Computational and Applied Mathematics (2023) 42:254
https://doi.org/10.1007/s40314-023-02394-8

An improved subgradient extragradient method with two different parameters for solving variational inequalities in reflexive Banach spaces

Huanqin Wu[1] · **Zhongbing Xie**[2] · **Min Li**[3]

Received: 25 February 2023 / Revised: 4 June 2023 / Accepted: 27 June 2023 /
Published online: 18 July 2023

Abstract

In this paper, we combine the classical subgradient extragradient method with the Bregman projection method for solving variational inequality problems in reflexive Banach spaces. Specifically, we set two different parameters in the two-step projections, as opposed to consistent parameters in other results. In addition, the application of the inertial technique accelerates the iteration efficiency. Finally, we compare the proposed algorithm with other known results and find that our method effectively improves the convergence process.

Keywords Banach space · Bregman projection · Strong convergence · Subgradient extragradient method · Variational inequality

Mathematics Subject Classification 47H05 · 47H07 · 47H10 · 54H25

1 Introduction

The main purpose of this paper is to study the variational inequality problem in Banach spaces. It is well known that the variational inequality problem consists in finding $p \in \mathcal{C}$ such that

$$\langle \mathscr{F} p, q - p \rangle \geq 0, \ \forall q \in \mathcal{C}, \tag{1}$$

✉ Zhongbing Xie
xzbmath@163.com

Huanqin Wu
hqwu@cug.edu.cn

Min Li
lm7math@163.com

1 School of Computer Science, China University of Geosciences, Wuhan 430074, China

2 School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China

3 College of Basic Education, Chongqing Industry and Trade Polytechnic, Chongqing 408000, China

where $\mathscr{F} : \mathscr{B} \to \mathscr{B}^*$ is a nonlinear operator, $\langle t, s \rangle : \mathscr{B}^* \times \mathscr{B} \to \mathbb{R}$ is the duality pairing for $t \in \mathscr{B}^*$ and $s \in \mathscr{B}$, $\mathscr{B}$ is a real Banach space with the dual $\mathscr{B}^*$ and $\mathcal{C}$ is a nonempty closed convex subset of $\mathscr{B}$. Throughout the paper, to simplify the narrative, we denote the variational inequality problem and its solution set by **VIP** and $\mathcal{S}$ respectively.

The theory of variational inequalities can be traced back to the 1960s, when the problem first appeared in the Signorini problem proposed by Antonio Signorini, and then gradually obtained complete results in the research of Fichera, Stampacchia and Lions. With the continuous scientific and technological innovations, many researchers applied this theory to different fields such as mechanics, economics and mathematics. Particularly, in mathematics, the VIP is closely related to saddle-point, equilibrium and fixed-point problems, see for example, Yao et al. (2020), Ceng et al. (2014), Yao et al. (2011), Barbagallo and Di Vincenzo (2015), Lions (1977), Jitpeera and Kumam (2010) and the references therein.

In recent years, many different methods have been developed to solve VIP. One of the simplest of these methods is the projected gradient method (for short, **PGM**), which can be expressed as follows:

$$w_{i+1} = \mathscr{P}_{\mathcal{C}}(w_i - \tau \mathscr{A} w_i), \tag{PGM}$$

where $\mathscr{P}_{\mathcal{C}}$ denotes the metric projection from Hilbert space $\mathscr{H}$ onto the feasible set $\mathcal{C}$, the operator $\mathscr{A} : \mathscr{H} \to \mathscr{H}$ is strongly monotone. Since the PGM is limited by strong assumptions, this will greatly affect its applicability. Therefore, Korpelevich (1976) proposed the extragradient method (for short, **EGM**) using the double-projection iteration, thereby weakening this condition and improving the applicability of the iterative method. The EGM is as follows:

$$\begin{cases} t_i = \mathscr{P}_{\mathcal{C}}(w_i - \tau \mathscr{A} w_i), \\ w_{i+1} = \mathscr{P}_{\mathcal{C}}(w_i - \tau \mathscr{A} t_i), \end{cases} \tag{EGM}$$

where $\mathscr{P}_{\mathcal{C}} : \mathscr{H} \to \mathcal{C}$ denotes the metric projection, $\mathscr{A}$ is monotone L-Lipschitz continuous operator and $\tau \in (0, \frac{1}{L})$. Under reasonable assumptions, Korpelevich obtained a weak convergence theorem for the sequence generated by EGM. Since then, the authors have combined various techniques to improve the convergence based on the EGM and have obtained further results, see for example, Jitpeera and Kumam (2010), Hieu et al. (2020), Xie et al. (2021), Dong et al. (2016), Tan et al. (2022) and the references therein.

However, although the EGM weakens the constraints based on the PGM, the projection onto the feasible set $\mathcal{C}$ must to be computed twice during each iteration of the loop. To reduce the computational effort, Censor et al. (2011) in a later study proposed the following subgradient extragradient method (for short, **SEGM**):

$$\begin{cases} t_i = \mathscr{P}_{\mathcal{C}}(w_i - \tau \mathscr{A} w_i), \\ w_{i+1} = \mathscr{P}_{T_i}(w_i - \tau \mathscr{A} t_i), \end{cases} \tag{SEGM}$$

where $\mathscr{P}_{\mathcal{C}}$, $\mathscr{A}$ and τ are the same as defined in EGM. In particular, they constructed a half-space $T_i := \{x \in \mathscr{H} : \langle w_i - \tau \mathscr{A} w_i - t_i, x - t_i \rangle \leq 0\}$ to replace the feasible set $\mathcal{C}$ in the second step of the projection process. By the definition of T_i, the projection is easy to calculate. The weak convergence theorem of SEGM in Hilbert spaces has been proved under some appropriate assumptions. At the same time, SEGM has also attracted the interest of many authors, see for instance, Jolaoso et al. (2021), Yao et al. (2022), Yang et al. (2020), Abubakar et al. (2022) and the references therein.

Noting that the above methods all yield the corresponding weak convergence theorems, it is natural to further consider the strong convergence. Kraikaew and Saejung (2014) combined the SEGM with the Halpern method and obtained the following algorithm:

$$\begin{cases} t_i = \mathscr{P}_C(w_i - \tau \mathscr{A} w_i), \\ s_i = \mathscr{P}_{T_i}(w_i - \tau \mathscr{A} t_i), \\ w_{i+1} = \alpha_i w_1 + (1-\alpha_i) s_i, \end{cases} \tag{HSEGM}$$

where the metric projection $\mathscr{P}_C$, the operator $\mathscr{A}$, the parameter τ and the half-space T_i are defined in the same way as in the SEGM, $\{\alpha_i\}$ is a sequence in (0,1) satisfying $\lim_{i\to\infty} \alpha_i = 0$ and $\sum_{i=1}^{\infty} \alpha_i = \infty$. By choosing appropriate values for the parameters, they obtained a strong convergence theorem for an iterative method for solving VIP in Hilbert space.

