TAIWANESE JOURNAL OF MATHEMATICS
Vol. 17, No. 6, pp. 2117-2133, December 2013
DOI: 10.11650/tjm.17.2013.3266
This paper is available online at http://journal.taiwanmathsoc.org.tw

AN INEXACT PROXIMAL POINT ALGORITHM FOR NONMONOTONE EQUILIBRIUM PROBLEMS IN BANACH SPACES

Guo-ji Tang* and Xing Wang

Abstract. In this paper, we continue to investigate the inexact hybrid proximal point algorithm proposed by Mashreghi and Nasri for equilibrium problems in Banach spaces. Under some classes of generalized monotone conditions, we prove that the sequence generated by the method is strongly convergent to a solution of the problem, which is closest to the initial iterate, in the sense of Bregman distance. As an application, we obtain some analogues for some classes of generalized monotone variational inequalities. The results presented in this paper generalize and improve some recent results in literatures.

1. Introduction

Let $\mathbb{B}$ be a reflexive Banach space and $K \subset \mathbb{B}$ a nonempty, closed and convex set. Given $f : K \times K \to \mathbb{R}$ such that
P1: $f(x, x) = 0$ for all $x \in K$,
P2: $f(x, \cdot) : K \to \mathbb{R}$ is convex and lower semicontinuous for all $x \in K$,
P3: $f(\cdot, y) : K \to \mathbb{R}$ is upper semicontinuous for all $y \in K$,
the equilibrium problem, denoted by EP(f, K), consists of finding $x^* \in K$ such that

$$f(x^*, y) \geq 0, \quad \forall y \in K. \tag{1.1}$$

Such an x^* is called a solution of EP(f, K). The set of solutions of EP(f, K) will be denoted by S.

Equilibrium problem theory has emerged as an interesting branch of applicable mathematics. This theory has become a rich source of inspiration and motivation

Received May 2, 2013, accepted June 6, 2013.
Communicated by Jen-Chih Yao.
2010 *Mathematics Subject Classification*: 90C30, 47H05.
Key words and phrases: Proximal point algorithm, Strong convergence, Generalized monotonicity, Equilibrium problem, Variational inequality, Banach space.
This work is supported by Guangxi Natural Science Foundation (2013GXNSFBA019015).
*Corresponding author.

for the study of a large number of problems arising in economics, optimization and operations research in a general and unified way.

The equilibrium problem has been rather widely studied. Many researchers consider computational methods for the problem (see, for example, [1, 2, 7, 15-17, 20, 23, 25, 26]) as well as some researchers concentrate on dealing with conditions for the existence of solutions (see, for example, [4, 5, 6, 9, 10, 13, 14]). Among these methods, some researchers study the problem in finite dimensional spaces [1, 2, 20]; some researchers require the problem to be monotone [25, 26]. In terms of the nonmonotone and infinite dimensional settings, only a few references can be found. Recently, Iusem et al. [17] and Mashreghi and Nasri [23] proposed inexact proximal point algorithms, whose origins can be traced back to [21, 24], attained its basic formulation in the work of Rockafellar [28], for some classes of generalized monotone equilibrium problems in Banach spaces. If the equilibrium problem has solutions, then the sequence $\{x^j\}$ generated by the method is shown to be weakly (strongly) convergent to a solution of the problem under the assumptions that f satisfies A2 and any one of A3, A4' and A4* (see Section 2 for their definitions), and $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$. In these settings, f is not necessarily monotone. However, if we would get rid of the weak upper semicontinuity of $f(\cdot, y)$ for $y \in K$, then f is asked to be monotone again on K (see Section 6 of [17] and Section 4 of [23]).

Motivated and inspired by the research work mentioned above, in this paper, we continue to study the inexact hybrid proximal Bregman projection method proposed by Mashreghi and Nasri [23]. Under some classes of generalized monotone conditions, we prove that the sequence generated by the method is strongly convergent to a solution of the problem in Banach spaces, which is closest to the initial iterate, in the sense of Bregman distance. We would like to point out that the techniques used in this paper are very different from those presented in the related references [15, 17, 23].

The main contribution of this paper lies in the following aspects:

(i) Compared with [23], we get rid of the condition that $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$;

(ii) Compared with [17], not only the strong convergence is obtained, but also the condition that $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$ is removed;

(iii) As an application, we obtain some analogues for some classes of generalized monotone variational inequalities.

Therefore, the results presented in this paper generalize, extend and improve some known results in Mashreghi and Nasri [23], Iusem et al. [17] and Farouq [8].

2. Preliminaries

From now on, $\mathbb{B}$ is a reflexive real Banach space.We will use the notation $\langle v, x\rangle$ for the duality product $v(x)$ of $x \in \mathbb{B}$ and $v \in \mathbb{B}^*$. Convergence in the weak (respectively strong) topology of a sequence will be indicated by $\rightharpoonup$ (respectively $\rightarrow$). Let $g : B \to \mathbb{R}$ be a strictly convex, lower semicontinuous and G-differentiable function. The Bregman distance associated to g is defined as $D_g : \mathbb{B} \times \mathbb{B} \to \mathbb{R}$,

$$D_g(x, y) = g(x) - g(y) - \langle g'(y), x - y\rangle. \tag{2.1}$$

From this definition, it is straightforward to verify that D_g satisfies the three-point equality (see, e.g., [27]):

$$D_g(x, y) = D_g(z, y) + D_g(x, z) + \langle g'(y) - g'(z), z - x\rangle. \tag{2.2}$$

for any $x, y, z \in \mathbb{B}$. As g is strictly convex, the function $D_g(\cdot, y)$ is nonnegative, strictly convex and $D_g(x, y) = 0$ if and only if $x = y$.

