CARPATHIAN J. MATH.
Volume **41** (2025), No. 4,
Pages 913-936

Online version at https://www.carpathian.cunbm.utcluj.ro/
ISSN 1584-2851 (Print edition); ISSN 1843-4401 (Electronic)
DOI: https://doi.org/10.37193/CJM.2025.04.04

Convergence analysis of Bregman projection methods with a new extrapolation technique for solving variational inequalities in reflexive Banach spaces

ZHONG-BAO WANG[1,2], ZHONG-CHENG ZHANG[1], ZHANG-YOU CHEN[1,2],
AND PONGSAKORN SUNTHRAYUTH[3,4]

ABSTRACT. In this paper, we firstly propose a Bregman projection algorithm with a new extrapolation technique for solving quasi-monotone variational inequalities in reflexive Banach spaces. We establish the weak convergence and non-asymptotic $O\left(\frac{1}{\sqrt{n}}\right)$ convergence rate of the algorithm under appropriate and mild assumptions. Secondly, we introduce the second algorithm and demonstrate its linear convergence under stronger conditions. Our numerical experiments show that our methods outperform existing algorithms in the literature.

1. INTRODUCTION

Throughout the paper, let E be a reflexive Banach space with the dual space E^*, the dual pair between E^* and E be denoted by $\langle \cdot, \cdot \rangle$. Suppose C is a nonempty, closed and convex subset of E, and $A : C \to E^*$ is a nonlinear mapping. The variational inequalities (VI) have the following form: find a point $x^* \in C$ such that

$$\langle Ax^*, y - x^* \rangle \geq 0, \ \forall \, y \in C. \tag{1.1}$$

We denote by $VI(C, A)$ the solution set of VI (1.1). Meanwhile, the solution set of the dual variational inequalities is represented by

$$DVI(C, A) = \{x^* : \langle Ay, y - x^* \rangle \geq 0, \ \forall \, y \in C\}, \tag{1.2}$$

which is a closed and convex subset in C, see [8].

Since variational inequalities were introduced firstly by Fichera ([9], [10]) in 1963, they have become a useful tool in studying various linear and nonlinear problems from elasticity, economics, transportation, optimization, network analysis, control theory, and engineering sciences, see, for example, [8, 22, 23, 45]. Due to the importance of variational inequalities in modern scientific research, numerous numerical methods for solving variational inequalities have been presented.

In particular, the extragradient technique (EGM) for solving variational inequalities, which was put forth by Korpelevich [24] and is among the most well-liked and extensively

Received: 10.09.2024. In revised form: 27.01.2025. Accepted: 26.05.2025
2020 *Mathematics Subject Classification.* 47H05, 47H07, 47H10, 54H25.
Key words and phrases. *variational inequalities, Bregman projection, extrapolation accelerations, convergence analysis, Banach spaces.*
Corresponding author: Pongsakorn Sunthrayuth; pongsakorn_su@rmutt.ac.th

applied projection-type methods. It can be described simply as follows:

$$\begin{cases} v_n = P_C(u_n - \lambda A u_n), \\ u_{n+1} = P_C(u_n - \lambda A v_n). \end{cases}$$

where $\lambda > 0$ is a suitable step size and $P_C(x) = \arg\min \left\{ \frac{1}{2}\|y - x\|^2 : y \in C \right\}$.

However, EGM needs to compute two projections onto C in each iterative step. This may have a bad influence on the convergence speed of the algorithm in the case of the set C is very complicated. To overcome this shortcoming, Tseng [42] introduced the following method

$$\begin{cases} v_n = P_C\left(u_n - \lambda A u_n\right), \\ u_{n+1} = v_n - \lambda\left(A v_n - A u_n\right), \end{cases}$$

which is usually called the Tseng extragradient method (TEGM). And TEGM only requires to compute the metric projection once in each iteration, which significantly reduces the complexity of computations and improves its overall computational efficiency. Over time, both EGM and TEGM have undergone numerous improvements and enhancements in Hilbert spaces and more general Banach spaces, see [1, 15, 16, 18, 19, 36, 38, 48, 46].

Additionally, Voung [43] proposed the relaxed inertial projection algorithm (RIPA) by a discrete version of the proposed dynamical system for solving variational inequalities in Hilbert spaces:

$$\begin{cases} w_n = v_n + \theta(v_n - v_{n-1}), \\ v_{n+1} = (1-\rho)\, w_n + \rho P_C\left(w_n - \lambda A w_n\right). \end{cases}$$

If A is strongly pseudo-monotone and Lipschitz continuous in a Hilbert space, then the linear convergence of RIPA is proved under suitable choices of parameters.

Under the Euclidean distance, the golden ratio algorithm for variational inequalities was introduced by Malitsky [26] and its name comes from the fact that this algorithm owns a new extrapolation that is $\frac{\varphi-1}{\varphi}x_k + \frac{1}{\varphi}w_{k-1}$ and $\varphi = \frac{1+\sqrt{5}}{2}$. Inspired by [26], recently, Oyewole and Reich [31] introduced the new extrapolation into the subgradient extragradient method for solving pseudomonotone variational inequalities and presented two algorithms and proved the weak convergence and linear convergence of the algorithms, respectively. Although the coefficient $\varphi \in (1, +\infty)$ and is no longer $\frac{1+\sqrt{5}}{2}$, they still call this extrapolation the golden ratio technique.

Many well-known distance concepts, such as the square of Euclidean distance, the Kullback-Leibler distance, and the Squared Mahalanobis distance, are examples of Bregman distances generated by various types of functions. Replacing the Euclidean distance with a more general Bregman distance is a useful way to potentially improve extragradient algorithms for variational inequalities in Hilbert spaces or Banach spaces. For instance, Wang et al. [46] introduced three new Bregman projection methods with non-monotone adaptive step sizes in real Hilbert spaces. Under suitable conditions, they proved the weak convergence and strong convergence of these methods. In reflexive Banach spaces, Jolaoso et al. [17] introduced a single Bregman projection method with adaptive step sizes and Izuchukwu et al. [16] proposed a one-step Bregman projection method with adaptive step sizes and proved some weak and strong convergence results for the proposed method under suitable conditions.Tam and Uteda [38] extended the golden ratio algorithms of [26] to the Bregman fixed step size version and Bregman-adaptive step size version, and established the linear convergence rate of the Bregman golden ratio algorithm with fixed step sizes. Under the Bregman distance framework, other extragradient-type algorithms for solving variational inequalities can be found in [13, 40, 41, 44]. It's worth noting that there is little literature for discussing the convergence rate of the extragradient type algorithm with adaptive step sizes under the Bregman distance.

Driven by the previously mentioned research, we firstly provide a new Bregman projection algorithm for solving quasi-monotone variational inequalities in Banach spaces. The algorithm incorporates the new extrapolation technique whose ideas are derived from the golden ratio technique of [31, 26] and non-monotone adaptive step sizes strategy, which results in a fast convergence rate without the need for prior information on the Lipschitz constant of the mapping A. The weak convergence and non-asymptotic $O\left(\frac{1}{\sqrt{n}}\right)$ convergence rate for the algorithm is proved under appropriate and mild assumptions. Furthermore, the second algorithm is showcased, and its linear convergence rate is determined. Some numerical experiments show that the proposed algorithms are more effective than some existing ones.

The structure of the paper is as follows: Section 2 introduces the necessary definitions and lemmas for this article. In Section 3, we present the main results of the proposed algorithm. Section 4 includes several numerical experiments to showcase the performance of our algorithms. Finally, Section 5 provides concluding remarks.

2. PRELIMINARIES

In this section, we present some definitions and preliminary results that is needed in our convergence analysis. In the sequel, we denote the weak convergence and strong convergence of a sequence $\{x_n\}$ to x by $x_n \rightharpoonup x$ and $x_n \to x$, respectively.

Definition 2.1. *The mapping $A : C \to E^*$ is said to be*

(i) monotone on C if

$$\langle Ay - Ax, y - x\rangle \geq 0,\ \forall x, y \in C;$$

(ii) pseudo-monotone on C if

$$\langle Ax, y - x\rangle \geq 0 \Longrightarrow \langle Ay, y - x\rangle \geq 0,\ \forall x, y \in C;$$

(iii) quasi-monotone on C if

$$\langle Ax, y - x\rangle > 0 \Longrightarrow \langle Ay, y - x\rangle \geq 0,\ \forall x, y \in C;$$

(iv) sequentially weakly continuous at x if $Ax_n \rightharpoonup Ax$ whenever $x_n \rightharpoonup x$;

(v) L-Lipschitz continuous on C if there exists some constant $L > 0$ satisfying

$$\|Ax - Ay\| \leq L\|x - y\|,\ \forall x, y \in C;$$

(vi) β-strongly pseudo-monotone on C if there exists some constant $\beta > 0$ satisfying

$$\langle Ax, y - x\rangle \geq 0 \Longrightarrow \langle Ay, y - x\rangle \geq \beta\|x - y\|^2,\ \forall x, y \in C.$$

Remark 2.1. *From the above definition, we can see that (i)$\Longrightarrow$ (ii)$\Longrightarrow$(iii) and (vi)$\Longrightarrow$ (ii). But the converse is generally false, see* [39].

The domain of f denotes by $\mathrm{dom}f := \{x \in E : f(x) < \infty\}$ and the interior of the domain of f is represented by $\mathrm{int}(\mathrm{dom}f)$.

Definition 2.2. *The function $f : E \to (-\infty, +\infty]$ is said to be:*

(1) proper if $\mathrm{dom}\, f \neq \emptyset$*;*

(2) lower semi-continuous if $\{x \in \mathrm{dom}\, f : f(x) \leq a\}$ *is closed set for* $\forall a \in \mathbb{R}$*;*

(3) convex if $f(tx + (1-t)y) \leq tf(x) + (1-t) f(y)\,,\ \forall x,\ y \in \mathrm{dom}f,\ t \in [0,1]$ *and it is said to be strictly convex if* $f(tx + (1-t)y) < tf(x) + (1-t) f(y)\,,\ \forall x, y \in \mathrm{dom}f$ *with* $x \neq y$ *and* $t \in (0,1)$*;*

(4) uniformly convex if there exists a nondecreasing and continuous function $g : [0,\infty) \to [0,\infty)$ *with* $g(0) = 0$ *such that* $f(tx + (1-t)y) \leq tf(x) + (1-t) f(y) - t(1-t)g(\|x-y\|)\,,\ \forall x,\ y \in \mathrm{dom}f$ *and* $t \in [0,1]$*;*

(5) bounded on bounded sets if $f(U)$ is bounded for each bounded subset U of E;

(6) uniformly smooth if there exists a nondecreasing and continuous function $h : [0,\infty) \to [0,\infty)$ *with* $h(0) = 0$ *such that* $f(tx + (1-t)y) \geq tf(x) + (1-t) f(y) - t(1-t)h(\|x-y\|)\,,\ \forall x, y \in \mathrm{dom}f$ *and* $t \in [0,1]$.

Assume that $f : E \to (-\infty, +\infty]$ is a proper and convex function. The Fenchel conjugate of f is defined by $f^*(x^*) = \sup_{x\in E} \{\langle x^*, x\rangle - f(x)\}\,,\ \forall x^* \in E^*$.

