

RESEARCH

Open Access



Weak and strong convergence Bregman extragradient schemes for solving pseudo-monotone and non-Lipschitz variational inequalities

Lateef Olakunle Jolaoso^{1*}  and Maggie Aphane¹

*Correspondence:

jollatanu@yahoo.co.uk

¹Department of Mathematics and Applied Mathematics, Sefako Makgatho Health Sciences University, P.O. Box 94 Medunsa 0204, Pretoria, South Africa

Abstract

In this paper, we introduce Bregman subgradient extragradient methods for solving variational inequalities with a pseudo-monotone operator which are not necessarily Lipschitz continuous. Our algorithms are constructed such that the stepsizes are determined by an Armijo line search technique, which improves the convergence of the algorithms without prior knowledge of any Lipschitz constant. We prove weak and strong convergence results for approximating solutions of the variational inequalities in real reflexive Banach spaces. Finally, we provide some numerical examples to illustrate the performance of our algorithms to related algorithms in the literature.

MSC: 65K15; 47J25; 65J15; 90C33

Keywords: Pseudo-monotone; Extragradient method; Line search; Variational inequalities; Bregman distance; Numerical result

1 Introduction

In this paper, we consider the *Variational Inequalities (VIs)* of Fichera [17] and Stampacchia [40] in a unified framework. The *VI* is defined as finding a point $x^\dagger \in C$ such that

$$\langle Ax^\dagger, y - x^\dagger \rangle \geq 0, \quad \forall y \in C, \quad (1.1)$$

where $C \subseteq E$ is a nonempty, closed and convex set in a real Banach space E with dual E^* , $\langle \cdot, \cdot \rangle$ denotes the duality pairing on E , and $A : E \rightarrow E^*$ is a given operator. We denote the set of solutions of the *VI* by $VI(C, A)$. The *VI* is a fundamental problem in optimization theory and captures various applications such as partial differential equations, optimal control, and mathematical programming (see, for instance, [19, 29]). A vast literature which deals on iterative methods for solving *VIs* can be found, for instance, in [1, 12, 14–16, 24–28, 41].

A classical method for solving the *VI* is the projection method given by

$$x_{n+1} = P_C(x_n - \alpha_n Ax_n), \quad (1.2)$$

© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

where P_C is the metric projection onto $C \subset \mathbb{R}^N$. The projection method is a natural extension of the projected gradient method for solving optimization problems, originally proposed by Goldstein [20], and Levitin and Polyak [31]. Under the assumption that A is η -strongly monotone and L -Lipschitz continuous with $\alpha_n \in (0, \frac{2}{L})$, the projection method converges to a solution of the VI. But if these conditions are weakened, say for example, the strong monotonicity is reduced to monotonicity, the situation becomes complicated and yields a divergent sequence independent of the stepsize α_n [13]. As a result of this setback, Korpelevich [30] and Antipin [3] introduced the *Extragradient Method (EM)* which is a two-projection process and is defined as

$$\begin{cases} y_n = P_C(x_n - \alpha_n Ax_n), \\ x_{n+1} = P_C(x_n - \alpha_n Ay_n), \end{cases} \tag{1.3}$$

where $\alpha_n \in (0, 1/L)$ and L is the Lipschitz constant of A in finite-dimensional settings. The *EM* has received great attention in recent days and many authors have introduced several improvements and modifications of the method. Observing that in the *EM*, there is need to calculate the projection onto the feasible set C twice per each iteration, and in a case where the set C is not simple to project onto, a minimal distance problem needs to be solved twice to obtain the next iterate (which can fact can affect the efficiency and applicability of the *EM*), Censor et al. [9] introduced the *Subgradient Extragradient Method (SEM)* in a real Hilbert space H , while the second projection onto C is replaced with a projection onto a half-space which can be calculated explicitly. In particular, the *SEM* is defined as

$$\begin{cases} y_n = P_C(x_n - \alpha_n Ax_n), \\ x_{n+1} = P_{T_n}(x_n - \alpha_n Ay_n), \\ \text{where} \\ T_n = \{w \in H : \langle x_n - \alpha_n Ax_n - y_n, w - y_n \rangle \leq 0\}, \end{cases} \tag{1.4}$$

and $\alpha_n \in (0, 1/L)$. Since the inception of the *SEM*, many authors have proposed various modifications of the *SEM*; see for instance [7, 8, 14, 41, 42, 44, 45].

Many of the results which deal with *EM* and *SEM* for solving the VI used the Euclidean norm distance and metric projections, which in certain cases do not allow for an application to the structure of generally feasible sets and efficient problem-solving. A possible way out is to use the Bregman divergence (or Bregman distance) introduced by Bregman [6] where he proposed a method of the type of cyclic projection for finding the common point of a convex set. His paper has initiated a new field in mathematical programming and non-linear analysis. For solving the VI, one of the modern variants of the *EM* is the Nomirovski prox-method [34], which can be interpreted as a variant of the *EM* with projection understood in the sense of Bregman divergence. It sometimes allows for the consideration of the structure of a general feasible set of the problem. For example, for a simplex, it is possible to take the Kullback–Leibler divergence (Bregman divergence on negative entropy) as the distance and to obtain explicitly calculated operator of projection onto a simplex. See [37] for more details of the Bregman divergence.

Recently, Nomirovskii et al. [35] proposed the following two-stage method using the Bregman divergence with operator $f : H \rightarrow \mathbb{R} \cup \{+\infty\}$ being continuously differentiable and σ -strongly convex.

Algorithm 1.1 Choose $x_0, y_0 \in C$ and positive number λ . Put $n = 1$.

Step 0: Calculate $x_1 = P_{x_0}^C(-\lambda y_0)$, $y_1 = P_{x_1}^C(-\lambda A y_0)$.

Step 1: Calculate $x_{n+1} = P_{x_n}^{T_n}(-\lambda A y_n)$ and $y_{n+1} = P_{x_{n+1}}^C(-\lambda A y_n)$ where

$$T_n = \{z \in H : \langle \nabla f(x_n) - \lambda A y_{n-1} - \nabla f(y_n), z - y_n \rangle \leq 0\}.$$

Step 2: If $x_{n+1} = x_n$ and $y_{n+1} = y_n = y_{n-1}$, then STOP and $y_n \in VI(C, A)$. Otherwise put $n := n + 1$ and go to **Step 1**, where P_x^C is the projection defined by

$$P_x^C(a) = \operatorname{argmin}_{y \in C} \{-\langle a, y - x \rangle + D_f(y, x)\}, \quad \forall a \in H, x \in \operatorname{int}(\operatorname{dom} f)$$

and $D_f(y, x)$ is the Bregman divergence between x and y .

They proved a weak convergence result for the VI with a pseudo-monotone and L -Lipschitz continuous operator in a finite-dimensional linear normed space provided the stepsize λ satisfies

$$\lambda \in \left(0, (\sqrt{2} - 1) \frac{\sigma}{L}\right).$$

Also, Gibali [18] introduced the following Bregman subgradient extragradient method for solving the VI with monotone and Lipschitz continuous operator in a real Hilbert space.

Algorithm 1.2 Choose $x_0, y_0 \in H$, $\lambda > 0$. Given the current iterates x_n and y_n , and also y_{n-1} , if $\nabla f(x_n) - \lambda A y_{n-1} \neq \nabla f(y_n)$, construct the half-space

$$T_n = \{w \in H : \langle \nabla f(x_n) - \lambda A y_{n-1} - \nabla f(y_n), w - y_n \rangle \leq 0\}$$

and if $\nabla f(x_n) - \lambda A y_{n-1} = \nabla f(y_n)$, take $T_n = H$. Now compute the next iterates via

$$\begin{cases} x_{n+1} = \Pi_{T_n}((\nabla f)^{-1}(\nabla f(x_n) - \lambda A y_n)), \\ y_{n+1} = \Pi_C((\nabla f)^{-1}(\nabla f(x_{n+1}) - \lambda A y_n)), \end{cases}$$

where Π_C denotes the Bregman projection onto C (see Definition 2.2).

Gibali [18] proved a weak convergence result for the sequence generated by Algorithm 1.2 provided the stepsize satisfies the condition $\lambda \in (0, \frac{\sqrt{2}-1}{L})$, where L is the Lipschitz constant of A .

It is obvious that the stepsize used in the EM and SEM has an essential role in the convergence of the two methods. An obvious disadvantage of Algorithms 1.1 and 1.2, which impedes their wide usage, is the assumption that the Lipschitz constant of the operator is known or admits a simple estimate. Moreover, in many problems, operators may not satisfy the Lipschitz condition and the operator may not even be monotone as our Example 4.2 shows in Sect. 4.

Motivated by the above results, in this paper, we present modified Bregman subgradient extragradient algorithms with line search technique for solving VI with a pseudo-monotone operator without necessarily satisfying the Lipschitz condition. The stepsize of the algorithm is determined via an Armijo line search technique which helps us to avoid finding a prior estimate of the Lipschitz constant as well as improve the convergence of the algorithm by finding optimum stepsize for each iteration. We proved weak and strong convergence results for approximating solutions of VI in real reflexive Banach spaces and provide numerical examples to illustrate the performance of our algorithms. Our results improve and extend the corresponding results of [7, 12, 18, 34, 35] in the literature.

2 Preliminaries

In this section, we give some definitions and basic results which will be used in our subsequent analysis. Let E be a real Banach space with dual E^* and C be a nonempty, closed and convex subset of E . We denote the strong and the weak convergences of a sequence $\{x_n\} \subseteq H$ to a point $p \in E$ by $x_n \rightarrow p$ and $x_n \rightharpoonup p$, respectively.

