Acta Math Vietnam
DOI 10.1007/s40306-013-0032-1

ALGORITHMS FOR A CLASS OF BILEVEL PROGRAMS INVOLVING PSEUDOMONOTONE VARIATIONAL INEQUALITIES

Bui Van Dinh · Le Dung Muu

Received: 4 May 2012 / Revised: 29 May 2013 / Accepted: 31 May 2013

Abstract We propose algorithms for finding the projection of a given point onto the solution set of the pseudomonotone variational inequality problem. This problem arises in the Tikhonov regularization method for pseudomonotone variational inequality. Since the solution set of the variational inequality is not given explicitly, the available methods of mathematical programming and variational inequalities cannot be applied directly.

Keywords Bilevel variational inequality · Pseudomonotonicity · Projection method · Armijo linesearch · Convergence

Mathematics Subject Classification (2010) 49M37 · 90C26 · 65K15

1 Introduction

Variational inequality (VI) is a fundamental topic in applied mathematics. VIs are used for formulating and solving various problems arising in mathematical physics, economics, engineering and other fields. Theory, methods and applications of VIs can be found in some comprehensive books and monographs (see e.g. [8, 9, 15, 16]). Mathematical programs with variational inequality constraints can be considered as one of the further development directions of variational inequality [18]. Recently, these problems have received much attention of researchers due to their vast applications.

In this paper, we are concerned with a special case of VIs with variational inequality constraints. Namely, we consider the bilevel variational inequality problem (BVI):

$$\begin{cases} \min\{\|x - x^g\| : x \in S\}, \\ \text{where } x^g \in C \text{ and } S = \{u \in C : \langle F(u), y - u\rangle \geq 0 \ \forall y \in C\} \end{cases} \tag{1.1}$$

B.V. Dinh
Le Quy Don Technical University, Hanoi, Vietnam
e-mail: vandinhb@gmail.com

L.D. Muu (✉)
Institute of Mathematics, VAST, Hanoi, Vietnam
e-mail: ldmuu@math.ac.vn

Published online: 08 November 2013

i.e., S is the solution set of the variational inequality VI(C, F) defined as

$$\text{Find } x^* \in C \text{ such that } \langle F(x^*), y - x^* \rangle \geq 0 \ \forall y \in C. \tag{1.2}$$

Throughout the paper, we suppose that C is a nonempty closed convex subset in the Euclidean space $\mathbb{R}^n$ and $F : \mathbb{R}^n \to \mathbb{R}^n$. We call problem (1.1) the upper problem and (1.2) the lower one.

It should be noticed that the solution set S of the lower problem (1.2) is convex whenever F is pseudomonotone on C. However, the main difficulty is that, even if the constrained set S is convex, it is not given explicitly as in a standard mathematical programming problem, and therefore the available methods of convex optimization and variational inequality cannot be applied directly to problem (1.1).

In the literature, there exist several solution methods that can be used to solve bilevel variational inequality problem (1.1) (see e.g. [1, 3, 6, 7, 14, 19, 20, 22, 28, 30] and the references cited therein): penalty function methods, regularization methods and hybrid fixed point-projection methods. Most of these methods can be used only when F is monotone. In the penalty function and regularization methods [6, 7, 14, 22, 28, 30], main subproblems to be solved are variational inequalities whose cost operators are the sum of F and some strongly monotone operator depending on a parameter. These cost operators are strongly monotone when F is monotone, but they may not be pseudomonotone when F is pseudomonotone, and therefore the subproblems cannot be solved by available algorithms. In hybrid fixed point methods with projections, the main subproblems to be solved are strongly convex programs; however for convergence some additional assumptions on nonexpansiveness of the operators involved are needed (see e.g. [1, 6, 19, 28, 30]). For problem (1.1), the required nonexpansiveness is satisfied when F is monotone, but it may fail to hold when F is pseudomonotone.

