Nonlinear Functional Analysis and Applications
Vol. 28, No. 2 (2023), pp. 337-363
ISSN: 1229-1595(print), 2466-0973(online)

https://doi.org/10.22771/nfaa.2023.28.02.02
http://nfaa.kyungnam.ac.kr/journal-nfaa

THE SUBGRADIENT EXTRAGRADIENT METHOD FOR SOLVING MONOTONE BILEVEL EQUILIBRIUM PROBLEMS USING BREGMAN DISTANCE

Roushanak Lotfikar[1], Gholamreza Zamani Eskandani[2] and Jong Kyu Kim[3]

[1]Faculty of Basic Science, Ilam University,
P. O. Box 69315-516, Ilam, Iran
e-mail: `r.lotfikar@ilam.ac.ir`

[2]Department of Pure Mathematics, Faculty of Mathematical Sciences,
University of Tabriz, Tabriz, Iran
e-mail: `zamani@tabrizu.ac.ir`

[3]Department of Mathematics Education, Kyungnam University,
Changwon Gyeongnam, 51767, Korea
e-mail: `jongkyuk@kyungnam.ac.kr`

Abstract. In this paper, we propose a new subgradient extragradient algorithm for finding a solution of monotone bilevel equilibrium problem in reflexive Banach spaces. The strong convergence of the algorithm is established under monotone assumptions of the cost bifunctions with Bregman Lipschitz-type continuous condition. Finally, a numerical experiments is reported to illustrate the efficiency of the proposed algorithm.

1. Introduction

Let X be a reflexive real Banach space and C be a nonempty, closed and convex subset of X. Throughout this paper, we shall denote the dual space of X by X^*. The norm and the duality pairing between X and X^* are respectively

[0]Received June 8, 2022. Revised October 11, 2022. Accepted November 2, 2022.
[0]2020 Mathematics Subject Classification: 47H05, 47H09, 47H10, 47J25, 65K10, 65K15.
[0]Keywords: Bilevel equilibrium problem, Bregman Lipschitz-type condition, Strongly monotone, subgradient extragradient method.
[0]Corresponding author: R. Lotfikar(`r.lotfikar@ilam.ac.ir`).

denoted by $\|.\|$ and $\langle .,. \rangle$, and $\mathbb{R}$ stands for the set of all real numbers. Let $g{:}C{\times}C{\to}\mathbb{R}$ be a bifunction satisfying the condition $g(x,x){=}0$ for every $x{\in}C$.

The equilibrium problem, shortly $EP(C,g)$, for the bifunction g is stated as follows:

$$\text{Find } \hat{x} \in C \quad \text{such that} \;\; g(\hat{x},y) \geq 0, \;\; \forall y \in C.$$

The solution set of $EP(C,g)$ is denoted by $Sol(C,g)$ ([23, 24]).

In this paper, we consider the equilibrium problem whose constraints are the solution sets of another equilibrium problem which usually is called bilevel equilibrium problem, shortly BEPs. That is

$$\text{Find } x^* \in Sol(C,g) \quad \text{such that} \;\; f(x^*,y) \geq 0, \;\; \forall y \in Sol(C,g),$$

where $f : C \times C \to \mathbb{R} \cup \{+\infty\}$ such that $f(x,x) = 0$ for every $x \in C$.

The bilevel equilibrium problems were introduced by Chadli et al. [16] in 2000. This kind of problems is very important and interesting because it is a generalization class of problems such as optimization problems over equilibrium constraints, variational inequality over equilibrium constraints, hierarchical minimization problems, and complementarity problems. Furthermore, the particular case of the bilevel equilibrium can be applied to a real word model such as the variational inequality over the fixed point set of a firmly nonexpansive mapping applied to the power control problem of CDMA networks which were introduced by Iiduka [20]. For more on the relation of bilevel equilibrium with particular cases, see [19, 21, 29].

Methods for solving BEPs have been studied extensively by many authors. In 2010, Moudafi [27] introduced a simple proximal method and proved the weak convergence to a solution of BEPs. In 2014, Quy [31] introduced an algorithm by combining the proximal method with the Halpern method for solving bilevel monotone equilibrium and fixed point problem. For more details and most recent works on the methods for solving BEPs, we refer the reader to [9, 17, 36]. The authors considered the method for monotone and pseudo monotone equilibrium problem. If a bifunction is more generally monotone, we cannot use the above methods for solving BEPs.

The extragradient method first introduced by Quoc et al. [30] in which only two strongly auxiliary convex problems onto the feasible set C are performed. The extragradient method is written as follows:

$$\begin{cases} x_0 \in C, \\ y_k = \arg\min\{\lambda g(x_k,x) + \frac{1}{2}\|x - x_k\|^2 : \, x \in C\}, \\ x_{k+1} = \arg\min\{\lambda g(y_k,x) + \frac{1}{2}\|y - x_k\|^2 : \, y \in C\}. \end{cases}$$

In the special case, if $g(x,y) = \langle G(x), y-x\rangle$ where $G : C \to H$ is a cost mapping then the problem $EP(C,g)$ becomes the variational inequality $VI(C,G)$ ([2]): Finding $x^* \in C$ such that

$$\langle G(x^*), x - x^*\rangle \geq 0, \quad \forall x \in C.$$

Recently, inspired by the extragradient method, Censor et al. [15] presented an algorithm, which is called the subgradient extragradient method, for solving $VI(C,G)$ in Hilbert spaces (see, [1, 28]). The Censor's subgradient extragradient method is as the following:

$$\begin{cases} x_0 \in H, \\ y_k = P_C\big(x_k - \lambda G(x_k)\big), \\ T_K = \big\{w \in H : \langle x_k - \lambda G(x_k) - y_k, w - y_k\rangle \leq 0\big\}, \\ x_{k+1} = p_{T_k}\big(x_k - \lambda G(y_k)\big), \end{cases} \tag{1.1}$$

where P_C is the metric projection from H to C. Under certain assumptions, the weakly convergence of the sequences $\{x_n\}$ has been established.

Recently, Anh et al. [4] presented the following subgradient extragradient method for solving monotone BEPs:

$$\begin{cases} x_0 \in C, \\ y_k = \arg\min\{\lambda_k g(x_k, y) + \frac{1}{2}\|y - x_k\|^2 : y \in C\}, \\ z_k = \arg\min\{\lambda_k g(y_k, z) + \frac{1}{2}\|z - x_k\|^2 : z \in T_k\}, \\ x_{k+1} = \arg\min\{\beta_k f(z_k, t) + \frac{1}{2}\|t - z_k\|^2 : t \in C\}, \end{cases} \tag{1.2}$$

where $T_k = \{v \in H : \langle x_k - \lambda_k w_k - y_k, v - y_k\rangle \leq 0\}$ and $w_k \in \partial_2 g(x_k, y_k)$. Under certain the conditions, they proved that $\{x_n\}$,$\{y_n\}$ and $\{z_n\}$ converge strongly to the unique solution point of BEPs.

Inspired by the above works, in this paper, using Bregman distance we propose a subgradient extragradient algorithm for finding a solution for BEPs where f is strongly monotone and Bregman Lipschitz-type continuous and g is pseudomonotone in reflexive Banach space X.

The paper is organized as follows: In Section 2, we recall some definitions and preliminaries that would be needed in the paper. The third section proposes a new algorithm and analyzes its convergence. In section 4, we will illustrate some applications for our algorithm. Finally, a numerical example for validity our main theorem will be exposed.

2. Preliminaries

Let $f : X \to (-\infty, \infty]$ be a proper convex and lower semicontinuous function. The set of minimizers of f is denoted by Argmin$f$. If Argmin$f$ is a singleton, its unique element is denoted by $\arg\min_{x\in X} \; f(x)$. Also we denote by dom f, the domain of f, that is the set $\{x \in X : f(x) < \infty\}$. Let $x \in \text{int dom} f$. Then subdifferential of f at x is the convex set defined by:

$$\partial f(x) = \{\xi \in X^* : f(x) + \langle y - x, \xi\rangle \leq f(y), \;\; \forall y \in X\},$$

and the Fenchel conjugate of f is the convex function

$$f^* : X^* \to (-\infty, \infty], \qquad f^*(\xi) = \sup\{\langle x, \xi\rangle - f(x) : x \in X\}.$$

It is well known that $\xi \in \partial f(x)$ is equivalent to

$$f(x) + f^*(\xi) = \langle x, \xi\rangle. \tag{2.1}$$

It is not difficult to check that f^* is proper convex and lower semicontinuous function. The function f is said to be cofinite if dom $f^* = X^*$.

For any convex mapping $f : X \to (-\infty, +\infty]$, we denote by $f^\circ(x, y)$ the right-hand derivative of f at $x \in \text{int dom} f$ in the direction y, that is,

$$f^\circ(x, y) := \lim_{t\downarrow 0} \frac{f(x + ty) - f(x)}{t}. \tag{2.2}$$

If the limit as $t \to 0$ in (2.2) exists for each y, then the function f is said to be Gâteaux differentiable at x. In this case, the gradient of f at x is the linear function $\nabla f(x)$, which is defined by $\langle y, \nabla f(x)\rangle := f^\circ(x, y)$ for all $y \in X$. The function f is said to be Gâteaux differentiable if it is Gâteaux differentiable at each $x \in \mathit{int\ dom} f$. When the limit as $t \to 0$ in (2.2) is attained uniformly for any $y \in X$ with $\|y\| = 1$, we say that f is Fréchet differentiable at x. Finally, f is said to be uniformly Fréchet differentiable for $x \in E$ and $\|y\| = 1$.

The function f is said to be Legendre if it satisfies the following two conditions:

(L1) int dom$f \neq \emptyset$ and ∂f is single-valued on its domain,
(L2) int dom$f^* \neq \emptyset$ and ∂f^* is single-valued on its domain.

