MATHEMATICS OF OPERATIONS RESEARCH
Vol. 22, No. 4, November 1997
Printed in U.S.A.

NECESSARY OPTIMALITY CONDITIONS FOR OPTIMIZATION PROBLEMS WITH VARIATIONAL INEQUALITY CONSTRAINTS

J. J. YE AND X. Y. YE

In this paper we study optimization problems with variational inequality constraints in finite dimensional spaces. Kuhn-Tucker type necessary optimality conditions involving coderivatives are given under certain constraint qualifications including one that ensures nonexistence of nontrivial abnormal multipliers. The result is applied to bilevel programming problems to obtain Kuhn-Tucker type necessary optimality conditions. The Kuhn-Tucker type necessary optimality conditions are shown to be satisfied without any constraint qualification by the class of bilevel programming problems where the lower level is a parametric linear quadratic problem.

1. Introduction. *An optimization problem with variational inequality constraints* (OPVIC) is a special class of optimization problems over variables x and y in which some or all its constraints are defined by a parametric variational inequality with y as its primary variable and x the parameter. In this paper we consider optimization problems with variational inequality constraints in finite dimensional spaces defined as follows:

$$\text{(OPVIC)} \qquad \text{minimize} f(x, y) \quad \text{subject to } x \in \Omega_1 \quad \text{and} \quad y \in S(x),$$

where $f : R^{n+m} \to R$, Ω_1 is a nonempty subset of R^n, and for each $x \in \Omega_1 S(x)$ is the *solution set of a variational inequality with parameter* x, i.e.,

$$S(x) = \{y \in \Omega_2(x) : \langle F(x, y), y - z\rangle \geq 0 \ \forall z \in \Omega_2(x)\},$$

where $F : R^{n+m} \to R^m$ and $\Omega_2 : R^n \rightrightarrows R^m$ is a set-valued map. The above problem is also refered to as generalized bilevel programming problems or mathematical programs with equilibrium constraints (see, e.g., Luo, Pang, Ralph and Wu 1996, Ye, Zhu and Zhu 1997). Throughout this paper, we shall make the blanket assumption that the set $\{(x, y) : x \in \Omega_1, y \in S(x)\}$ is nonempty.

In the case where the set-valued map $\Omega_2(x)$ is convex-valued and $F(x, y)$ is a gradient of a real-valued differentiable and pseudo-convex function, i.e., $F(x, y) = -\nabla_y g(x, y)$, where $g : R^{n+m} \to R$ is differentiable and pseudo-convex in y, the optimization problem with variational inequality constraints (OPVIC) is the following *bilevel programming problem*, or so called Stackelberg game (see, e.g., Anandalingam and Friesz 1992, Dempe 1992, Loridan and Morgan 1989, Outrata 1993, Von Stackelberg 1952, Ye and Zhu 1995, 1996 and Ye 1996):

$$\text{(BP)} \qquad \text{minimize} f(x, y) \quad \text{subject to } x \in \Omega_1 \quad \text{and} \quad y \in S(x)$$

where $S(x)$ is the set of solutions of the problem (P_x):

Received March 17, 1996; revised January 1, 1997, March 6, 1997.
AMS 1991 subject classification. Primary: 49K99, 90C99; Secondary: 90D65.
OR/MS Index 1978 subject classification. Primary: Programming/Variational inequality constraints; Secondary: Programming/Bilevel.
Key words. Optimization problems, variational inequality constraints, necessary optimality conditions, derivatives of set-valued maps, nonsmooth analysis.

0364-765X/97/2204/0977/$05.00

$$(P_x) \qquad \text{minimize } g(x, y) \quad \text{subject to } y \in \Omega_2(x).$$

The last decade has witnessed a growing interest in the theory of (OPVIC). See Anandalingam and Friesz (1992), Barbu (1984), Friesz, Tobin, Cho and Mehta (1990), Outrata (1994), Shi (1988, 1990), Ye, Zhu and Zhu (1997). A natural approach to obtain necessary optimality conditions for (OPVIC) is to reduce (OPVIC) to a ordinary (single level) mathematical programming problem and use the existing necessary optimality conditions for the single level problem. There are several equivalent single level formulations for (OPVIC). To illustrate we assume that

$$\Omega_2(x) := \{y \in R^m : \psi(x, y) \leq 0\}$$

where $\psi : R^{n+m} \to R^q$. The Karush-Kuhn-Tucker (KKT) approach is to interpret the variational inequality constraint $y \in S(x)$ as y being a solution of the following optimization problem:

$$\min\langle -F(x, y), z\rangle \quad \text{s.t. } z \in \Omega_2(x),$$

and replace it by the KKT necessary optimality conditions for the above optimization problem with is also sufficient if $\Omega_2(x)$ is convex for each x. Using this approach, under the assumption that $\Omega_2(x)$ is convex, $F(x, y)$ is differentiable and a certain constraint qualification condition holds for the above optimization problem, $(\bar{x}, \bar{y})$ is a solution of (OPVIC) if and only if there exists $\bar{u} \in R^q$ such that $(\bar{x}, \bar{y}, \bar{u})$ is a solution of the following problem:

$$(\text{KS}) \qquad \begin{aligned} \min \quad & f(x, y) \\ \text{s.t.} \quad & -F(x, y) + \nabla_y\psi(x, y)^\top u = 0, \\ & \langle u, \psi(x, y)\rangle = 0, \\ & u \geq 0, \qquad \psi(x, y) \leq 0, \\ & x \in \Omega_1, \qquad y \in R^m. \end{aligned}$$

Other approaches for single level formulation of (OPVIC) include the value function approach and the gap function approach and many others (see Ye and Zhu 1995, Ye, Zhu and Zhu 1997). A common phenomenon for these equivalent single level formulations is the existence of a nontrivial abnormal multiplier (see Proposition 3.2 of Ye and Zhu 1995 and Proposition 1.1 of Ye, Zhu and Zhu 1997) which is equivalent to saying that the usual constraint qualification conditions such as Mangasarian-Fromovitz condition will never be satisfied (see Proposition 3.1 of Ye and Zhu 1995) for these single level problems.

In this paper we formulate problem (OPVIC) where $\Omega_2 := \Omega_2(x)$ is convex and independent of x as the following optimization problem with a generalized equation constraint:

$$(\text{GP}) \qquad \begin{aligned} \min \quad & f(x, y) \\ \text{s.t.} \quad & (x, y) \in \Omega_1 \times R^m, \\ & 0 \in -F(x, y) + N(y, \Omega_2), \end{aligned}$$

where

$$N(y, \Omega_2) := \begin{cases} \text{the normal cone of } \Omega_2 \text{ at } y & \text{if } y \in \Omega_2 \\ \varnothing & \text{if } y \notin \Omega_2 \end{cases}$$

is the normal cone operator. Note that the assumption that $\Omega_2(x) = \Omega_2$ is independent of x is not very restrictive since for example any bilevel programming problem with convex lower level problem where the constraints are systems of inequalities and equalities satisfying a certain constraint qualification can be reformulated as (GP) (see the discussion in §4).

In this paper we introduce the concept of pseudo upper-Lipschitz continuity for a set-valued map (see Definition 2.8) which is implied by either upper-Lipschitz continuity (see Definition 2.7) or pseudo-Lipschitz continuity (see Definition 2.6). Under suitable regularity conditions on the problem data we show that if $(\bar{x}, \bar{y})$ is a solution of (OPVIC) and the set-valued map

$$\Sigma(v) := \{(x, y) \in \Omega_1 \times R^m : v \in -F(x, y) + N(y, \Omega_2)\}$$

is pseudo upper-Lipschitz continuous at $(0, (\bar{x}, \bar{y}))$, then there exist $\eta \in R^m$ such that

$$0 \in \partial_x f(\bar{x}, \bar{y}) - \nabla_x F(\bar{x}, \bar{y})^\top \eta + N(\bar{x}, \Omega_1)$$

$$0 \in \partial_y f(\bar{x}, \bar{y}) - \nabla_y F(\bar{x}, \bar{y})^\top \eta + D^* N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta),$$

where ∂ denotes the limiting subgradient (see Definition 2.3), N_{Ω_2} denote the set-valued map $y \rightrightarrows N(y, \Omega_2)$ and D^* denotes the coderivative of a set-valued map (see Definition 2.5). This is in fact in the form of the optimality condition given by Shi (1988, 1990) with the paratingent coderivative of the set-valued map N_{Ω_2} replaced by the Mordukhovich coderivative.

We shall call a vector $\eta \in R^m$ an abnormal multiplier for (GP) if it satisfies

$$0 \in -\nabla_x F(\bar{x}, \bar{y})^\top \eta + N(\bar{x}, \Omega_1)$$

$$0 \in -\nabla_y F(\bar{x}, \bar{y})^\top \eta + D^* N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta).$$

Sufficient conditions for pseudo-Lipschitz continuity of the set-valued map Σ include local strong monotonicity of $F(x, y)$ in variable y uniformly in x and nonexistence of nontrivial abnormal multipliers for (GP). In particular for the bilevel programming problem where the lower level problem is a parametric linear quadratic problem, the set-valued map Σ is a polyhedral set-valued map hence upper-Lipschitz continuous according to Robinson (1981). Therefore the Kuhn-Tucker condition derived in this paper is always satisfied by the class of the bilevel programming problem where the lower level problem is a parametric linear quadratic problem *without any constraint qualification.*

We organize the paper as follows. Section 1 contains background material on nonsmooth analysis and preliminary results. In §2 we derive the Kuhn-Tucker type necessary optimality conditions for (OPVIC). Applications to bilevel programming problems are given in §3 where we show that the bilevel programming problem where the lower level is a parametric quadratic programming problem always satisfies the constraint qualification. An example is given to illustrate the application of the theory.

