SIAM J. OPTIM.
Vol. 15, No. 1, pp. 252–274

NONDIFFERENTIABLE MULTIPLIER RULES FOR OPTIMIZATION AND BILEVEL OPTIMIZATION PROBLEMS*

JANE J. YE†

Abstract. In this paper we study optimization problems with equality and inequality constraints on a Banach space where the objective function and the binding constraints are either differentiable at the optimal solution or Lipschitz near the optimal solution. Necessary and sufficient optimality conditions and constraint qualifications in terms of the Michel–Penot subdifferential are given, and the results are applied to bilevel optimization problems.

Key words. necessary optimality conditions, sufficient optimality conditions, constraint qualifications, bilevel optimization problems, Michel–Penot subdifferentials

AMS subject classifications. 49K10, 90C30, 91A65

DOI. 10.1137/S1052623403424193

1. Introduction.

In this paper we study Lagrange multiplier rules and constraint qualifications (CQs) for the following optimization problem with equality and inequality constraints:

$$\text{(P)} \qquad \begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \le 0, \quad i = 1, 2, \ldots, I, \\ & h_j(x) = 0, \quad j = 1, 2, \ldots, J, \end{array}$$

where $f, g_i (i = 1, 2, \ldots, I), h_j (j = 1, 2, \ldots, J)$ are functions from a Banach space X to R and I, J are given integers. Generally one has $I \ge 1, J \ge 1$, but we allow I or $J = 0$ to signify the case in which there are no explicit constraints of the type. For any feasible solution $\bar{x}$ of problem (P), we denote by $I(\bar{x}) := \{i : g_i(\bar{x}) = 0\}$ the index set of the binding constraints.

The classical Lagrange multiplier rule (see, e.g., [4, 16]) usually requires that the objective function and the inequality constraints be Fréchet differentiable and the equality constraints be continuously differentiable. Most extensions of the classical Lagrange multipliers are given under two different assumptions: differentiability and Lipschitz continuity. On one hand, the classical multiplier rule was extended in the direction of eliminating the smoothness assumption while keeping the differentiability assumption such as in Halkin [9]. On the other hand, the classical multiplier rule was generalized in the direction of replacing the usual gradient by certain generalized gradients under Lipschitz assumptions such as in Rockafellar [22], Clarke [7], Michel and Penot [17, 18], Ioffe [11, 12], Mordukhovich [19], and Treiman [23, 24].

It is known that differentiability and Lipschitz continuity are two different kinds of assumptions and may not imply each other in general. Hence for nonlinear programming problems with mixed assumptions of differentiability and Lipschitz continuity, the only applicable optimality conditions in the literature were fuzzy multiplier

*Received by the editors May 12, 2003; accepted for publication (in revised form) March 25, 2004; published electronically December 9, 2004. The research of this paper is partially supported by NSERC and a University of Victoria internal research grant.
http://www.siam.org/journals/siopt/15-1/42419.html
†Department of Mathematics and Statistics, University of Victoria, P.O. Box 3045 STN CSC, Victoria, BC, V8W 3P4 Canada (janeye@math.uvic.ca).

rules for optimization problem with lower semicontinuous data (see, e.g., Borwein, Treiman, and Zhu [6] and Ngai and Théra [20]). Although in a finite-dimensional space the fuzzy multiplier rule reduces to an exact multiplier rule, it involves the singular subdifferential of the non-Lipschitz functions. Another issue involved is the size of the subdifferential. It is known that the Clarke generalized gradient of a differentiable function which is not strictly differentiable may contain other elements which are not the usual derivative. Our purpose is to provide an exact (not fuzzy) multiplier rule where the usual derivative (not the generalized gradient even if it is also Lipschitz continuous) is used when a function is differentiable and the generalized gradient is used when a function is not differentiable but Lipschitz continuous. Among various convex-valued generalized gradients which coincide with the usual derivative when a function is Gâteaux differentiable, including the B-generalized gradient of Treiman [23], the Michel–Penot (M-P) subdifferential is the smallest one, and hence we aim to provide a multiplier rule in terms of the M-P subdifferential. The multiplier rules in terms of other bigger generalized gradients follow immediately.

In Ye [27], under the mixed assumptions of Fréchet differentiability and Lipschitz continuity, and Fritz John and KKT Lagrange multiplier rules under generalized Mangasarian–Fromovitz, metric regularity and calmness CQs were given where the usual derivative is used when a function is differentiable.

In this paper we continue the study by considering the problem with mixed assumptions of Gâteaux differentiability, Fréchet differentiability, Hadamard differentiability (see, e.g., Definition 2.1), and Lipschitz continuity under other CQs that were not considered in [27]. Our main result includes the following generalized Lagrange multiplier rule, which summarizes the results obtained in Theorem 3.1 and Propositions 3.1–3.7.

THEOREM 1.1 (nondifferentiable KKT necessary optimality condition). *Let $\bar{x}$ be a local optimal solution of* (P). *Consider the following CQs at $\bar{x}$:*

(1) *the nondifferentiable weak reverse convex CQ as in Definition* 3.12;
(2) *the nondifferentiable weak Slater CQ as in Definition* 3.11;
(3) *the nondifferentiable Arrow–Hurwicz–Uzawa CQ as in Definition* 3.10;
(4) *the generalized Zangwill CQ as in Definition* 3.4;
(5) *the nondifferentiable linear independence CQ as in Definition* 3.8;
(6) *the nondifferentiable Slater CQ as in Definition* 3.9;
(7) *the nondifferentiable Mangasarian–Fromovitz CQ as in Definition* 3.7;
(8) *the nondifferentiable Kuhn–Tucker CQ as in Definition* 3.6;
(9) *the nondifferentiable Abadie CQ as in Definition* 3.3.

If f is either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, then the KKT condition in terms of the M-P subdifferential holds at $\bar{x}$ under one of the CQs (1)–(4). *If f is either Fréchet differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, then the KKT condition in terms of the M-P subdifferential holds at $\bar{x}$ under one of the CQs* (1)–(9). *That is, there exist scalars $\alpha_i \geq 0$ $(i \in I(\bar{x})), \beta_j \geq 0, \gamma_j \geq 0$ $(j = 1, 2, \ldots, J)$ such that*

$$0 \in \partial^\diamond f(\bar{x}) + \sum_{i \in I(\bar{x})} \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \partial^\diamond h_j(\bar{x}) - \sum_{j=1}^{J} \gamma_j \partial^\diamond h_j(\bar{x}),$$

where $\partial^\diamond$ denotes the M-P subdifferential.

The relationships between the various constraint qualifications are given in the following diagram:

$$
\begin{array}{ccccc}
 & & \text{WRC (1)} & & \text{weak Slater (2)} \\
 & & \Downarrow & & \Downarrow \\
\text{LICQ (5)} & \text{Slater (6)} & & \text{AHU (3)} & \\
\Downarrow & \Downarrow & & \Downarrow & \\
 & \text{MFCQ (7)} & & \text{Zangwill (4)} & \\
 & \Downarrow & & \Downarrow & \\
 & & \text{Kuhn–Tucker (8)} & & \\
 & & \Downarrow & & \\
 & & \text{Abadie (9)} & &
\end{array}
$$

Note that Theorem 1.1 under CQ (7) was given in [27, Theorem 4.1]. The above KKT condition, however, provides a nondifferentiable KKT condition under all CQs (1)–(9). Moreover, the relationships between various CQs are given. Over the years, many papers have been devoted to extensions of classical CQs of type (5)–(7) to nonsmooth optimization problems (see, e.g., [10, 14, 31]). To the best of the author's knowledge, CQs of type (1)–(4), (8)–(9) have never been extended to allow nondifferentiability in the literature. One of the purposes of this paper is to fill this gap since these nondifferentiable CQs are needed for studies of bilevel programming problems.

In Theorem 3.2 we also prove that the above KKT condition in terms of the M-P subdifferential becomes sufficient when the objective function is M-P pseudoconvex, the inequality constraints are M-P regular and quasiconvex, and the equality constraints are Gâteaux differentiable and quasiaffine at the optimal solution $\bar{x}$.

In the last section of this paper we apply the results obtained to the bilevel optimization problem. One may reformulate the bilevel optimization as a single level optimization problem by using either the value function or the KKT condition for the lower level problem. The difficulty is that the usual CQ, such as the linear independence CQ, the Slater CQ, and the Mangasarian–Fromovitz CQ, does not hold for such a single level optimization problem. In this paper we show that the rest of the CQs (1), (3), (4), and (8)–(9) may hold for bilevel optimization problems. In particular, no CQ is required for the generalized linear bilevel optimization problem, which generalizes the known result that no CQ is needed for the linear bilevel programming problem. When the lower level problem is convex the relationship between the multiplier rule for the single level formulation by the value function approach and the one by the KKT approach is compared. It is found that the multiplier rule for the single level formulation by the value function approach is sharper than the one by the KKT approach.

We organize the paper as follows. In the next section, we provide preliminaries and preliminary results to be used in the rest of the paper. Section 3 is devoted to the discussion of CQs and the KKT necessary and sufficient optimality conditions. In section 4, applications to the bilevel optimization problem are given. In this paper unless otherwise specified, we denote by X a Banach space and by X^* its dual space equipped with the weak-star topology w^*. For $A \subseteq X$, we denote by co$A$, cl$A$ its convex hull and its closure, respectively. We denote by $B(v, \delta)$ the open ball centered at $v \in X$ with radius $\delta > 0$.

2. Preliminaries and preliminary results. We first recall some definitions of the usual derivatives.

Definition 2.1 (usual derivatives). *Let X, Y be Banach spaces, let $\bar{x} \in X$, and let $f : X \to Y$. The usual directional derivative of f at $\bar{x}$ in the direction $v \in X$ is* *given by*

$$f'(\bar{x}; v) := \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

when this limit exists. f is said to be Gâteaux differentiable if there exists $Df(\bar{x})$, an element of the space $L(X, Y)$ of continuous linear functionals from X to Y such that for every $v \in X$, $f'(\bar{x}; v) = \langle Df(\bar{x}), v \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the canonic pairing. f is said to be Hadamard differentiable at $\bar{x}$ if $Df(\bar{x}) \in L(X, Y)$ and, for every $v \in X$,

$$\lim_{t \downarrow 0, v' \to v} \frac{f(\bar{x} + tv') - f(\bar{x})}{t} = \langle Df(\bar{x}), v \rangle.$$

f is said to be Fréchet differentiable at $\bar{x}$ if $Df(\bar{x}) \in L(X, Y)$ and the convergence in

$$f'(\bar{x}; v) := \lim_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t} = \langle Df(\bar{x}), v \rangle$$

is uniform with respect to v in bounded sets.

