

# The Generalized Mangasarian-Fromowitz Constraint Qualification and Optimality Conditions for Bilevel Programs

Stephan Dempe · Alain B. Zemkoho

Published online: 27 August 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** We consider the optimal value reformulation of the bilevel programming problem. It is shown that the Mangasarian-Fromowitz constraint qualification in terms of the basic generalized differentiation constructions of Mordukhovich, which is weaker than the one in terms of Clarke's nonsmooth tools, fails without any restrictive assumption. Some weakened forms of this constraint qualification are then suggested, in order to derive Karush-Kuhn-Tucker type optimality conditions for the aforementioned problem. Considering the partial calmness, a new characterization is suggested and the link with the previous constraint qualifications is analyzed.

**Keywords** Bilevel programming · Optimal value function · Basic generalized differentiation · Constraint qualification · Necessary optimality conditions

## 1 Introduction

The optimistic bilevel programming problem is an hierarchical optimization problem, where the so-called *upper level* problem is defined as

$$\min F(x, y) \quad \text{s.t. } (x, y) \in X \times \mathbb{R}^m, \quad y \in \Psi(x); \quad (1)$$

---

Communicated by V.F. Demyanov.

The authors are grateful to an anonymous referee for pointing an error in an example of an earlier version and for some useful comments, which helped to improve the quality of the paper. We would also like to acknowledge the valuable comments and suggestions from the editors, which vastly contributed to improve the presentation of the paper. Finally, the second author acknowledges the support by the Deutscher Akademischer Austausch Dienst (DAAD).

---

S. Dempe (✉) · A.B. Zemkoho  
Department of Mathematics and Computer Science, Technische Universität Bergakademie Freiberg,  
Freiberg, Germany  
e-mail: [dempe@math.tu-freiberg.de](mailto:dempe@math.tu-freiberg.de)

A.B. Zemkoho  
e-mail: [zemkoho@student.tu-freiberg.de](mailto:zemkoho@student.tu-freiberg.de)

with  $X$  being a closed subset of  $\mathbb{R}^n$  and  $\Psi(x)$  the solution set of the following parametric optimization problem called *lower level* problem:

$$\min f(x, y) \quad \text{s.t. } y \in K(x). \tag{2}$$

The functions  $F, f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  are continuous and throughout the paper,  $X$  and  $K(x)$  will be taken respectively as:

$$X := \{x \mid G(x) \leq 0, H(x) = 0\} \quad \text{and} \quad K(x) := \{y \mid g(x, y) \leq 0, h(x, y) = 0\},$$

where the functions  $G : \mathbb{R}^n \rightarrow \mathbb{R}^k, H : \mathbb{R}^n \rightarrow \mathbb{R}^l, g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^p$  and  $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q$  are also continuous.

If we assume that problem (1) has at least one feasible point, then a reformulation of this problem as a one level optimization problem is possible as:

$$\min F(x, y) \quad \text{s.t. } x \in X, y \in K(x), f(x, y) \leq \varphi(x), \tag{3}$$

where the optimal value function  $\varphi$  is defined by

$$\varphi(x) := \min\{f(x, y) \mid y \in K(x)\}.$$

Problem (3) is called the *optimal value reformulation* of problem (1). The two problems are globally and locally equivalent [1]. Our main concern in this paper is to derive KKT type optimality conditions for problem (3).

For many years now, most of the work on bilevel programming has been focussed on the KKT reformulation of (1). To obtain this reformulation, the lower level problem is replaced by its KKT conditions under the assumptions that the problem be convex and an appropriate constraint qualification be satisfied. Not only the KKT reformulation is not equivalent to the initial problem (1) for local solutions, but it is more demanding in terms of differentiation. In fact, deriving first order necessary optimality conditions for (1) using the KKT reformulation usually requires the computation of second order derivatives for the functions defining the lower level problem. More details on how to derive necessary optimality conditions for (1) using the KKT reformulation can be found in [2–5].

The price to pay for the optimal value reformulation (3) is due to the nature of the optimal value function itself, which is source of nonsmoothness and the failure of well-known constraint qualifications like the Mangasarian-Fromowitz constraint qualification (MFCQ), the linear independence constraint qualification (LICQ) and the Slater constraint qualification. Already in 1995, Ye and Zhu [6] showed that the nonsmooth version, in terms of Clarke’s generalized subdifferential of the MFCQ, fails for problem (3). In the same work, these authors introduced the weaker constraint qualification (CQ) named *partial calmness* after the stronger concept of calmness introduced by Clarke [7]. Using the Clarke subdifferential, necessary optimality conditions of KKT type were derived in [6, 8] under the partial calmness. Also Ye [9] extended the Abadie, Kuhn-Tucker, Zangwill, Arrow-Hurwicz-Uzawa, weak reverse and weak Slater constraint qualifications to problem (3). KKT type optimality conditions were then derived for problem (3) using the Michel-Penot subdifferential. Recently again Ye [10] considered problem (3), where the constraint function

$f(x, y) - \varphi(x)$  was replaced by a new function  $\psi(x, y)$  Lipschitz continuous near the optimal solution, thus exempting some common requirements on the lower level problem. Also the KKT optimality conditions were then derived in terms of Clarke's subdifferential under some of the CQs already used in [9].

Necessary optimality conditions for problem (3), of the KKT type, were also obtained recently by Babahadda and Gadhi [11] using convexificators. The constraint qualification used here is in the line of the regularity condition introduced in [12] using the concept of *approximation*, which is a nonsmooth differentiation tool.

More recent on the subject is the work of Dempe et al. [13]. In this paper, optimality conditions of KKT type are obtained for problem (3) with some new features, that is without dependence of the conditions on the partial derivatives of the lower level objective function with respect to the parameters from the upper level problem. It is important to mention that this has been possible thanks to the inner semicontinuity of  $\Psi$  required for the estimation of the subdifferential of  $\varphi$ . In fact, for most of the previous works (see e.g. [2]) on the subject upper semicontinuity has been used. Again the partial calmness was used here as constraint qualification.

In the present paper, we come back to the MFCQ. First we show that the Mordukhovich version of this CQ fails to hold for problem (3). Then, based on some works by Henrion and Outrata [14] and Henrion et al. [15] in the framework of the calmness of set-valued mappings, we observe that it is possible to weaken the MFCQ in order to derive KKT type optimality conditions for problem (3) using the basic subdifferential of Mordukhovich. Later in the paper we also suggest new sufficient conditions for the partial calmness based on a more general notion of the weak sharp minimum.

The rest of the paper is organized as follows. In next section, we recall some definitions of the nonsmooth tools needed in this work. The preliminaries are mainly concerned with relevant properties of the basic differentiation tools (i.e. the normal cone, subdifferential and coderivative) of Mordukhovich. Also of importance are some Lipschitz properties of multifunctions. In Sect. 3, we investigate a direct approach to optimality for problem (3) using the mentioned weakened form of the MFCQ. Section 4 is devoted to KKT conditions for (3) using the exact penalization. Also in this section we give new sufficient conditions for the *partial calmness*, which in some sense are equivalent to the *partial exact penalization* for problem (3).

In Sect. 5, we consider the strong stability in the sense of Kojima, for the parametric programming problem (2). We notice that the inner semicompactness of  $\Psi$  and the Aubin property of  $K$  needed to estimate the basic subdifferential of  $\varphi$  are no longer sufficient. The constant rank constraint qualification appears to be essential in order to derive necessary optimality conditions of KKT type for problem (3), under strong stability. Throughout the paper many examples are considered to illustrate some key points.

## 2 Preliminaries

In this section we present the basic concepts and notations used in the paper. More details on the material, briefly discussed here, can be found in [16–18]. We first consider

some notations: Let  $A$  be a subset of  $\mathbb{R}^n$ ,  $\text{co } A$ ,  $\text{cl } A$  and  $\text{int } A$  denote the convex hull, the closure and the interior of  $A$ , respectively. For a matrix  $B$ ,  $B^\top$  is the transposed matrix of  $B$  and, finally,  $\|\cdot\|$  denotes an arbitrary norm in  $\mathbb{R}^n$ .

Next we assume that  $A$  be a closed subset of  $\mathbb{R}^n$ . The contingent and Clarke’s tangent cone to  $A$  at some point  $\bar{x} \in A$  are defined respectively by

$$T_A(\bar{x}) := \{d \in \mathbb{R}^n \mid \exists t_k \downarrow 0, d_k \rightarrow d : \bar{x} + t_k d_k \in A\},$$

$$T_A^c(\bar{x}) := \{d \in \mathbb{R}^n \mid \forall t_k \downarrow 0, x_k \rightarrow \bar{x} (x_k \in A); \exists d_k \rightarrow d : x_k + t_k d_k \in A\}.$$

The respective normal cones, i.e. the regular and the Clarke normal cones, are obtained as

$$\widehat{N}_A(\bar{x}) := \{d^* \in \mathbb{R}^n \mid \langle d^*, d \rangle \leq 0, \forall d \in T_A(\bar{x})\},$$

$$N_A^c(\bar{x}) := \{d^* \in \mathbb{R}^n \mid \langle d^*, d \rangle \leq 0, \forall d \in T_A^c(\bar{x})\}.$$

Meanwhile, the basic normal cone introduced by Mordukhovich is defined as

$$N_A(\bar{x}) := \{d^* \in \mathbb{R}^n \mid \exists d_k^* \rightarrow d^*, x_k \rightarrow \bar{x} (x_k \in A) : d_k^* \in \widehat{N}_A(x_k)\}.$$

In contrast to the regular and the Clarke normal cones, which are convex, the basic normal cone is generally nonconvex. For this reason it cannot be polar to any tangential approximation of  $A$  [16]. But, in the context of optimality conditions, the basic normal cone presents one major advantage: it gives sharper conditions. To see this we mention the following relation between the basic and the Clarke normal cones

$$N_A^c(\bar{x}) = \text{cl co } N_A(\bar{x}), \tag{4}$$

which shows that the basic normal cone is included in the Clarke normal cone.