The Tikhonov regularization is a fundamental method for monotone variational inequality problems. Hao in [11] studied the Tikhonov method for pseudomonotone VIs and answered in the affirmative a question posed in [8, p. 1229]. It is shown in [11] (see also [12, 26]) that every Tikhonov trajectory defined by the regularized problem VI(C, F_ϵ), where $F_\epsilon := F + \epsilon I$ with $\epsilon > 0$ and I being the identity operator, tends to the least norm solution of the pseudomonotone variational inequality VI(C, F) as $\epsilon \to 0$. So the problem of finding the limit in the Tikhonov method applying to pseudomonotone VIs leads to the problem of the form (1.1) with S being the solution set of the original problem and $x^g = 0$. It is well known that if F is monotone, then $F + \epsilon I$ is strongly monotone for every $\epsilon > 0$. However, when F is pseudomonotone, the operator $F + \epsilon I$ may not be strongly monotone, even not pseudomonotone for any $\epsilon > 0$ (see the Counterexample 2.1 in [26]). This example raises further an interesting question posed in [26] for the Tikhonov regularization method "Why one has to replace the original pseudomonotone VI by the sequence of auxiliary problems VI(C, F_ϵ), $\epsilon > 0$, none of which is pseudomonotone?". This question suggests to us to develop algorithms for solving bilevel problem (1.1). Namely, in this paper, we propose algorithms for solving bilevel variational inequality problem (1.1) when the lower problem is pseudomonotone with respect to its solution set. The latter property is somewhat more general than the pseudomonotonicity. The proposed algorithms can be considered as a combination of the well-known extragradient method using the auxiliary problem principle with the cutting plane technique previously used in some papers (see e.g. [28] and the references cited therein). The proposed algorithms show that with the help of the auxiliary problem VI(C, T_ϵ), the same limit point of every Tikhonov trajectory can be obtained by solving the bilevel problem (1.1).

2 Preliminaries

As usual, by P_C we denote the projection operator onto the closed convex set C with the norm $\|.\|$, that is,

$$P_C(x) \in C : \|x - P_C(x)\| \le \|x - y\| \quad \forall y \in C.$$

The following well-known results on the projection operator will be used in the sequel.

Lemma 2.1 [8, Lemma 12.1.13] *Suppose that C is a nonempty closed convex set in $\mathbb{R}^n$. Then*

(i) *$P_C(x)$ is singleton and well-defined for every x;*
(ii) *$\pi = P_C(x)$ if and only if $\pi \in C$, $\langle x - \pi, y - \pi\rangle \le 0\ \forall y \in C$;*
(iii) *$\|P_C(x) - P_C(y)\|^2 \le \|x - y\|^2 - \|P_C(x) - x + y - P_C(y)\|^2\ \forall x, y \in C$.*

We recall some well-known definitions on monotonicity (see e.g. [8, 16]).

Definition 2.1 An operator $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is said to be

(a) strongly monotone on C with modulus γ, if

$$\langle \phi(x) - \phi(y), x - y\rangle \ge \gamma\|x - y\|^2 \quad \forall x, y \in C;$$

(b) monotone on C if

$$\langle \phi(x) - \phi(y), x - y\rangle \ge 0 \quad \forall x, y \in C;$$

(c) pseudomonotone on C if

$$\langle \phi(y), x - y\rangle \ge 0 \Longrightarrow \langle \phi(x), x - y\rangle \ge 0 \quad \forall x, y \in C;$$

(d) pseudomonotone on C with respect to $x^* \in C$ if

$$\langle \phi(x^*), x - x^*\rangle \ge 0 \Longrightarrow \langle \phi(x), x - x^*\rangle \ge 0 \quad \forall x \in C.$$

The operator ϕ is pseudomonotone on C with respect to a set $A \subseteq C$ if it is pseudomonotone on C with respect to every point $x^* \in A$.

From the definitions it follows that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\forall x^* \in C$.
In what follows we need the following assumptions on F:

(A1) F is continuous on its domain;
(A2) F is pseudomonotone on C with respect to every solution of problem VI(C, F).

Lemma 2.2 *Suppose that assumptions* (A1), (A2) *are satisfied and that the variational inequality* (1.2) *admits a solution. Then the solution set of* (1.2) *is closed, convex.*

The proof of this lemma when F is pseudomonotone on C can be found in [8, 16, 24]. When F is pseudomonotone with respect to its solution set, the proof can be done in the same way.

Following the auxiliary problem principle [5, 21], let us define a bifunction $L : C \times C \to \mathbb{R}$ such that

(B1) $L(x,x)=0, \exists\beta>0: L(x,y)\geq \frac{\beta}{2}\|x-y\|^2\ \forall x,y\in C$;
(B2) L is continuous, $L(x,.)$ is differentiable, strongly convex on C for every $x\in C$ and $\nabla_2 L(x,x)=0$ for every $x\in C$.

An example for such a bifunction is the Bregman distance (see e.g. [4])

$$L(x,y):=g(y)-g(x)-\langle\nabla g(x), y-x\rangle$$

with g being any differentiable, strongly convex function on C with modulus $\beta>0$, for instance, $g(x)=\frac{1}{2}\|x\|^2$. The following lemma is well known from the auxiliary problem principle for VIs [21]. Since this lemma was not proved in [21], we give here a short proof for it.