Remark 2.1. It is clear from the above definition that Fréchet differentiability is stronger than Hadamard differentiability, which in turn is stronger than Gâteaux differentiability.

DEFINITION 2.2 (M-P subdifferential). *Let $\bar{x} \in X$ and let $f : X \to R$ be any function. The M-P directional derivative of f at $\bar{x}$ in the direction $v \in X$ introduced in* [17] *is given by*

$$f^{\diamond}(\bar{x}; v) := \sup_{w \in X} \limsup_{t \downarrow 0} \frac{f(\bar{x} + t(v + w)) - f(\bar{x} + tw)}{t},$$

and the M-P subdifferential of f at $\bar{x}$ is given by the set

$$\partial^{\diamond} f(\bar{x}) := \{x^* \in X^* : \langle x^*, v \rangle \leq f^{\diamond}(\bar{x}; v) \quad \forall v \in X\}.$$

The M-P subdifferential is a natural generalization of the Gâteaux derivative since it is known (see [17, Proposition 1.3]) that when a function f is Gâteaux differentiable at $\bar{x}$, $f^{\diamond}(\bar{x}; v) = f'(\bar{x}; v)$ and $\partial^{\diamond} f(\bar{x}) = \{Df(\bar{x})\}$. Moreover when a function f is convex, the M-P subdifferential coincides with the subdifferential in the sense of convex analysis.

Whenever the Clarke generalized directional derivative $f^{\circ}(\bar{x}; v)$ and the Clarke generalized gradient $\partial^{\circ} f(\bar{x})$ exist, one always has

$$f^{\diamond}(\bar{x}; v) \leq f^{\circ}(\bar{x}; v), \qquad \partial^{\diamond} f(\bar{x}) \subseteq \partial^{\circ} f(\bar{x}).$$

Note that the above inequality and the inclusion may be strict even in the case when f is Lipschitz continuous. For example, the function

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0 \end{cases} \tag{1}$$

on R is Lipschitz near 0 and Fréchet differentiable at 0 with $Df(0) = 0$, and hence $\partial^{\diamond} f(0) = \{0\}$ and $f^{\diamond}(0; v) = f'(0, v) = 0$. However, $\partial^{\circ} f(0) = [-1, 1]$ and $f^{\circ}(0; v) = |v|$.

Similar to the Clarke regularity [7], the following regularity concept was introduced in [5] as semiregularity (see also [25]) for Lipschitz continuous functions. We now extend the definition to any functions so that a Gâteaux differentiable function is also M-P regular.

DEFINITION 2.3 (M-P regularity). *Let $f : X \to R$ be a function on X and let $\bar{x} \in X$. We say that f is M-P regular at $\bar{x}$ if the usual directional derivative $f'(\bar{x}; v)$ exists and $f'(\bar{x}; v) = f^\diamond(\bar{x}; v)$ for all $v \in X$.*

The following properties of the M-P directional derivative and the M-P subdifferential will be useful.

PROPOSITION 2.1 (see [17, 18, 5]). *Let X be a Banach space, let $\bar{x} \in X$, and let $f, g : X \to R$ be either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$. Then the following hold:*

(i) *The function $v \to f^\diamond(\bar{x}; v)$ is finite, positively homogeneous, and subadditive on X.*

(ii) *For any scalar λ, $\partial^\diamond(\lambda f)(\bar{x}) = \lambda \partial^\diamond f(\bar{x})$, and for every $v \in X, f^\diamond(\bar{x}; -v) = (-f)^\diamond(\bar{x}; v)$.*

(iii) *$\partial^\diamond(f+g)(\bar{x}) \subseteq \partial^\diamond f(\bar{x}) + \partial^\diamond g(\bar{x})$ and $(f+g)^\diamond(\bar{x}; v) \le f^\diamond(\bar{x}; v) + g^\diamond(\bar{x}; v)$ for all $v \in X$. The equalities hold if both f and g are M-P regular at $\bar{x}$.*

(iv) *$\partial^\diamond f(\bar{x})$ is a nonempty, convex, weak*-compact subset of X^*, and for every v in X, one has $f^\diamond(\bar{x}; v) = \max\{\langle \xi^*, v\rangle : \xi^* \in \partial^\diamond f(\bar{x})\}$.*

(v) *If $\bar{x}$ is a local minimum of f, then $0 \in \partial^\diamond f(\bar{x})$ and $f^\diamond(\bar{x}; v) \ge 0$ for all $v \in X$.*

In [26, Proposition 3.1], it was shown that a Lipschitz function f is strictly differentiable if and only if both f and $-f$ are Clarke regular. Similarly we have the following conclusion.

PROPOSITION 2.2. *Let $f : X \to R$ be a function which is Lipschitz near $\bar{x} \in X$. Then f and $-f$ are both M-P regular if and only if f is Gâteaux differentiable at $\bar{x}$.*

Proof. It is obvious that if f is Gâteaux differentiable at $\bar{x}$, then both f and $-f$ are M-P regular. Now suppose that both f and $-f$ are M-P regular; then by (iii) of Proposition 2.1, one has

$$\partial^\diamond f(\bar{x}) + \partial^\diamond(-f)(\bar{x}) = \partial^\diamond(f-f)(\bar{x}) = \{0\},$$

which implies that $\partial^\diamond f(\bar{x})$ is a singleton since both $\partial^\diamond f(\bar{x})$ and $\partial^\diamond(-f)(\bar{x})$ are nonempty. Let $\partial^\diamond f(\bar{x}) = \{\xi\}$. Since $\xi \in X^*$, to prove that f is Gâteaux differentiable at $\bar{x}$ it suffices to prove that $f^\diamond(\bar{x}; v) = \langle \xi, v\rangle$ for each $v \in X$. By (iv) of Proposition 2.1, for each $v \in X$, $f^\diamond(\bar{x}; v) \ge \langle \xi, v\rangle$. By the Hahn–Banach theorem there exists $\xi' \in X^*$ majorized by $f^\diamond(\bar{x}; \cdot)$ and agreeing with $f^\diamond(\bar{x}; \cdot)$ at v. It follows that $\xi' \in \partial^\diamond f(\bar{x})$, and we have $f^\diamond(\bar{x}; v) = \langle \xi', v\rangle \ge \langle \xi, v\rangle$. If $\langle \xi, v\rangle$ were less than $f^\diamond(\bar{x}; v)$, then $\xi \ne \xi'$, contrary to the fact that $\partial^\diamond f(\bar{x}) = \{\xi\}$, and hence f is Gâteaux differentiable at $\bar{x}$. ☐

Based on the M-P subdifferential, we extend the notions of pseudoconvexity and pseudoconcavity to allow nondifferentiability. For a definition of this kind of generalization to a class of generalized gradients, we refer the reader to [21].

DEFINITION 2.4 (M-P pseudoconvexity and pseudoconcavity). *Let f be a function defined on a Banach space X. f is said to be M-P pseudoconvex at $\bar{x} \in X$ if for all $x \in X$,*

$$f^\diamond(\bar{x}; x - \bar{x}) \ge 0 \ \Rightarrow \ f(x) \ge f(\bar{x}).$$

f is said to be M-P pseudoconcave at $\bar{x} \in X$ if for all $x \in X$,

$$f^\diamond(\bar{x}; x - \bar{x}) \le 0 \ \Rightarrow \ f(x) \le f(\bar{x}).$$

f is said to be M-P pseudoconvex (pseudoconcave) if it is M-P pseudoconvex (pseudoconcave) at all $x \in X$.

f is said to be M-P pseudoaffine if it is both M-P pseudoconvex and M-P pseudoconcave.

Remark 2.2. It is obvious that if f is Gâteaux differentiable at $\bar{x}$, then f is M-P pseudoconvex at $\bar{x}$ if and only if $-f$ is M-P pseudoconcave at $\bar{x}$. Using Proposition 2.1 it is easy to show that if f is Lipschitz near $\bar{x}$ and M-P pseudoconvex at $\bar{x}$, then $-f$ is M-P pseudoconcave at $\bar{x}$. However, the definitions for M-P pseudoconvexity and pseudoconcavity for a nondifferentiable function are not symmetric since M-P pseudoconcavity of f at x may not imply M-P pseudoconvexity of $-f$ at $\bar{x}$. For example, $\|x\|$ is both M-P pseudoconvex and pseudoconcave and hence M-P pseudoaffine, but $-\|x\|$ is M-P pseudoconcave but not M-P pseudoconvex.

As in the differentiable case we have the following necessary and sufficient optimality condition under the M-P pseudoconvexity.

THEOREM 2.1. *Let $\bar{x} \in X$ and f be M-P pseudoconvex at $\bar{x}$. Then $\bar{x}$ is a global minimum of the function $f(x)$ if and only if $f^{\diamond}(\bar{x}; x-\bar{x}) \geq 0$ for all $x \in X$, i.e., $0 \in \partial^{\diamond} f(\bar{x})$.*

Proof. Assume that $\bar{x}$ is a global minimum of the function $f(x)$; then for any $t \in (0,1)$ one has

$$\frac{f(\bar{x}+t(x-\bar{x}))-f(\bar{x})}{t} \geq 0,$$

and hence

$$f^{\diamond}(\bar{x}; x-\bar{x}) \geq 0 \quad \forall x \in X.$$

Conversely if the above inequality holds, then by definition of M-P pseudoconvexity, one has $f(x) \geq f(\bar{x})$ and the proof is complete. □

We now recall the definition for strictly quasiconvex (also referred to as semistrictly quasiconvex) functions and quasiconvex functions.

DEFINITION 2.5 (quasiconvexity and strict quasiconvexity). *Let f be a function defined on a Banach space X. f is said to be quasiconvex at $\bar{x} \in X$ if for all $x \in X$,*

$$f(x) \leq f(\bar{x}), 0 < \lambda < 1 \Rightarrow f((1-\lambda)\bar{x}+\lambda x) \leq f(\bar{x}).$$

f is said to be strictly quasiconvex at $\bar{x} \in X$ if for all $x \in X$,

$$f(x) < f(\bar{x}), 0 < \lambda < 1 \Rightarrow f((1-\lambda)\bar{x}+\lambda x) < f(\bar{x}).$$

f is said to be quasiconcave (strictly quasiconcave) at $\bar{x}$ if $-f$ is quasiconvex (strictly quasiconvex) at $\bar{x}$.

f is said to be (strictly) quasiconvex (quasiconcave) if it is (strictly) quasiconvex (quasiconcave) at all $x \in X$.

f is said to be quasiaffine if it is both quasiconvex and quasiconcave.

We relate M-P pseudoconvex functions to strictly quasiconvex functions and quasiconvex functions in the following proposition, which can be proved similarly to the proof of [16, Theorem 9.5].