For a lower semicontinuous function  $\vartheta : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ , the Mordukhovich (or basic) and the Clarke subdifferential of  $\vartheta$  are, respectively, defined by

$$\partial \vartheta(\bar{x}) := \{x^* \in \mathbb{R}^n \mid (x^*, -1) \in N_{\text{epi } \vartheta}(\bar{x}, \vartheta(\bar{x}))\},$$

$$\partial^c \vartheta(\bar{x}) := \{x^* \in \mathbb{R}^n \mid (x^*, -1) \in N_{\text{epi } \vartheta}^c(\bar{x}, \vartheta(\bar{x}))\};$$

where  $\text{epi } \vartheta$  is the epigraph of  $\vartheta$ . A relation (where the closure is omitted), similar to the one mentioned above between the basic and the Clarke normal cones, can be established between the basic and the Clarke subdifferentials when  $\vartheta$  is Lipschitz continuous. Hence the following *convex hull property*

$$\text{co } \partial(-\vartheta)(\bar{x}) = -\text{co } \partial \vartheta(\bar{x}), \tag{5}$$

holds true when  $\vartheta$  is Lipschitz continuous. Also the basic subdifferential is nonempty and compact in the latter case. In addition, if  $\vartheta$  is continuously differentiable, then

$$\partial \vartheta(\bar{x}) = \{\nabla \vartheta(\bar{x})\}.$$

Given a set-valued mapping  $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , let us denote by  $\text{gph } M$  its graph:

$$\text{gph } M := \{(u, v) \in \mathbb{R}^n \times \mathbb{R}^m \mid v \in M(u)\}.$$

The coderivative of  $M$  at  $(\bar{x}, \bar{y}) \in \text{gph } M$  is a positively homogeneous mapping  $D^*M(\bar{x}, \bar{y}) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$  with the values

$$D^*M(\bar{x}, \bar{y})(y^*) := \{x^* \in \mathbb{R}^n \mid (x^*, -y^*) \in N_{\text{gph } M}(\bar{x}, \bar{y})\}, \quad \forall y^* \in \mathbb{R}^m,$$

where the argument  $\bar{y}$  is omitted, if  $M$  is single-valued. Moreover, for a single-valued Lipschitzian mapping  $M : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,

$$D^*M(\bar{x})(y^*) = \partial \langle y^*, M \rangle(\bar{x}), \quad \forall y^* \in \mathbb{R}^m, \tag{6}$$

with  $\langle y^*, M \rangle(\bar{x}) := \langle y^*, M(\bar{x}) \rangle$  and  $\partial$  being the basic subdifferential defined above. Again let us mention that, if the single-valued mapping  $M$  is continuously differentiable around  $\bar{x}$ , then

$$D^*M(\bar{x})(y^*) = \{\nabla M(\bar{x})^\top y^*\}, \quad \forall y^* \in \mathbb{R}^m,$$

where  $\nabla M(\bar{x})$  is the Jacobian matrix of  $M$ .

Given a set-valued mapping  $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  and a point  $\bar{x}$  with  $M(\bar{x}) \neq \emptyset$ , we say that  $M$  is *inner semicompact* at  $\bar{x}$ , if and only if, for every sequence  $x_k \rightarrow \bar{x}$  with  $M(x_k) \neq \emptyset$  there is a sequence of  $y_k \in M(x_k)$  that contains a convergent subsequence as  $k \rightarrow \infty$ . It follows that the inner semicompactness holds whenever  $M$  is *uniformly bounded* around  $\bar{x}$ , i.e. there exist a neighborhood  $U$  of  $\bar{x}$  and a bounded set  $A \subset \mathbb{R}^m$ , such that  $M(x) \subseteq A, \forall x \in U$ .

A set-valued mapping  $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  satisfies the Aubin/Lipschitz-like property around the point  $(\bar{x}, \bar{y}) \in \text{gph } M$ , if and only if, there are neighborhoods  $U$  of  $\bar{x}$ ,  $V$  of  $\bar{y}$  and a constant  $L > 0$ , such that

$$d(y, M(x_2)) \leq L \|x_1 - x_2\|, \quad \forall x_1, x_2 \in U, \forall y \in M(x_1) \cap V,$$

where  $d$  stands for a distance on  $\mathbb{R}^m \times \mathbb{R}^m$ . The Lipschitz modulus is the infimum of the numbers  $L$  satisfying Aubin’s property. Estimations of the Lipschitz modulus can be found in [16].

A weaker concept of Lipschitz-like behavior of set-valued mappings is that of calmness, which is satisfied at some point  $(\bar{x}, \bar{y}) \in \text{gph } M$ , if and only if, there exist neighborhoods  $U$  of  $\bar{x}$ ,  $V$  of  $\bar{y}$  and a constant  $L > 0$ , such that

$$d(y, M(\bar{x})) \leq L \|x - \bar{x}\|, \quad \forall x \in U, \forall y \in M(x) \cap V.$$

The modulus of calmness of  $M$  at  $(\bar{x}, \bar{y})$  (defined in the like manner to the Lipschitz modulus) is denoted by  $L(M; \bar{x}, \bar{y})$ . Following the very recent work by Zheng and Ng [18],  $L(M; \bar{x}, \bar{y})$  can be computed using the *inner norm*. Given a positively homogeneous mapping  $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ , and a cone  $A$  in  $\mathbb{R}^n$ , the inner norm of  $M$  on  $A$  is defined by

$$\|M|_A\|^- := \sup_{x \in \mathbb{B} \cap A} \{\inf \|y\| : y \in M(x)\}; \tag{7}$$

where  $\mathbb{B}$  denotes the unit ball of  $\mathbb{R}^n$ . We have the following result from Zheng and Ng [18].

**Theorem 2.1** *If the set-valued mapping  $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is calm at  $(\bar{x}, \bar{y})$ , then*

$$L(M; \bar{x}, \bar{y}) \geq \limsup_{v \in M(\bar{x}), v \rightarrow \bar{y}} \|D^*M(\bar{x}, v)|_{-\widehat{N}_{M(\bar{x})}(v)}\|^-.$$

*Furthermore, if gph  $M$  is convex, then equality holds true.*

This result could be very useful (see Sect. 4) in order to obtain the *uniform calmness* of a certain family of set-valued mappings, which is sufficient to achieve the partial calmness mentioned in the Introduction.

### 3 A Direct Approach to Optimality

Consider a classical Lipschitz optimization problem

$$\min f(x) \quad \text{s.t. } x \in A, g(x) \leq 0, h(x) = 0, \tag{8}$$

where  $A$  is a closed subset of  $\mathbb{R}^n$  and the functions  $f : \mathbb{R}^n \rightarrow \mathbb{R}, g : \mathbb{R}^n \rightarrow \mathbb{R}^r$  and  $h : \mathbb{R}^n \rightarrow \mathbb{R}^s$  are different from those considered in Sect. 1. Let us consider a feasible point  $\bar{x}$  of problem (8); then it follows from Ye and Zhu [6] that the nonsmooth MFCQ for (8) at  $\bar{x}$ , has the dual form:

$$[0 \in \partial^c \Phi(\bar{x})^\top \zeta + N_A^c(\bar{x}), \zeta \in N_D^c(\Phi(\bar{x}))] \implies \zeta = 0, \tag{9}$$

where  $N^c$  and  $\partial^c$  denote Clarke’s normal cone and generalized Jacobian, respectively; and  $\Phi(x) := (g(x), h(x))^\top, D := \mathbb{R}_-^r \times \{0_s\}$ . Even though the Mordukhovich subdifferential does not take its source from a directional derivative, the MFCQ in terms of the basic generalized constructions can be defined similarly. For a feasible point  $\bar{x}$  of (8), the MFCQ, in terms of the basic generalized differentiation tools of Mordukhovich that we denote by MMFCQ, is satisfied at  $\bar{x}$ , if and only if

$$[0 \in \partial \langle \zeta, \Phi \rangle(\bar{x}) + N_A(\bar{x}), \zeta \in N_D(\Phi(\bar{x}))] \implies \zeta = 0, \tag{10}$$

where  $\partial$  and  $N$  stand for the basic subdifferential and basic normal cone, respectively. By applying (6) to (10), the MMFCQ can be rewritten using the coderivative. The condition (10) is closely related to the well-known coderivative criterion of a certain set-valued mapping; see e.g. Henrion and Outrata [14]. For more on the coderivative criterion and its applications in nonsmooth analysis, we refer the interested reader to the books by Rockafellar and Wets [17] and Mordukhovich [16]. It may also be important to mention that, if the functions  $g$  and  $h$  are continuously differentiable and the set  $A := \mathbb{R}^n$ , then conditions (9) and (10) coincide with the dual form of the well-known MFCQ of smooth functions [19].