2. Preliminaries. This section contains some background material on nonsmooth analysis and preliminary results which will be used in the next section. We only give concise definitions that will be needed in the paper. For more detail informations on the subject, our references are Aubin and Ekeland (1984), Aubin and Frankowska (1990), Clarke (1983), Mordukhovich (1994a, b).

First we give some concepts for various normal cones.

DEFINITION 2.1. Let Ω be a nonempty subset of R^n. Given $\bar{z} \in \mathrm{cl}\Omega$, the closure of set Ω, the following convex cone

$$N^\pi(\bar{z}, \Omega) := \{\xi \in R^n : \exists M > 0 \text{ s.t. } \langle \xi, z - \bar{z} \rangle \leq M\|z - \bar{z}\|^2, \forall z \in \Omega\}$$

is called the proximal normal cone to set Ω at point $\bar{z}$.

Equivalently, $\xi \in N^\pi(\bar{z}, \Omega)$ if and only if there exists $t > 0$ such that

$$d_\Omega(\bar{z} + t\xi) = t\|\xi\|$$

where $d_\Omega(z) := \inf\{\|z - z'\| : \forall z' \in \Omega\}$ denotes the distance of a point z to a set Ω.

DEFINITION 2.2. Given $\bar{z} \in \mathrm{cl}\Omega$, the following closed cone

$$N(\bar{z}, \Omega) := \left\{\lim_{i \to \infty} \xi_i : \xi_i \in N^\pi(z_i, \Omega), z_i \to \bar{z}\right\}$$

is called the limiting normal cone to Ω at point $\bar{z}$ and the closed convex hull of the limiting normal cone

$$N_C(\bar{z}, \Omega) := \mathrm{clco} N(\bar{z}, \Omega)$$

is called the Clarke normal cone to set Ω at point $\bar{z}$.

The following calulus for limiting normal cones will be useful later.

PROPOSITION 2.1 (COROLLARY 4.7 OF MORDUKHOVICH 1994b). *Let Ω_1 and Ω_2 be closed subsets of R^n and let $\bar{z} \in \Omega_1 \cap \Omega_2$. If $N(\bar{z}, \Omega_1) \cap (-N(\bar{z}, \Omega_2)) = \{0\}$, then one has the inclusion*

$$N(\bar{z}, \Omega_1 \cap \Omega_2) \subset N(\bar{z}, \Omega_1) + N(\bar{z}, \Omega_2).$$

PROPOSITION 2.2. *Let $\Phi : R^n \Rightarrow R^q$ be a set-valued map and $\Phi^{-1} : R^q \Rightarrow R^n$ be the inverse set-valued map to Φ defined by $\Phi^{-1}(v) := \{z \in R^n : v \in \Phi(z)\}$. Then*

(a) $(z, v) \in gph\Phi$ *iff* $(v, z) \in gph\Phi^{-1}$

(b) $(\xi, \eta) \in N((z, v), gph\Phi)$ *iff* $(\eta, \xi) \in N((v, z), gph\Phi^{-1})$,

where $gph\,\Phi := \{(z, v) : z \in R^n, v \in \Phi(z)\}$ *is the graph of* Φ.

PROOF. (a) follows by the definition of an inverse map. As to (b), by Definition of a proximal normal cone (see Definition 2.1) and (a) we have $(\xi, \eta) \in N^\pi((z, v), \mathrm{gph}\Phi)$ if only if $(\eta, \xi) \in N^\pi((v, z), \mathrm{gph}\Phi^{-1})$. (b) then follows from the definition of a limiting normal cone (see Definition 2.2). □

Using the definitions for normal cones, we now give definitions for subgradients of a single-valued map.

DEFINITION 2.3. Let $f : R^n \to R \cup \{+\infty\}$ be lower semicontinuous and finite at $\bar{z} \in R^n$. The limiting subgradient of f at $\bar{z}$ is defined by

$$\partial f(\bar{z}) := \{\xi : (\xi, -1) \in N((\bar{z}, f(\bar{z})), \mathrm{epi}(f))\},$$

and the Clarke generalized gradient of f at $\bar{z}$ is defined by

$$\partial_C f(\bar{z}) := \{\xi : (\xi, -1) \in N_C((\bar{z}, f(\bar{z})), \text{epi}(f))\},$$

where $\text{epi}(f) := \{(x, r) \in R^n \times R : f(x) \le r\}$ is the epigraph of f.

Let $f: R^{n+m} \to R$ be Lipschitz continuous near $(\bar{x}, \bar{y})$. In general there is no relationship between subgradients and partial subgradients. However under a certain regularity condition, certain relationships exist. We first define the regularity.

DEFINITION 2.4. Ω, a subset of R^n is said to be regular at $\bar{z}$ if $N(\bar{z}, \Omega) = N^\pi(\bar{z}, \Omega)$. $f: R^n \to R$ is said to be regular at $\bar{z}$ if its epigraph is regular at $\bar{z}$.

REMARK 2.1. If Ω is regular at $\bar{z}$, then $N(\bar{z}, \Omega)$ is convex and hence $N(\bar{z}, \Omega) = N_C(\bar{z}, \Omega)$. Apparently, if $f(z)$ is regular at $\bar{z}$, then $\partial_C f(\bar{z}) = \partial f(\bar{z})$.

The following proposition is a relationship between the limiting subgradients and their partial limiting subgradients.

PROPOSITION 2.3 (PROPOSITION 2.3.15 OF CLARKE 1983). *Let $f: R^{n+m} \to R$ be Lipschitz continuous near $(\bar{x}, \bar{y})$. If f is regular at $(\bar{x}, \bar{y})$ then we have*

$$\partial f(\bar{x}, \bar{y}) \subset \partial_x f(\bar{x}, \bar{y}) \times \partial_y f(\bar{x}, \bar{y}).$$

For set-valued maps, the definition for limiting normal cone leads to the definition of coderivative of a set-valued map introduced by Mordukhovich (see, e.g., Mordukhovich 1994b).

DEFINITION 2.5. Let $\Phi : R^n \rightrightarrows R^q$ be an arbitrary set-valued map (assigning to each $z \in R^n$ a set $\Phi(z) \subset R^q$ which may be empty) and $(\bar{x}, \bar{y}) \in \text{cl gph}\Phi$. The set-valued map $D^*\Phi(\bar{z}, \bar{v})$ from R^q into R^n defined by

$$D^*\Phi(\bar{z}, \bar{v})(\eta) = \{\xi \in R^n : (\xi, -\eta) \in N((\bar{z}, \bar{v}), \text{gph}\Phi)\},$$

is called the coderivative of Φ at point $(\bar{z}, \bar{v})$. By convention for $(\bar{z}, \bar{v}) \notin \text{cl gph}\Phi$ we define $D^*\Phi(\bar{z}, \bar{v})(\eta) = \emptyset$. The symbol $D^*\Phi(\bar{z})$ is used when Φ is single-valued at $\bar{z}$ and $\bar{v} = \Phi(\bar{z})$.

In the special case when a set-valued map is single-valued, the coderivative is related to the limiting subgradient in the following way.

PROPOSITION 2.4 (PROPOSITION 2.11 OF MORDUKHOVICH 1994b). *Let $\Phi : R^n \to R^q$ be single-valued and Lipschitz continuous near $\bar{z}$. Then*

$$D^*\Phi(\bar{z})(\eta) = \partial\langle \eta, \Phi\rangle(\bar{z}) \qquad \forall \eta \in R^q.$$

The following proposition is a sum rule for coderivatives.

PROPOSITION 2.5 (COROLLARY 4.4 OF MORDUKHOVICH 1994b). *Let $\Phi_1 : R^n \to R^q$ be strictly differentiable at $\bar{z}$ with the Jacobian $\nabla\Phi_1(\bar{z}) \in R^{q\times n}$, i.e.,*

$$\lim_{z,z'\to\bar{z}} \frac{\Phi_1(z) - \Phi_1(z') - \nabla\Phi_1(\bar{z})(z - z')}{\|z - z'\|} = 0$$

and $\Phi_2 : R^n \rightrightarrows R^q$ be an arbitrary closed set-valued map. Then for any $\bar{v} \in \Phi_1(\bar{z}) + \Phi_2(\bar{z})$ and $\eta \in R^q$ one has

$$D^*(\Phi_1 + \Phi_2)(\bar{z}, \bar{v})(\eta) = \nabla\Phi_1(\bar{z})^\top\eta + D^*\Phi_2(\bar{z}, \bar{v} - \Phi_1(\bar{z}))(\eta).$$

We now discuss Lipschitz behavior of a set-valued map. The following concept for locally Lipschitz behavior was introduced by Aubin (1994).

DEFINITION 2.6. A set-valued map $\Phi : R^n \rightrightarrows R^q$ is said to be pseudo-Lipschitz continuous around $(\bar{z}, \bar{v}) \in \text{gph}\Phi$ with modulus $\mu \geq 0$ if there exist a neighborhood U of $\bar{z}$, a neighborhood V of $\bar{v}$ such that

$$\Phi(z) \cap V \subset \Phi(z') + \mu\|z' - z\|B, \qquad \forall z', z \in U,$$

where B denotes the closed unit ball in the appropriate space.

On the other hand the following upper-Lipschitz behavior was studied by Robinson (1975, 1976).