The directional derivative of f at $x \in \mathrm{int}\,(\mathrm{dom}f)$ in the direction $y \in E$ is defined by

$$f'(x,y) := \lim_{t\to 0} \frac{f(x+ty) - f(x)}{t}. \tag{2.1}$$

We say that f is Gâteaux differentiable at x if for all $y \in E$, the directional derivative of f at x exists and $f'(x,y) = \langle \nabla f(x), y\rangle$, where $\nabla f(x)$ is the value of the gradient of f at x. If f is Gâteaux differentiable at each $x \in \mathrm{int}\,(\mathrm{dom}f)$, then f is said to be Gâteaux differentiable.

We say that f is Fréchet differentiable at x if the limit (2.1) is attained uniformly for every $y \in E$ with $\|y\| = 1$. Furthermore, if the limit (2.1) is attained uniformly for each $x \in C \subset E$ and $y \in E$ with $\|y\| = 1$, then f is said to be uniformly Fréchet differentiable on C. It is clear that if f is a Fréchet differentiable function, then it is Gâteaux differentiable, see [32].

Definition 2.3 ([34, 35])**.** *Let $f : E \to (-\infty, +\infty]$ be a proper, lower semi-continuous and convex function. It is said to be Legendre if f such that*

(L1) $\mathrm{int}(\mathrm{dom}f) \neq \emptyset$, *$f$ is Gâteaux differentiable on* $\mathrm{int}(\mathrm{dom}f)$ *and* $\mathrm{dom}\nabla f = \mathrm{int}(\mathrm{dom}f)$*;*

(L2) $\mathrm{int}(\mathrm{dom}f^*) \neq \emptyset$, *$f^*$ is Gâteaux differentiable on* $\mathrm{int}(\mathrm{dom}f^*)$ *and* $\mathrm{dom}\nabla f^* = \mathrm{int}(\mathrm{dom}f^*)$.

If f is a Legendre function, then ∇f is a bijection from $\mathrm{int}(\mathrm{dom}f)$ *into* $\mathrm{int}(\mathrm{dom}f^*)$ *such that* $\nabla f^* = (\nabla f)^{-1}$*, see* [2].

Definition 2.4 ([3])**.** *If $f : E \to (-\infty, +\infty]$ is convex and Gâteaux differentiable, then D_f :* $\mathrm{dom}f \times \mathrm{int}(\mathrm{dom}f) \to [0,\infty)$ *is referred to the Bregman distance in terms of f, where*

$$D_f(x,y) = f(x) - f(y) - \langle \nabla f(y), x - y\rangle\,, \forall x \in \mathrm{dom}f,\ y \in \mathrm{int}\,(\mathrm{dom}f)\,.$$

The geometric of Bregman distance associated with f can be found in [19, 36].
The function D_f owns the following properties:

(a) $D_f(x,y) - D_f(z,y) = f(x) - f(z) - \langle \nabla f(y), x - z\rangle, \forall x, z \in \mathrm{dom} f,\ y \in \mathrm{int}(\mathrm{dom} f)$;
(b) *three-point identity, that is,*

$$D_f(x,y) - D_f(x,z) - D_f(z,y) = \langle \nabla f(z) - \nabla f(y), x - z\rangle, \forall x, z \in \mathrm{dom} f,\ y \in \mathrm{int}(\mathrm{dom} f);$$

(c) *let $x \in \mathrm{dom} f,\ y,\ u,\ v \in \mathrm{int}(\mathrm{dom} f)$ and $a \in \mathbb{R}$, if $\nabla f(y) = a\nabla f(u) + (1-a)\nabla f(v)$, we have*

$$D_f(x,y) = a\left[D_f(x,u) - D_f(y,u)\right] + (1-a)\left[D_f(x,v) - D_f(y,v)\right].$$

The proof of the Property (c) can be found in [38].

The Gâteaux differentiable function f is called to be strongly convex with a fixed constant $\kappa > 0$, if

$$\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq \kappa \|x - y\|^2, \forall x, y \in \mathrm{dom} f.$$

This implies

$$f(x) \geq f(y) + \langle \nabla f(y), x - y\rangle + \frac{\kappa}{2}\|x - y\|^2, \forall x, y \in \mathrm{dom} f,$$

and then

$$D_f(x,y) \geq \frac{\kappa}{2}\|x - y\|^2, \forall x, y \in \mathrm{dom} f.$$

In addition, let Gâteaux differentiable function $f : \mathrm{dom} f \to (-\infty, \infty]$ be strongly convex, and C be a nonempty, closed and convex subset of dom$f$. We refer to the unique vector $P_C^f(u) \in C$ as Bregman projection of f of $x \in \mathrm{int}(\mathrm{dom} f)$ onto C if

$$D_f\left(P_C^f(x), x\right) = \inf\left\{D_f(v,x), v \in C\right\}.$$

Moreover, a function $f : E \to (-\infty, +\infty]$ is said to be strongly coercive if

$$\lim_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = +\infty.$$

Definition 2.5 ([34, 35]). *The function $v_f : \mathrm{int}(\mathrm{dom} f) \times [0, +\infty) \to [0, +\infty)$ is defined by*

$$v_f(x,t) := \inf\left\{D_f(y,x) : y \in \mathrm{int}(\mathrm{dom} f), \|y - x\| = t\right\},$$

where f is a convex and Gâteaux differentiable function. We say that the function f is totally convex at a point $x \in \mathrm{int}(\mathrm{dom} f)$ if $v_f(x,t) > 0$, $\forall\, t > 0$. Moreover, we say that the function f is totally convex on bounded subsets of E if $v_f(X,t) > 0$ for every bounded subset X of E and for any $t > 0$, where

$$v_f(X,t) = \inf\left\{v_f(x,t) : x \in X \cap \mathrm{int}(\mathrm{dom} f)\right\}.$$

It is evident that if f is strongly convex, then f is totally convex. In addition, it is well known that f is totally convex on bounded subsets if and only if f is uniformly convex on bounded subsets, see Theorem 2.10 of [4], and if f is uniformly convex, then f is totally convex, but the converge is generally false, see Section 1.3 of [5].

Lemma 2.1 ([20]). *If A is continuous, then $DVI(C,A) \subset VI(C,A)$ and if A is a pseudo-monotone mapping too, then $DVI(C,A) = VI(C,A)$.*

Lemma 2.2 ([33]). *Let K be a given nonempty bounded subset of E. Suppose that the bounded function $f : E \to \mathbb{R}$ is convex and uniformly Fréchet differentiable on K, then ∇f is uniformly continuous on K.*

Lemma 2.3 ([27]). *Let the function $f : E \to \mathbb{R}$ be Gâteaux differentiable and totally convex on E, if $\{x_n\} \subset E$ and $x_0 \in E$, $\{D_f(x_0, x_n)\}$ is bounded, then $\{x_n\}$ is also bounded.*

Lemma 2.4 ([7]). *Let $x \in E$ and $f : E \to \mathbb{R}$ be a totally convex and Gâteaux differentiable function. Then the following conclusions hold:*

(a) $p = P^f_C(x) \Leftrightarrow \langle \nabla f(x) - \nabla f(p), z - p\rangle \leq 0,\ \forall z \in C$;
(b) $D_f(u, P^f_C(x)) + D_f(P^f_C(x), x) \leq D_f(u, x),\ \forall u \in C$.

Lemma 2.5 ([29]). *Suppose that $f : E \to \mathbb{R}$ is a convex function and f is bounded on bounded subsets of E. Then the following three statements are equivalent:*

(a) *f is uniformly convex and strongly coercive on bounded subsets of E;*
(b) $\mathrm{dom} f^* = E^*$, *f^* is uniformly smooth and bounded on bounded subsets of E^*;*
(c) $\mathrm{dom} f^* = E^*$, *f^* is Fréchet differentiable, moreover, ∇f^* is uniformly norm-to-norm continuous on bounded subsets of E^*.*

Lemma 2.6 ([5]). *Let $f : E \to \mathbb{R}$ be strongly coercive function. Then the following conclusions hold:*

(a) *$\nabla f : E \to E^*$ is one-to-one, onto and norm-to-weak* continuous;*
(b) *the set $\{x \in E : D_f(x, y) \leq \gamma\}$ is bounded for each $y \in E$ and $\gamma > 0$;*
(c) $\mathrm{dom} f^* = E^*$, *f^* is Gâteaux differentiable and $\nabla f^* = (\nabla f)^{-1}$.*

Lemma 2.7 ([29]). *Suppose that $f : E \to \mathbb{R}$ is a Gâteaux differentiable function that is uniformly convex on bounded subsets of E. If $\{u_n\}$ and $\{v_n\}$ are two bounded sequences in E, then $\lim_{n\to\infty} D_f(v_n, u_n) = 0$ if and only if $\lim_{n\to\infty} \|v_n - u_n\| = 0$.*

Assume that $f : E \to \mathbb{R} \cup \{+\infty\}$ is a Legendre function. The bifunction $V_f : \mathrm{dom} f \times \mathrm{dom} f^* \to [0, +\infty)$ is defined by

$$V_f(y, y^*) = f(y) - \langle y, y^*\rangle + f^*(y^*), \forall y \in \mathrm{dom} f,\ y^* \in \mathrm{dom} f^*.$$

Thus

$$V_f(y, y^*) = D_f(y, \nabla f^*(y^*)), \forall y \in \mathrm{dom} f,\ y^* \in \mathrm{dom} f^*$$

and

$$V_f(y, y^*) + \langle \nabla f^*(y^*) - y, z^*\rangle \leq V_f(y, y^* + z^*), \forall y \in \mathrm{dom} f,\ y^*,\ z^* \in \mathrm{dom} f^*.$$

From [28], we know that V_f is convex for the second variable. So we have

$$D_f\left(y, \nabla f^*\left(\sum_{j=1}^N a_j \nabla f(y_j)\right)\right) \leq \sum_{j=1}^N a_j D_f(y, y_j), \tag{2.2}$$

where $y \in E$, $y_j \in E$ and $a_j \in (0,1)$ with $\sum_{j=1}^N a_j = 1$.

Lemma 2.8 ([11]). *Let a Gâteaux differentiable function $f : E \to (-\infty, \infty]$ be proper and strictly convex on $\operatorname{int}(\operatorname{dom} f)$ in E, and the sequence $\{x_n\} \subset \operatorname{dom} f$ such that $x_n \rightharpoonup x$ for some $x \in \operatorname{int}(\operatorname{dom} f)$. Thus*

$$\liminf_{n\to\infty} D_f(x, x_n) < \liminf_{n\to\infty} D_f(y, x_n), \ \forall y \in \operatorname{dom} f, \ y \neq x.$$

Lemma 2.9 ([6]). *Let $\{a_n\}$ and $\{b_n\}$ be two non-negative real sequences. If there exists $N \in \mathbb{N}$ satisfying*

$$a_{n+1} \leq a_n - b_n, \ \forall n \geq N,$$

then $\lim\limits_{n\to\infty} b_n = 0$ and $\lim\limits_{n\to\infty} a_n$ exists.

Lemma 2.10 ([30]). *Let $\{\beta_n\}$, $\{\theta_n\}$ and $\{\gamma_n\}$ be three non-negative real sequences satisfying*

$$\gamma_{n+1} \leq \beta_n \gamma_n + \theta_n, \ \forall n \in \mathbb{N}.$$

If $\{\beta_n\} \subset [1, +\infty)$, $\sum\limits_{n=1}^{\infty} (\beta_n - 1) < +\infty$, and $\sum\limits_{n=1}^{\infty} \theta_n < +\infty$, then $\lim\limits_{n\to\infty} \gamma_n$ exists.