We now introduce some necessary structure to formulate our algorithm. Let $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ be a function satisfying the following:

- (i) $\text{int}(\text{dom}f) \subseteq E$ is a nonempty convex set;
- (ii) f is continuously differentiable on $\text{int}(\text{dom}f)$;
- (iii) f is strongly convex with strong convexity constant $\sigma > 0$, i.e.,

$$f(x) \geq f(y) - \langle \nabla f(y), x - y \rangle + \frac{\sigma}{2} \|x - y\|^2, \quad \forall x \in \text{dom}f \text{ and } y \in \text{int}(\text{dom}f). \tag{2.1}$$

The subdifferential set of f at a point x denoted by ∂f is defined by

$$\partial f(x) = \{ \xi \in E^* : f(y) - f(x) \geq \langle \xi, y - x \rangle, \forall y \in E \}. \tag{2.2}$$

Each element $\xi \in \partial f(x)$ is called a subgradient of f at x . Since f is continuously differentiable, $\partial f(x) = \{ \nabla f(x) \}$, which is the gradient of f at x . The Fenchel conjugate of f is the convex functional $f^* : E^* \rightarrow \mathbb{R} \cup \{+\infty\}$ defined by $f^*(\xi) = \sup\{ \langle \xi, x \rangle - f(x) : x \in E \}$. Let E be a reflexive Banach space, the function f is said to be Legendre if and only if it satisfies the following two conditions:

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

The Bregman divergence (or Bregman distance) corresponding to the function f is defined by (see [6])

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y \rangle, \quad \forall x \in \text{dom}f \text{ and } y \in \text{int}(\text{dom}f). \tag{2.3}$$

Remark 2.1 Example of practically important Bregman divergence can be found in [5]. We consider the following two: for $f(\cdot) = \frac{1}{2} \|\cdot\|^2$, we have $D_f(x, y) = \frac{1}{2} \|x - y\|^2$ which is the Euclidean norm distance. Also, if $f(x) = -\sum_{i=1}^m x_i \log(x_i)$ which is the Shannon entropy for the non-negative orthant $\mathbb{R}_{++}^m := \{x \in \mathbb{R}^m : x_i > 0\}$, we obtain the Kullback–Leibler cross entropy defined by

$$D_f(x, y) = \sum_{i=1}^m \left(x_i \log\left(\frac{x_i}{y_i}\right) - 1 \right) + \sum_{i=1}^m y_i.$$

It is well known that the Bregman distance is not a metric, however, it satisfies the following three point identity:

$$D_f(x, z) = D_f(x, y) + D_f(y, z) + \langle \nabla f(y) - \nabla f(z), x - y \rangle. \tag{2.4}$$

Also from the strong convexity of f , we have

$$D_f(x, y) \geq \frac{1}{2} \|x - y\|^2. \tag{2.5}$$

Definition 2.2 The Bregman projection (see e.g. [38]) with respect to f of $x \in \text{int}(\text{dom} f)$ onto a nonempty closed convex set $C \subset \text{int}(\text{dom} f)$ is the unique vector $\Pi_C(x) \in C$ satisfying

$$\Pi_C(x) := \inf\{D_f(x, y) : y \in C\}.$$

The Bregman projection is characterized by the inequality

$$z = \Pi_C(x) \iff \langle \nabla f(x) - \nabla f(z), y - z \rangle \leq 0, \quad \forall y \in C. \tag{2.6}$$

Also

$$D_f(y, \Pi_C(x)) + D_f(\Pi_C(x), x) \leq D_f(y, x), \quad \forall y \in C, x \in \text{int}(\text{dom} f). \tag{2.7}$$

Following [2, 10], we define the function $V_f : E \times E \rightarrow [0, \infty)$ associated with f by

$$V_f(x, y) = f(x) - \langle x, y \rangle + f^*(y), \quad \forall x, y \in E. \tag{2.8}$$

V_f is non-negative and $V_f(x, y) = D_f(x, \nabla f(y))$ for all $x, y \in E$. Moreover, by the subdifferential inequality, it is easy to see that

$$V_f(x, y) + \langle z, \nabla f^*(y) - x \rangle \leq V_f(x, z + y) \tag{2.9}$$

for all $x, y, z \in E$. In addition, if $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper lower semicontinuous, then $f^* : E \rightarrow \mathbb{R} \cup \{+\infty\}$ is proper weak lower semicontinuous and convex. Hence, V_f is convex in second variable, i.e.,

$$D_f\left(z, \nabla f^*\left(\sum_{i=1}^N t_i \nabla f(x_i)\right)\right) \leq \sum_{i=1}^N t_i D_f(z, x_i), \tag{2.10}$$

where $\{x_i\} \subset E$ and $\{t_i\} \subset (0, 1)$ with $\sum_{i=1}^N t_i = 1$.

Definition 2.3 (see [32, 33]) The Minty Variational Inequality Problem (MVI) is defined as finding a point $\bar{x} \in C$ such that

$$\langle Ay, y - \bar{x} \rangle \geq 0, \quad \forall y \in C. \tag{2.11}$$

We denote by $M(C, A)$ the set of solutions of (2.11). Some existence results for the MVI have been presented in [32]. Also, the assumption that $M(C, A) \neq \emptyset$ has already been used for solving $VI(C, A)$ in finite-dimensional spaces (see e.g. [39]). It is not difficult to prove that pseudo-monotonicity implies property $M(C, A) \neq \emptyset$, but the converse is not true.

Lemma 2.4 (see [33]) *Consider the VI (1.1). If the mapping $h : [0, 1] \rightarrow E^*$ defined as $h(t) = A(tx + (1 - t)y)$ is continuous for all $x, y \in C$ (i.e., h is hemicontinuous), then $M(C, A) \subset VI(C, A)$. Moreover, if A is pseudo-monotone, then $VI(C, A)$ is closed, convex and $VI(C, A) = M(C, A)$.*

The following lemma was proved for the case of metric projection in [11] and can also be extended to Bregman projection.

Lemma 2.5 *For every $x \in E$ and $\alpha \geq \beta > 0$, the following inequalities hold:*

$$\frac{\|x - P_C(x - \alpha Ax)\|}{\alpha} \leq \frac{\|x - P_C(x - \beta Ax)\|}{\beta} \tag{2.12}$$

and

$$\|x - P_C(x - \beta Ax)\| \leq \|x - P_C(x - \alpha Ax)\|.$$

One powerful tool for deriving weak or strong convergence of iterative sequence is due to Opial [36]. A Banach space E is said to satisfy Opial property [36] if for any weakly convergent sequence $\{x_n\}$ in E with weak limit x , we have

$$\limsup_{n \rightarrow \infty} \|x_n - x\| < \limsup_{n \rightarrow \infty} \|x_n - y\|$$

for all y in E with $y \neq x$. Note that all Hilbert space, all finite-dimensional Banach space and the Banach space l^p ($1 \leq p < \infty$) satisfy the Opial property. However, not every Banach space satisfies the Opial property; see for example [21]. But, the following Bregman Opial-like inequality for every Banach space E has been proved in [23].

Lemma 2.6 ([23]) *Let E be a Banach space and let $f : E \rightarrow (-\infty, \infty]$ be a proper strictly convex function so that it is Gâteaux differentiable and $\{x_n\}$ is a sequence in E such that $x_n \rightharpoonup u$ for some $u \in E$. Then*

$$\limsup_{n \rightarrow \infty} D_f(u, x_n) < \limsup_{n \rightarrow \infty} D_f(v, x_n)$$

for all v in the interior of $\text{dom} f$ with $u \neq v$.

Lemma 2.7 ([46]) *Let $\{a_n\}$ be a non-negative real sequence satisfying $a_{n+1} \leq (1 - \alpha_n)a_n + \alpha_n b_n$, where $\{\alpha_n\} \subset (0, 1)$, $\sum_{n=0}^\infty \alpha_n = \infty$ and $\{b_n\}$ is a sequence such that $\limsup_{n \rightarrow \infty} b_n \leq 0$. Then $\lim_{n \rightarrow \infty} a_n = 0$.*

3 Main results

In this section, we present our algorithms and discuss the convergence analysis. The main advantages of our algorithms are that the stepsize is determined by an Armijo line search technique and does not require the prior knowledge of any Lipschitz constant. We assume that the following conditions are satisfied:

- (A1) C is nonempty, closed and convex subset of a reflexive Banach space E ;
- (A2) the operator $A : E \rightarrow E^*$ is pseudo-monotone, i.e., for all $x, y \in E$, $\langle Ax, y - x \rangle \geq 0$ implies $\langle Ay, y - x \rangle \geq 0$;
- (A3) the operator A is weakly sequentially continuous, i.e., if for any sequence $\{x_n\} \subset E$, we have $x_n \rightharpoonup x$ implies $Ax_n \rightharpoonup Ax$;
- (A4) the set $VI(C, A)$ is nonempty;
- (A5) the function $f : E \rightarrow \mathbb{R} \cup \{+\infty\}$ is Legendre, uniformly Gâteaux differentiable, strongly convex, bounded on bounded subsets of E and its gradient ∇f is weak-weak continuous, i.e., $x_n \rightharpoonup x$ implies that $\nabla f(x_n) \rightharpoonup \nabla f(x)$.

First, we introduce a weak convergence Bregman subgradient extragradient method for approximating solutions of the VI in real Banach spaces.

Algorithm 3.1

Step 0: Given $\gamma > 0$, $l \in (0, 1)$, $\mu \in (0, 1)$. Let $x_1 \in E$ and set $n = 1$.

Step 1: Compute

$$y_n = \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_n Ax_n)), \tag{3.1}$$

where $\alpha_n = \gamma l^{k_n}$, with k_n being the smallest non-negative integer k satisfying

$$\gamma l^k \|Ax_n - Ay_n\| \leq \mu \|x_n - y_n\|. \tag{3.2}$$

If $x_n = y_n$ or $Ay_n = 0$, stop, y_n is a solution of the VI. Else, do Step 2.

Step 2: Compute

$$x_{n+1} = \Pi_{T_n}(\nabla f^*(\nabla f(x_n) - \alpha_n Ay_n)), \tag{3.3}$$

where T_n is the half-space defined by

$$T_n = \{w \in E : \langle \nabla f(x_n) - \alpha_n x_n - \nabla f(y_n), w - y_n \rangle \leq 0\}. \tag{3.4}$$

Set $n := n + 1$ and return to Step 1.

Remark 3.2 We note that our Algorithm 3.1 is proposed in real Banach spaces while that of Nomirovski et al. [34] and Denisov et al. [12] were proposed in finite-dimensional spaces. Furthermore, our method is more general than that of [12, 18, 34, 35] which used a fixed stepsize for all iterates. We will see in the following result that the stepsize rule defined by (3.2) is well defined.