PROPOSITION 2.3. *Let f be a continuous and Gâteaux differentiable function on X. If f is M-P pseudoconvex (M-P pseudoconcave), then f is strictly quasiconvex (quasiconcave) and hence also quasiconvex (quasiconcave).*

3. Nondifferentiable multiplier rules and constraint qualifications. We first recall the notions of the contingent cone (also called the cone of tangents) and the cone of feasible directions.

DEFINITION 3.1 (contingent cone). *Let $\Omega \subseteq X$ and $\bar{x} \in \mathrm{cl}\Omega$. The contingent cone of Ω at $\bar{x}$ is the closed cone defined by*

$$T_\Omega(\bar{x}) := \{v \in X : \exists t_n \downarrow 0, v_n \to v \text{ s.t. } \bar{x} + t_n v_n \in \Omega \ \forall n\}.$$

DEFINITION 3.2 (cone of feasible directions). *Let $\Omega \subseteq X$ and $\bar{x} \in \mathrm{cl}\Omega$. The cone of feasible directions of Ω at $\bar{x}$ is the cone defined by*

$$D_\Omega(\bar{x}) := \{v \in X : \exists \delta > 0 \text{ s.t. } \bar{x} + tv \in \Omega \ \forall t \in (0, \delta)\}.$$

Based on the notions of contingent cone and M-P subdifferential, we extend the Abadie CQ introduced in [2] to our nondifferentiable setting.

DEFINITION 3.3 (nondifferentiable Abadie CQ). *Let $\bar{x} \in \Omega := \{x \in X : g_i(x) \leq 0,\ i = 1, 2, \ldots, I, h_j(x) = 0,\ j = 1, 2, \ldots, J\}$. We say that the nondifferentiable Abadie CQ holds at $\bar{x}$ if g_i $(i \in I(\bar{x}))$ and h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, the convex cone generated by*

$$A := \bigcup_{i \in I(\bar{x})} \partial^\diamond g_i(\bar{x}) \cup \bigcup_{j=1}^{J} \partial^\diamond h_j(\bar{x}) \cup \bigcup_{j=1}^{J} [-\partial^\diamond h_j(\bar{x})] \tag{2}$$

is closed, and

$$\left.\begin{array}{l} g_i^\diamond(\bar{x}; v) \leq 0 \ \forall i \in I(\bar{x}), \\ h_j^\diamond(\bar{x}; v) = 0 \ \forall j = 1, 2, \ldots, J \end{array}\right\} \Rightarrow v \in T_\Omega(\bar{x}).$$

Based on the notions of the cone of feasible directions and the M-P subdifferential, we extend the Zangwill CQ introduced in [32] from inequality constraints to inequality and equality constraints in the nondifferentiable setting.

DEFINITION 3.4 (generalized Zangwill CQ). *Let $\bar{x} \in \Omega := \{x \in X : g_i(x) \leq 0,\ i = 1, 2, \ldots, I, h_j(x) = 0,\ j = 1, 2, \ldots, J\}$. We say that the generalized Zangwill CQ holds at $\bar{x}$ if g_i $(i \in I(\bar{x}))$ and h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, the convex cone generated by the set A defined by (2) is closed, and*

$$\left.\begin{array}{l} g_i^\diamond(\bar{x}; v) \leq 0 \ \forall i \in I(\bar{x}), \\ h_j^\diamond(\bar{x}; v) = 0 \ \forall j = 1, 2, \ldots, J \end{array}\right\} \Rightarrow v \in \mathrm{cl}D_\Omega(\bar{x}).$$

LEMMA 3.1. *Let Ω be a closed subset of X and let $f : X \to R$ be either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$. If $\bar{x}$ is a local minimum of f over Ω, then*

$$f^\diamond(\bar{x}; v) \geq 0 \quad \forall v \in \mathrm{cl}D_\Omega(\bar{x}). \tag{3}$$

Moreover if f is either Fréchet differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, then

$$f^\diamond(\bar{x}; v) \geq 0 \quad \forall v \in T_\Omega(\bar{x}). \tag{4}$$

Proof. We first show that (3) holds. Suppose there exists $v \in D_\Omega(\bar{x})$ such that $f^\diamond(\bar{x}; v) < 0$. Then

$$\limsup_{t \downarrow 0} \frac{f(\bar{x} + tv) - f(\bar{x})}{t} \leq f^\diamond(\bar{x}; v) < 0,$$

which implies that

$$f(\bar{x}+tv)-f(\bar{x})<0 \quad \forall t>0 \text{ small enough.}$$

But this contradicts the fact that $\bar{x}$ is a local minimum of f over Ω, and hence

$$f^{\diamond}(\bar{x};v)\geq 0 \quad \forall v\in D_{\Omega}(\bar{x}).$$

Consequently (3) follows from the continuity of $f^{\diamond}(\bar{x};\cdot)$ (see (i) of Proposition 2.1).

Now suppose that there exists $v\in T_{\Omega}(\bar{x})$ such that $f^{\diamond}(\bar{x};v)<0$. Then there exist $r>0,\epsilon>0$ such that

$$f(\bar{x}+tv)-f(\bar{x})\leq -rt \quad \forall t\in(0,\epsilon).$$

If f is Lipschitz near $\bar{x}$, then there exists $\delta>0$ such that

$$f(\bar{x}+tv')-f(\bar{x}+tv)\leq L_f t\|v'-v\| \quad \forall v'\in B(v,\delta),$$

where L_f is the Lipschitz constant. By definition of the contingent cone, there exists $t_n\downarrow 0, v_n\to v$ such that $\bar{x}+t_nv_n\in\Omega$ for all n. Therefore for n large enough, one has

$$\|v_n-v\|<\frac{r}{2L_f}, \quad t_n\in(0,\epsilon)$$

and

$$\begin{aligned}
&f(\bar{x}+t_nv_n)-f(\bar{x})\\
&=f(\bar{x}+t_nv_n)-f(\bar{x}+t_nv)+f(\bar{x}+t_nv)-f(\bar{x})\\
&\leq L_f t_n\|v_n-v\|-rt_n\\
&<-\frac{r}{2}t_n<0.
\end{aligned}$$

But this contradicts the fact that $\bar{x}$ is a local minimum of f over Ω, and hence (4) holds.

We omit the proof for (4) under the Fréchet differentiability assumption since it is a classical result (see, e.g., [13, Theorem 4.14]). □

We now show that under the nondifferentiable Abadie CQ and the generalized Zangwill CQ, the KKT condition holds. It is interesting to note that although the nondifferentiable Abadie CQ is weaker than the generalized Zangwill CQ, the KKT condition under the nondifferentiable Abadie CQ requires stronger assumptions on the objective function.

THEOREM 3.1 (KKT conditions under Abadie CQ and Zangwill CQ). *Let $\bar{x}$ be a local optimal solution of* (P). *Under one of the following conditions,*

(i) *the generalized Zangwill CQ holds at $\bar{x}$, and f is either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$;*

(ii) *the nondifferentiable Abadie CQ holds at $\bar{x}$, and f is either Fréchet differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$,*

the KKT condition holds at $\bar{x}$, i.e., there exist scalars $\alpha_i\geq 0$ ($i\in I(\bar{x})$), $\beta_j\geq 0$, $\gamma_j\geq 0$ ($j=1,2,\ldots,J$) such that

$$0\in\partial^{\diamond}f(\bar{x})+\sum_{i\in I(\bar{x})}\alpha_i\partial^{\diamond}g_i(\bar{x})+\sum_{j=1}^{J}\beta_j\partial^{\diamond}h_j(\bar{x})-\sum_{j=1}^{J}\gamma_j\partial^{\diamond}h_j(\bar{x}).$$

Proof. Under the assumptions of the theorem, by Lemma 3.1 and Proposition 2.1 it is easy to show that $f^\diamond(\bar{x}; v) \geq 0$ for all v satisfying the following system:

$$\begin{aligned} g_i^\diamond(\bar{x}; v) &\leq 0 \quad \forall i \in I(\bar{x}), \\ h_j^\diamond(\bar{x}; v) &\leq 0 \quad \forall j = 1, 2, \ldots, J, \\ (-h_j)^\diamond(\bar{x}; v) &\leq 0 \quad \forall j = 1, 2, \ldots, J. \end{aligned}$$

Since g_i $(i \in I(\bar{x}))$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$ by (iv) of Proposition 2.1, v satisfies the above system if and only if $\max_{a \in A}\langle a, v\rangle \leq 0$, where A is the set defined by (2). Consequently,

$$f^\diamond(\bar{x}; v) \geq 0 \text{ whenever } \max_{a \in C}\langle a, v\rangle \leq 0,$$

where C denotes the convex cone generated by A. Thus the function $f^\diamond(\bar{x}; \cdot) + \delta_{C^0}(\cdot)$ attains its minimum at 0, where $C^0 := \{v \in X : \langle v, c\rangle \leq 0 \text{ for all } v \in C\}$ is the polar cone of C and δ_{C^0} is the indicator function of set C^0. Since $f^\diamond(\bar{x}; \cdot)$ is convex and continuous by virtue of Proposition 2.1(i), by the sum rule (see, e.g., [7, Corollary 1, p. 105]), one has

$$0 \in \partial\phi(0) + \partial\delta_{C^0}(0),$$

where $\phi(\cdot) := f^\diamond(\bar{x}; \cdot)$ and $\partial\phi$ is the subdifferential in the sense of convex analysis. Since $\partial\phi(0) = \partial^\diamond f(\bar{x})$ and $\partial\delta_{C^0}(0) = C^{00} = C$ the above inclusion is the same as

$$0 \in \partial^\diamond f(\bar{x}) + C.$$

Therefore there exist some $\xi^* \in \partial^\diamond f(\bar{x})$, $\eta_i^* \in \partial^\diamond g_i(\bar{x})$ $(i \in I(\bar{x}))$, $\zeta_j^*, \nu_j^* \in \partial^\diamond h_j(\bar{x})$ $(j = 1, 2, \ldots, J)$, $\alpha_i \geq 0$ $(i \in I(\bar{x}))$, $\beta_j \geq 0, \gamma_j \geq 0$ $(j = 1, 2, \ldots, J)$ such that

$$0 = \xi^* + \sum_{i \in I(\bar{x})} \alpha_i \eta_i^* + \sum_{j=1}^{J} \beta_j \zeta_j^* - \sum_{j=1}^{J} \gamma_j \nu_j^*,$$

which implies that the KKT condition holds. □

In the following theorem, we extend the classical KKT sufficient condition as given in [4, Theorem 4.3.8] to our nondifferentiable setting.

