MATHEMATICS OF OPERATIONS RESEARCH
Vol. 36, No. 1, February 2011, pp. 165–184
ISSN 0364-765X | EISSN 1526-5471 | 11 | 3601 | 0165

DOI 10.1287/moor.1100.0480

Necessary Optimality Conditions for Multiobjective Bilevel Programs

Jane J. Ye
Department of Mathematics and Statistics, University of Victoria, Victoria, British Columbia V8W 3R4, Canada,
janeye@uvic.ca, http://www.math.uvic.ca/faculty/janeye/index.html

The multiobjective bilevel program is a sequence of two optimization problems, with the upper-level problem being multiobjective and the constraint region of the upper level problem being determined implicitly by the solution set to the lower-level problem. In the case where the Karush-Kuhn-Tucker (KKT) condition is necessary and sufficient for global optimality of all lower-level problems near the optimal solution, we present various optimality conditions by replacing the lower-level problem with its KKT conditions. For the general multiobjective bilevel problem, we derive necessary optimality conditions by considering a combined problem, with both the value function and the KKT condition of the lower-level problem involved in the constraints. Most results of this paper are new, even for the case of a single-objective bilevel program, the case of a mathematical program with complementarity constraints, and the case of a multiobjective optimization problem.

Key words: multiobjective optimization; preference; necessary optimality condition; partial calmness; constraint qualification; nonsmooth analysis; value function; bilevel programming problem
MSC2000 subject classification: Primary: 90C29, 90C46; secondary: 90C26, 90C33
OR/MS subject classification: Primary: programming/complementarity; secondary: programming/multiple criteria
History: Received March 4, 2010; revised August 28, 2010.

1. Introduction. Let W be a finite-dimensional Banach space and let $\prec$ be a (nonreflexive) preference for vectors in W. We consider the following multiobjective bilevel programming problem (BLPP):

$$\text{(BLPP)}\quad \min_{x,y}\ F(x,y) \quad \text{s.t. } y\in S(x),\quad G(x,y)\le 0,$$

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

$$(P_x):\quad \min_{y}\ f(x,y) \quad \text{s.t. } g(x,y)\le 0$$

and F: $R^n\times R^m\to W$, f: $R^n\times R^m\to R$, G: $R^n\times R^m\to R^q$, g: $R^n\times R^m\to R^p$. We allow p or q to be zero to signify the case in which there are no explicit inequality constraints. In these cases it is clear that certain references to such constraints are simply to be deleted.

We say that (x,y) is a feasible point for problem (BLPP) if $y\in S(x)$ and $G(x,y)\le 0$. We say that $(\bar x,\bar y)$ is a local solution to (BLPP), provided that it is a feasible point for (BLPP) and there exists no other feasible point (x,y) in the neighborhood of $(\bar x,\bar y)$ such that $F(x,y)\prec F(\bar x,\bar y)$.

The preference we use in this paper is very general. It includes the following commonly used preferences as special cases. Let K be a closed cone in W. The preference relation for two vectors $x,y\in W$ is defined by $x\prec y$ if and only if $x-y\in K$ and $x\neq y$. We call such a preference a generalized Pareto preference. In particular, if $W=R^N$ and $K=R^N_-:=\{z\in R^N\colon z$ has nonpositive components$\}$, then the generalized Pareto preference in this case is simply called the Pareto preference. When $F(x,y)$ is a scalar function, the preference is $<$; the problem becomes a single-objective BLPP.

In this paper we assume that all preferences are closed at any $F(\bar x,\bar y)$, where $(\bar x,\bar y)$ is a local minimum of (BLPP) as defined below.

Definition 1.1 (Mordukhovich [18, Definition 5.55]). Let $l(r):=\{t\in W\colon t\prec r\}$ denote the level set at $r\in W$ with respect to the given preference $\prec$. We say that a preference $\prec$ is closed around $\bar r\in W$ provided that:

(H1) It is locally satiated around $\bar r$; i.e., for any $r\in W$, $r\in \overline{l(r)}$ for all r in some neighborhood of $\bar r$, where $\overline{l(r)}$ denotes the closure of the level set $l(r)$.

(H2) It is almost transitive on W; i.e., $t\in\overline{l(r)}$ and $r\prec s$, implies $t\prec s$.

The local satiation propety holds for any reasonable preference, and the almost transitivity requirement also holds for many preferences. For example, both (H1) and (H2) hold for the generalized Pareto preference when the closed cone K is convex and pointed (Mordukhovich [18, Proposition 5.56]). However, the almost transitive property may be restrictive in some applications. For example, it is known that the preference described by the lexicographical order is not almost transitive (see Mordukhovich [18, Example 5.57]). The reader is referred to a recent paper (Bao and Mordukhovich [1]) for results concerning the set-valued optimization in welfare economics, where the preference is not almost transitive.

The bilevel programming has been an important research area, and many researchers have made contributions. The origin of the BLPP can be traced back to von Stackelberg [23], who used it to model the market economy in 1934. BLPP has been successfully used to model the so-called leader-follower game or the moral hazard model of the principal-agent problem in political science and economics (see e.g., Mirrlees [16]). The reader is referred to several monographs (Bard [3], Dempe [6], Shimizu et al. [21]) for more applications of bilevel programming and to a bibliography review (Dempe [7], Vicente and Calamai [22]).

The classical Karush-Kuhn-Tucker (KKT) approach (also called the first-order approach) to solve a single-objective BLPP is to replace the lower-level problem with its KKT condition and solve the resulting mathematical programming problem with equilibrium constraints (MPEC). For MPECs, it is well known that the usual nonlinear programming constraint qualification—the Mangasarian-Fromovitz constraint qualification (MFCQ)—does not hold (see Ye et al. [34, Proposition 1.1]). Because MFCQ is a standard constraint qualification and a standard assumption for many numerical algorithms to work, the classical KKT condition may not hold, and the classical numerical algorithms may fail if we treat an MPEC as a standard nonlinear programming problem with equality and inequality constraints. By reformulating MPECs in different ways, various alternative stationary concepts such as Clarke, Mordukhovich, Strong, and Bouligand (also known as Piecewise) (C-, M-, S-, B-(P-)) stationary points arise (see, e.g., Scheel and Scholtes [20], Ye [28]), and constraint qualifications under which a local optimal solution of MPEC is a stationary point in the various sense have been given (see, e.g., Luo et al. [15], Ye [28]).

By using the KKT approach, one would hope to find candidates for optimal solutions of BLPP. This, however, may not always possible. Even for the case where the lower-level problem is convex, a recent paper of Dempe and Dutta [8] gives an example of BLPP with a convex lower level that has a local solution to the corresponding MPEC, whose (x, y) components are not a local solution to the original BLPP. Therefore, for the general case of BLPP with a not necessarily convex lower-level problem, there may not even exist any relationships between the original bilevel program and its KKT reformulation. To clarify this point, let us examine the following simple example taken from Ye and Zhu [32, Example 4.3].

$$\begin{aligned}(P)\quad &\min\quad (x-0.5)^2+(y-2)^2\\ &\text{s.t.}\quad y\in S(x):=\arg\min_y\{y^3-3y\colon\, y\ge x-3\},\\ &\qquad\ 0\le x\le 4.\end{aligned}$$

It is easy to verify that the set of optimal solution for the lower-level problem is

$$S(x)=\begin{cases}\{x-3\} & \text{if } 0\le x<1,\\ \{-2,1\} & \text{if } x=1,\\ \{1\} & \text{if } 1<x\le 4,\end{cases}$$

and $(\bar{x},\bar{y})=(1,1)$ is the unique solution to the the bilevel program (P). Replacing the lower-level problem with its KKT condition, we get the following problem:

$$\begin{aligned}\min\quad &(x-0.5)^2+(y-2)^2\\ \text{s.t.}\quad &3y^2-3-\lambda=0,\\ &x-3-y\le 0, &(1)\\ &\lambda\ge 0,\quad (x-3-y)\lambda=0,\\ &0\le x\le 4. &(2)\end{aligned}$$

At $(\bar{x},\bar{y})=(1,1)$, because the constraints (1)–(2) are not binding, the KKT condition would imply the existence of a real number u such that

$$\begin{pmatrix}0\\0\end{pmatrix}=2\begin{pmatrix}\bar{x}-0.5\\ \bar{y}-2\end{pmatrix}+\begin{pmatrix}0\\ 6\bar{y}\end{pmatrix}u.$$

But this is impossible. This example reveals a striking fact: *the optimal solution of the original BLPP may not even be a stationary point of the resulting single-level problem by the KKT approach*! Therefore, if the KKT approach is not used properly, the true optimal solution of the bilevel program may be missed completely!

In Ye and Zhu [30, 31], the following value function approach is taken to reformulate the BLPP. Define the *value function* of the lower-level problem as an extended valued function $V: R^n \to \bar{R}$ by

$$V(x) := \inf_y \{f(x,y): g(x,y) \le 0\},$$

where $\bar{R} := R \cup \{-\infty\} \cup \{+\infty\}$ is the extended real line and $\inf\{\emptyset\} = +\infty$ by convention. Then it is obvious that problem (BLPP) can be reformulated as the following problem involving the value function:

$$\begin{aligned}\text{(VP)}\quad \min_{x,y}\ & F(x,y)\\ \text{s.t.}\ & f(x,y) - V(x) \le 0,\\ & g(x,y) \le 0,\\ & G(x,y) \le 0.\end{aligned}$$

It is easy to think that because the reformulation (VP) is exactly equivalent to the original BLPP, the problem will be solved if the nonsmooth necessary optimality condition is used on the problem (VP). The problem turns out to be not so simple, because the nonsmooth MFCQ does not hold at any feasible solution of the problem (VP), so the KKT condition may not hold. To deal with this difficulty, Ye and Zhu [30, 31] proposed the partial calmness condition. The value function approach was further developed in Ye [27, 29] using other constraint qualifications, such as the Abadie constraint qualification. For the case where the value function is convex, Ye [27, 29] showed that the resulting KKT condition takes a simpler form in which only one solution of the lower-level optimal problem is involved. Under the partial calmness condition, this simpler KKT condition was proved to hold under the assumption of inner semicontinuity of the solution mapping of the lower-level program (Dempe et al. [9]).

In a recent paper (Ye and Zhu [32]), Ye and Zhu observed that the partial calmness condition may be too strong for many nonconvex BLPPs. For our simple problem (P), the value function of the lower-level problem can be easily derived:

$$V(x) = \begin{cases} -2 & \text{if } 1 \le x \le 4\\ (x-3)^3 - 3(x-3) & \text{if } 0 \le x \le 1.\end{cases}$$

By using the value function, the bilevel program (P) is obviously equal to

$$\begin{aligned}\min\ & (x-0.5)^2 + (y-2)^2\\ \text{s.t.}\ & y^3 - 3y - V(x) \le 0,\\ & y \ge x - 3,\\ & 0 \le x \le 4.\end{aligned}$$

The KKT condition (in the component y) for the above problem would imply the existence of a nonnegative number u such that

$$0 = 2(\bar{y} - 2) + (3\bar{y}^2 - 3\bar{x})u,$$

which is impossible for $(\bar{x}, \bar{y}) = (1, 1)$ to hold. Therefore, the value function approach is not useful for this problem either.