Lemma 3.3 *There exists a non-negative integer k satisfying (3.2). In addition $0 < \alpha_n \leq \gamma$.*

Proof If $x_n \in VI(C, A)$, then $x_n = \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_n Ax_n))$ and $k_n = 0$. Hence, we consider the case where $x_n \notin VI(C, A)$ and assume the contrary, i.e. for $k > 0$,

$$\gamma l^k \|Ax_n - A\Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| > \mu \|x_n - \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\|.$$

This implies that

$$\|Ax_n - A\Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| > \frac{\|x_n - \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\|}{\gamma l^k}. \tag{3.5}$$

Next, we consider two possibilities, namely, when $x_n \in C$ and when $x_n \notin C$.

First, if $x_n \in C$, then $x_n = \Pi_C(x_n)$. Since Π_C and A are continuous,

$$\lim_{k \rightarrow \infty} \|x_n - \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| = 0.$$

Consequently, by the continuity of A on bounded sets, we get

$$\lim_{k \rightarrow \infty} \|Ax_n - A\Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| = 0. \tag{3.6}$$

Combining (3.5) and (3.6), we have

$$\lim_{k \rightarrow \infty} \frac{\|x_n - \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\|}{\gamma l^k} = 0. \tag{3.7}$$

Moreover, from the uniform continuity of ∇f on bounded subsets, we have

$$\lim_{k \rightarrow \infty} \frac{\|\nabla f(x_n) - \nabla f(\Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n)))\|}{\gamma l^k} = 0. \tag{3.8}$$

Now let $z_n = \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))$, then, by (2.6), we obtain

$$\langle \nabla f(z_k) - \nabla f(x_n) + \gamma l^k Ax_n, x - z_n \rangle \geq 0, \quad \forall x \in C.$$

This means that

$$\langle Ax_n, x - z_n \rangle \geq \left\langle \frac{\nabla f(x_n) - \nabla f(z_n)}{\gamma l^k}, x - z_n \right\rangle, \quad \forall x \in C.$$

Hence by taking the limit as $n \rightarrow \infty$ and from (3.8), we get

$$\langle Ax_n, x - x_n \rangle \geq 0, \quad \forall x \in C.$$

Therefore, $x_n \in VI(C, A)$. This is a contradiction.

On the other hand, if $x_n \notin C$, then

$$\lim_{k \rightarrow \infty} \|x_n - \Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| = \|x_n - \Pi_C(x_n)\| > 0 \tag{3.9}$$

and

$$\lim_{k \rightarrow \infty} \gamma l^k \|Ax_n - A\Pi_C(\nabla f^*(\nabla f(x_n) - \gamma l^k Ax_n))\| = 0. \tag{3.10}$$

Combining (3.5), (3.9) and (3.10), we get a contradiction. □

We now present the convergence analysis of Algorithm 3.1. The following lemma will be used in the sequel.

Lemma 3.4 *Assume Conditions (A1)–(A5) hold. Let $\{x_n\}$ and $\{y_n\}$ be the sequences generated by Algorithm 3.1. Then*

$$D_f(p, x_{n+1}) \leq D_f(p, x_n) - (1 - \mu)D_f(y_n, x_n) - (1 - \mu)D_f(x_{n+1}, y_n)$$

for any $p \in VI(C, A)$.

Proof Since $x_{n+1} \in T_n$, it follows from (2.7) that

$$\begin{aligned} D_f(p, x_{n+1}) &= D_f(p, \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_n Ay_n))) \\ &\leq D_f(p, \nabla f^*(\nabla f(x_n) - \alpha_n Ay_n)) - D_f(x_{n+1}, \nabla f^*(\nabla f(x_n) - \alpha_n Ay_n)) \\ &= V_f(p, \nabla f(x_n) - \alpha_n Ay_n) - V_f(x_{n+1}, \nabla f(x_n) - \alpha_n Ay_n) \\ &= f(p) - \langle p, \nabla f(x_n) - \alpha_n Ay_n \rangle + f^*(\nabla f(x_n) - \alpha_n Ay_n) \\ &\quad - f(x_{n+1}) + \langle x_{n+1}, \nabla f(x_n) - \alpha_n Ay_n \rangle - f^*(\nabla f(x_n) - \alpha_n Ay_n) \\ &= f(p) - \langle p, \nabla f(x_n) \rangle + f^*(x_n) - f(x_{n+1}) + \langle x_{n+1}, \nabla f(x_n) \rangle - f^*(x_n) \\ &\quad + \alpha_n \langle p, Ay_n \rangle - \alpha_n \langle x_{n+1}, Ay_n \rangle \\ &= D_f(p, x_n) - D_f(x_{n+1}, x_n) + \alpha_n \langle p - x_{n+1}, Ay_n \rangle. \end{aligned} \tag{3.11}$$

Since A is pseudo-monotone and $p \in VI(C, A)$,

$$\langle Ay_n, y_n - p \rangle \geq 0. \tag{3.12}$$

Hence, using (3.12) and (2.4) in (3.11), we get

$$\begin{aligned} D_f(p, x_{n+1}) &\leq D_f(p, x_n) - D_f(x_{n+1}, x_n) + \alpha_n \langle p - y_n, Ay_n \rangle + \alpha_n \langle y_n - x_{n+1}, Ay_n \rangle \\ &\leq D_f(p, x_n) - D_f(x_{n+1}, x_n) + \alpha_n \langle y_n - x_{n+1}, Ay_n \rangle \\ &= D_f(p, x_n) - D_f(x_{n+1}, y_n) - D_f(y_n, x_n) + \langle \nabla f(x_n) - \nabla f(y_n), x_{n+1} - y_n \rangle \\ &\quad + \alpha_n \langle y_n - x_{n+1}, Ay_n \rangle \\ &= D_f(p, x_n) - D_f(x_{n+1}, y_n) - D_f(y_n, x_n) \\ &\quad + \langle \nabla f(x_n) - \alpha_n Ay_n - \nabla f(y_n), x_{n+1} - y_n \rangle. \end{aligned} \tag{3.13}$$

Now, we estimate the last variable in (3.13) as follows:

$$\begin{aligned} \langle \nabla f(x_n) - \alpha_n A y_n - \nabla f(y_n), x_{n+1} - y_n \rangle &= \langle \nabla f(x_n) - \alpha_n A x_n - \nabla f(y_n), x_{n+1} - y_n \rangle \\ &\quad + \alpha_n \langle A x_n - A y_n, x_{n+1} - y_n \rangle \\ &\leq \alpha_n \langle A x_n - A y_n, x_{n+1} - y_n \rangle. \end{aligned} \tag{3.14}$$

Using the Cauchy–Schwartz inequality and (3.2), we have

$$\begin{aligned} \langle \nabla f(x_n) - \alpha_n A y_n - \nabla f(y_n), x_{n+1} - y_n \rangle &\leq \alpha_n \|A x_n - A y_n\| \|x_{n+1} - y_n\| \\ &\leq \mu \|x_n - y_n\| \|x_{n+1} - y_n\| \\ &\leq \frac{\mu}{2} (\|y_n - x_n\|^2 + \|x_{n+1} - y_n\|^2) \\ &\leq \mu (D_f(y_n, x_n) + D_f(x_{n+1}, y_n)). \end{aligned} \tag{3.15}$$

Substituting (3.15) into (3.13), we get

$$D_f(p, x_{n+1}) \leq D_f(p, x_n) - (1 - \mu)D_f(y_n, x_n) - (1 - \mu)D_f(x_{n+1}, y_n). \quad \square$$

Theorem 3.5 *Assume that Conditions (A1)–(A5) holds and $\liminf_{n \rightarrow \infty} \alpha_n > 0$. Then any sequence $\{x_n\}$ generated by Algorithm 3.1 converges weakly to an element of $VI(C, A)$.*

Proof Claim 1: $\{x_n\}$ is bounded. Indeed, let $p \in VI(C, A)$, we have from Lemma 3.4

$$D_f(p, x_{n+1}) \leq D_f(p, x_n).$$

This implies that $\{D_f(p, x_n)\}$ is bounded and nonincreasing, thus, $\lim_{n \rightarrow \infty} D_f(p, x_n)$ exists. Hence

$$\lim_{n \rightarrow \infty} (D_f(p, x_n) - D_f(p, x_{n+1})) = 0. \tag{3.16}$$

Moreover, due to (2.5), we see that $\{x_n\}$ is bounded. Consequently $\{y_n\}$ is bounded.

Claim 2: $\lim_{n \rightarrow \infty} \|x_n - y_n\| = \lim_{n \rightarrow \infty} \|x_{n+1} - y_n\| = \lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0$. Indeed, from Lemma 3.4, we get

$$(1 - \mu)D_f(y_n, x_n) + (1 - \mu)D_f(x_{n+1}, y_n) \leq D_f(p, x_n) - D_f(p, x_{n+1}).$$

Since $\mu \in (0, 1)$, it follows from (3.16) that

$$\lim_{n \rightarrow \infty} D_f(y_n, x_n) = \lim_{n \rightarrow \infty} D_f(x_{n+1}, y_n) = 0. \tag{3.17}$$

Hence from (2.5), we get

$$\lim_{n \rightarrow \infty} \|x_n - y_n\| = \lim_{n \rightarrow \infty} \|x_{n+1} - y_n\| = 0, \tag{3.18}$$

and

$$\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0.$$

Claim 3: $\{x_n\}$ weakly converges to an element of $VI(C, A)$. Indeed, since $\{x_n\}$ is a bounded sequence, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup z \in E$. From the fact that $\lim_{n \rightarrow \infty} \|x_n - y_n\| = 0$, we obtain $y_{n_k} \rightharpoonup z$, where $\{y_{n_k}\}$ is a subsequence of $\{y_n\}$. Since

$$y_{n_k} = \Pi_C(\nabla f^*(\nabla f(x_{n_k}) - \alpha_{n_k}Ax_{n_k})),$$

it follows from (2.6) that

$$\langle \nabla f(x_{n_k}) - \alpha_{n_k}Ax_{n_k} - \nabla f(y_{n_k}), y - y_{n_k} \rangle \leq 0, \quad \forall y \in C. \tag{3.19}$$

This means that

$$\frac{1}{\alpha_{n_k}} \langle \nabla f(x_{n_k}) - \nabla f(y_{n_k}), y - y_{n_k} \rangle \geq \langle Ax_{n_k}, y - y_{n_k} \rangle, \quad \forall y \in C.$$