Lemma 2.3 [21] *Suppose that F satisfies* (A1), (A2) *and L satisfies* (B1), (B2). *Then, for every $\rho>0$, the following statements are equivalent*:

(a) x^* *is a solution to VI*(C,F);
(b) $x^*\in C: \langle F(x^*), y-x^*\rangle+\frac{1}{\rho}L(x^*,y)\geq 0\ \forall y\in C$;
(c) $x^*=\operatorname{argmin}\{\langle F(x^*), y-x^*\rangle+\frac{1}{\rho}L(x^*,y): y\in C\}$;
(d) $x^*\in C: \langle F(y), y-x^*\rangle\geq 0\ \forall y\in C$.

Proof First we show that (a), (b) and (c) are equivalent. Indeed, since $L(x,y)\geq 0$ for every $x,y\in C$, so (a) follows (b). However, (b) holds if and only if (c) holds, that is,

$$x^*=\operatorname{argmin}\left\{\langle F(x^*), y-x^*\rangle+\frac{1}{\rho}L(x^*,y): y\in C\right\}.$$

In fact, by necessary and sufficient optimality condition for convex programming, the latter is equivalent to

$$0\in F(x^*)+\frac{1}{\rho}\nabla_2 L(x^*,x^*)+N_C(x^*)=F(x^*)+N_C(x^*),$$

which is equivalent to (a).

To see that (d) implies (a), we suppose on the contrary that there exists some $w\in C$ such that $\langle F(x^*), w-x^*\rangle<0$. Then take $y_t:=(1-t)x^*+tw$ with $0<t<1$. By the continuity of F we have $\langle F(y_{t_*}), w-x^*\rangle<0$ for some $0<t_*<1$. Then it follows from (d) that

$$0\leq\langle F(y_{t_*}), y_{t_*}-x^*\rangle=\langle F(y_{t_*}), t_*x^*+(1-t_*)w-x^*\rangle=(1-t_*)\langle F(y_{t_*}), w-x^*\rangle<0;$$

a contradiction.

The fact that (a) implies (d) is immediate from the pseudomonotonicity of F. □

3 The algorithms and their convergence

In what follows we suppose that the solution set S of the lower variational inequality (1.2) is nonempty and that F is continuous, pseudomonotone on C with respect to S. In this case, S is closed and convex. The following algorithm can be considered as a combination of the extragradient method [8, 13, 17, 25] and the cutting techniques [27, 28] to the bilevel problem (1.1).

Algorithm 1 Choose $\rho > 0$ and $\eta \in (0, 1)$. Starting from $x^1 := x^g \in C$ (x^g plays the role of a guessed solution).

Iteration k ($k = 1, 2, \ldots$) Having x^k, perform the following steps:

Step 1. Solve the strongly convex program

$$\min\left\{\langle F(x^k), y - x^k\rangle + \frac{1}{\rho} L(x^k, y) : y \in C\right\} \qquad CP(x^k)$$

to obtain its unique solution y^k.

If $y^k = x^k$, take $u^k = y^k$ and go to *Step 3*. Otherwise, go to *Step 2*.

Step 2. (Armijo linesearch) Find m_k as the smallest nonnegative integer number m satisfying

$$z^{k,m} := (1 - \eta^m)x^k + \eta^m y^k, \tag{3.1}$$

$$\langle F(z^{k,m}), y^k - z^{k,m}\rangle + \frac{1}{\rho} L(x^k, y^k) \le 0. \tag{3.2}$$

Set $\eta_k = \eta^{m_k}$, $z^k := z^{k,m_k}$ and compute

$$\sigma_k = \frac{-\eta_k \langle F(z^k), y^k - z^k\rangle}{(1 - \eta_k)\|F(z^k)\|^2}, \qquad u^k := P_C(x^k - \sigma_k F(z^k)). \tag{3.3}$$

Step 3. Having x^k and u^k, construct two halfspaces

$$C_k := \{y \in \mathbb{R}^n : \|u^k - y\|^2 \le \|x^k - y\|^2\};$$
$$D_k := \{y \in \mathbb{R}^n : \langle x^g - x^k, y - x^k\rangle \le 0\}.$$

Let $B_k := C_k \cap D_k \cap C$.

Step 4. Compute $x^{k+1} := P_{B_k}(x^g)$.

If $x^{k+1} \in S$, terminate: x^{k+1} solves the bilevel problem (1.1). Otherwise, increase k by one and go to iteration k.