Our simple example demonstrates that neither the KKT nor the value function approach is applicable. To cope with this difficulty, Ye and Zhu [32] suggest that a combination of the classical KKT and the value function approach should be taken in this case. For our simple problem (P), the combined approach means that we add the KKT condition for the lower-level problem into the constraints of the problem (VP) and consider the following combined problem:

$$\begin{aligned}\min\ & (x-0.5)^2 + (y-2)^2\\ \text{s.t.}\ & y^3 - 3y - V(x) \le 0,\end{aligned}$$

$$3y^2-3-\lambda=0,$$
$$x-3-y\le 0,$$
$$\lambda\ge 0,\quad (x-3-y)\lambda=0,$$
$$0\le x\le 4.$$

At first glance, it seems that the KKT condition for the lower-level problem is superfluous. However, the resulting necessary optimality condition derived from such a combined problem is much more likely to hold because now there are multipliers corresponding to both the value function constraint and the KKT condition constraints, which provide more freedom to choose the multipliers. In the case of the multiplier corresponding to the value function being zero, the approach reduces to the KKT approach. In the case where the multiplier corresponding to the KKT condition is zero, the approach reduces to the value function approach. For our simple problem (P), it was shown in Ye and Zhu [32, Example 4.3] that the the nonsmooth multiplier rule holds for the combined problem, with both multipliers corresponding to the value function constraint and the KKT condition of the lower-level problem nonzero. Hence, both the classical KKT approach and the value function approach fail, and only the combined approach works for problem (P).

Various concepts of stationary conditions have their own usages. The S-stationary condition is known to be equivalent to the classical stationary condition and is hence the sharpest of all. However, it requires a very strong constraint qualification. In the case when the S-stationary condition does not hold, the M-stationary condition is the next sharpest condition, and it holds under relatively weak conditions. In particular, the M-stationary condition is very useful in the sensitivity analysis (see Lucet and Ye [13, 14]). C- and P-type stationary conditions are usually weaker, but many numerical algorithms converge to them. Thus, it is important to study all concepts of stationary conditions.

Note that in the single-objective bilevel programming paper (Ye and Zhu [32]), to concentrate on the main idea of the combined approach, the C-, M-type stationary conditions were left out, and in a multiobjective bilevel programming paper (Ye and Zhu [33]) the C-, S-, and P-stationary conditions were not studied for the KKT approach and the combined approach has not been taken to study the general problem. To fill this gap in this paper, we use the combined approach introduced in Ye and Zhu [32] to derive various C-, M-, S-, and P-stationary conditions for the multiobjective BLPP.

In Mordukhovich [18, §5.3], necessary optimality conditions for a class of multiobjective MPECs with an alternative criteria of optimality called *the generalized order optimality* have been derived. Results of this paper may be similarly extended to this class of multiobjective MPECs using the results of Mordukhovich [18, §5.3].

We organize the paper as follows. In the next section we provide the notations and the background materials on variational analysis to be used throughout the paper. In this section we also introduce the concepts of C-, S-, and P-stationary conditions for multiobjective MPECs and provide constraint qualifications under which a local optimal solution to the multiobjective MPEC is C-, S-, and P-stationary points. In §3 we concentrate on the KKT approach, and in §4 we use the combined approach to study a general BLPP.

2. Preliminaries and preliminary results. In this paper we adopt the following standard notation. For any two vectors a, b in a finite-dimensional Banach spaces Z, we denote by $\langle a, b\rangle$ the inner product. Given a function $F\colon R^n\to R^m$, we denote its Jacobian by $\nabla F(z)\in R^{m\times n}$. If $m=1$, the gradient $\nabla F(z)\in R^n$ is considered as a column vector. For a subset $A\subseteq R^n$, we denote by $\operatorname{int} A$, $\bar{A}$, $\operatorname{co} A$ the interior, the closure, and the convex hull of A, respectively. For a matrix $A\in R^{n\times m}$, A^T is its transpose.

2.1. Background in variational analysis. We present some background materials on variational analysis that will be used throughout the paper. Detailed discussions on these subjects can be found in Clarke [4]; Clarke et al. [5]; Mordukhovich[17, 18]; and Rockafellar and Wets [19].

Definition 2.1 (Normal Cones). Let Ω be a nonempty subset of a finite-dimensional space Z. Given $z\in\Omega$, the convex cone

$$N^{\pi}(z;\Omega):=\{\zeta\in Z\colon \exists\sigma>0,\ \text{such that } \langle\zeta, z'-z\rangle\le\sigma\|z'-z\|^2\ \forall z'\in\Omega\}$$

is called the *proximal normal cone* to set Ω at point z; the closed cone

$$N(z;\Omega)=\left\{\lim_{k\to\infty}\zeta_k\colon \zeta_k\in N^{\pi}(z_k;\Omega),\quad z_k\in\Omega,\ z_k\to z\right\}$$

is called the *limiting normal cone* (also known as Mordukhovich normal cone or basic normal cone) to Ω at point z. The *Clarke normal cone* can be obtained by taking the closure of the convex hull of the limiting normal cone, i.e.,

$$N^c(z;\Omega)=\overline{\text{co}}\,N(z;\Omega).$$

Note that alternatively the Fréchet (also called regular) normal cone (see Mordukhovich [17, Definition 1.1(ii)]) can be used to construct the limiting normal cone because the two definitions coincide in the finite-dimensional space (see Mordukhovich [17, Commentary to Chap. 1] or Rockafellar and Wets [19, p. 345] for a discussion). When Ω is convex, the proximal normal cone, the limiting normal cone, and the Clarke normal cone coincide with the normal cone in the sense of the convex analysis, i.e.,

$$N^\pi(z;\Omega)=N^c(z;\Omega)=N(z;\Omega)=\{\zeta\in Z\colon \langle\zeta,z'-z\rangle\le 0,\ \forall z'\in\Omega\}.$$

Definition 2.2 (Limiting Normal Cones to Moving Sets Mordukhovich [18, Definition 5.69]). Let $S\colon Z\rightrightarrows W$ be a set-valued mapping from a finite-dimensional space Z into another finite-dimensional space W, and let $(r,z)\in gphS$, where $gphS=\{(\gamma,z)\colon z\in S(\gamma)\}$ denotes the graph of the set-valued map S. Then

$$N_+(z;S(r)):=\left\{\lim_{k\to\infty}\zeta_k\colon \zeta_k\in N^\pi(z_k;S(r_k)),\ \ z_k\in S(r_k),\, z_k\to z,\, r_k\to r\right\}$$

is the extended normal cone to $S(r)$ at z. The mapping S is normally semicontinuous at (r,z) if

$$N_+(z;S(r))=N(z;S(r)).$$

Definition 2.3 (Clarke Generalized Gradients). Let $f\colon R^n\to R$ be Lipschitz continuous near $\bar x$. The Clarke generalized directional derivative of f at $\bar x$ in direction $d\in R^n$ is defined by

$$f^\circ(\bar x;d):=\limsup_{x\to\bar x,\,t\downarrow 0}\frac{f(x+td)-f(x)}{t},$$

and the Clarke generalized gradient at $\bar x$ is a convex and compact subset of R^n defined by

$$\partial^c f(\bar x):=\{\xi\in R^n\colon \langle\xi,d\rangle\le f^\circ(\bar x;d)\ \ \forall d\in R^n\}.$$

Definition 2.4 (Limiting Subdifferential). Let $f\colon R^n\to\bar R$ be a lower semicontinuous function and finite at $\bar x\in R^n$. The proximal subdifferential (Rockafellar and Wets [19, Definition 8.45]) of f at $\bar x$ is defined as

$$\partial^\pi f(\bar x):=\{\zeta\in R^n\colon \exists\sigma>0,\delta>0\ \text{ such that } f(x')\ge f(\bar x)+\langle\zeta,x'-\bar x\rangle-\sigma\|x'-\bar x\|^2\ \forall x'\in B(\bar x,\delta)\},$$

and the limiting (Mordukhovich or basic Mordukhovich [17]) subdifferential of f at $\bar x$ is defined as

$$\partial f(\bar x):=\left\{\lim_{k\to\infty}\xi_k\colon \xi_k\in\partial^\pi f(x_k),\ \ x_k\to\bar x,\ f(x_k)\to f(\bar x)\right\}.$$

When f is Lipschitz continuous near $\bar x$, the Clarke generalized gradient can be obtained by taking the convex hull of the limiting subdifferential, i.e.,

$$\partial^c f(\bar x)=\text{co}\,\partial f(\bar x).$$

The following calculation rules for Clarke generalized gradients will be useful in the paper.

Proposition 2.1 (See Clarke [4], Clarke et al. [5]). *Let $f,g\colon R^n\to R$ be Lipschitz continuous near $\bar x\in R^n$ and α,β be any real numbers. Then*

$$\partial^c(\alpha f+\beta g)(\bar x)\subseteq\alpha\partial^c f(\bar x)+\beta\partial^c g(\bar x).$$

Note that for limiting subdifferentials, in general, the above calculation rule holds only when α and β are nonnegative.

2.2. Necessary optimality conditions for MPECs. In this subsection we consider the multiobjective MPEC defined as follows:

$$\begin{aligned} \text{MPEC} \quad \min \ & f(z) \\ \text{s.t.} \ & g(z) \le 0, \quad h(z) = 0 \\ & 0 \le G(z) \perp H(z) \ge 0, \end{aligned}$$

where W is a finite-dimensional Banach space, f: $R^n \to W$, G, H: $R^n \to R^m$, g: $R^n \to R^p$, h: $R^n \to R^q$, and $a \perp b$ means that the vector a is perpendicular to vector b. For simplicity and easy reference in this section, we assume that f is Lipschitz near z^* and all other functions are continuously differentiable.

Given a feasible vector z^* of MPEC, we define the following index sets:

$$\begin{aligned} I_g &:= I_g(x^*) = \{i\colon g_i(z^*) = 0\}, \\ \alpha &:= \alpha(z^*) = \{i\colon G_i(z^*) = 0, H_i(z^*) > 0\}, \\ \beta &:= \beta(z^*) = \{i\colon G_i(z^*) = 0, H_i(z^*) = 0\}, \\ \gamma &:= \gamma(z^*) = \{i\colon G_i(z^*) > 0, H_i(z^*) = 0\}. \end{aligned}$$

Definition 2.5 (MPEC Stationary Conditions). A feasible point z^* of MPEC is called a Clarke stationary point (C-stationary point) if there exists a unit vector $\lambda \in N_+(f(z^*); \overline{l(f(z^*))})$ and $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that the following conditions hold:

$$0 \in \partial\langle \lambda, f\rangle(z^*) + \sum_{i\in I_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)], \tag{3}$$

$$\begin{gathered} \lambda_i \ge 0 \quad i \in I_g, \qquad \lambda_i^G = 0 \quad i \in \gamma, \qquad \lambda_i^H = 0 \quad i \in \alpha, \\ \lambda_i^G \lambda_i^H \ge 0 \quad i \in \beta. \end{gathered} \tag{4}$$

A feasible point z^* of MPEC is called a Mordukhovich stationary point (M-stationary point) if there exists a unit vector $\lambda \in N_+(f(z^*); \overline{l(f(z^*))})$ and $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that (3)–(4) and the following condition holds:

$$\text{either} \quad \lambda_i^G > 0, \ \lambda_i^H > 0 \quad \text{or} \quad \lambda_i^G \lambda_i^H = 0 \qquad \forall\, i \in \beta.$$

A feasible point z^* of MPEC is called a strong stationary point (S-stationary point) if there exists a unit vector $\lambda \in N_+(f(z^*); \overline{l(f(z^*))})$ and $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that (3)–(4) and the following condition holds:

$$\lambda_i^G \ge 0, \ \lambda_i^H \ge 0 \qquad \forall\, i \in \beta.$$

A feasible point z^* of MPEC is called a piecewise stationary point (P-stationary point) if for each partition of the index set β into P, Q, there exists a unit vector $\lambda \in N_+(f(z^*); \overline{l(f(z^*))})$ and $(\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that (3)–(4) and the following condition holds:

$$\lambda_i^G \ge 0 \quad \forall\, i \in P, \qquad \lambda_i^H \ge 0 \quad \forall\, i \in Q.$$

Remark 2.1. In the case where $\prec$ is the Pareto preference, by Zhu [35] the preference $\prec$ is regular in the sense that it is closed and the set-valued mapping $S(z) := \overline{l(f(z))}$ is normally semicontinuous; moreover

$$N_+(f(z^*); \overline{l(f(z^*))}) = N(f(z^*); \overline{l(f(z^*))}) = R_+^N$$

and in the case where $\prec$ is the generalized Pareto preference with a closed cone K,

$$N_+(f(z^*); \overline{l(f(z^*))}) = N(f(z^*); \overline{l(f(z^*))}) = K^- := \{s \in W\colon \langle s, t\rangle \le 0, \ t \in K\}.$$

Remark 2.2. As for single-level MPECs, it is not hard to show that the S-stationary condition is equivalent to the classical KKT condition for MPEC. For the special case of a single-level smooth MPEC, the P-stationary point is equivalent to a Bouligand stationary (B-stationary) point in the sense of Scheel and Scholtes [20] and is equivalent to a B-stationary point in the classical sense of Luo et al. [15] if a certain constraint qualification for each branch of the MPEC holds.