THEOREM 3.2 (nondifferentiable KKT sufficient condition). *Let $\bar{x}$ be a feasible solution of* (P). *Suppose that f, g_i $(i \in I(\bar{x}))$, h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$ and there exist scalars $\alpha_i \geq 0$ $(i \in I(\bar{x}))$, β_j $(j = 1, 2, \ldots, J)$ such that*

$$0 \in \partial^\diamond f(\bar{x}) + \sum_{i \in I(\bar{x})} \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \partial^\diamond h_j(\bar{x}). \tag{5}$$

Let $J^+ = \{j : \beta_j > 0\}$ and $J^- = \{j : \beta_j < 0\}$. Further suppose that f is M-P pseudoconvex at $\bar{x}$, g_i $(i \in I(\bar{x}))$, h_j $(j \in J^+)$, and $-h_j$ $(j \in J^-)$ are M-P regular and quasiconvex at $\bar{x}$. Then $\bar{x}$ is a global optimal solution of (P).

Proof. Note that (5) is equivalent to the existence of

$$\begin{aligned} &\xi^* \in \partial^\diamond f(\bar{x}), \quad \eta_i^* \in \partial^\diamond g_i(\bar{x}) \ (i \in I(\bar{x})), \\ &\gamma_j^* \in \partial^\diamond h_j\ (\bar{x})\ (j \in J^+), \quad \zeta_j^* \in \partial^\diamond(-h_j)(\bar{x})\ (j \in J^-) \end{aligned}$$

such that

$$0 = \xi^* + \sum_{i \in I(\bar{x})} \alpha_i \eta_i^* + \sum_{j \in J^+} \beta_j \gamma_j^* - \sum_{j \in J^-} \beta_j \zeta_j^*. \tag{6}$$

Let x be any feasible solution of (P); then for any $i \in I(\bar{x})$,

$$g_i(x) \leq 0 = g_i(\bar{x}).$$

By the quasiconvexity of g_i at $\bar{x}$ it follows that

$$g_i(\bar{x} + \lambda(x - \bar{x})) = g_i(\lambda x + (1 - \lambda)\bar{x}) \leq g_i(\bar{x}) \tag{7}$$

for all $\lambda \in (0, 1)$. This implies that

$$g_i^\diamond(\bar{x}; x - \bar{x}) = g_i'(\bar{x}; x - \bar{x}) \leq 0 \qquad \forall i \in I(\bar{x}) \tag{8}$$

by the M-P regularity. Similarly since $h_j(j \in J^+)$ and $-h_j(j \in J^-)$ are M-P regular and quasiconvex at $\bar{x}$, we have

$$h_j^\diamond(\bar{x}; x - \bar{x}) \leq 0 \quad \forall i \in J^+, \tag{9}$$

$$(-h_j)^\diamond(\bar{x}; x - \bar{x}) \leq 0 \quad \forall i \in J^-. \tag{10}$$

Note that (8)–(10) imply

$$\langle \eta_i^*, x - \bar{x} \rangle \leq 0 \quad \forall i \in I(\bar{x}), \tag{11}$$

$$\langle \gamma_j^*, x - \bar{x} \rangle \leq 0 \quad \forall j \in J^+, \tag{12}$$

$$\langle \zeta_j^*, x - \bar{x} \rangle \leq 0 \quad \forall j \in J^-. \tag{13}$$

Multiplying (11), (12), and (13) by $\alpha_i \geq 0$ $(i \in I(\bar{x}))$, $\beta_j > 0$ $(j \in J^+)$, and $-\beta_j > 0$ $(j \in J^-)$, respectively, and adding we get

$$\left\langle \sum_{i \in I(\bar{x})} \alpha_i \eta_i^* + \sum_{j \in J^+} \beta_j \gamma_j^* - \sum_{j \in J^-} \beta_j \zeta_j^*, x - \bar{x} \right\rangle \leq 0.$$

By virtue of (6), the above inequality implies that

$$\langle \xi^*, x - \bar{x} \rangle \geq 0,$$

which implies by (iv) of Proposition 2.1 that

$$f^\diamond(\bar{x}; x - \bar{x}) \geq \langle \xi^*, x - \bar{x} \rangle \geq 0$$

since $\xi^* \in \partial^\diamond f(\bar{x})$. By the M-P pseudoconvexity of f at $\bar{x}$, we must have $f(x) \geq f(\bar{x})$, and the proof is complete. □

We now extend the Kuhn–Tucker CQ introduced by Kuhn and Tucker in [15] to the nondifferentiable setting.

DEFINITION 3.5 (cone of attainable directions). *Let $\Omega \subseteq X$ and $\bar{x} \in \mathrm{cl}\Omega$. We say that $v \in A_\Omega(\bar{x})$, the cone of attainable directions of Ω at $\bar{x}$ if there exist $\delta > 0$, and a mapping $\alpha : R \to X$ such that $\alpha(\tau) \in \Omega$ for all $\tau \in (0, \delta), \alpha(0) = \bar{x}$, and $\lim_{\tau \downarrow 0} \frac{\alpha(\tau) - \alpha(0)}{\tau} = v$.*

The cone of attainable directions is also known as the adjacent cone (see, e.g., [1]) or the incident cone. In fact

$$A_\Omega(\bar{x}) = \liminf_{\tau \downarrow 0} \frac{\Omega - \bar{x}}{\tau}$$

and hence is a closed set.

Definition 3.6 (nondifferentiable Kuhn–Tucker CQ). *Let* $\bar{x} \in \Omega := \{x \in X : g_i(x) \leq 0,\ i = 1, 2, \ldots, I, h_j(x) = 0,\ j = 1, 2, \ldots, J\}$*. We say the nondifferentiable Kuhn–Tucker CQ is satisfied at* $\bar{x}$ *if* g_i $(i \in I(\bar{x}))$ *and* h_j $(j = 1, 2, \ldots, J)$ *are either Gâteaux differentiable at* $\bar{x}$ *or Lipschitz near* $\bar{x}$*, the convex cone generated by the set* (2) *is closed, and*

$$\left.\begin{aligned} g_i^\diamond(\bar{x}; v) \leq 0\ \forall i \in I(\bar{x}), \\ h_j^\diamond(\bar{x}; v) = 0\ \forall j = 1, 2, \ldots, J \end{aligned}\right\} \Rightarrow v \in A_\Omega(\bar{x}).$$

It is easy to see that $\text{cl}D_\Omega(\bar{x}) \subseteq A_\Omega(\bar{x}) \subseteq T_\Omega(\bar{x})$, and hence the following relationship among the generalized Zangwill CQ, the nondifferentiable Kuhn–Tucker CQ, and the nondifferentiable Abadie CQ is obvious.

Proposition 3.1. *The nondifferentiable Zangwilll CQ implies the nondifferentiable Kuhn–Tucker CQ, and the nondifferentiable Kuhn–Tucker CQ implies the nondifferentiable Abadie CQ. That is,*

$$\textit{Zangwill CQ} \Longrightarrow \textit{Kuhn–Tucker CQ} \Longrightarrow \textit{Abadie CQ}.$$

Definition 3.7 (nondifferentiable Mangasarian–Fromovitz CQ). *Let* $\bar{x}$ *be a feasible solution of* (P)*. We say that the nondifferentiable Mangasarian–Fromovitz CQ is satisfied if* g_i $(i \in I(\bar{x}))$ *are either Hadamard differentiable at* $\bar{x}$ *or Lipschitz near* $\bar{x}$*,* g_i $(i \notin I(\bar{x}))$ *are continuous at* $\bar{x}$*,* h_j $(j = 1, 2, \ldots, J)$ *are Fréchet differentiable at* $\bar{x}$ *and continuous in a neighborhood of* $\bar{x}$*,* $\{Dh_1(\bar{x}), \ldots, Dh_J(\bar{x})\}$ *are linearly independent, and there exists* $v \in X$ *such that*

$$g_i^\diamond(\bar{x}; v) < 0 \quad \forall i \in I(\bar{x}), \tag{14}$$

$$\langle Dh_j(\bar{x}), v\rangle = 0 \quad \forall j = 1, 2, \ldots, J. \tag{15}$$

Lemma 3.2. *If* g_i $(i \in I(\bar{x}))$ *are either Hadamard differentiable at* $\bar{x}$ *or Lipschitz near* $\bar{x}$*,* g_i $(i \notin I(\bar{x}))$ *are continuous at* $\bar{x}$*,* h_j $(j = 1, 2, \ldots, J)$ *are Fréchet differentiable at* $\bar{x}$ *and continuous in a neighborhood of* $\bar{x}$*, then the nondifferentiable Mangasarian–Fromovitz CQ is equivalent to the nonexistence of* $(\alpha, \beta) \in R_+^I \times R^J$ *such that* $(\alpha, \beta) \neq 0$ *and*

$$0 \in \sum_{i \in I(\bar{x})} \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^J \beta_j Dh_j(\bar{x}). \tag{16}$$

Proof. We prove the lemma by contradiction.

Suppose the nondifferentiable Mangasarian–Fromovitz CQ holds but there exists a nonzero vector $(\alpha, \beta) \in R_+^I \times R^J$ such that

$$0 = \sum_{i \in I(\bar{x})} \alpha_i \xi_i^* + \sum_{j=1}^J \beta_j Dh_j(\bar{x}) \tag{17}$$

for some $\xi_i^* \in \partial^\diamond g_i(\bar{x}),\ i \in I(\bar{x})$. Since the vectors $Dh_j(\bar{x})$ are linearly independent, at least one α_i is nonzero. By (17), for v which is a solution of (14), (15),

$$\sum_{i\in I(\bar{x})} \alpha_i \langle \xi_i^*, v\rangle = -\sum_{j=1}^{J} \beta_j \langle Dh_j(\bar{x}), v\rangle.$$

But this is impossible since the right-hand side of the equation is zero while the left-hand side of the equation is nonzero. Therefore the nondifferentiable Mangasarian–Fromovitz CQ implies that there is no nonzero vector $(\alpha, \beta) \in R_+^I \times R^J$ such that (16) holds.