Because the space X is assumed to be reflexive, we always have $(\partial f)^{-1}{=}\partial f^*$ (see[10], p. 83). This fact, when combined with the conditions (L1) and (L2), implies the following equalities:

$$\nabla f = (\nabla f^*)^{-1},$$
$$\text{ran}\; \nabla f = \text{dom}\; \nabla f^* = \text{int dom}\; f^*,$$
$$\text{ran}\; \nabla f^* = \text{dom}\; \nabla f = \text{int dom}\; f.$$

Also, the conditions (L1) and (L2), in conjunction with Theorem 5.4 of [8], imply that the functions f and f^* are strictly convex on the interior of their respective domains and f is Legendre if and only if f^* is Legendre. Several interesting examples of Legendre functions are presented in [6, 8]. Among them are the functions $\frac{1}{p}\|\cdot\|^p$ with $p \in (1,\infty)$, where the Banach space X is smooth and strictly convex.

In 1967, Bregman [11] introduced the concept of Bregman distance, and he discovered an elegant and effective technique for the use of the Bregman distance in the process of designing and analyzing feasibility and optimization algorithms.

From now on, we assume that $f : X \to (-\infty,+\infty]$ is also Legendre. The Bregman distance with respect to f, or simply, Bregman distance is the bi-function $D_f : \mathrm{dom} f \times \mathrm{int}\, \mathrm{dom} f \to [0,+\infty]$, defined by:

$$D_f(y,x) := f(y) - f(x) - \langle y - x, \nabla f(x)\rangle.$$

It should be noted that D_f is not a distance in the usual sense of the term. Clearly, $D_f(x,x) = 0$, but $D_f(y,x) = 0$ may not imply $x = y$. In our case, when f is Legendre this indeed holds (see [8], Theorem 7.3(vi), p. 642). In general, D_f is not symmetric and does not satisfy the triangle inequality. However, D_f satisfies the three point identity

$$D_f(x,y) + D_f(y,z) - D_f(x,z) = \langle x - y, \nabla f(z) - \nabla f(y)\rangle,$$

and four point identity

$$D_f(x,y) + D_f(w,z) - D_f(x,z) - D_f(w,y) = \langle x - w, \nabla f(z) - \nabla f(y)\rangle,$$

for any $x, w \in \mathrm{dom} f$ and $y, z \in \mathrm{int}\, \mathrm{dom} f$. During the last 30 years, Bregman distances have been studied by many researchers (see [7, 8, 12, 13, 25]).

We will use the following lemmas in the proof of our results.

Lemma 2.1. ([32]) *If $f : X \to \mathbb{R}$ is uniformly Fréchet differentiable and bounded on bounded subsets of X, then ∇f is uniformly continuous on bounded subsets of X from the strong topology of X to the strong topology of X^*.*

Recall that the function f is called sequentially consistent (see [14]) if for any two sequences $\{x_n\}$ and $\{y_n\}$ in X such that the first one is bounded, if

$$\lim_{n\to\infty} D_f(y_n, x_n) = 0 \quad \Rightarrow \quad \lim_{n\to\infty} \|y_n - x_n\| = 0.$$

Lemma 2.2. ([13]) *If* dom f *contains at least two points, then the function f is totally convex on bounded sets if and only if the function f is sequentially consistent.*

Lemma 2.3. ([33]) *Let $f : X \to \mathbb{R}$ be a Gâteaux differentiable and totally convex function. If $x_0 \in X$ and the sequence $\{D_f(x_n, x_0)\}$ is bounded, then the sequence $\{x_n\}$ is also bounded.*

Lemma 2.4. ([34]) *Let $f : X \to \mathbb{R}$ be a Legendre function such that ∇f^* is bounded on bounded subsets of* $\operatorname{int} \operatorname{dom} f^*$. *Let $x_0 \in X$, if $\{D_f(x_0, x_n)\}$ is bounded then the sequence $\{x_n\}$ is bounded too.*

Lemma 2.5. ([37]) *Let C be a nonempty convex subset of X and $f : C \to \mathbb{R}$ be a convex and subdifferentiable function on C. Then f attains its minimum at $x \in C$ if and only if $0 \in \partial f(x) + N_C(x)$, where $N_C(x)$ is the normal cone of C at x, that is*

$$N_C(x) := \{x^* \in X^* : \langle x - z, x^* \rangle \geq 0, \ \ \forall z \in C\}.$$

Lemma 2.6. ([18]) *If f and g are two convex functions on X such that there is a point $x_0 \in \operatorname{dom} f \cap \operatorname{dom} g$ where f is continuous, then*

$$\partial(f + g)(x) = \partial f(x) + \partial g(x), \quad \forall x \in X.$$

A function $g : C \times C \to (-\infty, +\infty]$, where C is a closed and convex subset of X, such that $g(x, x) = 0$ for all $x \in C$ is called a bifunction. Throughout this paper we consider bifunctions with the following properties:

(A1) g is pseudomonotone on C, that is, for all $x, y \in C$,
$$g(x, y) \geq 0 \ \ \Rightarrow \ \ g(y, x) \leq 0.$$
(A2) g is Bregman-Lipschitz-type continuous on C, that is, there exist two positive constants c_1 and c_2 such that
$$g(x, y) + g(y, z) \geq g(x, z) - c_1 D_f(y, x) - c_2 D_f(z, y), \quad \forall x, y, z \in C,$$
where $f : X \to (-\infty, +\infty]$ is a Legendre function. The constants c_1 and c_2 are called Bregman-Lipschitz coefficients with respect to f.
(A3) g is weakly continuous on $C \times C$, that is, if $x, y \in C$ and $\{x_n\}$ and $\{y_n\}$ are two sequences in C converging weakly to x and y, respectively, then $g(x_n, y_n) \to (x, y)$.
(A4) $g(x, .)$ is convex, lower semicontinuous and subdifferentiable on C for every fixed $x \in C$.
(A5) For each $x, y, z \in C$, $\limsup_{t \downarrow 0} g(tx + (1 - t)y, z) \leq g(y, z)$.
(A6) f is η-strongly monotone on C, that is, for all $x, y \in C$,
$$f(x, y) + f(y, x) \leqslant -\eta \|x - y\|^2.$$
(A7) g is monotone on C, that is, if for all $x, y \in C$, $g(x, y) + g(y, x) \leq 0$.

It is obvious that any monotone bifunction is a pseudomonotone one, but not vice versa. A mapping $A : C \to X^*$ is pseudomonotone if and only if the bifunction $g(x, y) = \langle A(x), y - x \rangle$ is pseudomonotone on C (see [38]).

Lemma 2.7. ([40]) *If a bifunction g satisfying the conditions $A1$, $A3$-$A5$, then $Sol(C, g)$ is closed and convex.*

We need the following technical lemmas.

Lemma 2.8. ([26]) *Let $\{a_n\}$ be a sequence of nonnegative real numbers. Suppose that for any integer m, there exists an integer p such that $p \geq m$ and $a_p \leq a_{p+1}$. Let k_0 be an integer such that $a_{k_0} \leq a_{k_0+1}$ and define, for all integer $k \geq k_0$,*

$$\tau(k) = \max\{i \in \mathbb{N} : k_0 \leq i \leq k,\, a_i \leq a_{i+1}\}.$$

Then, $0 \leq a_k \leq a_{\tau(k)+1}$ for all $k \geq k_0$. Furthermore, the sequence $\{\tau(k)\}_{k \geq k_0}$ is nondecreasing and tends to $+\infty$ as $k \to \infty$.

Lemma 2.9. ([5]) *Let X and Y be two sets, G be a compact set-valued map from Y to X, and W be a real valued function defined on $X \times Y$. The marginal function M is defined by*

$$M(y) = \big\{x^* \in G(y) : \, W(x^*, y) = \sup\{W(x, y) : \, x \in G(y)\}\big\}.$$

If W and G are continuous, then M is upper semicontinuous.

Let us assume that $H : X \to (-\infty, +\infty]$ is also Legendre and two bifunctions $f : C \times C \to \mathbb{R} \cup \{+\infty\}$ and $g : C \times C \to \mathbb{R} \cup \{+\infty\}$ satisfy the following conditions:

Assumptions on g:

- (g_1) $Sol(C, g)$ is nonempty.
- (g_2) g is monotone and Bregman Lipschitz-type continuous with constants c_1, c_2 and g is weakly continuous.

Assumptions on f:

- (f_1) f is weakly continuous and Bregman η-strongly monotone on C, i.e., for all $x, y \in C$, $f(x,y) + f(y,x) \leqslant -\eta\Big(D_H(y,x) + D_H(x,y)\Big)$.
- (f_2) There exists the mappings $\overline{f_i} : C \times C \to X$ and $\hat{f}_i : C \to X$ for each $i \in \{1, \cdots, m\}$ such that $\overline{f_i}(x,y) + \overline{f_i}(y,x) = 0$, $\|\overline{f_i}(x,y)\|^2 \leqslant \overline{L_i}^2 D_H(y,x)$ and $\|\nabla \hat{f}_i(x) - \nabla \hat{f}_i(y)\|^2 \leqslant \hat{L_i}^2 D_H(y,x)$ for all $x, y \in C$, also

$$f(x,y) + f(y,z) \geqslant f(x,z) + \sum_{i=1}^{m} \Big\langle \overline{f_i}(x,y), \nabla \hat{f}_i(y) - \nabla \hat{f}_i(z) \Big\rangle, \ \forall x, y, z \in C.$$

- (f_3) For any sequence $\{y_k\} \subset C$ such that $x_k \to d$, we have

$$\limsup_{k \to \infty} \frac{|f(d, y_k)|}{\sqrt{D_H(y_k, d)}} < +\infty.$$

Remark 2.10. Suppose that the bifunction f satisfies the condition (f_2). Then

$$\begin{aligned} f(x,y)+f(y,z) &\geqslant f(x,z)+\sum_{i=1}^{m}\left\langle \overline{f_i}(x,y),\nabla\hat{f}_i(y)-\nabla\hat{f}_i(z)\right\rangle \\ &\geqslant f(x,z)-\sum_{i=1}^{m}\left|\left\langle \overline{f_i}(x,y),\nabla\hat{f}_i(y)-\nabla\hat{f}_i(z)\right\rangle\right| \\ &\geqslant f(x,z)-\sum_{i=1}^{m}\left\|\overline{f_i}(x,y)\right\|\left\|\nabla\hat{f}_i(y)-\nabla\hat{f}_i(z)\right\| \\ &\geqslant f(x,z)-\sum_{i=1}^{m}\overline{L}_i\hat{L}_i\sqrt{D_H(z,y)}\sqrt{D_H(y,x)} \\ &\geqslant f(x,z)-\frac{1}{2}\sum_{i=1}^{m}\overline{L_i}\hat{L}_iD_H(z,y)-\frac{1}{2}\sum_{i=1}^{m}\overline{L}_i\hat{L}_iD_H(y,x) \\ &= f(x,z)-c_1D_H(z,y)-c_2D_H(y,x). \end{aligned}$$

Thus, f is Bregman Lipschitz-type continuous with constants $c_1=c_2=\frac{1}{2}\sum_{i=1}^{m}\overline{L}_i\hat{L}_i$.