The modulus of convexity of g at x is the function $\nu_g : \mathbb{B} \times \mathbb{R}_+ \to \mathbb{R}$, defined as

$$\nu_g(x, t) = \inf\{D_g(y, x) : \|y - x\| = t\}. \tag{2.3}$$

A function g is totally convex if $\nu_g(x, t) > 0$ for all $x \in \mathbb{B}$ and $t > 0$. Additionally, if $\inf_{x\in C} \nu_g(x, t) > 0$ for each bounded subset $C \subset \mathbb{B}$ then g is called uniformly totally convex. If g is totally convex, then

$$\nu_g(x, st) \geq s\nu_g(x, t), \quad \forall s \geq 1, t \geq 0, x \in B. \tag{2.4}$$

The assumptions on g to be considered in the sequel are the following:

H1: The level sets of $D_g(x, \cdot)$ are bounded for all $x \in \mathbb{B}$.

H2: g is uniformly totally convex.

H3: g' is uniformly continuous on bounded subsets of $\mathbb{B}$.

H4: g' is onto, i.e., for all $y \in \mathbb{B}^*$, there exists $x \in B$ such that $g'(x) = y$.

H5: $\lim\limits_{\|x\|\to\infty} [g(x) - \rho\|x - z\|] = +\infty$ for all $z \in K$ and $\rho \geq 0$.

It is important to check that functions satisfying these properties are available in a wide class of Banach spaces. We have the following result.

Proposition 2.1. [17]. *If $\mathbb{B}$ is a uniformly smooth and uniformly convex Banach space, then $g(x) = r\|x\|^s$ satisfies H1-H5 for all $r > 0$ and all $s > 1$.*

The next result deals with Bregman projections onto closed and convex sets.

Proposition 2.2. [17]. *If g satisfies H1-H2 and $C \subset \mathbb{B}$ is closed and convex, then for all $\bar{x} \in \mathbb{B}$ there exists a unique solution $\hat{x}$ of the problem* $\min D_g(x, \bar{x})$ *s.t.* $x \in C$*, which satisfies*

$$\langle g'(\bar{x}) - g'(\hat{x}), y - \hat{x}\rangle \leq 0, \quad \forall y \in C. \tag{2.5}$$

The point $\hat{x}$ is said to be the Bregman projection of $\bar{x}$ onto C, denoted by $\Pi_C^g(\bar{x})$.

Proposition 2.3. [18]. *Assume that g satisfies H2. Let $\{x^k\}, \{y^k\} \subset \mathbb{B}$ be two sequences such that at least one of them is bounded. If $\lim\limits_{k\to\infty} D_g(y^k, x^k) = 0$, then $\lim\limits_{k\to\infty} \|x^k - y^k\| = 0$.*

Proposition 2.4. [18]. *If g satisfies H3, then g and g' are bounded on bounded subsets of $\mathbb{B}$.*

We recall some monotonicity properties of the bifunction f, which will be used in the sequel.

Definition 2.1. Let $\mathbb{B}$ be a reflexive Banach space and $K \subset \mathbb{B}$ be a nonempty closed and convex set. $f : K \times K \to \mathbb{R}$ is a bifunction. f is said to be

A1: monotone, if

$$f(x, y) + f(y, x) \leq 0, \quad \forall x, y \in K; \tag{2.6}$$

A2: undermonotone, if there exists $\theta > 0$ such that

$$f(x, y) + f(y, x) \leq \theta\langle g'(x) - g'(y), x - y\rangle, \quad \forall x, y \in K, \tag{2.7}$$

where θ is called the undermonotonicity constant of f;

A3: pseudomonotone, if

$$f(x, y) \geq 0 \Rightarrow f(y, x) \leq 0, \quad \forall x, y \in K; \tag{2.8}$$

A4: properly quasimonotone, if for all $x^1, \cdots, x^n \in K$ and all $\lambda_1, \cdots, \lambda_n \geq 0$ such that $\sum\limits_{i=1}^{n} \lambda_i = 1$ it holds that

$$\min_{1\leq i\leq n} f(x^i, \sum_{j=1}^{n} \lambda_j x^j) \leq 0; \tag{2.9}$$

We will consider also the following two variants of properly quasimonotonicity:

A4': f satisfies A4, with strict inequality in (2.9) if the x^is are pairwise distinct and the λ_is are all strictly positive;

A4*: For all $x^1, \cdots, x^n \in K$ and all $\lambda_1, \cdots, \lambda_n \geq 0$ such that $\sum\limits_{i=1}^{n} \lambda_i = 1$ it holds that

$$\sum_{i=1}^{n} \lambda_i f(x^i, \sum_{j=1}^{n} \lambda_j x^j) \leq 0. \tag{2.10}$$

Remark 2.1. (i) In A2, if $\mathbb{B}$ is a Hilbert space and $g = \frac{1}{2}\|\cdot\|^2$, then (2.7) reduces to

$$f(x,y) + f(y,x) \leq \theta\|x-y\|^2, \quad \forall x, y \in K. \tag{2.11}$$

(ii) The condition A1 is widely used in proving the convergence results of most algorithms for solving (1.1) (see, for example, [25, 26]).

(iii) It is not hard to check that both A2 (i.e., undermonotonicity) and A3 (i.e., pseudomonotonicity) are weaker than A1 (i.e., monotonicity).

(iv) As mentioned by Iusem et al. in [13], A4 (i.e., proper quasimonotonicity) was introduced by Zhou and Chen [30]. In the case of variational inequalities, it is stronger than quasimonotonicity, as defined, for example, in [19]. In the general case, as considered here A4, it neither implies quasimonotonicity nor is implied by it (approximate examples are given in [4]).

(v) A4' and A4* as variants of A4 were introduced by Iusem et al. in [13]. Both A4' and A4* are weaker than A1 (i.e., monotonicity), for example, [17].

(vi) Some examples were given in [13] showing that A4', A4* and A3 (i.e., pseudomonotonicity) are mutually independent in the sense that none of them implies any of the remaining ones.

In this paper, we will use any one of three combination conditions: A2-A3, A2-A4' and A2-A4*, instead of A1. As discussed above, we conclude that the function set with A1 is properly contained in the one with any one among of A2-A3, A2-A4' and A2-A4*.

Under the condition that f satisfies A2, the authors of [17] proved the unique existence of the proximal subproblem.