Conversely suppose that there is no nonzero vector $(\alpha, \beta) \in R_+^I \times R^J$ such that (16) holds. It is obvious that under this assumption, $Dh_1(\bar{x}), \ldots, Dh_J(\bar{x})$ are linearly independent. We first prove that for any given $i \in I(\bar{x})$, there exists $v \in X$ such that

$$g_i^\diamond(\bar{x}; v) < 0, \tag{18}$$

$$\langle Dh_j(\bar{x}), v\rangle = 0 \quad \forall j = 1, 2, \ldots, J. \tag{19}$$

If, on the contrary, the above system has no solution, then $v = 0$ is a solution to the following optimization problem:

$$\begin{aligned} \min \quad & g_i^\diamond(\bar{x}; v) \\ \text{s.t.} \quad & \langle Dh_j(\bar{x}), v\rangle = 0 \quad \forall j = 1, \ldots, J. \end{aligned}$$

Since the objective function is convex and the constraints are linear, by the Lagrange multiplier rule and the fact that $\partial\phi(0) = \partial^\diamond g_i(\bar{x})$ for $\phi(\cdot) := g_i^\diamond(\bar{x}; \cdot)$ there must exist $\beta \in R^J$ such that

$$0 \in \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j Dh_j(\bar{x}),$$

which is a contradiction. Now we can show that for any two given $i, i' \in I(\bar{x})$, there exists $v \in X$ such that

$$\begin{aligned} g_{i'}^\diamond(\bar{x}; v) &< 0, \\ g_i^\diamond(\bar{x}; v) &< 0, \\ \langle Dh_j(\bar{x}), v\rangle &= 0 \quad \forall j = 1, 2, \ldots, J. \end{aligned}$$

On the contrary, suppose that the above system does not have a solution. Then $g_{i'}^\diamond(\bar{x}; v) \geq 0$ for all v satisfying the system (18)–(19), which implies that $v = 0$ is a solution to the following optimization problem with convex constraints:

$$\begin{aligned} \min \quad & g_{i'}^\diamond(\bar{x}; v) \\ \text{s.t.} \quad & g_i^\diamond(\bar{x}; v) \leq 0, \\ & \langle Dh_j(\bar{x}, v) = 0 \quad \forall j = 1, \ldots, J. \end{aligned}$$

Indeed, let v be any feasible solution of the above problem and let u be a solution of (18)–(19); then for any $t > 0$, $u + tv$ is a solution of (18)–(19), and hence $g_{i'}^\diamond(\bar{x}; v+tu) \geq 0$ by the assumption, which implies that $g_{i'}^\diamond(\bar{x}; v) \geq 0$ after taking limits as $t \to 0$. By the Lagrange multiplier rule, since the Slater condition holds for the above optimization problem, there must exist $(\alpha_i, \beta) \in R_+ \times R^J$ such that

$$0 \in \partial^\diamond g_{i'}(\bar{x}) + \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j D h_j(\bar{x}),$$

which is a contradiction. The rest of the proof follows from the mathematical induction. ∎

Definition 3.8 (nondifferentiable linear independence CQ). *Let $\bar{x}$ be a feasible solution of* (P). *We say that the nondifferentiable linear independence CQ is satisfied if g_i $(i \in I(\bar{x}))$ are either Hadamard differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$, h_j $(j = 1, 2, \ldots, J)$ are Fréchet differentiable at $\bar{x}$ and continuous in a neighborhood of $\bar{x}$, and for any $\xi_i^* \in \partial^\diamond g_i(\bar{x})$ $\{\xi_i^*(i \in I(\bar{x})), Dh_1(\bar{x}), \ldots, Dh_J(\bar{x})\}$ are linearly independent.*

The following is a straightforward consequence of Lemma 3.2.

Proposition 3.2 (LICQ implies MFCQ). *The nondifferentiable linear independence CQ implies the nondifferentiable Mangasarian–Fromovitz CQ.*

Definition 3.9 (nondifferentiable Slater CQ). *Let $\bar{x}$ be a feasible solution of* (P). *We say that the nondifferentiable Slater CQ is satisfied at $\bar{x}$ if g_i $(i \in I(\bar{x}))$ are M-P pseudoconvex at $\bar{x}$ and either Hadamard differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$; g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$; h_j $(j = 1, 2, \ldots, J)$ are Fréchet differentiable at $\bar{x}$, continuous in a neighborhood of $\bar{x}$, and quasiaffine at $\bar{x}$; $\{Dh_1(\bar{x}), \ldots, Dh_J(\bar{x})\}$ are linearly independent; and there exists $\hat{x} \in X$ such that*

$$\begin{aligned} g_i(\hat{x}) &< 0 \quad \forall i \in I(\bar{x}), \\ h_j(\hat{x}) &= 0 \quad \forall j = 1, 2, \ldots, J. \end{aligned}$$

Proposition 3.3 (Slater CQ implies MFCQ). *The nondifferentiable Slater CQ implies the nondifferentiable Mangasarian–Fromovitz CQ.*

Proof. Since $g_i(\hat{x}) < g_i(\bar{x})$ for all $i \in I(\bar{x})$, by the M-P pseudoconvexity of g_i $(i \in I(x))$ we have

$$g_i^\diamond(\bar{x}; \hat{x} - \bar{x}) < 0, \quad i \in I(\bar{x}).$$

Also since $h_j(\hat{x}) = h_j(\bar{x})$ $(j = 1, \ldots, J)$, quasiconvexity and quasiconcavity of h_j at $\bar{x}$ implies that

$$\langle Dh_j(\bar{x}), \hat{x} - \bar{x} \rangle = 0, \quad j = 1, \ldots, J.$$

Thus the system (14)–(15) has a solution $v = \hat{x} - \bar{x}$ and the nondifferentiable Mangasarian–Fromovitz CQ is satisfied. ∎

Proposition 3.4 (MFCQ implies Kuhn–Tucker CQ). *The nondifferentiable Mangasarian–Fromovitz CQ implies the nondifferentiable Kuhn–Tucker CQ.*

Proof. We first show that the convex cone generated by the set

$$A = \bigcup_{i \in I(\bar{x})} \partial^\diamond g_i(\bar{x}) \cup \bigcup_{j=1}^{J} \{\pm Dh_j(\bar{x})\}$$

is closed. It is easy to see that

$$cone A = cone \bigcup_{i \in I(\bar{x})} \partial^\diamond g_i(\bar{x}) + \left\{ \sum_{j=1}^{J} \beta_j Dh_j(\bar{x}) : \beta_j \in R \right\},$$

where *coneA* denotes the convex cone generated by set A. Since $\mathrm{co}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x})$ is a nonempty convex weak*-compact subset of X^* not containing zero, as it is shown in [22, Corollary 9.6.1], *cone* $\mathrm{co}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x})$ is closed. But

$$\mathit{cone}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x}) = \mathit{cone}\ \mathrm{co}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x});$$

hence $\mathit{cone}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x})$ is a closed convex cone. By Lemma 3.2, for any nonzero $\beta\in R^J$,

$$\sum_{j=1}^{J}\beta_j Dh_j(\bar{x}) \notin \mathit{cone}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x}).$$

By the linear independence of $\{Dh_1(\bar{x}),\dots,Dh_J(\bar{x})\}$, $0\neq\sum_{j=1}^{J}\beta_j Dh_j(\bar{x})$ for any nonzero $\beta\in R^J$. Therefore any nonzero element in the cone $\{\sum_{j=1}^{J}\beta_j Dh_j(\bar{x}):\beta_j\in R\}$ is not in the closed convex cone $\mathit{cone}\bigcup_{i\in I(\bar{x})}\partial^{\diamond}g_i(\bar{x})$, which implies that *coneA* is closed by [22, Corollary 9.1.2] (which obviously holds in any general Banach space as well).

Let v be a solution of (14) and (15). Since $\{Dh_1(\bar{x}),\dots,Dh_J(\bar{x})\}$ are linearly independent, by the correction theorem of Halkin [9, Theorem F], there exist a neighborhood U of $\bar{x}$ and a continuous mapping ζ from U into X such that $\zeta(\bar{x})=0$, ζ is Fréchet differentiable at $\bar{x}$ with $D\zeta(\bar{x})=0$, and

$$h_j(x+\zeta(x)) = \langle Dh_j(\bar{x}), x-\bar{x}\rangle \quad \forall x\in U, j=1,2,\dots,J. \tag{20}$$

For all $t\in R$ such that $\bar{x}+tv\in U$, denote

$$\alpha(t) := \bar{x}+tv+\zeta(\bar{x}+tv).$$

Then $h_j(\alpha(t))=0,\ j=1,2,\dots,J$, for all $t>0$ small enough. Let $i\in I(\bar{x})$. If g_i is Hadamard differentiable at $\bar{x}$, then since $\lim_{t\downarrow 0}\zeta(\bar{x}+tv)/t = D\zeta(\bar{x})=0$ by (14)

$$g_i(\alpha(t))<0 \quad \forall t>0 \text{ small enough.}$$

Now suppose that g_i is Lipschitz near $\bar{x}$. By (14) one has that there exists $r>0$ such that for $t>0$ small enough

$$g_i(\bar{x}+tv)-g_i(\bar{x}) < -rt.$$

Since g_i is Lipschitz near $\bar{x}$, for $t>0$ small enough and any $v'\in X$

$$g_i(\bar{x}+tv')-g_i(\bar{x}+tv) \le L_{g_i}t\|v'-v\|.$$

Since $D\zeta(\bar{x})=0, \zeta(\bar{x})=0$, one has

$$\|\zeta(\bar{x}+tv)\| < \frac{r}{2L_{g_i}}t \quad \forall t>0 \text{ small enough,}$$

and hence

$$\begin{aligned} g_i(\alpha(t)) &= g_i(\bar{x}+tv+\zeta(\bar{x}+tv)) \\ &= g_i\left(\bar{x}+t\left[v+\frac{\zeta(\bar{x}+tv)}{t}\right]\right) - g_i(\bar{x}+tv) + g_i(\bar{x}+tv) - g_i(\bar{x}) \\ &\le L_{g_i}\|\zeta(\bar{x}+tv)\| - rt \\ &< -\frac{r}{2}t < 0 \quad \text{ for } t>0 \text{ small enough.} \end{aligned}$$

By the continuity assumptions on $g_i(i \notin I(\bar{x}))$, one also has

$$g_i(\alpha(t)) < 0 \quad \forall t > 0 \text{ small enough}, \quad i \notin I(\bar{x}).$$

Hence $v \in A_\Omega(\bar{x})$.

Now let v satisfy

$$g_i^\diamond(\bar{x}; v) \le 0 \quad \forall i \in I(\bar{x}),$$
$$\langle Dh_j(\bar{x}), v\rangle = 0 \quad \forall j = 1, 2, \ldots, J.$$

By the assumption of the Mangasarian–Fromovitz CQ, there exists a sequence $\{v_k\}$ such that

$$g_i^\diamond(\bar{x}; v_k) < 0 \quad \forall i \in I(\bar{x}),$$
$$\langle Dh_j(\bar{x}), v_k\rangle = 0 \quad \forall j = 1, 2, \ldots, J,$$

and $v = \lim_{k\to\infty} v_k$. By the proof above, $v_k \in A_\Omega(\bar{x})$, and so $v = \lim_{k\to\infty} v_k \in A_\Omega(\bar{x})$. □

We now extend the Arrow–Hurwicz–Uzawa CQ introduced in [3] to the nondifferentiable setting.