3. Main results

In this section, assume that Assumptions on f and g hold. We first establish some lemmas and then propose a subgradient extragradient algorithm for finding a solution of BEPs. We also assume that the bifunction f is η-strongly monotone and Bregman Lipschitz-type continuous and g is pseudomonotone.

Algorithm A:

Initialization: Choose $x_0\in C$, the tolerance $\epsilon>0$, $s:=\sum_{i=1}^{m}\overline{L}_i\hat{L}_i$ and the sequences $\{\lambda_k\}$ and $\{\beta_k\}$ such that

$$\begin{cases} \{\lambda_k\}\subset(a,b)\subset(0,\min\{\frac{1}{c_1},\ \frac{1}{c_2}\}),\ \lim_{k\to\infty}\lambda_k=\lambda, \\ \beta_k\searrow 0,\ \ 1-\beta_k\eta+\beta_k^2s^2<1,\ \sum_{k=1}^{\infty}\beta_k=+\infty, \\ 0<\tau<\min\{\eta,s\},\ 0<\beta_k<\min\left\{\frac{1}{\tau},\frac{2\eta-2\tau}{s^2-\tau^2},\frac{2\eta}{s^2}\right\}. \end{cases} \tag{3.1}$$

Set $k=0$ and go to Step 1.

Step 1. Compute $(k=0,1,...)$

$$y_k = \arg\min\Big\{\lambda_k g(x_k, y) + D_H(y, x_k) : y \in C\Big\},$$

$$z_k = \arg\min\Big\{\lambda_k g(y_k, z) + D_H(z, x_k) : z \in T_k\Big\},$$

$$x_{k+1} = \arg\min\Big\{\beta_k f(z_k, t) + D_H(t, z_k) : t \in C\Big\},$$

where $T_k = \{v \in X \mid \langle v - y_k, \nabla H(x_k) - \lambda_k w_k - \nabla H(y_k)\rangle \leqslant 0\}$ and $w_k \in \partial_2 g(x_k, y_k)$.

Step 2. If $\max\big\{D_H(y_k, x_{k+1}), D_H(y_k, x_k)\big\} \leqslant \epsilon$ then stop.

Step 3. Otherwise, set $k := k+1$ and go back to Step 1.

To prove the convergence of Algorithm A, we need the following lemmas.

Lemma 3.1. ([3, Lemma 1]) *If* $x_k = y_k$ *then* $x_k \in \mathrm{Sol}(C, g)$.

Lemma 3.2. *Let* $\overline{x} \in \mathrm{Sol}(C, g)$. *Then*

$$D_H(\overline{x}, z_k) \leqslant D_H(\overline{x}, x_k) - (1 - \lambda_k c_1) D_H(y_k, x_k) - (1 - \lambda_k c_2) D_H(z_k, y_k).$$

Proof. Since $y_k = \arg\min\{\lambda_k g(x_k, y) + D_H(y, x_k) : y \in C\}$, by Lemmas 2.5 and 2.6, we have

$$0 \in \lambda_k \partial_2 g(x_k, y_k) + \nabla_1 D_H(y_k, x_k) + N_C(y_k).$$

Hence, there exists $w_k \in \partial_2 g(x_k, y_k)$ such that

$$\nabla H(x_k) - \lambda_k w_k - \nabla H(y_k) \in N_C(y_k).$$

Then, by the definition of $N_C(y_k)$, we get

$$\Big\langle x - y_k, \nabla H(x_k) - \lambda_k w_k - \nabla H(y_k)\Big\rangle \leqslant 0, \ \forall x \in C. \tag{3.2}$$

Using the definition of $w_k \in \partial_2 g(x_k, y_k)$, we obtain

$$\lambda_k\big(g(x_k, x) - g(x_k, y_k)\big) \geqslant \big\langle x - y_k, \lambda_k w_k\big\rangle. \tag{3.3}$$

Adding (3.2) and (3.3), we get

$$\lambda_k\big(g(x_k, x) - g(x_k, y_k)\big) + \big\langle x - y_k, \nabla H(y_k) - \nabla H(x_k)\big\rangle \geqslant 0, \ \forall x \in C. \tag{3.4}$$

Since $z_k \in T_k$, from the definition of T_k, we have

$$\Big\langle z_k - y_k, \nabla H(x_k) - \lambda_k w_k - \nabla H(y_k)\Big\rangle \leqslant 0.$$

Therefore

$$\big\langle z_k - y_k, \nabla H(x_k) - \nabla H(y_k)\big\rangle \leqslant \big\langle z_k - y_k, \lambda_k w_k\big\rangle. \tag{3.5}$$

Replacing $z_k = x$ into (3.3), we get

$$\lambda_k\big(g(x_k, z_k) - g(x_k, y_k)\big) \geqslant \big\langle z_k - y_k, \lambda_k w_k\big\rangle,$$

from (3.5), we obtain

$$\lambda_k\big(g(x_k, z_k) - g(x_k, y_k)\big) \geqslant \big\langle z_k - y_k, \nabla H(x_k) - \nabla H(y_k)\big\rangle. \tag{3.6}$$

Similarly, since $z_k = \arg\min\{\lambda_k g(y_k, z) + D_H(z, x_k) : z \in T_k\}$, there exists $q_k \in \partial_2 g(y_k, z_k)$ and $h_k \in N_{T_k}(z_k)$ such that

$$0 = \lambda_k q_k + \nabla H(z_k) - \nabla H(y_k) + h_k,$$

and hence

$$\begin{aligned}\big\langle y - z_k, \nabla H(x_k) - \nabla H(z_k)\big\rangle &= \big\langle y - z_k, h_k + \lambda_k q_k\big\rangle \\ &= \big\langle y - z_k, h_k\big\rangle + \big\langle y - z_k, \lambda_k q_k\big\rangle \\ &\leqslant \big\langle y - z_k, \lambda_k q_k\big\rangle, \quad \forall y \in T_k.\end{aligned}$$

It is easy to show that $C \subseteq T_k$. Substituting $y = \overline{x} \in C \subseteq T_k$ into two last inequalities and adding them, we get

$$\lambda_k\big(g(y_k, \overline{x}) - g(y_k, z_k)\big) \geqslant \big\langle \overline{x} - z_k, \nabla H(x_k) - \nabla H(z_k)\big\rangle.$$

Note that g is monotone, $\overline{x} \in \mathrm{Sol}(C, g)$ and $y_k \in C$, we have $g(y_k, \overline{x}) \leqslant 0$ and

$$-\lambda_k g(y_k, z_k) \geqslant \big\langle \overline{x} - z_k, \nabla H(x_k) - \nabla H(z_k)\big\rangle.$$

Since g is Bregman Lipschitz-type continues, we have

$$g(x_k, y_k) + g(y_k, z_k) \geqslant g(x_k, z_k) - c_1 D_H(y_k, x_k) - c_2 D_H(z_k, y_k),$$

from (3.6), we get

$$\begin{aligned}\big\langle z_k - \overline{x}, \nabla H(x_k) - \nabla H(z_k)\big\rangle &\geqslant \lambda_k g(y_k, z_k) \\ &\geqslant \lambda_k\Big(g(x_k, z_k) - g(x_k, y_k)\Big) - \lambda_k c_1 D_H(y_k, x_k) \\ &\quad - \lambda_k c_2 D_H(z_k, y_k) \\ &\geqslant \Big\langle z_k - y_k, \nabla H(x_k) - \nabla H(y_k)\Big\rangle \\ &\quad - \lambda_k c_1 D_H(y_k, x_k) - \lambda_k c_2 D_H(z_k, y_k).\end{aligned}$$

Applying the three point identity, we get the desired result. □

Lemma 3.3. *For each $x \in C$, we have*

$$D_H(x, x_{k+1}) \leqslant D_H(x, z_k) - D_H(x_{k+1}, z_k) + \beta_k\big(f(z_k, x) - f(z_k, x_{k+1})\big).$$

Proof. Since $x_{k+1} = \arg\min\{\beta_k f(z_k, t) + D_H(t, z_k) : t \in C\}$, by a similar way as in the proof of Lemma 3.2, there exists $v_k \in \partial_2 f(z_k, x_{k+1})$ such that

$$0 \in \beta_k v_k + \nabla H(x_{k+1}) - \nabla H(z_k) + N_C(x_{k+1}).$$

Using the definitions of the normal cone N_C and the subgradiant v_k, we get

$$\beta_k\big(f(z_k, x) - f(z_k, x_{k+1})\big) + \big\langle x - x_{k+1}, \nabla H(x_{k+1}) - \nabla H(z_k)\big\rangle \geqslant 0, \quad \forall x \in C. \tag{3.7}$$

Using the three point indentity, we get the desired result. □

Lemma 3.4. *Let x^* be a solution of the problem BEP_s. Then*

$$D_H(y^*_{k+1}, x_{k+1}) \leqslant \delta_k D_H(x^*, z_k) \leqslant (1-\tau\beta_k) D_H(x^*, z_k),$$

where

$$y^*_{k+1} = \arg\min\{\beta_k f(x^*, v) + D_H(\nu, x^*) : v \in C\}$$

and

$$\delta_k = 1 - \beta_k \eta + \beta_k^2 s^2.$$

Proof. Since $y^*_{k+1} = \arg\min \Big\{\beta_k f(x^*, v) + D_H(v, x^*) : v \in C\Big\}$, by a similar way as in the proof of (3.4), we get

$$\beta_k\big(f(x^*, x) - f(x^*, y^*_{k+1})\big) + \langle x - y^*_{k+1}, \nabla H(y^*_{k+1}) - \nabla H(x^*)\rangle \geqslant 0, \ \ \forall x \in C. \tag{3.8}$$