Hence

$$\frac{1}{\alpha_{n_k}} \langle \nabla f(x_{n_k}) - \nabla f(y_{n_k}), y - y_{n_k} \rangle + \langle Ax_{n_k}, y_{n_k} - x_{n_k} \rangle \leq \langle Ax_{n_k}, y - x_{n_k} \rangle, \quad \forall y \in C.$$

Now we show that

$$\liminf_{k \rightarrow \infty} \langle Ax_{n_k}, y - x_{n_k} \rangle \geq 0. \tag{3.20}$$

We consider two possible cases. First suppose that $\liminf_{k \rightarrow \infty} \alpha_{n_k} > 0$, since $\|x_{n_k} - y_{n_k}\| \rightarrow 0$ as $k \rightarrow \infty$, by the weak-weak continuity of ∇f , we have $\|\nabla f(x_{n_k}) - \nabla f(y_{n_k})\| \rightarrow 0$ as $k \rightarrow \infty$. Taking the limit of the above inequality as $k \rightarrow \infty$ we get

$$\liminf_{k \rightarrow \infty} \langle Ax_{n_k}, y - x_{n_k} \rangle \geq 0, \quad \forall y \in C. \tag{3.21}$$

On the other hand, suppose $\liminf_{k \rightarrow \infty} \alpha_{n_k} = 0$. Let $z_{n_k} = \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_{n_k}l^{-1}Ax_{n_k}))$, we have $\alpha_{n_k}l^{-1} > \alpha_{n_k}$ and by using Lemma 2.5, we obtain

$$\|x_{n_k} - z_{n_k}\| \leq \frac{1}{l} \|x_{n_k} - y_{n_k}\| \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

Furthermore, $z_{n_k} \rightharpoonup z \in C$, which implies that $\{z_{n_k}\}$ is a bounded sequence. By the uniform continuity of A , we have

$$\|Ax_{n_k} - Az_{n_k}\| \rightarrow 0 \quad \text{as } k \rightarrow \infty. \tag{3.22}$$

Using the Armijo line search rule, we get

$$\frac{1}{\mu} \|Ax_{n_k} - Az_{n_k}\| > \frac{\|x_{n_k} - z_{n_k}\|}{\alpha_{n_k}l^{-1}}. \tag{3.23}$$

Combining (3.22) and (3.23), we get

$$\lim_{k \rightarrow \infty} \frac{\|x_{n_k} - z_{n_k}\|}{\alpha_{n_k}l^{-1}} = 0.$$

Moreover,

$$\langle \nabla f(x_{n_k}) - \alpha_{n_k} l^{-1} Ax_{n_k} - \nabla f(z_{n_k}), x - z_{n_k} \rangle \leq 0, \quad \forall x \in C.$$

Hence

$$\frac{1}{\alpha_{n_k} l^{-1}} \langle \nabla f(x_{n_k}) - \nabla f(z_{n_k}), x - z_{n_k} \rangle + \langle Ax_{n_k}, z_{n_k} - x_{n_k} \rangle \leq \langle Ax_{n_k}, x - x_{n_k} \rangle, \quad \forall x \in C.$$

Taking the limit of the above inequality as $k \rightarrow \infty$, we get

$$\liminf_{k \rightarrow \infty} \langle Ax_{n_k}, x - x_{n_k} \rangle \geq 0.$$

Thus, the inequality is proven.

Now choose a sequence $\{\epsilon_k\} \subset (0, 1)$ such that $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$. For each $k \geq 1$, there exists a smallest number $N \in \mathbb{N}$ satisfying

$$\langle Ax_{n_k}, y - x_{n_k} \rangle + \epsilon_k \geq 0, \quad \forall k \geq N.$$

This implies that

$$\langle Ax_{n_k}, y + \epsilon_k t_{n_k} - x_{n_k} \rangle \geq 0, \quad \forall k \geq N$$

for some $t_{n_k} \in E$ satisfying $1 = \langle Ax_{n_k}, t_{n_k} \rangle$ (since $Ax_{n_k} \neq 0$). Since A is pseudo-monotone, we have

$$\langle A(y + \epsilon_k t_{n_k}), y + \epsilon_k t_{n_k} - x_{n_k} \rangle \geq 0, \quad \forall k \geq N.$$

This implies that

$$\langle Ay, y - x_{n_k} \rangle \geq \langle Ay - A(y + \epsilon_k t_{n_k}), y + \epsilon_k t_{n_k} - x_{n_k} \rangle - \epsilon_k \langle Ay, t_{n_k} \rangle, \quad \forall k \geq N. \tag{3.24}$$

Since $\epsilon_k \rightarrow 0$ and A is continuous, then the right hand side of (3.24) tends to zero. Thus, we obtain

$$\liminf_{k \rightarrow \infty} \langle Ay, y - x_{n_k} \rangle \geq 0, \quad \forall y \in C.$$

Hence

$$\langle Ay, y - z \rangle = \lim_{k \rightarrow \infty} \langle Ay, y - x_{n_k} \rangle \geq 0, \quad \forall y \in C.$$

Therefore, from Lemma 2.4, we obtain $z \in VI(C, A)$.

Finally, we show that z is unique. Assume the contrary, i.e., there exists a subsequence $\{x_{n_j}\}$ of $\{x_n\}$ such that $x_{n_j} \rightharpoonup \hat{z}$ with $\hat{z} \neq z$. Following a similar argument to the one above, we get $\hat{z} \in VI(C, A)$. It follows from the Bregman Opial-like property of H (more precisely,

Lemma 2.6) that

$$\begin{aligned} \lim_{n \rightarrow \infty} D_f(z, x_n) &= \lim_{k \rightarrow \infty} D_f(z, x_{n_k}) < \lim_{k \rightarrow \infty} D_f(\hat{z}, x_{n_k}) \\ &= \lim_{n \rightarrow \infty} D_f(\hat{z}, x_n) = \lim_{j \rightarrow \infty} D_f(\hat{z}, x_{n_j}) \\ &< \lim_{j \rightarrow \infty} D_f(z, x_{n_j}) = \lim_{n \rightarrow \infty} D_f(z, x_n), \end{aligned}$$

which is a contradiction. Thus, we have $z = \hat{z}$ and the desired result follows. This completes the proof. □

Next, we propose a strong convergence Bregman subgradient extragradient algorithm with Halpern iterative method [22] for solving the VI (1.1) in real Banach spaces. This is important for supporting the infinite-dimensional setting of our work.

Algorithm 3.6

Step 0: Given $\gamma > 0, l \in (0, 1), \mu \in (0, 1), \{\delta_n\} \subset (0, 1)$. Let $x_1, u \in E$ and set $n = 1$.

Step 1: Compute

$$y_n = \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_n A x_n)), \tag{3.25}$$

where $\alpha_n = \gamma l^{k_n}$, with k_n being the smallest non-negative integer k satisfying

$$\gamma l^k \|A x_n - A y_n\| \leq \mu \|x_n - y_n\|. \tag{3.26}$$

If $x_n = y_n$ or $A y_n = 0$, stop, y_n is a solution of the VI. Else, do Step 2.

Step 2: Compute

$$z_n = \Pi_{T_n}(\nabla f^*(\nabla f(x_n) - \alpha_n A y_n)), \tag{3.27}$$

where T_n is the half-space defined by

$$T_n = \{w \in E : \langle \nabla f(x_n) - \alpha_n x_n - \nabla f(y_n), w - y_n \rangle \leq 0\}. \tag{3.28}$$

Step 3: Compute

$$x_{n+1} = \nabla f^*(\delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n)). \tag{3.29}$$

Set $n := n + 1$ and return to Step 1.

For proving the convergence of Algorithm 3.6, we assume that the following condition is satisfied.

(C1) $\lim_{n \rightarrow \infty} \delta_n = 0$ and $\sum_{n=0}^{\infty} \delta_n = +\infty$.

We first prove the following lemmas which are crucial for our main theorem.

Lemma 3.7 *The sequence $\{x_n\}$ generated by Algorithm 3.6 is bounded.*

Proof Let $p \in VI(C, A)$, then

$$\begin{aligned}
 D_f(p, z_n) &= D_f(p, \Pi_C(\nabla f^*(\nabla f(x_n) - \alpha_n Ay_n))) \\
 &\leq D_f(p, \nabla f^*(\nabla f(x_n) - \alpha_n Ay_n)) - D_f(z_n, \nabla f^*(\nabla f(x_n) - \alpha_n Ay_n)) \\
 &= V_f(p, \nabla f(x_n) - \alpha_n Ay_n) - V_f(z_n, \nabla f(x_n) - \alpha_n Ay_n) \\
 &= f(p) - \langle p, \nabla f(x_n) - \alpha_n Ay_n \rangle + f^*(\nabla f(x_n) - \alpha_n Ay_n) \\
 &\quad - f(z_n) + \langle z_n, \nabla f(x_n) - \alpha_n Ay_n \rangle - f^*(\nabla f(x_n) - \alpha_n Ay_n) \\
 &= f(p) - \langle p, \nabla f(x_n) \rangle + f^*(x_n) - f(z_n) + \langle z_n, \nabla f(x_n) \rangle - f^*(x_n) \\
 &\quad + \alpha_n \langle p, Ay_n \rangle - \alpha_n \langle z_n, Ay_n \rangle \\
 &= D_f(p, x_n) - D_f(z_n, x_n) + \alpha_n \langle p - z_n, Ay_n \rangle. \tag{3.30}
 \end{aligned}$$

Since p is a solution of VI (1.1), we have $\langle Ap, x - p \rangle \geq 0$ for all $x \in C$. By the pseudo-monotonicity of A on H , we get $\langle Ax, x - p \rangle \geq 0$ for all $x \in C$. Hence $\langle Ay_n, p - y_n \rangle \geq 0$. Thus, we have

$$\langle Ay_n, p - z_n \rangle = \langle Ay_n, p - y_n \rangle + \langle Ay_n, y_n - z_n \rangle \leq \langle Ay_n, y_n - z_n \rangle. \tag{3.31}$$