Proposition 2.5. [17]. *Consider f satisfying P1-P3 and A2. Fix $\bar{x} \in \mathbb{B}$, $e \in \mathbb{B}^*$ and $\gamma > \theta$, where θ is the undermonotonicity constant in A2. Take $g : \mathbb{B} \to \mathbb{R}$ satisfying H1-H2 and H5. If $\tilde{f} : K \times K \to \mathbb{R}$ is defined as*

$$\tilde{f}(x,y) = f(x,y) + \gamma\langle g'(x) - g'(\bar{x}), y - x\rangle - \langle e, y - x\rangle, \tag{2.12}$$

then EP$(\tilde{f}, K)$ has a unique solution.

In order to prove the convergence result of our method, we need to show that the solution set of (1.1) is a closed and convex subset of $\mathbb{B}$. For this purpose, we are interested now in the following dual equilibrium problem (to be denoted DEP(f,K)) and relation with EP(f,K).

DEP(f,K)) consists of finding $x^* \in K$ such that

$$f(y, x^*) \leq 0, \quad \forall y \in K. \tag{2.13}$$

The set of solutions of DEP(f,K)) will be denoted by S^d.

Proposition 2.6. [17]. *Assume that any one among A3, A4' and A4* holds. If P2 is satisfied, then* $f(y, x^*) \leq 0$ *for all* $y \in K$ *and all* $x^* \in S$.

Proposition 2.7. *If* f *satisfies P1-P3 and any one among A3, A4' and A4*, then the solution sets of EP*(f, K) *and DEP*(f, K)*) coincide, i.e.,* $S = S^d$.

Proof. $S \subset S^d$ is immediate by Proposition 2.6. On the other inclusion relation, let $x^* \in S^d$, and take any $w \in K$. For each $t \in (0, 1)$, define $w_t := tw + (1-t)x^*$. Since K is convex, w_t belongs to K. Since x^* solves (2.13), we get $f(w_t, x^*) \leq 0$ for all $t \in (0, 1)$. Adding $tf(x^*, w)$ to two sides of this inequality and using P1, we obtain

$$f(w_t, x^*) + tf(x^*, w) - \frac{1}{1-t} f(w_t, w_t) \leq tf(x^*, w). \tag{2.14}$$

By P2, $f(w_t, \cdot)$ is convex, it follows that $f(w_t, w_t) = f(w_t, tw + (1-t)x^*) \leq tf(w_t, w) + (1-t)f(w_t, x^*)$. Thus,

$$\begin{aligned} & tf(x^*, w) \\ \geq\ & f(w_t, x^*) + tf(x^*, w) - \frac{1}{1-t}[tf(w_t, w) + (1-t)f(w_t, x^*)] \\ =\ & tf(x^*, w) - \frac{t}{1-t} f(w_t, w). \end{aligned}$$

By dividing t from the leftmost and rightmost hand side, we get

$$f(x^*, w) \geq f(x^*, w) - \frac{1}{1-t} f(w_t, w).$$

Taking limits as $t \to 0^+$, we conclude, using P3, that $f(x^*, w) \geq 0$ for all $w \in K$. This implies that $S^d \subset S$. Therefore, the desired result follows.

From Proposition 2.7, the solution set S will be characterized as

$$S = S^d = \bigcap_{y \in K} \{x^* \in K : f(y, x^*) \leq 0\}.$$

Note that for each $y \in K$, the set $\{x^* \in K : f(y, x^*) \leq 0\}$ is a nonempty, closed and convex subset of K, because $f(y, y) = 0$ for all $y \in K$, K is closed and convex, and f is lower semicontinuous and convex in the second argument. Thus, S is closed and convex set if it is nonempty.

3. Algorithm

We restate the algorithm of [23] as follows:

Algorithm 3.1.

Step 0. Choose any $x^0 \in K$.

Step j. For $x^j \in K$, find a pair $\tilde{x}^j \in K$, $e^j \in \mathbb{B}^*$ such that $\tilde{x}^j$ solves EP(f_j^e, K) with

(3.1) $$f_j^e(x, y) = f(x, y) + \gamma_j \langle g'(x) - g'(x^j), y - x\rangle - \langle e^j, y - x\rangle,$$

i.e.,

(3.2) $$f_j^e(\tilde{x}^j, y) \geq 0, \quad \forall y \in K,$$

and e^j satisfies

(3.3) $$\|e^j\| \leq \begin{cases} \sigma\gamma_j D_g(\tilde{x}^j, x^j), & \text{if } \|x^j - \tilde{x}^j\| < 1, \\ \sigma\gamma_j \nu_g(x^j, 1), & \text{if } \|x^j - \tilde{x}^j\| \geq 1. \end{cases}$$

Let

(3.4) $$v^j = \gamma_j[g'(x^j) - g'(\tilde{x}^j)] + e^j.$$

If $v^j = 0$ or $\tilde{x}^j = x^j$, then stop. Otherwise, take

(3.5) $$H_j = \{x \in \mathbb{B} : \langle v^j, x - \tilde{x}^j\rangle \leq 0\},$$

(3.6) $$W_j = \{x \in \mathbb{B} : \langle g'(x^0) - g'(x^j), x - x^j\rangle \leq 0\},$$

and define

(3.7) $$x^{j+1} = \text{argmin}_{x \in H_j \cap W_j \cap K} D_g(x, x^0) = \Pi^g_{H_j \cap W_j \cap K}(x^0).$$

In order to show the well-definedness of the algorithm, we just need to ensure existence of solution for the proximal subproblems (i.e., (3.2)-(3.3)) and nonemptyness of the closed and convex set $H_j \cap W_j \cap K$. Proposition 2.5 ensures the existence of a unique solution for EP(f_j^e, K), say $\tilde{x}^j$. That is to say (3.2)-(3.3) is feasible. The following proposition says that, if the solution set is nonempty, then $S \subset H_j \cap W_j \cap K$.