Definition 3.10 (nondifferentiable Arrow–Hurwicz–Uzawa CQ). *Let $\bar{x}$ be a feasible solution of* (P). *We say that the nondifferentiable Arrow–Hurwicz–Uzawa CQ is satisfied at $\bar{x}$ if g_i $(i \in I(\bar{x}))$, h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$, h_j $(j = 1, 2, \ldots, J)$ are M-P pseudoaffine at $\bar{x}$, the convex cone generated by the set* (2) *is closed, and there exists $v \in X$ such that*

$$g_i^\diamond(\bar{x}; v) < 0 \quad \forall i \in W, \tag{21}$$
$$g_i^\diamond(\bar{x}; v) \le 0 \quad \forall i \in V, \tag{22}$$
$$h_j^\diamond(\bar{x}; v) = 0 \quad \forall j = 1, 2, \ldots, J, \tag{23}$$

where

$$V := \{i \in I(\bar{x}) : g_i \text{ is M-P pseudoconcave at } \bar{x}\},$$
$$W := I(\bar{x})\backslash V.$$

Proposition 3.5 (AHUCQ implies Zangwill CQ). *The nondifferentiable Arrow–Hurwicz–Uzawa CQ implies the generalized Zangwill CQ.*

Proof. Suppose that v satisfies (21)–(23). For any $i \in W$ by virtue of (21), for all $\tau \in (0, 1]$ small enough,

$$g_i(\bar{x} + \tau v) < g_i(\bar{x}) = 0.$$

For $i \in V$ by virtue of (22), $g_i^\diamond(\bar{x}; v) \le 0$, which implies by the definition of the M-P pseudoconcavity that $g_i(\bar{x}+\tau v) \le g_i(\bar{x})$ for all $\tau \ge 0$ small enough. By the continuity assumptions at $\bar{x}$ for g_i $(i \notin I(\bar{x}))$, for all $\tau \in (0, 1]$ small enough,

$$g_i(\bar{x} + \tau v) < 0 \quad \forall i \notin I(\bar{x}).$$

Hence for all $\tau > 0$ small enough,

$$g_i(\bar{x} + \tau v) \le 0, \qquad i = 1, 2, \ldots, I,$$
$$h_j(\bar{x} + \tau v) = 0 \qquad \forall j = 1, 2, \ldots, J,$$

which implies that $v \in D_\Omega(\bar{x})$ and the proof of the proposition is complete due to the continuity of $g_i^\diamond(\bar{x}, \cdot)$ $(i \in I(\bar{x}))$ and $h_j^\diamond(\bar{x}, \cdot)$ $(j = 1, 2, \ldots, J)$. ∎

DEFINITION 3.11 (nondifferentiable weak Slater CQ). *Let $\bar{x}$ be a feasible solution of* (P). *We say the nondifferentiable weak Slater CQ holds at $\bar{x}$ if g_i $(i \in I(\bar{x}))$ are M-P pseudoconvex at $\bar{x}$ and either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$; g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$; h_j $(j = 1, 2, \ldots, J)$ are Gâteaux differentiable, continuous, and M-P pseudoaffine; the convex cone generated by the set* (2) *is closed; and there exists $\hat{x} \in X$ such that*

$$\begin{aligned} g_i(\hat{x}) &< 0, \quad i \in I(\bar{x}), \\ h_j(\hat{x}) &= 0, \quad j = 1, 2, \ldots, J. \end{aligned}$$

PROPOSITION 3.6 (weak Slater CQ implies AHUCQ). *The nondifferentiable weak Slater CQ implies the nonsmooth Arrow–Hurwicz–Uzawa CQ.*

Proof. Take $V = \emptyset$ and $W = I(\bar{x})$. Since $g_i(\hat{x}) < g_i(\bar{x})$ for all $i \in I(\bar{x})$, by the M-P pseudoconvexity of g_i $(i \in I(x))$ we have

$$g_i^\diamond(\bar{x}; \hat{x} - \bar{x}) < 0, \quad i \in I(\bar{x}). \tag{24}$$

By Proposition 2.3, h_j $(j = 1, 2, \ldots, J)$ are quasiaffine. Since $h_j(\hat{x}) = h_j(\bar{x})$ $(j = 1, \ldots, J)$, quasiconvexity and quasiconcavity of h_j at $\bar{x}$ implies that

$$h_j'(\bar{x}; \hat{x} - \bar{x}) = 0.$$

Hence the system (21)–(23) has a solution $v = \hat{x} - \bar{x}$. The proof of the proposition is complete. ∎

DEFINITION 3.12 (nondifferentiable weak reverse convex CQ). *We say that the nondifferentiable weak reverse convex CQ holds at $\bar{x}$ if g_i $(i \in I(\bar{x}))$, h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, g_i $(i \in I(\bar{x}))$ are M-P pseudoconcave at $\bar{x}$, g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$, h_j $(j = 1, 2, \ldots, J)$ are M-P pseudoaffine, and the convex cone generated by the set* (2) *is closed.*

Since (22)–(23) always has a solution $v = 0$, the following relationship between the nondifferentiable weak reverse convex constraint CQ and the nondifferentiable Arrow–Hurwicz–Uzawa CQ is immediate.

PROPOSITION 3.7 (weak reverse convex CQ implies AHUCQ). *The nondifferentiable weak reverse convex constraint CQ implies the nondifferentiable Arrow–Hurwicz–Uzawa CQ.*

Finally we end this section with an equivalent condition to the nondifferentiable Arrow–Hurwicz–Uzawa CQ. We omit the proof since the proof is similar to that of Lemma 3.2.

PROPOSITION 3.8. *Suppose that g_i $(i \in I(\bar{x}))$, h_j $(j = 1, 2, \ldots, J)$ are either Gâteaux differentiable at $\bar{x}$ or Lipschitz near $\bar{x}$, g_i $(i \notin I(\bar{x}))$ are continuous at $\bar{x}$, h_j $(j = 1, 2, \ldots, J)$ are M-P pseudoaffine at $\bar{x}$, and the convex cone generated by the set* (2) *is closed. The nondifferentiable Arrow–Hurwicz–Uzawa CQ is equivalent to the nonexistence of $(\alpha, \beta) \in R^I_+ \times R^J$ such that $\alpha_W \neq 0$ and*

$$0 \in \sum_{i \in W} \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{i \in V} \alpha_i \partial^\diamond g_i(\bar{x}) + \sum_{j=1}^{J} \beta_j \partial^\diamond h_j(\bar{x}), \tag{25}$$

where

$$\begin{aligned} V &:= \{i \in I(\bar{x}) : g_i \text{ is M-P pseudoconcave at } \bar{x}\}, \\ W &:= I(\bar{x}) \backslash V. \end{aligned}$$

4. Bilevel optimization. In this section we apply the results obtained in the previous section to the bilevel optimization problem,

$$\text{(BP)} \qquad \begin{aligned} \min \quad & F(x,y) \\ \text{s.t.} \quad & y \in S(x), \\ & G_i(x,y) \le 0, \quad i = 1,2,\dots,I, \\ & H_j(x,y) = 0, \quad j = 1,2,\dots,J, \end{aligned}$$

where $S(x)$ denotes the set of solutions of the lower level problem:

$$(P_x) \qquad \begin{aligned} \min_y \quad & f(x,y) \\ \text{s.t.} \quad & g_i(x,y) \le 0, \quad i = 1,2,\dots,m, \\ & h_j(x,y) = 0, \quad j = 1,2,\dots,n, \end{aligned}$$

and F, G_i, H_j, f, g_i, h_j are functions on the Banach space $X \times Y$. For simplicity we assume that $S(x)$ is nonempty for all $x \in X$.

Define the value function of the lower level problem by

$$V(x) := \min_y \{f(x,y) : g_i(x,y) \le 0,\ h_j(x,y) = 0,\ i = 1,2,\dots,m,\ j = 1,2,\dots,n\}.$$

Then (BP) can be reformulated as the following single level optimization problem:

$$\text{(SP)} \qquad \begin{aligned} \min \quad & F(x,y) \\ \text{s.t.} \quad & f(x,y) - V(x) \le 0, \\ & g_i(x,y) \le 0, \quad i = 1,2,\dots,m, \\ & h_j(x,y) = 0, \quad j = 1,2,\dots,n, \\ & G_i(x,y) \le 0, \quad i = 1,2,\dots,I, \\ & H_j(x,y) = 0, \quad j = 1,2,\dots,J. \end{aligned}$$

It is known that $V(x)$ may not be differentiable in general, even in the case where all problem data f, g_i, h_j are continuously differentiable, and hence a nonsmooth multiplier rule must be used as in [28, 29]. Moreover it was shown in [28, Proposition 3.2] that the CQs such as the linear independence CQ, the Slater CQ, and the Mangasarian–Fromovitz CQ do not hold for (SP). It is obvious that the nondifferentiable weak Slater CQ will never be satisfied since the inequality constraint $f(x,y) - V(x) \le 0$ is actually an equality constraint. In this section, we show that it is possible for the nondifferentiable weak reverse convex CQ to hold; hence the nondifferentiable Arrow–Hurwicz–Uzawa CQ, the generalized Zangwill CQ, the nondifferentiable Kuhn–Tucker CQ, and the nondifferentiable Abadie CQ are also applicable CQs for (SP).

Excluding the CQs that will never hold for (SP) such as (2), (5)–(7) in Theorem 1.1, we derive the KKT condition for (SP) by using the calculus rules for the M-P subdifferential in (ii)–(iii) of Proposition 2.1 as follows.

THEOREM 4.1. *Let* $(\bar{x}, \bar{y})$ *be a local optimal solution of* (SP)*. Suppose that the objective function* $F(x,y)$ *is either Gâteaux differentiable at* $(\bar{x}, \bar{y})$ *or Lipschitz near* $(\bar{x}, \bar{y})$ *and the value function* $V(x)$ *is Lipschitz near* $\bar{x}$*. If one of the CQs such as the nondifferentiable weak reverse convex CQ, the nondifferentiable Arrow–Hurwicz–Uzawa CQ, the generalized Zangwill CQ, the nondifferentiable Kuhn–Tucker CQ, and* *the nondifferentiable Abadie CQ holds at* $(\bar{x}, \bar{y})$*, then the KKT condition holds; i.e., there exist scalars* $\lambda \geq 0$*,* $\alpha_i \geq 0$ $(i = 1, 2, \ldots, m)$*,* β_j $(j = 1, 2, \ldots, p)$*,* γ_i $(i = 1, 2, \ldots, I)$*,* η_j $(j = 1, 2, \ldots, J)$ *such that*

$$
\begin{aligned}
&0 \in \partial^{\diamond} F(\bar{x}, \bar{y}) + \lambda(\partial^{\diamond} f(\bar{x}, \bar{y}) - \partial^{\diamond} V(\bar{x}) \times \{0\}) + \sum_{i=1}^{m} \alpha_i \partial^{\diamond} g_i(\bar{x}, \bar{y}) \\
&\qquad + \sum_{j=1}^{n} \beta_j \partial^{\diamond} h_j(\bar{x}, \bar{y}) + \sum_{i=1}^{I} \gamma_i \partial^{\diamond} G_i(\bar{x}, \bar{y}) + \sum_{j=1}^{J} \eta_j \partial^{\diamond} H_j(\bar{x}, \bar{y}), \\
&\alpha_i g_i(\bar{x}, \bar{y}) = 0, \quad i = 1, 2, \ldots, m, \\
&\gamma_i G_i(\bar{x}, \bar{y}) = 0, \quad i = 1, 2, \ldots, I.
\end{aligned}
$$

In the above KKT condition, we need to give an upper estimate for the term $\partial^{\diamond} V(\bar{x})$. Such an estimate usually involves a convex combination of solutions and multipliers for the lower level problem as in [28, Proposition 2.1], and a growth hypothesis assumption is usually needed [7, Theorem 6.2]. However, in the case when the value function is convex, no such convex combination and growth hypothesis are needed, as the following result indicates.

