Optimization,
2007, 1–30, iFirst

Extragradient algorithms extended to equilibrium problems¶

D. QUOC TRAN†, M. LE DUNG*‡ and VAN HIEN NGUYEN§

†National University of Hanoi, Vietnam
‡Hanoi Institute of Mathematics, Vietnam
§University of Namur, Belgium

(*Received 5 September 2005; in final form 17 October 2006*)

We make use of the auxiliary problem principle to develop iterative algorithms for solving equilibrium problems. The first one is an extension of the extragradient algorithm to equilibrium problems. In this algorithm the equilibrium bifunction is not required to satisfy any monotonicity property, but it must satisfy a certain Lipschitz-type condition. To avoid this requirement we propose linesearch procedures commonly used in variational inequalities to obtain projection-type algorithms for solving equilibrium problems. Applications to mixed variational inequalities are discussed. A special class of equilibrium problems is investigated and some preliminary computational results are reported.

Keywords: Equilibrium problem; Extragradient method; Linesearch; Auxiliary problem principle; Variational inequality

Mathematics Subject Classifications 2000: 65K10; 90C25

1. Introduction and the problem statement

Let K be a nonempty closed convex subset of the n-dimensional Euclidean space $\mathbb{R}^n$ and let $f: K \times K \to \mathbb{R} \cup \{+\infty\}$. Consider the following equilibrium problem in the sense of Blum and Oettli [6]:

$$\text{Find } x^* \in K \text{ such that} \quad f(x^*, y) \geq 0 \quad \text{for all } y \in K \tag{PEP}$$

where $f(x, x) = 0$ for every $x \in K$. As usual, we call a bifunction satisfying this property an *equilibrium bifunction* on K.

Equilibrium problems have been considered by several authors (see e.g. [6,12,13,21,22] and the references therein). It is well known (see e.g. [13,21,23]) that various classes of mathematical programing problems, variational inequalities, fixed point problems, Nash equilibrium in noncooperative games theory and minimax problems can be formulated in the form of (PEP).

*Corresponding author. Email: ldmuu@math.ac.vn
¶This article is dedicated to the Memory of W. Oettli.

Optimization
ISSN 0233-1934 print: ISSN 1029-4945 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/02331930601122876

The proximal point method was first introduced by Martinet in [16] for solving variational inequalities and then extended by Rockafellar [28] to the problem of finding a zero of a maximal monotone operator. Moudafi [20] further extended the proximal point method to monotone equilibrium problems. Konnov [14] used the proximal point method for solving Problem (PEP) with f being a weakly monotone equilibrium function.

Another strategy is to use, as for variational inequality problems, a gap function in order to convert an equilibrium problem into an optimization problem [14,18]. In general, the transformed mathematical programing problem is not convex.

The auxiliary problem principle, first introduced for solving optimization problems, by Cohen in [7], and then extended to variational inequalities in [8], becomes a useful tool for analyzing and developing efficient algorithms for the solution to various classes of mathematical programming and variational inequality problems (see e.g. [1,2,7–9,11,24,29] and the references cited therein). Recently, Mastroeni in [17] further extended the auxiliary problem principle to equilibrium problems involving strongly monotone equilibrium bifunctions satisfying some Lipschitz-type condition. Noor in [25] used the auxiliary problem principle to develop iterative methods for solving problems where the equlibrium bifunctions are supposed to be partially relaxed strongly monotone. As in the proximal point method, the subproblems needed to solve in these methods are strongly monotone equilibrium problems. In a recent article, Nguyen *et al.* [31] developed a bundle method for solving problems where the equilibrium functions satisfy a certain cocoercivity condition. A continuous extragradient method is proposed in [3] for solving equilibrium problems with skew bifunctions.

It is well known that algorithms based upon the auxiliary problem principle, in general, are not convergent for monotone variational inequalities that are special cases of the monotone equilibrium problem (PEP). To overcome this drawback, the extragradient method, first introduced by Korpelevich [15] for finding saddle points, is used to solve monotone, even pseudomonotone, variational inequalities [9,23,24].

In this article, we use the auxiliary problem principle to extend the extragradient method to equilibrium problems. By this way, we obtain extragradient algorithms for solving Problem (PEP). Convergence of the proposed algorithms does not require f to satisfy any type of monotonicity, but it must satisfy a certain Lipschitz-type condition as introduced in [17]. In order to avoid this requirement, we use a linesearch technique to obtain convergent algorithms for solving (PEP).

The rest of the article is organized as follows. In the next section, we give fixed-point formulations to Problem (PEP). We then use these formulations in the third section to describe an extragradient algorithm for (PEP). Section four is devoted to presentation of linesearch algorithms and their convergence results avoiding the aforementioned Lipschitz-type condition. In section five, we discuss applications of the proposed algorithms to mixed multivalued variational inequalities. The last section contains some preliminary computational results and experiments.

2. Fixed point formulations

First we recall some well-known definitions on monotonicity that we need in the sequel.

Definition 2.1 Let M and K be nonempty convex sets in $\mathbb{R}^n$, $M \subseteq K$, and let $f: K \times K \to \mathbb{R} \cup \{+\infty\}$. The bifunction f is said to be

(a) strongly monotone on M with constant $\tau > 0$ if for each pair $x, y \in M$, we have

$$f(x,y) + f(y,x) \leq -\tau\|x-y\|^2;$$

(b) strictly monotone on M if for all distinct $x, y \in M$, we have

$$f(x,y) + f(y,x) < 0;$$

(c) *monotone* on M if for each pair $x, y \in M$, we have

$$f(x,y) + f(y,x) \leq 0;$$

(d) *pseudomonotone* on M if for each pair $x, y \in M$ it holds that

$$f(x,y) \geq 0 \quad \text{implies} \quad f(y,x) \leq 0;$$

From the definition above we obviously have the following implications:

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d).$$

Following [14], associated with (PEP) we consider the following dual problem of (PEP)

$$\text{Find } x^* \in K \text{ such that} \quad f(y, x^*) \leq 0 \quad \forall y \in K. \tag{DEP}$$

For each $x \in K$, let

$$L_f(x) := \{y \in K : f(x,y) \leq 0\}.$$

Clearly, x^* is a solution to (DEP) if and only if $x^* \in \bigcap_{x \in K} L_f(x)$.

We will denote by K^* and K^d the solution sets of (PEP) and (DEP), respectively. Conditions under which (PEP) and (DEP) have solutions can be found, for example, in [6,12,13,30] and the references therein. Since $K^d = \bigcap_{x \in K} L_f(x)$, the solution set K^d is closed convex if $f(x, \cdot)$ is closed convex on K. In general, K^* may not be convex. However, if f is closed convex on K with respect to the second variable and hemi-continuous with respect to the first variable, then K^* is convex and $K^d \subseteq K^*$. Moreover, if f is pseudomonotone on K, then $K^* = K^d$ (see [14,21]). In what follows, we suppose that $K^d \neq \emptyset$.

The following lemma gives a fixed-point formulation for (PEP).

Lemma 2.1 ([17,23]) *Let $f: K \times K \to \mathbb{R} \cup \{+\infty\}$ be an equilibrium bifunction. Then the following statements are equivalent*:

(i) *x^* is a solution to (PEP)*;
(ii) *x^* is a solution to the problem*

$$\min_{y \in K} f(x^*, y). \tag{2.1}$$

The main drawback of the fixed-point formulation given by Lemma 2.1 is that Problem (2.1), in general, may not have a solution, and if it does, the solution may not be unique. To avoid this situation, it is very helpful to use another auxiliary equilibrium problem that is equivalent to (PEP).

Let $L : K \times K \to \mathbb{R}$ be a nonnegative differentiable convex bifunction on K with respect to the second argument y (for each fixed $x \in K$) such that

(i) $L(x, x) = 0$ for all $x \in K$,
(ii) $\nabla_2 L(x, x) = 0$ for all $x \in K$

where, as usual, $\nabla_2 L(x, x)$ denotes the gradient of the function $L(x, \cdot)$ at x. An important example for such a function is $L(x, y) := \frac{1}{2}\|y - x\|^2$.

We consider the auxiliary equilibrium problem defined as

$$\text{Find } x^* \in K \text{ such that} \quad \rho f(x^*, y) + L(x^*, y) \geq 0 \quad \text{for all } y \in K \tag{AuPEP}$$

where $\rho > 0$ is a regularization parameter.

Applying Lemma 2.1 to the equilibrium function $\rho f + L$ we see that x^* is a minimizer of the convex program

$$\min_{y \in K} \{\rho f(x^*, y) + L(x^*, y)\}. \tag{2.2}$$

Equivalence between (PEP) and (AuPEP) is stated in the following lemma.

LEMMA 2.2 ([17,23]) Let $f : K \times K \to \mathbb{R} \cup \{\{+\infty\}\}$ *be an equilibrium bifunction, and let* $x^* \in K$. *Suppose that* $f(x^*, \cdot) : K \to \mathbb{R}$ *is convex and subdifferentiable on* K. *Let* $L : K \times K \to \mathbb{R}_+$ *be a differentiable convex function on* K *with respect to the second argument* y *such that*

(i) $L(x^*, x^*) = 0$,
(ii) $\nabla_2 L(x^*, x^*) = 0$.

Then $x^* \in K$ *is a solution to* (*PEP*) *if and only if* x^* *is a solution to* (*AuPEP*).

We omit the proof for this nondifferentiable case because it is similar to the one given in [17,23] for differentiable case.

3. An extragradient algorithm for EP

As we have mentioned, if $f(x, \cdot)$ is closed convex on K and $f(\cdot, y)$ is upper hemicontinuous on K, then the solution set of (DEP) is contained in that of (PEP). In the following algorithm, as in [17], we use the auxiliary bifunction given by

$$L(x, y) := G(y) - G(x) - \langle \nabla G(x), y - x \rangle, \tag{3.1}$$

where $G : \mathbb{R}^n \to \mathbb{R}$ is a strongly convex (with modulus $\beta > 0$) and continuously differentiable function; for example $G(x) = \frac{1}{2}\|x\|^2$.

