P) has been intensively studied in the literature. In particular, the KM- and KN-stationarity conditions have been derived under various assumptions in [11, 12, 14, 16, 40, 60]. For the other conditions see, e.g., [13, 20] and their references.

In the next theorem, which we consider as one of the major achievements of this paper, the term "WP$_o$-stationarity" unifies the notions of KM-stationarity and weak

M-, C-, S-, and KN-stationarity for problem (P_o); the term "WP-stationarity" stands for the corresponding notions for problem (P).

THEOREM 6.9 (comparison between necessary optimality conditions for problem (P) and (P_o)). *If $\overline{x}$ is a WP_o-stationary point, then there exists $\overline{y} \in S_o(\overline{x})$ such that $(\overline{x}, \overline{y})$ is a WP-stationary one. Conversely, if $(\overline{x}, \overline{y})$ is WP-stationary for some $\overline{y} \in S_o(\overline{x})$, then the point $\overline{x}$ is WP_o-stationary in the corresponding sense.*

Proof. It follows from the direct comparison of the new necessary optimality/stationarity conditions obtained above for the original bilevel formulation (P_o) and the conventional one (P) in optimistic bilevel programming. $\square$

It follows from the above theorem that the weak stationarity conditions for (P_o) are in fact equivalent to those for (P) under the assumptions made, while we cannot make such a conclusion for the corresponding "strong" notions.

Regarding the inner semicompactness setting for the optimal solution map in the optimality conditions, recall the following result established in [11]: if $\overline{x}$ is a local optimal solution to (P_o) with $X := \mathbb{R}^n$ and if S_o is uniformly bounded around $\overline{x}$, then the pair $(\overline{x}, \overline{y})$ with $\overline{y} \in S(\overline{x})$ and $\varphi_o(\overline{x}) = F(\overline{x}, \overline{y})$ is a local optimal solution to problem (P). This actually means that the inner semicompactness assumption is needed anyway whenever one intends to derive optimality conditions for the original optimistic bilevel program (P_o) by first deriving those of the auxiliary problem (P).

Finally, it is worth mentioning that the results on the sensitivity analysis of two-level value functions obtained in this paper can readily be applied for the sensitivity analysis of the auxiliary problem (P) and also to derive necessary optimality conditions for the pessimistic bilevel program. The latter issue will be developed in more detail in a future research.

REFERENCES

[1] W. Achtziger and C. Kanzow. Mathematical programs with vanishing constraints: Optimality conditions and constraint qualifications. *Math. Program.*, 114:69-99, 2008.

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

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

[4] M.Červinka, C. Matonoha and J.V. Outrata. On the computation of relaxed pessimistic solutions to MPEC, submitted.

[5] F.H. Clarke. *Optimization and nonsmooth analysis.* Wiley, New York, 1983.

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

[7] S. Dassanayaka. *Methods of variational analysis in pessimistic bilevel programming.* PhD thesis, Department of Mathematics, Wayne State University, 2010.

[8] S. Dempe. *Foundations of bilevel programming.* Kluwer, Dordrecht, 2002.

[9] 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. Springer, Berlin, 2010.

[10] S. Dempe and J. Dutta. Is bilevel programming a special case of mathematical programming with equilibrium constraints? *Math. Program.*, to appear (DOI 10.1007/s10107-010-0342-1).

[11] S. Dempe, J. Dutta, and B.S. Mordukhovich. New necessary optimality conditions in optimistic bilevel programming. *Optimization*, 56(5-6):577–604, 2007.

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

[13] S. Dempe and A.B. Zemkoho. On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem. *Nonlinear Analysis*, to appear (Doi:10.1016/j.na.2011.05.097).

[14] S. Dempe and A.B. Zemkoho. The bilevel programming problem: reformulations, constraint qualifications and optimality conditions, submitted.

[15] A. Dhara and A. Mehra. Approximate optimality conditions for minimax programming problems. *Optimization*, 59(7-8):1013–1032, 2010.

[16] N. Dinh, B.S. Mordukhovich, and T.T.A. Nghia. Subdifferentials of value functions and optimality conditions for DC and bilevel infinite and semi-infinite programs. *Math. Program.*, 123(1):101–138, 2010.

[17] J. Dutta and S. Dempe. Bilevel programming with convex lower level problems. S. Dempe and V. Kalashnikov (eds.), Optimization with multivalued mappings. Springer, New York, 2006.

[18] M.C. Ferris. *Weak sharp minima and penalty functions in mathematical programming.* PhD thesis, Churchill College, University of Cambridge, 1989.

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

[20] M. Flegel. *Constraint qualification and stationarity concepts for mathematical programs with equilibrium constraints.* PhD thesis, Institute of Applied Mathematics and Statistics, University of Würzburg, 2005.

[21] M.L. Flegel and C. Kanzow. A direct proof for M-stationarity under MPEC-GCQ for mathematical programs with equilibrium constraints. S. Dempe and V. Kalashnikov (eds), Optimization with multivalued mappings. Springer, New York, 2006.

[22] M.L. Flegel, C. Kanzow, and J.V. Outrata. Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. *J. Set-Valued Anal.*, 15(2):139–162, 2007.