Before considering validity and convergence of the algorithm, we would like to emphasize that the main difference between the just described algorithm with the other available ones [6, 7, 19, 20, 28, 30] applied to the bilevel problem (1.1) is that the lower VI in (1.1) is pseudomonotone with respect to its solution rather than monotone as in the above mentioned papers. Moreover, the subproblems to be solved in our algorithms are strongly convex programs, which seem numerically easier than strongly monotone VIs as in the above mentioned algorithms for the monotone case. As we have mentioned, for pseudomonotonicity case, the latter subvariational inequality is no longer strongly monotone, even not pseudomonotone. In our recent paper [1], the lower variational inequality can be pseudomonotone. However, the algorithm proposed there consists of two loops, and for the convergence, we have to assume that the inner loop must terminate after a finite number of projections. For monotonicity case this assumption is satisfied, but the number of the projections, although finite, cannot be estimated. It is well known that the solution set S of $\mathrm{VI}(C,F)$ is just the fixed point set of an operator T defined in some way, for example $T(x) := P_C(x-\rho F(x))$ with $\rho>0$ or $T(x) := T_{\mathrm{prox}}(x)$, where

$$T_{\mathrm{prox}}(x) := \left\{u \in C : \left\langle F(u), y-u\right\rangle + c\langle y-u, u-x\rangle \geq 0\ \forall y \in C\right\}$$

with $c>0$. Thus the bilevel variational inequality (1.1) belongs to the class of variational inequality problems over the fixed point set of an operator (mapping) T. The latter problem with T being a nonexpansive operator is solved by some algorithms based upon the fixed point approach (see e.g. [10, 20, 29] and the references therein). In [30], a hybrid steepest descent method was proposed for variational inequality over the fixed point set of a quasi-shrinking nonexpansive mapping. We note that the algorithms proposed in this paper do not belong to the steepest descent method and, for convergence, the sequences of iterates generated by our algorithms are not required to satisfy any additional condition as the algorithms in [30]. Moreover, the operator $T(x) := P_C(x-\rho F(x))$ may not be necessarily quasi-shrinking nonexpansive whereas T_{prox} may have nonconvex values when F is pseudomonotone with respect to the solution set S.

Remark 3.1 (i) The linesearch in *Step 2* is well defined. Indeed, otherwise for all nonnegative integer numbers m one has

$$\left\langle F\left(z^{k,m}\right), y^k - z^{k,m}\right\rangle + \frac{1}{\rho} L\left(x^k, y^k\right) > 0. \tag{3.4}$$

Thus letting $m \to \infty$, by continuity of F, and $z^{k,m} = (1-\eta^m)x^k + \eta^m y^k \to x^k$, we have

$$\left\langle F\left(x^k\right), y^k - x^k\right\rangle + \frac{1}{\rho} L\left(x^k, y^k\right) \geq 0,$$

which, together with

$$\left\langle F\left(x^k\right), x^k - x^k\right\rangle + \frac{1}{\rho} L\left(x^k, x^k\right) = 0,$$

implies that x^k is the solution of the strongly convex program $CP(x^k)$. Thus $x^k = y^k$ which contradicts the fact that the linesearch is performed only when $y^k \neq x^k$.

Note that $m_k > 0$. Indeed, if $m_k = 0$, then by the Armijo rule, we have $z^k = y^k$, and therefore

$$\frac{1}{\rho} L\left(x^k, y^k\right) = \left\langle F\left(z^k\right), y^k - z^k\right\rangle + \frac{1}{\rho} L\left(x^k, y^k\right) \leq 0,$$

which, together with nonnegativity of L, implies $L(x^k, y^k) = 0$. Since $L(x^k, y^k) \geq \frac{\beta}{2}\|x^k - y^k\|^2$, one has $x^k = y^k$.

(ii) The step size σ_k defined by (3.3) is positive whenever $x^k \neq y^k$.

Lemma 3.1 *Under the assumptions of Lemma 2.3, it holds that*

$$\left\|u^k - x^*\right\|^2 \leq \left\|x^k - x^*\right\|^2 - \sigma_k^2 \left\|F\left(z^k\right)\right\|^2 \quad \forall x^* \in S\ \forall k. \tag{3.5}$$

Proof The proof of this lemma can be done similarly as the proof of Lemma 12.1.10 in [8] (see also [23]). So we give here only a sketch. For the simplicity of notation, we write F^k for $F(z^k)$ and v^k for $x^k - \sigma_k F^k$. Since $u^k = P_C(v^k)$, by the nonexpansiveness of the projection we have

$$\begin{aligned}\|u^k - x^*\|^2 = \|P_C(v^k) - P_C(x^*)\|^2 \le \|v^k - x^*\|^2 \\ = \|x^k - x^* - \sigma_k F^k\|^2 \\ = \|x^k - x^*\|^2 + \sigma_k^2 \|F^k\|^2 - 2\sigma_k \langle F^k, x^k - x^*\rangle. \end{aligned} \tag{3.6}$$