3. Algorithm and convergence analysis

In this section, we firstly present a novel Bregman projection algorithm with the new extrapolation technique for solving quasi-monotone variational inequalities in Banach spaces, and prove the weak convergence, nonasymptotic $O\left(\frac{1}{\sqrt{n}}\right)$ convergence rate of the method, respectively. For convenience, we introduce the following conditions:

- (A_1) the function $f : E \to \mathbb{R}$ is bounded, κ- strongly convex, uniformly Fréchet differentiable and Legendre;
- (A_2) the function $f : E \to \mathbb{R}$ is strongly coercive;
- (A_3) the mapping $A : E \to E^*$ is quasi-monotone and Lipschitz continuous with $L > 0$;
- (A_4) $DVI(C, A) \neq \emptyset$;
- (A_5) the mapping $A : E \to E^*$ satisfies the following property:

$$\text{whenever } \{q_n\} \subset C, \ q_n \rightharpoonup x, \text{ one has } \|Ax\| \leq \liminf_{n\to\infty} \|Aq_n\|.$$

If $A : E \to E^*$ is sequentially weakly continuous, then the Condition (A_5) holds. But the converse is not true, for details, see Remark 2 in [46]. Therefore, the Condition (A_5) is weaker than the Condition (A2) of [12] and the Condition (C3) of [47], respectively. Now we introduce our algorithm.

Algorithm 1.

Initialization: *Take $w_0, \ x_1 \in C, \ \gamma_1 > 0, \ \sigma \in (0, \min\{1, \kappa\})$. Set $n := 1$. Choose real non-negative sequences $\{\beta_n\}, \ \{t_n\}, \ \{\theta_n\}$ and $\{\alpha_n\}$ such that the following conditions hold:*

(1) $\{\beta_n\} \subset [1, +\infty)$, $\sum\limits_{n=1}^{\infty} (\beta_n - 1) < +\infty$ *and* $\sum\limits_{n=1}^{\infty} \theta_n < +\infty$;

(2) $0 < \varphi_1 \leq t_n \leq 1$;

(3) $0 < a \leq \alpha_{n+1} \leq \alpha_n \leq b < 1$.

Iterative steps: *Having x_n and w_{n-1}, compute the next iterate x_{n+1} as follows:*

Step 1. Compute

$$w_n = \nabla f^* \left((1 - \alpha_n) \nabla f(x_n) + \alpha_n \nabla f(w_{n-1}) \right).$$

Step 2. Compute

$$y_n = P_C^f \left(\nabla f^* \left(\nabla f\left(w_n\right) - \gamma_n A w_n\right)\right).$$

Step 3. Compute

$$x_{n+1} = \nabla f^* \left(t_n \nabla f\left(y_n\right) - \left(t_n - 1\right) \nabla f\left(w_n\right) - t_n \gamma_n \left(A y_n - A w_n\right)\right),$$

where

$$\gamma_{n+1} = \begin{cases} \min\left\{\frac{\sigma\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \beta_n \gamma_n + \theta_n\right\}, & \text{if } Aw_n \neq Ay_n, \\ \beta_n \gamma_n + \theta_n, & \text{otherwise.} \end{cases} \tag{3.1}$$

Update n and go to iterative step.

Remark 3.2. *Assume that the mapping $A : E \to E^*$ is Lipschitz continuous with $L > 0$. Let $\{\gamma_n\}$ be the sequence generated by (3.1). Then we have*

$$\lim_{n\to\infty} \gamma_n = \gamma \geq \min\left\{\frac{\sigma}{L}, \gamma_1\right\}.$$

Proof. According to the definition of $\{\gamma_n\}$ and the choice of β_n and θ_n, using Lemma 2.10, we have $\lim_{n\to\infty} \gamma_n = \gamma$. Since A is Lipschitz continuous with $L > 0$, we have

$$\frac{\sigma\|w_n - y_n\|}{\|Aw_n - Ay_n\|} \geq \frac{\sigma\|w_n - y_n\|}{L\|w_n - y_n\|} = \frac{\sigma}{L}.$$

Moreover, thanks to $\beta_n \geq 1$ and $\theta_n \geq 0$, we obtain $\beta_n \gamma_n + \theta_n \geq \gamma_n$. Therefore, we have $\gamma_n \geq \min\left\{\frac{\sigma}{L}, \gamma_1\right\} > 0$, $\forall\, n \in \mathbb{N}$. □

Remark 3.3. *(i) Step 1 of Algorithm 1 is different from the inertial acceleration term in* [46]. *If $f(x) = \frac{1}{2}\|x\|^2$, $\alpha_n = \frac{2}{\sqrt{5}+1}$ and $E = \mathbb{R}^m$, then Step 1 of Algorithm 1 becomes (22) of the golden ratio algorithm, that is, (22) of Algorithm 1 in* [26]. *Specially, if $f(x) = \frac{1}{2}\|x\|^2$, $\phi \in (0, +\infty)$, $\alpha_n = \frac{1}{\phi}$ and E is a Hilbert space, then Step 1 of Algorithm 1 is called to be the golden ratio technique studied in* [31, 49].

(ii) The sequence of step sizes $\{\gamma_n\}$ may be non-monotone and only need a simple calculation of known information without any prior estimation of parameters, such as the Lipschitz constant of the underlying operator.

(iii) Condition (A_2) is used commonly for Bregman projection methods of variational inequalities in reflexive Banach spaces, see [12, 16, 47]. *By Lemma 2.6, we know that ∇f and ∇f^* are one-to-one, onto and norm-to-weak* continuous under Condition (A_2). So Condition (A_2) can assure the sequences $\{w_n\}$, $\{y_n\}$ and $\{x_n\}$ generated by Algorithm 1 are well defined.*

To discuss the convergence and convergence rate of Algorithm 1, we next state some lemmas.

Lemma 3.11. *Assume that Conditions $(A_1)-(A_4)$ are satisfied. Let $\{x_n\}$ be a sequence generated by Algorithm 1. Then the following inequality holds:*

$$D_f\left(p, x_{n+1}\right) \leq D_f\left(p, w_n\right) - \sigma_n, \ \forall n \in \mathbb{N},$$

where $p \in DVI\left(C, A\right)$, $\sigma_n := t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f\left(y_n, w_n\right) + t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f\left(x_{n+1}, y_n\right).$

Proof. Let $p \in DVI(C, A)$. By the definition of x_{n+1}, we have

(3.2)
$$\begin{aligned} D_f(p, x_{n+1}) &= f(p) - f(x_{n+1}) - \langle t_n \nabla f(y_n) - (t_n - 1)\nabla f(w_n), p - x_{n+1}\rangle \\ &+ \gamma_n t_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle \\ &= t_n (f(p) - f(x_{n+1}) - \langle \nabla f(y_n), p - x_{n+1}\rangle) \\ &- (t_n - 1)(f(p) - f(x_{n+1}) - \langle \nabla f(w_n), p - x_{n+1}\rangle) \\ &+ \gamma_n t_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle \\ &= t_n (D_f(p, y_n) - D_f(x_{n+1}, y_n)) - (t_n - 1)(D_f(p, w_n) - D_f(x_{n+1}, w_n)) \\ &+ \gamma_n t_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle. \end{aligned}$$

Thanks to the three-point identity, we have

$$D_f(p, y_n) = D_f(p, w_n) - D_f(y_n, w_n) + \langle \nabla f(w_n) - \nabla f(y_n), p - y_n\rangle. \tag{3.3}$$

Substituting (3.3) into (3.2), we have

(3.4)
$$\begin{aligned} D_f(p, x_{n+1}) &= D_f(p, w_n) - t_n (D_f(y_n, w_n) + D_f(x_{n+1}, y_n)) - (1 - t_n) D_f(x_{n+1}, w_n) \\ &+ t_n \langle \nabla f(w_n) - \nabla f(y_n), p - y_n\rangle + \gamma_n t_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle. \end{aligned}$$

By the definition of y_n, $y_n \in C$, $p \in DVI(C, A)$ and Lemma 2.4, we have

$$\langle \nabla f(w_n) - \gamma_n Aw_n - \nabla f(y_n), p - y_n\rangle \leq 0, \tag{3.5}$$

and so

$$\langle \nabla f(w_n) - \nabla f(y_n), p - y_n\rangle \leq \langle \gamma_n Aw_n, p - y_n\rangle. \tag{3.6}$$

We can conclude that

$$\begin{aligned} &\langle \nabla f(w_n) - \nabla f(y_n), p - y_n\rangle + \gamma_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle \\ &\leq \gamma_n \langle Aw_n, p - y_n\rangle + \gamma_n \langle Ay_n - Aw_n, p - x_{n+1}\rangle \\ &= \gamma_n \langle Aw_n - Ay_n, x_{n+1} - y_n\rangle - \gamma_n \langle Ay_n, y_n - p\rangle. \end{aligned} \tag{3.7}$$

Combining (3.4), (3.7) and $t_n \in (0, 1]$, we obtain

$$\begin{aligned} D_f(p, x_{n+1}) &\leq D_f(p, w_n) - t_n (D_f(y_n, w_n) + D_f(x_{n+1}, y_n)) \\ &+ t_n \gamma_n \langle Aw_n - Ay_n, x_{n+1} - y_n\rangle - t_n \gamma_n \langle Ay_n, y_n - p\rangle. \end{aligned} \tag{3.8}$$

Thanks to $p \in DVI(C, A)$, we can get $\langle Ay_n, y_n - p\rangle \geq 0$. This and Remark 3.2 implies that

(3.9)
$$D_f(p, x_{n+1}) \leq D_f(p, w_n) - t_n (D_f(y_n, w_n) + D_f(x_{n+1}, y_n)) + t_n \gamma_n \langle Aw_n - Ay_n, x_{n+1} - y_n\rangle.$$

Since f is κ-strongly convex and the definition of γ_n, we have

$$
\begin{aligned}
D_f(p, x_{n+1}) &\leq D_f(p, w_n) - t_n\left(D_f(y_n, w_n) + D_f(x_{n+1}, y_n)\right) \\
&\quad + t_n\gamma_n \|Ay_n - Aw_n\| \|x_{n+1} - y_n\| \\
&\leq D_f(p, w_n) - t_n\left(D_f(y_n, w_n) + D_f(x_{n+1}, y_n)\right) \\
&\quad + \frac{\sigma t_n \gamma_n}{\gamma_{n+1}} \|y_n - w_n\| \|x_{n+1} - y_n\| \\
&\leq D_f(p, w_n) - t_n\left(D_f(y_n, w_n) + D_f(x_{n+1}, y_n)\right) \\
&\quad + \frac{\sigma t_n \gamma_n}{2\gamma_{n+1}} \|y_n - w_n\|^2 + \frac{\sigma t_n \gamma_n}{2\gamma_{n+1}} \|x_{n+1} - y_n\|^2 \\
&\leq D_f(p, w_n) - \sigma_n,
\end{aligned}
\tag{3.10}
$$

where $\sigma_n = t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f(y_n, w_n) + t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f(x_{n+1}, y_n)$.
Then the conclusion of Lemma 3.11 holds. This completes the proof. □

Lemma 3.12. *Suppose that Conditions $(A_1) - (A_4)$ hold. Then the sequence $\{x_n\}$ generated by Algorithm 1 is bounded and $\lim_{n\to\infty} D_f(p, x_n)$ exist for each $p \in DVI(C, A)$.*