From (3.30) and (3.31), we have

$$\begin{aligned}
 D_f(p, z_n) &\leq D_f(p, x_n) - D_f(z_n, x_n) + \langle Ay_n, y_n - z_n \rangle \\
 &= D_f(p, x_n) - D_f(z_n, y_n) - D_f(y_n, x_n) + \langle \nabla f(x_n) - \nabla f(y_n), z_n - y_n \rangle \\
 &\quad + \alpha_n \langle y_n - z_n, Ay_n \rangle \\
 &= D_f(p, x_n) - D_f(z_n, y_n) - D_f(y_n, x_n) \\
 &\quad + \langle \nabla f(x_n) - \alpha_n Ay_n - \nabla f(y_n), z_n - y_n \rangle. \tag{3.32}
 \end{aligned}$$

Note that

$$\begin{aligned}
 \langle \nabla f(x_n) - \alpha_n Ay_n - \nabla f(y_n), z_n - y_n \rangle &= \langle \nabla f(x_n) - \alpha_n Ax_n - \nabla f(y_n), z_n - y_n \rangle \\
 &\quad + \alpha_n \langle Ax_n - Ay_n, z_n - y_n \rangle \\
 &\leq \alpha_n \langle Ax_n - Ay_n, z_n - y_n \rangle \\
 &\leq \alpha_n \|Ax_n - Ay_n\| \|z_n - y_n\| \\
 &\leq \mu \|x_n - y_n\| \|z_n - y_n\| \\
 &\leq \frac{\mu}{2} (\|y_n - x_n\|^2 + \|z_n - y_n\|^2) \\
 &\leq \mu (D_f(y_n, x_n) + D_f(z_n, y_n)). \tag{3.33}
 \end{aligned}$$

Hence from (3.32) and (3.33), we get

$$\begin{aligned}
 D_f(p, z_n) &\leq D_f(p, x_n) - (1 - \mu)D_f(y_n, x_n) - (1 - \mu)D_f(z_n, y_n) \\
 &\leq D_f(p, x_n). \tag{3.34}
 \end{aligned}$$

Furthermore, from (3.29), we have

$$\begin{aligned}
 D_f(p, x_{n+1}) &= D_f(p, \nabla f^*(\delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n))) \\
 &\leq \delta_n D_f(p, u) + (1 - \delta_n) D_f(p, z_n) \\
 &\leq \delta_n D_f(p, u) + (1 - \delta_n) D_f(p, x_n) \\
 &\leq \max\{D_f(p, u), D_f(p, x_n)\} \\
 &\vdots \\
 &\leq \max\{D_f(p, u), D_f(p, x_1)\}.
 \end{aligned}$$

This implies that $\{D_f(p, x_n)\}$ is bounded. Hence, $\{x_n\}$ is bounded. Consequently, we see that $\{\nabla f(x_n)\}, \{y_n\}, \{z_n\}$ are bounded. □

Lemma 3.8 *The sequence $\{x_n\}$ generated by Algorithm 3.6 satisfies the following estimates:*

- (i) $s_{n+1} \leq (1 - \delta)s_n + \delta_n b_n,$
- (ii) $-1 \leq \limsup_{n \rightarrow \infty} b_n < +\infty,$

where $s_n = D_f(p, x_n), b_n = \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle$ for all $p \in VI(C, A)$.

Proof From (2.9) we have

$$\begin{aligned}
 D_f(p, x_{n+1}) &\leq D_f(p, \nabla f^*(\delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n))) \\
 &= V_f(p, \delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n)) \\
 &\leq V_f(p, \delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n) - \delta_n (\nabla f(u) - \nabla f(p))) \\
 &\quad - \langle -\delta_n (\nabla f(u) - \nabla f(p)), x_{n+1} - p \rangle \\
 &= V_f(p, \delta_n \nabla f(p) + (1 - \delta_n) \nabla f(z_n)) + \delta_n \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle \\
 &\leq \delta_n V_f(p, p) + (1 - \delta_n) V_f(p, \nabla f(z_n)) + \delta_n \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle \\
 &= (1 - \delta_n) D_f(p, z_n) + \delta_n \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle \\
 &\leq (1 - \delta_n) D_f(p, x_n) + \delta_n \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle. \tag{3.35}
 \end{aligned}$$

This established (i). Next, we show (ii). Since $\{x_n\}$ is bounded and $\delta_n \in (0, 1)$, we have

$$\sup_{n \geq 0} b_n \leq \sup_{n \geq 0} \|\nabla f(u) - \nabla f(p)\| \|x_{n+1} - p\| < \infty.$$

We now show that $\limsup_{n \rightarrow \infty} b_n \geq -1$. Assume the contrary, i.e., there exists $n_0 \in \mathbb{N}$ such that $b_n \geq -1$ for all $n \geq n_0$. Hence it follows that

$$\begin{aligned}
 s_{n+1} &\leq (1 - \delta_n) s_n + \delta_n b_n \\
 &< (1 - \delta_n) s_n - \delta_n \\
 &= s_n - \delta_n (s_n + 1) \\
 &\leq s_n - \delta_n.
 \end{aligned}$$

By induction, we obtain

$$s_{n+1} \leq s_{n_0} - \sum_{i=1}^n \delta_i, \quad \forall n \geq n_0.$$

Taking the lim sup of the above inequality, we have

$$\limsup_{n \rightarrow \infty} s_n \leq s_{n_0} - \lim_{n \rightarrow \infty} \sum_{i=1}^n \delta_i = -\infty.$$

This contradicts the fact that $\{s_n\}$ is a non-negative real sequence. Thus $\limsup_{n \rightarrow \infty} b_n \geq -1$. □

Next, we present our strong convergence theorem.

Theorem 3.9 *Assume Conditions (A1)–(A5) and (C1) hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.6 converges strongly to an element in $VI(C, A)$.*

Proof Let $p \in VI(C, A)$ and $\Gamma_n = D_f(p, x_n)$. We divide the proof into two cases.

Case I: Suppose that there exists $n_0 \in \mathbb{R}$ such that $\{\Gamma_n\}$ is monotonically non-increasing for $n \geq n_0$. Since $\{\Gamma_n\}$ is bounded (see Lemma 3.7), $\{\Gamma_n\}$ converges and therefore

$$\Gamma_n - \Gamma_{n+1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

From (3.34), we have

$$\begin{aligned} D_f(p, x_{n+1}) &\leq \delta_n D_f(p, u) + (1 - \delta_n) D_f(p, z_n) \\ &\leq \delta_n D_f(p, u) + (1 - \delta_n) [D_f(p, x_n) - (1 - \mu) D_f(z_n, y_n) - (1 - \mu) D_f(y_n, x_n)]. \end{aligned}$$

This implies that

$$(1 - \delta_n)(1 - \mu) [D_f(z_n, y_n) + D_f(y_n, x_n)] \leq \delta_n (D_f(p, u) - D_f(p, x_n)) + D_f(p, x_n) - D_f(p, x_{n+1}).$$

Since $\delta_n \rightarrow 0$, we have

$$\lim_{n \rightarrow \infty} (1 - \mu) [D_f(z_n, y_n) + D_f(y_n, x_n)] = 0.$$

Hence

$$\lim_{n \rightarrow \infty} D_f(z_n, y_n) = \lim_{n \rightarrow \infty} D_f(y_n, x_n) = 0.$$

Using (2.5), we obtain

$$\lim_{n \rightarrow \infty} \|z_n - y_n\| = \lim_{n \rightarrow \infty} \|y_n - x_n\| = 0.$$

This implies that

$$\lim_{n \rightarrow \infty} \|z_n - x_n\| = 0. \tag{3.36}$$

Furthermore

$$\begin{aligned}
 Df(z_n, x_{n+1}) &= Df(z_n, \nabla f^*(\delta_n \nabla f(u) + (1 - \delta_n) \nabla f(z_n))) \\
 &\leq \delta_n Df(z_n, u) + (1 - \delta_n) Df(z_n, z_n) \rightarrow 0.
 \end{aligned}$$

Thus

$$\lim_{n \rightarrow \infty} \|z_n - x_{n+1}\| = 0. \tag{3.37}$$

Therefore from (3.36) and (3.37), we get

$$\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = 0.$$

Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightarrow \bar{x}$. We now show that $\bar{x} \in VI(C, A)$. From

$$y_{n_k} = \Pi_C(\nabla f^*(\nabla f(x_{n_k}) - \alpha_{n_k} Ax_{n_k})),$$

we have

$$\langle \nabla f(x_{n_k}) - \alpha_{n_k} Ax_{n_k} - \nabla f(y_{n_k}), x - y_{n_k} \rangle \leq 0, \quad \forall x \in C. \tag{3.38}$$

This implies that

$$\frac{1}{\alpha_{n_k}} \langle \nabla f(x_{n_k}) - \nabla f(y_{n_k}), x - y_{n_k} \rangle \geq \langle Ax_{n_k}, x - y_{n_k} \rangle, \quad \forall x \in C.$$

Hence

$$\frac{1}{\alpha_{n_k}} \langle \nabla f(x_{n_k}) - \nabla f(y_{n_k}), x - y_{n_k} \rangle + \langle Ax_{n_k}, y_{n_k} - x_{n_k} \rangle \leq \langle Ax_{n_k}, x - x_{n_k} \rangle, \quad \forall x \in C.$$

Following a similar approach to Theorem 3.5, we can show that

$$\liminf_{k \rightarrow \infty} \langle Ax_{n_k}, x - x_{n_k} \rangle \geq 0, \quad \forall x \in C. \tag{3.39}$$

Let $\{\epsilon_k\}$ be a sequence in $(0, 1)$ such that $\epsilon_k \rightarrow 0$ as $k \rightarrow \infty$. For each $k \geq 1$, there exists a smallest number $N \in \mathbb{N}$ satisfying

$$\langle Ax_{n_k}, x - x_{n_k} \rangle + \epsilon_k \geq 0, \quad \forall k \geq N.$$

This implies that

$$\langle Ax_{n_k}, x + \epsilon_k t_{n_k} - x_{n_k} \rangle \geq 0, \quad \forall k \geq N$$

for some $t_{n_k} \in E$ satisfying $1 = \langle Ax_{n_k}, t_{n_k} \rangle$ (since $Ax_{n_k} \neq 0$). Since A is pseudo-monotone, we have

$$\langle A(x + \epsilon_k t_{n_k}), x + \epsilon_k t_{n_k} - x_{n_k} \rangle \geq 0, \quad \forall k \geq N.$$