DEFINITION 2.6 (MPEC CONSTRAINT QUALIFICATIONS). Let z^* be a feasible point of MPEC. We say that the no nonzero abnormal C-multiplier constraint qualification (NNACMCQ) holds at z^* if there is no nonzero vector $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that

$$0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)],$$
$$\lambda_i \geq 0 \quad i \in I_g, \qquad \lambda_i^G = 0 \quad i \in \gamma, \qquad \lambda_i^H = 0 \quad i \in \alpha,$$
$$\lambda_i^G \lambda_i^H \geq 0 \quad \forall\, i \in \beta.$$

We say that the MPEC no nonzero abnormal multiplier constraint qualification (MPEC NNAMCQ) holds at z^* if there is no nonzero vector $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that

$$0 = \sum_{i \in I_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)],$$
$$\lambda_i \geq 0 \quad i \in I_g, \qquad \lambda_i^G = 0 \quad i \in \gamma, \qquad \lambda_i^H = 0 \quad i \in \alpha,$$
$$\text{either} \quad \lambda_i^G > 0,\ \lambda_i^H > 0, \quad \text{or} \quad \lambda_i^G \lambda_i^H = 0 \qquad \forall\, i \in \beta.$$

We say that the MPEC linear independence constraint qualification (MPEC LICQ) holds at z^* if the gradient vectors

$$\nabla g_i(z^*) \quad i \in I_g, \qquad \nabla h_i(z^*) \quad i = 1, \ldots, q, \qquad \nabla G_i(z^*) \quad i \in \alpha \cup \beta, \qquad \nabla H_i(z^*) \quad i \in \gamma \cup \beta$$

are linearly independent.

We say that the error bound constraint qualification holds at z^* if there exist positive constants μ, δ, and ε such that

$$d(z, \mathcal{F}) \leq \mu \|(\alpha, \beta, u, v)\| \quad \forall\, (\alpha, \beta, u, v) \in \varepsilon B$$
$$z \in \mathcal{F}(\alpha, \beta, u, v) \cap B_\delta(z^*),$$

where $d(z, \mathcal{F})$ is the distance of z to the feasible region $\mathcal{F}$ and

$$\mathcal{F}(\alpha, \beta, u, v) := \left\{ z\colon g(z) + \alpha \leq 0,\ h(z) + \beta = 0,\ 0 \leq G(z) + v \perp H(z) + u \geq 0 \right\}$$

is the perturbed feasible region of MPEC.

We say that the MPEC linear constraint qualification (MPEC linear CQ) holds if all functions G, H, g, h are affine. We say that MPEC piecewise MFCQ holds at z^* if MFCQ holds at z^* for each branch of MPEC corresponding to partition P, Q of index set β defined as

$$\begin{aligned} \text{MPEC}_{P \cup Q} \quad & \min \ f(z) \\ & \text{s.t.} \ G_i(z) = 0 \quad i \in \alpha, \qquad H_i(z) = 0 \quad i \in \gamma, \\ & \quad G_i(z) \geq 0, \quad H_i(z) = 0 \qquad i \in P, \\ & \quad G_i(z) = 0, \quad H_i(z) \geq 0 \qquad i \in Q, \\ & \quad g(z) \leq 0, \quad h(z) = 0. \end{aligned}$$

REMARK 2.3. By Ye [28, Proposition 2.1], MPEC NNAMCQ is equivalent to the MPEC generalized MFCQ (MPEC GMFCQ), a MPEC version of the MFCQ. We refer the reader to the definition of MPEC GMFCQ in Ye [28, Definition 2.11].

It is known that for a single-objective MPEC with smooth problem data, a local optimal solution of MPEC must be an S-stationary point under MPEC LICQ. The proof of the results used the fact that under the MPEC LICQ each branch of MPEC has a unique multiplier (see Ye [25]). But this proof cannot be used for our case because the objective function is only assumed to be Lipschitz continuous.

To derive the S-stationary condition under the MPEC LICQ, we will need the following result, which is also of independent interest.

Proposition 2.2. *Let $\mathcal{F}$ denote the feasible region of MPEC and $z^* \in \mathcal{F}$. Suppose that MPEC LICQ holds at z^*, and let ξ be an element of the normal cone $N(z^*; \mathcal{F})$. Then there exists $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H) \in R^{p+q+2m}$ such that*

$$\xi = \sum_{i\in I_g} \lambda_i^g \nabla g_i(z^*) + \sum_{i=1}^{q} \lambda_i^h \nabla h_i(z^*) - \sum_{i=1}^{m} [\lambda_i^G \nabla G_i(z^*) + \lambda_i^H \nabla H_i(z^*)],$$

$$\lambda_i^g \geq 0 \quad i \in I_g, \qquad \lambda_i^G = 0 \quad i \in \gamma, \qquad \lambda_i^H = 0 \quad i \in \alpha,$$

$$\lambda_i^G \geq 0, \quad \lambda_i^H \geq 0 \qquad i \in \beta.$$

Proof. By the definition of the limiting normal cone, $\xi = \lim_{k\to\infty} \xi_k$ with $\xi_k \in N^\pi(z_k; \mathcal{F})$, $z_k \in \mathcal{F}$, $z_k \to z^*$. By the definition of the proximal normal cone, there exists $M_k > 0$ such that

$$\langle \xi_k, z - z_k \rangle \leq M_k \|z - z_k\|^2 \quad \forall\, z \in \mathcal{F},$$

which implies that $z = z_k$ is a minimizer of the following problem:

$$\begin{aligned} \min\ & -\langle \xi_k, z\rangle + M_k\|z - z_k\|^2 \\ \text{s.t.}\ & z \in \mathcal{F}. \end{aligned}$$

The above problem is an MPEC with continuously differentiable problem data. Because MPEC LICQ holds at z^* and $z_k \to z^*$, MPEC LICQ holds at z_k as well. Therefore, z_k is an S-stationary point for the above MPEC. That is, there exists a unique multiplier $\lambda^k = (\lambda^{gk}, \lambda^{hk}, \lambda^{Gk}, \lambda^{Hk}) \in R^{p+q+2m}$ such that

$$\xi_k = \sum_{i\in I_g} \lambda_i^{gk} \nabla g_i(z_k) + \sum_{i=1}^{q} \lambda_i^{hk} \nabla h_i(z_k) - \sum_{i=1}^{m} [\lambda_i^{Gk} \nabla G_i(z_k) + \lambda_i^{Hk} \nabla H_i(z_k)],$$

$$\lambda_i^{gk} \geq 0 \quad i \in I_g^k, \qquad \lambda_i^{Gk} = 0 \quad i \in \gamma_k, \qquad \lambda_i^{Hk} = 0 \quad i \in \alpha_k,$$

$$\lambda_i^{Gk} \geq 0, \quad \lambda_i^{Hk} \geq 0 \qquad i \in \beta_k,$$

where

$$\begin{aligned} I_g^k &:= I_g(z_k) = \{i\colon\ g_i(z_k) = 0\}, \\ \alpha_k &:= \alpha(z_k) = \{i\colon\ G_i(z_k) = 0,\ H_i(z_k) > 0\}, \\ \beta_k &:= \beta(z_k) = \{i\colon\ G_i(z_k) = 0,\ H_i(z_k) = 0\}, \\ \gamma_k &:= \gamma(z_k) = \{i\colon\ G_i(z_k) > 0,\ H_i(z_k) = 0\}. \end{aligned}$$

Let $k \to \infty$ and $z_k \to z^*$; then by the MPEC LICQ, λ_k is bounded and hence there exists a convergent subsequence. Without loss of generality, assume that the limit of the sequence $\lambda^k = (\lambda^{gk}, \lambda^{hk}, \lambda^{Gk}, \lambda^{Hk})$ is $\lambda = (\lambda^g, \lambda^h, \lambda^G, \lambda^H)$. Taking limits as $k \to \infty$, because $z_k \to z^*$, $\xi_k \to \xi$, and $\lambda^k \to \lambda$, we have the desired conclusion. □

We are now in a position to develop the necessary optimality conditions for our multiobjective MPEC.

Theorem 2.1 (MPEC Necessary Optimality Conditions). *Let z^* be a local optimal solution for MPEC. Then the following statements are true.*

(I) *Under NNACMCQ, z^* is C-stationary.*

(II) *Under one of the following constraint qualifications, z^* is M-stationary*:

(i) *MPEC NNAMCQ (or equivalently MPEC GMFCQ) holds at z^*.*

(ii) *The MPEC linear CQ holds.*

(iii) *The error bound constraint qualification holds at z^*.*

(III) *If MPEC LICQ holds, then z^* is S-stationary.*

(IV) *If either MPEC linear CQ or MPEC piecewise MFCQ holds at z^*, then z^* is P-stationary.*

Proof. (I) It is easy to see that z^* is also a local solution of the following nonsmooth multiobjective nonlinear programming problem:

$$\begin{aligned}
\text{(MPEC)}\quad \min\ & f(z)\\
\text{s.t. }\ & G_i(z)=0 \quad i\in\alpha, \qquad H_i(z)=0 \quad i\in\gamma,\\
& \min\{G_i(z),H_i(z)\}=0 \quad i\in\beta,\\
& g(z)\le 0, \quad h(z)=0.
\end{aligned}$$

Note that from the proof of Ye and Zhu [33, Theorem 1.2], with the absence of the normal semicontinuity of the set-valued mapping $\overline{l(f(z))}$ (i.e., the preference is closed but not regular), it is easy to see that the Fritz John type necessary optimality condition in Ye and Zhu [33, Theorem 1.3] holds, with the limiting normal cone $N(f(z^*);\overline{l(f(z^*))})$ replaced by the extended normal cone $N_+(f(z^*);\overline{l(f(z^*))})$ (see also Mordukhovich [18, §5]).

By the Fritz John type necessary optimality condition in Ye and Zhu [33, Theorem 1.3] with $N(f(z^*);\overline{l(f(z^*))})$ replaced by the extended normal cone $N_+(f(z^*);\overline{l(f(z^*))})$ and the nonsmooth calculus rule for the nonsmooth function $\min\{G_i(z),H_i(z)\}$ (as in Scheel and Scholtes [20, Lemma 1]), we find that there exist $\mu_0\in\{0,1\}$, a unit vector $\lambda\in N_+(f(z^*);\overline{l(f(z^*))})$ and $(\lambda^g,\lambda^h,\lambda^G,\lambda^H)\in R^{p+q+2m}$, not all equal to zero, such that the following conditions hold:

$$0\in\mu_0\partial\langle\lambda,f\rangle(z^*)+\sum_{i\in I_g}\lambda_i^g\nabla g_i(z^*)+\sum_{i=1}^{q}\lambda_i^h\nabla h_i(z^*)-\sum_{i=1}^{m}[\lambda_i^G\nabla G_i(z^*)+\lambda_i^H\nabla H_i(z^*)],$$
$$\lambda_i\ge 0\quad i\in I_g,\qquad \lambda_i^G=0\quad i\in\gamma,\qquad \lambda_i^H=0\quad i\in\alpha,$$
$$\lambda_i^G\lambda_i^H\ge 0\quad i\in\beta.$$

By the virtue of NNAMCCQ, μ_0 can be taken as 1 and hence the conclusion for (I) follows.