Proof. Let $p \in DVI(C, A)$. By Remark 3.2 and $\sigma \in (0, \kappa)$, we have $\lim_{n\to\infty}\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) = 1 - \frac{\sigma}{\kappa} > 0$. As a consequence, there exists $n_1 \in \mathbb{N}$ such that $1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} > 0,\ \forall n \geq n_1$. This implies that

$$
\sigma_n := t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f(y_n, w_n) + t_n\Big(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\Big) D_f(x_{n+1}, y_n) \geq 0,\ \forall n \geq n_1.
$$

By the definition of w_n and f is Legendre, we have

$$
\nabla f(w_n) = (1 - \alpha_n)\nabla f(x_n) + \alpha_n \nabla f(w_{n-1}). \tag{3.11}
$$

Thus

$$
\nabla f(x_n) = \frac{1}{1 - \alpha_n}\nabla f(w_n) - \frac{\alpha_n}{1 - \alpha_n}\nabla f(w_{n-1}). \tag{3.12}
$$

It follows from (3.12) and the Property (c) of the Bregman distance that

$$
D_f(p, x_n) = \frac{1}{1 - \alpha_n}\Big(D_f(p, w_n) - D_f(x_n, w_n)\Big) - \frac{\alpha_n}{1 - \alpha_n}\Big(D_f(p, w_{n-1}) - D_f(x_n, w_{n-1})\Big). \tag{3.13}
$$

Combining (3.13) with (3.10), we can get

$$
\begin{aligned}
D_f(p, x_{n+1}) - D_f(p, x_n) &\leq -\frac{\alpha_n}{1 - \alpha_n} D_f(p, w_n) + \frac{\alpha_n}{1 - \alpha_n} D_f(p, w_{n-1}) \\
&\quad + \frac{1}{1 - \alpha_n} D_f(x_n, w_n) - \frac{\alpha_n}{1 - \alpha_n} D_f(x_n, w_{n-1}) - \sigma_n.
\end{aligned}
\tag{3.14}
$$

It follows from (3.14) and $0 < \alpha_{n+1} \leq \alpha_n < 1$ that

$$
\begin{aligned}
D_f(p, x_{n+1}) + \frac{\alpha_{n+1}}{1 - \alpha_{n+1}} D_f(p, w_n) &\leq D_f(p, x_{n+1}) + \frac{\alpha_n}{1 - \alpha_n} D_f(p, w_n) \\
&\leq D_f(p, x_n) + \frac{\alpha_n}{1 - \alpha_n} D_f(p, w_{n-1}) \\
&\quad + \frac{1}{1 - \alpha_n} D_f(x_n, w_n) - \frac{\alpha_n}{1 - \alpha_n} D_f(x_n, w_{n-1}) - \sigma_n.
\end{aligned}
\tag{3.15}
$$

Setting

$$a_n := D_f(p, x_n) + \frac{\alpha_n}{1-\alpha_n} D_f(p, w_{n-1})$$

and

$$b_n := -\frac{1}{1-\alpha_n} D_f(x_n, w_n) + \frac{\alpha_n}{1-\alpha_n} D_f(x_n, w_{n-1}) + \sigma_n.$$

We infer that

$$a_{n+1} \leq a_n - b_n.$$

It follows from (3.11) and the Property (c) of the Bregman distance that

(3.16)

$$\begin{aligned} D_f(x_n, w_n) &= (1-\alpha_n)\left(D_f(x_n, x_n) - D_f(w_n, x_n)\right) + \alpha_n\left(D_f(x_n, w_{n-1}) - D_f(w_n, w_{n-1})\right) \\ &= -(1-\alpha_n) D_f(w_n, x_n) + \alpha_n\left(D_f(x_n, w_{n-1}) - D_f(w_n, w_{n-1})\right) \\ &\leq \alpha_n D_f(x_n, w_{n-1}). \end{aligned}$$

This implies that

$$b_n = D_f(w_n, x_n) + \frac{\alpha_n}{1-\alpha_n} D_f(w_n, w_{n-1}) + \sigma_n. \tag{3.17}$$

It is easy that $a_n \geq 0$ and $b_n \geq 0$, $\forall n \geq n_1$. Using Lemma 2.9, we can get $\lim\limits_{n\to\infty} b_n = 0$ and $\lim\limits_{n\to\infty} a_n$ exists. Thus, it follows from (3.17), $0 < a \leq \alpha_n \leq b < 1$, $t_n \geq \varphi_1 > 0$, the definition of σ_n and $\lim\limits_{n\to\infty} b_n = 0$ that

$$\lim_{n\to\infty} D_f(w_n, x_n) = 0 = \lim_{n\to\infty} D_f(w_n, w_{n-1}) \tag{3.18}$$

and

$$\lim_{n\to\infty} D_f(y_n, w_n) = 0 = \lim_{n\to\infty} D_f(x_{n+1}, y_n). \tag{3.19}$$

Since f is κ-strongly convex, we can also obtain

$$\lim_{n\to\infty} \|x_n - w_n\| = 0 = \lim_{n\to\infty} \|w_n - w_{n-1}\| \tag{3.20}$$

and

$$\lim_{n\to\infty} \|y_n - w_n\| = 0 = \lim_{n\to\infty} \|x_{n+1} - y_n\|. \tag{3.21}$$

Owing to (3.13) and (3.16), we have

$$\begin{aligned} a_n &= \frac{1}{1-\alpha_n}\Big(D_f(p, w_n) - D_f(x_n, w_n)\Big) + \frac{\alpha_n}{1-\alpha_n} D_f(x_n, w_{n-1}) \\ &= \frac{1}{1-\alpha_n} D_f(p, w_n) + D_f(w_n, x_n) + \frac{\alpha_n}{1-\alpha_n} D_f(w_n, w_{n-1}). \end{aligned} \tag{3.22}$$

From (3.18), $0 < a \leq \alpha_n \leq b < 1$ and $\lim\limits_{n\to\infty} a_n$ exists, it is easy to see that $\lim\limits_{n\to\infty} D_f(p, w_n)$ exists. Thus the definition of a_n implies $\lim\limits_{n\to\infty} D_f(p, x_n)$ exists. Lemma 2.3 ensures that $\{x_n\}$ and $\{w_n\}$ are bounded. We know that $\{y_n\}$ is bounded via (3.21). We thus complete the proof.

□

Lemma 3.13. *Assume that Conditions $(A_1) - (A_5)$ are satisfied. Let $\{w_n\}$ and $\{y_n\}$ be two sequences given by Algorithm 1. If $\lim\limits_{n\to\infty} \|w_n - y_n\| = 0$ and a subsequence of $\{w_n\}$ converges weakly to w^*, then $w^* \in DVI(C, A)$ or $Aw^* = 0$.*

Proof. The proof of Lemma 3.13 is the similar as the one of Lemma 3.4 in [46]. Thus we omit it. □

Theorem 3.1. *Let Conditions $(A_1)-(A_5)$ hold. Then the sequence $\{x_n\}$ generated by Algorithm 1 converges weakly to some point of $DVI(C,A) \subset VI(C,A)$.*

Proof. Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup z \in C$. In view of (3.20), we have $w_{n_k} \rightharpoonup z \in C$. It follows from Lemma 3.13 and (3.21) that $z \in DVI(C,A)$.

We now claim that $x_n \rightharpoonup z$. In fact it is sufficient to show that the weak cluster point of the sequence $\{x_n\}$ is unique. Assume that $\{x_{m_k}\}$ is another subsequence of $\{x_n\}$ such that $x_{m_k} \rightharpoonup z_1$ and $z_1 \neq z$. By utilizing the same arguments in getting $z \in DVI(C,A)$, we can obtain $z_1 \in DVI(C,A)$. On account of Lemma 3.12, we have $\lim_{n\to\infty} D_f(p,x_n)$ exists for each $p \in DVI(C,A)$. It follows from Lemma 2.8 that

$$\begin{aligned}\lim_{n\to\infty} D_f(z,x_n) &= \lim_{k\to\infty} D_f(z,x_{n_k}) = \liminf_{k\to\infty} D_f(z,x_{n_k}) \\ &< \liminf_{k\to\infty} D_f(z_1,x_{n_k}) = \lim_{n\to\infty} D_f(z_1,x_n).\end{aligned}$$

In addition, in a similar way, we can get $\lim_{n\to\infty} D_f(z_1,x_n) < \lim_{n\to\infty} D_f(z,x_n)$. This leads to a contradiction. Thus $z = z_1$ and the sequence $\{x_n\}$ converges weakly to z. The proof is finished. □

Remark 3.4. *If $\alpha_n = 0$, that is, $w_n = x_n$, we can get the same results as Lemma 3.12. In fact, using (3.10) and $w_n = x_n$, we have*

$$D_f(p,x_{n+1}) \leq D_f(p,x_n) - t_n\left(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right)(D_f(y_n,x_n) + D_f(x_{n+1},y_n)).$$

Set

$$a_n := D_f(p,x_n)$$

and

$$b_n := t_n\left(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right)(D_f(y_n,x_n) + D_f(x_{n+1},y_n)).$$

The next proof is similar to the proof of Lemma 3.12. We can yield the desired conclusion. Furthermore, resembling the proof of Theorem 3.1, it is easy to see that if $\alpha_n = 0$, then the result of Theorem 3.1 still holds.

If $\alpha_n = 0$ and $\beta_n = t_n = 1$, then Algorithm 1 becomes Algorithm 3.3 of [17]. *Theorem 3.1 only requires A is quasi-monotone and satisfies the Condition (A_5) other than being pseudo-monotone and weakly sequentially continuous as in Theorem 3.8 of* [17].

Next, we show that the nonasymptotic $O\left(\frac{1}{\sqrt{n}}\right)$ convergence rate with " $\min_{N\leq i\leq n}$ " of Algorithm 1.

Theorem 3.2. *Suppose that Conditions $(A_1)-(A_5)$ are satisfied. Let $\{x_n\}$ be a sequence generated by Algorithm 1. Then for each $p \in DVI(C,A)$, there exist some constant $\eta > 0$ and $N \in \mathbb{N}$ such that*

$$\min_{N\leq i\leq n} \|y_i - w_i\| \leq \left(\frac{2}{\eta\kappa} \cdot \frac{D_f(p,x_N) + \frac{\alpha_N}{1-\alpha_N} D_f(p,w_{N-1})}{n-N+1}\right)^{\frac{1}{2}}.$$

Proof. Set $\eta := \frac{\varphi_1}{2}\left(1-\frac{\sigma}{\kappa}\right)$. Obviously $\eta > 0$. By Remark 3.2 and $\sigma \in (0,\kappa)$, we have $\lim\limits_{n\to\infty}\left(1-\frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right) = 1-\frac{\sigma}{\kappa} > 0$. As a consequence, there exists $N \in \mathbb{N}$ such that $1-\frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} \geq \frac{1}{2}\left(1-\frac{\sigma}{\kappa}\right) > 0$, $\forall n \geq N$. It follows from $0 < a \leq \alpha_n \leq b < 1$, $a_{n+1} \leq a_n - b_n$, (3.17), $t_n \geq \varphi_1$ and the definition of σ_n that

$$\begin{aligned} a_{n+1} &\leq a_n - D_f(w_n,x_n) - \tfrac{\alpha_n}{1-\alpha_n} D_f(w_n,w_{n-1}) - \sigma_n \leq a_n - t_n\left(1-\tfrac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right) D_f(y_n,w_n) \\ &\leq a_n - \eta D_f(y_n,w_n),\ \forall n \geq N. \end{aligned}$$