We now consider our bilevel programming problem in the optimal value reformulation (3). We let

$$\Omega := \{(x, y) \mid x \in X, y \in K(x)\}$$

define an abstract constraint for the problem and from here on, we define  $\mathcal{G}(x, y) := f(x, y) - \varphi(x)$ . Then (3) has the form

$$\min F(x, y) \quad \text{s.t. } (x, y) \in \Omega, \mathcal{G}(x, y) \leq 0, \quad (11)$$

and one can easily verify that the MMFCQ for (11) at  $(\bar{x}, \bar{y})$  is

$$\partial\mathcal{G}(\bar{x}, \bar{y}) \cap (-N_{\Omega}(\bar{x}, \bar{y})) = \emptyset. \quad (12)$$

From now on, (12) will also be denoted as MMFCQ. An interesting thing about the MMFCQ for problem (11), in comparison to the nonsmooth MFCQ in terms of Clarke's tools, is that the earlier CQ is weaker. In fact, let  $\mathcal{G}$  be Lipschitz continuous; then, as mentioned in Sect. 2, we have  $\partial\mathcal{G}(\bar{x}, \bar{y}) \subseteq \partial^c\mathcal{G}(\bar{x}, \bar{y})$ . We also have from (4) that  $N_{\Omega}(\bar{x}, \bar{y}) \subseteq N_{\Omega}^c(\bar{x}, \bar{y})$ . Hence, it follows that:  $\partial^c\mathcal{G}(\bar{x}, \bar{y}) \cap (-N_{\Omega}^c(\bar{x}, \bar{y})) = \emptyset$  implies  $\partial\mathcal{G}(\bar{x}, \bar{y}) \cap (-N_{\Omega}(\bar{x}, \bar{y})) = \emptyset$ . Next, we show that the MMFCQ is violated at any feasible point of problem (11). The only condition needed here is a regularity condition, which ensures that the sum rule is applicable to the function  $\mathcal{G} + \delta_{\Omega}$ , where  $\delta_{\Omega}$  is the indicator function of the set  $\Omega$ .

**Theorem 3.1** *Let  $(\bar{x}, \bar{y})$  be an arbitrary feasible point of problem (11). Assume that*

$$\partial(\mathcal{G} + \delta_{\Omega})(\bar{x}, \bar{y}) \subseteq \partial\mathcal{G}(\bar{x}, \bar{y}) + \partial\delta_{\Omega}(\bar{x}, \bar{y}); \quad (13)$$

*then the MMFCQ fails at  $(\bar{x}, \bar{y})$ .*

*Proof* Assume that  $(\bar{x}, \bar{y})$  be a feasible point of (11). Then  $(\bar{x}, \bar{y}) \in \text{gph } \Psi$  and it follows that  $\mathcal{G}(\bar{x}, \bar{y}) = 0$ . On the other hand, for all  $x \in X$ , we have  $f(x, y) \geq \varphi(x)$ ,  $\forall y \in K(x)$ . It follows that  $\mathcal{G}(x, y) \geq 0$ ,  $\forall (x, y) \in \Omega$ . That is,  $\mathcal{G}(\bar{x}, \bar{y}) = 0 \leq \mathcal{G}(x, y)$ ,  $\forall (x, y) \in \Omega$ . Hence, from (13),  $0 \in \partial\mathcal{G}(\bar{x}, \bar{y}) + N_{\Omega}(\bar{x}, \bar{y})$  or equivalently  $\partial\mathcal{G}(\bar{x}, \bar{y}) \cap (-N_{\Omega}(\bar{x}, \bar{y})) \neq \emptyset$ .  $\square$

*Remark 3.2* A simple situation, where (13) holds, is when the lower level cost function  $f$  is continuously differentiable and the optimal value function  $\varphi$  is locally Lipschitz continuous [16]. For more details on regularity conditions for (13) to hold, we refer the reader to the books [16, 17].

Ye and Zhu [6] proved a similar result to Theorem 3.1, in terms of the Clarke's subdifferential, under the assumption that  $\varphi$  be locally Lipschitz continuous and equality holds in the estimation of Clarke's subdifferential as given by Gauvin and Dubeau [20, Theorem 5.3]. These assumptions seem to be restrictive and our result shows that the MFCQ fails for problem (11) in a broader sense.

It is possible to have a weaker CQ for problem (11). Recall that for a set  $A \subseteq \mathbb{R}^n$ , the topological boundary of  $A$  is defined as

$$\text{bd } A := \text{cl } A \setminus \text{int } A. \quad (14)$$

We consider the weaker CQ closely related to MMFCQ that we denote by WMFCQ:

$$\partial\mathcal{G}(\bar{x}, \bar{y}) \cap (-\text{bd } N_{\Omega}(\bar{x}, \bar{y})) = \emptyset. \quad (15)$$

It does not make any doubt that the WMFCQ be implied by the MMFCQ since the normal cone  $N_{\Omega}(\bar{x}, \bar{y})$  is closed and  $\text{bd } N_{\Omega}(\bar{x}, \bar{y}) \subseteq N_{\Omega}(\bar{x}, \bar{y})$ , cf. (14).

Assuming that the WMFCQ has some chances to be satisfied for problem (11), we can derive KKT type optimality conditions for this problem, in a direct way under WFMFQ. Before heading to that, the following result is crucial. For this, denote by  $C := \{(x, y) \in \Omega \mid \mathcal{G}(x, y) \leq 0\}$  the feasible set of (11).

**Lemma 3.3** *Let  $(\bar{x}, \bar{y})$  be a feasible point of problem (11). Assume that  $\Omega$  be convex and  $\mathcal{G}$  be Lipschitz continuous around  $(\bar{x}, \bar{y})$ . Then*

$$N_C(\bar{x}, \bar{y}) \subseteq \bigcup_{r \geq 0} r \partial \mathcal{G}(\bar{x}, \bar{y}) + N_{\Omega}(\bar{x}, \bar{y})$$

provided that WMFCQ holds true at  $(\bar{x}, \bar{y})$ .

*Proof* Assume that  $\Omega$  be convex. Then  $\Omega$  is regular in the sense of Clarke, and semi-smooth at  $(\bar{x}, \bar{y})$ . Thus the result follows from Henrion et al. [15, Theorem 4.1] and Henrion and Outrata [14, Theorem 3.1]. For the definition of semismoothness and some related properties, the interested reader is referred for example to [14].  $\square$

Unless otherwise stated, we assume from now on that the functions  $F, G, H, f, g$  and  $h$  defining problem (1) be continuously differentiable. The Lipschitz continuity of  $\mathcal{G}$  can then be achieved when  $\varphi$  is Lipschitz continuous. It appears that the following result due to Mordukhovich and Nam [21] and Mordukhovich [16] be needed to ensure the Lipschitz continuity of  $\varphi$  and the estimation of its basic subdifferential.

**Theorem 3.4** *Assume that  $\Psi$  be inner semicompact at  $\bar{x}$  and  $K$  satisfies Aubin’s property around  $(\bar{x}, y) \in \text{gph } K$ , for all  $y \in \Psi(\bar{x})$ . Then  $\varphi$  is Lipschitz continuous around  $\bar{x}$ . Furthermore the basic subdifferential of  $\varphi$  is estimated as*

$$\partial \varphi(\bar{x}) \subseteq \bigcup_{y \in \Psi(\bar{x})} \left[ \bigcup_{(\lambda, \mu) \in \Lambda(\bar{x}, y)} \left\{ \nabla_x f(\bar{x}, y) + \nabla_x g(\bar{x}, y)^\top \lambda + \nabla_x h(\bar{x}, y)^\top \mu \right\} \right].$$

The set  $\Lambda(\bar{x}, y)$  of Lagrange multipliers for the lower level problem (2), when the parameter is fixed at  $\bar{x}$ , is defined as

$$\Lambda(\bar{x}, y) := \left\{ (\lambda, \mu) \in \mathbb{R}^p \times \mathbb{R}^q \mid \nabla_y f(\bar{x}, y) + \nabla_y g(\bar{x}, y)^\top \lambda + \nabla_y h(\bar{x}, y)^\top \mu = 0, \right. \\ \left. \lambda \geq 0, \lambda^\top g(\bar{x}, y) = 0 \right\}.$$

The Lipschitz continuity of  $\varphi$  is also achievable under the inner semicontinuity of  $\Psi$  [21]. Klatté and Kummer [22] also established the Lipschitz continuity of  $\varphi$  under the upper semicontinuity of  $\Psi$  and an estimation of  $\partial^c \varphi$  has been given by Gauvin and Dubeau [20], which differs from the one given for  $\partial \varphi$  by the fact that instead the convex hull is considered in the right hand side of the inclusion in Theorem 3.4.

We are now ready to state our result on the KKT type necessary optimality conditions of problem (11) under the WMFCQ. For this reason, the following regularity



conditions are necessary. Consider a feasible point  $(\bar{x}, \bar{y})$  to problem (11), the conditions

$$\begin{aligned} & \left[ \nabla G(\bar{x})^\top \lambda' + \nabla H(\bar{x})^\top \mu' = 0, (\mu', \lambda') \in \mathbb{R}^l \times N_{\mathbb{R}^k} (G(\bar{x})) \right] \implies \\ & (\lambda', \mu') = 0, \\ & \left[ \nabla_y g(\bar{x}, \bar{y})^\top \lambda + \nabla_y h(\bar{x}, \bar{y})^\top \mu = 0, (\mu, \lambda) \in \mathbb{R}^q \times N_{\mathbb{R}^p} (g(\bar{x}, \bar{y})) \right] \implies \\ & (\lambda, \mu) = 0 \end{aligned}$$

define the *upper level* and *lower level* regularity condition at  $\bar{x}$  and  $(\bar{x}, \bar{y})$ , respectively. Let us remind that this is nothing, but the MFCQ as given in (9)–(10), for the upper level and the unperturbed lower level constraints, respectively.