On the other hand, it is known that inertial technique can effectively accelerate the iterative process of algorithms. Therefore, authors added inertia term to the algorithm for solving VIP. For example, Thong and Hieu (2018) combined the inertial technique with the SEGM and thus obtained the following approach:

$$\begin{cases} w_i = x_i + \alpha_i(x_i - x_{i-1}), \\ t_i = \mathscr{P}_C(w_i - \tau \mathscr{A} w_i), \\ x_{i+1} = \mathscr{P}_{T_i}(w_i - \tau \mathscr{A} t_i), \end{cases} \tag{ISEGM}$$

where $T_i := \{x \in \mathscr{H} : \langle w_i - \tau \mathscr{A} w_i - t_i, x - t_i \rangle \le 0\}$, $\{\alpha_i\}$ is non-decreasing sequence and $0 \le \alpha_i \le \alpha \le \sqrt{5} - 2$, $\tau L \le \frac{\frac{1}{2} - 2\alpha - \frac{1}{2}\alpha^2 - \delta}{\frac{1}{2} - \alpha + \frac{1}{2}\alpha^2}$ for some $0 < \delta < \frac{1}{2} - 2\alpha - \frac{1}{2}\alpha^2$. With a suitable choice of parameters, they proved that the sequence generated by the algorithm weakly converges to an element of the solution set $\mathcal{S}$ and that the operator $\mathscr{A}$ involved is monotone and L-Lipschitz continuous. Of course, the algorithm proposed by Thong and Hieu (2018) can be combined with the Halpern method, based on the ideas mentioned above, to obtain the corresponding strong convergence theorem.

Since Banach spaces have more general properties than Hilbert spaces, some authors have solved VIP in certain Banach spaces using tools based on existing results in Hilbert spaces. For example, Cai et al. (2018) combined the SEGM with the Halpern method and solved variational inequalities in 2-uniformly convex Banach spaces. The algorithm is iterated as follows:

$$\begin{cases} t_i = \Pi_C(J w_i - \tau \mathscr{A} w_i), \\ s_i = \Pi_{T_i}(J w_i - \tau \mathscr{A} t_i), \\ w_{i+1} = J^{-1}(\alpha_i J w_1 + (1-\alpha_i) J s_i), \end{cases} \tag{BSEGM}$$

where $\Pi_C : \mathscr{B}^* \to \mathcal{C} \subset \mathscr{B}$ is the generalized projection operator, $J : \mathscr{B} \to 2^{\mathscr{B}^*}$ is the normalized duality mapping and $\mathscr{B}$ is a real 2-uniformly convex Banach space with dual $\mathscr{B}^*$. The step size τ satisfies $0 < \tau < \frac{1}{\mu L}$, where $\mu \ge 1$ is the 2-uniform convexity constant of $\mathscr{B}$ and L is the Lipschitz constant of $\mathscr{A}$. $\{\alpha_i\}$ is a sequence in (0,1) satisfying $\lim_{i\to\infty} \alpha_i = 0$ and $\sum_{i=1}^{\infty} \alpha_i = \infty$. Cai et al. (2018) obtained a strong convergence theorem by choosing the appropriate parameters.

Furthermore, some authors have recently considered solving VIP in reflexive Banach spaces, see for instance, Jolaoso and Shehu (2022), Jolaoso et al. (2022), Oyewole et al. (2022), Reich et al. (2021), Abass et al. (2022) and the references therein. For example, Jolaoso and Shehu (2022) generalized the classical Tseng's extragradient method proposed by Tseng (2000) to reflexive Banach spaces using Bregman projection, thus obtaining strong and weak convergence theorems under different conditions, respectively. In addition, Jolaoso et al. (2022) combined Popov's method Popov (1980) with the SEGM to obtain that the sequence generated by the method converges to an element of $\mathcal{S}$ in reflexive Banach spaces by selecting appropriate parameters with the help of Bregman projection. In Reich et al. (2021), Reich et al. applied the inertial technique to the reflexive Banach space based on the hybrid and shrinking projection method and Tseng's extragradient method and obtained a strong convergence theorem under reasonable assumptions. These results have a common feature in that they all achieve a generalization of the known results from Hilbert spaces to reflexive Banach spaces.

Motivated by the above works, in this paper, we propose a new inertial Bregman projection method for solving VIP in real reflexive Banach space. The modifications are as follows:

- Reich et al. (2021) successfully combined the inertial technique with the Tseng's extragradient method and applied it to reflexive Banach spaces. Based on this, we found that adding inertial terms to the SEGM can also be achieved.
- In contrast to the above mentioned methods, we obtain a strong convergence theorem in reflexive Banach spaces by Halpern-type iteration and the concepts of Bregman distance and Bregman projection under suitable conditions. It is well known that strong convergence can be used to infer weak convergence, but the reverse is not necessarily true.
- We modify the step size parameter based on the SEGM. We set two different constant parameters to control the step size, which allows the algorithm to improve the convergence process of the iterative sequence without the restriction that the two parameters must be equal.

2 Preliminaries

This section collects several background material to facilitate the study that follows.

Let $f : \mathscr{B} \to \mathbb{R}$ be a proper convex and lower semicontinuous function. For simplicity, dom f denotes the domain of f, dom $f := \{s \in \mathscr{B} : f(s) < \infty\}$. If $s \in \text{int}(\text{dom } f)$, then

(i) the subdifferential of f at s is the convex set given by

$$\partial f(s) := \{s^* \in \mathscr{B}^* : f(s) + \langle t - s, s^* \rangle \leq f(t),\ \forall t \in \mathscr{B}\}. \tag{2}$$

(ii) the Fenchel conjugate of f is the convex function $f^* : \mathscr{B}^* \to \mathbb{R}$ with

$$f^*(s^*) := \sup\{\langle s, s^* \rangle - f(s) : s \in \mathscr{B}\}.$$

(iii) the directional derivative of f at s is defined as

$$f^\circ(s, t) := \lim_{n \to 0} \frac{f(s + nt) - f(s)}{n},\ \forall t \in \mathscr{B}. \tag{3}$$

(iv) f is called Gâteaux differentiable at s, if the limit of (3) exists. At this time, the gradient of f at s is the linear function $\nabla f(s)$ satisfying

$$\langle \nabla f(s), t \rangle := f^\circ(s, t),\ \forall t \in \mathscr{B}.$$

(v) f is said to Fréchet differentiable at s, if $n \to 0$ in (3) is attained uniformly for any $\|t\| = 1$. f is called uniformly Fréchet differentiable on a subset $\mathcal{C}$, if $n \to 0$ in (3) is attained uniformly for any $\|t\| = 1$ and $s \in \mathcal{C} \subset \mathscr{B}$.