Since G is strongly convex on the closed convex set K, the problem

$$\min_{y \in K} \left\{ \rho f(x, y) + G(y) - G(x) - \langle \nabla G(x), y - x \rangle \right\} \tag{Cx}$$

always admits a unique solution.

Lemma 2.2 gives a fixed-point formulation for Problem (PEP) that suggests an iterative method for solving (PEP) by setting $x^{k+1} = s(x^k)$ where $s(x^k)$ is the unique solution of the strongly convex problem (Cx^k). Unfortunately, it is well known (see also [9]) that, for monotone variational inequality problems, which are special cases of monotone equilibrium problem (PEP), the sequence $\{x^k\}$ may not be convergent. This fact suggested the use of the extragadient method introduced by Korpelevich in [15], first for finding saddle points, to monotone variational inequalities [9,23]. For the singlevalued variational inequality problem given as

$$\text{Find } x^* \in K \text{ such that } \quad \langle F(x^*), x - x^* \rangle \geq 0 \quad \text{for all } x \in K \tag{VIP}$$

the extragradient (or double projection) method constructs two sequences $\{x^k\}$ and $\{y^k\}$ by setting

$$y^k := \Pi_K\left(x^k - \rho F(x^k)\right) \quad \text{and} \quad x^{k+1} := \Pi_K\left(x^k - \rho F(y^k)\right)$$

where $\rho > 0$ and Π_K denotes the Euclidean projection onto K.

Now we further extend the extragradient method to equilibrium problem (PEP). Throughout the rest of the article, we suppose that the function $f(x, \cdot)$ is closed, convex and subdifferentiable on K for each $x \in K$. Under this assumption, subproblems needed to solve in the algorithms below are convex programs with strongly convex objective functions. In Algorithm 1 we are going to describe, in order to be able to obtain its convergence, the regularization ρ must satisfy some condition (see convergence Theorem 3.2).

Algorithm 1

Step 0 Take $x^0 \in K$, $\rho > 0$ and set $k := 0$.
Step 1 Solve the strongly convex program

$$\min_{y \in K} \left\{ \rho f(x^k, y) + G(y) - \langle \nabla G(x^k), y - x^k \rangle \right\} \tag{3.2}$$

to obtain its unique optimal solution y^k.
If $y^k = x^k$, then stop: x^k is a solution to (PEP). Otherwise, go to Step 2.
Step 2 Solve the strongly convex program

$$\min_{y \in K} \left\{ \rho f(y^k, y) + G(y) - \langle \nabla G(x^k), y - x^k \rangle \right\} \tag{3.3}$$

to obtain its unique solution x^{k+1}.
Step 3 Set $k := k + 1$, and go back to Step 1.

The following lemma shows that, if Algorithm 1 terminates after a finite number of iterations, then a solution to (PEP) has already been found.

LEMMA 3.1 *If the algorithm terminates at some iterate point x^k, then x^k is a solution of (PEP).*

Proof If $y^k = x^k$, then, by the fact that $f(x, x) = 0$, we have

$$\rho f\left(x^k, y^k\right) + G\left(y^k\right) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y^k - x^k\right\rangle = 0.$$

Since $y^k = x^k$ is the solution of (3.2), we have

$$\begin{aligned} 0 &= \rho f\left(x^k, y^k\right) + G\left(y^k\right) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y^k - x^k\right\rangle \\ &\le \rho f\left(x^k, y\right) + G(y) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y - x^k\right\rangle \quad \forall y \in K. \end{aligned}$$

Thus, by Lemma 2.2, x^k is a solution to (PEP). ■

The following theorem establishes the convergence of the algorithm.

THEOREM 3.2 *Suppose that*

(i) *G is strongly convex with modulus $\beta > 0$ and continuously differentiable on an open set Ω containing K.*
(ii) *There exist two constants $c_1 > 0$ and $c_2 > 0$ such that*

$$f(x, y) + f(y, z) \ge f(x, z) - c_1\|y - x\|^2 - c_2\|z - y\|^2 \quad \forall x, y, z \in K. \tag{3.4}$$

Then

(a) *For every $x^* \in K^d$, it holds true*

$$l\left(x^k\right) - l\left(x^{k+1}\right) \ge \left(\frac{\beta}{2} - \rho c_1\right)\left\|y^k - x^k\right\|^2 + \left(\frac{\beta}{2} - \rho c_2\right)\left\|x^{k+1} - y^k\right\|^2 \tag{3.5}$$

where $l(y) := G(x^*) - G(y) - \langle \nabla G(y), x^* - y\rangle$ for each $y \in K$.

(b) *Suppose in addition that f is lower semicontinuous on $K \times K$, $f(\cdot, y)$ is upper semicontinuous on K, and $0 < \rho < \min\{\{(\beta/2c_1)\}, (\beta/2c_2)$, then the sequence $\{x^k\}$ is bounded, and every cluster point of $\{x^k\}$ is a solution to (DEP).*
Moreover, if $K^d = K^*$ (in particular, if f is pseudomonotone on K), then the whole sequence $\{x^k\}$ converges to a solution of (PEP).

Proof (a) Take any $x^* \in K^d$. By the definition of l and since $x^k, x^{k+1} \in K$, we have

$$\begin{aligned} l\left(x^k\right) - l\left(x^{k+1}\right) &= G\left(x^{k+1}\right) - G\left(x^k\right) + \left\langle \nabla G\left(x^{k+1}\right), x^* - x^{k+1}\right\rangle \\ &\quad - \left\langle \nabla G\left(x^k\right), x^* - x^k\right\rangle \\ &= G\left(x^{k+1}\right) - G\left(x^k\right) + \left\langle \nabla G\left(x^{k+1}\right) - \nabla G\left(x^k\right), x^* - x^{k+1}\right\rangle \\ &\quad - \left\langle \nabla G\left(x^k\right), x^{k+1} - x^k\right\rangle. \end{aligned} \tag{3.6}$$

Using the well-known necessary and sufficient condition for optimality of convex programing [27] we see that x^{k+1} solves the convex program

$$\min_{y\in K}\{\rho f(y^k,y)+G(y)-G(x^k)-\langle\nabla G(x^k),y-x^k\rangle\},$$

if and only if

$$0\in\partial_2\Big\{\rho f\big(y^k,x^{k+1}\big)+G\big(x^{k+1}\big)-G\big(x^k\big)-\big\langle\nabla G\big(x^k\big),x^{k+1}-x^k\big\rangle\Big\}+N_K\big(x^{k+1}\big).$$

where $N_K(x)$ is the (outward) normal cone of K at $x\in K$.

Thus, since $f(y^k,\cdot)$ is subdifferentiable and G is strongly convex, differentiable on K, by the well-known Moreau-Rockafellar theorem [xxvii], there exists $w\in\partial_2 f(y^k,x^{k+1})$ such that

$$\big\langle\nabla G\big(x^{k+1}\big)-\nabla G\big(x^k\big),y-x^{k+1}\big\rangle\geq\rho\langle w,x^{k+1}-y\rangle\quad\forall y\in K.$$

By the definition of subgradient we have, from the latter inequality, that

$$\big\langle\nabla G\big(x^{k+1}\big)-\nabla G\big(x^k\big),y-x^{k+1}\big\rangle\geq\rho f\big(y^k,x^{k+1}\big)-\rho f(y^k,y)\quad\forall y\in K.$$

With $y=x^*$, this inequality becomes

$$\big\langle\nabla G\big(x^{k+1}\big)-\nabla G\big(x^k\big),x^*-x^{k+1}\big\rangle\geq\rho f\big(y^k,x^{k+1}\big)-\rho f\big(y^k,x^*\big).$$

Since x^* is a solution to (DEP), $f(y^k,x^*)\leq 0$. Thus

$$\big\langle\nabla G\big(x^{k+1}\big)-\nabla G\big(x^k\big),x^*-x^{k+1}\big\rangle\geq\rho f\big(y^k,x^{k+1}\big).\tag{3.7}$$

Now applying Assumption (3.4) with $x=x\hat{}k$, $y=y^k$ and $z=x^{k+1}$, it follows from (3.7) that

$$\begin{aligned}\big\langle\nabla G\big(x^{k+1}\big)-\nabla G\big(x^k\big),x^*-x^{k+1}\big\rangle\geq{}&\rho f\big(x^k,x^{k+1}\big)-\rho f\big(x^k,y^k\big)\\&-\rho c_1\big\|y^k-x^k\big\|^2-\rho c_2\big\|x^{k+1}-y^k\big\|^2.\end{aligned}\tag{3.8}$$

On the other hand, by Step 1, as y^k is the solution to the convex program

$$\min_{y\in K}\{\rho f(x^k,y)+G(y)-G(x^k)-\langle\nabla G(x^k),y-x^k\rangle\},$$

we have

$$0\in\partial_2\{\rho f\big(x^k,y^k\big)+G\big(y^k\big)-G\big(x^k\big)-\big\langle\nabla G\big(x^k\big),y^k-x^k\big\rangle\}+N_K\big(y^k\big).$$

Similarly, we can show that

$$\rho f(x^k, y) - \rho f(x^k, y^k) \geq \langle \nabla G(y^k) - \nabla G(x^k), y^k - y \rangle \quad \forall y \in K$$

With $y = x^{k+1}$ we obtain

$$\rho f(x^k, x^{k+1}) - \rho f(x^k, y^k) \geq \langle \nabla G(y^k) - \nabla G(x^k), y^k - x^{k+1} \rangle. \tag{3.9}$$

It follows from (3.6), (3.8), and (3.9) that

$$\begin{aligned} l(x^k) - l(x^{k+1}) &\geq G(x^{k+1}) - G(x^k) - \langle \nabla G(x^k), x^{k+1} - x^k \rangle \\ &\quad + \langle \nabla G(y^k) - \nabla G(x^k), y^k - x^{k+1} \rangle \\ &\quad - \rho c_1 \|y^k - x^k\|^2 - \rho c_2 \|x^{k+1} - y^k\|^2 \\ &= G(x^{k+1}) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \\ &\quad + \langle \nabla G(y^k), y^k - x^{k+1} \rangle \\ &\quad - \rho c_1 \|y^k - x^k\|^2 - \rho c_2 \|x^{k+1} - y^k\|^2 \\ &= \left[G(x^{k+1}) - G(y^k) - \langle \nabla G(y^k), x^{k+1} - y^k \rangle \right] \\ &\quad + \left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \right] \\ &\quad - \rho c_1 \|y^k - x^k\|^2 - \rho c_2 \|x^{k+1} - y^k\|^2. \end{aligned} \tag{3.10}$$

Since G is strongly convex with modulus $\beta > 0$, for every x and y, one has

$$G(y) - G(x) - \langle \nabla G(x), y - x \rangle \geq \frac{\beta}{2} \|y - x\|^2 \quad \forall x, y \in K. \tag{3.11}$$

Applying (3.11) first with x^{k+1}, y^k and then with y^k, x^k we obtain from (3.10) that

$$l(x^k) - l(x^{k+1}) \geq \left(\frac{\beta}{2} - \rho c_1 \right) \|y^k - x^k\|^2 + \left(\frac{\beta}{2} - \rho c_2 \right) \|x^{k+1} - y^k\|^2 \quad \forall k \geq 0 \tag{3.12}$$

which proves (a).