Replacing x with $y^*_{k+1} \in C$ in (3.7) and x with $x_{k+1} \in C$ in (3.8), we have

$$\beta_k\big(f(z_k, y^*_{k+1}) - f(z_k, x_{k+1})\big) + \langle y^*_{k+1} - x_{k+1}, \nabla H(x_{k+1}) - \nabla H(z_k)\rangle \geqslant 0,$$
$$\beta_k\big(f(x^*, x_{k+1}) - f(x^*, y^*_{k+1})\big) + \langle x_{k+1} - y^*_{k+1}, \nabla H(y^*_{k+1}) - \nabla H(x^*)\rangle \geqslant 0.$$

Therefore, we obtain

$$\begin{aligned}\beta_k\Big(f(z_k, y^*_{k+1}) &- f(z_k, x_{k+1}) + f(x^*, x_{k+1}) - f(x^*, y^*_{k+1})\Big)\\ &+ \langle y^*_{k+1} - x_{k+1}, \nabla H(x_{k+1}) - \nabla H(z_k)\rangle\\ &+ \langle x_{k+1} - y^*_{k+1}, \nabla H(y^*_{k+1}) - \nabla H(x^*)\rangle \geqslant 0.\end{aligned}$$

Using the three point identity, we have

$$\begin{aligned}\beta_k\Big(f(z_k, y^*_{k+1}) &- f(z_k, x_{k+1}) + f(x^*, x_{k+1}) - f(x^*, y^*_{k+1})\Big)\\ &+ D_H(y^*_{k+1}, z_k) - D_H(y^*_{k+1}, x_{k+1}) - D_H(x_{k+1}, z_k)\\ &+ D_H(x_{k+1}, x^*) - D_H(y^*_{k+1}, x^*) - D_H(x_{k+1}, y^*_{k+1}) \geqslant 0.\end{aligned} \tag{3.9}$$

Using Assumption (f_2) we get

$$f(z_k, y^*_{k+1}) - f(x^*, y^*_{k+1}) \leqslant f(z_k, x^*) - \sum_{i=1}^m \big\langle \overline{f_i}(z_k, x^*), \nabla \hat{f}_i(x^*) - \nabla \hat{f}_i(y^*_{k+1})\big\rangle$$

and

$$f(x^*, x_{k+1}) - f(z_k, x_{k+1}) \leqslant f(x^*, z_k) - \sum_{i=1}^m \big\langle \overline{f_i}(x^*, z_k), \nabla \hat{f}_i(z_k) - \nabla \hat{f}_i(x_{k+1})\big\rangle. \tag{3.10}$$

Under Assumptions (f_1) and (f_2), we have

$$\begin{aligned}
&f(z_k,y^*_{k+1})-f(x^*,y^*_{k+1})+f(x^*,x_{k+1})-f(z_k,x_{k+1})\\
&\quad\leqslant -\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&\qquad+\sum_{i=1}^{m}\big\langle\overline{f_i}(z_k,x^*),\nabla\hat{f}_i(z_k)-\nabla\hat{f}_i(x_{k+1})-\nabla\hat{f}_i(x^*)+\nabla\hat{f}_i(y^*_{k+1})\big\rangle\\
&\quad\leqslant -\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&\qquad+\sum_{i=1}^{m}\big\|\overline{f_i}(z_k,x^*)\big\|\big\|\nabla\hat{f}_i(z_k)-\nabla\hat{f}_i(x_{k+1})-\nabla\hat{f}_i(x^*)+\nabla\hat{f}_i(y^*_{k+1})\big\|\\
&\quad\leqslant -\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&\qquad+\sum_{i=1}^{m}\big\|\overline{f_i}(z_k,x^*)\big\|\Bigg(\big\|\nabla\hat{f}_i(z_k)-\nabla\hat{f}_i(x_{k+1})\big\|+\big\|\nabla\hat{f}_i(y^*_{k+1})-\nabla\hat{f}_i(x^*)\big\|\Bigg)\\
&\quad\leqslant -\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&\qquad+s\sqrt{D_H(x^*,z_k)}\Bigg(\sqrt{D_H(x_{k+1},z_k)}+\sqrt{D_H(x^*,y^*_{k+1})}\Bigg). \qquad (3.11)
\end{aligned}$$

Combining (3.9) and (3.11), we get

$$\begin{aligned}
0\leqslant&-\beta_k\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&+s\beta_k\sqrt{D_H(x^*,z_k)}\Big(\sqrt{D_H(x_{k+1},z_k)}+\sqrt{D_H(x^*,y^*_{k+1})}\Big)\\
&+D_H(y^*_{k+1},z_k)-D_H(y^*_{k+1},x_{k+1})-D_H(x_{k+1},z_k)\\
&+D_H(x_{k+1},x^*)-D_H(y^*_{k+1},x^*)-D_H(x_{k+1},y^*_{k+1}),
\end{aligned}$$

then

$$\begin{aligned}
D_H(y^*_{k+1},x_{k+1})\leqslant&-\beta_k\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&+s\beta_k\Big(\sqrt{D_H(x_{k+1},z_k)}\sqrt{D_H(x^*,z_k)}\Big)\\
&+s\beta_k\Big(\sqrt{D_H(x^*,y^*_{k+1})}\sqrt{D_H(x^*,z_k)}\Big)\\
&+D_H(y^*_{k+1},z_k)-D_H(x_{k+1},z_k)+D_H(x_{k+1},x^*)\\
&-D_H(y^*_{k+1},x^*)-D_H(x_{k+1},y^*_{k+1})
\end{aligned}$$

$$\begin{aligned}
&\leqslant -\beta_k\eta\Big(D_H(x^*,z_k)+D_H(z_k,x^*)\Big)\\
&\quad-\Big(s\beta_k\sqrt{D_H(x^*,z_k)}-\frac{1}{2}\sqrt{D_H(x_{k+1},z_k)}\Big)^2\\
&\quad+s^2\beta_k^2D_H(x^*,z_k)+\frac{1}{4}D_H(x_{k+1},z_k)\\
&\quad-\Big(\sqrt{D_H(x^*,z_k)}-\frac{1}{2}s\beta_k\sqrt{D_H(x^*,y^*_{k+1})}\Big)^2\\
&\quad+D_H(x^*,z_k)+\frac{1}{4}s^2\beta_k^2D_H(x^*,y^*_{k+1})+D_H(y^*_{k+1},z_k)\\
&\quad-D_H(x_{k+1},z_k)+D_H(x_{k+1},x^*)-D_H(y^*_{k+1},x^*)\\
&\quad-D_H(x_{k+1},y^*_{k+1}).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
D_H(y^*_{k+1},x_{k+1})\leqslant&\Big(1-\eta\beta_k+s^2\beta_k^2\Big)D_H(x^*,z_k)+\frac{1}{4}s^2\beta_k^2D_H(x^*,y^*_{k+1})\\
&+D_H(y^*_{k+1},z_k)+D_H(x_{k+1},x^*).
\end{aligned}\tag{3.12}$$

The rest of the proof will be divided into two parts:

Case A: Suppose that $D_H(y^*_{k+1},x_{k+1})\leqslant\big(1-\eta\beta_k+s^2\beta_k^2\big)D_H(x^*,z_k)$. Then the proof is clear.

Case B: Suppose that $D_H(y^*_{k+1},x_{k+1}) > \big(1-\eta\beta_k+s^2\beta_k^2\big)D_H(x^*,z_k)$.

Applying Lemma 2.9 for the following data:

$$X:=C,\ Y:=\mathbb{R},\ G(x):=C,\ \forall x\in C$$

$$y_k:=\beta_k,\ W(x,y)=-yf(x^*,x)-D_H(x^*,x),$$

we get that

$$\begin{aligned}
M(\beta_k)&=\Big\{\hat{x}\in G(\beta_k)\Big|\,W(\hat{x},\beta_k)=\sup\big\{-\beta_kf(x^*,x)-D_H(x^*,x):\ x\in G(\beta_k)\big\}\Big\}\\
&=\arg\max\{W(x,\beta_k):x\in C\}\\
&=\ \arg\min\big\{\beta_kf(x^*,x)+D_H(x^*,x)\big\}\\
&=y^*_{k+1}.
\end{aligned}$$

Hence

$$\limsup_{k\to\infty}y^*_{k+1}=\limsup_{k\to\infty}M(\beta_k)\subseteq \mathrm{cl}M(0)=\mathrm{cl}(x^*),$$

this implies that

$$\limsup_{k\to\infty}y^*_{k+1}=x^*.\tag{3.13}$$

So

$$\limsup_{k\to\infty}D_H(y^*_{k+1},x_{k+1})>\limsup_{k\to\infty}\big(1-\eta\beta_k+s^2\beta_k^2\big)D_H(x^*,z_k),$$

then from (3.1) we obtain

$$\limsup_{k\to\infty} D_H(x^*, x_{k+1}) \geq \limsup_{k\to\infty} D_H(x^*, z_k).$$

And this is a contradiction with Lemma 3.3. □

Lemma 3.5. *The sequences $\{x_k\}$, $\{y_k\}$ and $\{z_k\}$ are bounded.*

Proof. By a similar way as in the proof of Lemma 3.4, $\limsup_{k\to\infty} y^*_{k+1} = x^*$. Since f is continuous on C, so

$$\lim_{k\to\infty} f(x^*, y^*_{k+1}) = \limsup_{k\to\infty} f(x^*, y^*_{k+1}) = f(x^*, x^*) = 0.$$

By the Assumption (f_3), there exists a constant $\overline{M}(x^*) > 0$ such that

$$\left|f(x^*, y^*_{k+1})\right| \leqslant \overline{M}(x^*)\sqrt{D_H(x^*, y^*_{k+1})}, \quad \forall k > 0. \tag{3.14}$$