So we have

$$\eta D_f(y_n,w_n) \leq a_n - a_{n+1},\ \forall n \geq N.$$

Thus

$$\eta \sum_{i=N}^{n} D_f(y_i,w_i) \leq \sum_{i=N}^{n}(a_i - a_{i+1}) = a_N - a_{n+1} \leq a_N = D_f(p,x_N) + \frac{\alpha_N}{1-\alpha_N} D_f(p,w_{N-1}).$$

This implies that

$$\min_{N\leq i\leq n} D_f(y_i,w_i) \leq \frac{1}{\eta}\frac{D_f(p,x_N) + \frac{\alpha_N}{1-\alpha_N} D_f(p,w_{N-1})}{n-N+1}.$$

Combining f is κ-strongly convex, we have

$$\min_{N\leq i\leq n} \frac{\kappa}{2}\|y_i - w_i\|^2 \leq \min_{N\leq i\leq n} D_f(y_i,w_i) \leq \frac{1}{\eta}\frac{D_f(p,x_N) + \frac{\alpha_N}{1-\alpha_N} D_f(p,w_{N-1})}{n-N+1},$$

which implies that

$$\min_{N\leq i\leq n} \|y_i - w_i\| \leq \left(\frac{2}{\eta\kappa}\cdot\frac{D_f(p,x_N) + \frac{\alpha_N}{1-\alpha_N} D_f(p,w_{N-1})}{n-N+1}\right)^{\frac{1}{2}}.$$

This proof is finished. □

Remark 3.5. *It is clear that $y_n = w_n$ can imply that $y_n \in VI(C,A)$. This and the fact*

$$\lim_{n\to 0}\|y_n - w_n\| = 0$$

imply that the estimation of error provided in Theorem 3.2 can be considered as a nonasymptotic convergence rate of Algorithm 1.

We regard that the sequence $\{x_n\} \subset E$ converges Q-linearly to some point $p \in E$ if there exists some $q \in (0,1)$ such that $\|x_{n+1} - p\| \leq q\|x_n - p\|$ for all n sufficiently large. We say that the sequence $\{x_n\} \subset E$ converges R-linearly if $\|x_n - p\| \leq b_n$ for all n sufficiently large, where $\{b_n\} \subset \mathbb{R}$ converges Q-linearly to zero.

Next, to obtain a linear convergence rate, by changing Step 3 of Algorithm 1, we get Algorithm 2 as follows:

Algorithm 2.
Initialization: *Take $w_0,\ x_1 \in C$, $\gamma_1 > 0$ and $\sigma \in \left(0,\frac{\kappa}{2}\right)$. Choose real non-negative sequences $\{\beta_n\}$, $\{t_n\}$, $\{\theta_n\}$ and $\{\alpha_n\}$ such that the following conditions hold:*

(1) $\{\beta_n\} \subset [1,+\infty)$, $\sum\limits_{n=1}^{\infty}(\beta_n - 1) < +\infty$ *and* $\sum\limits_{n=1}^{\infty}\theta_n < +\infty$;
(2) $0 < \varphi_1 \leq t_n \leq 1$;
(3) $0 < a \leq \alpha_{n+1} \leq \alpha_n \leq b < 1$.

Iterative steps: *Having x_n and w_{n-1}, compute the next iterate x_{n+1} as follows:*

Step 1. Compute

$$w_n = \nabla f^* \left((1-\alpha_n) \nabla f (x_n) + \alpha_n \nabla f (w_{n-1}) \right).$$

Step 2. Compute

$$y_n = P_C^f \left(\nabla f^* \left(\nabla f (w_n) - \gamma_n A w_n \right) \right).$$

Step 3. Compute

$$x_{n+1} = \nabla f^* \left(t_n \nabla f (y_n) - (t_n - 1) \nabla f (w_n) \right),$$

where

$$\gamma_{n+1} = \begin{cases} \min\left\{ \frac{\sigma \|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \beta_n \gamma_n + \theta_n \right\}, & \text{if } Aw_n \neq Ay_n, \\ \beta_n \gamma_n + \theta_n, & \text{otherwise}. \end{cases} \tag{3.23}$$

Update n and go to the iterative step.

Remark 3.6. *Assume that the mapping $A : E \to E^*$ is Lipschitz continuous with $L > 0$. Let $\{\gamma_n\}$ be the sequence generated by (3.23), using the similar proof as in Remark 3.2, we have*

$$\lim_{n\to\infty} \gamma_n = \gamma \geq \min\left\{ \frac{\sigma}{L}, \gamma_1 \right\}.$$

We first assume that, for some $\beta > 0$, the mapping A satisfies

$$\langle Ax, y - x \rangle \geq 0 \Rightarrow \langle Ay, y - x \rangle \geq \beta D_f (x, y), \forall x \in \mathrm{dom} f, y \in \mathrm{int}(\mathrm{dom} f). \tag{3.24}$$

If we take $f(x) = x \log(x)$ and $Ax = e^x$, then (3.24) holds. Indeed, if $x, y > 0$ satisfy $\langle Ax, y - x \rangle \geq 0$, then we have $y \geq x$. Hence

$$\langle Ay, y - x \rangle = e^y (y - x) \geq y - x \geq D_f (x, y) = x \log\left(\frac{x}{y} \right) + y - x.$$

In addition, if $f(x) = \frac{1}{2}\|x\|^2$, then (3.24) reduces to the case that A is a β-strongly pseudo-monotone mapping, which has been frequently used in the literature, see [31, 43, 49].

Theorem 3.3. *Assume Conditions $(A_1) - (A_4)$ hold and the mapping A satisfies (3.24) with $\beta > \frac{1}{2\min\left\{\frac{\sigma}{L}, \gamma_1\right\}}$. Let $\{x_n\}$ be a sequence generated by Algorithm 2. Then $\{x_n\}$ converges R-linearly to p, where $p \in VI(C, A)$.*

Proof. Since f is κ-strongly convex and satisfies (3.24), the mapping A is a $\frac{\beta\kappa}{2}$-strongly pseudo-monotone mapping. This yields $VI(C, A)$ has no more than one solution, see [21]. Lemma 2.1 implies that there exists $p \in DVI(C, A) = VI(C, A)$.

By the definition of y_n, $y_n \in C$, $p \in VI(C, A)$ and Lemma 2.4, we have

$$\langle \nabla f (w_n) - \gamma_n A w_n - \nabla f (y_n), p - y_n \rangle \leq 0.$$

Combing the three-point identity, we can get

$$\begin{aligned} \gamma_n \langle Aw_n, y_n - p \rangle &\leq \langle \nabla f (w_n) - \nabla f (y_n), y_n - p \rangle \\ &= D_f (p, w_n) - D_f (p, y_n) - D_f (y_n, w_n). \end{aligned} \tag{3.25}$$

It follows from (3.24), the κ-strong convexity of f and the definition of γ_n that

$$\begin{aligned}\langle \gamma_n Aw_n, p-y_n\rangle &= \gamma_n \langle Aw_n - Ay_n, p-y_n\rangle + \gamma_n \langle Ay_n, p-y_n\rangle \\ &\le \gamma_n \|Aw_n - Ay_n\| \|p-y_n\| - \beta\gamma_n D_f(p,y_n) \\ &\le \frac{\sigma\gamma_n}{\gamma_{n+1}} \|w_n - y_n\| \|p-y_n\| - \beta\gamma_n D_f(p,y_n) \\ &\le \frac{\sigma\gamma_n}{2\gamma_{n+1}} \left(\|w_n-y_n\|^2 + \|p-y_n\|^2\right) - \beta\gamma_n D_f(p,y_n) \\ &\le \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} \left(D_f(y_n,w_n) + D_f(p,y_n)\right) - \beta\gamma_n D_f(p,y_n).\end{aligned} \tag{3.26}$$

Combining (3.25) and (3.26), we can get

$$D_f(p,w_n) - D_f(p,y_n) - D_f(y_n,w_n) \ge \left(\beta\gamma_n - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right) D_f(p,y_n) - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} D_f(y_n,w_n),$$

thus yields

$$\left(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} + \beta\gamma_n\right) D_f(p,y_n) \le D_f(p,w_n) - \left(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right) D_f(y_n,w_n). \tag{3.27}$$

By $\beta > \frac{1}{2\min\left\{\frac{\sigma}{L},\gamma_1\right\}}$ and Remark 3.6, we have $\lim_{n\to\infty}\left(1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}}\right) = 1 - \frac{\sigma}{\kappa} > \frac{1}{2}$ and $\lim_{n\to\infty} \beta\gamma_n = \beta\gamma > \frac{1}{2}$. This implies that there exists $N \in \mathbb{N}$ and some number $\tau \in \left(\frac{1}{2}, \min\{1-\frac{\sigma}{\kappa}, \beta\gamma\}\right)$ satisfying

$$1 - \frac{\sigma\gamma_n}{\kappa\gamma_{n+1}} > \tau,\ \forall n \ge N,$$

and

$$\beta\gamma_n > \tau,\ \forall n \ge N.$$

Therefore, we have

$$2\tau D_f(p,y_n) \le D_f(p,w_n),\ \forall n \ge N. \tag{3.28}$$

It follows from the definition of x_{n+1}, $t_n \in (0,1]$ and the Property (c) of the Bregman distance that

$$\begin{aligned}D_f(p,x_{n+1}) &= t_n\left(D_f(p,y_n) - D_f(x_{n+1},y_n)\right) + (1-t_n)\left(D_f(p,w_n) - D_f(x_{n+1},w_n)\right) \\ &\le t_n D_f(p,y_n) + (1-t_n) D_f(p,w_n).\end{aligned} \tag{3.29}$$

Combining $1 \ge t_n \ge \varphi_1 > 0$, $\tau > \frac{1}{2}$, (3.28) and (3.29), we can get

$$D_f(p,x_{n+1}) \le \left(\frac{t_n}{2\tau} + 1 - t_n\right) D_f(p,w_n) \le \left(\frac{\varphi_1}{2\tau} + 1 - \varphi_1\right) D_f(p,w_n),\ \forall n > N. \tag{3.30}$$

Shifting the index in (3.13) (by taking $n \equiv n+1$) and rearranging terms, we have

$$\begin{aligned}D_f(p,x_{n+1}) &= \frac{1}{1-\alpha_{n+1}} D_f(p,w_{n+1}) - \frac{\alpha_{n+1}}{1-\alpha_{n+1}} D_f(p,w_n) \\ &\quad + \frac{\alpha_{n+1}}{1-\alpha_{n+1}}\left(D_f(x_{n+1},w_n) - \frac{1}{\alpha_{n+1}} D_f(x_{n+1},w_{n+1})\right).\end{aligned} \tag{3.31}$$

Shifting the index in (3.16) (by taking $n \equiv n+1$) and combing (3.31) and $0 < a \leq \alpha_{n+1} \leq \alpha_n \leq b < 1$, we have

$$D_f\left(p, x_{n+1}\right) \geq \frac{1}{1-\alpha_{n+1}} D_f\left(p, w_{n+1}\right) - \frac{\alpha_{n+1}}{1-\alpha_{n+1}} D_f\left(p, w_n\right). \tag{3.32}$$

Using (3.30) and (3.32), we can get

$$D_f\left(p, w_{n+1}\right) \leq \left(\left(\varphi_1 - \frac{\varphi_1}{2\tau}\right)\alpha_{n+1} + \frac{\varphi_1}{2\tau} + 1 - \varphi_1\right) D_f\left(p, w_n\right), \ \forall n \geq N.$$