Now we prove (b). By Assumption $0 < \rho < \min\{(\beta/2c_1), (\beta/2c_2)\}$, we have

$$\frac{\beta}{2} - \rho c_1 > 0 \quad \text{and} \quad \frac{\beta}{2} - \rho c_2 > 0.$$

Thus, from the inequality (3.12), we deduce that

$$l(x^k) - l(x^{k+1}) \geq \left(\frac{\beta}{2} - \rho c_1 \right) \|y^k - x^k\|^2 \geq 0 \quad \forall k. \tag{3.13}$$

Thus $\{l(x^k)_{k\geq 0}\}$ is a nonincreasing sequence. Since it is bounded below by 0, it converges to l^*. Passing to the limit as $k \to \infty$ it is easy to see from (3.13) that

$$\lim_{k\to\infty} \|y^k - x^k\| = 0. \tag{3.14}$$

Note that, since G is β-strongly convex, by the definition of $l(x^k)$, we can write

$$0 \leq \frac{\beta}{2}\|x^* - x^k\|^2 \leq l(x^k) \quad \forall k.$$

Thus, since $\{l(x^k)\}$ is convergent, we can deduce that the sequence $\{x^k\}_{k\geq 0}$ is bounded, so it has at least one cluster point. Let $\overline{x} \in K$ be any cluster point and $\{x^{k_i}{}_{i\geq 0}\}$ be the subsequence such that

$$\lim_{i\to\infty} x^{k_i} = \overline{x}.$$

Then, it follows from (3.14) that

$$\lim_{i\to\infty} y^{k_i} = \overline{x}.$$

Again by Step 1 of the algorithm, we have

$$\begin{aligned} &\rho f(x^{k_i}, y) + G(y) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y - x^{k_i} \rangle \\ &\geq \rho f(x^{k_i}, y^{k_i}) + G(y^{k_i}) - G(x^{k_i}) - \langle \nabla G(x^{k_i}), y^{k_i} - x^{k_i} \rangle \quad \forall y \in K. \end{aligned}$$

Since f is lower semicontinuous on $K \times K$, $f(\cdot, y)$ is upper semicontinuous on K and $f(\overline{x}, \overline{x}) = 0$, letting $i \to \infty$ we obtain from the last inequality that

$$\rho f(\overline{x}, y) + G(y) - G(\overline{x}) - \langle \nabla G(\overline{x}), y - \overline{x} \rangle \geq 0 \quad \forall y \in K,$$

which shows that $\overline{x}$ is a solution of the (AuPEP) corresponding to $L(x, y) = G(y) - G(x) - \langle \nabla G(x), y - x \rangle$. Then, by Lemma 2.2, $\overline{x}$ is a solution to (PEP).

Suppose now $K^d = K^*$. We claim that the whole sequence $\{x^k\}_{k\geq 0}$ converges to $\overline{x}$. Indeed, using the definition of $l(x\hat{}k)$ with $x^* = \overline{x} \in K^d$, we have $l(\overline{x}) = 0$. Thus, as G is β-strongly convex, we can write

$$l(x^k) - l(\overline{x}) = G(\overline{x}) - G(x^k) - \langle \nabla G(x^k), x^k - \overline{x} \rangle \geq \frac{\beta}{2}\|x^k - \overline{x}\|^2 \quad \forall k \geq 0. \tag{3.15}$$

On the other hand, since the $\{l(x^k)\}_{k\geq 0}$ is nonincreasing and as $l(x^{k_i}) \to l(\overline{x})$, we must have $l(x^k) \to l(\overline{x})$ when $k \to \infty$. Thus, by (3.15), $\lim_{k\to\infty} x^k = \overline{x} \in K^*$. ■

Remark 3.1 The condition (3.4) does not necessarily imply that f is continuous. In fact, if $f(x, y) := \varphi(y) - \varphi(x)$, then clearly that (3.4) holds true for any $c_1 \geq 0$, $c_2 \geq 0$ and for any function φ.

4. Linesearch algorithms

Algorithm 1 requires that f satisfies the Lipschitz-type condition (3.4) which in some cases is not known. In order to avoid this requirement, in this section we modify Algorithm 1 by using a linesearch. The linesearch technique has been used widely in descent methods for mathematical programing problems as well as for variational inequalities [9,13].

First, we begin with the following definition.

Definition 4.1 ([13]) Let K be a nonempty closed set in $\mathbb{R}^n$. A mapping $P : \mathbb{R}^n \to \mathbb{R}^n$ is said to be

(i) *feasible with respect* to K if

$$P(x) \in K \quad \forall x \in \mathbb{R}^n,$$

(ii) *quasi-nonexpansive with respect* to K if for every $x \in \mathbb{R}^n$,

$$\left\| P(x) - y \right\| \le \|x - y\| \quad \forall y \in K. \tag{4.1}$$

Note that, if Π_K is the Euclidean projection on K, then Π_K is a feasible quasi-nonexpansive mappings. We denote by $\mathcal{F}(K)$ the class of feasible quasi-nonexpansive mappings with respect to K.

Next, we choose a sequence $\{\gamma_k\}_{k\ge 0}$ such that

$$\gamma_k \in (0,2) \quad \forall k = 0,1,2,\ldots \quad \text{and} \quad \liminf_{k\to\infty} \gamma_k(2-\gamma_k) > 0. \tag{4.2}$$

The algorithm then can be described as follows.

Algorithm 2

Data $x^0 \in K, \alpha \in (0,1), \theta \in (0,1)$ and $\rho > 0$.

Step 0 Set $k := 0$.

Step 1 Solve the following strongly convex optimization problem

$$\min_{y\in K}\left\{ f(x^k, y) + \frac{1}{\rho}\left[G(y) - \left\langle \nabla G(x^k), y - x^k \right\rangle \right] \right\} \tag{4.3}$$

to obtain its unique solution y^k.

If $y^k = x^k$, stop: x^k is a solution to (PEP). Otherwise, go to Step 2.

Step 2

Step 2.1 Find the smallest positive integer m such that

$$\begin{cases} z^{k,m} = (1-\theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \left\langle \nabla G(x^k), y^k - x^k \right\rangle \right] \le 0. \end{cases} \tag{4.4}$$

Step 2.2 Set $\theta_k = \theta^m, z^k = z^{k,m}$. If $0 \in \partial_2 f(z^k, z^k)$, stop: z^k is a solution to (PEP). Otherwise, go to Step 3.

Step 3 Select $g^k \in \partial_2 f(z^k, z^k)$, and compute

$$\sigma_k = \frac{-\theta_k f(z^k, y^k)}{(1-\theta_k)\|g^k\|^2} \quad \text{and} \quad x^{k+1} = P_k(x^k - \gamma_k \sigma_k g^k), \tag{4.5}$$

where $P_k \in \mathcal{F}(K)$.

Step 4 Set $k := k+1$, and go back to Step 1.

The following lemma indicates that if Algorithm 2 terminates at Step 1 or Step 2.2, then indeed a solution of (PEP) has been found.

Lemma 4.1 *If Algorithm 2 terminates at Step* 1 (*resp. Step* 2.2), *then* x^k (*resp.* z^k) *is a solution to* (*PEP*).

Proof If the algorithm terminates at Step 1, then $x^k = y^k$. Since y^k is the solution to the convex optimization problem (4.3), we have

$$\begin{aligned} &f(x^k, y) + \frac{1}{\rho}\left[G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k\rangle\right] \\ &\quad \geq f(x^k, y^k) + \frac{1}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle \quad \forall y \in K. \end{aligned}$$

By the same argument as in the proof of Lemma 2.2, we can show that x^k is a solution to (PEP).

If the algorithm terminates at Step 2.2, then $0 \in \partial_2 f(z^k, z^k)$. Since $f(z^k, \cdot)$ is convex, it implies that $f(z^k, z^k) \leq f(z^k, y)$ for all $y \in K$. Alternatively, since $f(z^k, z^k) = 0$, it shows that z^k is a solution to (PEP). ■

The next lemma shows that there always exists a positive integer m such that Condition (4.4) in Step 2.1 is satisfied.

Lemma 4.2 *Suppose that f is upper semicontinuous on K with respect to the first variable, and $y^k \neq x^k$. Then*

(i) There exists an integer $m > 0$ such that the inequality in (4.4) holds.
(ii) $f(z^k, y^k) < 0$.