PROPOSITION 4.1. *Let* $x \in X$ *and* $y \in S(x)$*. Suppose that* f, g_i $(i = 1, 2, \ldots, I)$*,* h_j $(j = 1, 2, \ldots, J)$ *are Gâteaux differentiable at* (x, y) *and the KKT condition holds for problem* (26)*. If the value function* $V(x)$ *is convex, then for any* $y \in S(x)$*,*

$$
\partial V(x) \subseteq \left\{ D_x f(x, y) + \sum_{i=1}^{m} \nu_i D_x g_i(x, y) + \sum_{j=1}^{n} \pi_j D_x h_j(x, y) : (\nu, \pi) \in M^1(y) \right\},
$$

where $\partial V(x)$ *denotes the subdifferential in the sense of convex analysis and* $M^1(y)$ *is the set of multipliers for* (P_x)*:*

$$
M^1(y) := \left\{ (\nu, \pi) \in R^m_+ \times R^n : \begin{array}{l} D_y f(x, y) + \sum_{i=1}^{m} \nu_i D_y g_i(x, y) + \sum_{j=1}^{n} \pi_j D_y h_j(x, y) = 0 \\ \nu_i g_i(x, y) = 0, \quad i = 1, 2, \ldots, m \end{array} \right\}.
$$

Proof. Now let $\xi \in \partial V(x)$. Then by definition of the subdifferential in the sense of convex analysis,

$$
V(x') - V(x) \geq \langle \xi, x' - x \rangle \quad \forall x' \in X,
$$

which implies by definition of the value function that for all (x', y') satisfying the constraints

$$
g_i(x', y') \leq 0, \quad i = 1, 2, \ldots, m, \quad h_j(x', y') = 0, \quad j = 1, 2, \ldots, n,
$$

one has

$$
f(x', y') - f(x, y) \geq \langle \xi, x' - x \rangle.
$$

That is, $(x', y') = (x, y)$ is a solution to the following optimization problem:

$$
(26) \qquad \begin{cases} \min_{x', y'} & f(x', y') - \langle \xi, x' \rangle, \\ & g_i(x', y') \leq 0, \quad i = 1, 2, \ldots, m, \\ & h_j(x', y') = 0, \quad i = 1, 2, \ldots, n. \end{cases}
$$

By the KKT condition there exists $(\nu, \pi) \in R^m_+ \times R^p$ such that

$$
\begin{aligned}
&0 = D_x f(x,y) - \xi + \sum_{i=1}^m \nu_i D_x g_i(x,y) + \sum_{j=1}^n \pi_j D_x h_j(x,y),\\
&0 = D_y f(x,y) + \sum_{i=1}^m \nu_i D_y g_i(x,y) + \sum_{j=1}^n \pi_j D_y h_j(x,y),\\
&\nu_i g_i(x,y) = 0, \quad i = 1,2,\ldots,m,
\end{aligned}
$$

and hence the proof is complete. ∎

Combining Theorem 4.1 and Proposition 4.1, we now have a KKT condition for (SP) which does not involve the value function.

THEOREM 4.2. *Let* $(\bar{x}, \bar{y})$ *be a local optimal solution of* (SP). *Suppose that* $F(x,y)$ *is either Gâteaux differentiable at* $(\bar{x},\bar{y})$ *or Lipschitz continuous near* $(\bar{x},\bar{y})$, f, g_i $(i = 1,2,\ldots,I)$, h_j $(j = 1,2,\ldots,J)$ *are Gâteaux differentiable; the KKT condition for problem* (26) *holds; and the value function* $V(x)$ *is convex. If one of the CQs, such as the nondifferentiable weak reverse convex CQ, the nondifferentiable Arrow–Hurwicz–Uzawa CQ, the generalized Zangwill CQ, the nondifferentiable Kuhn–Tucker CQ, or the nondifferentiable Abadie CQ, holds at* $(\bar{x},\bar{y})$, *then there exist scalars* $\lambda \geq 0, \alpha_i \geq 0, \nu_i \geq 0$ $(i = 1,2,\ldots,m)$, β_j, π_j $(j = 1,2,\ldots,n)$, $\gamma_i \geq 0$ $(i = 1,2,\ldots,I)$, η_j $(j = 1,2,\ldots,J)$ *such that*

$$
\begin{aligned}
&0 \in \partial^\diamond F(\bar{x},\bar{y}) + \sum_{i=1}^m (\alpha_i - \lambda\nu_i) D g_i(\bar{x},\bar{y}) + \sum_{j=1}^n \beta_j D h_j(\bar{x},\bar{y})\\
&\qquad + \sum_{i=1}^I \gamma_i \partial^\diamond G_i(\bar{x},\bar{y}) + \sum_{j=1}^J \eta_j \partial^\diamond H_j(\bar{x},\bar{y}),\\
&0 = D_y f(\bar{x},\bar{y}) + \sum_{i=1}^m \nu_i D_y g_i(\bar{x},\bar{y}) + \sum_{j=1}^n \pi_j D_y h_j(\bar{x},\bar{y}),\\
&\nu_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m,\\
&\alpha_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m,\\
&\gamma_i G_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,I.
\end{aligned}
$$

Proof. Applying Theorem 4.1 and Proposition 4.1, we find scalars $\lambda \geq 0, \alpha_i \geq 0, \nu_i \geq 0$ $(i = 1,2,\ldots,m)$, ζ_j, π_j $(j = 1,2,\ldots,n)$, $\gamma_i \geq 0$ $(i = 1,2,\ldots,I)$, η_j $(j = 1,2,\ldots,J)$ such that

$$
\begin{aligned}
&0 \in \partial^\diamond F(\bar{x},\bar{y}) - \lambda \left[\sum_{i=1}^m \nu_i D_x g_i(\bar{x},\bar{y}) + \sum_{j=1}^n \pi_j D_x h_j(\bar{x},\bar{y}) \right] \times \{-D_y f(\bar{x},\bar{y})\}\\
&\qquad + \sum_{i=1}^m \alpha_i D g_i(\bar{x},\bar{y}) + \sum_{j=1}^n \zeta_j D h_j(\bar{x},\bar{y}) + \sum_{i=1}^I \gamma_i \partial^\diamond G_i(\bar{x},\bar{y}) + \sum_{j=1}^J \eta_j \partial^\diamond H_j(\bar{x},\bar{y}), \qquad (27)
\end{aligned}
$$

$$
0 = D_y f(\bar{x},\bar{y}) + \sum_{i=1}^m \nu_i D_y g_i(\bar{x},\bar{y}) + \sum_{j=1}^n \pi_j D_y h_j(\bar{x},\bar{y}), \qquad (28)
$$

$$
\begin{aligned}
&\nu_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m,\\
&\alpha_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m,\\
&\gamma_i G_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,I.
\end{aligned}
$$

From (28),

$$-D_y f(\bar{x},\bar{y}) = \sum_{i=1}^{m} \nu_i D_y g_i(\bar{x},\bar{y}) + \sum_{j=1}^{n} \pi_j D_y h_j(\bar{x},\bar{y}).$$

Substituting the above into (27) and denoting $\beta_j = \zeta_j - \lambda\pi_j$ completes the proof. □

We now consider the special case when the lower level problem is linear; i.e., the functions $f(x,y), g_i(x,y), h_j(x,y)$ are all jointly linear. It is known that for linear bilevel programming problems, which are the bilevel optimization problems where the lower level problem is jointly linear and there is no upper level constraints, no CQs are needed. By Theorem 4.2 and the weak reverse convex CQ, we have the following KKT condition for the following "generalized linear" bilevel optimization problem where no constraint qualification is needed.

COROLLARY 4.1. *Let $(\bar{x},\bar{y})$ be a local optimal solution of* (SP). *Suppose that $F(x,y)$ is either Gâteaux differentiable at $(\bar{x},\bar{y})$ or Lipschitz continuous near $(\bar{x},\bar{y})$, $f(x,y), g_i(x,y)$ $(i=1,2,\ldots,m)$, $h_j(x,y)$ $(j=1,2,\ldots,n)$ are jointly linear; $G_i(x,y)$ $(i=1,2,\ldots,I)$ are Gâteaux differentiable and M-P pseudoconcave at $(\bar{x},\bar{y})$; and $H_j(x,y)(j=1,2,\ldots,J)$ are Gâteaux differentiable and M-P pseudoaffine at $(\bar{x},\bar{y})$. Then there exist scalars $\lambda \geq 0, \alpha_i \geq 0, \nu_i \geq 0$ $(i=1,2,\ldots,m)$, β_j, π_j $(j=1,2,\ldots,n)$, $\gamma_i \geq 0$ $(i=1,2,\ldots,I)$, η_j $(j=1,2,\ldots,J)$ such that*

$$\begin{aligned}
&0 \in \partial^{\diamond} F(\bar{x},\bar{y}) + \sum_{i=1}^{m} (\alpha_i - \lambda\nu_i) Dg_i(\bar{x},\bar{y}) + \sum_{j=1}^{n} \beta_j Dh_j(\bar{x},\bar{y}) \\
&\quad + \sum_{i=1}^{I} \gamma_i DG_i(\bar{x},\bar{y}) + \sum_{j=1}^{J} \eta_j DH_j(\bar{x},\bar{y}), \\
&0 = D_y f(\bar{x},\bar{y}) + \sum_{i=1}^{m} \nu_i D_y g_i(\bar{x},\bar{y}) + \sum_{j=1}^{n} \pi_j D_y h_j(\bar{x},\bar{y}), \\
&\nu_i g_i(\bar{x},\bar{y}) = 0, \quad i=1,2,\ldots,m, \\
&\alpha_i g_i(\bar{x},\bar{y}) = 0, \quad i=1,2,\ldots,m, \\
&\gamma_i G_i(\bar{x},\bar{y}) = 0, \quad i=1,2,\ldots,I.
\end{aligned}$$

Proof. By Theorem 4.2 and the weak reverse convex CQ, under the assumptions of the corollary it suffices to prove that the convex cone generated by the set

$$A := [Df(\bar{x},\bar{y}) - \partial V(\bar{x}) \times \{0\}] \cup \bigcup_{i=1}^{m} \{Dg_i\} \cup \bigcup_{j=1}^{n} \{\pm Dh_j\} \cup \bigcup_{i=1}^{I} \{DG_i\} \cup \bigcup_{j=1}^{J} \{\pm DH_j\}$$

is closed, where the dependence of the derivatives on $(\bar{x},\bar{y})$ is omitted whenever there is no confusion.