The Banach space $\mathscr{B}$ is said to be reflexive, if $\mathscr{J}(\mathscr{B}) = \mathscr{B}^{**}$, where $\mathscr{J} : \mathscr{B} \to \mathscr{B}^{**}$ is the standard embedding operator. Simultaneously, f is said to be Legendre if and only if it has the following forms:

(P1) f is Gâteaux differentiable, dom∇f=int(dom f), int(dom f)$\neq \emptyset$,
(P2) f^* is Gâteaux differentiable, dom∇f^*=int(dom f^*), int(dom f^*)$\neq \emptyset$.

The reflexivity of $\mathscr{B}$ yields $(\partial f)^{-1} = \partial f^*$. And combining (P1) and (P2) shows that $\nabla f = (\nabla f^*)^{-1}$, ran$\nabla f$ = dom∇f^* = int(domf^*), ran∇f^* = dom∇f = int(dom$f$). Moreover, in the interior of their respective domains, f and f^* are strictly convex and f is Legendre if and only if f^* is Legendre.

Assume that f is Gâteaux differentiable, the Bregman distance associated to f is the function $\mathcal{D}_f$:dom$f\times$int(dom$f$)$\to [0, +\infty)$ given by

$$\mathcal{D}_f(t, s) := f(t) - f(s) - \langle \nabla f(s), t - s\rangle.$$

One readily observes the following properties concerning to $\mathcal{D}_f$:

(i) three point identity:

$$\mathcal{D}_f(t, w) + \mathcal{D}_f(w, s) - \mathcal{D}_f(t, s) = \langle \nabla f(w) - \nabla f(s), w - t\rangle,$$

(ii) four point identity:

$$\mathcal{D}_f(w, x) + \mathcal{D}_f(t, s) - \mathcal{D}_f(w, s) - \mathcal{D}_f(t, x) = \langle \nabla f(s) - \nabla f(x), w - t\rangle,$$

for any $s, t, w, x \in \mathscr{B}$.

We say that a Gâteaux differentiable function f belongs to β-strongly convex if

$$\langle \nabla f(s) - \nabla f(t), s - t\rangle \geq \beta\|s - t\|^2, \ \forall s, t \in \text{dom } f,$$

namely,

$$f(t) \geq f(s) + \langle \nabla f(s), t - s\rangle + \frac{\beta}{2}\|s - t\|^2, \ \forall s, t \in \text{dom } f.$$

This gives that

$$\mathcal{D}_f(s, t) \geq \frac{\beta}{2}\|s - t\|^2 \tag{4}$$

for every $s \in$ dom f, $t \in$int(dom f).

The modulus of total convexity at $s \in$ int(dom f) denoted $v_f(s, \cdot) : [0, +\infty) \to [0, +\infty]$ is given by

$$v_f(s, w) := \inf\{\mathcal{D}_f(t, s) : t \in \text{dom} f, \|t - s\| = w\}.$$

We say that f is totally convex at $u \in$ int(dom f), if $v_f(s, w) > 0, \ \forall w > 0$. Furthermore, the modulus of total convexity of f on nonempty subset $\mathcal{C} \subset \mathscr{B}$ is given by

$$v_f(\mathcal{C}, w) := \inf\{v_f(s, w) : s \in \mathcal{C} \cap \text{int(dom } f)\}.$$

For given bounded subset $\mathcal{C}$ and $w > 0$, the hypothesis of $v_f(\mathcal{C}, w) > 0$ yields that f is totally convex. Moreover, strongly convex function can derive totally convexity. Specially, when f is a Legendre function, we know that totally convexity is consistent with uniformly convexity of f. Besides, f belongs to strongly coercive if $\lim_{\|s\|\to\infty} \dfrac{f(s)}{\|s\|} = +\infty$.

Given a nonempty closed convex subset $\mathcal{C} \subset \mathscr{B}$, the operator $\mathscr{F}$ is

(i) monotone on $\mathcal{C}$, if

$$\langle \mathscr{F}s - \mathscr{F}t, s - t\rangle \geq 0, \ \forall s, t \in \mathcal{C}.$$

(ii) pseudomonotone on $\mathcal{C}$, if

$$\langle \mathscr{F}s, t-s\rangle \geq 0 \Rightarrow \langle \mathscr{F}t, t-s\rangle \geq 0,\ \forall s,t\in\mathcal{C}.$$

(iii) L-Lipschitz continuous on $\mathcal{C}$, if there is an absolute constant L with

$$\|\mathscr{F}s-\mathscr{F}t\| \leq L\|s-t\|,\ \forall s,t\in\mathcal{C}.$$

Assume that $B_r := \{t\in\mathscr{B} : \|t\| < r, r > 0\}$, $S_{\mathscr{B}} := \{t\in\mathscr{B} : \|t\| = 1\}$ and $\rho_r : [0,+\infty)\to[0,+\infty]$ is the gauge of uniformly convexity of f,

$$\rho_r := \inf_{x,y\in B_r, \|x-y\|=t, \alpha\in(0,1)} \frac{\alpha f(x)+(1-\alpha)f(y)-f(\alpha x+(1-\alpha)y)}{\alpha(1-\alpha)}.$$

Then $f : \mathscr{B}\to\mathbb{R}$ is uniformly convex on bounded subsets of $\mathscr{B}$ if $\rho_r(t) > 0,\ \forall r,t > 0$. It is known that a strongly convex function is uniformly convex.

We will need several lemmas to obtain our main results.

Lemma 1 (Naraghirad and Yao 2013) *If f is uniformly convex on bounded subsets of $\mathscr{B}$, then*

$$f\left(\sum_{i=0}^{k}\alpha_i t_i\right) \leq \sum_{i=0}^{k}\alpha_i f(t_i) - \alpha_i\alpha_j\rho_r(\|t_i-t_j\|),$$

for all $r > 0$, $i,j\in\{0,1,2,\ldots,k\}$, $t_i\in B_r$, $\alpha_i\in(0,1)$ with $\sum_{i=0}^{k}\alpha_i = 1$, where ρ_r is the gauge of uniform convexity of f.