Substituting $x = x^*$ into (3.8) and note that $f(x^*, x^*) = 0$, we obtain

$$\beta_k\big(f(x^*, x^*) - f(x^*, y^*_{k+1})\big) + \big\langle x^* - y^*_{k+1}, \nabla H(y^*_{k+1}) - \nabla H(x^*)\big\rangle \geqslant 0,$$

therefore

$$\begin{aligned} -\beta_k f(x^*, y^*_{k+1}) &\geqslant \big\langle x^* - y^*_{k+1}, \nabla H(x^*) - \nabla H(y^*_{k+1})\big\rangle \\ &= D_H(x^*, y^*_{k+1}) + D_H(y^*_{k+1}, x^*) \\ &\geqslant D_H(x^*, y^*_{k+1}), \end{aligned}$$

using (3.14), we get

$$D_H(x^*, y^*_{k+1}) \leqslant -\beta_k f(x^*, y^*_{k+1}) \leqslant \beta_k \overline{M}(x^*)\sqrt{D_H(x^*, y^*_{k+1})}, \quad \forall k \geqslant 0,$$

then

$$\sqrt{D_H(x^*, y^*_{k+1})} \leqslant \beta_k \overline{M}(x^*). \tag{3.15}$$

Using the three point identity, we have

$$\begin{aligned} D_H(x^*, y^*_{k+1}) + D_H(y^*_{k+1}, x_{k+1}) + \Big\langle y^*_{k+1} - x^*, \nabla H(x_{k+1}) - \nabla H(y^*_{k+1})\Big\rangle \\ = D_H(x^*, x_{k+1}), \end{aligned}$$

from (3.15) and Lemma 3.4, we obtain

$$D_H(x^*, x_{k+1}) \leqslant \beta_k^2 \overline{M}^2(x^*) + \delta_k D_H(x^*, x_k) + \big\langle y^*_{k+1} - x^*, \nabla H(x_{k+1}) - \nabla H(y^*_{k+1})\big\rangle.$$

By using (3.1) and (3.13), we get

$$\limsup_{k\to\infty} D_H(x^*, x_{k+1}) \leqslant \limsup_{k\to\infty} D_H(x^*, x_k) \leqslant \ldots \leqslant \limsup_{k\to\infty} D_H(x^*, x_0) < \infty.$$

Therefore, the sequence $\{D_H(x^*, x_k)\}$ is bounded and by Lemma 2.4, the sequence $\{x_k\}$ is bounded too. From Lemma 3.3, it follows that $D_H(x^*, z_k) \leqslant D_H(x^*, x_k) < \infty$, so $\{z_k\}$ is also bounded. Since

$$D(y_k, x_k) \leqslant \frac{1}{1-\lambda_k c_k}\big(D(\overline{x}, x_k) - D(\overline{x}, z_k)\big),$$

so $\{y_k\}$ is bounded. □

Lemma 3.6. *Let $\{x_{k_i}\}$ be a subsequence of $\{x_k\}$ and converges weakly to $\hat{x}$ and $\lim\limits_{i\to\infty} D_H(y_{k_i}, x_{k_i}) = 0$. Then, $\hat{x} \in \mathrm{Sol}(C, g)$.*

Proof. Since $\{x_k\}\subset C$ and $\{x_{k_i}\}$ converges weakly to $\hat{x}$ and C is closed and convex, therefore $\hat{x} \in C$. On the other hand, since $\lim\limits_{i\to\infty} D_H(y_{k_i}, x_{k_i})=0$, we have $\lim\limits_{i\to\infty} \|y_{k_i} - x_{k_i}\|=0$, therefore $y_{k_i} \rightharpoonup \hat{x}$. In view of (3.4), we have

$$\lambda_{k_i}\big(g(x_{k_i}, x) - g(x_{k_i}, y_{k_i})\big) + \big\langle x - y_{k_i}, \nabla H(x_{k_i}) - \nabla H(y_{k_i})\big\rangle \geqslant 0, \quad \forall x \in C.$$

Passing to the limit in the last inequality as $i \to \infty$ and using $\lim\limits_{k\to\infty} \lambda_k = \lambda > 0$, boundedness of $\{z_{k_i}\}$ and weak continuity of g, we obtain

$$\lambda g(\hat{x}, x) \geqslant 0, \quad \forall x \in C,$$

therefore, we get $\hat{x} \in \mathrm{Sol}(C, g)$. □

Theorem 3.7. *Assume that Assumptions on f and g hold. The parameters are satisfied by the conditions* (3.1)*. Then, the sequences $\{x_k\}$, $\{y_k\}$ and $\{z_k\}$ generated by Algorithm A converge strongly to the unique solution point x^* of the BEP_s.*

Proof. Assume that $a_k = D_H(x^*, x_k)$. Using Lemmas 3.2, 3.3 and 3.5, we have

$$\begin{aligned} D_H(x^*, x_{k+1}) \leqslant & D_H(x^*, z_k) - D_H(x_{k+1}, z_k) + \beta_k\big(f(z_k, x^*) - f(z_k, x_{k+1})\big) \\ \leqslant & D_H(x^*, x_k) - (1-\lambda_k c_1) D_H(y_k, x_k) - (1-\lambda_k c_2) D_H(z_k, y_k) \\ & - D_H(x_{k+1}, z_k) + \beta_k\big(f(z_k, x^*) - f(z_k, x_{k+1})\big) r \qquad (3.16) \\ \leqslant & D_H(x^*, x_k) - (1-\lambda_k c_1) D_H(y_k, x_k) - (1-\lambda_k c_2) D_H(z_k, y_k) \\ & - D_H(x_{k+1}, z_k) + \beta_k K, \end{aligned}$$

where $K = \sup\limits_k \big\{f(z_k, x^*) - f(z_k, x_{k+1})\big\} < \infty$.

The rest of the proof will be divided into two parts:

Case 1: In this case, we suppose that there exists k_0 such that $D_H(x^*, x_{k+1}) \leqslant D_H(x^*, x_k)$ for all $k \geq k_0$. Then, from Lemma 2.8, we have $\lim\limits_{k\to\infty} D_H(x^*, x_k) = A < \infty$. By (3.16), we get

$$D_H(y_k, x_k) = D_H(z_k, y_k) = D_H(x_{k+1}, z_k) = 0. \qquad (3.17)$$

Since $\{z_k\}$ is bounded and X is a reflexive Banach space, so there exists a subsequence $\{z_{k_i}\}$ of $\{z_k\}$ such that $z_{k_i} \rightharpoonup \bar{z}$, $\bar{z} \in C$ and

$$\liminf_{k\to\infty} \big(f(x^*, z_k) + f(z_k, x_{k+1})\big) = \liminf_{i\to\infty} \big(f(x^*, z_{k_i}) + f(z_{k_i}, x_{k_i+1})\big).$$

It implies that $x_{k_i+1} \rightharpoonup \overline{z}$, $x_{k_i} \rightharpoonup \overline{z}$ and $y_{k_i} \rightharpoonup \overline{z}$. Then by Lemma 3.6, we have $\overline{z} \in \mathrm{Sol}(C, g)$ and

$$\liminf_{k\to\infty} \big(f(x^*, z_k) + f(z_k, x_{k+1})\big) = f(x^*, \overline{z}) \geqslant 0. \tag{3.18}$$

Then, by Assumption (f_1), f is η-strongly monotone and we have

$$\begin{aligned}\liminf_{k\to\infty} \big(f(x^*, z_k)+f(z_k, x_{k+1})\big) &\leqslant \liminf_{k\to\infty} -\eta\Big(D_H(x^*, z_k)+(D_H z_k, x^*)\Big)\\ &= -\,\eta A.\end{aligned} \tag{3.19}$$

Combining (3.18) and (3.19), we obtain

$$\begin{aligned}\liminf_{k\to\infty} \big(f(z_k, x^*) - f(z_k, x_{k+1})\big) &= \liminf_{k\to\infty} \big(f(z_k, x^*) + f(x^*, z_k) - f(z_k, x_{k+1})\big)\\ &\leqslant \liminf_{k\to\infty} \big(f(z_k, x^*) + f(x^*, z_k)\big)\\ &\quad - \liminf_{k\to\infty} \big(f(x^*, z_k) + f(z_k, x_{k+1})\big)\\ &\leqslant -\,\eta A.\end{aligned}$$

Now, we have to show that $A = 0$. If not, assume that $A > 0$. There exists k_0 such that

$$f(z_k, x^*) - f(z_k, x_{k+1}) \leqslant \frac{-\eta A}{2}. \tag{3.20}$$

Applying Lemma 3.3 for $x = x^*$ and using Lemma 3.2 for all $k > k_0$, we get

$$\begin{aligned}D_H(x^*, x_{k+1}) &\leqslant D_H(x^*, z_k) - D_H(x_{k+1}, z_k)+\beta_k\big(f(z_k, x^*)-f(z_k, x_{k+1})\big)\\ &\leqslant D_H(x^*, x_{k+1}) - \beta_k\eta A.\end{aligned} \tag{3.21}$$

Hence

$$D_H(x^*, x_{k+1}) - D_H(x^*, x_{k_i}) \leqslant -\,\eta A \sum_{j=k_0}^{k} \beta_j, \;\; \forall k \geqslant k_0. \tag{3.22}$$

Since $\sum_{k=1}^{\infty} \beta_k = +\infty$, we have $\lim\limits_{k\to\infty} D_H(x^*, x_k) = 0$. So $x_k \to x^*$.