Thus

$$\langle Ax, x - x_{n_k} \rangle \geq \langle Ax - A(x + \epsilon_k t_{n_k}), x + \epsilon_k t_{n_k} - x_{n_k} \rangle - \epsilon_k \langle Ax, t_{n_k} \rangle, \quad \forall k \geq N. \tag{3.40}$$

Since $\epsilon_k \rightarrow 0$ and A is continuous, the right hand side of (3.40) tends to zero. Thus, we obtain

$$\liminf_{k \rightarrow \infty} \langle Ax, x - x_{n_k} \rangle \geq 0, \quad \forall x \in C.$$

Hence

$$\langle Ax, x - \bar{x} \rangle = \lim_{k \rightarrow \infty} \langle Ax, x - x_{n_k} \rangle \geq 0, \quad \forall x \in C.$$

Therefore, from Lemma 2.4, we obtain $\bar{x} \in VI(C, A)$. We now show that $\{x_n\}$ converges strongly to p . It suffices to show that $\limsup_{n \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle \leq 0$. To do this, choose a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that

$$\limsup_{n \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle = \lim_{k \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{n_k+1} - p \rangle. \tag{3.41}$$

Since $\|x_{n_k+1} - x_{n_k}\| \rightarrow 0$ and $x_{n_k} \rightarrow \bar{x}$ as $k \rightarrow \infty$, we have from (2.6) and (3.41)

$$\begin{aligned} \limsup_{n \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{n+1} - p \rangle &= \lim_{k \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{n_k+1} - p \rangle \\ &= \langle \nabla f(u) - \nabla f(p), \bar{x} - p \rangle \\ &\leq 0. \end{aligned} \tag{3.42}$$

Using Lemma 2.7, Lemma 3.8(i) and (3.42), we obtain $\lim_{n \rightarrow \infty} D_f(p, x_n) = 0$. This implies that $\|x_n - p\| \rightarrow 0$, hence, $\{x_n\}$ converges strongly to p . Consequently, $\{y_n\}$ and $\{z_n\}$ converges strongly to p .

Case II: Suppose $\{D_f(p, x_n)\}$ is not monotonically decreasing. Let $\tau : \mathbb{N} \rightarrow \mathbb{N}$ for all $n \geq n_0$ (for some n_0 large enough) be defined by

$$\tau_n = \max\{k \in \mathbb{N} : \tau_k \leq \tau_{k+1}\}.$$

Clearly, τ is nondecreasing, $\tau(n) \rightarrow \infty$ as $n \rightarrow \infty$ and

$$0 \leq D_f(p, x_{\tau(n)}) \leq D_f(p, x_{\tau(n)+1}), \quad \forall n \geq n_0.$$

Following a similar argument to Case I, we obtain

$$\|x_{\tau(n)} - y_{\tau(n)}\| \rightarrow 0, \quad \|v_{\tau(n)} - Tv_{\tau(n)}\| \rightarrow 0, \quad \|x_{\tau(n)+1} - x_{\tau(n)}\| \rightarrow 0$$

as $n \rightarrow \infty$ and $\Omega_w(x_{\tau(n)}) \subset VI(C, A)$, where $\Omega_w(x_{\tau(n)})$ is the weak subsequential limit of $\{x_{\tau(n)}\}$. Also,

$$\limsup_{n \rightarrow \infty} \langle \nabla f(u) - \nabla f(p), x_{\tau(n)+1} - p \rangle \leq 0. \tag{3.43}$$

From Lemma 3.8(i), we have

$$D_f(p, x_{\tau(n+1)}) \leq (1 - \delta_{\tau(n)})D_f(p, x_{\tau(n)}) + \delta_{\tau(n)}\langle \nabla f(u) - \nabla f(p), x_{\tau(n+1)} - p \rangle.$$

Since $D_f(p, x_{\tau(n)}) \leq D_f(p, x_{\tau(n+1)})$,

$$\begin{aligned} 0 &\leq D_f(p, x_{\tau(n+1)}) - D_f(p, x_{\tau(n)}) \\ &\leq (1 - \delta_{\tau(n)})D_f(p, x_{\tau(n)}) + \delta_{\tau(n)}\langle \nabla f(u) - \nabla f(p), x_{\tau(n+1)} - p \rangle - D_f(p, x_{\tau(n)}). \end{aligned}$$

Hence, from (3.42), we get

$$D_f(p, x_{\tau(n)}) \leq \langle \nabla f(u) - \nabla f(p), x_{\tau(n+1)} - p \rangle \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

As a consequence, we obtain, for all $n \geq n_0$,

$$0 \leq D_f(p, x_n) \leq \max\{D_f(p, x_{\tau(n)}), D_f(p, x_{\tau(n+1)})\} = D_f(p, x_{\tau(n+1)}).$$

Thus

$$D_f(p, x_n) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Therefore, from (2.5)

$$\lim_{n \rightarrow \infty} \|x_n - p\| = 0.$$

This implies that $\{x_n\}$ converges strongly to p . This completes the proof. □

Remark 3.10

- (i) We note that our results extend the results of [12, 35] from finite-dimensional spaces to real Banach spaces.
- (ii) We also extend the result of Gibali [18] to solving pseudo-monotone variational inequalities and real Banach spaces.
- (iii) Moreover, our algorithms does not require any prior estimate of the Lipschitz constant for their convergence. This improves the corresponding results of [7, 8, 12, 18, 35] in the literature.
- (iv) The strong convergence theorem proved in this paper is more desirable for solving optimization problems.

Remark 3.11 The operator A is said to be monotone if $\langle Ax - Ay, x - y \rangle \geq 0$ for all $x, y \in E$. It is easy to see that every monotone operator is pseudo-monotone but the converse is not true (see Example 4.2). Moreover, we give the following example of a variational inequality problem satisfying assumptions (A2)–(A4), but not Lipschitz continuous (see also [43]).

Example 3.12 Let $E = \ell_2(\mathbb{R})$, $C = \{x = (x_1, x_2, \dots, x_i, \dots) \in E : |x_i| \leq \frac{1}{i}, \forall i = 1, 2, \dots\}$ and

$$Ax = \left(\|x\| + \frac{1}{\|x\| + 1} \right) x.$$

It is easy to see that $VI(C, A) \neq \emptyset$ since $0 \in VI(C, A)$, A is pseudo-monotone, sequentially weakly continuous but not Lipschitz continuous on E . Indeed, let $u, v \in C$ be such that $\langle Ax, y - x \rangle \geq 0$. This means that $\langle x, y - x \rangle \geq 0$. Consequently,

$$\begin{aligned} \langle Ay, y - x \rangle &= \left(\|y\| + \frac{1}{\|y\| + 1} \right) \langle y, y - x \rangle \\ &> \left(\|y\| + \frac{1}{\|y\| + 1} \right) (\langle y, y - x \rangle - \langle x, y - x \rangle) \\ &= \left(\|y\| + \frac{1}{\|y\| + 1} \right) \|y - x\|^2 > 0. \end{aligned}$$

Hence, A is pseudo-monotone. Also, since A is compact, it is uniformly continuous and thus sequentially weakly continuous on E . To see that A is not L -Lipschitz continuous, let $x = (L, 0, \dots, 0, \dots)$ and $y = (0, 0, \dots, 0, \dots)$, then

$$\|Ax - Ay\| = \|Ax\| = \left(\|x\| + \frac{1}{\|x\| + 1} \right) \|x\| = \left(L + \frac{1}{L + 1} \right) L.$$

Moreover, $\|Ax - Ay\| \leq L\|x - y\|$ is equivalent to

$$\left(L + \frac{1}{L + 1} \right) L \leq L^2.$$

This implies that

$$\frac{1}{L + 1} \leq 0.$$

This is a contradiction, and thus A is not Lipschitz continuous on E .

4 Numerical illustrations

In this section, we present some numerical examples to illustrate the convergence and efficiency of the proposed algorithms. The projection onto C is computed effectively by using the function *quadprog* in Matlab optimization toolbox, while the projection onto the half-space is calculated explicitly. All program computation are performed on a Lenovo PC Intel(R) Core i7, 4.00 GB RAM. The stopping criterion used for the examples is $\frac{\|x_{l+1} - x_l\|^2}{\|x_2 - x_1\|^2} < \varepsilon$, where ε is stated in each example.

Example 4.1 Consider an operator $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ defined by $Ax = Mx + q$ with q being a vector in \mathbb{R}^m and

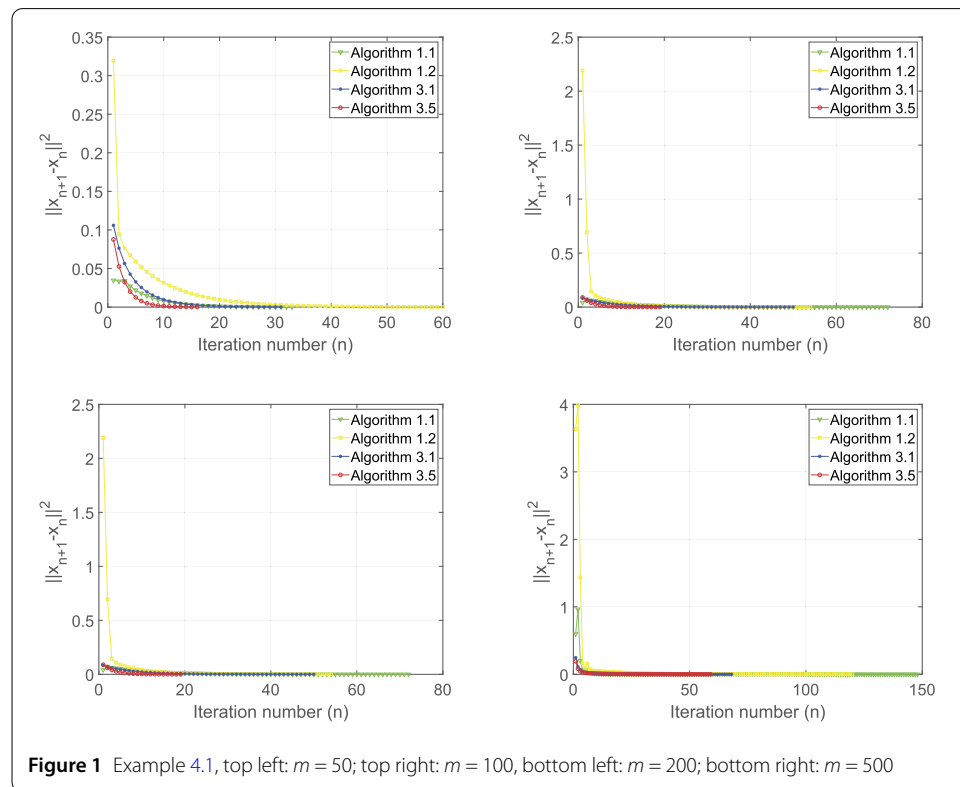
$$M = NN^T + S + D,$$

where N is a $m \times m$ matrix, S is a $m \times m$ skew-symmetric matrix and D is a $m \times m$ diagonal matrix with its diagonal entries being non-negative (so that M is positive semidefinite). The feasible set C in this case is defined by

$$C := \{x = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m : -5 \leq x_i \leq 5, i = 1, 2, \dots, m\}.$$