Lemma 2 (Butnariu and Iusem 2000) *The function f is totally convex on bounded subsets of $\mathscr{B}$ if and only if, for $\{t_i\}\subset$ int(dom f) and $\{s_i\}\subset$ dom f, such that the first one is bounded,*

$$\lim_{i\to\infty}\mathcal{D}_f(s_i,t_i) = 0 \Rightarrow \lim_{n\to\infty}\|s_i-t_i\| = 0.$$

Lemma 3 (Reich and Sabach 2009) *If f is convex, bounded and uniformly Fréchet differentiable on bounded subsets of $\mathscr{B}$, then ∇f is uniformly continuous on bounded subsets of $\mathscr{B}$ from the strong topology of $\mathscr{B}$ to the strong topology of $\mathscr{B}^*$.*

Lemma 4 (Zălinescu 2002) *If f is convex and bounded on bounded subsets of $\mathscr{B}$. Then the following are equivalent:*

1. *f is strongly coercive and uniformly convex on bounded subsets of $\mathscr{B}$.*
2. *dom $f^* = \mathscr{B}^*$, f^* is bounded on bounded subsets and uniformly smooth on bounded subsets of $\mathscr{B}^*$.*
3. *dom $f^* = \mathscr{B}^*$, f^* is Fréchet differentiable and ∇f^* is uniformly norm-to-norm continuous on bounded subsets of $\mathscr{B}^*$.*

Lemma 5 (Butnariu and Iusem 2000) *If f is strongly coercive, then*

1. *$\nabla f : \mathscr{B}\to\mathscr{B}^*$ is one-to-one, onto and norm-to-weak* continuous.*
2. *$\{t\in\mathscr{B} : \mathcal{D}_f(t,s)\leq r\}$ is bounded for all $s\in\mathscr{B}$ and $r > 0$.*
3. *dom $f^* = \mathscr{B}^*$, f^* is Gâteaux differentiable and $\nabla f^* = (\nabla f)^{-1}$.*

Lemma 6 (Martín-Márquez et al. 2013) *Let f be Gâteaux differentiable on int(dom f) such that ∇f^* is bounded on bounded subsets of dom f^*. Let $t_0\in\mathscr{B}$ and $\{t_i\}\subset$ int($\mathscr{B}$), if $\{\mathcal{D}_f(t_0,t_i)\}$ is bounded, then the sequence $\{t_i\}$ is bounded too.*

Lemma 7 (Butnariu and Iusem 2000) *Let $\mathcal{C}$ be a nonempty closed convex subset of a reflexive Banach space $\mathscr{B}$. A Bregman projection of $t \in int(dom\ f)$ onto $\mathcal{C} \subset int(dom\ f)$ is the unique vector $Proj_{\mathcal{C}}^{f}(t) \in \mathcal{C}$ which satisfies*

$$\mathcal{D}_f(Proj_{\mathcal{C}}^{f}(t), t) = inf\{\mathcal{D}_f(s, t) : s \in \mathcal{C}\}.$$

Lemma 8 (Alber 1996; Censor and Lent 1981; Phelps 1993) *Let $\mathcal{C}$ be a nonempty closed convex subset of a reflexive Banach space $\mathscr{B}$ and $t \in \mathscr{B}$. Let f be Gâteaux differentiable and totally convex function. Then*

1. $p = Proj_{\mathcal{C}}^{f}(t) \Leftrightarrow \langle \nabla f(t) - \nabla f(p), s - p\rangle \le 0,\ \forall s \in \mathcal{C}$.
2. $\mathcal{D}_f(s, Proj_{\mathcal{C}}^{f}(t)) + \mathcal{D}_f(Proj_{\mathcal{C}}^{f}(t), t) \le \mathcal{D}_f(s, t),\ \forall s \in \mathcal{C}$.

Define the bifunction $V_f : \mathscr{B} \times \mathscr{B}^* \to [0, +\infty)$ by

$$V_f(t, t^*) := f(t) - \langle t, t^*\rangle + f^*(t^*),\ \forall t \in \mathscr{B}, t^* \in \mathscr{B}^*.$$

Then

$$V_f(t, t^*) = \mathcal{D}_f(t, \nabla f^*(t^*)),\ \forall t \in \mathscr{B}, t^* \in \mathscr{B}^*, \tag{5}$$

and

$$V_f(t, t^*) + \langle \nabla f^*(t^*) - t, s^*\rangle \le V_f(t, t^* + s^*),\ \forall t \in \mathscr{B}, t^*, s^* \in \mathscr{B}^*. \tag{6}$$

In addition, if f is a proper lower semicontinuous function, then f^* is a proper weak* lower semicontinuous and convex function. Hence, V_f is convex in the second variable. And

$$\mathcal{D}_f(t, \nabla f^*(\sum_{i=1}^{N} s_i \nabla f(t_i))) \le \sum_{i=1}^{N} s_i \mathcal{D}_f(t, t_i) \tag{7}$$

for all $x \in \mathscr{B}$, where $\{t_i\}_{i=1}^{N} \subset \mathscr{B}$ and $\{s_i\}_{i=1}^{N} \subset (0, 1)$ with $\sum_{i=1}^{N} s_i = 1$.

Lemma 9 (Maingé 2008) *Let $\{w_i\}$ be a sequence of non-negative real number. If there is a subsequence $\{w_{i_j}\}$ of $\{w_i\}$ satisfies $w_{i_j} < w_{i_j+1}$ for all $j \in \mathbb{N}$, then there exists a non-decreasing sequence $\{m_k\} \subset \mathbb{N}$ such that $\lim_{k\to\infty} m_k = \infty$ and for all (sufficiently large) number $k \in \mathbb{N}$,*

$$w_{m_k} \le w_{m_k+1}, \quad w_k \le w_{m_k+1}.$$

In fact, $m_k := \max\{i \le k : w_i < w_{i+1}\}$.

Lemma 10 (Xu 2002) *Let $\{a_i\}, \{b_i\}, \{c_i\}, \{\alpha_i\}$ and $\{\beta_i\}$ be sequences of non-negative real numbers such that $\{\alpha_i\} \subset (0, 1), \sum_{i=1}^{\infty} \alpha_i = \infty, \limsup_{i\to\infty} b_i \le 0, \{\beta_i\} \subset [0, \frac{1}{2}], \sum_{i=1}^{\infty} c_i < \infty$, for $i \ge 1$,*

$$a_{i+1} \le (1 - \alpha_i - \beta_i)a_i + \beta_i a_{i-1} + \alpha_i b_i + c_i.$$

Then $\lim_{i\to\infty} a_i = 0$.