DEFINITION 2.7. A set-valued map $\Phi : R^n \rightrightarrows R^q$ is said to be (locally) upper-Lipschitz continuous at $\bar{z} \in R^n$ with modulus $\mu \geq 0$ if there exists a neighborhood U of $\bar{z}$ such that

$$\Phi(z) \subset \Phi(\bar{z}) + \mu\|z - \bar{z}\|B, \qquad \forall z \in U.$$

It is easy to see that in general, pseudo-Lipschitz continuity of Φ around $(\bar{z}, v)$ for all $v \in \Phi(\bar{z})$ does not imply upper-Lipschitz continuity of Φ at $\bar{z}$ and vice versa. In order to establish the relationship between the upper-Lipschitz continuity with the pseudo-Lipschitz continuity, we introduce the following weaker concept of locally upper-Lipschitz continuity.

DEFINITION 2.8. A set-valued map $\Phi : R^n \rightrightarrows R^q$ is said to be pseudo upper-Lipschitz continuous at $(\bar{z}, \bar{v}) \in \text{gph}\Phi$ with modulus $\mu \geq 0$ if there exist a neighborhood U of $\bar{z}$, a neighborhood V of $\bar{v}$ such that

$$\Phi(z) \cap V \subset \Phi(\bar{z}) + \mu\|z - \bar{z}\|B, \qquad \forall z \in U.$$

It turns out that the pseudo upper-Lipschitz continuity is implied by either upper-Lipschitz continuity or pseudo-Lipschitz continuity.

PROPOSITION 2.6. (1) *If Φ is upper-Lipschitz continuous at $\bar{z}$ then Φ is pseudo upper-Lipschitz continuous at $(\bar{z}, v)$ for all $v \in \Phi(\bar{z})$.*

(2) *If Φ is pseudo-Lipschitz continuous around $(\bar{z}, \bar{v}) \in gph\Phi$, then Φ is pseudo upper-Lipschitz continuous at $(\bar{z}, \bar{v})$.*

The following useful criterion for pseudo-Lipschitz continuity of a set-valued map was given in Proposition 3.5 of Mordukhovich (1994c).

PROPOSITION 2.7. *Let $\Phi : R^n \rightrightarrows R^q$ be a set-valued map with a closed graph. Then Φ is pseudo-Lipschitz continuous around $(\bar{z}, \bar{v}) \in gph\Phi$ if and only if*

$$D^*\Phi(\bar{z}, \bar{v})(0) = \{0\}.$$

Using this criterion we give conditions to ensure pseudo-Lipschitz continuity of the solution map for a perturbed generalized equation which will be useful in §3.

PROPOSITION 2.8. *Let Ω be a closed subset of R^n and $\bar{z} \in \Omega$. Suppose that function $h : R^n \to R$ is strictly differentiable around $\bar{z}$ and $Q : R^n \rightrightarrows R^q$ is a set-valued map with a close graph. Further assume that*

(a) $D^*Q(\bar{z}, h(\bar{z}))(0) \cap \{-N(\bar{z}, \Omega)\} = \{0\}$.

(b) *There is no nonzero vector $\eta \in R^n$ such that*

$$0 \in -\nabla h(\bar{z})^{\top}\eta + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega).$$

Then the solution map for the perturbed generalized equation

$$\Sigma(v) := \{z \in \Omega : v \in -h(z) + Q(z)\}$$

is pseudo-Lipschitz continuous around $(0, \bar{z})$.

PROOF. Denote by $G(z) := -h(z) + Q(z)$. Then $\Sigma(v) = G^{-1}(v) \cap \Omega$. By Proposition 2.7 it is sufficient to show $D^*\Sigma(0, \bar{z})(0) = \{0\}$.

First we show that

$$(1) \qquad N((0, \bar{z}), \mathrm{gph}G^{-1}) \cap (-N((0, \bar{z}), R^n \times \Omega)) = \{0\}.$$

Indeed, for any $(0, q) \in N((0, \bar{z}), \mathrm{gph}G^{-1})$ where $q \in -N(\bar{z}, \Omega)$, by Proposition 2.2, one has $(q, 0) \in N((\bar{z}, 0), \mathrm{gph}G)$ and hence from Proposition 2.5,

$$q \in D^*G(\bar{z}, 0)(0) = -\nabla h(\bar{z})^{\top}0 + D^*Q(\bar{z}, h(\bar{z}))(0) = D^*Q(\bar{z}, h(\bar{z}))(0).$$

Hence equality (1) follows from assumption (a).

Now suppose that $\eta \in D^*\Sigma(0, \bar{z})(0)$, which means by the definition of coderivatives that $(\eta, 0) \in N((0, \bar{z}), \mathrm{gph}\Sigma)$. It is easy to see that $\mathrm{gph}\Sigma = \mathrm{gph}G^{-1} \cap (R^n \times \Omega)$. Since (1) holds, we can apply Proposition 2.1 and obtain

$$\begin{aligned}(\eta, 0) &\in N((0, \bar{z}), \mathrm{gph}G^{-1}) + N((0, \bar{z}), R^n \times \Omega) \\ &= N((0, \bar{z}), \mathrm{gph}G^{-1}) + \{0\} \times N(\bar{z}, \Omega).\end{aligned}$$

That is, there exist $q_1, q_2 \in R^n$ such that $(\eta, q_1) \in N((0, \bar{z}), \mathrm{gph}G^{-1})$, $q_2 \in N(\bar{z}, \Omega)$ and $q_1 + q_2 = 0$. By virtue of Proposition 2.2, $(q_1, \eta) \in N((\bar{z}, 0), \mathrm{gph}G)$ which implies that $q_1 \in D^*G(\bar{z}, 0)(-\eta)$. By Proposition 2.5,

$$q_1 \in D^*G(\bar{z}, 0)(-\eta) = \nabla h(\bar{z})^{\top}\eta + D^*Q(\bar{z}, h(\bar{z}))(-\eta).$$

Since $-q_1 = q_2 \in N(\bar{z}, \Omega)$ the above inclusion becomes

$$0 \in -\nabla h(\bar{z})^{\top}(-\eta) + D^*Q(\bar{z}, h(\bar{z}))(-\eta) + N(\bar{z}, \Omega).$$

By assumption (b) we deduce from above inclusion that $\eta = 0$. That is $D^*\Sigma(0, \bar{z})(0) = \{0\}$. Hence Σ is pseudo-Lipschitz continuous around $(0, \bar{z})$ by Proposition 2.7. □

3. Necessary optimality conditions. The purpose of this section is to derive necessary optimality conditions involving coderivatives for optimization problem with variational inequality constraints (OPVIC).

First, we consider the following optimization problem with a generalized equation constraint:

$$(P_0) \qquad \begin{aligned} &\min \quad f(z) \\ &\text{s.t.} \quad 0 \in \Phi(z), \\ &\qquad\quad z \in \Omega, \end{aligned}$$

where $\Omega \subset R^{n+m}$, $f : \Omega \to R$ and $\Phi : R^{n+m} \rightrightarrows R^q$ is a set-valued map. It is obvious that (P_0) can be rewritten as the following problem:

$$(P'_0) \qquad \begin{aligned} \min \quad & f(z) \\ \text{s.t.} \quad & v \in \Phi(z), \qquad v = 0, \\ & z \in \Omega. \end{aligned}$$

In the following lemma, we show that the problem (P'_0) is actually equivalent to its localized penalized problem (P_r) when the function f is Lipschitz continuous and the set-valued map $\Phi^{-1} \cap \Omega$ is pseudo upper-Lipschitz continuous.

LEMMA 3.1. *Suppose f is Lipschitz continuous of rank $L_f \geq 0$. Assume that $\bar{z}$ solves (P_0) and the set-valued map $\Phi^{-1} \cap \Omega$ defined by $(\Phi^{-1} \cap \Omega)(v) := \Phi^{-1}(v) \cap \Omega$ is pseudo upper-Lipschitz continuous with modulus $\mu \geq 0$ at $(0, \bar{z})$. Then there exist a neighborhood U of $\bar{z}$ and a neighborhood V of 0 such that $(v, z) = (0, \bar{z})$ solves the following localized penalized problem of (P'_0) for all $r \geq L_f \mu$.:*

$$(P_r) \qquad \begin{aligned} \min \quad & f(z) + r\|v\| \\ s.t. \quad & v \in \Phi(z) \cap V, \\ & z \in \Omega \cap U. \end{aligned}$$

PROOF. Since $\Phi^{-1} \cap \Omega$ is pseudo upper-Lipschitz continuous at $(0, \bar{z})$, there exist U, a neighborhood of $\bar{z}$ and V, a neighborhood of 0 such that for any $z \in \Phi^{-1}(v) \cap \Omega \cap U$ there exists $z^* \in \Phi^{-1}(0) \cap \Omega$ such that

$$\|z - z^*\| \leq \mu\|v\|, \qquad \forall v \in V. \tag{2}$$

Thus we have for any $z \in \Phi^{-1}(v) \cap \Omega \cap U, v \in V$ (i.e., $v \in \Phi(z) \cap V, z \in \Omega \cap U$),

$$\begin{aligned} f(\bar{z}) &\leq f(z^*) \quad \text{since } \bar{z} \text{ is a solution of } (P_0) \\ &= f(z) + (f(z^*) - f(z)) \\ &\leq f(z) + L_f\|z^* - z\| \quad \text{by Lipschitz continuity of } f \\ &\leq f(z) + L_f \mu\|v\| \quad \text{by virtue of } (2) \\ &\leq f(z) + r\|v\|, \end{aligned}$$

for all $r \geq L_f \mu$. The proof is complete. □

We now give a Kuhn-Tucker type necessary optimality condition for problem (P_0). A similar result in Proposition 4.3 of Zhang (1994) was proved under upper Lipschitz continuity assumption. The use of pseudo upper-Lipschitz continuity assumption significantly enlarges applicability of the result since pseudo-Lipschitz continuity implies pseudo upper-Lipschitz continuity while upper-Lipschitz continuity may not imply pseudo upper-Lipschitz continuity.