Table 1 Computation result for Example 4.1

		Algorithm 1.1	Algorithm 1.2	Algorithm 3.1	Algorithm 3.6
$m = 50$	No of Iter.	33	60	31	16
	CPU time (sec)	0.0074	0.0184	0.0099	0.0040
$m = 100$	No of Iter.	60	48	29	19
	CPU time (sec)	0.0169	0.0143	0.0094	0.0063
$m = 200$	No of Iter.	72	54	50	19
	CPU time (sec)	0.0258	0.0183	0.0170	0.0034
$m = 500$	No of Iter.	148	120	68	59
	CPU time (sec)	0.0609	0.0469	0.0193	0.0172



Clearly A is monotone (hence, pseudo-monotone). We set $q = 0 \in \mathbb{R}^m$ and choose the entries of N and S to be randomly generated in $(-2, 2)$ while that of D are randomly generated in $(0, 1)$. It is easy to see that $VI(C, A) = \{0 \in \mathbb{R}^m\}$. For the sake of simplicity, we define $f(x) = \frac{1}{2} \|x\|_2^2$, we take $\sigma = 0.57$, $\gamma = 0.3$, $l = 5$, $\mu = 0.02$ and test our algorithm for $m = 50, 100, 200$ and 500 . We compare the performance of our algorithms with the algorithms of Nomirovski [35] and Gibali [18], taking $\varepsilon = 10^{-4}$. The numerical result can be found in Table 1 and Fig. 1.

Next, we give an example in infinite-dimensional space to support the strong convergence of our algorithm. We take $f(x) = \frac{1}{2} \|x\|^2$.

Example 4.2 Let $E = L^2([0, 1])$ with norm $\|x\| = (\int_0^1 |x(t)|^2 dt)^{\frac{1}{2}}$ and inner product $\langle x, y \rangle = \int_0^1 x(t)y(t) dt$, $x, y \in E$. Let C be the unit ball in E defined by $C = \{x \in E : \|x\| \leq 1\}$. Let $B : C \rightarrow \mathbb{R}$ be an operator defined by $B(u) = \frac{1}{1+\|u\|^2}$ and $F : L^2([0, 1]) \rightarrow L^2([0, 1])$ be the

Volterra integral operator defined by $F(u)(t) = \int_0^t u(s) ds$ for all $u \in L^2([0, 1])$ and $t \in [0, 1]$. F is bounded, linear and monotone (cf. Exercise 20.12 in [4]). Now define $A : C \rightarrow L^2([0, 1])$ by $A(u)(t) = (B(u)F(u))(t)$. Suppose $\langle Au, v - u \rangle \geq 0$ for all $u, v \in C$, then $\langle Fu, v - u \rangle \geq 0$. Hence

$$\begin{aligned} \langle Av, v - u \rangle &= \langle BvFv, v - u \rangle \\ &= Bv\langle Fv, v - u \rangle \\ &\geq Bv(\langle Fv, v - u \rangle - \langle Fu, v - u \rangle) \\ &= Bv\langle Fv - Fu, v - u \rangle \geq 0. \end{aligned} \tag{4.1}$$

Thus, A is pseudo-monotone. To see that A is not monotone, choose $v = 1$ and $u = 2$, then

$$\langle Av - Au, v - u \rangle = -\frac{1}{10} < 0.$$

Now consider the VI in which the underlying operator A is as defined above. Clearly, the unique solution of the VI is $0 \in L^2([0, 1])$. Choosing $\sigma = 0.15$, $\gamma = 0.7$, $l = 7$, $\mu = 0.34$ and $\varepsilon < 10^{-4}$. We plot the graph of $\|x_{n+1} - x_n\|^2$ against number of iteration for Algorithm 3.6 and Algorithm 1.2 of [18]. We choose the same x_1 for both algorithms and take u in Algorithm 3.6 to be y_1 in Algorithm 1.2 as follows:

- Case I: $x_1 = \sin(t)$, $u = t^2 + t/5$,
- Case II: $x_1 = \exp(2t)/40$, $u = \exp(3t)/7$.

The numerical result is reported in Fig. 2 and Table 2.

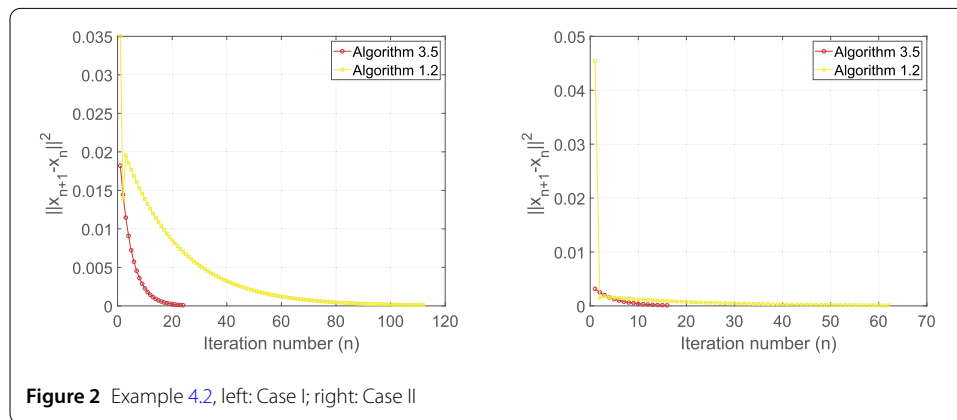


Table 2 Computation result for Example 4.2

		Algorithm 3.6	Algorithm 1.2
Case I	No of Iter.	24	112
	CPU time (sec)	1.3052	69.5319
Case II	No of Iter.	16	62
	CPU time (sec)	1.8171	88.4272

Table 3 Computation result for Example 4.3

		Algorithm 3.1	Algorithm 3.6
N = 50	No of Iter.	17	9
	CPU time (sec)	0.0284	0.0122
N = 100	No of Iter.	17	9
	CPU time (sec)	0.0832	0.0428
N = 200	No of Iter.	18	9
	CPU time (sec)	0.3606	0.1757
N = 500	No of Iter.	18	9
	CPU time (sec)	3.1267	1.4778

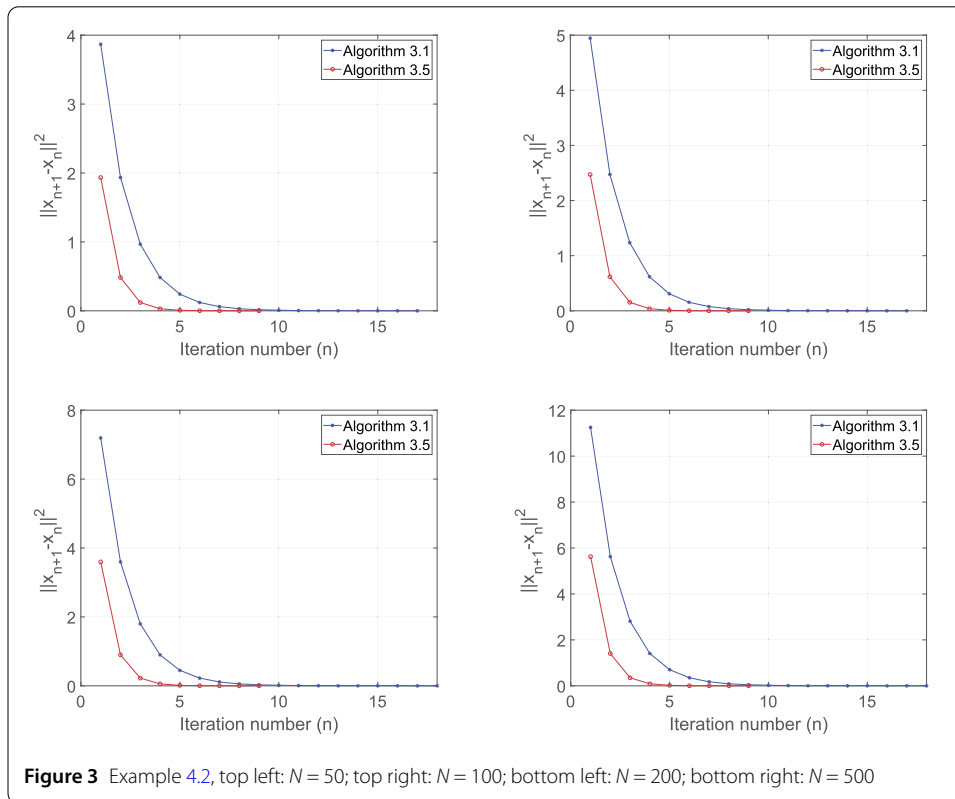


Figure 3 Example 4.2, top left: $N = 50$; top right: $N = 100$; bottom left: $N = 200$; bottom right: $N = 500$

Example 4.3 Finally, we consider $E = \mathbb{R}^N$ with $C = \{x \in \mathbb{R}^N : x_i \geq 0, \sum_{i=1}^N x_i = 1\}$ and $f(x) = -\sum_{i=1}^N x_i \log(x_i), x \in \mathbb{R}_{++}^N$. The projection onto C in this case is given by (see [5, 12])

$$\Pi_C(x) = \left[\frac{x_1 e_1}{\sum_{j=1}^N x_j e_j}, \frac{x_2 e_2}{\sum_{j=1}^N x_j e_j}, \dots, \frac{x_m e_m}{\sum_{j=1}^N x_j e_j} \right],$$

where $e = (e_1, e_2, \dots, e_N)$ is the standard basis of \mathbb{R}^N . We define the operator A as

$$A(x) = \max\{x, 0\}.$$

It is not difficult to see that A is monotone (hence pseudo-monotone) and $VI(C, A) = \{0\}$. Taken $\sigma = 0.02, \gamma = 0.5, l = 5, \mu = 0.9$ and $\varepsilon < 10^{-5}$. We apply Algorithm 3.1 and 3.6 for solving the variational inequality with respect to the above operator A using different ran-

domly generated initial point x_1 for $N = 50$, $N = 100$, $N = 200$ and $N = 500$. The numerical results can be found in Table 3 and Fig. 3.