Lemma 11 (Mashreghi and Nasri 2010) *If the mapping $h : [0, 1] \to \mathscr{B}^*$ defined as $h(z) = \mathscr{F}(zs + (1 - z)t)$ is continuous for all $s, t \in \mathcal{C}$, then $\mathscr{M}(\mathcal{C}, \mathscr{F}) := \{s^* \in \mathcal{C} : \langle \mathscr{F}t, t - s^*\rangle, \forall t \in \mathcal{C}\} \subset \mathcal{S}$. Moreover, if $\mathscr{F}$ is pseudomonotone, then $\mathscr{M}(\mathcal{C}, \mathscr{F})$ is closed, convex and $\mathscr{M}(\mathcal{C}, \mathscr{F}) = \mathcal{S}$.*

3 Main results

In this section, we propose an inertial subgradient extragradient method with Bregman distance for solving pseudomonotone variational inequality problems in reflexive Banach spaces. First, we give the following assumptions:

(C1) The feasible set $\mathcal{C}$ is a nonempty closed convex subset of real reflexive Banach space $\mathscr{B}$. The proper lower semicontinuous function $f : \mathscr{B} \to \mathbb{R}$ is strongly coercive Legendre which is bounded, uniformly Fréchet differentiable and β-strongly convex on bounded subsets of $\mathscr{B}$.

(C2) The operator $\mathscr{F} : \mathscr{B} \to \mathscr{B}^*$ is pseudomonotone, L-Lipschitz continuous and satisfies the following condition:

$$\{q_i\} \subset \mathcal{C}, q_i \rightharpoonup q \Rightarrow \|\mathscr{F}q\| \leq \liminf_{i\to\infty} \|\mathscr{F}q_i\|. \tag{8}$$

The solution set $\mathcal{S}$ is nonempty.

(C3) Let $\{\epsilon_i\}$ be a positive sequence such that $\lim_{i\to\infty} \frac{\epsilon_i}{\alpha_i} = 0$, where $\{\alpha_i\} \subset (0, 1)$, $\lim_{i\to\infty} \alpha_i = 0$ and $\sum_{i=1}^{\infty} \alpha_i = \infty$.

Now, we introduce the following algorithm.

Algorithm 3

Initialization: Let $\theta \in (0, \frac{1}{2})$, $\tau \in (0, \frac{\beta}{L})$, $\xi \in (0, \tau]$ and $w_0, w_1 \in \mathscr{B}$.
Iterative steps: Given the current iterates w_{i-1} and w_i $(i \geq 1)$.
Step 1. Set $x_i = \nabla f^*(\nabla f(w_i) + \theta_i(\nabla f(w_{i-1}) - \nabla f(w_i)))$, where

$$\theta_i = \begin{cases} \min\left\{\frac{\epsilon_i}{\|w_i - w_{i-1}\|}, \theta\right\}, & \text{if } w_i \neq w_{i-1}, \\ \theta, & \text{otherwise,} \end{cases}$$

and evaluate

$$t_i = Proj_{\mathcal{C}}^f\big(\nabla f^*(\nabla f(x_i) - \tau\mathscr{F}x_i)\big).$$

If $t_i = x_i$ or $\mathscr{F}t_i = 0$, then stop. Otherwise go to **Step 2**.
Step 2. Compute

$$s_i = Proj_{T_i}^f\big(\nabla f^*(\nabla f(x_i) - \xi\mathscr{F}t_i)\big),$$

where

$$T_i := \{x \in \mathscr{B} : \langle \nabla f(x_i) - \tau\mathscr{F}x_i - \nabla f(t_i), x - t_i\rangle \leq 0\}.$$

Step 3. Calculate

$$w_{i+1} = \nabla f^*\big(\alpha_i \nabla f(w_1) + (1 - \alpha_i)\nabla f(s_i)\big).$$

Set $i := i + 1$ and return to **Step 1**.

Remark 1 : 1. Note that $\lim_{i\to\infty}\frac{\theta_i}{\alpha_i}\|w_i-w_{i-1}\|=0$. Indeed, for all i we have $\theta_i\le\frac{\epsilon_i}{\|w_i-w_{i-1}\|}$, it follows from (C3) that

$$\lim_{i\to\infty}\frac{\theta_i}{\alpha_i}\|w_i-w_{i-1}\|\le\lim_{i\to\infty}\frac{\epsilon_i}{\alpha_i}=0.$$

2. If $t_i=x_i$, then $t_i\in\mathcal{S}$. Indeed, the definition of $\{t_i\}$ and Lemma 8(1) show that

$$\langle\nabla f(x_i)-\tau\mathscr{F}x_i-\nabla f(t_i),y-t_i\rangle\le 0,\ \forall y\in\mathcal{C}.$$

Due to $t_i=x_i$, we get that

$$\tau\langle\mathscr{F}t_i,y-t_i\rangle\ge 0,\ \forall y\in\mathcal{C}.$$

Since $\tau>0$, then $t_i\in\mathcal{S}$. If $\mathscr{F}t_i=0$, it is easy to see that $t_i\in\mathcal{S}$.

Lemma 12 *Assume that the conditions (C1–C3) hold. Let $\{t_i\}$, $\{s_i\}$ and $\{x_i\}$ be the sequences produced by Algorithm 3. Then*

$$\begin{aligned}\mathcal{D}_f(p,s_i)\le\mathcal{D}_f(p,x_i)&-\left(1-\frac{\xi}{\tau}\right)\mathcal{D}_f(s_i,x_i)-\frac{\xi}{\tau}\left(1-\frac{\tau L}{\beta}\right)\mathcal{D}_f(t_i,x_i)\\&-\frac{\xi}{\tau}\left(1-\frac{\tau L}{\beta}\right)\mathcal{D}_f(s_i,t_i)\end{aligned}$$

for every $p\in\mathcal{S}$.

Proof Bearing in mind Lemma 8(1), the definition of $\{s_i\}$ yields that

$$\langle\nabla f(x_i)-\xi\mathscr{F}t_i-\nabla f(s_i),y-s_i\rangle\le 0,\ \forall y\in T_i.$$

Due to $p\in\mathcal{S}\subset T_i$, we deduce that

$$\langle\nabla f(x_i)-\xi\mathscr{F}t_i-\nabla f(s_i),p-s_i\rangle\le 0,$$

namely

$$\langle\nabla f(x_i)-\nabla f(s_i),p-s_i\rangle\le\xi\langle\mathscr{F}t_i,p-s_i\rangle.\tag{9}$$

In light of the three point identity, we see that

$$\mathcal{D}_f(p,s_i)+\mathcal{D}_f(s_i,x_i)-\mathcal{D}_f(p,x_i)=\langle\nabla f(x_i)-\nabla f(s_i),p-s_i\rangle.\tag{10}$$