Proof To prove (i) we suppose by contradiction that for every positive integer m such that

$$z^{k,m} = (1-\theta^m)x^k + \theta^m y^k$$

we have

$$f(z^{k,m}, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] > 0.$$

Since $f(\cdot, y^k)$ is upper semicontinuous, passing to the limit $m \to \infty$ we have

$$f(x^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \geq 0. \tag{4.6}$$

On the other hand, since y^k is a solution to the convex optimization problem (4.3), we can write

$$f(x^k, y^k) + \frac{1}{\rho}\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big]$$
$$\leq f(x^k, y) + \frac{1}{\rho}\Big[G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k\rangle\Big] \quad \forall y \in K.$$

With $y = x^k$ the latter inequality becomes

$$f(x^k, y^k) + \frac{1}{\rho}\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big] \leq 0. \tag{4.7}$$

It follows from (4.6) and (4.7) that

$$\frac{1}{\rho}\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big]$$
$$\leq \frac{\alpha}{\rho}\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big].$$

Since

$$\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big] \geq 0,$$

we deduce that either $[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle] = 0$ or $\alpha \geq 1$. The first case implies that $x^k = y^k$, since G is strongly convex. Hence, both cases contradict the assumption. So (i) holds true.

The statement (ii) is immediate from the rule for determination of z^k as

$$f(z^k, y^k) + \frac{\alpha}{\rho}\Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\Big] \leq 0$$

and

$$G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle \geq \frac{\beta}{2}\|x^k - y^k\|^2 > 0,$$

G being strongly convex and $x^k \neq y^k$. ■

In order to prove the convergence of Algorithm 2, we give the following key property of the sequence $\{x^k\}_{k\geq 0}$ generated by the algorithm.

LEMMA 4.3 *If $f(x, \cdot)$ is convex and subdifferentiable on K, then the following statements hold true*:

(i) *For every solution x^* of (DEP) one has*

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k(2 - \gamma_k)(\sigma_k\|g^k\|)^2. \tag{4.8}$$

(ii) $\sum_{k=0}^{\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < \infty$.
(iii) *Suppose that the algorithm does not terminate. Then if in addition f is continuous with respect to the second argument and finite on an open set containing K, then the sequence $\{g^k\}$ is bounded.*

Proof First, we prove (i). Take any $x^* \in K^d$. By property (4.1) of P_k and (4.5), setting $w^k = z^k - \gamma_k\sigma_k g^k$, as $x^{k+1} = P_k(w^k)$ we have

$$\begin{aligned}\|x^{k+1}-x^*\|^2 &= \|P_k(w^k)-x^*\|^2\\ &\le \|w^k - x^*\|^2\\ &= \|x^k - \gamma_k\sigma_k g^k - x^*\|^2\\ &= \|x^k-x^*\|^2 - 2\gamma_k\sigma_k\langle g^k, x^k-x^*\rangle + \left(\gamma_k\sigma_k\|g^k\|\right)^2. \end{aligned} \tag{4.9}$$

Since $g^k \in \partial_2 f(z^k, z^k)$ and $f(z^k,\cdot)$ is convex on K, we have

$$\begin{aligned}\langle g^k, x^k - x^*\rangle &= \langle g^k, x^k - z^k + z^k - x^*\rangle\\ &\ge \langle g^k, x^k - z^k\rangle + f(z^k, z^k) - f(z^k, x^*),\end{aligned}$$

where $f(z^k, x^*) \le 0$ because $x^* \in K^d$. Thus, it follows from the latter inequality that

$$\langle g^k, x^k - x^*\rangle \ge \langle g^k, x^k - z^k\rangle. \tag{4.10}$$

Using (4.4) we can write

$$x^k - z^k = \frac{\theta_k}{1-\theta_k}\left(z^k - y^k\right).$$

Thus

$$\begin{aligned}\langle g^k, x^k - z^k\rangle &= \frac{\theta_k}{1-\theta_k}\langle g^k, z^k - y^k\rangle\\ &\ge \frac{\theta_k}{1-\theta_k}\left[f(z^k,z^k) - f(z^k,y^k)\right]\\ &= \frac{-\theta_k}{(1-\theta_k)} f(z^k, y^k).\end{aligned}$$

From (ii) of Lemma 4.2 and (4.5) it follows that

$$\frac{-\theta_k}{(1-\theta_k)} f(z^k, y^k) = \sigma_k\|g^k\|^2 > 0 \tag{4.11}$$

which, together with (4.9) and (4.10), implies

$$\|x^{k+1}-x^*\|^2 \le \|x^k - x^*\|^2 - \gamma_k(2-\gamma_k)\left(\sigma_k\|g^k\|\right)^2 \quad \forall x^* \in K^d.$$

To prove (ii) we apply the latter inequality for every k from 0 to m to obtain

$$\sum_{k=0}^{m} \gamma_k(2-\gamma_k)\big(\sigma_k\|g^k\|\big)^2 \le \|x^0 - x^*\|^2 - \|x^{m+1} - x^*\|^2.$$

Since $\{\|x^m - x^*\|\}_{m\geq 0}$ is convergent, taking the limit $m \to \infty$ we obtain

$$\sum_{k=0}^{\infty} \gamma_k(2-\gamma_k)(\sigma_k\|g^k\|)^2 < \infty.$$

We finally prove (iii). To prove that $\{g^k\}$ is bounded we first observe that $\{y^k\}$ is bounded. Indeed, since y^k is the unique solution of Problem (4.3) whose objective function is continuous and the feasible set is constant, by the Maximum Theorem (Proposition 23 in [4], see also [5]), the mapping $x^k \to s(x^k) = y^k$ is continuous. Since $\{x^k\}$ is bounded, $\{y^k\}$ is bounded, and therefore, $\{z^k\}$ is bounded too because z^k is a convex combination of x^k and y^k. Thus, without loss of generality we may assume that $z^k \to z^*$ as $k \to +\infty$. By continuity of the convex function $f(z^k, \cdot)$, the sequence $\{f(z^k, \cdot)\}$ converges pointwise to $f(z^*, \cdot)$. Since $g^k \in \partial_2 f(z^k, z^k)$, we can deduce, from Theorem 24.5 in [27], that $\{g^k\}$ is bounded. ∎

We are now in a position to prove the following convergence theorem for Algorithm 2. As we have seen in Lemma 4.1 that if Algorithm 2 terminates then a solution to (PEP) has already been found. Otherwise, if the algorithm does not terminate, we have the following convergence results.

THEOREM 4.4 *In addition to the assumptions of Lemmas* 4.2 *and* 4.3 *we assume that f is continuous on $K \times K$. Then*

(i) *The sequence $\{x^k\}$ is bounded, and every cluster point of $\{x^k\}$ is a solution to (PEP).*
(ii) *If $K^* = K^d$ (in particular, if f is pseudomonotone on K), then the whole sequence $\{x^k\}$ converges to a solution of (PEP). In addition, if $\gamma_k = \gamma \in (0, 2)$ for all $k \geq 0$, then*

$$\liminf_{k\to\infty}(\sigma_k\|g^k\|\sqrt{k+1}) = 0 \tag{4.12}$$

holds true.

Proof The boundedness of $\{x^k\}$ follows immediately from (i) and (ii) of Lemma 4.3. Again by (ii) of Lemma 4.3 we have

$$\gamma_k(2-\gamma_k)\big(\sigma_k\|g^k\|\big)^2 \to 0 \quad \text{as } k \to \infty.$$

By (4.2), $\liminf_{k\to\infty} \gamma_k(2-\gamma_k) > 0$. Thus $\sigma_k\|g^k\| \to 0$ as $k \to \infty$. Then

$$\sigma_k\|g^k\| = \frac{-\theta_k}{(1-\theta_k)\|g^k\|} f(z^k, y^k) \to 0.$$

Since $\{g^k\}$ is bounded by (iii) of Lemma 4.3, we deduce that

$$\frac{-\theta_k}{1-\theta_k} f\left(z^k, y^k\right) \to 0 \quad \text{as } k \to \infty. \tag{4.13}$$

On the other hand, since G is β-strongly convex, according to the rule (4.4) we have

$$\begin{aligned}\frac{\alpha\beta}{\rho}\left\|x^k - y^k\right\|^2 &\leq \frac{\alpha}{\rho}\left[G\left(y^k\right) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y^k - x^k\right\rangle\right] \\ &\leq -f(z^k, y^k).\end{aligned} \tag{4.14}$$

We will consider two cases:

Case 1 $\limsup_{k\to\infty}\theta_k > 0$. Then there exists $\overline{\theta} > 0$ and a subsequence $N^* \subseteq \mathbb{N}$ such that $\theta_k \geq \overline{\theta}$ for every $k \in N^*$. From (4.13) and (4.14), we deduce that

$$\lim_{k\to\infty,\ k\in N^*}\left\|y^k - x^k\right\| = 0. \tag{4.15}$$

Let x^* be any cluster point of $\{x^k\}$. Suppose that the sequence $\left\{x^k : k \in N^*\right\}$ converges to x^*. Using (4.15) we see that the corresponding subsequence $\left\{y^k : k \in N^*\right\}$ also converges to $y^* = x^*$. Hence, from Step 1 of Algorithm 2, since y^k is the solution of problem (4.2), we have

$$\begin{aligned}&\rho f\left(x^k, y^k\right) + \left[G\left(y^k\right) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y^k - x^k\right\rangle\right] \\ &\quad\leq \rho f\left(x^k, y\right) + \left[G(y) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), y - x^k\right\rangle\right] \quad \forall y \in K.\end{aligned}$$

Letting $k \to +\infty$, $k \in N^*$, by continuity of f and $x^* = y^*$, we obtain

$$\begin{aligned}0 &= \rho f(x^*, y^*) + \left[G(y^*) - G(x^*) - \left\langle \nabla G(x^*), y^* - x^*\right\rangle\right] \\ &\leq \rho f(x^*, y) + \left[G(y) - G(x^*) - \left\langle \nabla G(x^*), y - x^*\right\rangle\right] \quad \forall y \in K.\end{aligned}$$

Thus, in virtue of Lemma 2.2, x^* is a solution to (PEP).