5 Conclusion

The aim of the research is to study new subgradient extragradient methods for solving variational inequalities using Bregman distance approach in real reflexive Banach spaces. One of the advantages of the new methods is the use of an Armijo-like line search technique which prevents finding a prior estimate of the Lipschitz constant of the pseudo-monotone operator involved in the variational inequalities. Weak and strong convergence theorems were proved under mild conditions and some numerical experiments were performed to show the computational advantages of the new methods. The results in this paper improve and extend many recent results in the literature.

Acknowledgements

The authors acknowledge, with gratitude, the Department of Mathematics and Applied Mathematics at the Sefako Makgatho Health Sciences University for making their facilities available for the research.

Funding

This first author is supported by the Postdoctoral research grant from the Sefako Makgatho Health Sciences University, South Africa.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that there is not competing interest on the paper.

Authors' contributions

All authors worked equally on the results, read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 March 2020 Accepted: 23 July 2020 Published online: 31 July 2020

References

1. Alakoya, T.O., Jolaoso, L.O., Mewomo, O.T.: Modified inertial subgradient extragradient method with self-adaptive stepsize for solving monotone variational inequality and fixed point problems. *Optimization* (2020). <https://doi.org/10.1080/02331934.2020.1723586>
2. Alber, Y.I.: Metric and generalized projection operators in Banach spaces: properties and applications. In: Kartsatos, A.G. (ed.) *Theory and Applications of Nonlinear Operator of Accretive and Monotone Type*, pp. 15–50. Marcel Dekker, New York (1996)
3. Antipin, A.S.: On a method for convex programs using a symmetrical modification of the Lagrange function. *Èkon. Mat. Metody* **12**, 1164–1173 (1976)
4. Bauschke, H.H., Combettes, P.L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, New York (2011)
5. Beck, A.: *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, Philadelphia (2017)
6. Bregman, L.M.: The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math. Math. Phys.* **7**, 200–217 (1967)
7. Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalities in Hilbert spaces. *J. Optim. Theory Appl.* **148**, 318–335 (2011)
8. Censor, Y., Gibali, A., Reich, S.: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optim. Methods Softw.* **26**, 827–845 (2011)
9. Censor, Y., Gibali, A., Reich, S.: Extensions of Korpelevich's extragradient method for variational inequality problems in Euclidean space. *Optimization* **61**, 1119–1132 (2012)
10. Censor, Y., Lent, A.: An iterative row-action method for interval convex programming. *J. Optim. Theory Appl.* **34**, 321–353 (1981)
11. Denisov, S.V., Semenov, V.V., Chabak, L.M.: Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators. *Cybern. Syst. Anal.* **51**, 757–765 (2015)
12. Denisov, S.V., Semenov, V.V., Stetsynk, P.I.: Bregman extragradient method with monotone rule of step adjustment. *Cybern. Syst. Anal.* **55**(3), 377–383 (2019)
13. Dong, Q.L., Jiang, D., Gibali, A.: A modified subgradient extragradient method for solving the variational inequality problem. *Numer. Algorithms* **79**, 927–940 (2018)

14. Dong, Q.L., Lu, Y.Y., Yang, J.: The extragradient algorithm with inertial effects for solving the variational inequality. *Optimization* **65**(12), 2217–2226 (2016)
15. Facchinei, F., Pang, J.S.: *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. II. Springer Series in Operations Research. Springer, New York (2003)
16. Fang, C., Chen, S.: Some extragradient algorithms for variational inequalities. In: *Advances in Variational and Hemivariational Inequalities*. *Advances in Mechanics and Mathematics*, vol. 33, pp. 145–171. Springer, Cham (2015)
17. Fichera, G.: Sul problema elastostatico di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat.* **34**, 138–142 (1963)
18. Gibali, A.: A new Bregman projection method for solving variational inequalities in Hilbert spaces. *Pure Appl. Funct. Anal.* **3**(3), 403–415 (2018)
19. Glowinski, R., Lions, J.L., Trémolières, R.: *Numerical Analysis of Variational Inequalities*. North-Holland, Amsterdam (1981)
20. Goldstein, A.A.: Convex programming in Hilbert space. *Bull. Am. Math. Soc.* **70**, 709–710 (1964)
21. Gossez, J.P., Lami Dozo, E.: Some geometric properties related to the fixed point theory for nonexpansive mappings. *Pac. J. Math.* **40**, 565–573 (1972)
22. Halpern, B.: Fixed points of nonexpanding maps. *Proc. Am. Math. Soc.* **73**, 957–961 (1967)
23. Huang, Y.Y., Jeng, J.C., Kuo, T.Y., Hong, C.C.: Fixed point and weak convergence theorems for point-dependent λ -hybrid mappings in Banach spaces. *Fixed Point Theory Appl.* **2011**, Article ID 105 (2011)
24. Jolaoso, L.O., Alakoya, T.O., Taiwo, A., Mewomo, O.T.: An inertial extragradient method via viscosity approximation approach for solving equilibrium problem in Hilbert spaces. *Optimization* (2020). <https://doi.org/10.1080/02331934.2020.1716752>
25. Jolaoso, L.O., Ogbuisi, F.U., Mewomo, O.T.: An iterative method for solving minimization, variational inequality and fixed point problems in reflexive Banach spaces. *Adv. Pure Appl. Math.* **9**(3), 167–184 (2017)
26. Jolaoso, L.O., Taiwo, A., Alakoya, T.O., Mewomo, O.T.: A self adaptive inertial subgradient extragradient algorithm for variational inequality and common fixed point of multivalued mappings in Hilbert spaces. *Demonstr. Math.* **52**, 183–203 (2019)
27. Jolaoso, L.O., Taiwo, A., Alakoya, T.O., Mewomo, O.T.: A unified algorithm for solving variational inequality and fixed point problems with application to the split equality problem. *Comput. Appl. Math.* (2019). <https://doi.org/10.1007/s40314-019-1014-2>
28. Jolaoso, L.O., Taiwo, A., Alakoya, T.O., Mewomo, O.T.: A strong convergence theorem for solving pseudo-monotone variational inequalities using projection methods in a reflexive Banach space. *J. Optim. Theory Appl.* (2020). <https://doi.org/10.1007/s10957-020-01672-3>
29. Kinderlehrer, D., Stampacchia, G.: *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York (1980)
30. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. *Èkon. Mat. Metody* **12**, 747–756 (1976) (in Russian)
31. Levitin, E.S., Polyak, B.T.: Constrained minimization problems. *USSR Comput. Math. Math. Phys.* **6**, 1–50 (1966)
32. Lin, L.J., Yang, M.F., Ansari, Q.H., Kassay, G.: Existence results for Stampacchia and Minty type implicit variational inequalities with multivalued maps. *Nonlinear Anal., Theory Methods Appl.* **61**, 1–19 (2005)
33. Mashreghi, J., Nasri, M.: Forcing strong convergence of Korpelevich's method in Banach spaces with its applications in game theory. *Nonlinear Anal.* **72**, 2086–2099 (2010)
34. Nemirovski, A.: Prox-method with rate of convergence $O(1/t)$ for variational inequalities with Lipschitz continuous monotone operators and smooth convex–concave saddle point problems. *SIAM J. Optim.* **15**, 229–251 (2004)
35. Nomirowskii, D.A., Rublyov, B.V., Semenov, V.V.: Convergence of two-step method with Bregman divergence for solving variational inequalities. *Cybern. Syst. Anal.* **55**(3), 359–368 (2019)
36. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73**, 591–597 (1967)
37. Reem, D., Reich, S., De Pierro, A.: Re-examination of Bregman functions and new properties of their divergences. *Optimization* **68**, 279–348 (2019)
38. Reich, S., Sabach, S.: A strong convergence theorem for proximal type-algorithm in reflexive Banach spaces. *J. Nonlinear Convex Anal.* **10**, 471–485 (2009)
39. Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequality problems. *SIAM J. Control Optim.* **37**, 765–776 (1999)
40. Stampacchia, G.: Formes bilinéaires coercitives sur les ensembles convexes. *C. R. Math. Acad. Sci. Paris* **258**, 4413–4416 (1964)
41. Taiwo, A., Jolaoso, L.O., Mewomo, O.T.: A modified Halpern algorithm for approximating a common solution of split equality convex minimization problem and fixed point problem in uniformly convex Banach spaces. *Comput. Appl. Math.* **38**(2), Article ID 77 (2019)
42. Thong, D.V., Hieu, D.V.: New extragradient methods for solving variational inequality problems and fixed point problems. *J. Fixed Point Theory Appl.* **20**, Article ID 129 (2018). <https://doi.org/10.1007/s11784-018-0610-x>
43. Thong, D.V., Shehu, Y., Iyiola, O.S.: A new iterative method for solving pseudomonotone variational inequalities with non-Lipschitz operators. *Comput. Appl. Math.* **39**, Article ID 108 (2020)
44. Thong, D.V., Vinh, N.T., Cho, Y.J.: A strong convergence theorem for Tseng's extragradient method for solving variational inequality problems. *Optim. Lett.* (2019). <https://doi.org/10.1007/s115900-019-01391-3>
45. Thong, D.V., Vuong, P.T.: Modified Tseng's extragradient methods for solving pseudo-monotone variational inequalities. *Optimization* **68**(11), 2207–2226 (2019)
46. Xu, H.K.: Iterative algorithms for nonlinear operators. *J. Lond. Math. Soc.* **66**, 240–256 (2002)