(II) It is well known that MPEC NNAMCQ and the MPEC linear CQ both imply the error bound constraint qualification (see Ye [26, Theorems 4.3 and 4.4]). Hence it suffices to show (II)(iii). By virtue of Ye and Zhu [33, Theorem 1.3], treating the problem (MPEC) as the following optimization problem with an abstract constraint set:

$$\min f(z)\qquad \text{s.t. } z\in\mathcal{F},$$

we conclude that there exists a unit vector $\lambda\in N_+(f(z^*);\overline{l(f(z^*))})$ such that

$$0\in\partial\langle\lambda,f\rangle(z^*)+N(z^*;\mathcal{F}).$$

Rewrite the feasible region $\mathcal{F}$ as

$$\mathcal{F}=\{z\colon \varphi(z)\in Q\}$$

where $\varphi(z)=(g(z),h(z),G(z),H(z))$ and $Q=R_-^p\times\{0\}\times\Omega$ with $\Omega:=\{(x,y)\colon 0\le x\perp y\ge 0\}$. Because the error bound constraint qualification at z^* is equivalent to the calmness of the set-valued map $\mathcal{F}(\alpha,\beta,u,v)$ at $(0,z^*)$, using the recent result of Ioffe and Outrata [12, Proposition 3.4], we obtain

$$N(z^*;\mathcal{F})\subseteq\{\nabla\varphi(z^*)^Ty^*\colon y^*\in N(\varphi(z^*);Q)\}.$$

Because

$$N(\varphi(z^*);Q)=N(g(z^*);R_-^p)\times R^q\times N(G(z^*),H(z^*);\Omega)$$

and

$$N(G(z^*),H(z^*);\Omega)=\left\{(\lambda^G,\lambda^H)\colon \begin{array}{ll} \lambda_i^G=0 & \text{if } i\in\gamma\\ \lambda_i^H=0 & \text{if } i\in\alpha\\ \text{either } \lambda_i^G>0,\ \ \lambda_i^H>0,\ \text{ or }\ \lambda_i^G\lambda_i^H=0 & \text{if } i\in\beta\end{array}\right\}$$

(see, e.g., Ye [26, Proposition 3.7]), the desired assertion follows.

(III) By virtue of Ye and Zhu [33, Theorem 1.3], treating the problem (MPEC) as the following optimization problem with an abstract constraint set:

$$\min f(z) \qquad \text{s.t. } z \in \mathcal{F},$$

we conclude that there exists a unit vector $\lambda \in N_+(f(z^*); \overline{l(f(z^*))})$ such that

$$0 \in \partial\langle \lambda, f\rangle(z^*) + N(z^*; \mathcal{F}).$$

If MPEC LICQ holds at z^*, then by Proposition 2.2, we conclude that z^* is S-stationary.

(IV) It is easy to see that for each partition (P, Q) of the index set β, z^* is a local solution of the subproblem $\text{MPEC}_{P\cup Q}$. Hence, if either MPEC linear CQ or MPEC piecewise MFCQ holds at z^*, then z^* is P-stationary. □

Remark 2.4. All results in Theorem 2.1 are new except II(i), which was derived in Ye and Zhu [33, Theorem 4.2]. (II)(iii) and (II)(ii) have improved the results of Ye and Zhu [33, Theorems 4.8 and 4.9], respectively, in that the preference is not required to be the Pareto preference (which was called the preference in the weak Pareto sense in Ye and Zhu [33]). In the case where there are no equilibrium constraints, (II)(iii) and (II)(ii) provide the error bound and the linear constraint qualifications for the classical multiobjective optimization problem under the general preference studied in this paper. To our knowledge, even these results are new. Recently Bao and Mordukhovich [1] have introduced some notions of relative Pareto minimizers for constrained multiobjective optimization problems and derived necessary optimality conditions for these problems. Using the necessary optimality conditions in Bao and Mordukhovich [1, Theorem 5.3] and our proof technique, one can show that (II)(iii) and (II)(ii) can also serve as constraint qualifications for the necessary optimality conditions for MPECs with various notions of relative Pareto minimizers, as defined in Bao and Mordukhovich [1].

3. The KKT approach. If it works, the KKT approach provides a simple characterization of optimality for BLPP. However, as discussed in the introduction, the KKT approach may be misleading if it is not used properly. In this section we try to explore the possibility of using the KKT approach to solve BLPPs. The following result provides a relationship between local solutions of (BLPP) and problem (KP).

Proposition 3.1. *Let $(\bar{x}, \bar{y})$ be a solution of (BLPP) on $U(\bar{x}, \bar{y})$, where $U(\bar{x}, \bar{y})$ is a neighborhood of $(\bar{x}, \bar{y})$. Suppose that for each $(x, y) \in U(\bar{x}, \bar{y})$, the KKT condition is necessary and sufficient for y to be a global optimal solution of the lower-level problem (P_x) and $\bar{u}$ is a corresponding multiplier associated with $(\bar{x}, \bar{y})$; i.e.,*

$$\nabla_y f(\bar{x}, \bar{y}) + \bar{u}\nabla_y g(\bar{x}, \bar{y}) = 0, \quad \bar{u} \geq 0, \qquad \langle g(\bar{x}, \bar{y}), \bar{u}\rangle = 0,$$

where $u\nabla_y g(x, y) := \sum_{i=1}^p u_i \nabla_y g_i(x, y)$. Then $(\bar{x}, \bar{y}, \bar{u})$ is a local optimal solution (on $U(\bar{x}, \bar{y}) \times R^p$) of the following one-level multiobjective optimization problem, where the lower-level problem has been replaced by its KKT conditions:

$$\begin{aligned}
(\text{KP}) \quad \min_{x,y,u} \ & F(x, y) \\
\text{s.t. } & \nabla_y f(x, y) + u\nabla_y g(x, y) = 0, \\
& g(x, y) \leq 0, \quad u \geq 0, \quad \langle g(x, y), u\rangle = 0, \\
& G(x, y) \leq 0.
\end{aligned}$$

Conversely, suppose that $(\bar{x}, \bar{y}, \bar{u})$ is a local optimal solution to (KP) restricting on $U(\bar{x}, \bar{y}) \times R^p$, the KKT condition is necessary and sufficient for $\bar{y}$ to be a global optimal solution of the lower-level problem $(P_{\bar{x}})$, and the KKT condition holds at each $y \in S(x)$ for all $(x, y) \in U(\bar{x}, \bar{y})$. Then $(\bar{x}, \bar{y})$ is a local solution of (BLPP).

Proof. Let $(\bar{x}, \bar{y})$ be an optimal solution to (BLPP) restricting on $B(\bar{x}, \bar{y})$. Then $\bar{y}$ must be a global optimal solution of the lower-level problem $P_{\bar{x}}$. By the assumption, the KKT condition holds and $\bar{u}$ is a corresponding multiplier. Hence, $(\bar{x}, \bar{y}, \bar{u})$ is a feasible solution to problem (KP). To show that $(\bar{x}, \bar{y}, \bar{u})$ is a local optimal solution of (KP), it suffices to show that there is no other feasible point (x, y, u) of (KP) on $U(\bar{x}, \bar{y}) \times R^p$ such that

$$F(x, y) \prec F(\bar{x}, \bar{y}). \tag{5}$$

We show this by contradiction. Suppose that there is a feasible point (x, y, u) of (KP) on $U(\bar{x}, \bar{y}) \times R^p$ such that (5) holds. Then by the assumption, y must be a global optimal solution of P_x and hence (x, y) is obviously a feasible solution of the (BLPP); this contradicts the fact that $(\bar{x}, \bar{y})$ is an optimal solution to (BLPP) on $U(\bar{x}, \bar{y})$.

Conversely, suppose that $(\bar{x}, \bar{y}, \bar{u})$ is an optimal solution to (KP) on $U(\bar{x}, \bar{y}) \times R^p$. Then there is no other feasible solution (x, y, u) that lies in $U(\bar{x}, \bar{y}) \times R^p$ such that

$$F(x, y) \prec F(\bar{x}, \bar{y}). \tag{6}$$

We now prove that $(\bar{x}, \bar{y})$ is an optimal solution to (BLPP) on $U(\bar{x}, \bar{y})$ by contradiction. First by the assumption, the KKT condition is necessary and sufficient for $\bar{y}$ to be a global optimal solution of the lower-level problem $P_{\bar{x}}$. Consequently, $\bar{y} \in S(\bar{x})$ and hence $(\bar{x}, \bar{y})$ is a feasible solution to (BLPP). Suppose that $(\bar{x}, \bar{y})$ is not an optimal solution of (BLPP) on $U(\bar{x}, \bar{y})$. Then there exists (x, y), a feasible solution of (BLPP) on $U(\bar{x}, \bar{y})$, such that (6) holds. But by the assumption, the KKT condition holds at (x, y), which means that there exists u such that (x, y, u) is a feasible solution of problem (KP). This contradicts the optimality of $(\bar{x}, \bar{y}, \bar{u})$. □

Remark 3.1. (i) Note that the converse statement of Proposition 3.1 is not the same as saying that the (x, y) component of a local solution of (KP) must be a local solution of (BLPP), because $(\bar{x}, \bar{y}, \bar{u})$ is required to be a local optimal solution to (KP) locally for (x, y) but globally for all u component. In fact, Dempe and Dutta [8, Example 3.1] have given an example for which the (x, y) component of a local solution of (KP) is not a local solution of (BLPP) that has a *convex* lower-level problem. Moreover, they showed that LICQ of the lower-level problem is not a generic condition, and hence this situation is not just an exception. Actually, the converse statement of Proposition 3.1 for the case of a single-objective bilevel program with a convex lower-level problem and the Slater condition was given by Dempe and Dutta [8, Theorem 3.2].

(ii) Although it is obvious that for the case where the lower-level problem is convex and the Slater condition holds for $P_{\bar{x}}$; the KKT condition is necessary and sufficient for all lower-level problems near the optimal solution. There are a few more situations where this condition holds for not necessarily convex lower-level problems, for example, when the lower-level problem is generalized convex, i.e., when $f(x, \cdot)$ is a differentiable pseudoconvex function, $g_i(x, \cdot)$ are differentiable quasiconvex functions and a certain constraint qualification is satisfied for all lower-level problems near the optimal solution. Another case when this happens is when *all* lower-level problems near the optimal solution have a unique KKT point and the optimal solution exists.