Since the lower level problem is linear, by [8, Proposition 2.13] (which obviously holds in any general Banach space as well), the value function $V(x)$ is a polyhedral convex function, which implies by [22, Theorem 23.10] (which obviously holds in any general Banach space as well) that $\partial V(\bar{x})$ is a polyhedral convex set. Since by the assumptions for the problem (SP) $V(x)$ is finite and convex on X, $\partial V(\bar{x})$ is bounded and hence $Df(\bar{x},\bar{y}) - \partial V(\bar{x}) \times \{0\}$ is a bounded polyhedral convex set. Therefore, by definition, $Df(\bar{x},\bar{y}) - \partial V(\bar{x}) \times \{0\}$ is a convex hull of a finite set of points. Consequently, the convex hull of the set $[A \cup \{0\}]$ is a polyhedral convex set containing the origin. By [22, Corollary 19.7.1] (which also holds in any general Banach space) the convex cone generated by $\text{co}[A \cup \{0\}]$ is polyhedral. But the convex cone generated by A is the same as the convex cone generated by $\text{co}[A \cup \{0\}]$, so it is also polyhedral and hence closed. The proof of the corollary is therefore complete. □

It is interesting to compare the value function approach with the classical approach, in which the lower level problem is replaced by the KKT condition of the lower level problem. Suppose the KKT condition is necessary and sufficient for optimality of the lower level problem; then BP is equivalent to the following single level optimization problem:

$$
\text{(KP)} \qquad
\begin{aligned}
\min_{x,y,\nu,\pi} \quad & F(x,y) \\
\text{s.t.} \quad & 0 = D_y f(x,y) + \sum_{i=1}^{m} \nu_i D_y g_i(x,y) + \sum_{j=1}^{n} \pi_j D_y h_j(x,y), \\
& \sum_{i=1}^{m} \nu_i g_i(x,y) \geq 0 \\
& \nu_i \geq 0, \quad g_i(x,y) \leq 0, \quad i = 1,2,\ldots,m, \\
& h_j(x,y) = 0, \quad j = 1,2,\ldots,n, \\
& G_i(x,y) \leq 0, \quad i = 1,2,\ldots,I, \\
& H_j(x,y) = 0, \quad j = 1,2,\ldots,J.
\end{aligned}
$$

It is known [30, Proposition 1.1] that usual CQs such as the Mangasarian–Fromovitz CQ do not hold for problem (KP). However, if a suitable CQ is satisfied and the functions f, g_i $(i = 1,\ldots,I)$, h_j $(j = 1,\ldots,J)$ are second order continuously differentiable, then the KKT condition for problem (KP) is the existence of scalars $\lambda \geq 0, \mu, \alpha_i \geq 0, \nu_i \geq 0$ $(i = 1,2,\ldots,m)$, β_j, π_j $(j = 1,2,\ldots,n)$, $\gamma_i \geq 0$ $(i = 1,2,\ldots,I)$, η_j $(j = 1,2,\ldots,J)$ such that

$$
\begin{aligned}
0 \in\ & \partial^{\diamond} F(\bar{x},\bar{y}) + \sum_{i=1}^{m} (\alpha_i - \lambda\nu_i) D g_i(\bar{x},\bar{y}) + \sum_{j=1}^{n} \beta_j D h_j(\bar{x},\bar{y}) \\
& + \sum_{i=1}^{I} \gamma_i \partial^{\diamond} G_i(\bar{x},\bar{y}) + \sum_{j=1}^{J} \eta_j \partial^{\diamond} H_j(\bar{x},\bar{y}) \\
& + \mu D\left[D_y f + \sum_{i=1}^{m} \nu_i D_y g_i + \sum_{j=1}^{n} \pi_j D_y h_j \right](\bar{x},\bar{y}), \\
0 =\ & D_y f(\bar{x},\bar{y}) + \sum_{i=1}^{m} \nu_i D_y g_i(\bar{x},\bar{y}) + \sum_{j=1}^{n} \pi_j D_y h_j(\bar{x},\bar{y}), \\
& \nu_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m, \\
& \alpha_i g_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,m, \\
& \gamma_i G_i(\bar{x},\bar{y}) = 0, \quad i = 1,2,\ldots,I.
\end{aligned}
$$

Comparing the KKT condition for (KP) with the KKT condition for (SP) in Theorem 4.2, it is easy to see that if the value function is convex, then the fact that the KKT condition holds for problem (SP) implies that the KKT condition for problem

(KP) holds with $\mu = 0$, and in the case when the lower level problem is linear, the KKT condition for (SP) coincides with the KKT condition for problem (KP). This establishes the relationship between the two approaches. Hence the nondifferentiable Arrow–Hurwicz–Uzawa CQ is also an applicable CQ for problem (KP).

Acknowledgments. I would like to thank Jean-Paul Penot for pointing out an error in an earlier version and for many helpful conversations and suggestions about the paper. I would also like to thank the anonymous referees for their extremely valuable comments and suggestions. In particular, I thank one of the referees for suggesting the current version of the statement and the proof of Theorem 3.1.

REFERENCES

[1] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser Boston, Boston, 1990.

[2] J. M. Abadie, *On the Kuhn-Tucker theorem*, in Nonlinear Programming, J. Abadie, ed., John Wiley, New York, 1967, pp. 21–36.

[3] K. J. Arrow, L. Hurwicz, and H. Uzawa, eds., *Studies in Linear and Nonlinear Programming*, Stanford University Press, Stanford, CA, 1958.

[4] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming Theory and Algorithms*, 2nd ed., John Wiley, New York, 1993.

[5] J. R. Birge and L. Qi, *Semiregularity and generalized subdifferentials with applications to optimization*, Math. Oper. Res., 18 (1993), pp. 982–1005.

[6] J. M. Borwein, J. S. Treiman, and Q. J. Zhu, *Necessary conditions for constrained optimization problems with semicontinuous and continuous data*, Trans. Amer. Math. Soc., 350 (1998), pp. 2409–2429.

[7] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York, 1983.

[8] A. V. Fiacco and J. Kyparisis, *Convexity and concavity properties of the optimal value function in parametric nonlinear programming*, J. Optim. Theory Appl., 48 (1986), pp. 95–126.

[9] H. Halkin, *Implicit functions and optimization problems without continuous differentiability of the data*, SIAM J. Control, 12 (1974), pp. 229–236.

[10] M. S. Gowda and M. Teboulle, *A comparison of constraint qualifications in infinite-dimension convex programming*, SIAM J. Control Optim., 28 (1990), pp. 925–935.

[11] A. D. Ioffe, *Necessary conditions in nonsmooth optimization*, Math. Oper. Res., 9 (1984), pp. 159–189.

[12] A. D. Ioffe, *A Lagrange multiplier rule with small convex-valued subdifferentials for nonsmooth problems of mathematical programming involving equality and nonfunctional constraints*, Math. Programming, 58 (1993), pp. 137–145.

[13] J. Jahn, *Introduction to the Theory of Nonlinear Optimization*, Springer, Berlin, 1996.

[14] A. Jourani, *Constraint qualifications and Lagrange multipliers in nondifferentiable programming problems*, J. Optim. Theory Appl., 81 (1994), pp. 533–548.

[15] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, in Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed., University of California Press, Berkeley, CA, 1951, pp. 481–492.

[16] O. L. Mangasarian, *Nonlinear Programming*, McGraw-Hill, New York, 1969; reprinted as Classics. Appl. Math. 10, SIAM, Philadelphia, 1994.

[17] P. Michel and J.-P. Penot, *Calcul sous-différentiel pour des fonctions lipschitziennes et non lipschitziennes*, C. R. Acad. Sci. Paris Sér. I Math., 12 (1984), pp. 269–272.

[18] P. Michel and J.-P. Penot, *A generalized derivative for calm and stable functions*, Differential Integral Equations, 5 (1992), pp. 433–454.

[19] B. S. Mordukhovich, *On necessary conditions for an extremum in nonsmooth optimization*, Soviet Math. Dokl., 283 (1985), pp. 215–220.

[20] H. V. Ngai and M. Théra, *A fuzzy necessary optimality condition for non-Lipschitz optimization in Asplund spaces*, SIAM J. Optim., 12 (2002), pp. 656–668.

[21] J.-P. Penot, *Are generalized derivatives useful for generalized convex functions?*, in Generalized Convexity, Generalized Monotonicity: Recent Results, J. P. Crouzeix et al., eds., Kluwer Academic, Dordrecht, The Netherlands, 1998, pp. 3–59.

[22] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[23] J. S. Treiman, *Shrinking generalized gradients*, Nonlinear Analysis, 12 (1988), pp. 1429–1450.

[24] J. S. Treiman, *Lagrange multipliers for nonconvex generalized gradients with equality, inequality, and set constraints*, SIAM J. Control Optim., 37 (1999), pp. 1313–1329.
[25] Z. Wu and J. J. Ye, *Some results on integration of subdifferentials*, Nonlinear Anal., 39 (2000), pp. 955–976.
[26] Z. Wu and J. J. Ye, *Equivalence among various derivatives and subdifferentials of the distance function*, J. Math. Anal. Appl., 282 (2003), pp. 629–647.
[27] J. J. Ye, *Multiplier rules under mixed assumptions of differentiability and Lipschitz continuity*, SIAM J. Control Optim., 39 (2001), pp. 1441–1460.
[28] J. J. Ye and D. L. Zhu, *Optimality conditions for bilevel programming problems*, Optimization, 33 (1995), pp. 9–27.
[29] J. J. Ye and D. L. Zhu, *A note on optimality conditions for bilevel programming problems*, Optimization, 39 (1997), pp. 361–366.
[30] J. J. Ye, D. L. Zhu, and Q. J. Zhu, *Exact penalization and necessary optimality conditions for generalized bilevel programming problems*, SIAM J. Optim., 7 (1997), pp. 481–507.
[31] C. Zălinescu, *A comparison of constraint qualifications in infinite-dimensional convex programming revisited*, J. Austral. Math. Soc. Ser. B, 40 (1999), pp. 353–378.
[32] W. I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewood Cliffs, NJ, 1969.