**Theorem 3.5** *Let  $(\bar{x}, \bar{y})$  be a local optimal solution to problem (11). Assume that  $\Psi$  is inner semicompact at  $\bar{x}$ , the lower (resp. upper) level regularity is satisfied at  $(\bar{x}, y)$ , for all  $y \in \Psi(\bar{x})$  (resp.  $\bar{x}$ ). Moreover, we suppose that  $\Omega$  is convex and the WMFCQ holds true at  $(\bar{x}, \bar{y})$ . Then, there exist  $r \geq 0, \lambda, \mu, \lambda', \mu', \lambda_s, \mu_s, \eta_s \geq 0$ , and  $y_s \in \Psi(\bar{x}), s = 1, \dots, n + 1$  with  $\sum_{s=1}^{n+1} \eta_s = 1$ , such that:*

$$\begin{aligned} & \nabla_x F(\bar{x}, \bar{y}) + r \nabla_x f(\bar{x}, \bar{y}) - r \sum_{s=1}^{n+1} \eta_s \nabla_x f(\bar{x}, y_s) \\ & + \nabla_x g(\bar{x}, \bar{y})^\top \lambda - r \sum_{s=1}^{n+1} \eta_s \nabla_x g(\bar{x}, y_s)^\top \lambda_s \\ & + \nabla_x h(\bar{x}, \bar{y})^\top \mu - r \sum_{s=1}^{n+1} \eta_s \nabla_x h(\bar{x}, y_s)^\top \mu_s \\ & + \nabla G(\bar{x})^\top \lambda' + \nabla H(\bar{x})^\top \mu' = 0, \end{aligned} \tag{16}$$

$$\nabla_y F(\bar{x}, \bar{y}) + r \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^\top \lambda + \nabla_y h(\bar{x}, \bar{y})^\top \mu = 0, \tag{17}$$

$$\nabla_y f(\bar{x}, y_s) + \nabla_y g(\bar{x}, y_s)^\top \lambda_s + \nabla_y h(\bar{x}, y_s)^\top \mu_s = 0, \tag{18}$$

$$\lambda \geq 0, \lambda^\top g(\bar{x}, \bar{y}) = 0, \tag{19}$$

$$\lambda' \geq 0, \lambda'^\top G(\bar{x}) = 0, \tag{20}$$

$$\lambda_s \geq 0, \lambda_s^\top g(\bar{x}, y_s) = 0. \tag{21}$$

*Proof* Let  $(\bar{x}, \bar{y})$  be a local optimal solution of problem (11). Since  $F$  is continuously differentiable at  $\bar{z}$ , it follows from Mordukhovich [16, Proposition 5.1] that

$$0 \in \nabla F(\bar{x}, \bar{y}) + N_C(\bar{x}, \bar{y}).$$

Now let us mention that the lower level regularity at  $(\bar{x}, y)$ , for all  $y \in \Psi(\bar{x})$  implies that Aubin’s property holds around  $(\bar{x}, y)$ , for all  $y \in \Psi(\bar{x})$ . Thus in addition to the

inner semicompactness of  $\Psi$  at  $\bar{x}$ , it follows from Theorem 3.4 that  $\varphi$  is Lipschitz continuous around  $\bar{x}$ . Hence, the convexity of  $\Omega$  and the Lipschitz continuity of  $\varphi$  imply from Lemma 3.3 that

$$0 \in \nabla F(\bar{x}, \bar{y}) + \bigcup_{r \geq 0} r \partial \mathcal{G}(\bar{x}, \bar{y}) + N_{\Omega}(\bar{x}, \bar{y}).$$

It follows that there exist  $r \geq 0$  and  $u \in N_{\Omega}(\bar{x}, \bar{y})$  such that

$$-(\nabla F(\bar{x}, \bar{y}) + r \nabla f(\bar{x}, \bar{y}) + u) \in r \partial(-\varphi)(\bar{x}) \times \{0\},$$

implying that there exist  $r \geq 0$  and  $u \in N_{\Omega}(\bar{x}, \bar{y})$  such that

$$(\nabla F(\bar{x}, \bar{y}) + r \nabla f(\bar{x}, \bar{y}) + u) \in r \text{co} \partial \varphi(\bar{x}) \times \{0\} \tag{22}$$

from

$$\partial(-\varphi)(\bar{x}) \subseteq \text{co} \partial(-\varphi)(\bar{x}) = -\text{co} \partial \varphi(\bar{x}).$$

Since upper and lower level regularities are satisfied at  $\bar{x}$  and  $(\bar{x}, \bar{y})$  respectively, then  $u \in N_{\Omega}(\bar{x}, \bar{y})$  implies (cf. Rockafellar [23, Theorem 4.3]) that, there exist  $r \geq 0$  and  $(\lambda, \lambda', \mu, \mu') \in \mathbb{R}^p \times \mathbb{R}^k \times \mathbb{R}^q \times \mathbb{R}^l$  such that equality (17), (in)equalities (19)–(20) and

$$\begin{aligned} & (\nabla_x F(\bar{x}, \bar{y}) + r \nabla_x f(\bar{x}, \bar{y}) + \nabla_x g(\bar{x}, \bar{y})^\top \lambda + \nabla_x h(\bar{x}, \bar{y})^\top \mu \\ & + \nabla G(\bar{x})^\top \lambda' + \nabla H(\bar{x})^\top \mu') \in r \text{co} \partial \varphi(\bar{x}) \end{aligned} \tag{23}$$

are satisfied.

Picking  $v \in \text{co} \partial \varphi(\bar{x})$  and applying Carathéodory’s theorem (see e.g. Mangasarian [24]), we find  $\eta_s \in \mathbb{R}$  and  $v_s \in \mathbb{R}^n$  with  $s = 1, \dots, n + 1$  such that

$$v = \sum_{s=1}^{n+1} \eta_s v_s, \quad \sum_{s=1}^{n+1} \eta_s = 1, \quad \eta_s \geq 0, \quad v_s \in \partial \varphi(\bar{x}), \quad \text{for } s = 1, \dots, n + 1. \tag{24}$$

Applying Theorem 3.4 to (24), we have vectors  $y_s \in \Psi(\bar{x})$  and  $(\lambda_s, \mu_s) \in \Lambda(\bar{x}, y_s)$  with  $s = 1, \dots, n + 1$  such that,

$$v_s = \nabla_x f(\bar{x}, y_s) + \nabla_x g(\bar{x}, y_s)^\top \lambda_s + \nabla_x h(\bar{x}, y_s)^\top \mu_s, \tag{25}$$

and equality (18), (in)equality (21) are satisfied. The result then follows by combining (23)–(25). □

*Remark 3.6* As mentioned in the Introduction of this paper, it was observed by Dempe et al. [13] that under the inner semicontinuity of  $\Psi$ , the partial calmness CQ leads to KKT conditions for problem (3) without the partial derivative of the lower level cost function with respect to the upper level variable. In Theorem 3.5, if we replace the inner semicompactness assumption on  $\Psi$  by the inner semicontinuity, the WMFCQ will also lead to the same phenomenon.

We now consider the class of *simple convex bilevel programming problems* studied for example by Dempe et al. [25]:

$$\min F(z) \quad \text{s.t. } z \in S := \operatorname{argmin}\{f(z)|z \in \Omega\}, \tag{26}$$

where  $\Omega$  is a closed and convex set, and the upper and lower level objective functions  $F$  and  $f$  are also convex. Denoting by  $\alpha := \min\{f(z)|z \in \Omega\}$ , problem (26) can be reformulated as

$$\min F(z) \quad \text{s.t. } z \in \Omega, \quad f(z) \leq \alpha. \tag{27}$$

This is a convex optimization problem, but the MFCQ also fails at any feasible point [25]. KKT type necessary optimality conditions can be derived for problem (27) by applying the technique in Theorem 3.5. In this case, the WMFCQ reduces to:

$$\partial f(\bar{z}) \cap (-\operatorname{bd} N_{\Omega}(\bar{z})) = \emptyset. \tag{28}$$

The following example from the class of simple convex bilevel programming problems shows that the WMFCQ could be quite useful.

*Example 3.7* We consider the simple convex bilevel programming problem:

$$\min x^2 + y^2 \quad \text{s.t. } (x, y) \in S := \operatorname{argmin}\{x + y|x, y \geq 0\}.$$

We have  $f(x, y) := x + y$  and  $\Omega := \mathbb{R}_+^2$ . The point  $(0, 0)$  is the unique optimal solution of the problem and  $N_{\Omega}(0, 0) = \mathbb{R}_-^2$ . Hence,

$$\partial f(0, 0) = (1, 1) \notin \{(x, 0)|x \geq 0\} \cup \{(0, y)|y \geq 0\} = -\operatorname{bd} N_{\Omega}(0, 0).$$

Clearly, condition (28) is fulfilled at  $(0, 0)$ .

Continuing with the simple convex bilevel programming problem, the WMFCQ can further be weakened by also passing to the boundary of the subdifferential of the optimal value constraint function  $\mathcal{G}$ , which corresponds here to  $z \mapsto f(z) - \alpha$ . Precisely, unlike in (28), if we assume that the condition

$$\operatorname{bd} \partial f(\bar{z}) \cap (-\operatorname{bd} N_{\Omega}(\bar{z})) = \emptyset \tag{29}$$

be satisfied at a local optimal solution  $\bar{z}$  of problem (27), then KKT type optimality conditions can also be derived for this problem using the same technique as in Theorem 3.5. To proceed, one can state a result similar to Lemma 3.3, under condition (29); cf. Henrion and Outrata [14, Theorem 4.2]. It is quite obvious that the qualification condition in (28) implies the one in (29). Hence the latter is also satisfied for the problem in Example 3.7.

Another alternative to the MFCQ in the framework of the simple convex bilevel programming problem may be the qualification condition

$$\operatorname{int} \left[ \bigcup \{\gamma \partial f(\bar{z}) | \gamma \in [0, 1]\} \right] \cap (-N_{\Omega}(\bar{z})) \neq \emptyset. \tag{30}$$

KKT type optimality conditions can also be derived for problem (27) under the CQ (30) by applying the technique of Theorem 3.5, cf. [15, Theorem 3.6]. The following example shows that this may also be a quite useful CQ for the bilevel programming problem.

*Example 3.8* Let us consider the simple convex bilevel programming problem:

$$\min x^2 + 1 \quad \text{s.t. } x \in S := \operatorname{argmin}\{x \mid x \geq 0\}.$$

We have  $f(x) := x$  and  $\Omega := \mathbb{R}_+$ . We have 0 as the unique optimal solution of the problem. Also,  $\bigcup\{\gamma \partial f(0) \mid \gamma \in [0, 1]\} = [0, 1]$  and  $N_\Omega(0) = \mathbb{R}_-$ . One can easily verify that condition (30) is fulfilled at 0.