THEOREM 3.1. *Suppose f is Lipschitz continuous of rank $L_f \geq 0$. Let $\bar{z}$ be a solution of (P_0) and $\Phi^{-1} \cap \Omega$ be pseudo upper-Lipschitz continuous at $(0, \bar{z})$ with modulus μ. Suppose that $D^*\Phi(\bar{z}, 0)(0) \cap (-N(\bar{z}, \Omega)) = \{0\}$. Then for any $r \geq L_f\mu$ there exists $\eta \in rB_q$ such that*

$$0 \in \partial f(\bar{z}) + D^*\Phi(\bar{z}, 0)(\eta) + N(\bar{z}, \Omega), \tag{3}$$

where B_q denotes the closed unit ball in R^q.

PROOF. By Lemma 3.1 we know that $(v, z) = (0, \bar{z})$ is a solution of (P_r). Rewrite (P_r) in the following form:

$$(\bar{P}_r) \qquad \begin{aligned} &\min \quad f(z) + r\|v\| \\ &\text{s.t.} \quad (z, v) \in \text{gph}\Phi \cap (\Omega \times R^q) \cap (U \times V). \end{aligned}$$

By the well-known optimality condition we have

$$\begin{aligned} (0, 0) &\in \partial f(\bar{z}) \times (rB_q) + N((\bar{z}, 0), \text{gph}\Phi \cap (U \times V) \cap (\Omega \times R^q)) \\ &= \partial f(\bar{z}) \times (rB_q) + N((\bar{z}, 0), \text{gph}\Phi \cap (\Omega \times R^q)), \end{aligned} \tag{4}$$

since U and V are neighborhood of $\bar{z}$ and 0 respectively.

The assumption $D^*\Phi(\bar{z}, 0)(0) \cap (-N(\bar{z}, \Omega)) = \{0\}$ is equivalent to

$$N((\bar{z}, 0), \text{gph}\Phi) \cap (-N((\bar{z}, 0), \Omega \times R^q)) = \{0\}.$$

Thus by Proposition 2.1, (4) implies that

$$(0, 0) \in \partial f(\bar{z}) \times (rB_q) + N((\bar{z}, 0), \text{gph}\Phi) + N_\Omega(\bar{z}) \times \{0\}.$$

That is, there exist $(p_1, \eta) \in \partial f(\bar{z}) \times (rB_q)$, $(p_2, q) \in N((\bar{z}, 0), \text{gph}\Phi)$ and $p_3 \in N_\Omega(\bar{z})$ such that

$$p_1 + p_2 + p_3 = 0, \qquad \eta + q = 0. \tag{5}$$

By the definition of coderivatives, $(p_2, q) \in N((\bar{z}, 0), \text{gph}\Phi)$ implies that

$$p_2 \in D^*\Phi(\bar{z}, 0)(-q) = D^*\Phi(\bar{z}, 0)(\eta).$$

And hence there exists $\eta \in rB_q$ such that

$$0 = p_1 + p_2 + p_3 \in \partial f(\bar{z}) + D^*\Phi(\bar{z}, 0)(\eta) + N_\Omega(\bar{z}).$$

The proof is complete. □

Now we consider the optimization problems with variational inequality constraints (OPVIC) where Ω_2 is a closed convex subset of R^m. Since Ω_2 is a convex set, by the definition of a normal cone in the sense of convex analysis, it is easy to see that problem (OPVIC) can be rewritten as the optimization problem with generalized equation constraints (GP). By applying the previous theorem to the equivalent problem (GP), we obtain the following Kuhn-Tucker necessary optimality conditions for (OPVIC).

THEOREM 3.2. *Let Ω_2 be a closed and convex subset of R^m and $(\bar{x}, \bar{y})$ a solution of problem (OPVIC). Assume that f is Lipschitz continuous and regular at $(\bar{x}, \bar{y})$ and F is continuously differentiable around $(\bar{x}, \bar{y})$. Further assume that one of the following assumptions is satisfied*:

(a) *The set of solutions to the perturbed generalized equation*

$$\Sigma(v) := \{(x, y) \in \Omega_1 \times R^m : v \in -F(x, y) + N(y, \Omega_2)\} \tag{6}$$

is pseudo upper-Lipschitz continuous at $(0, (\bar{x}, \bar{y}))$.

(b) $-F$ *is locally strongly monotone in y uniformly in x with modulus* $\delta > 0$, *i.e., there exist neighborhoods* U_1 *of* $\bar{x}$ *and* U_2 *of* $\bar{y}$ *such that*

$$\langle -F(x, y) + F(x, z), y - z\rangle \geq \delta\|y - z\|^2 \qquad \forall y \in U_2 \cap \Omega_2, z \in \Omega_2, x \in U_1 \cap \Omega_1.$$

(c) Ω_1 *is a closed subset of* R^n *and there is no nonzero vector* $\eta \in R^m$ *such that*

$$0 \in -\nabla_x F(\bar{x}, \bar{y})^\top \eta + N(\bar{x}, \Omega_1),$$

$$0 \in -\nabla_y F(\bar{x}, \bar{y})^\top \eta + D^* N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta),$$

where N_{Ω_2} *denotes the set-valued map* $y \Rightarrow N(y, \Omega_2)$.

(d) *The set* $\Omega_1 = R^n$ *and the rank of the matrix* $\nabla_x F(\bar{x}, \bar{y})$ *is m*.

Then there exist $r > 0$, $\eta \in rB_m$ *such that*

$$0 \in \partial_x f(\bar{x}, \bar{y}) - \nabla_x F(\bar{x}, \bar{y})^\top \eta + N(\bar{x}, \Omega_1),$$

$$0 \in \partial_y f(\bar{x}, \bar{y}) - \nabla_y F(\bar{x}, \bar{y})^\top \eta + D^* N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta).$$

Before proving Theorem 3.2, we first prove the following lemmas under the same assumptions specified in the theorem.

LEMMA 3.2. (b) *implies that the set-valued map defined in* (6) *is pseudo-Lipschitz continuous around* $(0, (\bar{x}, \bar{y}))$.

PROOF. Fix any $x \in U_1 \cap \Omega_1$. Denote the projection of Σ by

$$\Sigma_x(v) := \{y \in R^m : v \in -F(x, y) + N(y, \Omega_2)\}.$$

For any $v, v' \in V$, a neighborhood of 0, let $y \in \Sigma_x(v) \cap U_2$ and $z \in \Sigma_x(v')$. Then $y \in U_2 \cap \Omega_2$, $z \in \Omega_2$ and

$$\langle v + F(x, y), z' - y\rangle \leq 0 \qquad \forall z' \in \Omega_2,$$

$$\langle v' + F(x, z), z' - z\rangle \leq 0 \qquad \forall z' \in \Omega_2.$$

In particular one has

$$\langle v + F(x, y), z - y\rangle \leq 0,$$

$$\langle v' + F(x, z), y - z\rangle \leq 0,$$

which implies that

$$\langle v + F(x,y) - v' - F(x,z), z - y\rangle \le 0 \qquad \forall v, v' \in V.$$

By strong monotonicity of $-F$, there exists $\delta > 0$ such that

$$\langle -F(x,y) + F(x,z), y - z\rangle \ge \delta\|y - z\|^2, \qquad \forall y \in U_2 \cap \Omega_2, z \in \Omega_2, x \in U_1 \cap \Omega_1.$$

Hence for all $y \in \Sigma_x(v) \cap U_2, z \in \Sigma_x(v'), x \in U_1 \cap \Omega_1$,

$$\begin{aligned}\|y - z\| &\le \frac{1}{\delta}\frac{\langle -F(x,y) + F(x,z), y - z\rangle}{\|y - z\|}\\ &\le \frac{1}{\delta}\frac{\langle v - v', y - z\rangle}{\|y - z\|}\\ &\le \frac{1}{\delta}\|v - v'\| \qquad \forall v, v' \in V.\end{aligned}$$

That is, $\Sigma_x(v)$ is pseudo-Lipschitz continuous around $(0, \bar{y})$ with the same modulus $\mu = 1/\delta$ for all $x \in U_1 \cap \Omega_1$. Notice that the modulus is independent of x, we have

$$\begin{aligned}\Sigma(v) \cap (U_1 \times U_2) &= \{(x,y) \in (U_1 \cap \Omega_1) \times U_2 : v \in -F(x,y) + N_{\Omega_2}(y)\}\\ &= \{(x,y) \in (U_1 \cap \Omega_1) \times U_2 : y \in \Sigma_x(v)\}\\ &\subset \{(x,y) \in (U_1 \cap \Omega_1) \times R^m : y \in \Sigma_x(v') + \mu\|v - v'\|B\}\\ &= \{(x,y') \in (U_1 \cap \Omega_1) \times R^m : y' \in \Sigma_x(v')\} + \{0\} \times \mu\|v - v'\|B\}\\ &\subset \Sigma(v') + \mu\|v - v'\|B \qquad \forall v, v' \in V.\end{aligned}$$

The proof of the lemma is complete. □

LEMMA 3.3. (c) *implies that the set-valued map defined in* (6) *is pseudo-Lipschitz continuous around* $(0, (\bar{x}, \bar{y}))$.