Proposition 3.1. *If $S \neq \emptyset$, then $S \subset H_j \cap W_j \cap K$, thus $H_j \cap W_j \cap K \neq \emptyset$.*

Proof. Since $S \neq \emptyset$, take any arbitrary $x^* \in S$. It follows that

$$f(x^*, y) \geq 0, \quad \forall y \in K.$$

From (3.2), we have $f_j^e(\tilde{x}^j, y) \geq 0$ for all $y \in K$, i.e.,

(3.8) $$f(\tilde{x}^j, y) - \langle v^j, y - \tilde{x}^j\rangle \geq 0, \quad \forall y \in K.$$

Thus, we have

$$\langle v^j, x^* - \tilde{x}^j\rangle \leq f(\tilde{x}^j, x^*) \leq 0,$$

where the first inequality follows from (3.8) by taking $y = x^*$, the second from Proposition 2.6. So, $x^* \in H_j$ by the definition of H_j, thus, $S \subset H_j$ by the fact that x^* is arbitrary in S.

From the analysis above, it is sufficient to prove that $S \subset W_j$ for all $j \geq 0$. The proof will be given by induction.

Obviously,

$$S \subset W_0 = \mathbb{B}.$$

Suppose that $S \subset W_j$, thus $S \subset H_j \cap W_j \cap K$. For any $x^* \in S$, using Proposition 2.2 and the fact that $x^{j+1} = \Pi_{H_j \cap W_j \cap K}(x^0)$, it holds that

$$\langle g'(x^0) - g'(x^{j+1}), x^* - x^{j+1} \rangle \leq 0,$$

and thus $S \subset W_{j+1}$. This shows that $S \subset W_j$ for all $j \geq 0$ and the desired result follows.

Proposition 3.2. [17]. *Let $\{x^j\}$, $\tilde{x}^j$, $\{\gamma_j\}$ and σ be in Algorithm 3.1 and assume that g satisfies H2. For all j, it holds that*

$$\|e^j\| \|x^j - \tilde{x}^j\| \leq \sigma \gamma_j D_g(\tilde{x}^j, x^j) \leq \gamma_j D_g(\tilde{x}^j, x^j). \tag{3.9}$$

4. Convergence Analysis

First we settle the issue of finite termination of Algorithm 3.1.

Theorem 4.1. *Suppose that Algorithm 3.1 stops after j steps. Then $\tilde{x}^j$ generated by Algorithm 3.1 is a solution of (1.1).*

Proof. Algorithm 3.1 stops at jth iteration in two cases: if $v^j = 0$, in which case, by (3.1), (3.2) and (3.4), $\tilde{x}^j$ solves (1.1), or if $\tilde{x}^j = x^j$, in which case, by (3.3), $e^j = 0$, which in turn implies, by (3.4), $v^j = 0$ and we are back to the first case. Consequently, $\tilde{x}^j$ is a solution of (1.1).

From now on we assume that the sequence $\{x^j\}$ is infinite.

Theorem 4.2. *Let $x^0 \in K$ and W_j be defined as in (3.6). Suppose that the algorithm, starting from x^0, reaches iteration j. Then*

(a) *For any $w \in W_j$ it holds*

$$D_g(w, x^0) \geq D_g(w, x^j) + D_g(x^j, x^0). \tag{4.1}$$

(b) *x^j is the Bregman projection, associated to g, of x^0 over W_j, i.e.,*

$$x^j = \Pi^g_{W_j}(x^0) = \mathit{argmin}_{x \in W_j} D_g(x, x^0). \tag{4.2}$$

Proof. To prove item (a), take any $w \in W_j$. From (3.6),

$$\langle g'(x^0) - g'(x^j), w - w^j \rangle \leq 0.$$

Using also the three-point property (see (2.2)), it follows that

$$\begin{aligned} & D_g(w, x^0) \\ &= D_g(w, x^j) + D_g(x^j, x^0) + \langle g'(x^0) - g'(x^j), x^j - w \rangle \\ &\geq D_g(w, x^j) + D_g(x^j, x^0), \end{aligned}$$

which proves item (a).

Item (b) follows (a) and nonnegativity and strict convexity of $D_g(\cdot, x^j)$. Just note that, in view of (3.6), $x^j \in W_j$.

Theorem 4.3. *Consider EP(f, K). Assume that f satisfies P1-P3, A2 and also any one among A3, A4' and A4*. Take $g : \mathbb{B} \to \mathbb{R}$ satisfies H1-H5 and an exogenous sequence $\{\gamma_j\} \subset (\theta, \bar{\gamma}]$, where θ is the undermonotonicity constant in A2 and $\bar{\gamma} > \theta$ is a some constant. Let $\{x^j\}$ be the sequence by Algorithm 3.1. If EP(f, K) has solutions, then*

(a) $\{D_g(x^j, x^0)\}$ *is nondecreasing and convergent.*

(b) $\{x^j\}$ *is bounded.*

(c) $\{x^{j+1} - x^j\}$ *converges strongly to 0.*

(d) $\{\gamma_j^{-1} e^j\}$ *is bounded.*

(e) $\{x^{j+1} - \tilde{x}^j\}$ *and* $\{\tilde{x}^j - x^j\}$ *converge strongly to 0.*

(f) $\{v^j\}$ *converges strongly to 0.*

Proof. (a) Since $x^{j+1} \in W_j$ (see (3.7)), from item (a) of Theorem 4.2 and nonnegativity of $D_g(\cdot, x^j)$ we have

$$D_g(x^{j+1}, x^0) \geq D_g(x^{j+1}, x^j) + D_g(x^j, x^0) \geq D_g(x^j, x^0) \tag{4.3}$$

which shows that $\{D_g(x^j, x^0)\}$ is nondecreasing. From (3.7) we know that

$$D_g(x^{j+1}, x^0) \leq D_g(x, x^0) \tag{4.4}$$

for all $x \in H_j \cap W_j \cap K$, particularly, for all $x^* \in S(\subset H_j \cap W_j \cap K)$. So, it holds

$$D_g(x^j, x^0) \leq D_g(x^*, x^0), \tag{4.5}$$

and hence the sequence $\{D_g(x^j, x^0)\}$ is convergent.

The remainder of the proof is similar to that of Proposition 4.2 of [23].