Given a feasible vector $(\bar{x}, \bar{y}, \bar{u})$ in the feasible region of (KP), we define the following index sets:

$$\begin{aligned} I_G = I_G(\bar{x}, \bar{y}) &:= \{i\colon G_i(\bar{x}, \bar{y}) = 0\} \\ I_+ = I_+(\bar{x}, \bar{y}, \bar{u}) &:= \{i\colon g_i(\bar{x}, \bar{y}) = 0, \bar{u}_i > 0\} \\ I_u = I_u(\bar{x}, \bar{y}, \bar{u}) &:= \{i\colon g_i(\bar{x}, \bar{y}) < 0, \bar{u}_i = 0\} \\ I_0 = I_0(\bar{x}, \bar{y}, \bar{u}) &:= \{i\colon g_i(\bar{x}, \bar{y}) = 0, \bar{u}_i = 0\}. \end{aligned}$$

Definition 3.1 (Stationary Conditions for (KP)). Let $(\bar{x}, \bar{y}, \bar{u})$ be a feasible solution to (KP). We say that $(\bar{x}, \bar{y}, \bar{u})$ is a C-stationary point if there exists a unit vector $\lambda \in N_+(F(\bar{x}, \bar{y}); \overline{l(F(\bar{x}, \bar{y}))})$ and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$, such that

$$0 \in \partial\langle \lambda, F\rangle(\bar{x}, \bar{y}) + \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T \beta + \nabla g(\bar{x}, \bar{y})^T \eta^g + \nabla G(\bar{x}, \bar{y})^T \eta^G, \tag{7}$$

$$\eta_i^G \geq 0 \quad i \in I_G, \qquad \eta_i^G = 0 \quad i \notin I_G, \tag{8}$$

$$\eta_i^g = 0 \quad i \in I_u, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \quad i \in I_+, \tag{9}$$

$$\eta_i^g (\nabla_y g(\bar{x}, \bar{y})^T \beta)_i \geq 0 \quad i \in I_0.$$

We say that $(\bar{x}, \bar{y}, \bar{u})$ is an M-stationary point if there exist a unit vector $\lambda \in N_+(F(\bar{x}, \bar{y}); \overline{l(F(\bar{x}, \bar{y}))})$ and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that (7)–(9) and the following condition holds:

$$\text{either } \eta_i^g > 0, (\nabla_y g(\bar{x}, \bar{y})\beta)_i > 0 \quad \text{or} \quad \eta_i^g (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \qquad i \in I_0.$$

We say that $(\bar{x}, \bar{y}, \bar{u})$ is an S-stationary point if there exist a unit vector $\lambda \in N_+(F(\bar{x}, \bar{y}); \overline{l(F(\bar{x}, \bar{y}))})$ and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$, such that (7)–(9) and the following condition holds:

$$\eta_i^g \geq 0, \quad (\nabla_y g(\bar{x}, \bar{y})\beta)_i \geq 0 \qquad i \in I_0.$$

We say that $(\bar{x}, \bar{y}, \bar{u})$ is a P-stationary point if for each partition of the index set I_0 into P, Q there exist a unit vector $\lambda \in N_+(F(\bar{x}, \bar{y}); \overline{l(F(\bar{x}, \bar{y}))})$ and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that (7)–(9) and the following condition holds:

$$\eta_i^g \geq 0 \quad i \in P, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i \geq 0 \quad i \in Q.$$

THEOREM 3.1. *Let $(\bar{x}, \bar{y})$ be a local optimal solution of (BLPP). Assume that F is Lipschitz continuous, G is C^1, and f, g are twice continuously differentiable around $(\bar{x}, \bar{y})$. Further assume that for each (x, y), which is sufficiently close to $(\bar{x}, \bar{y})$, the KKT condition is necessary and sufficient for y to be a global optimal solution of P_x and $\bar{u}$ is a corresponding multiplier associated with $(\bar{x}, \bar{y})$.*

(I) *$(\bar{x}, \bar{y}, \bar{u})$ is a C-stationary point if there is no nonzero vector $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that*

$$0 = \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T \beta + \nabla g(\bar{x}, \bar{y})^T \eta^g + \nabla G(\bar{x}, \bar{y})^T \eta^G,$$
$$\eta_i^G \geq 0 \quad i \in I_G, \qquad \eta_i^G = 0 \quad i \notin I_G,$$
$$\eta_i^g = 0 \quad i \in I_u, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \quad i \in I_+,$$
$$\eta_i^g (\nabla_y g(\bar{x}, \bar{y})\beta)_i \geq 0 \quad i \in I_0.$$

(II) *$(\bar{x}, \bar{y}, \bar{u})$ is an M-stationary point if one of the following constraint qualifications holds:*

(i) *There is no nonzero vector $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that*

$$0 = \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T \beta + \nabla g(\bar{x}, \bar{y})^T \eta^g + \nabla G(\bar{x}, \bar{y})^T \eta^G,$$
$$\eta_i^G \geq 0 \quad i \in I_G, \qquad \eta_i^G = 0 \quad i \notin I_G,$$
$$\eta_i^g = 0 \quad i \in I_u, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \quad i \in I_+,$$
$$\text{either} \quad \eta_i^g > 0, \ (\nabla_y g(\bar{x}, \bar{y})\beta)_i > 0 \quad \text{or} \quad \eta_i^g (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \qquad i \in I_0.$$

(ii) *$\nabla_y f, g, G$ are affine mappings.*

(iii) *The error-bound constraint qualification holds for (KP) at $(\bar{x}, \bar{y}, \bar{u})$.*

(iv) *There is no inequality constraint $G(x, y) \leq 0$. Furthermore, the second-order sufficient condition holds for the lower-level problem $P_{\bar{x}}$ at $\bar{y}$; i.e., for any nonzero v such that*

$$\nabla_y g_i(\bar{x}, \bar{y})^T v = 0, \quad i \in I_+, \qquad \nabla_y g_i(\bar{x}, \bar{y})^T v \leq 0, \quad i \in I_0$$

$\langle v, (\nabla_y^2 f(\bar{x}, \bar{y}) + \bar{u}\nabla_y^2 g(\bar{x}, \bar{y}))v\rangle > 0$.

(III) *$(\bar{x}, \bar{y}, \bar{u})$ is an S-stationary point if MPEC LICQ holds for KP.*

(IV) *$(\bar{x}, \bar{y}, \bar{u})$ is a P-stationary point if either $\nabla_y f, g, G$ is affine or MPEC piecewise MFCQ holds for KP.*

PROOF. By virtue of Proposition 3.1, under the assumptions of the theorem, $(\bar{x}, \bar{y}, \bar{u})$ is a local optimal solution of (KP). Because (KP) is a MPEC, (I), (II)(i)(ii)(iii), (III), and (IV) follow immediately from applying (I), (II)(i)(ii)(iii), (III), and (IV) of Theorem 2.1 to the problem (KP), respectively. It now remains to show that (II)(iv) implies the error-bound constraint qualification in (II)(iii). Indeed, the implication of (II)(iv) to the error-bound constraint qualification in (II)(ii) follows from the error-bound result of Hager and Gowda [11, Lemma 2]. □

REMARK 3.2. Necessary optimality conditions for multiobjective bilevel programs using the KKT approach have been studied in the literature by Ye and Zhu [33] for the smooth case and Bao et al. [2] for the nonsmooth case. All results in Theorem 3.1 are new except II(i), which was derived in Ye and Zhu [33, Theorem 5.3(b)]. (II)(ii) has improved the results of Ye and Zhu [33, Theorem 5.3(a)] in that the preference is not required to be the Pareto preference (which was called the preference in the weak Pareto sense in Ye and Zhu [33]). (II)(iv) has improved the results of Ye and Zhu [33, Theorem 5.3(c)] in that no LICQ is required for the lower-level problem. This result is even new for the case of a single-objective BLPP, in which case it improved the result of Ye [29, Theorem 5.1(d)] in that no LICQ is required for the lower-level problem. This new result significantly improves the existing results, because the LICQ is a pretty strong requirement. (The author thanks a referee for questioning the necessity of the LICQ, which motivated finding the answer to the question.)

4. Combined MPEC and the value function approach. Unfortunately, as demonstrated in the introduction section by using the example (P), optimal solutions of many *nonconvex* BLPPs where the lower-level problem is not convex do not satisfy the KKT conditions derived by using either the KKT approach or the value function approach (see more examples in Ye and Zhu [32]). According to Proposition 3.1, even when the lower-level problem is convex, the KKT condition is still required to hold for *all* points near the optimal solution for the KKT approach to work.

As proposed in Ye and Zhu [32], we should consider the following combined problem:

$$
\begin{aligned}
(\text{CP}) \quad \min_{x,y,u} \quad & F(x,y) \\
\text{s.t.} \quad & f(x,y) - V(x) \le 0, \\
& \nabla_y f(x,y) + u\nabla_y g(x,y) = 0, \\
& g(x,y) \le 0, \qquad u \ge 0, \qquad \langle g(x,y), u\rangle = 0, \\
& G(x,y) \le 0.
\end{aligned} \tag{10}
$$

The relationship of (CP) and (BLPP) is given in the following proposition. Note that using the combined problem, the KKT condition is only required to hold at the optimal solution $(\bar{x}, \bar{y})$.

Proposition 4.1. *Let $(\bar{x}, \bar{y})$ be a local (global) optimal solution to* (BLPP). *Suppose that at $\bar{y}$, the KKT condition holds for the lower-level problem $P_{\bar{x}}$. Then there exists $\bar{u}$ such that $(\bar{x}, \bar{y}, \bar{u})$ is a local (global) optimal solution of* (CP). *Conversely, suppose that $(\bar{x}, \bar{y}, \bar{u})$ is an optimal solution to (CP) restricting on $U(\bar{x}, \bar{y}) \times R^p$, where $U(\bar{x}, \bar{y})$ is a neighborhood of $(\bar{x}, \bar{y})$ and the KKT condition holds at $y \in S(x)$ for lower-level problem P_x for all (x, y) in $U(\bar{x}, \bar{y})$. Then $(\bar{x}, \bar{y})$ is a local solution of* (BLPP).

Proof. Let $(\bar{x}, \bar{y})$ be a local optimal solution to (BLPP). Then $\bar{y}$ must be a global optimal solution to the lower-level problem $P_{\bar{x}}$. By the assumption, the KKT condition holds; i.e., there exists a multiplier $\bar{u}$ such that

$$
0 = \nabla_y f(\bar{x}, \bar{y}) + \sum_{i=1}^{p} \bar{u}_i \nabla_y g_i(\bar{x}, \bar{y}),
$$

$$
\bar{u} \ge 0, \quad \sum_{i=1}^{p} \bar{u}_i g_i(\bar{x}, \bar{y}) = 0.
$$

Therefore, $(\bar{x}, \bar{y}, \bar{u})$ is a feasible solution to problem (CP). To show that $(\bar{x}, \bar{y}, \bar{u})$ is a local optimal solution of (CP), it suffices to show that there is no other feasible point (x, y, u) of (CP) in a neighborhood of $(\bar{x}, \bar{y}, \bar{u})$ such that

$$
F(x,y) \prec F(\bar{x}, \bar{y}). \tag{11}
$$

We show this by contradiction. Suppose that there is a feasible point (x, y, u) of (CP) in a neighborhood of $(\bar{x}, \bar{y}, \bar{u})$ such that (11) holds. Then, because the (x, y) components of the vector (x, y, u) are obviously a feasible solution of the (BLPP), this contradicts the fact that $(\bar{x}, \bar{y})$ is a local optimal solution to (BLPP).

Conversely, suppose that $(\bar{x}, \bar{y}, \bar{u})$ is an optimal solution to (CP) on $U(\bar{x}, \bar{y}) \times R^p$. Then there is no other feasible solution (x, y, u) that lies in $U(\bar{x}, \bar{y}) \times R^p$ such that

$$
F(x,y) \prec F(\bar{x}, \bar{y}). \tag{12}
$$

We now prove that $(\bar{x}, \bar{y})$ is an optimal solution to (BLPP) on $U(\bar{x}, \bar{y})$. To the contrary, suppose that $(\bar{x}, \bar{y})$ is not an optimal solution of (BLPP) on $U(\bar{x}, \bar{y})$. Then there exists (x, y), a feasible solution to (BLPP) on $U(\bar{x}, \bar{y})$, such that (12) holds. But by the assumption, the KKT condition holds at (x, y), which means that there exists u such that (x, y, u) is a feasible solution of the problem (CP). This contradicts the optimality of $(\bar{x}, \bar{y}, \bar{u})$. □

Suppose that the value function $V(x)$ is Lipschitz continuous near the optimal solution then the problem (CP) is an MPEC with continuously differentiable and Lipschitz-continuous problem data. However, because of the value function constraint (10), we can argue as in Ye and Zhu [32, Proposition 1.3] that the usual MPEC constraint qualifications such as MPEC LICQ and MPEC piecewise MFCQ will never hold. Because the value function is usually not linear, the MPEC linear CQ is unlikely to hold as well. We extend the following partial calmness condition for (CP) introduced in Ye and Zhu [32] to the multiobjective case.