PROOF. In order to apply Proposition 2.8, we denote $z := (x, y)$, $Q(z) := N(y, \Omega_2)$, $h(z) := F(x, y)$ and $\Omega = \Omega_1 \times R^m$.

Since $Q(z) = N(y, \Omega_2)$ is independent of x, it is straightforward to show that

$$(7) \qquad D^*Q(\bar{z}, h(\bar{z}))(\eta) = \{0_n\} \times D^*N_{\Omega_2}(\bar{y}, h(\bar{z}))(\eta),$$

where 0_n denotes the zero vector in R^n. Equation (7) implies that

$$\begin{aligned}&D^*Q(\bar{z}, h(\bar{z}))(0) \cap \{-N(\bar{z}, \Omega)\}\\ &\quad = \{\{0_n\} \times D^*N((\bar{y}, h(\bar{z})), \Omega_2)(0)\} \cap \{-N(\bar{x}, \Omega_1) \times \{0_m\}\}\\ &\quad = \{0_{n+m}\}.\end{aligned}$$

That is, condition (a) in Proposition 2.8 is satisfied.

As to condition (b) in Proposition 2.8, by virtue of (7), we have

$$D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega) = \{0\} \times D^*N_{\Omega_2}(\bar{y}, h(\bar{z}))(\eta) + N(\bar{x}, \Omega_1) \times \{0\}.$$

So

$$0 \in -\nabla h(\bar{z})^\top \eta + D^*Q(\bar{z}, h(\bar{z}))(\eta) + N(\bar{z}, \Omega)$$

implies that

$$0 \in -\nabla_x F(\bar{x}, \bar{y})^\top \eta + N(\bar{x}, \Omega_1),$$

$$0 \in -\nabla_y F(\bar{x}, \bar{y})^\top \eta + D^*N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta).$$

From assumption (c), we conclude from the above inclusions that $\eta = 0$. Applying Proposition 2.8, we conclude that the set-valued map Σ is pseudo-Lipschitz continuous around $(0, (\bar{x}, \bar{y}))$. □

Now we are ready to prove Theorem 3.2.

PROOF OF THEOREM 3.2. It is obvious that assumption (d) implies assumption (c). Hence by virtue of Lemmas 3.2 and 3.3 and Proposition 2.6, any one of assumptions (b), (c) and (d) imply assumption (a). Applying Theorem 3.1 to problem (GP), there exist $r > 0$, $\eta \in rB_q$ such that

$$0 \in \partial f(\bar{x}, \bar{y}) + D^*\Phi((\bar{x}, \bar{y}), 0)(\eta) + N((\bar{x}, \bar{y}), \Omega_1 \times R^m),$$

where $\Phi(x, y) := -F(x, y) + N(y, \Omega_2)$. By Proposition 2.5, the sum rule for coderivatives, we have

$$\begin{aligned} 0 &\in \partial f(\bar{x}, \bar{y}) - \nabla F(\bar{x}, \bar{y}) + \{0\} \times D^*N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta) + N((\bar{x}, \bar{y}), \Omega_1 \times R^m) \\ &\subset \partial_x f(\bar{x}, \bar{y}) \times \partial_y f(\bar{x}, \bar{y}) - \nabla_x F(\bar{x}, \bar{y}) \times \nabla_y F(\bar{x}, \bar{y}) \\ &\quad + \{0\} \times D^*N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))(\eta) + N(\bar{x}, \Omega_1) \times \{0\}, \end{aligned}$$

where the last inclusion follows from the regularity of f (c.f. Proposition 2.3). The proof of the theorem is complete. □

ANOTHER PROOF OF THEOREM 3.2 UNDER ASSUMPTION (c). Problem (OPVIC) can be also rewritten in the following form:

$$\begin{aligned} &\min \quad f(x, y) \\ &\text{s.t.} \quad (x, y) \in \operatorname{Gr} S \cap (\Omega_1 \times R^m). \end{aligned}$$

Therefore by the well-known optimality condition, we have

$$0 \in \partial f(\bar{x}, \bar{y}) + N((\bar{x}, \bar{y}), \operatorname{Gr} S \cap (\Omega_1 \times R^m)). \tag{8}$$

By Theorem 5.1 of Mordukhovich (1994b), we have

$$\begin{aligned} &D^*S(\bar{x}, \bar{y})(v) \\ &\quad \subset \{\nabla_x F(\bar{x}, \bar{y})^\top p : -v \in \nabla_y F(\bar{x}, \bar{y})^\top p + D^*N(\bar{y}, F(\bar{x}, \bar{y}), \Omega_2)(p)\}. \end{aligned} \tag{9}$$

Now let $\xi = (\xi_1, \xi_2) \in N((\bar{x}, \bar{y}), \mathrm{Gr}S) \cap (-N(\bar{x}, \Omega_1) \times \{0\})$. Then $\xi_2 = 0$ and $\xi_1 \in -N(\bar{x}, \Omega_1) \cap D^*S(\bar{x}, \bar{y})(0)$. That is, there exists η such that

$$0 \in -\nabla_y F(\bar{x}, \bar{y})^\top \eta + D^*N((\bar{y}, F(\bar{x}, \bar{y})), \Omega_2)(\eta)$$

and $\nabla_x F(\bar{x}, \bar{y})^\top \eta = \xi_1 \in N(\bar{x}, \Omega_1)$. But by assumption (c), such η can only be zero vector. So $\xi = 0$.

Therefore by virtue of Proposition 2.1, (8) implies that

$$0 \in \partial f(\bar{x}, \bar{y}) + N((\bar{x}, \bar{y}), \mathrm{Gr}S) + N(\bar{x}, \Omega_1) \times \{0\}.$$

Hence there exists $(\xi_1, \xi_2) \in N((\bar{x}, \bar{y}), \mathrm{Gr}S)$ such that

$$0 \in \partial f(\bar{x}, \bar{y}) + (\xi_1, \xi_2) + N(\bar{x}, \Omega_1) \times \{0\}.$$

By regularity of f, one has

$$0 \in \partial_x f(\bar{x}, \bar{y}) + D^*S(\bar{x}, \bar{y})(-\xi_2) + N(\bar{x}, \Omega_1) \tag{10}$$

$$0 \in \partial_y f(\bar{x}, \bar{y}) + \xi_2. \tag{11}$$

By using (9), we have the conclusion. □

REMARK 3.1. If we assume further that f is continuously differentiable, inclusions (10) and (11) become

$$0 \in \nabla_x f(\bar{x}, \bar{y}) + D^*S(\bar{x}, \bar{y})(\nabla_y f(\bar{x}, \bar{y})) + N(\bar{x}, \Omega_1).$$

If the strong regularity condition introduced by Robinson (1980) holds at the solution point, the solution map $S(x)$ will be locally single-valued and Lipschitz continuous around $(\bar{x}, \bar{y})$ according to Robinson (1980). In this case, by virtue of Proposition 2.4,

$$D^*S(\bar{x}, \bar{y})(\nabla_y f(\bar{x}, \bar{y})) = \partial\langle \nabla_y f(\bar{x}, \bar{y}), S\rangle(\bar{x}) \subset \partial_C S(\bar{x})^\top \nabla_y f(\bar{x}, \bar{y}).$$

As in the proof of Theorem 2.3 of Outrata (1994), there exists a suitable chosen index set $K(\bar{x})$ such that

$$\partial_C S(\bar{x})^\top \nabla_y f(\bar{x}, \bar{y}) \subset \nabla_x F(\bar{x}, \bar{y})^\top \mathrm{co}\{p_i : i \in K(\bar{x})\},$$

where the vectors p_i, $i \in K(\bar{x})$ solve the linearization of the generalized equation. Hence in the second proof of Theorem 3.2, using the above upper bound instead of the upper bound (9) Theorem 2.3 of Outrata (1994) can be derived. This establishes the relationship between our result and the one obtained by Outrata (1994).

Note that Theorem 3.2 involves the coderivative $D^*N_{\Omega_2}(\bar{y}, F(\bar{x}, \bar{y}))$ whose computation depends on the limiting normal cone $N_{\mathrm{gph}N_{\Omega_2}}(\bar{y}, F(\bar{x}, \bar{y}))$. In many application, Ω_2 can be chosen as $\Omega_2 = R^q_+$ for some positive integer q. In what follows we compute the set $N_{\mathrm{gph}N_{R^q_+}}(z)$ for the case $z \neq 0$ and $q \leq 2$ and provide an estimate for the set $N_{\mathrm{gph}N_{R^q_+}}(z)$ by a system of algebraic equation. Hence the optimality conditions given in Theorem 3.2 may be expressed as a systems of inequalities and equalities.