Case 2 $\lim_{k\to\infty}\theta_k = 0$. According to the algorithm we have

$$z^k = (1-\theta_k)x^k + \theta_k y^k.$$

As before, we may suppose that the subsequence $\{x^k : k \in N^* \subseteq \mathbb{N}\}$ converges to some point x^*. Since y^k is the solution of Problem (4.3), it follows from the lower semicontinuity of the objective function $\rho f(x^k, \cdot) + G(\cdot) - G\left(x^k\right) - \left\langle \nabla G\left(x^k\right), \cdot - x^k\right\rangle$ and the Maximum Theorem ([4] Proposition 19) that the sequence $\{y^k\}$ is bounded.

Thus, by taking a subsequence, if necessary, we may assume that the subsequence $\{y^k : k \in N^*\}$ converges to some point y^*. By the definition of y^k we have

$$\rho f(x^k, y^k) + \left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right]$$
$$\leq \rho f(x^k, y) + \left[G(y) - G(x^k) - \langle \nabla G(x^k), y - x^k\rangle\right] \quad \forall y \in K.$$

Taking the limit as $k \to \infty$, $k \in N^*$, since f is lower semicontinuous on $K \times K$ and $f(\cdot, y)$ is upper semicontinuous on K, we have

$$\rho f(x^*, y^*) + \left[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^*\rangle\right]$$
$$\leq \rho f(x^*, y) + \left[G(y) - G(x^*) - \langle \nabla G(x^*), y - x^*\rangle\right] \quad \forall y \in K. \tag{4.16}$$

On the other hand, by Step 2.2 of Algorithm 2, m is the smallest nonnegative integer satisfying (4.4), so we have

$$\rho f(z^{k,m-1}, y^k) + \alpha\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] > 0.$$

By taking subsequences, if necessary, we may assume that $\theta_k \to 0$. Then $z^{k,m-1} \to x^*$. By continuity of f at (x^*, y^*), we obtain in the limit $k \to \infty$ that

$$\rho f(x^*, y^*) + \alpha\left[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^*\rangle\right] \geq 0. \tag{4.17}$$

Substituting $y = x^*$ into (4.16) we get

$$\rho f(x^*, y^*) + \left[G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^*\rangle\right] \leq 0. \tag{4.18}$$

Since

$$G(y^*) - G(x^*) - \langle \nabla G(x^*), y^* - x^*\rangle \geq \frac{\beta}{2}\|x^* - y^*\|^2,$$

taking into account (4.17) and (4.18) we deduce that

$$(1 - \alpha)\|y^* - x^*\|^2 \leq 0$$

which together with $\alpha \in (0, 1)$ implies $x^* = y^*$. Then it follows from (4.16) that x^* is an optimal solution to the optimization problem

$$\min_{y \in K}\left\{\rho f(x^*, y) + \left[G(y) - G(x^*) - \langle \nabla G(x^*), y - x^*\rangle\right]\right\}.$$

Thus, again by Lemma 2.2, x^* is a solution to (PEP).

Now we suppose that $K^d = K^*$. From the discussion above, we know that the sequence $\{x^k\}$ has a cluster point $x^* \in K^*$. Since $K^d \equiv K^*$, $x^* \in K^d$. Applying (i) of Lemma 4.3 we see that the whole sequence $\{x^k - x^*\}$ is convergent. Hence, the whole sequence $\{x^k\}$ must converge to x^* because it has a subsequence converging to x^*.

Finally, we suppose that (4.12) does not hold. Then there exists a number $\mu > 0$ such that $\sigma_k \|g^k\| \geq (\mu/\sqrt{k+1})$ for all k. From (ii) of Lemma 4.3 we have

$$\mu^2 \sum_{k=0}^{\infty} \frac{1}{k+1} < \infty$$

which is a contradiction. Thus, the theorem is completed. ■

Remark 4.1 In practice, to implement the algorithm we take a tolerance $\epsilon > 0$ and we terminate the algorithm when either $\|x^k - y^k\| \leq \epsilon$ or $\|g^k\| \leq \epsilon$.

Remark 4.2 If f is convex with respect to the first variable x on K, then the rule (4.4) to determine z^k holds true automatically for any θ_k satisfying $0 < \theta_k \leq 1 - \alpha$ for all k.

Indeed, since $f(y^k, y^k) = 0$, by convexity of f with respect to the first variable, we can write

$$\begin{aligned} & f(z^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \\ & \quad = f((1-\theta_k)x^k + \theta_k y^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \\ & \quad \leq (1-\theta_k) f(x^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right]. \end{aligned}$$

On the other hand, since y^k is a solution to (4.3), we have

$$f(x^k, y^k) + \frac{1}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \leq 0.$$

Thus

$$\begin{aligned} & (1-\theta_k) f(z^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \\ & \quad \leq \frac{(\alpha - 1 + \theta_k)}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \\ & \quad \leq 0 \end{aligned}$$

which shows that the linesearch condition (4.4) is satisfied provided $0 < \theta_k \leq 1 - \alpha$. ■

By Remark 4.2, when $f(\cdot, y)$ is convex on K (for example, when f is a saddle function on $K \times K$), in order to perform Step 2 in Algorithm 2, we can simply take $z^k = (1-\theta_k)x^k + \theta_k y^k$ where θ_k can be any number between 0 and $1 - \alpha$. For gereral case, the linesearch at Step 2.1 of Algorithm 2 sometimes leads θ_k to 0 that may cause zigzag. To avoid this case, we propose another linesearch to obtain the following algorithm.

Algorithm 2a

Data $x^0 \in K, \alpha \in (0,1), \theta \in (0,1)$ and $\rho > 0$.

Step 0 Set $k = 0$.

Step 1 Find $y^k \in K$ as the unique solution to the strongly convex program

$$\min_{y \in K} \left\{ f(x^k, y) + \frac{1}{\rho} \left[G(y) - \langle \nabla G(x^k), y - x^k \rangle \right] \right\}. \tag{4.19}$$

If $y^k = x^k$, stop: x^k is a solution to (PEP). Otherwise, go to Step 2.

Step 2

Step 2.1 Find m as the smallest nonnegative integer such that

$$\begin{cases} z^{k,m} = (1-\theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, x^k) - f(z^{k,m}, y^k) \geq \dfrac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \right]. \end{cases} \tag{4.20}$$

Step 2.2 Set $\theta_k = \theta^m, z^k = z^{k,m}$ and go to Step 3.

Step 3 Select any $g^k \in \partial_2 f(z^k, x^k)$, and compute

$$\sigma_k = \frac{f(z^k, x^k)}{\|g^k\|^2} \quad \text{and} \quad x^{k+1} = P_k(x^k - \gamma_k \sigma_k g^k) \tag{4.21}$$

where $P_k \in \mathcal{F}(K)$.

Step 4 Set $k := k+1$, and go back to Step 1.

As for Algorithm 2, we have the following lemma for Algorithm 2a.

LEMMA 4.5 *Suppose that f is continuous on K with respect to the first variable, and $y^k \neq x^k$. Then*

(i) *There exists an integer $m \geq 0$ such that the inequality in* (4.20) *holds.*
(ii) $f(z^k, x^k) > 0$.
(iii) $0 \notin \partial_2 f(z^k, x^k)$.

Proof To prove (i) we suppose by contradiction that for every nonnegative integer m, we have

$$\begin{cases} z^{k,m} = (1-\theta^m)x^k + \theta^m y^k, \\ f(z^{k,m}, x^k) - f(z^{k,m}, y^k) < \dfrac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \right]. \end{cases}$$

Passing to the limit $m \to \infty$, since $f(\cdot, y^k)$ is continuous, $\theta^m \to 0$ and $f(x^k, x^k) = 0$, we have

$$0 \leq f(x^k, y^k) + \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \right].$$

The remainder of part (i) can be proved similarly as in the proof of Lemma 4.2.

We now prove (ii). By part (i), there is a nonnegative integer m such that (4.20) holds. Since f is convex with respect to the second argument, we have

$$\theta_k f(z^k, y^k) + (1-\theta_k) f(z^k, x^k) \geq f(z^k, z^k) = 0,$$

or

$$f(z^k, x^k) \geq \theta_k \Big[f(z^k, x^k) - f(z^k, y^k) \Big].$$

Using the inequality in (4.20), and the strong convexity of G, we obtain from the latter inequality that

$$f(z^k, x^k) \geq \frac{\alpha \theta_k}{\rho} \Big[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k \rangle \Big] > 0,$$

which proves (ii).

To prove (iii) we suppose by contradiction that $0 \in \partial f_2(z^k, x^k)$. Since $f(z^k, \cdot)$ is convex on K, the inclusion $0 \in \partial f_2(z^k, x^k)$ implies that

$$f(z^k, x^k) \leq f(z^k, y) \quad \forall y \in K.$$

Substituting $y = z^k \in K$ into the above inequality we obtain $f(z^k, x^k) \leq 0$ which contradicts (ii). So the proof of the lemma is complete. ■

Lemma 4.6 *Lemma 4.3 remains true to Algorithm* 2a.

Proof Take any $x^* \in K^d$. By using the same argument as in the proof of Lemma 4.3, we have

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - 2\gamma_k \sigma_k \langle g^k, x^k - x^* \rangle + \left(\gamma_k \sigma_k \|g^k\|\right)^2. \tag{4.22}$$

Since $g^k \in \partial_2 f(z^k, x^k)$, and $f(z^k, \cdot)$ is convex on K, we have

$$\langle g^k, x^k - x^* \rangle \geq f(z^k, x^k) - f(z^k, x^*).$$

On the other hand, since $x^* \in K^d$, we have $f(z^k, x^*) \leq 0$. Thus

$$\langle g^k, x^k - x^* \rangle \geq f(z^k, x^k) = \sigma_k \|g^k\|^2 > 0. \tag{4.23}$$

Using (4.22) and (4.23) we obtain

$$\|x^{k+1} - x^*\|^2 \leq \|x^k - x^*\|^2 - \gamma_k (2 - \gamma_k) \left(\sigma_k \|g^k\|\right)^2 \quad \forall x^* \in K^d.$$

The remainder of the proof can be done by the same way as in the proof of Lemma 4.3. ■

Using the same argument as in the proof of Theorem 4.4 we obtain the following convergence result for Algorithm 2a.