It is important to mention that necessary and sufficient optimality conditions for the simple convex bilevel programming problem were derived in [25] using other CQs.

An approach similar to the one we have used to derive optimality conditions for the optimal value reformulation of (1), can also be applied for the KKT reformulation. To see this we first assume that both the function  $f(x, \cdot)$  and the set  $K(x)$  be convex for all  $x \in X$ , and also that the lower level regularity holds at  $(x, y) \in \operatorname{gph} K$ , for all  $x \in X$ . Hence the feasible set of (1) can be replaced by

$$C' := \{(x, y) \in X \times \mathbb{R}^m \mid 0 \in \nabla_y f(x, y) + N_{K(x)}(y)\}.$$

We now assume  $K(x) := K$ , i.e. the feasible set of the lower level problem does not depend on the upper level variable. Then problem (1) can be reformulated in a one level optimization problem as

$$\min F(x, y) \quad \text{s.t. } (x, y) \in X \times \mathbb{R}^m, (y, -\nabla_y f(x, y)) \in \operatorname{gph} N_K, \quad (31)$$

where  $\operatorname{gph} N_K$  represents the graph of the set-valued mapping  $N_K$  defined from  $\mathbb{R}^m$  to  $\mathbb{R}^m$  as  $N_K(y)$  being the normal cone (in the sense of convex analysis) to  $K$  at  $y$ , when  $y \in K$ ; and  $N_K(y) := \emptyset$  otherwise. It follows from [23, Proposition 3.3] that  $\operatorname{gph} N_K$  is closed as a subset of  $K \times \mathbb{R}^m$ . Problem (31) is locally and globally equivalent to problem (1) [1]. Now let  $\Omega' := X \times \mathbb{R}^m$  and consider the regularity condition SCQ:

$$\left. \begin{aligned} & \left( -[\nabla_{xy}^2 f(\bar{x}, \bar{y})]^\top z, w - [\nabla_{yy}^2 f(\bar{x}, \bar{y})]^\top z \right) \in -\operatorname{bd} N_{\Omega'}(\bar{x}, \bar{y}) \\ & (w, z) \in N_{\operatorname{gph} N_K}(y, -\nabla_y f(\bar{x}, \bar{y})) \end{aligned} \right\} \Rightarrow (w, z) = 0$$

defined at a feasible point problem  $(\bar{x}, \bar{y})$  of (31). Then using SCQ, necessary optimality conditions for (1) can be obtained as follows.

**Theorem 3.9** *Assume that  $X$  be convex and let  $(\bar{x}, \bar{y})$  be an optimal solution of (31) that satisfies SCQ. Then, there exist  $(\bar{w}, \bar{z}) \in N_{\operatorname{gph} N_K}(y, -\nabla_y f(\bar{x}, \bar{y}))$  and  $\gamma \in N_X(\bar{x})$ , such that*

$$\nabla_x F(\bar{x}, \bar{y}) - (\nabla_{xy}^2 f(\bar{x}, \bar{y}))^\top \bar{z} + \gamma = 0,$$

$$\nabla_y F(\bar{x}, \bar{y}) - (\nabla_{yy}^2 f(\bar{x}, \bar{y}))^\top \bar{z} + \bar{w} = 0.$$

The SCQ was introduced in [14], in the framework of mathematical programming problems with complementarity constraints. As far as the bilevel programming problem is concerned, an approach similar to the one of Theorem 3.9 has been used in [5]; at the difference that instead of  $\text{bd } N_{\Omega'}$  (in the SCQ), the normal cone  $N_{\Omega'}$  was considered. Thus, making the resulting CQ stronger than the SCQ. The SCQ could be a quite effective CQ for the bilevel programming problem (see [1, 26]); however, the estimation of  $N_{\text{gph } N_K}$  may require very strong assumptions [27]. Some links between the KKT conditions in Theorem 3.9 and those in Theorem 3.5 can be found in [1, 26].

In the line of direct approaches to derive KKT optimality conditions for the bilevel programming problem, let us mention the work of Ye [9] where some classical CQs have been extended and used to obtain optimality conditions for the optimal value reformulation of problem (1) using the Michel-Penot subdifferential. The optimality conditions derived in [9] are similar to those obtained in Theorem 3.5.

To close this section, it is worth mentioning that, the CQs in (29) and (30) may well be extended to problem (11), provided that  $\mathcal{G}$  is regular at  $(\bar{x}, \bar{y})$ , in the sense of Clarke. In the next section, we introduce the concept of partial calmness, a CQ that may be used to design KKT type optimality conditions for problem (3) via (partial) exact penalization. It will be shown that the WMFCQ strictly implies the partial calmness. Another possible alternative to the MMFCQ may be the CQ introduced in [28, 29]. This CQ is obtained by separation arguments in terms of tangent cone of a suitable set in the image space. It has been shown to be weaker than the MFCQ and its closeness with calmness and other well-known CQs has been established in [28, 29]. The application of this CQ to the bilevel programming problem may constitute a topic for future research.

#### 4 Optimality Conditions via Exact Penalization

In this section, we are mainly concerned with the concept of *partial calmness* introduced by Ye and Zhu [6], and which has drawn a lot of attention (see e.g. [1, 6, 8, 13, 30–32]) in the study of optimality conditions for the bilevel programming problem (1), using the optimal value reformulation (3).

**Definition 4.1** Let  $(\bar{x}, \bar{y})$  be a feasible point of problem (3). The bilevel programming problem (1) is partially calm at  $(\bar{x}, \bar{y})$ , if and only if, there exist  $\alpha > 0$  and a neighborhood  $U$  of  $(\bar{x}, \bar{y}, 0) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ , such that:

$$\begin{aligned} F(x, y) - F(\bar{x}, \bar{y}) + \alpha|u| &\geq 0, \\ \forall (x, y, u) \in U : (x, y) \in \Omega, f(x, y) - \varphi(x) + u &= 0. \end{aligned}$$

The notion of partial calmness is closely related to that of *partial exact penalization* as shown by the following result from [6].

**Theorem 4.2** *Let  $(\bar{x}, \bar{y})$  be a local optimal solution of (1). Then, problem (1) is partially calm at  $(\bar{x}, \bar{y})$ , if and only if, there exists  $\alpha > 0$  such that,  $(\bar{x}, \bar{y})$  is a local optimal solution to the partially penalized problem:*

$$\min F(x, y) + \alpha(f(x, y) - \varphi(x)) \quad \text{s.t. } x \in X, y \in K(x).$$

To obtain sufficient conditions for partial calmness, the concept of weak sharp minimum has been used [6, 8]. We now introduce a notion of weak sharp minimum from Henrion et al. [15], which generalizes the one used by Ye and Zhu [6]. For this reason, we consider a general optimization problem

$$\min f(x) \quad \text{s.t. } x \in A, \tag{32}$$

where  $A$  is a closed subset of  $\mathbb{R}^n$ , the function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is continuous and we denote the solution set by  $S$ .

**Definition 4.3** In (32), the function  $f$  is said to have  $S$  as set of weak sharp minima, with respect to  $A \cap \mathcal{N}$ , if and only if, there exists  $\alpha > 0$  such that

$$f(x) - f_* \geq \alpha d(x, S), \quad \forall x \in A \cap \mathcal{N};$$

where  $f_* := \inf\{f(x)|x \in A\}$  and  $\mathcal{N}$  is a neighborhood of  $S$ .

By replacing the set  $\mathcal{N}$  by the whole space  $\mathbb{R}^n$ , we obtain the definition used by Ye and Zhu [6]. As we will see in what follows, Definition 4.3 can lead to a new sufficient condition for partial calmness. Before heading to that, it seems appropriate to first recall the link between the partial calmness and the notion of weak sharp minimum. For this, we bring the previous definition to the context of the parametric optimization problem (2).

The family of parametric problems  $\{(2)|x \in X\}$ , will be said to have a *uniformly weak sharp minimum*, if and only if, there exist  $\alpha > 0$  and a neighborhood  $\mathcal{N}(x)$  of  $\Psi(x)$ ,  $x \in X$  such that

$$f(x, y) - \varphi(x) \geq \alpha d(y, \Psi(x)), \quad \forall y \in K(x) \cap \mathcal{N}(x), \forall x \in X.$$

The term *uniformly weak sharp minimum* was first used by Ye and Zhu [6]. We now present the following result, without the proof, since it can easily be adapted from the one in [6, Proposition 5.1].

**Theorem 4.4** *Let  $(\bar{x}, \bar{y})$  be an optimal solution of problem (1). Assume that  $F$  is Lipschitz continuous in  $y$  uniformly in  $x \in X$ , and the family  $\{(2)|x \in X\}$  has a uniformly weak sharp minimum. Then problem (1) is partially calm at  $(\bar{x}, \bar{y})$ .*

Now, we define the family of functions  $h_x(y) := f(x, y) - \varphi(x)$  and multifunctions

$$M_x(z) := \{y \in K(x) | h_x(y) + z \leq 0\}, \quad \forall x \in X. \tag{33}$$

From the definition of the calmness of a multifunction given in Sect. 2, we recall that for a given parameter  $x \in X$ , the multifunction  $M_x$  will be said to be calm at  $(\bar{z}, \bar{y}) \in \text{gph } M_x$ , if and only if, there exist neighborhoods  $U_x$  of  $\bar{z}$ ,  $V_x$  of  $\bar{y}$  and a constant  $L_x > 0$  such that

$$d(y, M_x(\bar{z})) \leq L_x |z - \bar{z}|, \quad \forall z \in U_x, \forall y \in M_x(z) \cap V_x. \quad (34)$$

The family of multifunctions  $\{M_x | x \in X\}$  will be said to be *uniformly calm*, if and only if, for each  $x \in X$ ,  $M_x$  is calm on  $\{0\} \times \Psi(x)$  and there is a family  $\{L_x(y) | y \in \Psi(x), x \in X\}$  of calmness constants satisfying

$$L_x(y) \leq \alpha, \quad \forall y \in \Psi(x), x \in X; \quad (35)$$

where  $\alpha$  is a positive number. Hence, it should be clear in the definition of *uniform calmness* that, when we fix  $x \in X$  and a vector  $y \in \Psi(x)$ , we consider only a certain calmness constant  $L_x(y)$  satisfying (34) and such that  $L_x(y) \leq \alpha$ . It is also clear from the definition of  $M_x$  in (33) that  $M_x(0) = \Psi(x)$ . Hence, the following result ensuring the calmness of  $M_x$  on  $\{0\} \times \Psi(x)$  is a simple consequence of [14, Theorem 3.1].