We remind that a sequence $\{z^j\} \subset K$ is an asymptotically solving sequence for EP(f, K) if $\liminf\limits_{j \to \infty} f(z^j, y) \geq 0$ for all $y \in K$.

Theorem 4.4. *Consider EP(f, K). Assume that f satisfies P1-P3, A2 and also any one among A3, A4' and A4*. Take $g : \mathbb{B} \to \mathbb{R}$ satisfies H1-H5 and an exogenous sequence $\{\gamma_j\} \subset (\theta, \bar{\gamma}]$, where θ is the undermonotonicity constant in A2 and $\bar{\gamma} > \theta$ is a some constant. Let $\{x^j\}$ be the sequence by Algorithm 3.1. If EP(f, K) has solutions, then*

(a) $\{\tilde{x}^j\}$ *is an asymptotically solving sequence for EP(f, K).*

(b) *all weak cluster points of $\{x^j\}$ solve EP(f, K).*

(c) $\{D_g(x^j, x^0)\}$ *converges to* $\inf\limits_{z\in S} D_g(z, x^0)$ *and* $\{x^j\}$ *converges strongly to* $\hat{x} = \Pi_S^g(x^0) = \arg\min_{z\in S} D_g(z, x^0)$.

Proof. (a) Fix $y \in K$. Because $\{\tilde{x}^j\}$ solves EP(f_j^e, K), by the definition of f_j^e and Cauchy-Schwartz inequality, we have that

$$
\begin{aligned}
0 &\leq f_j^e(\tilde{x}^j, y) \\
&= f(\tilde{x}^j, y) + \langle \gamma_j[g'(\tilde{x}^j) - g'(x^j)] - e^j, y - \tilde{x}^j\rangle \\
&= f(\tilde{x}^j, y) + \langle -v^j, y - \tilde{x}^j\rangle \\
&\leq f(\tilde{x}^j, y) + \|v^j\|\|y - \tilde{x}^j\|.
\end{aligned} \tag{4.6}
$$

By item (b)and (e) of Theorem 4.3 , we know that $\{\tilde{x}^j\}$, and therefore, the sequence $\{y - \tilde{x}^j\}$, are bounded for each fixed y. Consequently, taking limits in (4.6) as $j \to \infty$ and using Theorem 4.3 (f), we get

$$0 \leq \liminf_{j\to\infty} f(\tilde{x}^j, y), \tag{4.7}$$

for all $y \in K$.

(b) By Theorem 4.3 (b) and (e), $\{x^j\}$ has weak cluster points, all of which are also weak cluster points of $\{\tilde{x}^j\}$. These weak cluster points belong to K, which, being closed and convex, is weakly closed. Let $\tilde{x}$ be a weak cluster point of $\{\tilde{x}^j\}$, say the weak limit of the subsequence $\{\tilde{x}^{j_k}\}$ of $\{\tilde{x}^j\}$. By our assumptions, we investigate the property of $\tilde{x}$ being a solution of (1.1) in three cases as follows.

The case of A3: By (4.7), it follows that

$$0 \leq \liminf_{k\to\infty} f(\tilde{x}^{j_k}, y), \quad \forall y \in K. \tag{4.8}$$

By the definition of lower limit, for each $y \in K$, there is a subsequence $\{\tilde{x}^{j_{k_i}}\}$, which also converges weakly to $\tilde{x}$, of $\{\tilde{x}^{j_k}\}$ such that $\lim\limits_{i\to\infty} f(\tilde{x}^{j_{k_i}}, y) = \liminf\limits_{k\to\infty} f(\tilde{x}^{j_k}, y)$. Using (4.8), we have $0 \leq \lim\limits_{i\to\infty} f(\tilde{x}^{j_{k_i}}, y)$, that is to say, there is a positive integer N, such that

$$f(\tilde{x}^{j_{k_i}}, y) \geq 0, \quad \forall i > N. \tag{4.9}$$

By A3, we get $f(y, \tilde{x}^{j_{k_i}}) \le 0$ for all $i > N$. Because f, being convex and lower semicontinuous in the second argument, is weakly lower semicontinuous in the same one, using $\tilde{x}^{j_{k_i}} \rightharpoonup \tilde{x}$, we obtain

$$f(y, \tilde{x}) \le \liminf_{i\to\infty} f(y, \tilde{x}^{j_{k_i}}) \le 0.$$

Observe that y is arbitrary in K. Consequently, $\tilde{x}$ is a solution of (2.13), by Proposition 2.7, $\tilde{x}$ also solves (1.1).

The case of A4': For each $y \in K$, similar to the proof of item (a), substituting $\tilde{x}^j$ by $\tilde{x}^{j_k}$, e^j by e^{j_k} and y by $\frac{1}{k}y + (1-\frac{1}{k})\tilde{x}^{j_k}$ $(\in K)$ in (4.6), we get

$$0 \le f(\tilde{x}^{j_k}, \frac{1}{k}y + (1-\frac{1}{k})\tilde{x}^{j_k}) + \|v^{j_k}\|\|\frac{1}{k}(y - \tilde{x}^{j_k})\|, \tag{4.10}$$

thus, (4.7) is replaced by

$$0 \le \liminf_{k\to\infty} f(\tilde{x}^{j_k}, \frac{1}{k}y + (1-\frac{1}{k})\tilde{x}^{j_k}). \tag{4.11}$$

By again the definition of lower limit, there is a subsequence $\{k_i\}$ such that

$$\lim_{i\to\infty} f(\tilde{x}^{j_{k_i}}, \frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}}) = \liminf_{k\to\infty} f(\tilde{x}^{j_k}, \frac{1}{k}y + (1-\frac{1}{k})\tilde{x}^{j_k}) \tag{4.12}$$

and there must be $\frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}} \rightharpoonup \tilde{x}$ as $i \to \infty$. Replacing (4.12) into (4.11), we conclude that there is a positive integer N such that

$$f(\tilde{x}^{j_{k_i}}, \frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}}) \ge 0, \quad \forall i > N. \tag{4.13}$$