Together (9) with (10),

$$\begin{aligned}\mathcal{D}_f(p,s_i)&\le\mathcal{D}_f(p,x_i)-\mathcal{D}_f(s_i,x_i)+\xi\langle\mathscr{F}t_i,p-s_i\rangle\\&=\mathcal{D}_f(p,x_i)-\mathcal{D}_f(s_i,x_i)+\xi\langle\mathscr{F}t_i,p-t_i\rangle+\xi\langle\mathscr{F}t_i,t_i-s_i\rangle.\end{aligned}\tag{11}$$

Since $t_i\in\mathcal{C}$ and $p\in\mathcal{S}$, we have $\langle\mathscr{F}p,t_i-p\rangle\ge 0$. The pseudomonotonicity of $\mathscr{F}$ implies $\langle\mathscr{F}t_i,t_i-p\rangle\ge 0$. Thus (11) can be transformed into

$$\mathcal{D}_f(p,s_i)\le\mathcal{D}_f(p,x_i)-\mathcal{D}_f(s_i,x_i)+\xi\langle\mathscr{F}t_i,t_i-s_i\rangle.\tag{12}$$

Next we estimate $\xi\langle\mathscr{F}t_i,t_i-s_i\rangle$. The three point identity shows that

$$\begin{aligned}&\mathcal{D}_f(s_i,x_i)-\tau\langle\mathscr{F}t_i,t_i-s_i\rangle\\&\quad=\mathcal{D}_f(s_i,t_i)+\mathcal{D}_f(t_i,x_i)-\langle\nabla f(x_i)-\nabla f(t_i)-\tau\mathscr{F}t_i,s_i-t_i\rangle\\&\quad=\mathcal{D}_f(s_i,t_i)+\mathcal{D}_f(t_i,x_i)-\langle\nabla f(x_i)-\tau\mathscr{F}x_i-\nabla f(t_i),s_i-t_i\rangle\\&\qquad-\tau\langle\mathscr{F}x_i-\mathscr{F}t_i,s_i-t_i\rangle.\end{aligned}\tag{13}$$

Applying the definition of T_i and $s_i \in T_i$, one readily observes

$$\langle \nabla f(x_i) - \tau \mathscr{F} x_i - \nabla f(t_i), s_i - t_i \rangle \leq 0. \tag{14}$$

It follows from (4) and (8) that

$$\begin{aligned}
\tau \langle \mathscr{F} x_i - \mathscr{F} t_i, s_i - t_i \rangle &\leq \tau \|\mathscr{F} x_i - \mathscr{F} t_i\| \|s_i - t_i\| \\
&\leq \tau L \|x_i - t_i\| \|s_i - t_i\| \\
&\leq \frac{\tau L}{2} \big(\|x_i - t_i\|^2 + \|s_i - t_i\|^2 \big) \\
&\leq \frac{\tau L}{\beta} \big(\mathcal{D}_f(t_i, x_i) + \mathcal{D}_f(s_i, t_i) \big).
\end{aligned} \tag{15}$$

Combining (13), (14) and (15), then

$$\tau \langle \mathscr{F} t_i, t_i - s_i \rangle \leq \mathcal{D}_f(s_i, x_i) - \left(1 - \frac{\tau L}{\beta}\right) \Big(\mathcal{D}_f(t_i, x_i) + \mathcal{D}_f(s_i, t_i) \Big). \tag{16}$$

Substituting (16) into (12), we obtain

$$\begin{aligned}
\mathcal{D}_f(p, s_i) &\leq \mathcal{D}_f(p, x_i) - \mathcal{D}_f(s_i, x_i) + \frac{\xi}{\tau} \mathcal{D}_f(s_i, x_i) \\
&\quad - \frac{\xi}{\tau} \left(1 - \frac{\tau L}{\beta}\right) \big(\mathcal{D}_f(t_i, x_i) + \mathcal{D}_f(s_i, t_i) \big) \\
&= \mathcal{D}_f(p, x_i) - \left(1 - \frac{\xi}{\tau}\right) \mathcal{D}_f(s_i, x_i) - \frac{\xi}{\tau} \left(1 - \frac{\tau L}{\beta}\right) \mathcal{D}_f(t_i, x_i) \\
&\quad - \frac{\xi}{\tau} \left(1 - \frac{\tau L}{\beta}\right) \mathcal{D}_f(s_i, t_i),
\end{aligned} \tag{17}$$

which is the desired inequality. □

Lemma 13 *Suppose the conditions (C1–C3) hold. The sequence $\{w_i\}$ produced by Algorithm 3 is bounded.*

Proof As $\tau \in \left(0, \frac{\beta}{L}\right)$ and $\xi \in (0, \tau]$, we see that

$$1 - \frac{\xi}{\tau} \geq 0, \ \frac{\xi}{\tau} \left(1 - \frac{\tau L}{\beta}\right) > 0. \tag{18}$$

With (18) and Lemma 12 in hand, we get

$$\mathcal{D}_f(p, s_i) \leq \mathcal{D}_f(p, x_i), \ \forall p \in \mathcal{S}. \tag{19}$$

The definition of $\{w_i\}$ and (7) show that

$$\begin{aligned}
\mathcal{D}_f(p, x_i) &= \mathcal{D}_f\Big(p, \nabla f^*\big(\nabla f(w_i) + \theta_i(\nabla f(w_{i-1}) - \nabla f(w_i))\big)\Big) \\
&= \mathcal{D}_f\Big(p, \nabla f^*\big((1 - \theta_i)\nabla f(w_i) + \theta_i(\nabla f(w_{i-1}))\big)\Big) \\
&\leq (1 - \theta_i)\mathcal{D}_f(p, w_i) + \theta_i \mathcal{D}_f(p, w_{i-1}).
\end{aligned}$$

Hence, with Lemma 1 in hand, the definition of V_f and the property of ρ_r^* yield that

$$\mathcal{D}_f(p, w_{i+1}) = \mathcal{D}_f\Big(p, \nabla f^*\big(\alpha_i \nabla f(w_1) + (1 - \alpha_i)\nabla f(s_i)\big)\Big)$$

$$\le \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\mathcal{D}_f(p, s_i)$$
$$\le \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\mathcal{D}_f(p, x_i)$$
$$\le \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\left[(1-\theta_i)\mathcal{D}_f(p, w_i) + \theta_i \mathcal{D}_f(p, w_{i-1})\right]$$
$$\le \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\max\left\{\mathcal{D}_f(p, w_i), \mathcal{D}_f(p, w_{i-1})\right\}$$
$$\le \max\left\{\mathcal{D}_f(p, w_1), \mathcal{D}_f(p, w_i), \mathcal{D}_f(p, w_{i-1})\right\}$$

By induction, we conclude that

$$\mathcal{D}_f(p, w_{i+1}) \le \max\left\{\mathcal{D}_f(p, w_1), \mathcal{D}_f(p, w_0)\right\}. \tag{20}$$

Consequently, the sequence $\{w_i\}$ is bounded by Lemma 6. □

Lemma 14 *Assume that the conditions (C1-C3) hold. Let* $\{x_{i_k}\}$ *be a subsequence of* $\{x_i\}$ *produced by Algorithm 3 such that* $x_{i_k} \rightharpoonup q$ *and* $\lim_{k\to\infty}\|x_{i_k} - t_{i_k}\| = 0$*, then* $q \in \mathcal{S}$.