Since $x^* \in S$, using assertion (d) of Lemma 2.3 we can write

$$\langle F^k, x^k - x^*\rangle = \langle F^k, x^k - z^k + z^k - x^*\rangle \ge \langle F^k, x^k - z^k\rangle. \tag{3.7}$$

Since $x^k - z^k = \frac{\eta_k}{1-\eta_k}(z^k - y^k)$,

$$\langle F^k, x^k - z^k\rangle = \frac{\eta_k}{1-\eta_k}\langle F^k, z^k - y^k\rangle = \sigma_k \|F^k\|^2. \tag{3.8}$$

The last equality comes from the definition of σ_k by (3.3) in the algorithm. Combining (3.6), (3.7) and (3.8) we obtain (3.5). □

The following theorem shows validity and convergence of the algorithm, with parts of the proof using a technique in [28].

Theorem 3.1 *Suppose that the assumptions* (A1), (A2) *and* (B1), (B2) *are satisfied and that VI*(C, F) *admits a solution. Then both the sequences* $\{x^k\}$, $\{u^k\}$ *converge to the unique solution of the original bilevel problem* (1.1).

Proof As we have remarked, the linesearch used in the algorithm is well defined. So, to see validity of the algorithm, it is sufficient to show that $S \subseteq B_k$ for every k. Indeed, from Lemma 3.1, it follows that $\|u^k - x^*\| \le \|x^k - x^*\|$ for every k and $x^* \in S$. Hence, by the definition of C_k, one has $\emptyset \ne S \subseteq C_k$ for every k. Moreover, $S \subseteq D_k$ for every k. In fact, since $x^1 = x^g$, $S \subseteq D_1 = \mathbb{R}^n$. By the definition of x^{k+1}, it follows, by induction, that if $S \subseteq D_k$, then, since $x^{k+1} = P_{B_k}(x^g)$, we have $S \subseteq D_{k+1}$. Consequently, $\emptyset \ne S \subseteq C_k \cap D_k \cap C = B_k$ for every k. Hence the projection onto B_k is well defined. Due to the definition of D_k, by the assertion (ii) of Lemma 2.1 we have $x^k = P_{D_k}(x^g)$.

Note that $x^{k+1} \in D_k$, we can write $\|x^k - x^g\| \le \|x^{k+1} - x^g\|$ for every k. Moreover, since $x^k = P_{D_k}(x^g)$ and $S \subset D_k$ for every k, we have $\|x^k - x^g\| \le \|x^* - x^g\|$ for any $x^* \in S$ and for every k, therefore $\{x^k\}$ is bounded.

From the boundedness of $\{x^k\}$ and the inequality $\|x^k - x^g\| \le \|x^{k+1} - x^g\|$ for every k, it follows that $\lim_k \|x^k - x^g\|$ exists and is finite. The sequence $\{x^k\}$ is asymptotically regular, i.e., $\|x^{k+1} - x^k\| \to 0$ as $k \to \infty$. Indeed, since $x^k \in D_k$ and $x^{k+1} \in D_k$, by the convexity of D_k one has $\frac{x^{k+1}+x^k}{2} \in D_k$. By the definition of D_k, we have $x^k = P_{D_k}(x^g)$. Then we can write

$$\begin{aligned}\|x^g - x^k\|^2 &\le \left\|x^g - \frac{x^{k+1} + x^k}{2}\right\|^2\\ &= \left\|\frac{x^g - x^{k+1}}{2} + \frac{x^g - x^k}{2}\right\|^2\\ &= \frac{1}{2}\|x^g - x^{k+1}\|^2 + \frac{1}{2}\|x^g - x^k\|^2 - \frac{1}{4}\|x^{k+1} - x^k\|^2,\end{aligned}$$

where the last equality comes from the property of the Euclidean norm (see e.g. [2, Corollary 2.14]). Thus we have

$$\frac{1}{2}\|x^{k+1} - x^k\|^2 \le \|x^g - x^{k+1}\|^2 - \|x^g - x^k\|^2.$$

Remembering that $\lim_k \|x^k - x^g\|$ does exist and is finite, we obtain $\|x^{k+1} - x^k\| \to 0$ as $k \to \infty$.

On the other hand, since $x^{k+1} \in B_k \subseteq C_k$, by the definition of C_k, we have

$$\|u^k - x^{k+1}\| \le \|x^{k+1} - x^k\|.$$

Thus,

$$\|u^k - x^k\| \le \|u^k - x^{k+1}\| + \|x^{k+1} - x^k\| \le 2\|x^{k+1} - x^k\|,$$

which, together with $\|x^{k+1} - x^k\| \to 0$, implies that $\|u^k - x^k\| \to 0$ as $k \to \infty$.