If y does not belong to $\{\tilde{x}^{j_{k_i}} : i > N\}$, then it follows from the property A4' and (4.13) that

$$f(y, \frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}}) < 0, \quad \forall i > N. \tag{4.14}$$

Since $\frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}} \rightharpoonup \tilde{x}$ as $i \to \infty$ and f is weakly lower semicontinuous in second argument, by taking limits as $i \to \infty$, we have

$$f(y, \tilde{x}) \le \liminf_{i\to\infty} f(y, \frac{1}{k_i}y + (1-\frac{1}{k_i})\tilde{x}^{j_{k_i}}) \le 0.$$

Otherwise, for each $i > N$, we consider

$$f(\tilde{x}^{j_{k_i}}, \tilde{x}) \le \liminf_{i\to\infty} f(\tilde{x}^{j_{k_i}}, \tilde{x}^{j_{k_i}}) = 0,$$

where the inequality follows from $\tilde{x}^{j_{k_i}} \rightharpoonup \tilde{x}$ and the weakly lower semicontinuity of f in the second argument, the equality follows from the assumption P1. Thus, we can conclude that

$$f(y, \tilde{x}) \leq 0, \quad \forall y \in K.$$

Consequently, by Proposition 2.7, $\tilde{x}$ solves (1.1).

The case of A4*: We copy some steps from the start point of the proof of the case A4' to (4.13). Now we continue to the proof. It follows from A4* that

$$\frac{1}{k_i} f(y, \frac{1}{k_i} y + (1 - \frac{1}{k_i}) \tilde{x}^{j_{k_i}}) + (1 - \frac{1}{k_i}) f(\tilde{x}^{j_{k_i}}, \frac{1}{k_i} y + (1 - \frac{1}{k_i}) \tilde{x}^{j_{k_i}}) \leq 0.$$

By (4.13) and $\frac{1}{k_i} > 0$, we conclude that

$$f(y, \frac{1}{k_i} y + (1 - \frac{1}{k_i}) \tilde{x}^{j_{k_i}}) \leq 0, \quad \forall i > N,$$

Since $\frac{1}{k_i} y + (1 - \frac{1}{k_i}) \tilde{x}^{j_{k_i}} \rightharpoonup \tilde{x}$ as $i \to \infty$ and f is weakly lower semicontinuous in second argument, by taking limits as $i \to \infty$, we have

$$f(y, \tilde{x}) \leq \liminf_{i \to \infty} f(y, \frac{1}{k_i} y + (1 - \frac{1}{k_i}) \tilde{x}^{j_{k_i}}) \leq 0.$$

Consequently, $\tilde{x}$ is a solution of (2.13), by again Proposition 2.7, $\tilde{x}$ solves (1.1).

(c) Note that the Bregman projection of the initial iterate x^0 over S, $\hat{x} = \Pi_S(x^0)$, exists because the solution set is closed, convex and we assumed it to be nonempty and g is totally convex. By taking $x^* = \hat{x}$ in (4.5), we have

$$D_g(x^j, x^0) \leq D_g(\hat{x}, x^0). \tag{4.15}$$

Because $\{D_g(x^j, x^0)\}$ is nondecreasing and convergent (see Theorem 4.3 (a)), we can let

$$\alpha = \lim_{j \to \infty} D_g(x^j, x^0) = \sup_j D_g(x^j, x^0), \tag{4.16}$$

and choose any weakly convergent subsequence $x^{j_k} \rightharpoonup x^\infty$. Then, from the above discussion, $x^\infty \in S$. On the other hand, since g is convex and lower semicontinuous, so is $D_g(\cdot, x^0)$. Consequently, $D_g(\cdot, x^0)$ is weakly lower continuous. This implies

$$D_g(\hat{x}, x^0) \leq D_g(x^\infty, x^0) \leq \liminf_{k \to \infty} D_g(x^{j_k}, x^0) = \alpha, \tag{4.17}$$

where the first inequality follows from the the fact that $\hat{x} = \Pi_S^g(x^0)$. From (4.15)-(4.17), we get

$$\alpha = D_g(\hat{x}, x^0) = D_g(x^\infty, x^0). \tag{4.18}$$

Consequently, it follows that

$$\begin{aligned}
0 &\le D_g(x^\infty, \hat{x}) \quad \text{(by nonnegativity)} \\
&= D_g(x^\infty, x^0) - D_g(\hat{x}, x^0) + \langle g'(x^0) - g'(\hat{x}), x^\infty - \hat{x} \rangle \qquad \text{(by (2.2))} \\
&= \langle g'(x^0) - g'(\hat{x}), x^\infty - \hat{x} \rangle \quad \text{(by (4.18))} \\
&\le 0. \quad \text{(by } \hat{x} = \Pi_S^g(x^0) \text{ and (2.5))}
\end{aligned}$$

So, $D_g(x^\infty, \hat{x}) = 0$, thus, $x^\infty = \hat{x}$. This implies that $\{x^j\}$ has a unique weak cluster point and converges weakly to $\hat{x}$. Moreover, from (4.1) with $w = \hat{x} \in W_j$, and taking limits, it follows

$$\limsup_{j\to\infty} D_g(\hat{x}, x^j) \le \limsup_{j\to\infty} [D_g(\hat{x}, x^0) - D_g(x^j, x^0)] = 0.$$

Thus, $\lim\limits_{j\to\infty} D_g(\hat{x}, x^j) = 0$. Then, (H2) ensures $x^j \to \hat{x}$, i.e., the convergence is strong.

Remark 4.1. We compare Theorem 4.4 with Theorem 5.5 of [17] in two folds: First, the two conditions in Theorem 5.5 of [17] that $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$ and either g satisfies H6 or EP(f, K) has a unique solution are removed, where

H6: If $\{y^j\}$ and $\{z^j\}$ are sequences in K that converges weakly to y and z, respectively and $y \ne z$, then

$$\liminf_{k\to\infty} |\langle g'(y^j) - g'(z^j), y - z \rangle| > 0.$$

Second, we prove the strong convergence of $\{x^j\}$ to the solution of (1.1) which is closest to the initial point x^0 in the sense of Bregman distance by adapting the ideas of [27] whereas the weak convergence is obtained in Theorem 5.5 of [17].