Case 2: Now, we suppose that there does not exist k_0 such that

$$D_H(x^*, x_{k+1}) \leqslant D_H(x^*, x_k),$$

for all $k > k_0$. Then, there exists an integer k_0 such that

$$D_H(x^*, x_{k_0}) \leqslant D_H(x^*, x_{k_0+1}).$$

Lemma 2.8 introduces a subsequence $\{D_H(x^*, x_{\tau(k)})\}$ of $\{D_H(x^*, x_k)\}$ which is defined as

$$\tau(k) = \max\left\{i \in \mathbb{N}\middle| k_0 \leqslant i \leqslant k,\ D_H(x^*, x_i) \leqslant D_H(x^*, x_{i+1})\right\}.$$

Then, $\tau(k) \to \infty$ and

$$0 \leqslant D_H(x^*, x_k) \leqslant D_H(x^*, x_{\tau(k)+1}),$$

so, we get

$$0 \leqslant D_H(x^*, x_{\tau(k)}) \leqslant D_H(x^*, x_{\tau(k)+1}),\ \forall k \geqslant k_0. \tag{3.23}$$

Now, let $M := \lim_{k\to\infty} D_H(x^*, x_{\tau(k)})$. By Lemma 3.5 we have $M<\infty$. From (3.16) and $\lim_{k\to\infty} \beta_k=0$, we get

$$\begin{aligned}\lim_{k\to\infty} D_H(z_{\tau(k)}, x_{\tau(k)+1}) &= \lim_{k\to\infty} D_H(y_{\tau(k)}, x_{\tau(k)}) \\ &= \lim_{k\to\infty} D_H(z_{\tau(k)}, y_{\tau(k)}) \\ &= 0. \end{aligned} \tag{3.24}$$

Since $\{z_k\}$ is bounded, so there exists any subsequence which converges weakly to $\overline{x}$. Without loss of generality, we can assume that $z_{\tau(k)} \rightharpoonup \overline{x}$. Then $x_{\tau(k)+1} \rightharpoonup \overline{x}$. Replacing $k = \tau_k$ into (3.16), using the condition (3.1) that $1 - \lambda_{\tau(k)}c_1 > 0$ and $1 - \lambda_{\tau(k)}c_2 > 0$, we have

$$\begin{aligned}\beta_{\tau(k)}\big(f(z_{\tau(k)}, x_{\tau(k)+1}) - f(z_{\tau(k)}, x^*)\big) \leqslant\ & D_H(x^*, x_{\tau(k)}) - D_H(x^*, x_{\tau(k)+1}) \\ & - D_H(z_{\tau(k)}, x_{\tau(k)+1}) \\ & - (1-\lambda_{\tau(k)}c_1) D_H(y_{\tau(k)}, x_{\tau(k)}) \\ & - (1 - \lambda_{\tau(k)}c_2) D_H(z_{\tau(k)}, y_{\tau(k)}) \leqslant 0,\end{aligned}$$

hence

$$f(z_{\tau(k)}, x_{\tau(k)+1}) - f(z_{\tau(k)}, x^*) \leqslant 0. \tag{3.25}$$

Since f is η-strongly monotone on C, we have

$$\eta\Big(D_H(x^*, z_{\tau(k)}) + D_H(z_{\tau(k)}, x^*)\Big) \leqslant -f(z_{\tau(k)}, x^*) - f(x^*, z_{\tau(k)}). \tag{3.26}$$

Combining (3.25) and (3.26), we get

$$\eta\Big(D_H(x^*, z_{\tau(k)}) + D_H(z_{\tau(k)}, x^*)\Big) \leqslant -f(z_{\tau(k)}, x_{\tau(k)+1}) - f(x^*, z_{\tau(k)}).$$

Considering the limsup in the last inequality as $k\to\infty$ and using the Assumption (f_1), we get

$$\begin{aligned}\eta \limsup_{k\to\infty}\Big(D_H(x^*, z_{\tau(k)})+D_H(z_{\tau(k)}, x^*)\Big) &\leqslant \limsup_{k\to\infty}\Big(-f(z_{\tau(k)}, x_{\tau(k)+1})\\ &\quad -f(x^*, z_{\tau(k)})\Big)\\ &= -f(\overline{x},\overline{x}) - f(x^*,\overline{x})\\ &= -f(x^*,\overline{x})\\ &\leqslant 0.\end{aligned}$$

Therefore, $\limsup_{k\to\infty} D_H(z_{\tau(k)}, x^*) = 0$ and $\limsup_{k\to\infty} D_H(x^*, z_{\tau(k)}) = 0$. Using the three point identity, we have

$$\begin{aligned}D_H(x^*, x_{\tau(k)+1}) = &D_H(x^*, z_{\tau(k)}) + D_H(z_{\tau(k)}, x_{\tau(k)+1})\\ &-\Big\langle x^*-z_{\tau(k)}, \nabla H(x_{\tau(k)+1})-\nabla H(z_{\tau(k)})\Big\rangle.\end{aligned}$$

Passing the limit in the last inequality as $k\to\infty$ and from (3.24), we obtain

$$\lim_{k\to\infty} D_H(x^*, x_{\tau(k)+1}) = 0.$$

From (3.23) we have

$$0 \leqslant D_H(x^*, x_k) \leqslant D_H(x^*, x_{\tau(k)+1}) \xrightarrow{k\to\infty} 0.$$

The proof is complete. □

4. Applications

In this section, we first introduce the bilevel variational inequality problems, and then apply Algorithm A for solving of these problems.

Let C be a nonempty, closed and convex subset of a reflexive Banach space. The classical variational inequality problems, shortly (VIP), are

$$\text{Find } z \in C \quad \text{such that} \quad \langle y - z, Az\rangle \geq 0, \ \ \forall y \in C,$$

where $A : C\to X^*$ is a mapping and we denote the set of solutions of these problems by $VIP(A, C)$. The VIP have been intensively studied and widely applied to some practical problems arising in economics, optimization problems, differential equations, transportation, net work and structural analysis, finance and game theory. The important results and properties of VIP can find in [22] and the references there in.

The bilevel variational inequality problems ($BVIP$) are formulated as follows:

$$\text{Find } x^* \in VIP(A,C) \quad \text{such that } \langle y - x^*, Bx^* \rangle \geq 0, \ \ \forall y \in VIP(A,C),$$

where $A, B : C \to X^*$ and $VIP(A,C)$ is the set of solutions of the following VIP,

$$\text{Find } x^* \in C \quad \text{such that} \quad \langle y - x^*, Ax^* \rangle \geq 0, \ \ \forall y \in C.$$

Lemma 4.1. ([40]) *Let C be a nonempty closed convex subset of a reflexive Banach space X, $A : C \to X^*$ be a mapping and $f : X \to \mathbb{R}$ be a Legendre function. Then*

$$\overleftarrow{Proj}_C^f(\nabla f^*[\nabla f(x) - \lambda A(y)]) = \arg\min_{w \in C} \big\{\lambda \langle w - y, A(y) \rangle + D_f(w,x)\big\},$$

for all $x \in X$, $y \in C$ and $\lambda \in (0, +\infty)$.

Let X be a real Banach space and $1 < q \leq 2 \leq p$ with $\frac{1}{p} + \frac{1}{q} = 1$. The modulus of convexity $\delta_X : [0,2] \to [0,1]$ is defined by:

$$\delta_X(\epsilon) = \inf \Big\{1 - \frac{\|x+y\|}{2} : \|x\| = \|y\| = 1, \|x-y\| \geq \epsilon\Big\},$$

X is called uniformly convex if $\delta_X(\epsilon) > 0$ for any $\epsilon \in (0,2]$, p-uniformly convex if there is a $c_p > 0$ so that $\delta_X(\epsilon) \geq c_p \epsilon^p$ for any $\epsilon \in (0,2]$. The modulus of smoothness $\rho_X(\tau) : [0,\infty) \to [0,\infty)$ is defined by:

$$\rho_X(\tau) = \sup \Big\{\frac{\|x + \tau y\| + \|x - \tau y\|}{2} - 1 : \|x\| = \|y\| = 1\Big\},$$

X is called uniformly smooth if $\displaystyle\lim_{\tau \to 0} \frac{\rho_X(\tau)}{\tau} = 0$.

For the p-uniformly convex space, the metric and Bregman distance have the following relation [35]:

$$\tau \|x - y\|^p \leq D_{\frac{1}{p}\|.\|^p}(x,y) \leq \langle x - y, J_X^p(x) - J_X^p(y) \rangle, \tag{4.1}$$

where $\tau > 0$ is fixed number and duality mapping $J_X^p : X \to 2^{X^*}$ is defined by:

$$J_X^p(x) = \{f \in X^*, \langle x, f \rangle = \|x\|^p, \|f\| = \|x\|^{p-1}\},$$

for every $x \in X$. We know that X is smooth if and only if J_X^p is single-valued mapping of X into X^*. We also know that X is reflexive if and only if J_X^p is surjective, and X is strictly convex if and only if J_X^p is one-to-one. Therefore, if X is smooth, strictly convex and reflexive Banach space, then J_X^p is a single-valued bijection and in this case, $J_X^p = (J_{X^*}^q)^{-1}$ where $J_{X^*}^q$ is the duality mapping of X^*.

For $p = 2$, the duality mapping J_X^p, is called the normalized duality and is denoted by J. The function $\phi : X^2 \to \mathbb{R}$ is defined by:

$$\phi(y, x) = \| y \|^2 - 2\langle y, Jx\rangle + \| x \|^2,$$

for all $x, y \in X$. The generalized projection Π_C from X onto C is defined by:

$$\Pi_C(x) = \arg\min_{y\in C} \phi(y, x), \quad \forall x \in X,$$

where C is a nonempty, closed and convex subset of X.

Let X be a uniformly smooth and uniformly convex Banach space, and $f = \frac{1}{2}\|.\|^2$. So $\nabla f = J$, $D_{\frac{1}{2}\|.\|^2}(x, y) = \dfrac{1}{2}\phi(x, y)$ and $\overleftarrow{Proj}_C^{\frac{1}{2}\|.\|^2} = \Pi_C$. In particular if X is a Hilbert space, then $\nabla f = I$, $D_{\frac{1}{2}\|.\|^2}(x, y) = \dfrac{1}{2}\|x - y\|^2$ and $\overleftarrow{Proj}_C^{\frac{1}{2}\|.\|^2} = P_C$, where P_C is the metric projection.

Now, we consider the bilevel equilibrium problem corresponding to the functions g and f defined by:

$$g(x, y) = \langle y - x, Bx\rangle, \quad \forall x, y \in C,$$
$$f(x, y) = \langle y - x, Ax\rangle, \quad \forall x, y \in C,$$

where $A, B : C \to X^*$.