Due to $0 < a \leq \alpha_n \leq b < 1$, we can get

$$D_f\left(p, w_{n+1}\right) \leq \left(\left(\varphi_1 - \frac{\varphi_1}{2\tau}\right) b + \frac{\varphi_1}{2\tau} + 1 - \varphi_1\right) D_f\left(p, w_n\right), \ \forall n \geq N.$$

It follows from (3.30), the κ-strong convexity of f and the above inequality that

$$\frac{\upsilon\kappa}{2}\left\|p - x_{n+2}\right\|^2 \leq \upsilon D_f\left(p, x_{n+2}\right) \leq D_f\left(p, w_{n+1}\right) \leq \rho D_f\left(p, w_n\right) \leq \ldots \leq \rho^{n-N+1} D_f\left(p, w_N\right),$$

where $\upsilon = \frac{1}{\frac{\varphi_1}{2\tau} + 1 - \varphi_1}$ and $\rho = \left(\varphi_1 - \frac{\varphi_1}{2\tau}\right) b + \frac{\varphi_1}{2\tau} + 1 - \varphi_1$. Thanks to $1 \geq \varphi_1 > 0$, $\tau > \frac{1}{2}$ and $0 < b < 1$, we can find that $\upsilon > 0$ and $0 < \rho < 1$. Consequently, $\{x_n\}$ converges R-linearly to p. $\square$

Theorem 3.4. *Assume Conditions $(A_1) - (A_4)$ hold and the mapping A satisfies (3.24). Let the step size sequence $\gamma_n = \ell_n$ satisfying $\lim_{n\to\infty} \ell_n = \ell \in \left(0, \frac{\kappa}{L}\right)$ and $\ell_n \geq 0$, $\forall n \in \mathbb{N}$, and $\kappa > \frac{L}{\beta}$. Then $\{x_n\}$ generated by Algorithm 2 converges R-linearly to p, where $p \in VI\left(C, A\right)$.*

Proof. it follows from (3.24), the κ-strong convexity of f and the definition of ℓ_n that

$$\begin{aligned} \left\langle \ell_n A w_n, p - y_n \right\rangle &= \ell_n \left\langle A w_n - A y_n, p - y_n \right\rangle + \ell_n \left\langle A y_n, p - y_n \right\rangle \\ &\leq \ell_n \left\|A w_n - A y_n\right\| \left\|p - y_n\right\| - \beta \ell_n D_f\left(p, y_n\right) \\ &\leq L\ell_n \left\|w_n - y_n\right\| \left\|p - y_n\right\| - \beta \ell_n D_f\left(p, y_n\right) \\ &\leq \frac{L\ell_n}{2}\left(\left\|w_n - y_n\right\|^2 + \left\|p - y_n\right\|^2\right) - \beta \ell_n D_f\left(p, y_n\right) \\ &\leq \frac{L\ell_n}{\kappa}\left(D_f\left(y_n, w_n\right) + D_f\left(p, y_n\right)\right) - \beta \ell_n D_f\left(p, y_n\right). \end{aligned} \tag{3.33}$$

Combining (3.25) and (3.33), we can get

$$\left(1 - \frac{L\ell_n}{\kappa} + \beta\ell_n\right) D_f\left(p, y_n\right) \leq D_f\left(p, w_n\right) - \left(1 - \frac{L\ell_n}{\kappa}\right) D_f\left(y_n, w_n\right).$$

By $\lim_{n\to\infty} \ell_n = \ell \in \left(0, \frac{\kappa}{L}\right)$ and $\kappa > \frac{L}{\beta}$, we have

$$\lim_{n\to\infty}\left(1 - \tfrac{L\ell_n}{\kappa}\right) = 1 - \tfrac{L\ell}{\kappa} > 0 \text{ and } \lim_{n\to\infty}\left(\beta\ell_n + 1 - \tfrac{L\ell_n}{\kappa}\right) = \beta\ell + 1 - \tfrac{L\ell}{\kappa} > 1.$$

This implies that there exists $N \in \mathbb{N}$ and some number $\tau \in \left(\frac{1}{2}, \frac{1}{2}\left(\beta\ell + 1 - \frac{L\ell}{\kappa}\right)\right)$ satisfying

$$\beta\ell_n + 1 - \frac{L\ell_n}{\kappa} > 2\tau, \ \forall n \geq N.$$

Therefore, we have

$$2\tau D_f\left(p, y_n\right) \leq D_f\left(p, w_n\right), \ \forall n \geq N.$$

The rest of the proof now follows the same arguments as those used in the proof of Theorem 3.3, so we omit it. We obtain this conclusion. $\square$

Remark 3.7. *From Theorem 3.4, it follows that the upper boundary of the limitation of the step size sequence $\gamma_n = \ell_n$ is controlled by $\frac{\kappa}{L}$. Note that if L is too large, then the step sizes become correspondingly small, which may affect the convergence rate of Algorithm 2, see Example 4.3. Therefore, in Algorithm 2 we sometimes still adopt the adaptive step sizes defined by* (3.23)*, though we need know the Lipschitz constant L of the mapping A in advance by Theorem 3.3.*

Remark 3.8. *As far as we know, Bregman projection algorithms using the golden ratio technique for solving variational inequalities do not have a linear convergence result in reflexive Banach spaces. Theorem 3.3 presents a new discovery in this regard.*

Additionally, if $\alpha_n = 0$, that is, $w_n = x_n$, then $\{x_n\}$ also converges R-linearly to p. In fact, it follows from $w_n = x_n$ and (3.30) that

$$D_f(p, x_{n+1}) \leq \left(\frac{\varphi_1}{2\tau} + 1 - \varphi_1\right) D_f(p, x_n), \forall n > N.$$

This means that

$$D_f(p, x_{n+1}) \leq \left(\frac{\varphi_1}{2\tau} + 1 - \varphi_1\right) D_f(p, x_n) \leq \ldots \leq \left(\frac{\varphi_1}{2\tau} + 1 - \varphi_1\right)^{n-N+1} D_f(p, x_N).$$

Combining $1 \geq \varphi_1 > 0$ and $\tau > \frac{1}{2}$, we have $0 < \frac{\varphi_1}{2\tau} + 1 - \varphi_1 < 1$ and so we can get the sequence $\{x_n\}$ converges R-linearly to p.

4. Numerical experiments

In this section, we showcase three numerical experiments for our proposed algorithm to illustrate its performance. We conduct all computations using Matlab 2023(b) on a PC equipped with 8.00GB RAM. Let "Iter" denote number of iteration and "Time" denote the CPU time in seconds.

Example 4.1. *Consider $E = l_2 = \left\{x = (x_1, x_2, \ldots) : \sum\limits_{i=1}^{\infty} |x_i|^2 < \infty\right\}$ with $C = \{x \in l_2 : \|x\| \leq 2\}$ and $f(x) = \frac{1}{2}\|x\|^2$. The Bregman distance associated with f is referred to as Euclidean distance given by $D_f(x, y) = \frac{1}{2}\|x - y\|^2$. Then we have*

$$P_C^f(x) = \frac{2}{\max\{\|x\|, 2\}} x,$$

and define

$$Ax := (3 - \|x\|)\, x.$$

Then A is quasi-monotone on C and L-Lipschitz continuous on E with $DVI(C, A) = \{\mathbf{0}\}$, for more details, see [1].

In this experiment, we compare our Algorithm 1 (namely, Alg1) with the Algorithms 1 and 2 (namely, wAlg1 and wAlg2) of Wang et al. [46]*, and the Algorithm 3.12 (namely, AMAlg) of Alakoya et al.* [1]*. The termination condition for the test is $E_n = \|x_n - \mathbf{0}\|^2 < 10^{-8}$. The parameters of each algorithm are set as follows:*

- *Alg1: $t_n = 1 - \frac{1}{n^3+2}$, $\alpha_n = \frac{1}{16.8} + \frac{1}{3n^2}$, $\sigma = 0.29$, $\theta_n = \frac{2}{7n^2+1}$, $\gamma_1 = 1.4$, $\beta_n = 1 + \frac{1}{n\sqrt{n}}$ and $w_0 = x_1 = (\underbrace{0.2, 0.2, \ldots, 0.2}_{m}, 0, \ldots, 0, \ldots)$;*
- *wAlg1: $\beta = 0.9$, $\delta = 0.9$, $\mu = 0.9$, $p_n = 0$, $\xi_n = \frac{0.1}{n^2}$, $\lambda_1 = 1$ and*

$x_0 = x_1 = (\underbrace{0.2, 0.2, \ldots, 0.2}_{m}, 0, \ldots, 0, \ldots)$;

• *wAlg2:* $\beta = 0.9,\ \delta = 0.9,\ \mu = 0.9,\ p_n = 0,\ \xi_n = \frac{0.1}{n^2},\ \lambda_1 = 1$ *and* $x_0 = x_1 = (\underbrace{0.2, 0.2, \ldots, 0.2}_{m}, 0, \ldots, 0, \ldots)$;

• *AMAlg:* $\lambda_0 = 0.9,\ \theta = 0.8,\ \mu = 0.6,\ \rho_n = \frac{1000}{(n+1)^2},\ f(x) = \frac{x}{3},\ \alpha_n = \frac{1}{2n},\ \xi_n = \frac{1}{(2n+1)^3}$ *and* $x_0 = x_1 = (\underbrace{0.2, 0.2, \ldots, 0.2}_{m}, 0, \ldots, 0, \ldots)$.

The results of experiments are reported in Table 1 and Figure 1.

TABLE 1. Numerical results of Example 4.1.

m	*Alg1*	*wAlg1*	*wAlg2*	*AMAlg*
	Iter	*Iter*	*Iter*	*Iter*
510	*38*	*49*	*64*	*53*
520	*37*	*49*	*63*	*40*
540	*28*	*50*	*60*	*56*

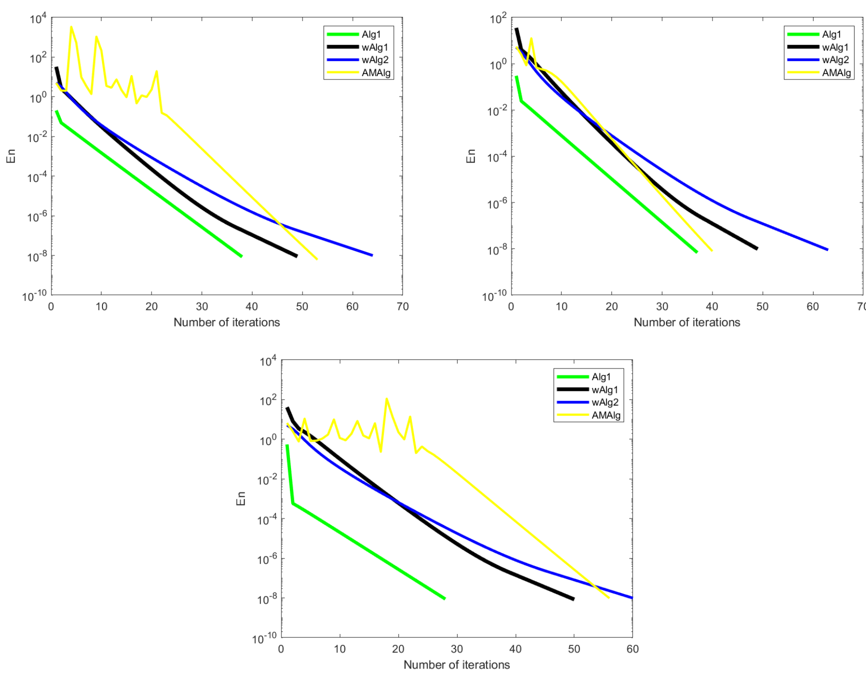

FIGURE 1. Numerical behaviour of E_n for Example 4.1, Top left: m=510; Top right: m=520; Bottom: m=540.