**Theorem 4.5** *Assume that  $f$  be Lipschitz continuous in  $y$  uniformly in  $x$ ,  $\varphi$  continuous and  $K$  be a convex-valued mapping. Then  $M_x$  is calm at  $(0, y)$  for all  $y \in \Psi(x)$ ,  $x \in X$ ; provided that the qualification condition*

$$\partial_y f(x, y) \cap -bd N_{K(x)}(y) = \emptyset, \quad \forall y \in \Psi(x), x \in X$$

*holds true.*

Interested readers are also referred to Heerda and Kummer [33] for other ways to characterize the calmness of the mappings  $M_x$ ,  $x \in X$ .

We are now able to give a new sufficient condition for problem (1) to be partially calm. The proof is inspired from [15] and the set-valued mapping  $M_x$  is defined as in (33).

**Theorem 4.6** *We assume that the family  $\{M_x | x \in X\}$  be uniformly calm and the set-valued mapping  $\Psi$  be compact-valued with  $\text{dom } \Psi = X$ . We further assume that the follower's cost function  $f$  be continuous in  $y$  uniformly in  $x$ . Then, there exists  $\alpha > 0$ , such that*

$$f(x, y) - \varphi(x) \geq \alpha^{-1} d(y, \Psi(x)), \quad \forall y \in K(x) \cap \mathcal{N}(x), x \in X,$$

where  $\mathcal{N}(x)$  is a neighborhood of  $\Psi(x)$ .

*Proof* Fix  $x_0 \in X$  and let  $y \in \Psi(x_0)$ .  $M_{x_0}$  is calm at  $(0, y)$  and it follows from the definition of calmness (34), that there exist  $\varepsilon_{x_0}, \delta_{x_0} > 0$  such that

$$d(y', M_{x_0}(0)) \leq L_{x_0}(y) |z|, \quad \forall z : |z| < \delta_{x_0}, \forall y' \in M_{x_0}(z) \cap B(y, \varepsilon_{x_0}). \quad (36)$$

Since  $f$  is continuous in  $y'$  uniformly in  $x$ , then we can choose  $\varepsilon_{x_0}$  small enough such that

$$|f(x_0, y') - f(x_0, y)| < \delta_{x_0}, \quad \forall y' \in B(y, \varepsilon_{x_0}),$$

with  $f(x_0, y) = \varphi(x_0)$ . Thus we have  $|h_{x_0}(y')| < \delta_{x_0}, \forall y' \in B(y, \varepsilon_{x_0})$ . Hence, by taking  $z = h_{x_0}(y')$  in (36) and observing that  $M_{x_0}(0) = \Psi(x_0)$ , we have

$$d(y', \Psi(x_0)) \leq L_{x_0}(y)|h_{x_0}(y')|, \quad \forall y' \in K(x_0) \cap B(y, \varepsilon_{x_0}).$$

Since the solution set-valued mapping  $\Psi$  is compact-valued; then, there is a finite number of vectors  $y^i \in \Psi(x_0)$  and real numbers  $\delta_{x_0}^i > 0$  and  $L_{x_0}(y^i) > 0$  such that  $\Psi(x_0) \subseteq \bigcup_i B(y^i, \delta_{x_0}^i)$ , and we have

$$d(y', \Psi(x_0)) \leq L_{x_0}(y^i)(f(x_0, y') - \varphi(x_0)), \quad \forall y' \in K(x_0) \cap B(y^i, \delta_{x_0}^i), \quad \forall i.$$

By taking  $\mathcal{N}(x_0) = \bigcup_i B(y^i, \delta_{x_0}^i)$ , it follows that

$$d(y', \Psi(x_0)) \leq c(x_0)(f(x_0, y') - \varphi(x_0)), \quad \forall y' \in K(x_0) \cap \mathcal{N}(x_0),$$

where  $c(x_0) = \max_i L_{x_0}(y^i)$ . Since the family of multifunctions  $\{M_x | x \in X\}$  is uniformly calm, we assume without loss of generality that the family of calmness constants  $\{L_x(y) | y \in \Psi(x), x \in X\}$  is chosen in such a way that, inequality (35) is satisfied. Hence, there exists  $\alpha > 0$  ( $c(x) < \alpha, \forall x \in X$ ) such that

$$f(x, y) - \varphi(x) \geq \alpha^{-1}d(y, \Psi(x)), \quad \forall y \in K(x) \cap \mathcal{N}(x), \quad \forall x \in X.$$

□

The next result gives a possible way to achieve the uniform calmness of the family of multifunctions  $\{M_x | x \in X\}$ , with the possibility of estimating the constant  $\alpha$ . We recall that  $L(M_x; 0, y)$  denotes the modulus of calmness of the mapping  $M_x$  at the point  $(0, y)$  and the restricted inner norm  $\|\cdot\|_A^-$  used here is defined in (7).

**Theorem 4.7** *Assume that the function  $f(x, \cdot)$  be convex for all  $x \in X$  and the set-valued mapping  $K$  be convex-valued. Also, we suppose that for all  $x \in X, M_x$  is calm on  $\{0\} \times \Psi(x)$ . Then, we have*

$$L(M_x; 0, y) = \limsup_{v \in \Psi(x), v \rightarrow y} \|D^*M_x(0, v)|_{-\widehat{N}_{\Psi(x)}(v)}\|^-, \quad \forall y \in \Psi(x), x \in X. \quad (37)$$

Furthermore, assume that for all  $y \in \Psi(x), x \in X, L(M_x; 0, y)$  be achieved as a minimum; and there is a number  $\alpha > 0$  such that

$$L(M_x; 0, y) \leq \alpha, \quad \forall y \in \Psi(x), x \in X. \quad (38)$$

Then, the family  $\{M_x | x \in X\}$  is uniformly calm.



*Proof* Fix  $x \in X$ , if  $f(x, \cdot)$  is convex and  $K(x)$  is convex, then  $\text{gph } M_x$  is convex. Hence, the equality in (37) follows from Theorem 2.1, given that  $M_x(0) = \Psi(x)$ . Finally, if  $L(M_x; 0, y)$  is achieved as a minimum for all  $y \in \Psi(x), x \in X$ , then we choose  $L_x(y) = L(M_x; 0, y), y \in \Psi(x), x \in X$  and it follows from (38) that the family  $\{M_x|x \in X\}$  is uniformly calm.  $\square$

A second approach to obtain a result similar to Theorem 4.6 for the family  $\{(2)|x \in X\}$  to have a uniformly weak sharp minimum is to consider but the family of multifunctions

$$M_x(z) := \{y \in K(x) | f(x, y) \leq z\}, \quad x \in X$$

instead of that in (33). Hence, in the definition of the uniform calmness of the family of multifunctions  $\{M_x|x \in X\}$ , consider but the calmness of each  $M_x$  on  $\{\varphi(x)\} \times \Psi(x)$  instead of  $\{0\} \times \Psi(x)$ . A result similar to Theorem 4.6 can then be stated and proven using [15, Lemma 4.7].

In the following example, we give an example of parametric optimization problem where each multifunction  $M_x$ , defined as in (33), is calm on  $\{0\} \times \Psi(x), x \in X$ .

*Example 4.8* We consider the parametric optimization problem

$$\min\{y|y \in K(x)\},$$

where  $K(x) := \{y|x \leq y\}$ , for all  $x \in X := \mathbb{R}$ . We have  $f(x, y) := y$ , and it is obvious that  $\Psi(x) = \{x\}$ , for all  $x \in X$ . Hence,  $\forall x \in X, \forall y \in \Psi(x)$ , we have  $\partial_y f(x, y) = 1$  and  $N_{K(x)}(y) := \mathbb{R}_-$ . Clearly,  $1 \notin \{0\} = \text{bd } \mathbb{R}_+$ . The calmness of  $M_x$  (see (33), with  $\varphi(x) := x$ ) on  $\{0\} \times \Psi(x)$  for all  $x \in X$  then follows from Theorem 4.5.

It has been shown in [1, Theorem 4.2], that the bilevel programming problem (1) is partially calm at any feasible point, provided that  $X := \mathbb{R}^n, \text{dom } \Psi = \mathbb{R}^n$ , the upper-level objective function  $F$  is Lipschitz continuous, and the functions  $f, g$  and  $h$  are linear w.r.t.  $y$ . Other sufficient conditions for the partial calmness of the bilevel programming problem have been given in [6, 8, 31].

We now present the optimality conditions for the bilevel programming problem (1) under the partial calmness and outline the difference with the optimality conditions obtained under the WMFCQ.