THEOREM 4.7 *The conclusions of Theorem* 4.4 *remain true for Algorithm* 2*a*.

5. Application to mixed (multivalued) variational inequalities

In this section, we discuss about applications of the proposed algorithms to the following variational inequality

$$\text{Find } \ x^* \in K,\ v^* \in F(x^*) \text{ such that} \\ \langle v^*, x - x^* \rangle + \varphi(x) - \varphi(x^*) \geq 0 \tag{MVIP}$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, and $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a closed proper convex function. We suppose that $F(x)$ is a nonempty compact set for each $x \in K$, and $K \subseteq \mathrm{dom}\varphi$ where $\mathrm{dom}\varphi$ denotes the effective domains of φ.

For each pair $x, y \in K$ we put

$$f(x, y) := \max_{u \in F(x)} \left\{ \left\langle u, y - x \right\rangle + \varphi(y) - \varphi(x) \right\}. \tag{5.1}$$

We can easily check that x^* is a solution to (MVIP) if and only if it is a solution to (PEP).

We will need the following definitions:

(1) The mapping F is said to be *φ-pseudomonotone* on K if for all $x, y \in K$ and all $u \in F(x)$, $v \in F(y)$ the inequality

$$\left\langle u, y - x \right\rangle + \varphi(y) - \varphi(x) \geq 0$$

implies

$$\left\langle v, y - x \right\rangle + \varphi(y) - \varphi(x) \geq 0.$$

(2) F is said to be *Lipschitz continuous* on K with constant L if for all $x, y \in K$ one has

$$\sup_{u \in F(x)} \inf_{v \in F(y)} \|u - v\| \leq L\|x - y\|.$$

If $h(A, B)$ denotes the Hausdorff distance between two sets A and B, then this definition means that

$$h\big(F(x), F(y)\big) \leq L\|x - y\| \quad \forall x, y \in K.$$

From the previous section, we know that Condition (3.4) does not imply the continuity of f. Conversely, however, if f is given by (5.1), F is L-Lipschitz continuous, and φ is continuous on K, then f satisfies Condition (3.4) as stated in the next lemma. This explains why we call (3.4) a Lipschitz-type condition.

LEMMA 5.1 *Let f be defined by* (5.1). *The following statements hold:*

(i) *If F, φ are continuous on K and $F(x)$ is compact for every $x \in K$, then f is continuous on $K \times K$.*
(ii) *If F is φ-pseudomonotone on K, then f is pseudomonotone on K.*
(iii) *If F is L-Lipschitz continuous on K then, for any $\mu > 0$,*

$$f(x, y) + f(y, z) \geq f(x, z) - \frac{L\mu}{2}\|x - y\|^2 - \frac{L}{2\mu}\|y - z\|^2 \quad \forall x, y, z \in K. \tag{5.2}$$

Proof The first statement follows from the Maximum Theorem (Proposition 23 in [4]).

The second statement is immediate from the definition.

To prove (iii) we suppose that F is L-Lipschitz continuous on K. Let $x, y, z \in K$. For any $u \in F(x)$ and $\epsilon > 0$, since F is L-Lipschitz continuous, by definition, there exists $v \in F(y)$ such that $\|u - v\| \leq L\|x - y\| + \epsilon$. Thus

$$\begin{aligned}
&\langle u, z - x\rangle - \sup_{v \in F(y)} \langle v, z - y\rangle - \sup_{w \in F(x)} \langle w, y - x\rangle \\
&\leq \langle u, z - x\rangle - \langle v, z - y\rangle - \langle u, y - x\rangle \\
&= \langle u, z - y\rangle - \langle v, z - y\rangle = \langle u - v, z - y\rangle \\
&\leq \|u - v\|\|z - y\| \leq \Big(L\|x - y\| + \epsilon\Big)\|z - y\|.
\end{aligned}$$

Since $\epsilon > 0$ and $u \in F(x)$ are arbitrary, we deduce that

$$f(x, z) - f(y, z) - f(x, y) \leq L\|x - y\|\|z - y\|.$$

Using the well-known inequality $2ab \leq (a^2/\mu) + \mu b^2$ that holds true for all $a, b \in \mathbb{R}$ and $\mu > 0$, we obtain (5.2). ■

Note that when F is Lipschitz continuous and singlevalued on K, Algorithm 1 with $f(x, y) := \langle F(x), y - x\rangle$, $\varphi \equiv 0$ becomes the well-known extragradient algorithm for variational inequalities (see e.g. [9] Chapter 12). When F is continuous, but not necessarily Lipschitz, Algorithm 2 and 2a coincide with the extragradient linesearch (hyperplane projection) algorithms (see e.g. [9] Chapter 12), but with a minor difference in the way of determining the linesearch and the stepsize. When F is multivalued, by Lemma 5.1, theoretically, Algorithm 1 as well as Algorithms 2 and 2a can be applied. Practically, to implement the algorithms, it is difficult to choose an approximation of $f(x, y)$. This issue remains, as far as the authors know, open.

6. Examples and numerical results

In this section, we illustrate the proposed algorithms by a class of equilibrium problems as defined by (PEP), where K is a polyhedral convex set given by

$$K := \{x \in \mathbb{R}^n \mid Ax \leq b\}, \tag{6.1}$$

and the equilibrium bifunction $f : K \times K \to \mathbb{R} \cup \{+\infty\}$ is of the form

$$f(x, y) = \langle F(x) + Qy + q, y - x\rangle, \tag{6.2}$$

with $F : K \to \mathbb{R}^n$, $Q \in \mathbb{R}^{n\times n}$ being a symmetric positive semidefinite matrix and $q \in \mathbb{R}^n$. Since Q is symmetric positive semidefinite, $f(x, \cdot)$ is convex for each fixed $x \in K$. The bifunction defined by (6.2) is a generalized form of the bifunction defined by the Cournot–Nash equilibrium model considered in [26].

For this class of equilibrium problems we have the following results.

LEMMA 6.1 *If $F : K \to \mathbb{R}^n$ is τ-strongly monotone on K. Then*

(i) *f is monotone on K whenever $\tau = \|Q\|$.*
(ii) *f is $\tau - \|Q\|$-strongly monotone on K whenever $\tau > \|Q\|$.*

Proof From the definition of f we have

$$f(x, y) + f(y, x) = \langle Q(y - x), y - x\rangle - \langle F(y) - F(x), y - x\rangle. \tag{6.3}$$

Note that

$$\langle Q(y - x), y - x\rangle \le \|Q\|\,\|y - x\|^2. \tag{6.4}$$

Since F is τ-strongly monotone on K, that is

$$\langle F(y) - F(x), y - x\rangle \ge \tau\|y - x\|^2,$$

we have, from (6.3) and (6.4) that $f(x, y) + f(y, x) \le 0$ whenever $\tau = \|Q\|$ and

$$f(x, y) + f(y, x) \le -(\tau - \|Q\|)\|y - x\|^2,$$

whenever $\tau > \|Q\|$. ■

LEMMA 6.2 *If F is L-Lipschitz continuous on K, i.e.,*

$$\|F(y) - F(x)\| \le L\|y - x\| \quad \forall x, y \in K,$$

then f satisfies the Lipschitz-type condition (3.4). *Namely,*

$$f(x, y) + f(y, z) \ge f(x, z) - c_1\|y - x\|^2 - c_2\|z - y\| \quad \forall x, y, z \in K,$$

for any $c_1 > 0, c_2 > 0$ satisfying

$$2\sqrt{c_1c_2} \ge L + \|Q\|.$$

Proof For every $x, y, z \in K$ we have

$$f(x, y) + f(y, z) - f(x, z) = \langle F(y) - F(x), z - y\rangle + \langle Q(y - z), y - x\rangle,$$

Applying the Cauchy–Schwartz inequality, we have

$$\langle Q(y - z), y - x\rangle \ge -\|Q\|\,\|z - y\|\,\|y - x\|$$

and

$$\langle F(y) - F(x), z - y\rangle \geq -\|F(y) - F(x)\|\,\|z - y\|.$$

Since F is L-Lipschitz, we can write

$$-\|F(y) - F(x)\|\,\|z - y\| \geq -L\|y - x\|\,\|z - y\|.$$

Thus, it follows from the last three inequalities that

$$f(x,y) + f(y,z) - f(x,z) \geq -(L + \|Q\|)\|y - x\|\,\|z - y\|.$$

Then, by hypothesis we have

$$-(L + \|Q\|)\|y - x\|\,\|z - y\| \geq -c_1\|y - x\|^2 - c_2\|z - y\|^2.$$

Hence

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|y - x\|^2 - c_2\|z - y\|^2. \qquad \blacksquare$$

For the special case when F is a linear mapping of the form $F(x) = Px$ with $P \in \mathbb{R}^{n\times n}$, the function f defined by (6.2) takes the form

$$f(x,y) = \langle Px + Qy + q, y - x\rangle. \tag{6.5}$$

We suppose that the matrices P, Q are chosen such that Q is symmetric positive semidefinite and $Q - P$ is negative semidefinite. Then f has the following properties:

(i) *f is monotone, $f(\cdot, y)$ is continuous and $f(x, \cdot)$ is differentiable convex on K.*
(ii) *For every $x, y, z \in K$ one has*

$$f(x,y) + f(y,z) \geq f(x,z) - c_1\|z - y\|^2 - c_2\|y - x\|^2, \tag{6.6}$$

where $c_1 = c_2 = \frac{1}{2}\|P - Q\|$.

Indeed, for every $x, y \in K$, since $Q - P$ is negative semidefinite, we have

$$f(x,y) + f(y,x) = \langle (Q - P)(y - x), y - x\rangle \leq 0.$$

Hence, f is monotone on K. Clearly, $f(x, \cdot)$ is differentiable, and since Q is symmetric positive semidefinite, $f(x, \cdot)$ is convex for each fixed x.