Definition 4.1 (Partial Calmness for (CP)). Let $(\bar{x}, \bar{y}, \bar{u})$ be a local solution of (CP) with $W = R^N$. We say that (CP) is partially calm at $(\bar{x}, \bar{y}, \bar{u})$ if there exists $\mu > 0$ such that $(\bar{x}, \bar{y}, \bar{u})$ is a local solution of the following partially penalized problem:

$$
\begin{aligned}
(\text{CP})_\mu \quad \min \quad & F(x,y) + \mu(f(x,y) - V(x)) \\
\text{s.t.} \quad & \nabla_y f(x,y) + u\nabla_y g(x,y) = 0, \\
& u \ge 0, \qquad g(x,y) \le 0, \qquad \langle g(x,y), u\rangle = 0 \\
& G(x,y) \le 0,
\end{aligned} \tag{13}
$$

where $F(x,y)+\mu(f(x,y)-V(x))$ denote the vector in $W=R^N$ with the ith component equal to $F_i(x,y)+\mu(f(x,y)-V(x))$.

Definition 4.2 (Stationary Conditions for (CP) Based on the Value Function). Let $(\bar{x},\bar{y},\bar{u})$ be a feasible solution to (CP) with $W=R^N$. Suppose that F,G are C^1 and f,g are C^2 around $(\bar{x},\bar{y})$. We say that $(\bar{x},\bar{y},\bar{u})$ is a C-stationary point based on the value function if there exists a unit vector $\lambda\in N_+(F(\bar{x},\bar{y});\overline{l(F(\bar{x},\bar{y}))})$ and $\mu\geq 0$, $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that

$$0\in\sum_{i=1}^N\lambda_i\nabla F_i(\bar{x},\bar{y})+\mu[\nabla f(\bar{x},\bar{y})-\partial^c V(\bar{x})\times\{0\}]$$

$$+\nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T\beta+\nabla g(\bar{x},\bar{y})^T\eta^g+\nabla G(\bar{x},\bar{y})^T\eta^G, \tag{14}$$

$$\eta_i^G\geq 0\quad i\in I_G,\qquad \eta_i^G=0\quad i\notin I_G, \tag{15}$$

$$\eta_i^g=0\quad i\in I_u,\qquad (\nabla_y g(\bar{x},\bar{y})\beta)_i=0\quad i\in I_+, \tag{16}$$

$$\eta_i^g(\nabla_y g(\bar{x},\bar{y})\beta)_i\geq 0\quad i\in I_0.$$

We say that $(\bar{x},\bar{y},\bar{u})$ is an M-stationary point based on the value function if there exist a unit vector $\lambda\in N_+(F(\bar{x},\bar{y});\overline{l(F(\bar{x},\bar{y}))})$ and $\mu\geq 0$, $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that (14)–(16) and the following condition holds:

$$\text{either}\quad \eta_i^g>0,\ (\nabla_y g(\bar{x},\bar{y})\beta)_i>0,\quad\text{or}\quad \eta_i^g(\nabla_y g(\bar{x},\bar{y})\beta)_i=0\qquad i\in I_0.$$

We say that $(\bar{x},\bar{y},\bar{u})$ is an S-stationary point based on the value function if there exist a unit vector $\lambda\in N_+(F(\bar{x},\bar{y});\overline{l(F(\bar{x},\bar{y}))})$ and $\mu\geq 0$, $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that (14)–(16) and the following condition holds:

$$\eta_i^g\geq 0,\quad (\nabla_y g(\bar{x},\bar{y})\beta)_i\geq 0\qquad i\in I_0.$$

We say that $(\bar{x},\bar{y},\bar{u})$ is a P-stationary point based on the value function if for each partition of the index set I_0 into P,Q, there exist a unit vector

$$\lambda\in N_+(F(\bar{x},\bar{y});\overline{l(F(\bar{x},\bar{y}))})$$

and $\mu\geq 0$, $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that (14)–(16) and the following condition holds:

$$\eta_i^g\geq 0\quad i\in P,\qquad (\nabla_y g(\bar{x},\bar{y})\beta)_i\geq 0\quad i\in Q.$$

According to Proposition 4.1, similarly as in the proof of Theorem 3.1, we may apply Theorem 2.1 to the problem $((CP))_\mu$ and obtain the following results.

Theorem 4.1. *Let $(\bar{x},\bar{y})$ be a local solution to (BLPP) with $W=R^N$. Suppose that F,G are C^1 and f,g are C^2 around $(\bar{x},\bar{y})$. Suppose that at $\bar{y}$, the KKT condition holds for the lower-level problem $P_{\bar{x}}$ and $\bar{u}$ is a corresponding multiplier. Moreover, suppose that the value function $V(x)$ is Lipschitz continuous near $\bar{x}$ and* (CP) *is partially calm at $(\bar{x},\bar{y},\bar{u})$.*

(I) $(\bar{x},\bar{y},\bar{u})$ *is a C-stationary point based on the value function if there is no nonzero vector $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that*

$$0=\nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T\beta+\nabla g(\bar{x},\bar{y})^T\eta^g+\nabla G(\bar{x},\bar{y})^T\eta^G,$$

$$\eta_i^G\geq 0\quad i\in I_G,\qquad \eta_i^G=0\quad i\notin I_G,$$

$$\eta_i^g=0\quad i\in I_u,\qquad (\nabla_y g(\bar{x},\bar{y})\beta)_i=0\quad i\in I_+,$$

$$\eta_i^g(\nabla_y g(\bar{x},\bar{y})\beta)_i\geq 0\quad i\in I_0.$$

(II) $(\bar{x},\bar{y},\bar{u})$ *is an M-stationary point based on the value function if one of the following constraint qualifications holds*:

(i) *There is no nonzero vector $\beta\in R^m$, $\eta^g\in R^p$, $\eta^G\in R^q$ such that*

$$0=\nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T\beta+\nabla g(\bar{x},\bar{y})^T\eta^g+\nabla G(\bar{x},\bar{y})^T\eta^G,$$

$$\eta_i^G\geq 0\quad i\in I_G,\qquad \eta_i^G=0\quad i\notin I_G,$$

$$\eta_i^g=0\quad i\in I_u,\quad (\nabla_y g(\bar{x},\bar{y})\beta)_i=0\quad i\in I_+,$$

$$\textit{either}\quad \eta_i^g>0,\ (\nabla_y g(\bar{x},\bar{y})\beta)_i>0,\quad \textit{or}\quad \eta_i^g(\nabla_y g(\bar{x},\bar{y})\beta)_i=0\qquad i\in I_0.$$

(ii) $\nabla_y f,g,G$ *are affine mappings.*

(iii) *The error-bound constraint qualification holds for $(CP)_\mu$ at $(\bar{x}, \bar{y}, \bar{u})$.*

(iv) *There is no inequality constraint $G(x, y) \le 0$, and the second-order sufficient condition holds for the lower-level problem $P_{\bar{x}}$ at $\bar{y}$; i.e., for any nonzero v such that*

$$\nabla_y g_i(\bar{x}, \bar{y})^T v = 0, \quad i \in I_+, \qquad \nabla_y g_i(\bar{x}, \bar{y})^T v \le 0, \quad i \in I_0$$

$\langle v, (\nabla_y^2 f(\bar{x}, \bar{y}) + \bar{u}\nabla_y^2 g(\bar{x}, \bar{y}))v\rangle > 0.$

(III) $(\bar{x}, \bar{y}, \bar{u})$ *is an S-stationary point based on the value function if MPEC LICQ holds for $(CP)_\mu$.*

(IV) $(\bar{x}, \bar{y}, \bar{u})$ *is a P-stationary point if either $\nabla_y f, g, G$ is affine or MPEC piecewise MFCQ holds for $(CP)_\mu$.*

In what follows, we give some sufficient conditions for partial calmness of the problem (CP) to hold. First we apply an error-bound result from Wu and Ye [24, Theorem 4.2] to obtain the following result.

Lemma 4.1. *Let $\widetilde{\mathcal{F}}$ and $\mathcal{F}$ denote the feasible regions of the problem CP_μ and (CP), respectively. If for some $c > 0$, $\varepsilon > 0$, and each $(x, y) \in \widetilde{\mathcal{F}}$ such that $0 < f(x, y) - V(x) < \varepsilon$ there exists a unit vector (d_x, d_y, d_u) that lies in the tangent cone of $\widetilde{\mathcal{F}}$ at (x, y, u) such that*

$$\nabla f(x, y)^T (d_x, d_y) - V^-(x; d_x) \le -c^{-1},$$

then

$$d((x, y, u), \mathcal{F}) \le c(f(x, y) - V(x)) \quad \forall\, (x, y, u) \in \widetilde{\mathcal{F}} \quad \text{such that } 0 < f(x, y) - V(x) < \varepsilon, \tag{17}$$

where

$$V^-(x; d_x) := \liminf_{d' \to d_x,\, t \downarrow 0} \frac{V(x + td') - V(x)}{t}$$

is the lower Dini derivative of V at x in direction d_x.

Proposition 4.2. *Assume that $(\bar{x}, \bar{y}, \bar{u})$ is a local solution of (CP) with $W = R^N$ and $\prec$ is the Pareto preference. Furthermore, suppose that for some $c > 0$, $\varepsilon > 0$, and each $(x, y) \in \widetilde{\mathcal{F}}$ such that $0 < f(x, y) - V(x) < \varepsilon$, there exists a unit vector (d_x, d_y, d_u) that lies in the tangent cone of $\widetilde{\mathcal{F}}$ at (x, y, u) such that*

$$\nabla f(x, y)^T (d_x, d_y) - V^-(x; d_x) \le -c^{-1}.$$

Then (CP) is partially calm at $(\bar{x}, \bar{y}, \bar{u})$ with $\mu = L_F c$, where L_F is the Lipschitz constant of F.

Proof. To the contrary, suppose that $(\bar{x}, \bar{y}, \bar{u})$ is not a local solution of CP_μ with $\mu = L_F c$. Then for all $\varepsilon > 0$, there is $(x, y, u) \in \widetilde{\mathcal{F}} \cap B((\bar{x}, \bar{y}, \bar{u}), \varepsilon)$ such that

$$F(x, y) + L_F c(f(x, y) - V(x)) \prec F(\bar{x}, \bar{y}).$$

Let $(\tilde{x}, \tilde{y}, \tilde{u})$ be the projection of (x, y, u) to $\mathcal{F}$; i.e., $(\tilde{x}, \tilde{y}, \tilde{u}) \in \mathcal{F}$ and

$$d((x, y, u), \mathcal{F}) = \|(x, y, u) - (\tilde{x}, \tilde{y}, \tilde{u})\|.$$

By Lemma 4.1, we can choose $\varepsilon > 0$ small enough such that the local error bound (17) holds and F is Lipschitz. Then

$$\begin{aligned} F(\tilde{x}, \tilde{y}) &\le F(x, y) + L_F\|(x, y) - (\tilde{x}, \tilde{y})\| \quad \text{by Lipschitz continuity of } F \\ &\le F(x, y) + L_F\|(x, y, u) - (\tilde{x}, \tilde{y}, \tilde{u})\| \\ &\le F(x, y) + L_F c(f(x, y) - V(x)) \quad \text{by local error bound (17)} \\ &\prec F(\bar{x}, \bar{y}). \end{aligned}$$

But this contradicts the fact that $(\bar{x}, \bar{y}, \bar{u})$ is a local solution of (CP). □

It can be shown easily that the linearization cone of the feasible region of (CP_μ) can be described as follows:

Definition 4.3 (Linearization Cone). Let $\widetilde{\mathscr{F}}$ denote the feasible region of the problem CP_μ. The linearization cone of $\widetilde{\mathscr{F}}$ at $(\bar{x}, \bar{y}, \bar{u})$ is the cone defined by

$$\mathscr{L}((\bar{x},\bar{y},\bar{u});\widetilde{\mathscr{F}}) := \left\{ (d,v)\in R^{n+m}\times R^p : \begin{array}{l} \nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T d+\nabla_y g(\bar{x},\bar{y})^T v=0 \\ \nabla G_i(\bar{x},\bar{y})^T d\le 0 \quad i\in I_G \\ \nabla g_i(\bar{x},\bar{y})^T d=0 \quad i\in I_+ \\ v_i=0 \quad i\in I_u \\ \nabla g_i(\bar{x},\bar{y})^T d\le 0, \quad v_i\ge 0 \;\; i\in I_0 \end{array} \right\}.$$

Because the feasible region of a MPEC may be nonconvex, it is unreasonable to expect that the usual linearization cone of the feasible region $\widetilde{\mathscr{F}}$ is equal to the tangent cone of the feasible region $\widetilde{\mathscr{F}}$. However, in the MPEC literature, it is known that under weak assumptions, the MPEC linearization cone defined as follows is equal to the tangent cone of the feasible region. When the tangent cone is equal to the MPEC linearization cone, it is said that MPEC Abadie constraint qualification holds. The reader is referred to Ye [28] for sufficient conditions for MPEC Abadie constraint qualification to hold.