Next, we show that any cluster point of the sequence $\{x^k\}$ is a solution to variational inequality $\mathrm{VI}(C, F)$. Indeed, let $\bar{x}$ be any cluster point of $\{x^k\}$. For the simplicity of notation, without loss of generality we may assume that $x^k \to \bar{x}$. We consider two distinct cases:

Case 1: The linesearch is performed only for finitely many k. In this case, by the algorithm, either $u^k = x^k$ or $u^k = z^k$ for infinitely many k. In the first case, x^k is a solution to $\mathrm{VI}(C, F)$, while in the latter case, u^k is a solution to $\mathrm{VI}(C, F)$ for infinitely many k. Hence, by $\|u^k - x^k\| \to 0$, we see that $\bar{x}$ is a solution to $\mathrm{VI}(C, F)$.

Case 2: The linesearch is performed for infinitely many k. Then, by taking a subsequence if necessary, we may assume that the linesearch is performed for every k.

We distinguish two possibilities:

(a) $\overline{\lim}_k \eta_k > 0$. From $x^k \to \bar{x}$ and $\|u^k - x^k\| \to 0$ it follows that $u^k \to \bar{x}$. Then applying (3.5) with some $x^* \in S$ we see that $\sigma_k \, \|F^k\|^2 \to 0$. Thus by the definition of σ_k, we have $-\frac{\eta_k}{1-\eta_k}\langle F^k, y^k - z^k\rangle \to 0$. Since $\overline{\lim}_k \eta_k > 0$, by taking again a subsequence if necessary, we may assume that $\langle F^k, y^k - z^k\rangle \to 0$. On the other hand, using assumption (B1) and the Armijo rule, we can write

$$0 \le \frac{\beta}{2\rho}\|x^k - y^k\|^2 \le \frac{1}{2\rho}L(x^k, y^k) \le -\langle F^k, y^k - z^k\rangle \to 0.$$

Hence $\|x^k - y^k\| \to 0$, which, together with $x^k \to \bar{x}$, implies that $y^k \to \bar{x}$. Since y^k is a solution of the problem

$$\min\left\{\langle F(x^k), y - x^k\rangle + \frac{1}{\rho}L(x^k, y) : y \in C\right\}, \qquad CP(x^k)$$

we have

$$\langle F(x^k), y - x^k\rangle + \frac{1}{\rho}L(x^k, y) \geq \langle F(x^k), y^k - x^k\rangle + \frac{1}{\rho}L(x^k, y^k) \quad \forall y \in C.$$

Letting k to infinity, by the continuity of F and L we obtain

$$\langle F(\bar{x}), y - \bar{x}\rangle + \frac{1}{\rho}L(\bar{x}, y) \geq \langle F(\bar{x}), \bar{x} - \bar{x}\rangle + \frac{1}{\rho}L(\bar{x}, \bar{x}) \quad \forall y \in C.$$

which means that $\bar{x}$ is a solution of CP$(\bar{x})$. Then, by Lemma 2.3, $\overline{x}$ is a solution of VI(C, F).

(b) $\lim_k \eta_k = 0$. In this case, the sequence $\{y^k\}$ is also bounded. Indeed, since y^k is the solution of $CP(x^k)$, we have

$$\langle F(x^k), y - x^k\rangle + \frac{1}{\rho}L(x^k, y) \geq \langle F(x^k), y^k - x^k\rangle + \frac{1}{\rho}L(x^k, y^k) \quad \forall y \in C.$$

In particular, with $y = x^k$, by (B1), one can write

$$0 \geq \langle F(x^k), y^k - x^k\rangle + \frac{1}{\rho}L(x^k, y^k) \geq \langle F(x^k), y^k - x^k\rangle + \frac{\beta}{2\rho}\|x^k - y^k\|^2,$$

which implies that $\|x^k - y^k\| \leq \frac{2\rho}{\beta}\|F(x^k)\|$. Since $\{x^k\}$ is bounded and F is continuous, $\{y^k\}$ is bounded. Thus, we may assume, taking a subsequence if necessary, that $y^k \to \overline{y}$ for some $\overline{y}$. By the same argument as before we have

$$\langle F(\overline{x}), \overline{y} - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, \overline{y}) \leq \langle F(\overline{x}), y - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, y) \quad \forall y \in C. \tag{3.9}$$

On the other hand, as m_k is the smallest natural number satisfying the Armijo linesearch rule, we have

$$\langle F(z^{k,m_k-1}), y^k - z^{k,m_k-1}\rangle + \frac{1}{\rho}L(x^k, y^k) > 0.$$