To see (ii) we observe that, for every $x, y, z \in K$,

$$\begin{aligned}
& f(x,y) + f(y,z) - f(x,z) \\
&\quad = \langle Px + Qy + q, y - x\rangle + \langle Py + Qz + q, z - y\rangle - \langle Px + Qz + q, z - x\rangle \\
&\quad = \langle Px + Qy, y - x\rangle + \langle Py + Qz, z - y\rangle - \langle Px + Qz, z - x\rangle \\
&\qquad + \langle q, y - x + z - y - (z - x)\rangle \\
&\quad = \langle Px + Qy, y - x\rangle + \langle Py + Qz, z - y\rangle - \langle Px + Qz, z - x\rangle \\
&\quad = \langle Px, y - x - (z - x)\rangle + \langle Qz, z - y - (z - x)\rangle + \langle Qy, y - x\rangle + \langle Py, z - y\rangle \\
&\quad = \langle Px, y - z\rangle + \langle Qz, x - y\rangle + \langle Qy, y - x\rangle + \langle Py, z - y\rangle \\
&\quad = \langle P(y - x), z - y\rangle + \langle Q(z - y), x - y\rangle \\
&\quad = \langle P(y - x), z - y\rangle + \langle Q^T(x - y), z - y\rangle \\
&\quad = \langle P(y - x), z - y\rangle + \langle Q(x - y), z - y\rangle \quad \text{since } Q = Q^T \\
&\quad = \langle (P - Q)(y - x), z - y\rangle,
\end{aligned}$$

from which it follows that

$$\begin{aligned}
f(x,y) + f(y,z) - f(x,z) &= \langle (P - Q)(y - x), z - y\rangle \\
&\geq -2\frac{\|P - Q\|}{2}\|y - x\|\|z - y\| \\
&\geq -\frac{\|P - Q\|}{2}\|y - x\|^2 - \frac{\|P - Q\|}{2}\|z - y\|^2.
\end{aligned}$$

By setting, for example, $c_1 = c_2 = \frac{1}{2}\|Q - P\|$, we obtain (6.6).

Now we use Algorithm 1 with the quadratic regularization function $G(x) := \frac{1}{2}\|x\|^2$ to solve Problem (PEP) where $f(x, y)$ is given by (6.5), and K is defined as in (6.1). In this case, the subproblem needed to solve at Step 1 is of the from

$$\min_{y \in K}\left\{\rho f(x,y) + \frac{1}{2}\|y - x\|^2\right\}.$$

Since, by (6.5), $f(x, \cdot)$ is convex quadratic, this subproblem can then be solved efficiently, for example, by the MATLAB Optimization Toolbox.

To illustrate our algorithms, we will consider two academic numerical tests of small size subsequently.

Test 1 $n = 5$ and the matrices P and Q (randomly generated) are

$$Q = \begin{bmatrix} 1.6 & 1 & 0 & 0 & 0 \\ 1 & 1.6 & 0 & 0 & 0 \\ 0 & 0 & 1.5 & 1 & 0 \\ 0 & 0 & 1 & 1.5 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}; \qquad P = \begin{bmatrix} 3.1 & 2 & 0 & 0 & 0 \\ 2 & 3.6 & 0 & 0 & 0 \\ 0 & 0 & 3.5 & 2 & 0 \\ 0 & 0 & 2 & 3.3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}.$$

With $q = (1, -2, -1, 2, -1)^T$,

$$K = \left\{ x \in R^5 \mid \sum_{i=1}^{5} x_i \geq -1, -5 \leq x_i \leq 5, \ i = 1, \ldots, 5 \right\},$$

$c_1 = c_2 = \frac{1}{2}\|Q - P\| = 1.4525, \rho = \frac{1}{2}c_1 = 0.7262,\ \ x_0 = (1, 3, 1, 1, 2)^T$ and $\epsilon = 10^{-3}$ we obtained the following iterates

Iter(k)	x_1^k	x_2^k	x_3^k	x_4^k	x_5^k
0	1.00000	3.00000	1.00000	1.00000	2.00000
1	−0.34415	1.59236	0.68742	−0.15427	0.63458
2	−0.67195	1.10393	0.65016	−0.57872	0.30562
3	−0.73775	0.92351	0.66742	−0.74459	0.22567
4	−0.74236	0.85341	0.68785	−0.81261	0.20624
5	−0.73668	0.82486	0.70195	−0.84184	0.20152
6	−0.73168	0.81276	0.71030	−0.85493	0.20037
7	−0.72864	0.80747	0.71491	−0.86100	0.20009
8	−0.72700	0.80511	0.71737	−0.86389	0.20002
9	−0.72617	0.80403	0.71865	−0.86529	0.20001
10	−0.72576	0.80354	0.71931	−0.86598	0.20000

The approximate solution obtained after 10 iterations is

$$x^{10} = (-0.72576, 0.80354, 0.71931, -0.86598, 0.20000)^T.$$

If we choose

$$P = \begin{bmatrix} 3.1 & 2.0 & 0.0 & 0.0 & 0.0 \\ 2.0 & 3.6 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 3.5 & 2.0 & 0.0 \\ 0.0 & 0.0 & 2.0 & 3.3 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 2.0 \end{bmatrix},$$

then the eigenvalues of the matrix $Q - P$ are: $-0.7192, -2.7808, 2.9050, -0.8950, 0.0000$. Thus, by (i) of Lemma 60, f is monotone. In this case, the computed iterates are

Iter(k)	x_1^k	x_2^k	x_3^k	x_4^k	x_5^k
0	1.00000	3.00000	1.00000	1.00000	2.00000
1	−0.34006	1.59892	0.69395	−0.14884	0.69814
2	−0.67118	1.10637	0.65254	−0.57720	0.36476
3	−0.73773	0.92446	0.66833	−0.74422	0.27939
4	−0.74245	0.85380	0.68821	−0.81255	0.25753
5	−0.73676	0.82503	0.70210	−0.84185	0.25193
6	−0.73172	0.81283	0.71037	−0.85495	0.25049
7	−0.72866	0.80751	0.71494	−0.86102	0.25013
8	−0.72701	0.80512	0.71738	−0.86390	0.25003
9	−0.72618	0.80404	0.71866	−0.86530	0.25001
10	−0.72577	0.80354	0.71932	−0.86599	0.25000

and an approximate solution is

$$x^{10} = (-0.72577, 0.80354, 0.71932, -0.86599, 0.25000)^T$$

with the tolerance $\epsilon = 10^{-3}$.

Test 2 We use Algorithm 2a with the same equilibrium bifunction and the quadratic regularization function as before. In this algorithm at Step 3 of iteration k, we choose x^{k+1} to be the Euclidean projection of $x^k - \gamma_k \sigma_k g^k$ onto K. It is well known that computing this projection leads to a convex quadratic program.

In this particular case, Algorithm 2a collapses to the following one (see Lemma 6.3).

Algorithm 2b

Data $\epsilon > 0$, $\alpha \in (0,1), \theta \in (0,1)$, and select $x^0 \in K$.

Step 0 Set $k = 0$.

Step 1 Find $y^k \in K$ as the unique solution to the quadratic programing

$$\min_{y \in K} \left\{ \frac{1}{2} \langle Hy, y \rangle + \langle h(x^k), y \rangle \right\}, \tag{6.7}$$

where $H = 2\rho Q + I$ and $h(x^k) = [\rho(P - Q) - I]x^k + \rho q$.
If $\|y^k - x^k\| \leq \epsilon$, then stop.

Step 2 Choose $\theta_k \in (0,1)$ such that

$$0 < \theta_k \leq \min \left\{ \frac{u(x^k, y^k)}{v(x^k, y^k)}, \theta \right\}. \tag{6.8}$$

where

$$u(x^k, y^k) = \frac{1}{2\rho} \langle (2\rho P - \alpha I)x^k + (2\rho Q + \alpha I)y^k + 2\rho q, x^k - y^k \rangle$$
$$\text{and} \quad v(x^k, y^k) = \langle (P - Q)(x^k - y^k), x^k - y^k \rangle.$$

Set $z^k = (1 - \theta_k)x^k + \theta_k y^k$ and go to Step 3.

Step 3 Compute $g^k := (P - Q)z^k + 2Qx^k + q$ and

$$\sigma_k := \frac{\langle Pz^k + Qx^k + q, x^k - z^k \rangle}{\|g^k\|^2}. \tag{6.9}$$

Solve the convex quadratic program

$$\min_{y \in K} \left\{ \frac{1}{2} \|y\|^2 + \langle c(x^k, z^k), y \rangle \right\}, \tag{6.10}$$

where $c(x^k, z^k) = \gamma_k \sigma_k g^k - x^k$, to obtain its unique solution x^{k+1}.

Step 4 Set $k := k + 1$, and go to Step 1.

LEMMA 6.3 *Let $f(x, y) = \langle Px + Qy + q, y - x\rangle$. Suppose that the matrix Q is symmetric positive semidefinite and $P - Q$ is symmetric positive definite. Then Algorithm 2a collapses to Algorithm 2b.*

Proof First, we show that the problem (4.19) collapses to the quadratic programing problem (6.7). Indeed, from (4.19), we have:

$$\begin{aligned} f(x^k, y) + \frac{1}{2\rho}\|y - x^k\|^2 &= \langle Px^k + Qy + q, y - x^k\rangle + \frac{1}{2\rho}\|y - x^k\|^2 \\ &= \frac{1}{2\rho}\langle (2\rho Q + I)y, y\rangle + \frac{1}{\rho}\langle [\rho(P - Q) - I]x^k + \rho q, y\rangle \\ &\quad + \frac{1}{2\rho}\|x^k\|^2 - \langle (Px^k + q), x^k\rangle. \end{aligned}$$

Thus, if we take $H = 2\rho Q + I, h(x^k) = [\rho(P - Q) - I]x^k + \rho q$, then it is easy to see that the quadratic programing problem (6.7) is equivalent to (4.19).