Remark 4.2. Compared with item (ii) of Proposition 4.3 and Theorem 4.4 in [23], the condition that $f(\cdot, y)$ is weakly upper semicontinuous for all $y \in K$ is removed; Compared with item (iii) of Proposition 4.3 and Theorem 4.4 in [23], one of combination conditions (A2 and A3, i.e., undermonotonicity and pseudomonotonicity) used in Theorem 4.4 in this paper is weaker than monotonicity presented in [23]. Besides this, we also provide other combination conditions (A2 and A4' or A2 and A4*) that guarantee the results hold. Moreover, there is a slight difference between Theorem 4.4 in this paper and Theorem 4.4 of [23], i.e., we do not directly make the assumption that $S = S^d$.

5. Application to Variational Inequality

The aim of this section is to apply the approach of the previous section to variational inequality problem, denoted by VIP, consisting of finding $x^* \in K$ such that

$$\langle F(x^*), y - x^* \rangle \ge 0, \quad \forall y \in K, \tag{5.1}$$

where $F : K \to \mathbb{B}^*$ is a nonlinear mapping and K is a nonempty closed and convex subset of $\mathbb{B}$, which was considered by many researchers (see, for example, [22, 12]), when F is monotone on K. However, rather few algorithms have been developed for (5.1) when F is not monotone, see, for example, [3, 8, 29].

Let $F : K \to \mathbb{B}^*$ be a nonlinear mapping with K a nonempty closed and convex subset of $\mathbb{B}$. If we set $f(x, y) = \langle F(x), y - x\rangle$, then all (generalized) monotonicity concepts on bifunction f in Definition 2.1 reduce to the familiar ones on mapping F. We recall them as follows:

Definition 5.1. Let $\mathbb{B}$ be a reflexive Banach space and $K \subset \mathbb{B}$ be a nonempty closed and convex set. Let $F : K \to \mathbb{B}^*$ be a nonlinear mapping. Then F is said to be

A1-F: monotone, if

$$\langle F(x) - F(y), x - y\rangle \geq 0, \quad \forall x, y \in K,$$

A2-F: undermonotone, if there exists $\theta > 0$ such that

$$\langle F(x) - F(y), x - y\rangle \geq -\theta\langle g'(x) - g'(y), x - y\rangle, \quad \forall x, y \in K,$$

where θ is called the undermonotonicity constant of F,

A3-F: pseudomonotone, if

$$\langle F(x), y - x\rangle \geq 0 \Rightarrow \langle F(y), y - x\rangle \geq 0, \quad \forall x, y \in K,$$

A4-F: properly quasimonotone, if for all $x^1, \cdots, x^n \in K$ and all $\lambda_1, \cdots, \lambda_n \geq 0$ such that $\sum_{i=1}^{n} \lambda_i = 1$ it holds that

$$\min_{1\leq i\leq n} \langle F(x^i), \sum_{j=1}^{n} \lambda_j x^j - x^i\rangle \leq 0. \tag{5.2}$$

A4'-F: f satisfies A4-F, with strict inequality in (5.2) if the x^is are pairwise distinct and the λ_is are all strictly positive,

A4*-F: For all $x^1, \cdots, x^n \in K$ and all $\lambda_1, \cdots, \lambda_n \geq 0$ such that $\sum_{i=1}^{n} \lambda_i = 1$ it holds that

$$\sum_{i=1}^{n} \lambda_i \langle F(x^i), \sum_{j=1}^{n} \lambda_j x^j - x^i\rangle \leq 0.$$

If $f(x, y) = \langle F(x), y - x\rangle$ in Algorithm 3.1, then (3.2) with (3.1) reduces to

$$\langle F(\tilde{x}^j) + \gamma_j[g'(\tilde{x}^j) - g'(x^j)] - e^j, y - \tilde{x}^j\rangle \geq 0, \quad \forall y \in K. \tag{5.3}$$

We denote the algorithm consisting of (5.3) and (3.3)-(3.7) as Algorithm 5.1.

Remark 5.1. We compare Algorithm 5.1 with Algorithm 3.1 of [8] as follows: If we $x^{j+1} = \tilde{x}^j$ in (5.3) instead of (3.4)-(3.7), in addition, $e^j = 0$ in (5.3), then Algorithm 5.1 reduces to Algorithm 3.1 of [8], which is known as exact proximal point algorithm for VIP.

We present the convergence theorem of Algorithm 5.1 for VIP as follows.

Theorem 5.1. *Consider VIP. Let $F : K \to \mathbb{B}^*$ be a continuous mapping. Assume that F satisfies A2-F and any one of A3-F, A4'-F and A4*-F. Take $g : \mathbb{B} \to \mathbb{R}$ satisfies H1-H5 and an exogenous sequence $\{\gamma_j\} \subset (\theta, \bar{\gamma}]$ for some $\bar{\gamma} > \theta$, where θ is the undermonotonicity constant in A2-F. Let $\{x^j\}$ be the sequence by Algorithm 5.1. If VIP has solutions, then*

(a) *$\{\tilde{x}^j\}$ is an asymptotically solving sequence for VIP.*
(b) *all weak cluster points of $\{x^j\}$ solve VIP.*
(c) *$\{D_g(x^j, x^0)\}$ converges to $\inf_{z \in S(VIP)} D_g(z, x^0)$ and $\{x^j\}$ converges strongly to $\hat{x} =$ argmin$_{z \in S(VIP)} D_g(z, x^0)$.*

Proof. It is not hard to see that Algorithm 5.1 is deduced by applying Algorithm 3.1 to $f(x, y) := \langle F(x), y - x\rangle$. If F is continuous on K, then $f(x, y) := \langle F(x), y - x\rangle$ satisfies P1-P3. Thus, Theorem 4.4 ensures this theorem.