Definition 4.4 (MPEC Linearization Cone). The MPEC linearization cone of $\widetilde{\mathscr{F}}$ at $(\bar{x}, \bar{y}, \bar{u})$ is the cone defined by

$$\mathscr{L}^{\text{MPEC}}((\bar{x},\bar{y},\bar{u});\widetilde{\mathscr{F}}) := \left\{ (d,v)\in R^{n+m}\times R^p : \begin{array}{l} \nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T d+\nabla_y g(\bar{x},\bar{y})^T v=0 \\ \nabla G_i(\bar{x},\bar{y})^T d\le 0 \quad i\in I_G \\ \nabla g_i(\bar{x},\bar{y})^T d=0 \quad i\in I_+ \\ v_i=0 \quad i\in I_u \\ \left\{ \begin{array}{l} \nabla g_i(\bar{x},\bar{y})^T d\cdot v_i=0 \\ \nabla g_i(\bar{x},\bar{y})^T d\le 0, \quad v_i\ge 0 \end{array} \right. \quad i\in I_0 \end{array} \right\}.$$

Theorem 4.2. *Let $(\bar{x}, \bar{y})$ be a local solution to (BLPP) with $W = R^N$. Suppose that F, G are C^1 and f, g are C^2 around $(\bar{x}, \bar{y})$. Suppose that at $\bar{y}$, the KKT condition holds for the lower-level problem $P_{\bar{x}}$ and $\bar{u}$ is a corresponding multiplier. Moreover, suppose that the value function $V(x)$ is Lipschitz continuous near $\bar{x}$.*

If for some $\mu > 0$ there is no $(d, v) \in \mathscr{L}^{\text{MPEC}}((\bar{x}, \bar{y}, \bar{u}); \widetilde{\mathscr{F}})$ such that

$$[F+\mu(f-V)]^\circ((\bar{x},\bar{y});d) \prec 0, \tag{18}$$

where $[F+\mu(f-V)]^\circ((\bar{x},\bar{y});d)$ denotes the vector with the ith component equal to $[F_i+\mu(f-V)]^\circ((\bar{x},\bar{y});d)$, then $(\bar{x}, \bar{y}, \bar{u})$ is an M- and P-stationary point based on the value function.

If for some $\mu > 0$ there is no $(d, v) \in \mathscr{L}((\bar{x}, \bar{y}, \bar{u}); \widetilde{\mathscr{F}})$ such that

$$[F+\mu(f-V)]^\circ((\bar{x},\bar{y});d) \prec 0 \tag{19}$$

then $(\bar{x}, \bar{y}, \bar{u})$ is an S-stationary point based on the value function.

Proof. By (18), $(d, v) = (0, v)$ is an optimal solution to the following linearized problem:

$$\begin{aligned} \min_{(d,v)} \quad & \Phi(d) := [F+\mu(f-V)]^\circ((\bar{x},\bar{y});d) \\ \text{s.t.} \quad & \nabla(\nabla_y f+\bar{u}\nabla_y g)(\bar{x},\bar{y})^T d+\nabla_y g(\bar{x},\bar{y})^T v=0, \\ & \nabla G_i(\bar{x},\bar{y})^T d\le 0 \quad i\in I_G, \\ & \nabla g_i(\bar{x},\bar{y})^T d=0 \quad i\in I_+, \\ & v_i=0 \quad i\in I_u, \\ & \left\{ \begin{array}{l} \nabla g_i(\bar{x},\bar{y})^T d\cdot v_i=0, \\ \nabla g_i(\bar{x},\bar{y})^T d\le 0, \quad v_i\ge 0 \end{array} \right. \quad i\in I_0. \end{aligned} \tag{20}$$

The objective function of the above problem is nonsmooth and convex and the constraint functions are all linear in variable (d, v). Hence, the MPEC linear CQ holds. Applying Theorem 2.1, we conclude that there exists a unit vector $\lambda \in N_+(F(\bar{x}, \bar{y}); \overline{l(F(\bar{x}, \bar{y}))})$, multipliers $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$, $\eta^u \in R^p$ such that

$$
\begin{gathered}
0 \in \partial\langle\lambda, \Phi\rangle(0) + \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T\beta + \nabla g(\bar{x}, \bar{y})^T\eta^g + \nabla G(\bar{x}, \bar{y})^T\eta^G,\\
0 = \nabla_y g(\bar{x}, \bar{y})\beta - \eta^u,\\
\eta_i^G \geq 0 \quad i \in I_G, \qquad \eta_i^G = 0 \quad i \notin I_G,\\
\eta_i^g = 0 \quad i \in I_u, \qquad \eta_i^u = 0 \quad i \in I_+,\\
\text{either} \quad \eta_i^g > 0 \ \eta_i^u > 0 \quad \text{or} \quad \eta_i^g\eta_i^u = 0 \qquad i \in I_0.
\end{gathered}
$$

By the calculus rules for Clarke generalized gradients in Proposition 2.1, one has

$$
\partial^c\Phi_i(0) \subseteq \nabla F_i(\bar{x}, \bar{y}) + \mu\{\nabla f(\bar{x}, \bar{y}) - \partial^c V(\bar{x}) \times \{0\}\}.
$$

Hence, we have by Proposition 2.1 that

$$
\begin{aligned}
\partial\langle\lambda, \Phi\rangle(0) &\subset \partial^c\langle\lambda, \Phi\rangle(0)\\
&\subset \sum_{i=1}^n \lambda_i\partial^c\Phi_i(0)\\
&\subset \sum_{i=1}^n \lambda_i\nabla F_i(\bar{x}, \bar{y}) + \mu\{\nabla f(\bar{x}, \bar{y}) - \partial^c V(\bar{x}) \times \{0\}\}.
\end{aligned}
$$

The conclusion that $(\bar{x}, \bar{y}, \bar{u})$ is an M-stationary point based on the value function follows from replacing η^u by $\nabla_y g(\bar{x}, \bar{y})\beta$. Similarly, we can prove that $(\bar{x}, \bar{y}, \bar{u})$ is a P-stationary point based on the value function.

Now suppose that (19) holds. Then $(d, v) = (0, v)$ is an optimal solution to the following linearized problem:

$$
\begin{aligned}
\min_{(d, v)} \quad & \Phi(d) := F_\mu^\circ((\bar{x}, \bar{y}); d)\\
\text{s.t.} \quad & \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T d + \nabla_y g(\bar{x}, \bar{y})^T v = 0,\\
& \nabla G_i(\bar{x}, \bar{y})^T d \leq 0 \quad i \in I_G,\\
& \nabla g_i(\bar{x}, \bar{y})^T d = 0 \quad i \in I_+,\\
& v_i = 0 \quad i \in I_u,\\
& \nabla g_i(\bar{x}, \bar{y})^T d \leq 0, \quad v_i \geq 0 \qquad i \in I_0.
\end{aligned}
$$

The above problem is a multiobjective optimization problem with linear constraints. The conclusion that $(\bar{x}, \bar{y}, \bar{u})$ is an S-stationary point based on the value function follows from applying Theorem 2.1 to the above optimization problem. □

The necessary optimality conditions obtained in Theorem 4.2 involve the Clarke generalized directional derivative and the Clarke generalized gradient of the value function, and $V(x)$ is required to be Lipschitz continuous. Let $x \in R^n$. For any $y \in S(x)$ we denote the set of KKT multipliers for the lower-level problem P_x at y as follows:

$$
M^1(x, y) := \left\{u \in R^p\colon 0 = \nabla_y f(x, y) + \sum_{i=1}^p u_i\nabla_y g_i(x, y),\ u \geq 0,\ \sum_{i=1}^p u_i g_i(x, y) = 0\right\}.
$$

Recall that a set-valued map Y is called uniformly bounded around $\bar{x}$ if there exists a neighborhood U of $\bar{x}$ such that the set $\bigcup_{x\in U} Y(x)$ is bounded. The following result can be found in Gauvin and Dubeau [10] (which is a special case of Clarke [4, Theorem 6.5.2]).

Proposition 4.3. *Assume that the set-valued map $Y(x) := \{y \in R^m\colon g(x, y) \leq 0\}$ is uniformly bounded around $\bar{x}$. Suppose that MFCQ holds at y' for all $y' \in S(\bar{x})$. Then the value function $V(x)$ is Lipschitz continuous near $\bar{x}$ and*

$$
\partial^c V(\bar{x}) \subseteq \operatorname{co} W(\bar{x}),
$$

where

$$
W(\bar{x}) := \{\nabla_x f(\bar{x}, y') + u'\nabla_x g(\bar{x}, y')\colon y' \in S(\bar{x}),\ u' \in M^1(\bar{x}, y')\}. \tag{21}
$$

In some practical circumstance, calculating the Clarke generalized gradients may be difficult or impossible. We now introduce two new conditions under which our new necessary optimality conditions hold. Our new conditions do not involve either the Clarke generalized directional derivative or the Clarke generalized gradient of the value function.

Definition 4.5. Given a feasible vector $(\bar{x}, \bar{y}, \bar{u})$ of (CP_μ). Suppose the preference is the Pareto preference. We say that (CP) is MPEC-weakly calm at $(\bar{x}, \bar{y}, \bar{u})$ with modulus μ if there is no $(d, v) \in \mathscr{L}^{\text{MPEC}}((\bar{x}, \bar{y}, \bar{u}); \widetilde{\mathscr{F}})$ such that

$$[\nabla F(\bar{x}, \bar{y}) + \mu \nabla f(\bar{x}, \bar{y})]^T d - \mu \min_{\xi \in W(\bar{x})} \xi dx \prec 0. \tag{22}$$

We say that (CP) is weakly calm at $(\bar{x}, \bar{y}, \bar{u})$ with modulus μ if there is no $(d, v) \in \mathscr{L}((\bar{x}, \bar{y}, \bar{u}); \widetilde{\mathscr{F}})$ such that

$$[\nabla F(\bar{x}, \bar{y}) + \mu \nabla f(\bar{x}, \bar{y})]^T d - \mu \min_{\xi \in W(\bar{x})} \xi dx \prec 0. \tag{23}$$

Because

$$(-V)^\circ(\bar{x}; dx) = \max_{\xi \in \partial(-V)^c(\bar{x})} \{\xi dx\} \le \max_{\xi \in -W(\bar{x})} \{\xi dx\} = \max_{\xi \in W(\bar{x})} \{-\xi dx\} = -\min_{\xi \in W(\bar{x})} \{\xi dx\},$$

the MPEC-weakly calmness condition and the weakly calmness condition are weaker than conditions (18) and (19), respectively.