Next, we see that the linesearch at Step 2 of Algorithm 2a and 2b are equivalent. Indeed, from (4.20) we have

$$\begin{aligned} & f(z^{k,m}, x^k) - f(z^{k,m}, y^k) - \frac{\alpha}{\rho}\left[G(y^k) - G(x^k) - \langle \nabla G(x^k), y^k - x^k\rangle\right] \\ &= \langle Px^k + Qy^k + q, x^k - y^k\rangle - \theta_k\langle (P - Q)(x^k - y^k), x^k - y^k\rangle - \frac{\alpha}{2\rho}\|x^k - y^k\|^2 \\ &= \frac{1}{2\rho}\langle 2\rho(Px^k + Qy^k + q) - \alpha(x^k - y^k), x^k - y^k\rangle \\ &\quad - \theta_k\langle (P - Q)(x^k - y^k), x^k - y^k\rangle \geq 0 \end{aligned} \tag{6.11}$$

Set

$$u(x^k, y^k) = \frac{1}{2\rho}\langle (2\rho P - \alpha I)x^k + (2\rho Q + \alpha I)y^k + 2\rho q, x^k - y^k\rangle,$$

and

$$v(x^k, y^k) = \langle (P - Q)(x^k - y^k), x^k - y^k\rangle.$$

Then it follows from (6.11) that

$$u(x^k, y^k) \geq \theta_k v(x^k, y^k).$$

Since $x^k \neq y^k$ and $P - Q$ is a symmetric positive definite matrix, $v(x^k, y^k) > 0$. Thus, we can choose

$$0 < \theta_k \leq \min\left\{\frac{u(x^k, y^k)}{v(x^k, y^k)}, \theta\right\} < 1,$$

which proves that (6.8) is equivalent to (4.20).

Finally, by (6.5), we have $g^k = \nabla_2 f(z^k, x^k) = (P - Q)z^k + 2Qx^k + q$ and $f(z^k, x^k) = \langle Pz^k + Qx^k + q, x^k - z^k \rangle$, which imply

$$\sigma_k := \frac{\langle Pz^k + Qx^k + q, x^k - z^k \rangle}{\|g^k\|^2}.$$

Using (4.19) in Step 3 of Algorithm 2a we obtain

$$x^{k+1} = P_K(x^k - \gamma_k \sigma_k g^k),$$

which shows that x^{k+1} is the unique solution of the problem

$$\min_{y \in K} \left\{ \left\| y - (x^k - \gamma_k \sigma_k g^k \right\|^2 \right\}$$

that can be rewritten as

$$\min_{y \in K} \left\{ \frac{1}{2} \|y\|^2 + \left\langle c\left(x^k, y^k\right), y \right\rangle \right\}$$

where $c(x^k, y^k) := \gamma_k \sigma_k g^k - x^k$. The proof is thus complete. ■

Since Algorithm 2b requires the matrix $P - Q$ to be symmetric positive definite, we choose matrix P, Q as in Test 1 that satisfy the condition of Lemma 6.3. With $\rho = 0.5, \alpha = 0.5, \theta = 0.5, \gamma_k = 1$ $(\forall k \geq 1)$ we obtained the following iterates

Iter(k)	x_1^k	x_2^k	x_3^k	x_4^k	x_5^k
0	1.00000	3.00000	1.00000	1.00000	2.00000
1	0.16459	2.08602	0.62354	0.45032	1.42838
2	−0.30068	1.56029	0.43500	0.10278	1.02996
3	−0.55734	1.25434	0.35314	−0.12691	0.74954
4	−0.69594	1.07287	0.33294	−0.28875	0.54864
5	−0.76570	0.96281	0.35151	−0.41320	0.40142
...	...	...	...	...	...
17	−0.72708	0.80471	0.71099	−0.85747	0.20000
18	−0.72657	0.80423	0.71355	−0.86008	0.20000
19	−0.72621	0.80389	0.71538	−0.86196	0.20000
20	−0.72596	0.80365	0.71670	−0.86330	0.20000
21	−0.72579	0.80349	0.71764	−0.86425	0.20000

and an approximate solution

$$x^{21} = (-0.72579, 0.80349, 0.71764, \; -0.86425, 0.20000)^T$$

with the tolerance $\epsilon = 10^{-3}$.

Acknowledgements

The work presented here was completed while the second author was on leave at the University of Namur, Belgium. He wishes to thank the University of Namur for financial support in the form of a FUNDP Research Scholarship. The authors wish to thank the referee for his useful comments and remarks that helped them very much in revising the article.

References

[1] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J., 2005, On the contraction and nonexpansiveness properties of the marginal mapping in generalized variational inequalities involving co-coercive operators. In: A. Eberhard, N. Hadjisavvas and D.T. Luc (Eds), *Generalized Convexity and Generalized Monotonicity and Applications*, Chapter 5 (New York: Springer), pp. 89–111.

[2] Anh, P.N., Muu, L.D., Nguyen, V.H. and Strodiot, J.J., 2005, Using the Banach contraction principle to implement the proximal point method for multivalued monotone variational inequalities. *Journal of Optimization Theory and Applications*, **124**, 285–306.

[3] Antipin, A.S., Budak, B.A. and Vasil'ev, F.P., 2002, A regularized continuous extragradient method of the first order with variable metric for problems of equilibrium programming. *Differential Equations*, **38**, 1683–1693.

[4] Aubin, J.P. and Ekeland, I., 1984, *Applied Nonlinear Analysis* (New York: Wiley).

[5] Berge, C., 1968, *Topological Spaces* (New York: MacMillan).

[6] Blum, E. and Oettli, W., 1994, From optimization and variational inequality to equilibrium problems. *The Mathematics Student*, **63**, 127–149.

[7] Cohen, G., 1980, Auxiliary problem principle and decomposition of optimization problems. *Journal of Optimization Theory and Applications*, **32**, 277–305.

[8] Cohen, G., 1988, Auxiliary principle extended to variational inequalities. *Journal of Optimization Theory and Applications*, **59**, 325–333.

[9] Facchinei, F. and Pang, J.S., 2003, *Finite-Dimensional Variational Inequalities and Complementary Problems* (NewYork: Springer-Verlag).

[10] El Farouq, N., 2001, Pseudomonotone variational inequalities: convergence of auxiliary problem method. *Journal of Optimization Theory and Applications*, **111**, 305–325.

[11] Hue, T.T., Strodiot, J.J. and Nguyen, V.H., 2004, Convergence of the approximate auxiliary problem method for solving generalized variational inequalities. *Journal of Optimization Theory and Applications*, **121**, 119–145.

[12] Iusem, A.N., Kassay, G. and Sosa, W., On certain conditions for the existence of solutions of equilibrium problems. *Applied Mathematics E-Notes* (to appear).

[13] Konnov, I.V., 2000, *Combined Relaxation Methods for Variational Inequalities* (Berlin: Springer-Verlag).

[14] Konnov, I.V., 2003, Application of the proximal point method to nonmonotone equilibrium problems. *Journal of Optimization Theory and Applications*, **119**, 317–333.

[15] Korpelevich, G.M., 1976, Extragradient method for finding saddle points and other problems. *Matecon*, **12**, 747–756.

[16] Martinet, B., 1970, Régularisation d'inéquations variationelles par approximations successives. *Revue Francaise d'Automatique et d'Informatique Recherche Opérationnelle*, **4**, 154–159.

[17] Mastroeni, G., 2000, On auxiliary principle for equilibrium problems. *Publicatione del Dipartimento di Mathematica dell'Universita di Pisa*, **3** 1244–1258.

[18] Mastroeni, G., 2004, Gap function for equilibrium problems. *Journal of Global Optimization*, **27**, 411–426.

[19] Moudafi, A., 2003, Second order differential proximal methods for equilibrium problems. *Journal of Inequalities in Pure and Applied Mathematics*, **4**, 1–7.

[20] Moudafi, A., 1999, Proximal point algorithm extended to equilibrium problem. *Journal of Natural Geometry*, **15**, 91–100.

[21] Muu, L.D., 1984, Stability property of a class of variational inequalities. *Optimization*, **15**, 347–353.

[22] Muu, L.D. and Oettli, W., 1992, Convergence of an adaptive penalty scheme for finding constraint equilibria. *Nonlinear Analysis: Theory, Methods and Applications*, **18**, 1159–1166.

[23] Nguyen, V.H., 2002, *Lecture Notes on Equilibrium Problems* (Nha Trang: CIUF-CUD Summer School on Optimization and Applied Mathematics).

[24] Noor, M.A., 2003, Extragradient methods for pseudomonotone variational inequalities. *Journal of Optimization Theory and Applications*, **117**, 475–488.

[25] Noor, M.A., 2004, Auxiliary principle technique for equilibrium problems. *Journal of Optimization Theory and Applications*, **122**, 371–386.

[26] Quy, N.V., Nguyen, V.H. and Muu, L.D., 2005, On the Cournot–Nash oligopolistic market equilibrium problem with concave cost function, Epreprint. Hanoi Institute of Mathematics, 2.

[27] Rockafellar, R.T., 1970, *Convex Analysis* (Princeton, NJ: Princeton University Press).

[28] Rockafellar, R.T., 1976, Monotone operators and the proximal point algorithm. *SIAM Journal Control Optimization*, **14**, 877–898.
[29] Salmon, G., Nguyen, V.H. and Strodiot, J.J., 2000, Coupling the auxiliary problem principle and the epiconvergence theory to solve generalized variational inequalities. *Journal of Optimization Theory and Applications*, **104**, 629–657.
[30] Tan, N.X., 2004, On the existence of solutions of quasi-variational inclusion problems. *Journal of Optimization Theory and Applications*, **123**, 619–638.
[31] Nguyen, T.T.V., Strodiot, J.J. and Nguyen, V.H., A Bundle method for solving equilibrium problems. *Mathematical Programming* (to appear).