Since that $z^{k,m_k-1} \to \overline{x}$ as $k \to \infty$ and since F and L are continuous, from the last inequality we obtain in the limit that

$$\langle F(\overline{x}), \overline{y} - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, \overline{y}) \geq 0. \tag{3.10}$$

Substituting $y = \overline{x}$ into (3.9) we get

$$\langle F(\overline{x}), \overline{y} - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, \overline{y}) \leq 0,$$

which, together with (3.10), yields

$$\langle F(\overline{x}), \overline{y} - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, \overline{y}) = 0. \tag{3.11}$$

From (3.11) and

$$\langle F(\overline{x}), \overline{x} - \overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, \overline{x}) = 0$$

it follows that both $\overline{x}$ and $\overline{y}$ are solutions of the strongly convex program

$$\min\left\{\langle F(\overline{x}), y-\overline{x}\rangle + \frac{1}{\rho}L(\overline{x}, y) : y \in C\right\}.$$

Hence $\overline{x} = \overline{y}$ and, therefore, by Lemma 2.3, $\overline{x}$ solves VI(C, F). Moreover, from $\|u^k - x^k\| \to 0$, we can also conclude that every cluster point of $\{u^k\}$ is a solution to VI(C, F).

Finally, we show that $\{x^k\}$ converges to the unique solution of the bilevel problem (1.1). To this end, let x^* be any cluster point of $\{x^k\}$. Then there exists a subsequence $\{x^{k_j}\}$ such that $x^{k_j} \to x^*$ as $j \to \infty$. Since, by the algorithm, $x^{k_j} = P_{B_{k_j-1}}(x^g)$, we have

$$\langle x^{k_j} - x^g, y - x^{k_j}\rangle \geq 0 \quad \forall y \in B_{k_j-1} \supseteq S \; \forall j.$$

Letting $j \to \infty$ we obtain

$$\langle x^* - x^g, y - x^*\rangle \geq 0 \quad \forall y \in S,$$

which, together with $x^* \in S$, implies that $x^* = P_S(x^g)$. Thus, the whole sequence must converge to the unique solution to the original bilevel problem (1.1). Then, since $\|x^k - u^k\| \to 0$, it follows that u^k converges to the solution of (1.1). □

Other Armijo rules can be chosen for the above algorithm. For example, one can use the linesearch in [27]. Then at each iteration k in Algorithm 1, Step 2 and Step 3 are replaced by Step 2a and Step 3a below.

Step 2a. Having $x^k \in C$, find m_k as the smallest nonnegative integer m such that

$$\langle F(x^k - \eta^m r(x^k)), r(x^k)\rangle \geq \sigma \|r(x^k)\|^2, \tag{3.12}$$

where $\sigma \in (0, 1)$, $r(x^k) = x^k - y^k$ with y^k being the unique solution of the problem $CP(x^k)$. Take $z^k := (1-\eta_k)x^k + \eta_k y^k$. If $F(z^k) = 0$, take $u^k := z^k$. Otherwise, go to Step 4 (in Algorithm 1).

Step 3a. Define

$$H_k := \{x \in C : \langle F(z^k), x - z^k\rangle \leq 0\}$$

and take $u^k := P_{H_k}(x^k)$.

It is easy to see that if F is pseudomonotone on C with respect to S, then $S \subseteq H_k$. In the same way as the linesearch rule (3.2) we can show that the linesearch (3.12) is well defined. Moreover, as in the linesearch (3.2), the following lemma can be proved for (3.12) by the same idea as in the proof of Theorem 2.1 in [27].

Lemma 3.2 *Under the assumptions of Lemma* 2.3, *it holds that*

$$\|u^k - x^*\|^2 \leq \|x^k - x^*\|^2 - \frac{\eta_k \sigma}{\|F(z^k)\|^2}\|r(x^k)\|^4 \quad \forall x^* \in S \; \forall k. \tag{3.13}$$

Using this lemma we can prove the following convergence theorem by the same arguments as in the proof of Theorem 3.2.

Theorem 3.2 *Suppose that assumptions* (A1), (A2) *and* (B1), (B2) *are satisfied and that VI(C, F) admits a solution. Then Algorithm* 1 *with Steps* 2 *and* 3 *replaced by Steps* 2*a and* 3*a respectively is well defined and both the sequences* $\{x^k\}$, $\{u^k\}$ *converge to the unique solution of* (1.1).

Acknowledgements We would like to thank the Associate Editor and the referees for their useful remarks and comments that helped us very much in revising the paper. This work is supported by the Vietnam National Foundation for Science Technology Development (NAFOSTED) under Grant 101-02-2011.19.