Remark 4.9. *It is clear from Table 1 and Figure 1 that our proposed algorithm has superior convergence regarding iteration counts to the referenced algorithms.*

Example 4.2. *Consider $E = \mathbb{R}^m$ with*

$$C = \left\{ x = (x_1, x_2, \ldots, x_m)^\top \in \mathbb{R}^m : x_i \geq \frac{1}{2\sqrt{m}} \text{ and } \sum_{i=1}^{m} x_i^2 \leq 1,\ i = 1, 2, \ldots, m \right\}$$

and the negative entropy $f(x) = \sum_{i=1}^{m} x_i \log(x_i)$, $x_i > 0$. Then

$$\nabla f(x) = (1 + \log(x_1), \ldots, 1 + \log(x_m))^\top$$

and

$$\nabla f^*(x) = \left(e^{x_1 - 1}, \ldots, e^{x_m - 1}\right)^\top$$

The Bregman distance associated with f is referred to as KL divergence given by $D_f(x, y) = \sum_{i=1}^{m} x_i \log\left(\frac{x_i}{y_i}\right) + y_i - x_i$. Here, we employ the function fmincon in Optimization Toolbox of MATLAB for computing Bregman projection onto C. The mapping $A : E \to E$ is defined by

$$\begin{aligned}
&Ax = F_1(x) + F_2(x), \\
&F_1(x) = (h_1(x), h_2(x), \ldots, h_m(x)), \\
&F_2(x) = Dx + d, \\
&h_i(x) = x_{i-1}^2 + x_i^2 + x_{i-1}x_i + x_i x_{i+1},\ i = 1, 2, \ldots, m, \\
&x_0 = x_{m+1} = 0,
\end{aligned}$$

where $d = (-1, -1, \ldots, -1)^\top$ and $D_{m \times m}$ given by

$$D_{ij} = \begin{cases} 4 \text{ if } i = j, \\ 1 \text{ if } i = j + 1, \\ -2 \text{ if } j = i + 1, \\ 0 \text{ otherwise.} \end{cases}$$

Then A is pseudo-monotone and L-Lipschitz continuous, see [37]. *Many authors have taken into consideration the variational inequality with A and C, see* [14, 46].

In this experiment, we compare our Algorithm 1 (namely, Alg1) with the Algorithm 1 and Algorithm 2 (namely, wAlg1 and wAlg2) of Wang et al. [46], *Algorithm 3.1 (namely, LYAlg) of Liu and Yang* [25]. *To terminate the iterations, all algorithms use $E_n = \|x_{n+1} - x_n\|^2 < 10^{-8}$. The parameters of each algorithm are set as follows:*

- *Alg1: $t_n = \frac{1}{n^2}$, $\alpha_n = \frac{1}{1.65} + \frac{1}{3n^2}$, $\sigma = 0.99$, $\theta_n = \frac{2}{7n^2+1}$, $\gamma_1 = 1$, $\beta_n = 1 + \frac{1}{n\sqrt{n}}$ and $w_0 = x_1 = (0.1, 0.1, \ldots, 0.1)^\top$;*
- *wAlg1: $\beta = 0.9$, $\delta = 0.9$, $\mu = 0.9$, $p_n = 0$, $\xi_n = \frac{0.1}{n^2}$, $\lambda_1 = 1$ and $x_0 = x_1 = (0.1, 0.1, \ldots, 0.1)^\top$;*
- *wAlg2: $\beta = 0.9$, $\delta = 0.9$, $\mu = 0.85$, $p_n = 0$, $\xi_n = \frac{0.1}{n^2}$, $\lambda_1 = 1$ and $x_0 = x_1 = (0.1, 0.1, \ldots, 0.1)^\top$;*
- *LYAlg: $\mu = 0.5$, $p_n = \frac{100}{(n+1)^{1.1}}$, $\lambda_0 = 1$ and $x_1 = (0.1, 0.1, \ldots, 0.1)^\top$.*

The results of experiments are reported in Table 2 and Figure 2.

Remark 4.10. *From Table 2 and Figure 2, we see that our method outperforms all other methods based on the negative entropy function.*

TABLE 2. Numerical results of Example 4.2.

m	*Alg1*	*wAlg1*	*wAlg2*	*LYAlg*
	Iter	*Iter*	*Iter*	*Iter*
35	*17*	*23*	*26*	*43*
45	*18*	*24*	*26*	*39*
55	*19*	*24*	*26*	*35*

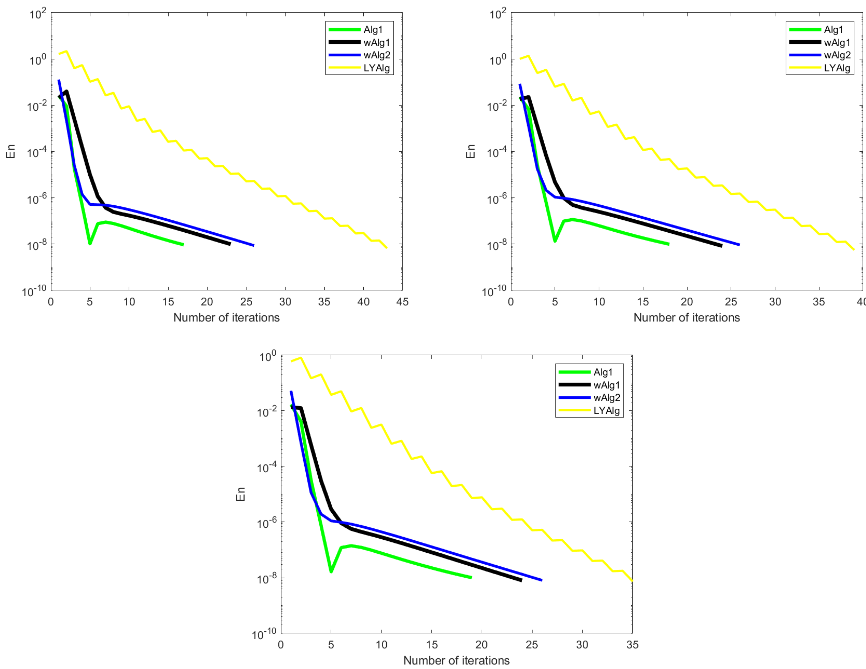

FIGURE 2. Numerical behaviour of E_n for Example 4.2, Top left: m=35; Top right: m=45; Bottom: m=55.

Example 4.3. *Let* $E = \mathbb{R}^3$ *with*

$$C = \left\{x = (x_1, x_2, x_3)^\top \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0, -5 \leq x_i \leq 5,\ i = 1, 2, 3\right\}.$$

Let $f(x) = \frac{1}{2}\|x\|^2$ *and*

$$Ax = \left(e^{-\|x\|^2} + 10\right) Mx,$$

where

$$M = \begin{bmatrix} 1.7 & 0 & 0 \\ 0 & 1.71 & 0 \\ 0 & 0 & 1.69 \end{bmatrix}.$$

Then A is θ-strongly pseudo-monotone on E with $\theta \approx 10\lambda_{\min} \approx 16.9$, where $\lambda_{\min}$ is the smallest eigenvalue of M (see [43]*), and A is L-Lipschitz continuous on E with $L \approx 32.39$ (by $\|\nabla Ax\| \leq L,\ \forall x \in E$), and we can know $VI(C, A) = \{\mathbf{0}\}$.*

Firstly, we consider two cases of step size for Algorithm 2, which are simply denoted as Alg21 and Alg22, respectively, and compare them with Algorithm 1 (namely, SEMGR) of Oyewole and Reich [31], *and RIPA of Vuong* [43]*. Our termination criterion is $E_n = \|x_n - \mathbf{0}\|^2 < 10^{-50}$. The parameters of each algorithm are set as follows:*

• *RIPA: $\rho = 1.26,\ \theta = 0,\ \eta = 1.5,\ \lambda = \frac{1.99\theta}{\eta L^2}$ and $x_0 = x_1 = (-3, 2, 3)^\top$;*

• *SEMGR: $\varphi = 150,\ \mu = 0.45,\ \delta_n = \frac{2}{7n^2+1},\ \lambda_1 = 0.07,\ \beta_n = 1 + \frac{1}{n\sqrt{n}}$ and $w_0 = x_1 = (-3, 2, 3)^\top$;*

•*Alg21: $t_n = 1 - \frac{2}{2n^4+1},\ \alpha_n = \frac{1}{150} + \frac{1}{150n^5},\ \sigma = 0.4998,\ \theta_n = \frac{2}{2n^2+1},\ \gamma_1 = 1,\ \beta_n = 1 + \frac{1}{n^2}$ and $w_0 = x_1 = (-3, 2, 3)^\top$;*

•*Alg22: $t_n = 1 - \frac{2}{2n^4+1},\ \alpha_n = \frac{1}{150} + \frac{1}{150n^5},\ \gamma_n = \frac{1}{L}\left(\frac{19}{20} - \frac{4}{5n}\right)$ and $w_0 = x_1 = (-3, 2, 3)^\top$.*

The results of experiments are reported in Table 3 and Figure 3.

TABLE 3. Numerical results of Example 4.3.

	RIPA	*SEMGR*	*Alg21*	*Alg22*
Iter	*88*	*124*	*15*	*84*
Time	*0.0183*	*0.0752*	*0.0068*	*0.0162*

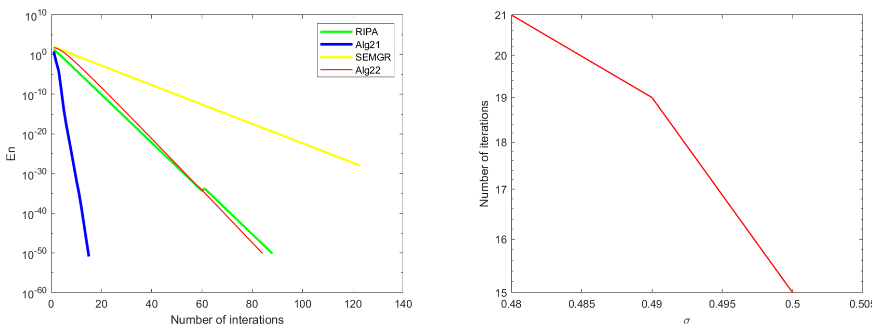

FIGURE 3. Numerical behaviour for Example 4.3, Left: E_n versus Iter; Right: Iter versus σ in Alg21.