PROPOSITION 3.1. *Let $z \in gphN_{R^q_+}$. Then $\eta \in N_{gphN_{R^q_+}}(z)$ implies that*

$$\eta_i z_i = 0 \qquad \forall i = 1, 2, \ldots, 2q.$$

PROOF. Let $q = 1$. Then $\text{gph}N_{R_+} = ((0, \infty) \times \{0\}) \cup (\{0\} \times (-\infty, 0])$. For any nonzero vector $z = (z_1, z_2) \in \text{gph}N_{R_+}$, it is easy to see that

$$N(z, \text{gph}N_{R_+}) = \begin{cases} \{0\} \times R & \text{if } z_1 > 0, z_2 = 0, \\ R \times \{0\} & \text{if } z_1 = 0, z_2 < 0. \end{cases}$$

Let $q = 2$. Then $\text{gph}N_{R^2_+} = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$, where

$$\Omega_1 := (0, \infty) \times (0, \infty) \times \{0\} \times \{0\},$$

$$\Omega_2 := \{0\} \times \{0\} \times (-\infty, 0] \times (-\infty, 0],$$

$$\Omega_3 := \{0\} \times (0, \infty) \times (-\infty, 0] \times \{0\},$$

$$\Omega_4 := (0, \infty) \times \{0\} \times \{0\} \times (-\infty, 0].$$

Case 1. Two of the components of $z \in \text{gph}N_{R^2_+}$ are nonzero. If $z \in \Omega_1$, then $z \notin \Omega_2 \cup \text{cl}\Omega_3 \cup \text{cl}\Omega_4$ and there exists a neighborhood of z that does not contain in $\Omega_2 \cup \text{cl}\Omega_3 \cup \text{cl}\Omega_4$. Therefore

$$N_{\text{gph}N_{R^2_+}} = N_{\text{cl}\Omega_1}(z) = \{0\} \times \{0\} \times R \times R.$$

Similarly we have

$$N_{\text{gph}N_{R^2_+}}(z) = \begin{cases} N_{\text{cl}\Omega_1}(z) = \{0\} \times \{0\} \times R \times R & \text{if } z_1 > 0, z_2 > 0, z_3 = 0, z_4 = 0, \\ N_{\Omega_2}(z) = R \times R \times \{0\} \times \{0\} & \text{if } z_1 = 0, z_2 = 0, z_3 < 0, z_4 < 0, \\ N_{\text{cl}\Omega_3}(z) = R \times \{0\} \times \{0\} \times R & \text{if } z_1 = 0, z_2 > 0, z_3 < 0, z_4 = 0, \\ N_{\text{cl}\Omega_4}(z) = \{0\} \times R \times R \times \{0\} & \text{if } z_1 > 0, z_2 = 0, z_3 = 0, z_4 < 0. \end{cases}$$

Case 2. One of the components of $z \in \text{gph}N_{R^2_+}$ are nonzero. If $z = (0, 0, 0, z_4) \in \text{gph}N_{R^2_+}$ where $z_4 < 0$, then $z \in \Omega_2 \cap \text{cl}\Omega_4$ and $z \notin \text{cl}\Omega_1 \cup \text{cl}\Omega_3$ and there exists a neighborhood of z that does not contain in the set $\text{cl}\Omega_1 \cup \text{cl}\Omega_3$. Hence $N_{\text{gph}N_{R^2_+}}(z) = N_{\Omega_2 \cup \text{cl}\Omega_4}(z)$. We first calculate the proximal normal cone $N^{\pi}_{\Omega_2 \cup \text{cl}\Omega_4}(z)$. Let $t > 0$ then $d_{\Omega_2 \cup \text{cl}\Omega_4}(t\xi + z) = \|t\xi\|$ if and only if

$$\sqrt{(t\xi_1)^2 + (t\xi_2)^2 + (t\xi_3)^2 + (t\xi_4)^2}$$

$$\leq \begin{cases} \sqrt{(t\xi_1)^2 + (t\xi_2)^2 + (t\xi_3 - z'_3)^2 + (t\xi_4 + z_4 - z'_4)^2} & \forall z' \in \Omega_2, \\ \sqrt{(t\xi_1 - z'_1)^2 + (t\xi_2)^2 + (t\xi_3)^2 + (t\xi_4 + z_4 - z'_4)^2} & \forall z' \in \text{cl}\Omega_4. \end{cases}$$

Such a $t > 0$ exists if and only if $\xi \in (-\infty, 0] \times R \times [0, \infty) \times \{0\}$. Therefore if $z_1 = 0$, $z_2 = 0$, $z_3 = 0$, $z_4 < 0$ then

$$N^{\pi}_{\Omega_2 \cup \text{cl}\Omega_4}(z) = (-\infty, 0] \times R \times [0, \infty) \times \{0\}.$$

Let $z_k = (z^1_k, z^2_k, z^3_k, z^4_k) \in \Omega_2 \cup \text{cl}\Omega_4$ and $z_k \to 0$. If $z^3_k < 0$, then

$$N^{\pi}_{\Omega_2\cup \mathrm{cl}\Omega_4}(z_k) = N^{\pi}_{\Omega_2}(z_k) = R \times R \times \{0\} \times \{0\}.$$

If $z^1_k > 0$, then

$$N^{\pi}_{\Omega_2\cup \mathrm{cl}\Omega_4}(z_k) = N^{\pi}_{\mathrm{cl}\Omega_4}(z_k) = \{0\} \times R \times R \times \{0\}.$$

Therefore by definition,

$$\begin{aligned} N_{\mathrm{gph} N_{R^2_+}}(z) &= N_{\Omega_2\cup \mathrm{cl}\Omega_4}(z) \\ &= ((-\infty, 0] \times R \times [0, \infty) \times \{0\}) \cup (R \times R \times \{0\} \times \{0\}) \\ &\quad \cup (\{0\} \times R \times R \times \{0\}). \end{aligned}$$

Similarly, we have

$$\begin{aligned} N^{\pi}_{\mathrm{cl}\Omega_1\cup \mathrm{cl}\Omega_4}(z) &= \{0\} \times (-\infty, 0] \times R \times (0, \infty) \quad \text{if } z_1 > 0, z_2 = 0, z_3 = 0, z_4 = 0, \\ N^{\pi}_{\mathrm{cl}\Omega_2\cup \mathrm{cl}\Omega_3}(z) &= R \times (-\infty, 0] \times \{0\} \times [0, \infty) \quad \text{if } z_1 = 0, z_2 = 0, z_3 < 0, z_4 = 0, \\ N^{\pi}_{\mathrm{cl}\Omega_1\cup \mathrm{cl}\Omega_3}(z) &= (-\infty, 0] \times \{0\} \times [0, \infty) \times R \quad \text{if } z_1 = 0, z_2 > 0, z_3 = 0, z_4 = 0. \end{aligned}$$

Therefore, If $z_1 > 0$, $z_2 = 0$, $z_3 = 0$, $z_4 = 0$ then

$$\begin{aligned} N_{\mathrm{gph} N_{R^2_+}}(z) &= N_{\mathrm{cl}\Omega_1\cup \mathrm{cl}\Omega_4}(z) \\ &= (\{0\} \times (-\infty, 0] \times R \times [0, \infty)) \cup (\{0\} \times \{0\} \times R \times R) \\ &\quad \cup (\{0\} \times R \times R \times \{0\}). \end{aligned}$$

If $z_1 = 0$, $z_2 = 0$, $z_3 < 0$, $z_4 = 0$ then

$$\begin{aligned} N_{\mathrm{gph} N_{R^2_+}}(z) &= N_{\Omega_2\cup \mathrm{cl}\Omega_3}(z) \\ &= (R \times (-\infty, 0] \times \{0\} \times [0, \infty)) \\ &\quad \cup (R \times R \times \{0\} \times \{0\}) \cup (R \times \{0\} \times \{0\} \times R). \end{aligned}$$

If $z_1 = 0$, $z_2 > 0$, $z_3 = 0$, $z_4 = 0$ then

$$\begin{aligned} N_{\mathrm{gph} N_{R^2_+}}(z) &= N_{\mathrm{cl}\Omega_1\cup \mathrm{cl}\Omega_3}(z) \\ &= ((-\infty, 0] \times \{0\} \times [0, \infty) \times R) \cup (\{0\} \times \{0\} \times R \times R) \\ &\quad \cup (R \times \{0\} \times \{0\} \times R). \end{aligned}$$

For general q, similarly we can show that for any vector $z \in \mathrm{gph} N_{R^q_+}$, $z_i \neq 0$ implies that $\eta_i = 0$, $\forall i = 1, 2, \ldots, 2q$. The proof of the proposition is complete. □

REMARK 3.2. Application of the above estimation will be given in Example 4.1.

4. Applications to bilevel programming problems. Now consider the bilevel programming problem (BP) where the constaint region for the lower level problem (P_x) is a system of inequalities, i.e., $\Omega_2(x) := \{y \in R^m : \psi(x, y) \leq 0\}$ where $\psi : R^{n+m} \to R^q$. Let $y \in S(x)$. If a certain constraint qualification holds at (x, y), then there exists $u \in R^q$ such that

$$\nabla_y g(x, y) + u\nabla_y \psi(x, y) = 0, \qquad \psi(x, y) \leq 0,$$

$$u \geq 0, \qquad \langle u, \psi(x, y)\rangle = 0,$$

where $u\nabla_y\psi := \Sigma\, u_k \nabla_y \psi_k(x, y)$. According to Robinson (1979), the above Kuhn-Tucker conditions for (P_x) can be written as the generalized equation

$$0 \in -F(x, z) + N(z, \Omega)$$

where $\Omega = R^m \times R^q_+$, $z = (y, u) \in R^{m+q}$, $F : R^{n+m+q} \to R^{m+q}$ given by

$$F(x, z) = \begin{bmatrix} -[\nabla_y g + u\nabla_y\psi]^\top(x, y) \\ \psi(x, y) \end{bmatrix}. \tag{12}$$

Applying Theorem 3.2 we now derive a Kuhn-Tucker type necessary optimality conditions for (BP).