References

1. Anh, P.N., Kim, J.K., Muu, L.D.: An extragradient algorithm for solving bilevel pseudomonotone variational inequalities. J. Glob. Optim. **52**, 627–639 (2012)
2. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Springer, New York (2011)
3. Ceng, L.C., Ansari, Q.H., Yao, J.-C.: Iterative methods for triple hierarchical variational inequalities in Hilbert spaces. J. Optim. Theory Appl. **151**, 489–512 (2011)
4. Censor, Y., Lent, A.: An iterative row-action method for interval convex programming. J. Optim. Theory Appl. **34**, 321–353 (1981)
5. Cohen, G.: Auxiliary problem principle extended to variational inequalities. J. Optim. Theory Appl. **59**, 325–333 (1988)
6. Ding, X.P.: Auxiliary principle and algorithm for mixed equilibrium problems and bilevel equilibrium problems in Banach spaces. J. Optim. Theory Appl. **146**, 347–357 (2010)
7. Dinh, B.V., Muu, L.D.: On penalty and gap function methods for bilevel equilibrium problems. J. Appl. Math. (2011). doi:10.1155/2011/646452
8. Facchinei, F., Pang, J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems. Springer, New York (2003)
9. Giannessi, F., Maugeri, A., Pardalos, P.M. (eds.): Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models. Nonconvex Optim. Appl., vol. 58. Kluwer Academic, Dordrecht (2001)
10. Giuseppe, M., Xu, H.-K.: Explicit hierarchical fixed point approach to variational inequalities. J. Optim. Theory Appl. **149**, 61–78 (2011)
11. Hao, N.T.: The Tikhonov regularization algorithm for pseudomonotone variational inequalities. Acta Math. Vietnam. **31**, 283–289 (2006)
12. Hung, P.G., Muu, L.D.: The Tikhonov regularization extended to equilibrium problems involving pseudomonotone bifunctions. Nonlinear Anal. **74**, 6121–6129 (2011)
13. Iusem, A.N., Swaiter, B.F.: A variant of Korpelevich's method for variational inequalities with a new search strategy. Optimization **42**, 309–321 (1997)
14. Kalashnikov, V.V., Klashnikova, N.I.: Solving two-level variational inequality. J. Glob. Optim. **8**, 289–294 (1996)
15. Kinderlehrer, D., Stampachia, G.: An Introduction to Variational Inequalities and Their Applications. Pure Appl. Math., vol. 88. Academic Press, New York (1980)
16. Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Lecture Notes in Economics and Mathematical Systems, vol. 495. Springer, Berlin (2001)
17. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Èkon. Mat. Metody **12**, 747–756 (1976)
18. Luo, J.Q., Pang, J.-S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge (1996)
19. Maingé, P.-E.: Projected subgradient techniques and viscosity methods for optimization with variational inequality constraints. Eur. J. Oper. Res. **205**, 501–506 (2010)
20. Marino, G., Xu, H.-K.: Explicit hierarchical fixed point approach to variational inequalities. J. Optim. Theory Appl. **149**, 61–78 (2011)
21. Mastroeni, G.: Gap functions for equilibrium problems. J. Glob. Optim. **27**, 411–426 (2003)
22. Moudafi, A.: Proximal methods for a class of bilevel monotone equilibrium problems. J. Glob. Optim. **47**, 287–292 (2010)
23. Quoc, T.D., Dung, L.M., Van Nguyen, H.: Extragradient algorithms extended to equilibrium problems. Optimization **57**, 749–776 (2008)
24. Muu, L.D., Oettli, W.: Convergence of an adaptive penalty scheme for finding constrained equilibria. Nonlinear Anal. **18**, 1159–1166 (1992)
25. Noor, M.A.: Extragradient methods for pseudomonotone variational inequalities. J. Optim. Theory Appl. **117**, 475–488 (2003)
26. Tam, N.N., Yao, J.C., Yen, N.D.: Solution methods for pseudomonotone variational inequalities. J. Optim. Theory Appl. **138**, 253–273 (2008)
27. Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequality problems. SIAM J. Control Optim. **37**, 765–776 (1999)

28. Tada, A., Takahashi, W.: Weak and strong convergence theorems for a nonexpansive mapping and an equilibrium problem. J. Optim. Theory Appl. **133**, 359–370 (2007)
29. Yamada, I., Ogura, N.: Hybrid steepest descent method for variational inequality problem over the fixed point of certain quasi-nonexpansive mappings. Numer. Funct. Anal. Optim. **25**, 619–655 (2004)
30. Yao, Y., Liou, Y.-C., Kang, S.M.: Minimization of equilibrium problems, variational inequality problems and fixed point problems. J. Glob. Optim. **48**, 643–656 (2010)