Proof Due to $t_{i_k} := Proj_{\mathcal{C}}^f\left(\nabla f^*(\nabla f(x_{i_k}) - \tau \mathscr{F} x_{i_k})\right)$, it follows from Lemma 8(1) that

$$\langle \nabla f(x_{i_k}) - \tau \mathscr{F} x_{i_k} - \nabla f(t_{i_k}), x - t_{i_k}\rangle \le 0, \ \forall x \in \mathcal{C},$$

that is

$$\frac{1}{\tau}\langle \nabla f(x_{i_k}) - \nabla f(t_{i_k}), x - t_{i_k}\rangle \le \langle \mathscr{F} x_{i_k}, x - t_{i_k}\rangle, \ \forall x \in \mathcal{C}.$$

Hence,

$$\frac{1}{\tau}\langle \nabla f(x_{i_k}) - \nabla f(t_{i_k}), x - t_{i_k}\rangle + \langle \mathscr{F} x_{i_k}, t_{i_k} - x_{i_k}\rangle \le \langle \mathscr{F} x_{i_k}, x - x_{i_k}\rangle, \ \forall x \in \mathcal{C}. \tag{21}$$

Since $\lim\limits_{k\to\infty}\|x_{i_k} - t_{i_k}\| = 0$, Lemma 3 yields that ∇f is uniformly continuous, then

$$\lim_{k\to\infty}\|\nabla f(x_{i_k}) - \nabla f(t_{i_k})\| = 0.$$

Putting $k \to \infty$ in (21), $\tau > 0$ shows that

$$\liminf_{k\to\infty}\langle \mathscr{F} x_{i_k}, x - x_{i_k}\rangle \ge 0, \ \forall x \in \mathcal{C}. \tag{22}$$

We then conclude that

$$\begin{aligned}\langle \mathscr{F} t_{i_k}, x - t_{i_k}\rangle &= \langle \mathscr{F} t_{i_k} - \mathscr{F} x_{i_k}, x - x_{i_k}\rangle \\ &\quad + \langle \mathscr{F} x_{i_k}, x - x_{i_k}\rangle + \langle \mathscr{F} t_{i_k}, x_{i_k} - t_{i_k}\rangle.\end{aligned} \tag{23}$$

Due to $\lim\limits_{k\to\infty}\|x_{i_k} - t_{i_k}\| = 0$ and the fact that $\mathscr{F}$ is Lipschitz continuous, we get

$$\liminf_{k\to\infty}\langle \mathscr{F} t_{i_k}, x - t_{i_k}\rangle \ge 0. \tag{24}$$

We now proceed to the proof of $q \in \mathcal{S}$. Given a sequence $\{\epsilon_k\}$ of positive numbers that decreases and tends to 0, for each k, it follows from (24) that there is the smallest positive integer N_k such that

$$\langle \mathscr{F} t_{i_j}, x - t_{i_j}\rangle + \epsilon_k \ge 0, \ \forall j \ge N_k, \tag{25}$$

$\{\epsilon_k\}$ is decreasing implies that the sequence $\{N_k\}$ is increasing. Besides, for each k, there is a bounded sequence $\{v_{N_k}\} \subset \mathscr{B}$ such that $\epsilon_k = \langle \mathscr{F} t_{N_k}, \epsilon_k v_{N_k}\rangle$. Thus, we have

$$\langle \mathscr{F} t_{N_k}, x + \epsilon_k v_{N_k} - t_{N_k}\rangle \ge 0.$$

The pseudomonotonicity of $\mathscr{F}$ shows that

$$\langle \mathscr{F}(x+\epsilon_k v_{N_k}), x+\epsilon_k v_{N_k} - t_{N_k}\rangle \geq 0,$$

which implies that

$$\langle \mathscr{F}x, x - t_{N_k}\rangle \geq \langle \mathscr{F}x - \mathscr{F}(x+\epsilon_k v_{N_k}), x+\epsilon_k v_{N_k} - t_{N_k}\rangle - \langle \mathscr{F}x, \epsilon_k v_{N_k}\rangle. \tag{26}$$

In light of $x_{i_k} \rightharpoonup q$ and $\lim_{k\to\infty}\|x_{i_k} - t_{i_k}\| = 0$, we have $t_{i_k} \rightharpoonup q$ $(k \to \infty)$ and thus $q \in \mathcal{C}$. According to (8),

$$0 < \|\mathscr{F}q\| \leq \liminf_{k\to\infty} \left\|\mathscr{F}t_{i_k}\right\|,$$

which implies that $\lim_{k\to\infty}\epsilon_k v_{N_k} = 0$. Thus, letting $k \to \infty$ in (26), the Lipschitz continuity of $\mathscr{F}$ and the boundedness of $\{x_{N_k}\}$ and $\{v_{N_k}\}$ show that

$$\liminf_{k\to\infty}\langle \mathscr{F}x, x - t_{N_k}\rangle \geq 0. \tag{27}$$

Hence, for all $x \in \mathcal{C}$,

$$\langle \mathscr{F}x, x - q\rangle = \lim_{k\to\infty}\langle \mathscr{F}x, x - t_{N_k}\rangle = \liminf_{k\to\infty}\langle \mathscr{F}x, x - t_{N_k}\rangle \geq 0. \tag{28}$$

By Lemma 11, we see therefore that $q \in \mathcal{S}$. □

Theorem 1 *Assume that the conditions (C1–C3) hold. Then the sequence $\{w_i\}$ produced by Algorithm 3 strongly converges to $p = Proj_{\mathcal{S}}^f(w_1)$.*