**Theorem 4.9** *Let  $(\bar{x}, \bar{y})$  be a local optimal solution to problem (11). Assume that  $\Psi$  be inner semicompact at  $\bar{x}$ , the lower (resp. upper) level regularity be satisfied at  $(\bar{x}, y)$ , for all  $y \in \Psi(\bar{x})$  (resp. at  $\bar{x}$ ). Moreover, we suppose that problem (11) be partially calm at  $(\bar{x}, \bar{y})$ . Then, there exist  $r > 0, \lambda, \mu, \lambda', \mu', \lambda_s, \mu_s, \eta_s \geq 0$ , and  $y_s \in \Psi(\bar{x}), s = 1, \dots, n + 1$  with  $\sum_{s=1}^{n+1} \eta_s = 1$  such that the optimality conditions in Theorem 3.5 hold.*

*Proof* Under the assumptions of inner semicompactness of  $\Psi$  at  $\bar{x}$  and lower level regularity at  $(\bar{x}, y)$ , for all  $y \in \Psi(\bar{x})$ , it follows from Theorem 3.4 that  $\varphi$  is Lipschitz

continuous around  $\bar{x}$ . Then applying Theorem 4.2 to problem (11), we have that there exists  $r > 0$  such that  $(\bar{x}, \bar{y})$  is a local optimal solution of

$$\min F(x, y) + r(f(x, y) - \varphi(x)) \quad \text{s.t. } (x, y) \in \Omega, \tag{39}$$

where  $\Omega$  is defined in (3).

Since  $F$  and  $f$  are continuously differentiable, the objective function of problem (39) is Lipschitz continuous. Hence, from Mordukhovich [16, Proposition 5.3], it follows that

$$0 \in \nabla F(\bar{x}, \bar{y}) + r\nabla f(\bar{x}, \bar{y}) + r\partial(-\varphi)(\bar{x}) \times \{0\} + N_{\Omega}(\bar{x}, \bar{y}).$$

From here the rest of the proof follows as that of Theorem 3.5. □

What holds our attention here is the nature of the multiplier  $r$ , which is positive in Theorem 4.9 and simply nonnegative in Theorem 3.5. Though surprising, given that the partial calmness is weaker than the WMFCQ, as we show in the next result, this is rather understandable considering the difference in the two processes. Another important point here is that, the convexity of set  $\Omega$  is not necessary in the latter result.

**Theorem 4.10** *Consider problem (3) and assume that,  $\Omega$  be convex and  $\mathcal{G}$  be Lipschitz continuous. Then, the following implications hold true at a local optimal solution of problem (3).*

$$MMFCQ \implies WMFCQ \implies \text{partial calmness.}$$

*Proof* Since the proof of the first implication is straightforward (see Sect. 3), we focus our attention on the second implication.

Under the convexity of  $\Omega$  and the Lipschitz continuity of  $\mathcal{G}$ , the WMFCQ is well-defined to be a CQ at a given point  $(\bar{x}, \bar{y})$ , as we can observe from Lemma 3.3. Moreover, it follows from [14] that the set-valued mapping

$$M(t) = \{(x, y) \in \Omega \mid \mathcal{G}(x, y) \leq t\}$$

is calm at  $(0, \bar{x}, \bar{y})$ . Also observe that the feasible set of (3) is  $C = M(0)$ . Hence, it follows from [7] that, if  $(\bar{x}, \bar{y})$  is a local optimal solution of (3), then there exists  $\alpha > 0$  and a neighborhood  $W$  of  $(\bar{x}, \bar{y})$ , such that

$$F(\bar{x}, \bar{y}) \leq F(x, y) + \alpha d((x, y), M(0)), \quad \forall (x, y) \in W. \tag{40}$$

On the other hand, the calmness of  $M$  at  $(0, \bar{x}, \bar{y})$  implies that there exist neighborhoods  $U$  of 0,  $V$  of  $(\bar{x}, \bar{y})$  and a constant  $L > 0$  such that

$$d((x, y), M(0)) \leq L|t|, \quad \forall t \in U, \forall (x, y) \in M(t) \cap V. \tag{41}$$

By combining (40) and (41), it follows that

$$F(x, y) - F(\bar{x}, \bar{y}) + \alpha L|t| \geq 0, \quad \forall t \in U, \forall (x, y) \in W \cap V \cap M(t),$$

which coincides with the definition of partial calmness in Definition 4.1. □

In the next example, we show that the converse of the second implication of this result is not always possible.

*Example 4.11* Consider the bilevel program

$$\min x^2 + y^2 \quad \text{s.t. } x \in \mathbb{R}, y \in \operatorname{argmin}\{x^2y + y \mid y \geq 0\}. \tag{42}$$

Set  $f(x, y) := x^2y + y$  and  $\Omega := \mathbb{R} \times \mathbb{R}_+$ . We have  $\varphi(x) = 0, \forall x \in X := \mathbb{R}$ , and  $\mathcal{G}(x, y) = x^2y + y$ . We can easily check that  $(0, 0)$  is an optimal solution of problem (42). Hence  $\partial\mathcal{G}(0, 0) = (0, 1)$  and  $N_\Omega(0, 0) = \{0\} \times \mathbb{R}_-$ . It follows that  $(0, 1) \in \{0\} \times \mathbb{R}_+ = -\operatorname{bd} N_\Omega(0, 0)$ . This means that WMFCQ fails at  $(0, 0)$ .

On the other hand, it follows from [1, Theorem 4.2] that problem (42) is partially calm at  $(0, 0)$ .

### 5 Optimality Conditions Under Strong Stability

We consider the bilevel programming problem (1). Let a feasible point  $(\bar{x}, \bar{y})$  be such that, the vector  $\bar{y} \in \Psi(\bar{x})$  is a strongly stable local optimal solution of the parametric optimization problem (2), in the sense of Kojima [34]; i.e. there exist open neighborhoods  $U(\bar{x})$  of  $\bar{x}$ , and  $V(\bar{y})$  of  $\bar{y}$ , and a uniquely determined continuous vector-valued function  $y(\cdot) : U(\bar{x}) \rightarrow V(\bar{y})$  such that,  $y(x)$  is the unique local optimal solution of problem (2) in  $V(\bar{y})$ , for all  $x \in U(\bar{x})$ . Then, the optimal value function takes the form

$$\varphi(x) = f(x, y(x)), \quad \forall x \in U(\bar{x}). \tag{43}$$

In this case, the mapping  $\Psi$  is obviously inner semicompact at  $\bar{x}$ , given that it is uniformly bounded around  $\bar{x}$ , from the definition we have in Sect. 2. In contrast to what we had in the previous sections, the estimation of the subdifferential of the optimal value function  $\varphi$  at the point  $\bar{x}$ , could face some difficulties, since the function  $y(\cdot) : U(\bar{x}) \rightarrow V(\bar{y})$  may not be Lipschitz continuous [35], even when the lower level regularity is achieved. Before tackling this problem, we give a result ensuring the strong stability in the sense of Kojima. For this, we assume for the rest of the section that the functions  $f, g$  and  $h$  are twice continuously differentiable; hence we can mention the following *strong sufficient condition of second order* (SSOC) which is needed:

SSOC is said to hold at  $(\bar{x}, \bar{y})$  if, for each  $(\lambda, \mu) \in \Lambda(\bar{x}, \bar{y})$  and for each  $d \neq 0$  satisfying

$$\begin{aligned} \nabla_y g_i(\bar{x}, \bar{y})d &= 0, \quad \text{for each } i \in J := \{j \mid \lambda_j > 0\} \\ \nabla_y h_j(\bar{x}, \bar{y})d &= 0, \quad \text{for } j = 1, \dots, q, \end{aligned}$$

we have

$$d^\top \nabla_{yy}^2 L(\bar{x}, \bar{y}, \lambda, \mu)d > 0;$$

where  $L(x, y, \lambda, \mu) := f(x, y) + \lambda^\top g(x, y) + \mu^\top h(x, y)$  is the Lagrangian function of problem (2). We can then state this theorem from Kojima [34].

**Theorem 5.1** *Let  $\bar{y} \in \Psi(\bar{x})$  be a local optimal solution of problem (2). Assume that the SSOC and the lower level regularity hold at  $(\bar{x}, \bar{y})$ , then the local optimal solution  $\bar{y}$  is strongly stable.*

In order to obtain the Lipschitz continuity of  $y(\cdot)$ , and hence to be able to compute the subdifferential of  $\varphi$ , the following *constant rank constraint qualification (CRCQ)* is crucial, as was pointed out by Ralph and Dempe [35]:

CRCQ is said to be satisfied for problem (2) at  $(\bar{x}, \bar{y})$  if, there exists an open neighborhood  $W$  of  $(\bar{x}, \bar{y})$  such that, for subsets

$$I \subseteq I(\bar{x}, \bar{y}) := \{i | g_i(\bar{x}, \bar{y}) = 0\}, \quad J \subseteq \{1, \dots, q\},$$

the family of gradient vectors  $\{\nabla_y g_i(x, y) | i \in I\} \cup \{\nabla_y h_j(x, y) | j \in J\}$ , has the same rank, for all  $(x, y) \in W$ .

We are now ready to state a result analogous to Theorem 3.4.

**Lemma 5.2** *Consider a local optimal solution  $\bar{y} \in \Psi(\bar{x})$  of problem (2). Assume that the lower level regularity, the SSOC and the CRCQ hold at  $(\bar{x}, \bar{y})$ . Then, it follows from Theorem 5.1 that,  $\bar{y}$  is strongly stable. Moreover, there exists a neighborhood  $U(\bar{x})$  of  $\bar{x}$  such that, the optimal value function defined by (43) is Lipschitz continuous around  $\bar{x}$  and*

$$\partial\varphi(\bar{x}) = \nabla_x f(\bar{x}, \bar{y}) + D^*y(\bar{x})(\nabla_y f(\bar{x}, \bar{y})). \tag{44}$$

*Proof* From Theorem 5.1, it follows that the local optimal solution  $\bar{y}$ , of the lower level problem (2), is strongly stable. Furthermore, from [2, Theorem 4.10],  $y(\cdot)$  is Lipschitz continuous around  $\bar{x}$  provided that, CRCQ is satisfied. Hence, equality (44) follows from the chain rule in [16, Theorem 1.110].  $\square$

In the next result, we give KKT type optimality conditions for the optimal value reformulation (3), of problem (1) under the strong stability of the parametric lower level problem, in the sense of Kojima.