Remark 5.2. We compare Theorem 5.1 with Theorem 4.1 of [8] in the following aspects:

(i) Strong convergence is obtained in Banach spaces in Theorem 5.1, while convergence result of Theorem 4.1 of [8] is in the setting of finite dimensional sapces.
(ii) The conditions on F are any one of three combination ones: (A2-F)-(A3-F), (A2-F)-(A4'-F) and (A2-F)-(A4*-F) in Theorem 5.1. If $\mathbb{B} = \mathbb{R}^n$, the first combination conditions (A2-F)-(A3-F) above reduce to the ones on mapping in Theorem 4.1 of [8]. That is, Theorem 5.1 shows that the convergence results also hold under the one of other two assumptions (A2-F)-(A4'-F) and (A2-F)-(A4*-F) besides (A2-F)-(A3-F). Thus, Theorem 5.1 generalizes and extends Theorem 4.1 of [8] from finite dimensional spaces to reflexive Banach spaces.

References

1. P. N. Anh and H. A. Le Thi, An Armijo-type method for pseudomonotone equilibrium problems and its applications, *J. Glob. Optim.*, (2012), doi: 10.1007/s10898-012-9970-8.
2. P. N. Anh and J. K. Kim, Outer approximation algorithms for pseudomonotone equilibrium problems, *Comput. Math. Appl.*, **61** (2011), 2588-2595.

3. D. Aussel and N. Hadjisavvas, On quasimonotone variational inequalities, *J. Optim. Theory Appl.*, **121** (2004), 445-450.

4. M. Bianchi and R. Pini, A note on equilibrium problems with properly quasimonotone bifunctions, *J. Global Optim.*, **20** (2001), 67-76.

5. M. Bianchi and R. Pini, Coercivity conditions for equilibrium problems, *J. Optim. Theory Appl.*, **124** (2005), 79-92.

6. M. Bianchi and S. Schaible, Generalized monotone bifunctions and equilibrium problems, *J. Optim. Theory Appl.*, **90** (1996), 31-43.

7. R. Burachik and G. Kassay, On a generalized proximal point method for solving equilibrium problems in Banach spaces, *Nonlinear Anal.*, **75** (2012), 6456-6464.

8. N. E. Farouq, Pseudomonotone variational inequalities: convergence of proximal methods, *J. Optim. Theory Appl.*, **109** (2001), 311-326.

9. F. Flores-Bazán, Existence theorems for generalized noncoercive equilibrium problems: quasiconvex case, *SIAM J. Optim.*, **11** (2000), 675-790.

10. F. Flores-Bazán, Existence theory for finite dimensional pseudomonotone equilibrium problems, *Acta Applicandae Mathematicae*, **77** (2003), 249-297.

11. O. Güler, On the convergence of the proximal point algorithm for convex minimization, *SIAM J. Control Optim.*, **29** (1991), 403-409.

12. B. S. He, Z. H. Yang and X. M. Yuan, An approximate proximal-extragradient type method for monotone variational inequalities, *J. Math. Anal. Appl.*, **300** (2004), 362-374.

13. A. N. Iusem, G. Kassay and W. Sosa, On certain conditions for the existence of solutions of equilibrium problems, *Math. Program.*, **116** (2009), 259-273.

14. A. N. Iusem and W. Sosa, New existence results for equilibrium problems, *Nonlinear Analysis*, **52** (2003), 621-635.

15. A. N. Iusem and W. Sosa, On the proximal point method for equilibrium problems in Hilbert spaces, *Optimization*, **59** (2010), 1259-1274.

16. A. N. Iusem and W. Sosa, Iterative algorithms for equilibrium problems, *Optimization*, **52** (2003), 301-316.

17. A. N. Iusem and M. Nasri, Inexact proximal point methods for equilibrium problems in Banach spaces, *Numer. Funct. Anal. Optim.*, **28** (2007), 1279-1308.

18. A. N. Iusem and Gárciga Otero, Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, *Numer. Funct. Anal. Optim.*, **22** (2001), 609-640.

19. S. Karamardian and S. Schaible, Seven kinds of monotone maps, *J. Optim. Theory Appl.*, **66** (1990), 36-46.

20. J. K. Kim, P. N. Anh and H. G. Hyun, A proximal point-type algorithm for pseudomonotone equilibrium problems, *Bull. Korean Math. Soc.*, **49** (2012), 749-759.

21. M. A. Krasnoselskii, Two observations about the method of succesive approximations, *Uspekhi Matematicheskikh Nauk*, **10** (1955), 123-127.

22. M. Li and M. H. Xu, Comparison of two proximal point algorithms for monotone variational inequalities, *Comput. Math. Appl.*, **52** (2006), 1543-1554.

23. J. Mashreghi and M. Nasri, Strong convergence of an inexact proximal point algorithm for equilibrium problems in Banach spaces, *Numer. Funct. Anal. Optim.*, **31** (2010), 1053-1071.

24. J. Moreau, Proximité et dualité dans un espace hilbertien, *Bulletin de la Societé Mathématique de France*, **93** (1965), 273-299.

25. A. Moudafi, On the convergence of splitting proximal methods for equilibrium problems in Hilbert spaces, *J. Math. Anal. Appl.*, **359** (2009), 508-513.

26. A. Moudafi, Proximal methods for a class of bilevel monotone equilibrium problems, *J. Glob. Optim.*, **47** (2010), 287-292.

27. R. G. Otero and B. F. Svaiter, A strongly convergent hybrid proximal method in Banach spaces, *J. Math. Anal. Appl.*, **289** (2004), 700-711.

28. R. T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM J. Control Optim.*, **14** (1976), 877-898.

29. S. Schaible, J. C. Yao and L. C. Zeng, A proximal method for pseudomonotone type variational-like inequalities, *Taiwan. J. Math.*, **10** (2006), 497-513.

30. J. X. Zhou and G. Y. Chen, Diagonal convexity conditions for problems in convex analysis and quasivariational inequalities, *J. Math. Anal. Appl.*, **132** (1988), 213-225.

Guo-ji Tang
School of Science
Guangxi University for Nationalities
Nanning, Guangxi 530006
P. R. China
E-mail: guojvtang@126.com

Xing Wang
School of Information Technology
Jiangxi University of Finance and Economics
Nanchang 330013
P. R. China