Remark 4.11. *Table 3 and Figure 3 (left) display Alg21 outperforms Alg22, whose reason may be the Lipschitz constant L of A is too large. As can be shown in Table 3 and Fig. 3, Algorithm 2 is far better than SEMGR of* [31] *and RIPA of* [43]*. Next, we take $t_n = 1 - \frac{2}{2n^4+1},\ \alpha_n = \frac{1}{150} + \frac{1}{150n^5},\ \theta_n = \frac{2}{2n^2+1},\ \gamma_1 = 1,\ \beta_n = 1 + \frac{1}{n^2}$, $w_0 = x_1 = (-3, 2, 3)^\top$, $E_n = \|x_n - \mathbf{0}\|^2 < 10^{-50},\ f(x) = \frac{1}{2}\|x\|^2$, and consider the influence of different σ on the performance of Algorithm 2. The numerical behavior is shown in Figure 3 (right).*

5. CONCLUSIONS

To address variational inequalities in reflexive Banach spaces, two Bregman projection algorithms with a new extrapolation technique are introduced. The weak convergence and non-asymptotic $O\left(\frac{1}{\sqrt{n}}\right)$ convergence rate of the Algorithm 1 are established under appropriate and mild assumptions. The linear convergence rate of Algorithm 2 is proved. Our results have the following benefits advantages:

(i) Algorithm 1 and Algorithm 2 both have non-monotone adaptive step sizes that only need a simple calculation of known information, eliminate the limitation of the Lipschitz constant of A and lessen the dependence on the initial point.

(ii) Instead of being pseudo-monotone and weakly sequentially continuous, like in Theorem 3.8 of [17], Theorem 3.1 only requires A is quasi-monotone and satisfies the Condition (A_5). Particularly, to the best of our knowledge, Bregman projection algorithms with the golden ratio technique for solving variational inequalities have no linear convergent result in reflexive Banach spaces. Theorem 3.3 and Theorem 3.4 are new results.

(iii) Different from the inertial acceleration, Step 1 of Algorithm 1 and Algorithm 2 is an extension of the golden ratio technique which can also speed up the convergence of algorithms, see [31, 49]. Our Algorithms 1 and 2 have been tested through numerical experiments and have shown to be more efficient than the corresponding algorithms presented in [1, 25, 31, 43, 46].

Acknowledgments

This study received financial support from Railway Foundation Joint Research Funds of the National Natural Science Foundation of China (U236820070), the National Natural Science Foundation of China (11701479, 11701478), the Chinese Postdoctoral Science Foundation (2018M643434), the Fundamental Research Funds for the Central Universities (2682021ZTPY040) and the National Science, Research and Innovation Fund, Thailand Science Research and Innovation (TSRI), through Rajamangala University of Technology Thanyaburi (FRB68E0617) (Grant No.: FRB680045/0168).

References

[1] Alakoya, T.O.; Mewomo O.T.; Shehu Y. Strong convergence results for quasimonotone variational inequalities. *Math. Methods Oper. Res.* **95** (2021), 249–279.

[2] Bauschke, H.H.; Borwein, J.M.; Combettes, P.L. Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces. *Commun. Contemp. Math.* **3** (2021), 615–647.

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

[4] Butnariu, D.; Resmerita, E. Bregman distances, totally convex functions, and a method for solving operator equations in Banach spaces. *Abstr. Appl. Anal.* **39** (2006), Art. ID 84919.

[5] Butnariu, D.; Iusem, A.N. Totally convex functions for fixed points computation and infinite dimensional optimization. Applied Optimization, Kluwer Academic Publishers: Dordrecht, 2000.

[6] Chang, X.K.; Yang, J.F. A golden ratio primal-dual algorithm for structured convex optimization. *J. Sci. Comput.* **87** (2021), 47.

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

[8] Facchinei, F.; Pang, J.S. Finite-dimensional variational inequalities and complementarity problems. Vol. I. Springer series in operations research, New York: Springer, 2003.

[9] Fichera, G. Sul problema elastostatico di Signorini con ambigue condizioni al contorno. *Rend. Accad. Naz. Lincei, s. VIII.* **34** (1963).

[10] Fichera, G. Problemi elastostatici con vincoli unilaterali: il problema di Signorini conambigue condizioni al contorno. *Rend. Accad. Naz. Lincei, s. VIII.* **7** (1964).

[11] Huang, Y.Y.; Jeng, J.C.; Kuo, T.Y.; et al. Fixed point and weak convergence theorems for point-dependent λ-hybrid mappings in Banach spaces. *Fixed Point Theory Appl.***105** (2011), 105.

[12] Hu, S.; Wang, Y.; Dong, Q.L. Convergence analysis of a new Bregman extragradient Method for solving fixed point problems and variational inequality problems in reflexive Banach spaces. *J. Sci. Comput.* **96** (2023), 19.

[13] Hieu, D.V.; Cho, Y.J.; Xiao, Y.; Kumam; P. Relaxed extragradient algorithm for solving pseudomonotone variational inequalities in Hilbert spaces. *Optimization* **69** (2020), 2279–2304.

[14] Hieu, D.V.; Cholamjiak, P. Modified extragradient method with Bregman distance for variational inequalities. *Appl. Anal.* **101** (2022), 655–670.

[15] Izuchukwu, C.; Shehu, Y.; Yao, J.C. New strong convergence analysis for variational inequalities and fixed point problems. *Optimization* (2024), 1–22.

[16] Izuchukwu, C.; Reich, S.; Shehu, Y. One-step Bregman projection methods for solving variational inequalities in reflexive Banach spaces. *Optimization* **73** (2023), 1519–1549.

[17] Jolaoso, L.O.; Shehu, Y. Single Bregman projection method for solving variational inequalities in reflexive Banach spaces. *Appl. Anal.* **101**(2022), 4807–4828.

[18] Jolaoso, L.O.; Aphane, M. Bregman subgradient extrageadient method with monotone sele-adjustment stepsize for solving pseudomonotone variational inequalities and fixed point problems. *J. Ind. Manag. Optim.* **18** (2022), 773–794.

[19] Jolaoso, L.O.; Sunthrayuth, P.; Cholamjiak, P.; Cho; Y.J. Analysis of two versions of relaxed inertial algorithms with Bregman divergences for solving variational inequalities. *Comput. Appl. Math.* **41** (2022), 300.

[20] Kien, B.T.; Yao, B.T.; Yen, N.D. On the solution existence of pseudomonotone variational inequalities. *J. Glob. Optim.* **41** (2007), 135–145.

[21] Kim, D.S.; Vuong, P.T.; Khanh, P.D. Qualitative properties of strongly pseudomonotone variational inequalities. *Optim. Lett.* **10** (2016), 1669–1679.

[22] Kinderlehrer, D.; Stampacchia, G. An introduction to variational inequalities and their applications. New York: Academic Press, 1980.

[23] Konnov, I.V. Equilibrium models and variational inequalities. Elsevier B V: Amsterdam, 2007.

[24] Korpelevich, G.M. The extragradient method for finding saddle points and other problems. *Ekno. Mat. Metod.* **12** (1976), 747–756.

[25] Liu, H.; Yang, J. Weak convergence of iterative methods for solving quasi-monotone variational inequalities. *Comput. Optim. Appl.* **77** (2020), 491–508.

[26] Malitsky, Y. Golden ratio algorithms for variational inequalities. *Math. Program.* **184** (2018), 383–410.

[27] Martín-Márquez, V.; Reich, S.; Sabach, S. Bregman strongly nonexpansive operators in reflexive Banach spaces. *J. Math. Anal. Appl.* **400** (2013), 597–614.

[28] Martín-Márquez, V.; Reich, S.; Sabach, S. Iterative Methods for approximating fixed points of Bregman nonexpansive operators. *Discrete Contin. Dyn. Syst. Ser. S.* **6** (2013), 1043–1063.

[29] Naraghirad, E.; Yao, J.C. Bregman weak ralatively nonexpansive mapping in Banach spaces. *Fixed Point Theory Appl.* **141** (2013), 43.

[30] Osilike, M.O.; Aniagbosor, S.C. Weak and strong convergence theorems for fixed points of asymptotically nonexpansive mappings. *Math. Comput. Modelling* **32** (2000), 1181–1191.

[31] Oyewole, O.; Reich, S. Two subgradient extragradient methods based on the golden ratio technique for solving variational inequality problems. *Numer. Algorithms* **97** (2024), 1215–1236.

[32] Pathak, H.K. An introduction to nonlinear analysis and fixed point theory, Springer: Singapore, 2018.

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

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

[35] Reich, S.; Sabach, S. Two strong convergence theorems for Bregman strongly nonexpansive operators in reflexive Banach spaces. *Nonlinear Anal.* **73** (2010), 122–135.

[36] Reich, S.; Tuyen, T.M.; Sunthrayuth; Cholamjiak, P. Two new inertial algorithms for solving variational inequalities in reflexive Banach spaces. *Numer. Funct. Anal. Optim.* **42** (2021), 1954–1984.

[37] Sun, D. A projection and contraction method for the nonlinear complementarity problems and its extensions. Math Number Sin. **16** (1994), 183–194.

[38] Tam, M.K.; Uteda, D.J. Bregman-Golden Ratio algorithms for variational inequalities. *J. Optim. Theory Appl.* **199** (2023), 993–1021.

[39] Tang, G.J.; Huang, N.J. Existence theorems of the variational-hemivariational inequalities. *J. Glob. Optim.* **56** (2013), 605–622.

[40] Thong, D.V.; Triet, N.A.; Li, X.H.; Dong, Q.L. Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems. *Numer. Algorithms* **83** (2020), 1123–1143.

[41] Thong, D.V.; Yang, J.; Cho, Y.J.; Rassias; T.M. Explicit extragradient-like method with adaptive stepsizes for pseudomonotone variational inequalities. *Optim. Lett.* **15** (2021), 2181–2199.

[42] Tseng, P. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control Optim.* **38** (2000), 431–446.

[43] Vuong, P.T. A second order dynamical system and its discretization for strongly pseudo-monotone variational inequalities. *SIAM J. Control Optim.* **59**, (2021), 2875–2897.

[44] Wang, K.; Wang, Y.H.; Shehu, Y.; Jiang, B. Double inertial forward-backward-forward method with adaptive step size for variational inequalities with quasi-monotonicity. *Commun. Nonlinear Sci. Numer. Simul.* **132** (2024), 107924.

[45] Wang, Z.B.; Chen, Z.Y.; Xiao, Y.B.; Zhang, C. A new projection-type method for solving multi-valued mixed variational inequalities without monotonicity. *Appl. Anal.* **99** (2020), 1453–1466.

[46] Wang, Z.B.; Sunthrayuth, P.; Adamu, A.; Cholamjiak, P. Modified accelerated Bregman projection methods for solving quasi-monotone variational inequalities. *Optimization* **73** (2024), 2053–2087.

[47] Xie, Z.B.,; Cai, G.; Dong, Q.L. Strong convergence of Bregman projection method for solving variational inequality problems in reflexive Banach spaces. *Numer. Algorithms* **93** (2023), 269–294.

[48] Yang, J.; Cholamjiak, P.; Sunthrayuth, P. Weak and strong convergence results for solving monotone variational inequalities in reflexive Banach spaces. *Optimization* **72** (2023), 2609–2634.

[49] Zhang, C.J.; Chu, Z.Y. New extrapolation projection algorithms based on the golden ratio for pseudo-monotone variational inequalities. *AIMS Math.* **8** (2023), 23291–23312.

[1] Department of Mathematics, Southwest Jiaotong University, Chengdu, Sichuan 611756, China

Email address: zhongbaowang@hotmail.com, 1332147624@qq.com, zhangyouchen@swjtu.edu.cn

[2] National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Chengdu, Sichuan 611756, China

[3] Department of Mathematics and Computer Science, Faculty of Science and Technology, Rajamangala University of Technology Thanyaburi, Thanyaburi, Pathumthani 12110, Thailand

[4] AI-Powered Digital Platform Research Unit, Rajamangala University of Technology Thanyaburi, Thanyaburi, Pathumthani 12110, Thailand

Email address: pongsakorn_su@rmutt.ac.th