THEOREM 4.1. *Assume that $f : R^{n+m} \to R$ is Lipschitz continuous and regular, $g : R^{n+m} \to R$ and $\psi : R^{n+m} \to R^q$ are twice continuously differentiable. Further assume that g is pseudoconvex in y, ψ is quasiconvex in y. Let $(\bar{x}, \bar{y})$ solve the problem (BP). Suppose that a certain constraint qualification holds for (P_x) and $\bar{u}$ is a corresponding multiplier associated with $(\bar{x}, \bar{y})$, i.e.,*

$$0 = [\nabla_y g + \bar{u}\nabla_y\psi]^\top(\bar{x}, \bar{y}) \qquad \bar{u} \geq 0, \qquad \langle \psi(\bar{x}, \bar{y}), \bar{u}\rangle = 0.$$

If one of the following conditions hold:

(a) *The set-valued map*

$$\Sigma(v) := \{(x, y, u) \in \Omega_1 \times R^m \times R^q : v \in -F(x, z) + N(z, R^m \times R^q_+)\} \tag{13}$$

is pseudo upper-Lipschitz continuous at $(0, \bar{x}, \bar{y}, \bar{u})$.

(b) *Ω_1 is a closed subset of R^n and there is no nonzero vector $\eta = (\eta_1, \eta_2)$ such that*

$$0 \in (\nabla^2_{xy} g + \bar{u}\nabla^2_{xy}\psi)(\bar{x}, \bar{y})^\top\eta_1 - \nabla_x\psi(\bar{x}, \bar{y})^\top\eta_2 + N(\bar{x}, \Omega_1), \tag{14}$$

$$0 = (\nabla^2_{yy} g + \bar{u}\nabla^2_{yy}\psi)(\bar{x}, \bar{y})^\top\eta_1 - \nabla_y\psi(\bar{x}, \bar{y})^\top\eta_2, \tag{15}$$

$$(-\nabla_y\psi(\bar{x}, \bar{y})\eta_1, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), gphN_{R^q_+}). \tag{16}$$

then there exist $r > 0$, $\eta = (\eta_1, \eta_2) \in rB_{m+q}$ such that

$$0 \in \partial_x f(\bar{x}, \bar{y}) + (\nabla^2_{xy} g + \bar{u}\nabla^2_{xy}\psi)(\bar{x}, \bar{y})^\top\eta_1 - \nabla_x\psi(\bar{x}, \bar{y})^\top\eta_2 + N(\bar{x}, \Omega_1), \tag{17}$$

$$0 = \partial_y f(\bar{x}, \bar{y}) + (\nabla^2_{yy} g + \bar{u}\nabla^2_{yy}\psi)(\bar{x}, \bar{y})^\top\eta_1 - \nabla_y\psi(\bar{x}, \bar{y})^\top\eta_2, \tag{18}$$

$$(19) \qquad (-\nabla_y \psi(\bar{x}, \bar{y})\eta_1, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), gphN_{R^q_+}),$$

where $u\nabla^2_{xy}\psi := \Sigma_k u_k \nabla^2_{xy}\psi_k$, *and*

$$\nabla^2_{xy} = \begin{bmatrix} \dfrac{\partial^2}{\partial x_1 \partial y_1} & \cdots & \dfrac{\partial^2}{\partial x_n \partial y_1} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial^2}{\partial x_1 \partial y_m} & \cdots & \dfrac{\partial^2}{\partial x_n \partial y_m} \end{bmatrix}.$$

PROOF. Since the objective function of the lower level problem g is pseudoconvex in y and the constraint ϕ is quasiconvex in y, by Theorem 4.2.11 of Bazaraa and Shetty (1979) the Kuhn-Tucker condition is a necessary and sufficient condition for optimality. Therefore from the discussion proceeding Theorem 4.1 we know that $(\bar{x}, \bar{y})$ is a solution of the following problem:

$$\begin{aligned} \min \quad & \tilde{f}(x, z) \\ \text{s.t.} \quad & 0 \in -F(x, z) + N(z, \Omega_2), \\ & (x, z) \in \Omega_1 \times R^m \times R^q \end{aligned}$$

where $z := (y, u) \in R^{m+q}$, $\Omega_2 = R^m \times R^q{}_+$, $\tilde{f}(x, z) = f(x, y)$ and $F(x, z)$ is defined as in (12).

It is straightforward to show that

$$\nabla_x F(\bar{x}, \bar{z})^\top = \begin{bmatrix} -[\nabla^2_{xy} g + \bar{u}\nabla^2_{xy}\psi](\bar{x}, \bar{y}) \\ \nabla_x \psi(\bar{x}, \bar{y}) \end{bmatrix},$$

and

$$(20) \qquad \nabla_z F(\bar{x}, \bar{z})^\top = \begin{bmatrix} -(\nabla^2_{yy} g + \bar{u}\nabla^2_{yy}\psi)(\bar{x}, \bar{y}) & \nabla_y \psi(\bar{x}, \bar{y})^\top \\ -\nabla_y \psi(\bar{x}, \bar{y}) & 0 \end{bmatrix}.$$

We now verify that assumption (c) in Theorem 3.2 is satisfied. Equation (14) is apparently equivalent to $0 \in -\nabla_x F(\bar{x}, \bar{z})^\top \eta + N(\bar{x}, \Omega_1)$.

Now let $(\xi_1, \xi_2) \in D^* N_{R^m \times R^q_+}(\bar{z}, F(\bar{x}, \bar{z}))(\eta)$ then by definition of coderivatives,

$$(\xi_1, \xi_2, -\eta) = (\xi_1, \xi_2, -\eta_1, -\eta_2) \in N((\bar{z}, F(\bar{x}, \bar{z})), \text{gph} N_{R^m \times R^q_+}).$$

Observing

$$\begin{aligned} \text{gph} N_{R^m \times R^q_+} &= \{(y, u, w_1, w_2) : y \in R^m, w_1 = 0, w_2 \in N(u, R^q_+)\} \\ &= R^m \times \{(u, 0, w_2) : w_2 \in N(u, R^q_+)\}, \end{aligned}$$

we obtain from the last inclusion that

$$\xi_1 = 0, \qquad (\xi_2, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), \mathrm{gph} N_{R^q_+}). \tag{21}$$

Since

$$\nabla_z F(\bar{x}, \bar{z})^\top \eta = (-(\nabla^2_{xy} g + \bar{u} \nabla^2_{xy} \psi)(\bar{x}, \bar{y})^\top \eta_1 + \nabla_x \psi(\bar{x}, \bar{y})^\top \eta_2, -\nabla_y \psi(\bar{x}, \bar{y}) \eta_1),$$

(15) and (16) are equivalent to

$$0 \in -\nabla_z F(\bar{x}, \bar{z})^\top \eta + D^* N_{R^m \times R^q_+}(\bar{z}, F(\bar{x}, \bar{z}))(\eta) \tag{22}$$

in light of (21). Therefore assumption (b) in Theorem 4.1 implies assumption (c) in Theorem 3.2. Applying Theorem 3.2, we conclude that there exist $r > 0$, $\eta \in rB_{m+q}$ such that

$$0 \in \partial_x \tilde{f}(\bar{x}, \bar{z}) - \nabla_x F(\bar{x}, \bar{z})^\top \eta + N(\bar{x}, \Omega_1), \tag{23}$$

$$0 \in \partial_z \tilde{f}(\bar{x}, \bar{z}) - \nabla_z F(\bar{x}, \bar{z})^\top \eta + D^* N_{R^m \times R^q_+}(\bar{z}, F(\bar{x}, \bar{z}))(\eta). \tag{24}$$

It is easy to see that (23) implies (17).

By virtue of (20), (24) and (21), we obtain

$$-(\nabla^2_{yy} g + \nabla^2_{yy} \psi)(\bar{x}, \bar{y}) \eta_1 + \nabla_y \psi(\bar{x}, \bar{y})^\top \eta_2 \in \partial_y f(\bar{x}, \bar{y}),$$

$$(-\nabla_y \psi(\bar{x}, \bar{y}) \eta_1, -\eta_2) \in N((\bar{u}, \psi(\bar{x}, \bar{y})), \mathrm{gph} N_{R^q_+}).$$

The proof is complete. □

We now consider the bilevel programming problem where the lower level problem is the following parametric quadratic programming problem:

$$\mathrm{QP}_x \qquad \begin{array}{ll} \min & \langle y, Px \rangle + \frac{1}{2} \langle y, Qy \rangle + p^t x + q^t y \\ \text{s.t.} & Ax + By - b \le 0, \end{array}$$

where $Q \in R^{m \times m}$ is a symmetric and positive semidefinite matrix, $p \in R^n$, $q \in R^m$, $P \in R^{m \times n}$, A and B are $q \times n$ and $q \times m$ matrices respectively and $b \in R^q$.

Recall that a set-valued map is called a polyhedral multifunction if its graph is unions of finitely many polyhedral convex sets. This class of set-valued maps are closed under (finite) addition, scalar multiplication, and (finite) composition. By Proposition 1 of Robinson (1981), a polyhedral multifunction is upper-Lipschitz. Hence the following result is straighforward.

PROPOSITION 4.1. *The set-valued map*

$$\Sigma(v) := \{(x, y, u) : v \in -F(x, y, u) + N(y, u, R^m \times R^q_+)\}$$

where

$$F(x, y, u) := \begin{bmatrix} -(Qy + Px + q + B^\top u) \\ Ax + By - b \end{bmatrix}$$

is upper-Lipschitz continuous around 0.

PROOF. Since the graph of $N_{R^m \times R^q_+}$ is a finite union of polyhedral convex sets. So $N_{R^m \times R^q_+}$ is polyhedral, so that $-F + N_{R^m \times R^q_+}$ (as the sum of $-F(\cdot)$ and $N_{R^m \times R^q_+}(\cdot)$) is polyhedral, and so therefore is its inverse map Σ. □

Proposition 4.1 leads to the following Kuhn-Tucker necessary optimality contions for (BP) where the lower level problem is a parametric quadratic programming problem. It is an easy corollary of Theorem 4.1.