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

[24] T. Hoheisel, C. Kanzow and J.V. Outrata. Exact penalty results for mathematical programs with vanishing constraints. *Nonlinear Anal.*, 72:2514-2526, 2010.

[25] R. Henrion, A. Jourani, and J.V. Outrata. On the calmness of a class of multifunctions. *SIAM J. Optim.*, 13(2):603–618, 2002.

[26] R. Henrion and J.V. Outrata. A subdifferential condition for calmness of multifunctions. *J. Math. Anal. Appl.*, 258(1):110–130, 2001.

[27] R. Henrion and T. Surowiec. On calmness conditions in convex bilevel programming. *Applicable Analysis*, 90(6):951–970, 2011.

[28] T. Hoheisel, C. Kanzow, and J.V. Outrata. Exact penalty results for mathematical programs with vanishing constraints. *Nonlinear Analysis*, 72:2514-2526, 2010.

[29] X. Hu and D. Ralph. A note on sensitivity of value functions of mathematical programs with complementarity constraints. *Math. Program.*, 93:265–279, 2002.

[30] S. Lohse. *Eine spezielle klasse von zwei-ebenen-optimierungsaufgaben.* PhD thesis, Fakultät für Mathematik und Informatik, Technische Universität Bergakademie Freiberg, 2010.

[31] Y. Lucet and J.J. Ye. Sensitivity analysis of the value function for optimization problems with variational inequality constraints. *SIAM J. Control Optim.*, 40(3):699–723, 2001. (with Erratum in *SIAM J. Control Optim.*, 41(4):1315-1319, 2002).

[32] Z.Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints.* Cambridge University Press. Cambridge, UK, 1996.

[33] O.L. Mangasarian and S. Fromovitz. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.*, 17:37–47, 1967.

[34] B.S. Mordukhovich. Maximum principle in problems of time optimal control with nonsmooth constraints. *J. Appl. Math. Mech.*, 40:690–669, 1976.

[35] B.S. Mordukhovich. Metric approximations and necessaery optimality conditions for general clases of extremal problems. *Soviet Math. Dokl.*, 22: 526–530, 1980.

[36] B.S. Mordukhovich. Nonsmooth analysis with nonconvex generalized differentials and conjugate mappings. *Dokl. Akad. Nauk BSSR*, 28:976–979, 1984.

[37] B. S. Mordukhovich. Complete characterizations of openness, metric regularity, and Lipschitzioan properties of multifunctions. *Trans. Amer. Math. Soc.*, 340:1–35, 1993.

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

[39] B.S. Mordukhovich and N.M. Nam. Variational stability and marginal functions via generalized differentiation. *Math. Oper. Res.*, 30(4):800–816, 2005.

[40] B.S. Mordukhovich, M.N. Nam and H.M. Phan. Variational analysis of marginal functions with applications to bilevel programming, *J. Optim. Theory Appl.*, 152 (2011), No. 3.

[41] B.S. Mordukhovich, N.M. Nam and N.D. Yen. Fréchet subdifferential calculus and optimality conditions in nondifferentiable programming. *Optimization*, 55(5-6):685–708, 2006.

[42] B.S. Mordukhovich, M.N. Nam and N.D. Yen. Subgradients of marginal functions in parametric mathematical programming. *Math. Program.*, 116(1-2):369–396, 2009.

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

[44] J.V. Outrata. On the numerical solution of a class of Stackekberg problems. *ZOR-Methods and Models Oper. Res.*, 34:255–277, 1990.

[45] J.V. Outrata. Optimality conditions for a class of mathematical programs with equilibrium constraints. *Math. Oper. Res.*, 24:627–644, 1999.

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

[47] B.T. Polyak. *Sharp minima.* Institute of Control Sciences Lecture Notes, Moscow, USSR, 1979; Presented at the IIASA Workshop on Generalized Lagrangians and Their Applications, IIASA, Laxenburg, Austria, 1979.

[48] S.M. Robinson. Some continuity properties of polyhedral multifunctions. *Math. Program. Study*, 14:206–214, 1981.

[49] R.T. Rockafellar. Marginal values and second-order necessary conditions for optimality. *Math. Program.*, 26:245–286, 1983.

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

[51] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity. *Math. Oper. Res.*, 25(1):1–22, 2000.

[52] W. Schirotzek. *Nonsmooth analysis.* Springer, Berlin, 2007.

[53] W.E. Schmitendorf. A simple derivation of necessary conditions for static minmax problems. *J. Math. Anal. Appl.*, 70:486–489, 1979.

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

[55] J.J. Ye. New uniform parametric error bounds. *J. Optim. Theory Appl.*, 98(1):197–219, 1998.

[56] J.J. Ye. Optimality conditions for optimization problems with complementarity constraints. *SIAM J. Optim.*, 9(2):374–387, 1999.

[57] J.J. Ye. Nondifferentiable multiplier rules for optimization and bilevel optimization problems. *SIAM J. Optim.*, 15(1):252–274, 2004.

[58] J.J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *J. Math. Anal. Appl.*, 307(1):350–369, 2005.

[59] J.J. Ye and X.Y. Ye. Necessary optimality conditions for optimization problems with variational inequality constraints. *Math. Oper. Res.*, 22(4):977-997, 1997.

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

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