Corollary 4.2. *Let C be a nonempty closed convex subset of a uniformly smooth and 2-uniformly convex Banach space X. Assume that Assumptions on f and g hold and the parameters are satisfied by the conditions* (3.1)*. Then, the sequences $\{x_k\}$, $\{y_k\}$ and $\{z_k\}$ generated by the following algorithm converge strongly to the unique solution point x^* of the $BVIP_s$.*

$$\begin{cases} y_k = \Pi_C\Big(J^{-1}(J(x_k) - \lambda_K B(x_k))\Big), \\ z_k = \Pi_{T_k}\Big(J^{-1}(J(x_k) - \lambda_K B(y_k))\Big), \\ T_k = \{v \in X |\, \big\langle v - y_k, J(x_k) - \lambda_n B(x_k) - J(y_k)\big\rangle \leqslant 0\}, \\ x_{k+1} = \Pi_C\Big(J^{-1}(J(z_k) - \lambda_K A(z_k))\Big). \end{cases}$$

5. Numerical experiment

In this subsection the numerical results will be presented in order to test our algorithm.

Let $X = \mathbb{R}^n$ and $C = \Big\{x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n : -5 \le x_i \le 5,\ 1 \le i \le n\Big\}$. We define the following bifunctions

$$g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$
$$g(x, y) = \langle Px + Qy, y - x\rangle,$$

where P and Q are orbitally symmetric positive semidefinite matrices such that $P-Q$ is positive semidefinite. We define

$$f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$
$$f(x,y) = \langle Ax - By, y - x\rangle,$$

with A and B being positive definite matrices defined by:

$$B = N^T N + nI_n \quad \text{and} \quad A = B + M^T M + nI_n,$$

where M and N are orbitally $n \times n$ matrices and I_n is the identity matrix.

From [39], we have the following properties

(1) g is pseudomonotone.
(2) g is Lipschitz- type continuous with constants $\frac{1}{2}\|P-Q\|$.
(3) $Sol(C,g) \neq \emptyset$.
(4) f is n-strongly monotone.

In addition, we define the function $H : \mathbb{R}^n \to \mathbb{R}$ as

$$H(x) = \frac{1}{2}\|A_1 x - b_1\|,$$

where A_1 is a orbitally matrix and b_1 is a orbitally vector. Therefore, we have

$$\langle y, \nabla H\rangle = \lim_{t\to\infty} \frac{H(x,ty) - H(x)}{t} = \langle A_1 y, A_1 x - b_1\rangle,$$

this mean is $\nabla H = A_1^2 x - A_1 b_1$.

Finally, we consider the condition (f_2) ,

$$\begin{aligned} f(x,y) + f(y,z) - f(x,z) =& \langle Ax - Ay, y - z\rangle + \langle By - Bz, y - x\rangle \\ =& \langle \bar{f}_1(x,y), \nabla \hat{f}_1(y) - \nabla \hat{f}_1(z)\rangle \\ &+ \langle \bar{f}_2(x,y), \nabla \hat{f}_2(y) - \nabla \hat{f}_2(z)\rangle, \end{aligned}$$

letting $\bar{f}_1(x,y) = A(x,y)$, $\bar{f}_2(x,y) = y - x$, $\nabla \hat{f}_1(y) - \nabla \hat{f}_1(z) = y - z$ and $\nabla \hat{f}_2(y) - \nabla \hat{f}_2(z) = B(y-z)$. Therefore, we have the following relations

$$\|\bar{f}_1(x,y)\| = \|A\|\|x-y\|, \quad \text{and} \quad \bar{L}_1 = \sqrt{2}\|A\|,$$

$$\|\bar{f}_2(x,y)\| = \|y-x\|, \quad \text{and} \quad \bar{L}_2 = \sqrt{2},$$

$$\|\nabla \hat{f}_1(y) - \nabla \hat{f}_1(z)\| = \|x-y\|, \quad \text{and} \quad \hat{L}_1 = \sqrt{2},$$

$$\|\nabla \hat{f}_2(y) - \nabla \hat{f}_2(z)\| = \|B\|\|x-y\|, \quad \text{and} \quad \hat{L}_2 = \sqrt{2}\|B\|.$$

Then, the our algorithm becomes

Initialization: Choose $x_0 \in C$, the tolerance $\epsilon > 0$ and the sequences $\{\lambda_k\}$ and $\{\beta_k\}$ such that

$$\begin{cases} \{\lambda_k\} = \dfrac{1}{2c_1 + 700k}, \quad \lim\limits_{k\to\infty} \lambda_k = 0, \\ \beta_k \searrow 0, \ 1 - \beta_k \eta + \beta_k^2 s^2 < 1, \ \eta = n, \ s = \sum\limits_{i=1}^{2} \hat{L}_i \bar{L}_i, \\ 0 < \tau < \min\{\eta, s\}, \ 0 < \beta_k < \min\left\{ \dfrac{1}{\tau}, \dfrac{2\eta - 2\tau}{s^2 - \tau^2}, \dfrac{2\eta}{s^2} \right\}. \end{cases} \tag{5.1}$$

Set $k = 0$ and go to Step 1.

Step 1: Compute $(k = 0, 1, ...)$

$$y_k = \arg\min\{\lambda_k g(x_k, y) + \frac{1}{2}\|y - x_k\|^2 : y \in C\},$$

$$z_k = \arg\min\{\lambda_k g(y_k, z) + \frac{1}{2}\|z - x_k\|^2 : z \in T_k\},$$

$$x_{k+1} = \arg\min\{\beta_k f(z_k, t) + \frac{1}{2}\|t - z_k\|^2 : t \in C\},$$

where $T_k = \Big\{ v \in X \mid \langle \nabla H(x_k) - \lambda_k w_k - \nabla H(y_k), v - y_k \rangle\} \leqslant 0 \Big\}$ and $w_k \in \{2Py_k + (P - Q)x_k\}$.

Step 2: If $\max\Big\{\|y_k - x_{k+1}\|, \|y_k - x_k\|\Big\} \leqslant \epsilon$ then stop.

Step 3: Otherwise, set $k := k + 1$ and go back to Step 1.

We will illustrate the numerical results by the following example.

Example 5.1. Suppose that $X = \mathbb{R}^3$ and $C = \Big\{ x = (x_1, x_2, x_3) \in \mathbb{R}^3 \ \Big| \ -5 \leqslant x_i \leqslant 5, \ \forall i = 1, 2, 3 \Big\}$. We choose the tolerance $\epsilon = 10^{-5}$, the starting point $x_0 = (4, 0, -1)$ and the orbitally vector $b = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$. We also choose the following matrices

$$P = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \ Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \ M = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix},$$

$$N = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 1 \end{bmatrix}, \ A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}.$$

The numerical results are showed in Table 1 and Table 2. The decreasing values of $\|x_m\|$ and also the values of $\|x_m - y_{m-1}\|$ are shown in the Figure 1 and Figure 2.

TABLE 1. Numerical results

step m	y_m	x_{m+1}
0	$(2.33333, -8.08441 \times 10^{-2}, -8.33333 \times 10^{-1})$	$(-2.93986, 1.80376 \times 10^{-2}, -1.42857 \times 10^{-1})$
1	$(-1.80915, 1.08225 \times 10^{-2}, -1.20879 \times 10^{-1})$	$(1.87599, -1.3964 \times 10^{-2}, -2.85714 \times 10^{-2})$
2	$(1.20599, -8.72747 \times 10^{-3}, -2.44898 \times 10^{-2})$	$(-7.17105 \times 10^{-1}, 5.90252 \times 10^{-3}, -9.52380 \times 10^{-3})$
3	$(-4.78070 \times 10^{-1}, 3.81928 \times 10^{-3}, -8.25396 \times 10^{-3})$	$(6.68996 \times 10^{-2}, -8.71873 \times 10^{-4}, -4.55486 \times 10^{-3})$
4	$(4.59935 \times 10^{-2}, -5.81249 \times 10^{-4}, -3.98550 \times 10^{-3})$	$(1.02730 \times 10^{-2}, -1.14135 \times 10^{-4}, -2.73292 \times 10^{-3})$
5	$(7.25155 \times 10^{-3}, -7.80923 \times 10^{-5}, -2.41140 \times 10^{-3})$	$(3.51114 \times 10^{-3}, -4.23808 \times 10^{-5}, -1.89202 \times 10^{-3})$
6	$(2.53582 \times 10^{-3}, -2.96663 \times 10^{-5}, -1.68179 \times 10^{-3})$	$(1.69705 \times 10^{-3}, -1.99866 \times 10^{-5}, -1.43793 \times 10^{-3})$
7	$(1.25046 \times 10^{-3}, -1.42762 \times 10^{-5}, -1.28657 \times 10^{-3})$	$(9.96384 \times 10^{-4}, -1.12190 \times 10^{-5}, -1.16404 \times 10^{-3})$
8	$(7.47293 \times 10^{-4}, -8.15923 \times 10^{-6}, -1.04764 \times 10^{-3})$	$(6.62020 \times 10^{-4}, -7.10045 \times 10^{-6}, -9.84957 \times 10^{-4})$
9	$(5.04396 \times 10^{-4}, -5.24816 \times 10^{-6}, -8.91151 \times 10^{-4})$	$(4.78560, \times 10^{-4} - 4.89727 \times 10^{-6}, -8.60544 \times 10^{-4})$
10	$(3.69796 \times 10^{-4}, -3.67295 \times 10^{-6}, -7.82313 \times 10^{-4})$	$(3.67454 \times 10^{-4}, -3.60493 \times 10^{-6}, -7.69961 \times 10^{-4})$
11	$(2.87573 \times 10^{-4}, -2.73975 \times 10^{-6}, -7.03008 \times 10^{-4})$	$(2.95060 \times 10^{-4}, -2.78598 \times 10^{-6}, -7.01516 \times 10^{-4})$

TABLE 2. Numerical results

step m	$\|x_m\|$	$\|y_{m-1} - x_m\|$
0	2.94339	5.31824
1	1.87625	3.68637
2	0.71719	1.92321
3	6.70601×10^{-2}	5.45002×10^{-1}
4	1.06309×10^{-2}	3.57454×10^{-2}
5	3.98868×10^{-3}	3.77647×10^{-3}
6	2.22442×10^{-3}	8.73553×10^{-4}
7	1.53229×10^{-3}	2.82094×10^{-4}
8	1.18679×10^{-3}	1.05840×10^{-4}
9	9.84673×10^{-4}	4.00550×10^{-5}
10	8.53156×10^{-4}	1.25725×10^{-5}
11	7.61048×10^{-4}	7.63405×10^{-6}