COROLLARY 4.1. *Let $(\bar{x}, \bar{y})$ be an optimal solution of BP where the lower level problem is the parametric quadratic programming problem. Suppose that f is Lipschitz continuous and regular. Then there exist $r > 0$, $\eta = (\eta_1, \eta_2) \in rB_{m+q}$, $\bar{u} \in R^q_+$ such that*

$$0 \in \partial_x f(\bar{x}, \bar{y}) + P^{\mathrm{T}}\eta_1 - A^{\mathrm{T}}\eta_2 + N(\bar{x}, \Omega_1),$$

$$0 \in \partial_y f(\bar{x}, \bar{y}) + Q^{\mathrm{T}}\eta_1 - B^{\mathrm{T}}\eta_2,$$

$$(-B\eta_1, -\eta_2) \in N((\bar{u}, A\bar{x} + B\bar{y} - b), \mathrm{gph} N_{R^q_+}),$$

$$0 = Q\bar{y} + P\bar{x} + q + B^{\mathrm{T}}\bar{u},$$

$$\langle A\bar{x} + B\bar{y} - b, \bar{u} \rangle = 0.$$

We now give an example to illustrate the application of the theory.

EXAMPLE 4.1. Consider the following classical bilevel problem:

$$\begin{aligned} \min \quad & x^2 - 2y \\ \text{s.t.} \quad & x \in R, \\ & y \in \operatorname{argmin}\{y^2 - 2xy : y - 1 \le 0, -y \le 0\}. \end{aligned}$$

This is a BP where the lower level problem is the parametric quadratic mathematical problem with $Q = 2$, $P = -2$, $p = q = 0$, $A = (0, 0)'$, $B = (1, -1)'$ and $b = (1, 0)'$.

It is easy to see that the solution for the above simple problem is $(x = 1, y = 1)$. To illustrate the application of the theory we now show that the solution can be actually solved from the Kuhn-Tucker necessary optimality conditions developed in this paper.

Suppose that $(\bar{x}, \bar{y})$ solves the problem. Then by Corollary 4.1 and Proposition 3.1, there exist $r > 0$, $\eta = (\eta_1, \eta_2^1, \eta_2^2) \in rB_3$, $\bar{u} = (\bar{u}_1, \bar{u}_2) \in R^2$ such that

$$0 = 2\bar{x} - 2\eta_1, \tag{25}$$

$$0 = -2 + 2\eta_1 - \eta_2^1 + \eta_2^2, \tag{26}$$

$$\eta_1\bar{u}_1 = 0, \quad \eta_1\bar{u}_2 = 0, \quad \eta_2^1(\bar{y} - 1) = 0, \quad \eta_2^2(-\bar{y}) = 0, \quad 0 \le \bar{y} \le 1, \tag{27}$$

$$2\bar{y} - 2\bar{x} + \bar{u}_1 - \bar{u}_2 = 0, \tag{28}$$

$$\bar{u}_1 \ge 0, \quad \bar{u}_2 \ge 0, \quad \bar{u}_1(\bar{y} - 1) = 0, \quad \bar{u}_2(-\bar{y}) = 0. \tag{29}$$

We now discuss possible cases:

Case (1) Suppose $\bar{y} \in (0, 1)$, then $\bar{u}_1 = \bar{u}_2 = 0$ by (29). Thus $\bar{y} = \bar{x}$ by (28). On the other hand, by (27) we have $\eta_2^1 = \eta_2^2 = 0$ and hence $\bar{x} = 1$ by (25) and (26). It is a contradiction.

Case (2) Suppose $\bar{y} = 0$. Then $\bar{u}_1 = 0$ by (29). By (27) we know $\eta_2^1 = 0$. Therefore, if $\bar{u}_2 > 0$, using (27) one has $\eta_1 = 0$ and hence $\bar{x} = 0$ which contradicts (28). If $\bar{u}_2 = 0$, by (28) we have $\bar{x} = \bar{y} = 0$.

Case (3) $\bar{y} = 1$. Then by (29) and (27) we have $\bar{u}_2 = 0$ and $\eta_2^2 = 0$. Suppose $\bar{u}_1 > 0$, then $\eta_1 = 0$, that is, $\bar{x} = 0$. By (28), it is easy to see that it is impossible. Therefore, $\bar{u}_1 = 0$ and by (28) one has $\bar{x} = \bar{y} = 1$. By (25) and (26) we have $\eta_1 = 1$, $\eta_2^1 = 0$.

Comparing the value for the objective function for the cases (2) and (3) we conclude that $\bar{x} = \bar{y} = 1$ is the optimal solution.

Acknowledgments. The work of the first author was supported by NSERC and a University of Victoria internal research grant.

References

Anandalingam, G., T. L. Friesz (1992). eds., Hierarchical Optimization, *Ann. of Oper. Res.* **34.**

Aubin, J.-P. (1994). Lipschitz behavior of solutions to convex minimization problems. *Math. Oper. Res.* **9** 87–111.

———, I. Ekeland (1984). *Applied Nonlinear Analysis*, Wiley-Interscience, New York.

———, H. Frankowska (1990). *Set-Valued Analysis*, Birkhauser, Boston.

Barbu, V. (1984). *Optimal Control of Variational Inequalities*, Pitman, Boston.

Bazaraa, M. S., C. M. Shetty (1979). *Nonlinear Programming Theory and Algorithms*, John Wiley and Sons.

Clarke, F. H. (1983). *Optimization and Nonsmooth Analysis*, Wiley-Interscience, New York.

Dempe, S. (1992). A necessary and sufficient optimality condition for bilevel programming problems. *Optimization* **25** 341–354.

Dussault, J.-P., P. Marcotte (1988). Conditions de régularité géométrique pour les inéquations variationnelles, *RAIRO Rech. Opér.* **23** 1–16.

Friesz, T. L., R. T. Tobin, H.-J. Cho, N. J. Mehta (1990). Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints. *Math. Programming* **48** 265–284.

Harker, P. T., J. S. Pang (1988). On the existence of optimal solutions to mathematical program with equilibrium constraints. *Oper. Res. Lett.* **7** 61–64.

———, ——— (1990). Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Math. Programming* **48** 161–220.

Ioffe, A. D. (1979). Regular points of Lipschitz functions. *Trans. Amer. Math. Soc.* **251** 61–69.

Loridan, P., J. Morgan (1989). A theoretical approximation scheme for Stackelberg problems. *J. Optim. Theory Appl.* **11** 95–110.

Luo, Z. Q., J. S. Pang, D. Ralph, S-Q. Wu (1996). Exact Penalization and stationarity conditions of mathematical programs with equilibrium constraints. *Math. Programming* **75** 19–76.

Marcotte, P., D. L. Zhu (1996). Exact and inexact penalty methods for the generalized bilevel programming problem. *Math. Programming* **74** 141–157.

Mordukhovich, B. S. (1994a). Sensitivity analysis for constraint and variational systems by means of set-valued differentiation. *Optimization* **31** 13–46.

——— (1994b). Generalized differential calculus for nonsmooth and set-valued mappings. *J. Math. Anal. Appl.* **183** 250–288.

——— (1994c). Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis. *Trans. Amer. Math. Soc.* **343** 609–655.

Outrata, J. V. (1993). Necessary optimality conditions for Stackelberg problems. *J. Optim. Theory Appl.* **76** 305–320.

——— (1994). On optimization problems with variational inequality constraints. *SIAM J. Optim.* **2** 340–357.

Robinson, S. M. (1975). Stability theory for systems of inequalities, part I: linear systems. *SIAM J. Numer. Anal.* **12** 754–769.

——— (1976). Stability theory for systems of inequalities, part II: defferentiable nonlinear systems. *SIAM J. Numer. Anal.* **13** 497–513.

——— (1979). Generalized equations and their solutions. *Math. Programming Stud.* **10** 128–141.

——— (1980). Strongly regular generalized equations. *Math. Oper. Res.* **5** 43–62.

——— (1981). Some continuity properties of polyhedral multifunctions. *Math. Programming Stud.* **14** 206–214.

Shi, S. (1988). Optimal control of strongly monotone variational inequalities. *SIAM J. Control Optim.* **26** 274–290.

——— (1990). Erratum: Optimal control of strongly monotone variational inequalities. *SIAM J. Control Optim.* **28** 243–249.

Von Stackelberg, H. (1952). *The Theory of the Market Economy*, Oxford University Press, Oxford, England.

Ye, J. J. (1996). New uniform parametric error bounds, *J. Optim. Theory Appl.* (to appear).

———, D. L. Zhu (1995). Optimality conditions for bilevel programming problems. *Optimization* **33** 9–27.

———, ——— (1997). A note on optimality conditions for bilevel programming problems. *Optimization* **39** 361–366.

———, ———, Q. J. Zhu (1997). Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM J. Optim.* **7** 481–507.

Zhang, R. (1990). *Problems of hierarchical optimization: Nonsmoothness and analysis of solutions*, Ph.D. Thesis, Department of Applied Mathematics, University of Washington, Seattle.

——— (1994). Problems of hierarchical optimization in finite dimensions. *SIAM J. Optim.* **4** 521–536.

J. J. Ye: Department of Mathematics and Statistics, University of Victoria, Victoria, B.C., Canada V8W 3P4; e-mail: janeye@uvic.ca

X. Y. Ye: Department of Biometry and Statistics, School of Public Health, SUNY at Albany, New York 12222-0001