Theorem 4.3. *Let $(\bar{x}, \bar{y})$ be a local solution to (BLPP) with $W = R^N$ and $\prec$ being the Pareto preference. Suppose that F, G are C^1 and f, g are C^2 around $(\bar{x}, \bar{y})$. Suppose that at $\bar{y}$, the KKT condition holds for the lower-level problem $P_{\bar{x}}$ and $\bar{u}$ is a corresponding multiplier. Moreover, suppose that the set $W(\bar{x})$ as defined in (21) is nonempty and compact.*

If (CP) is MPEC-weakly calm at $(\bar{x}, \bar{y}, \bar{u})$ with modulus $\mu \ge 0$, then there exist $\lambda_i \ge 0$, $i = 1, \ldots, N$, $\sum_{i=1}^N \lambda_i = 1$, $\alpha^i \ge 0$, $\sum_{i=1}^{n+1} \alpha^i = 1$, $y^i \in S(\bar{x})$, $u^i \in M^1(\bar{x}, y^i)$, $i = 1, 2, \ldots, n+1$, and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that

$$\begin{aligned} 0 = &\sum_{i=1}^N \lambda_i \nabla_x F_i(\bar{x}, \bar{y}) + \mu \sum_{i=1}^{n+1} \alpha^i (\nabla_x f(\bar{x}, \bar{y}) - \nabla_x f(\bar{x}, y^i) - u^i \nabla_x g(\bar{x}, y^i)) \\ &+ \nabla_x(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T \beta + \nabla_x g(\bar{x}, \bar{y})^T \eta^g + \nabla_x G(\bar{x}, \bar{y})^T \eta^G, \end{aligned} \tag{24}$$

$$0 = \sum_{i=1}^N \lambda_i \nabla_y F_i(\bar{x}, \bar{y}) + \mu \nabla_y f(\bar{x}, \bar{y}) + \nabla_y(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T \beta + \nabla_y g(\bar{x}, \bar{y})^T \eta^g + \nabla_y G(\bar{x}, \bar{y})^T \eta^G, \tag{25}$$

$$\eta_i^G \ge 0 \quad i \in I_G, \qquad \eta_i^G = 0 \quad i \notin I_G, \tag{26}$$

$$\eta_i^g = 0 \quad i \in I_u, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \quad i \in I_+, \tag{27}$$

$$\text{either} \quad \eta_i^g > 0,\ (\nabla_y g(\bar{x}, \bar{y})\beta)_i > 0 \quad \text{or} \quad \eta_i^g (\nabla_y g(\bar{x}, \bar{y})\beta)_i = 0 \qquad i \in I_0.$$

Also for each partition of the index set I_0 into P, Q there exist $\lambda_i \ge 0$, $i = 1, \ldots, N$, $\sum_{i=1}^N \lambda_i = 1$, $\alpha^i \ge 0$, $\sum_{i=1}^{n+1} \alpha^i = 1$, $y^i \in S(\bar{x})$, $u^i \in M^1(\bar{x}, y^i)$, $i = 1, 2, \ldots, n+1$, and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that (24)–(27) and the following condition hold:

$$\eta_i^g \ge 0 \quad i \in P, \qquad (\nabla_y g(\bar{x}, \bar{y})\beta)_i \ge 0 \quad i \in Q.$$

If (CP) is weakly calm with modulus $\mu \ge 0$ at $(\bar{x}, \bar{y}, \bar{u})$, then there exist $\lambda_i \ge 0, i = 1, \ldots, N$ $\sum_{i=1}^N \lambda_i = 1$, $\alpha^i \ge 0$, $\sum_{i=1}^{n+1} \alpha^i = 1$, $y^i \in S(\bar{x})$, $u^i \in M^1(\bar{x}, y^i)$, $i = 1, 2, \ldots, n+1$, and $\beta \in R^m$, $\eta^g \in R^p$, $\eta^G \in R^q$ such that (24)–(27) holds and

$$\eta_i^g \ge 0, \quad (\nabla_y g(\bar{x}, \bar{y})\beta)_i \ge 0 \ \ i \in I_0.$$

Proof. Suppose that (CP) is MPEC-weakly calm. Then (22) holds at $(\bar{x}, \bar{y}, \bar{u})$ for some $\mu \ge 0$. Therefore, $(d, v) = (0, v)$ is an optimal solution to the following linearized problem:

$$\begin{aligned} \min_{(h,v)} \quad & \varphi(x, y, d) \\ \text{s.t.} \quad & \nabla(\nabla_y f + \bar{u}\nabla_y g)(\bar{x}, \bar{y})^T d + \nabla_y g(\bar{x}, \bar{y})^T v = 0, \\ & \nabla G_i(\bar{x}, \bar{y})^T d \le 0 \quad i \in I_G, \end{aligned}$$

$$\begin{aligned}&\nabla g_i(\bar{x}, \bar{y})^T d = 0 \quad i \in I_+\\&v_i = 0 \quad i \in I_u\\&\begin{cases}\nabla g_i(\bar{x}, \bar{y})^T d \cdot v_i = 0\\ \nabla g_i(\bar{x}, \bar{y})^T d \le 0, \quad v_i \ge 0\end{cases} \quad i \in I_0,\end{aligned} \tag{28}$$

where

$$\varphi_i(x, y, d) := [\nabla F_i(\bar{x}, \bar{y}) + \mu \nabla f(\bar{x}, \bar{y})]^T d - \mu \min_{\xi \in W(\bar{x})} \xi^T dx.$$

Let $\phi(z) := \min_{\xi \in W(\bar{x})} \xi^T z$. Because the set $W(\bar{x})$ is assumed to be nonempty and compact, by Danskin's theorem one has $\partial \phi(0) = \operatorname{co} W(\bar{x})$. Therefore, by Proposition 2.1

$$\partial^c \varphi_i(\bar{x}, \bar{y}, 0) \subset \nabla F_i(\bar{x}, \bar{y}) + \mu[\nabla f(\bar{x}, \bar{y}) - \operatorname{co} W(\bar{x}) \times \{0\}].$$

By Carathéodory's theorem, the convex set $\operatorname{co} W(\bar{x}) \subseteq R^n$ can be represented by not more than $n+1$ elements at a time. Therefore,

$$\operatorname{co} W(\bar{x}) = \left\{ \begin{array}{l} \sum_{i=1}^{n+1} \alpha^i (\nabla_x f(\bar{x}, y^i) + u^i \nabla_x g(\bar{x}, y^i)) \colon y^i \in S(\bar{x}), \quad u^i \in M^1(\bar{x}, y^i), \\ \qquad\qquad\qquad\qquad \alpha^i \ge 0, \ \sum_{i=1}^{n+1} \alpha^i = 1 \end{array} \right\}.$$

As in the proof of Theorem 4.2, the desired result follows in applying Theorem 2.1, and we omit the proof.

The proof of the assertion under the weakly calmness condition is similar, and we omit it. □

Remark 4.1. (i) A sufficient but not necessary condition for the set $W(\bar{x})$ to be nonempty and compact is that the MFCQ holds at every optimal solution of the lower-level problem $P_{\bar{x}}$ and the set-valued map $Y(x) := \{y \in R^m \colon g(x, y) \le 0\}$ is uniformly bounded around $\bar{x}$.

(ii) The new M, S, or P type optimality conditions obtained in Theorem 4.3 are in general weaker than the M-, S-, or P-stationary conditions based on the value function defined as in Definition 4.2 respectively since by the sensitivity of the value function

$$\partial^c V(\bar{x}) \subseteq \operatorname{co} W(\bar{x}). \tag{29}$$

However they are the most suitable surrogates for the C-, M-, S-, or P-stationary conditions, because the equality in (29) holds under certain conditions.

Acknowledgments. The author thanks two anonymous referees, the associate editor, and the area editor for their valuable comments that helped to significantly improve the presentation of the material in this paper. The research of this author was partially supported by NSERC.

References

[1] Bao, T. Q., B. S. Mordukhovich. 2010. Relative Pareto minimizers for multiobjective problems: Existence and optimality conditions. *Math. Programming Ser. A* **122** 301–347.
[2] Bao, T. Q., P. Gupta, B. S. Mordukhovich. 2007. Necessary conditions in multiobjective optimization with equilibrium constraints. *J. Optim. Theory Appl.* **135** 179–203.
[3] Bard, J. F. 1998. *Practical Bilevel Optimization: Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
[4] Clarke, F. H. 1983. *Optimization and Nonsmooth Analysis*. Wiley-Interscience, New York.
[5] Clarke, F. H., Y. S. Ledyaev, R. J. Stern, P. R. Wolenski. 1998. *Nonsmooth Analysis and Control Theory*. Springer, New York.
[6] Dempe, S. 2002. *Foundations of Bilevel Programming*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
[7] Dempe, S. 2003. Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. *Optimization* **52** 333–359.
[8] Dempe, S., J. Dutta. 2010. Is bilevel programming a special case of mathematical program with equilibrium constraints? *Math. Programming*. Forthcoming.
[9] Dempe, S., J. Dutta, B. S. Mordukhovich. 2007. New necessary optimality conditions in optimistic bilevel programming. *Optim.* **56** 577–604.
[10] Gauvin, J., F. Dubeau. 1982. Differential properties of the marginal function in mathematical programming. R. W. Cottle, L. C. W. Dixon, B. Korte, T. L. Magnanti, M. J. Todd, E. L. Allgower, R. Bartels et al., eds. *Optimality and Stability in Mathematical Programming, Mathematical Programming Studies*, Vol. 19. Springer, Berlin/Heidelberg, 101–119.
[11] Hager, W. W., M. S. Gowda. 1999. Stability in the presence of degeneracy and error estimation. *Math. Programming* **85** 181–192.

[12] Ioffe, A. D., J. V. Outrata. 2008. On metric and calmness qualification conditions in subdifferential calculus. *Set-Valued Anal.* **16** 199–227.

[13] Lucet, Y., J. J. Ye. 2001. Sensitivity analysis of the value function for optimization problems with variational inequality constraints. *SIAM J. Control Optim.* **40** 699–723.

[14] Lucet, Y., J. J. Ye. 2002. Erratum: Sensitivity analysis of the value function for optimization problems with variational inequality constraints. *SIAM J. Control Optim.* **41** 1315–1319.

[15] Luo, Z. Q., J. S. Pang, D. Ralph. 1996. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, UK.

[16] Mirrlees, J. 1999. The theory of moral hazard and unobservable behaviour—Part I. *Rev. Econom. Stud.* **66** 3–22.

[17] Mordukhovich, B. S. 2006. *Variational Analysis and Generalized Differentiation*, Vol. 1: *Basic Theory*. Springer, Berlin.

[18] Mordukhovich, B. S. 2006. *Variational Analysis and Generalized Differentiation*, Vol. 2: *Applications*. Springer, Berlin.

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

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

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

[22] Vicente, L. N., P. H. Calamai. 1994. Bilevel and multilevel programming: A bibliography review. *J. Global Optim.* **5** 291–306.

[23] von Stackelberg, H. 1934. *Marktform and Gleichgewicht* [*The theory of the market economy*]. Springer-Verlag, Berlin.

[24] Wu, Z., J. J. Ye. 2003. First-order and second-order conditions for error bounds. *SIAM J. Optim.* **14** 621–645.

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

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

[27] Ye, J. J. 2004. Nondifferentiable multiplier rules for optimization and bilevel optimization problems. *SIAM J. Optim.* **15** 252–274.

[28] Ye, J. J. 2005. Necessary and sufficient optimality conditions for mathemtical programs with equilibrium constraints. *J. Math. Anal. Appl.* **307** 350–369.

[29] Ye, J. J. 2006. Constraint qualifications and KKT conditions for bilevel programming problems. *Math. Oper. Res.* **31** 811–824.

[30] Ye, J. J., D. L. Zhu. 1995. Optimality conditions for bilevel programming problems. *Optimization* **33** 9–27.

[31] Ye, J. J., D. L. Zhu. 1997. A note on optimality conditions for bilevel programming problems. *Optimization* **39** 361–366.

[32] Ye, J. J., D. L. Zhu. 2010. New necessary optimality conditions for bilevel programs by combining the MPEC and value function approaches. *SIAM J. Optim.* **20** 1885–1905.

[33] Ye, J. J., Q. J. Zhu. 2003. Multiobjective optimization problems with variational inequality constraints. *Math. Programming Ser. A* **96** 139–160.

[34] Ye, J. J., D. L. Zhu, Q. J. Zhu. 1997. Exact penalization and necessary optimality conditions for generalized bilevel programming problems. *SIAM J. Optim.* **2** 481–507.

[35] Zhu, Q. J. 2000. Hamiltonian necessary conditions for a multiobective optimal control problem with endpoint constraints. *SIAM J. Control Optim.* **39** 97–112.