**Theorem 5.3** *Let  $(\bar{x}, \bar{y})$  be a local optimal solution to problem (11). Assume that the SSOC, CRCQ, and lower (resp. upper) level regularity are satisfied at  $(\bar{x}, \bar{y})$  (resp.  $\bar{x}$ ). Moreover, we suppose that  $\Omega$  be convex and the WMFCQ holds true at  $(\bar{x}, \bar{y})$ . Then, there exist  $r \geq 0$  and  $\lambda, \mu, \lambda', \mu'$  such that:*

$$\begin{aligned} 0 &\in \nabla_x F(\bar{x}, \bar{y}) - r \partial y(\bar{x})^\top \nabla_y f(\bar{x}, \bar{y}) + \nabla_x g(\bar{x}, \bar{y})^\top \lambda + \nabla_x h(\bar{x}, \bar{y})^\top \mu \\ &\quad + \nabla G(\bar{x})^\top \lambda' + \nabla H(\bar{x})^\top \mu', \\ \nabla_y F(\bar{x}, \bar{y}) + r \nabla_y f(\bar{x}, \bar{y}) + \nabla_y g(\bar{x}, \bar{y})^\top \lambda + \nabla_y h(\bar{x}, \bar{y})^\top \mu &= 0, \\ \lambda &\geq 0, \lambda^\top g(\bar{x}, \bar{y}) = 0, \\ \lambda' &\geq 0, \lambda'^\top G(\bar{x}) = 0, \end{aligned}$$

where  $\partial y(\bar{x})$  denotes the Clarke generalized Jacobian of the function  $y(\cdot)$ .

*Proof* The proof is identical to that of Theorem 3.5, at the exception that in (23), one has to notice the following changes: Firstly, it follows from Lemma 5.2 that

$$\text{co } \partial \varphi(\bar{x}) = \nabla_x f(\bar{x}, \bar{y}) + \text{co } D^* y(\bar{x})(\nabla_y f(\bar{x}, \bar{y}))$$

and secondly, from Mordukhovich [36] one has

$$\text{co } D^* y(\bar{x})(\nabla_y f(\bar{x}, \bar{y})) = \partial y(\bar{x})^\top \nabla_y f(\bar{x}, \bar{y}). \quad \square$$

For the computation of the generalized Jacobian of the function  $y(\cdot)$ , the interested reader is referred to Dempe and Vogel [37]. Optimality conditions for problem (1), under strong stability, can also be derived by inserting the solution function  $y(\cdot)$  in the upper level objective function. Thus, the new problem to be solved is:

$$\min F(x, y(x)) \quad \text{s.t. } x \in X.$$

For more on this direction, the reader is referred to Dempe [2].

## 6 Conclusion

After observing that the Mangasarian-Fromowitz constraint qualification in terms of the basic generalized differentiation constructions of Mordukhovich (MMFCQ) is weaker than the MFCQ in terms of Clarke's tools, we have shown that the former CQ fails under a mild assumption. Knowing that the nonsmooth MFCQ in terms of Clarke is weaker than the nonsmooth linear independence CQ and the nonsmooth Slater CQ [6], we can conclude that under the sum rule (13), the latter CQs fail for problem (3). Using a weakened form of the MMFCQ (i.e. the WMFCQ), KKT type optimality conditions have been derived for problem (3). Other closely related CQs have also been suggested (see Sect. 3), which seem quite interesting and easier to check in the case of a simple convex bilevel programming problem.

In this work, we have also given a new sufficient condition for the partial calmness of problem (3) based on the new concept of *uniform calmness* of a certain family of set-valued mappings, but the difficulty here is how to choose the desired family of calmness constants. The first step in this direction is given in Theorem 4.7.

## References

1. Dempe, S., Zemkoho, A.B.: The bilevel programming problem: reformulations, constraint qualifications and optimality conditions. Submitted for publication
2. Dempe, S.: Foundations of Bilevel Programming. Kluwer Academic, Dordrecht (2002)
3. Dempe, S., Dutta, J., Lohse, S.: Optimality conditions for bilevel programming problems. *Optimization* **55**, 505–524 (2006)
4. Dempe, S., Kalashnikov, V.V., Kalashnykova, N.: Optimality conditions for bilevel programming problems. In: Dempe, S., Kalashnikov, V.V. (eds.) *Optimization with Multivalued Mappings*, vol. 2, pp. 3–28. Springer, New York (2006)
5. Dutta, J., Dempe, S.: Bilevel programming with convex lower level problems. In: Dempe, S., Kalashnikov, V.V. (eds.) *Optimization with Multivalued Mappings*, vol. 2, pp. 51–71. Springer, New York (2006)

6. Ye, J.J., Zhu, D.L.: Optimality conditions for bilevel programming problems. *Optimization* **33**, 9–27 (1995)
7. Clarke, F.H.: *Optimization and Nonsmooth Analysis*. SIAM Classics in Applied Mathematics, vol. 5. Wiley, New York (1983). Reprint, Philadelphia (1994)
8. Ye, J.J., Zhu, D.L.: A note on optimality conditions for bilevel programming problems. *Optimization* **39**, 361–366 (1997)
9. Ye, J.J.: Nondifferentiable multiplier rules for optimization and bilevel optimization problems. *SIAM J. Optim.* **15**, 252–274 (2004)
10. Ye, J.J.: Constraint qualifications and KKT conditions for bilevel programming problems. *Math. Oper. Res.* **31**, 811–824 (2006)
11. Babahadda, H., Gadhi, N.: Necessary optimality conditions for bilevel optimization problems using convexificators. *J. Glob. Optim.* **34**, 535–549 (2006)
12. Amahroq, T., Gadhi, N.: On the regularity condition for vector programming problems. *J. Glob. Optim.* **21**, 433–441 (2001)
13. Dempe, S., Dutta, J., Mordukhovich, B.S.: New necessary optimality conditions in optimistic bilevel programming. *Optimization* **56**, 577–604 (2007)
14. Henrion, R., Outrata, J.: A subdifferential condition for calmness of multifunctions. *J. Math. Anal. Appl.* **258**, 110–130 (2001)
15. Henrion, R., Jourani, A., Outrata, J.: On the calmness of a class of multifunctions. *SIAM J. Optim.* **13**, 603–618 (2002)
16. Mordukhovich, B.S.: *Variational Analysis and Generalized Differentiation. I: Basic Theory, II: Applications*. Springer, Berlin (2006)
17. Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Springer, Berlin (1998)
18. Zheng, X.Y., Ng, K.F.: Calmness for L-subsmooth multifunctions in Banach spaces. *SIAM J. Optim.* **19**, 1648–1673 (2009)
19. Mangasarian, O.L., Fromowitz, S.: The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.* **17**, 37–47 (1967)
20. Gauvin, J., Dubeau, F.: Differential properties of the marginal function in mathematical programming. *Math. Program. Stud.* **18**, 101–119 (1982)
21. Mordukhovich, B.S., Nam, N.M.: Variational stability and marginal functions via generalized differentiation. *Math. Oper. Res.* **30**, 800–816 (2005)
22. Klatté, D., Kummer, B.: Stability properties of infima and optimal solutions of parametric optimization problems. *Nondifferentiable optimization: motivations and applications*. In: Proc. IIASA Workshop, Sopron/Hung, 1984. Lect. Notes Econ. Math. Syst, vol. 255, pp. 215–229. Springer, Berlin (1985)
23. Rockafellar, R.T.: Lagrange multipliers and optimality. *SIAM Rev.* **35**, 183–238 (1993)
24. Mangasarian, O.L.: *Nonlinear Programming*. SIAM Classics in Applied Mathematics, vol. 10. McGraw-Hill, New York (1969). Reprint, Philadelphia (1994)
25. Dempe, S., Dinh, N., Dutta, J.: Optimality conditions for a simple convex bilevel programming problem. In: Burachik, R.S., Yao, J.-C. (eds.) *Variational Analysis and Generalized Differentiation in Optimization and Control*. Springer, Berlin (2010)
26. Dempe, S., Zemkoho, A.B.: On the Karush-Kuhn-Tucker reformulation of the bilevel optimization problem. Submitted for publication
27. Henrion, R., Outrata, J., Surowiec, T.: On the co-derivative of normal cone mappings to inequality systems. *Nonlinear Anal.* **71**, 1213–1226 (2009)
28. Moldovan, A., Pellegrini, L.: On regularity for constrained extremum problems. I: Sufficient optimality conditions. *J. Optim. Theory Appl.* **142**, 147–163 (2009)
29. Moldovan, A., Pellegrini, L.: On regularity for constrained extremum problems. II: Necessary optimality conditions. *J. Optim. Theory Appl.* **142**, 165–183 (2009)
30. Dempe, S., Zemkoho, A.B.: Bilevel road pricing: Theoretical analysis and optimality conditions. Submitted for publication
31. Ye, J.J.: New uniform parametric error bounds. *J. Optim. Theory Appl.* **98**, 197–219 (1998)
32. Ye, J.J., Zhu, D.L., Zhu, Q.J.: Exact penalization and necessary optimality conditions for generalized bilevel programming problems. *SIAM J. Optim.* **7**, 481–507 (1997)
33. Heerda, J., Kummer, B.: Characterization of calmness for Banach space mappings. Preprint, Humboldt-Universität zu Berlin (2006)
34. Kojima, M.: Strongly stable stationary solutions in nonlinear programs. Analysis and computation of fixed points. *Proc. Symp., Univ. Wis.* **1979**, 93–138 (1980)

35. Ralph, D., Dempe, S.: Directional derivatives of the solution of a parametric nonlinear program. *Math. Program.* **70**, 159–172 (1995)
36. Mordukhovich, B.S.: Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.* **183**, 250–288 (1994)
37. Dempe, S., Vogel, S.: The generalized Jacobian of the optimal solution in parametric optimization. *Optimization* **50**, 387–405 (2001)