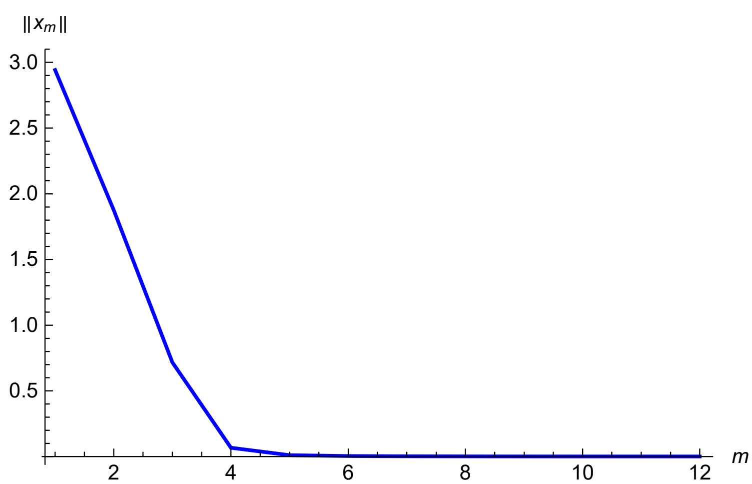

FIGURE 1. Plotting of $\|x_m\|$.

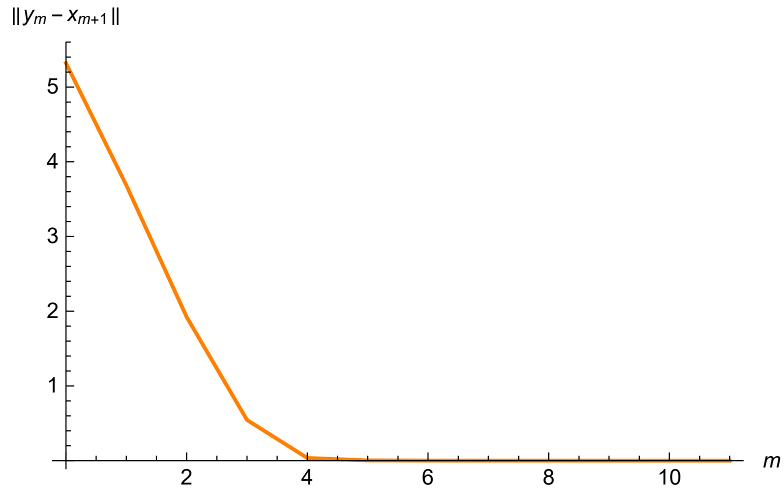

FIGURE 2. Plotting of $\|y_m - x_{m+1}\|$.

Acknowledgments: This work was supported by the Basic Science Research Program through the National Research Foundation(NRF) Grant funded by Ministry of Education of the republic of Korea (2018R1D1A1B07045427).

REFERENCES

[1] P.N. Anh, J.K Kim, N.D. Hien and N.V. Hong, *Strong convergence of inertial hybrid subgradient methods for solving equilibrium problems in Hilbert spaces,* J. Nonlinear Convex Anal. **24**(3) (2023), 499–514.

[2] P.N. Anh, H.T.C. Thach and J.K. Kim, *Proximal-like subgradient methods for solving multi-valued variational inequalities,* Nonlinear Funct. Anal. Appl., **25**(3) (2020), 437–451.

[3] P.N. Anh and H.A. Le Thi, *An Armijo-type method for pseudomonotone equilibrium problems and its applications*, J. Global Optim., **57**(3) (2013), 803–820.

[4] P.N. Anh and H.A. Le Thi, *New subgradient extragradient methods for solving monotone bilevel equilibrium problems,* Optimization, (2019) https://doi.org/ 10.1080/02331934.2019.1656204.

[5] J.P. Aubin and I. Ekeland, *Applied nonlinear analysis*, New York. Wiley, 1984.

[6] H.H. Bauschke and J.M. Borwein, *Legendre functions and the method of random Bregman projections,* J. Convex Anal., **4** (1997), 27–67.

[7] H.H. Bauschke, J.M. Borwein and P.L. Combettes, *Bregman monotone optimization algorithms,* SIAM J. Control Optim., **42** (2003), 596–636.

[8] H.H. Bauschke, J.M. Borwein and P.L. Combettes, *Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces,* Commun. Contemp. Math., **3** (2001), 615–647.

[9] G.C. Bento, J.X. CruzNeto, J.O. Lopes, P.A. Soares and A.Jr. Soubeyran, *Generalized proximal distances for bilevel equilibrium problems,* SIAM J. Optim., **26** (2016), 810–830.

[10] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems,* Springer-Verlag, New York, 2000.

[11] L.M. Bregman, *A relaxation method for finding the common point of convex sets and its application to the solution of problems in convex programming,* USSR Comput. Math. Math. Phys., **7** (1967), 200–217.

[12] D. Butnariu, Y. Censor and S. Reich, Îterative averaging of entropic projections for solving stochastic convex feasibility problems, Comput. Optim. Appl., **8** (1997), 21–39.

[13] D. Butnariu and A.N. Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization,* Kluwer Academic Publishers. Dordrecht. The Netherlands, 2000.

[14] D. Butnariu and E. Resmerita, *Bregman distances, totally convex functions and a method for solving operator equations in Banach spaces,* Abstr. Appl. Anal. Art., 2006, 1–39.

[15] Y. Censor, A. Gibali and S. Reich, *The subgradient extragradient method for solving variational inequalities in Hilbert space,* J. Optim. Theory Appl., **148** (2011), 318–335.

[16] O. Chadli, Z. Chbani and H. Riahi, *Equilibrium problems with generalized monotone bifunctions and applications to variational inequalities,* J. Optim. Theory Appl., **105** (2000), 299–323.

[17] Z. Chbani and H. Riahi, *Weak and strong convergence of proximal penalization and proximal splitting algorithms for two-level hierarchical Ky Fan minimax inequalities,* Optimization, **64** (2015), 1285–1303.

[18] I. Cioranescu, *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems,* Kluwer Academic Publishers, Netherlands, 1990.

[19] S. Dempe, *Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints,* Optimization, **52** (2003), 333–359.

[20] H. Iiduka, *Fixed point optimization algorithm and its application to power control in CDMA data networks,* Math. Program, **133** (2012), 227–242.

[21] H. Iiduka and I. Yamada, *A use of conjugate gradient direction for the convex optimization problem over the fixed point set of a nonexpansive mapping,* SIAM J. Optim., **19** (2009), 1881–1893.

[22] G. Kassay, S. Reich and S. Sabach, *Iterative methods for solving systems of variational inequalities in reflexive Banach spaces,* SIAM J. Optim.,

21 (2011), 1319–1344.

[23] J.K. Kim and P. Majee, *Modified Krasnoselski Mann iterative method for hierarchical fixed point problem and split mixed equilibrium problem,* J. Ineq. Appl., **(2020)** 2020:227, ISSN 1029-242X doi.org/10.1186/s13660-020-02493-8.

[24] J.K. Kim and Salahuddin, *Existence of solutions for multi-valued equilibrium problems,* Nonlinear Funct. Anal. Appl., **23**(4) (2018), 779-795.

[25] J.K. Kim and T.M. Tuyen, *A parallel iterative method for a finite family of Bregman strongly nonexpansive mappings in reflexive Banach spaces,* J. Kor. Math. Soc., **57**(3) (2020), 617–640. https://doi.org/10.4134/JKMS.j190268.

[26] P.E. Maingé, *A hybrid extragradient-viscosity method for monotone operators and fixed point problems,* SIAM J. Control Optim., **47** (2008), 1499–1515.

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

[28] K. Muangchoo, *A new explicit extragradient method for solving equilibrium problems with convex constraints,* Nonlinear Funct. Anal. Appl., **27**(1) (2022), 1–22.

[29] L.D. Muu and W. Oettli, *Optimization over equilibrium sets,* Optimization, **49** (2000), 179–189.

[30] TD. Quoc, LD. Muu and VH. Nguyen, *Extragradient algorithms extended to equilibrium problems,* Optimization, **57**(6) (2008), 749–776.

[31] N.V. Quy, *An algorithm for a bilevel problem with equilibrium and fixed point constraints,* Optimization, **64** (2014), 1–17.

[32] S. Reich and S. Sabach, *A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces,* J. Nonlinear Convex Anal., **10** (2009), 471–485.

[33] S. Reich and S. Sabach, *Two strong convergence theorems for a proximal method in reflexive Banach spaces,* Numer. Funct. Anal. Optim., **31** (2010), 22–44.

[34] S. Sabach, *Products of finitely many resolvents of maximal monotone mappings in reflexive Banach spaces,* SIAM J. Optim., **21** (2011), 1289–1308.

[35] F. Schöpfer, T. Schuster and A. K Louis, *An iterative regularization method for the solution of the split feasibility problem in Banach spaces,* Inverse Problems, **24** (2008).

[36] L.Q. Thuy and T.N. Hai, *A projected subgradient algorithm for bilevel equilibrium problems and applications,* J. Optim. Theory Appl., https://doi.org/10.1007/s10957-017-1176-2.

[37] J.V. Tiel, *Convex Analysis, An introductory text.* John Wiley & Sons Ltd, 1984.

[38] J.C. Yao, *Variational inequalities with generalized monotone operators,* Math. Oper. Research. **19** (1994), 691–705.

[39] T. Yuying, B.V. Dinh, D.S. Kim and S. Plubtieng, *Extragradient subgradient methods for solving bilevel equilibrium problems*, J. Ineq. Appl., https://doi.org/10.1186/s13660-018-1898-1.

[40] G. Zamani Eskandani, M. Raeisi and Themistocles M. Rassias, *A Hybrid extragradient method for pseudomonotone equilibrium problems by using Bregman distance,* Fixed Point Theory Appl., **27** (2018), 20–132.