Proof According to the definition of $\{w_i\}$, the formulas (5), (6) and (7) show that

$$\begin{aligned}
\mathcal{D}_f(p, w_{i+1}) &= \mathcal{D}_f\Big(p, \nabla f^*\big(\alpha_i \nabla f(w_1) + (1-\alpha_i)\nabla f(s_i)\big)\Big) \\
&= V_f\big(p, \alpha_i \nabla f(w_1) + (1-\alpha_i)\nabla f(s_i)\big) \\
&\leq V_f\big(p, \alpha_i \nabla f(w_1) + (1-\alpha_i)\nabla f(s_i) - \alpha_i(\nabla f(w_1) - \nabla f(p))\big) \\
&\quad + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \\
&= V_f\big(p, \alpha_i \nabla f(p) + (1-\alpha_i)\nabla f(s_i)\big) + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \\
&\leq \alpha_i \mathcal{D}_f(p, p) + (1-\alpha_i)\mathcal{D}_f(p, s_i) + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \\
&\leq (1-\alpha_i)\mathcal{D}_f(p, x_i) + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \\
&\leq (1-\alpha_i)\big[(1-\theta_i)\mathcal{D}_f(p, w_i) + \theta_i \mathcal{D}_f(p, w_{i-1})\big] \\
&\quad + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \\
&= \big[1 - \alpha_i - (1-\alpha_i)\theta_i\big]\mathcal{D}_f(p, w_i) + (1-\alpha_i)\theta_i \mathcal{D}_f(p, w_{i-1}) \\
&\quad + \alpha_i\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle.
\end{aligned} \tag{29}$$

We now proceed to the proof of $w_i \to p$ in two cases:

Case 1: There exists $N \in \mathbb{N}$ such that $\{\mathcal{D}_f(p, w_i)\}_{n=N}^{\infty}$ is nonincreasing. Thus $\{\mathcal{D}_f(p, w_i)\}$ is convergent and

$$\lim_{i\to\infty}\big(\mathcal{D}_f(p, w_i) - \mathcal{D}_f(p, w_{i+1})\big) = \lim_{i\to\infty}\big(\mathcal{D}_f(p, w_{i-1}) - \mathcal{D}_f(p, w_i)\big) = 0.$$

Moreover, it follows from Lemmas 12 and 13 that

$$\mathcal{D}_f(p, w_{i+1}) \leq \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\mathcal{D}_f(p, s_i)$$

$$\leq \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\big[\mathcal{D}_f(p, x_i) - (1-\frac{\xi}{\tau})\mathcal{D}_f(s_i, x_i) - \frac{\xi}{\tau}(1-\frac{\tau L}{\beta})\mathcal{D}_f(t_i, x_i) - \frac{\xi}{\tau}(1-\frac{\tau L}{\beta})\mathcal{D}_f(s_i, t_i)\big]$$
$$\leq \alpha_i \mathcal{D}_f(p, w_1) + (1-\alpha_i)\Big\{(1-\theta_i)\mathcal{D}_f(p, w_i) + \theta_i \mathcal{D}_f(p, w_{i-1}) - (1-\frac{\xi}{\tau})\mathcal{D}_f(s_i, x_i) - \frac{\xi}{\tau}(1-\frac{\tau L}{\beta})\big[\mathcal{D}_f(t_i, x_i) + \mathcal{D}_f(s_i, t_i)\big]\Big\}.$$

Namely

$$(1-\alpha_i)\frac{\xi}{\tau}\left(1-\frac{\tau L}{\beta}\right)\big[\mathcal{D}_f(t_i, x_i) + \mathcal{D}_f(s_i, t_i)\big] \leq \mathcal{D}_f(p, w_i) - \mathcal{D}_f(p, w_{i+1}) + (1-\alpha_i)\theta_n[\mathcal{D}_f(p, w_{i-1}) - \mathcal{D}_f(p, w_i)] + \alpha_n M_1, \quad (30)$$

for some $M_1 > 0$. Due to $\lim_{i\to\infty} \alpha_i = 0$, combining (30) and (18), we can conclude that

$$\lim_{i\to\infty} \mathcal{D}_f(t_i, x_i) = \lim_{i\to\infty} \mathcal{D}_f(s_i, t_i) = 0. \quad (31)$$

Owing to Lemma 2, we have

$$\lim_{i\to\infty} \|t_i - x_i\| = \lim_{i\to\infty} \|s_i - t_i\| = 0, \quad (32)$$

which implies that

$$\lim_{i\to\infty} \|s_i - x_i\| \leq \lim_{i\to\infty} \|s_i - t_i\| + \lim_{i\to\infty} \|t_i - x_i\| = 0. \quad (33)$$

At the same time, let $i \to \infty$, then

$$\|\nabla f(x_i) - \nabla f(w_i)\| = \alpha_i \frac{\theta_i}{\alpha_i}\|\nabla f(w_{i-1}) - \nabla f(w_i)\| = \alpha_i \frac{\theta_i}{\alpha_i}\|w_{i-1} - w_i\| \to 0,$$
$$\|\nabla f(w_{i+1}) - \nabla f(s_i)\| = \alpha_i \|\nabla f(w_1) - \nabla f(s_i)\| \leq \alpha_i M_2 \to 0,$$

for some $M_2 > 0$. From Lemma 4, the uniformly norm-to-norm continuity of ∇f^* implies that

$$\lim_{i\to\infty} \|x_i - w_i\| = \lim_{i\to\infty} \|w_{i+1} - s_i\| = 0. \quad (34)$$

Therefore, in light of (33) and (34),

$$\|w_{i+1} - w_i\| \leq \|w_{i+1} - s_i\| + \|s_i - x_i\| + \|x_i - w_i\| \to 0 \ (n \to \infty). \quad (35)$$

Since $\{w_i\}$ is bounded, there is a subsequence $\{w_{i_k}\} \subset \{w_i\}$ such that $w_{i_k} \rightharpoonup q$ and thus $x_{i_k} \rightharpoonup q$ as $k \to \infty$. Owing to Lemma 14, we can conclude that $q \in \mathcal{S}$. Combining with $p = Proj_{\mathcal{S}}^f(w_1)$, then

$$\begin{aligned}\limsup_{i\to\infty}\langle \nabla f(w_1) - \nabla f(p), w_i - p\rangle &= \lim_{k\to\infty}\langle \nabla f(w_1) - \nabla f(p), w_{i_k} - p\rangle \\ &= \lim_{k\to\infty}\langle \nabla f(w_1) - \nabla f(p), q - p\rangle \\ &\leq 0, \end{aligned} \quad (36)$$

which means that

$$\limsup_{i\to\infty}\langle \nabla f(w_1) - \nabla f(p), w_{i+1} - p\rangle \leq 0. \quad (37)$$

Combining (29), (37) with Lemma 10, we see therefore that $\mathcal{D}_f(p, w_i) \to 0$ as $i \to \infty$, that is $w_i \to p$ as $i \to \infty$.

Case 2: There is a subsequence $\{\mathcal{D}_f(p, w_{i_j})\} \subset \{\mathcal{D}_f(p, w_i)\}$ such that $\mathcal{D}_f(p, w_{i_j}) < \mathcal{D}_f(p, w_{i_j+1})$ for all $j \in \mathbb{N}$. It follows from Lemma 9 that there exists a non-decreasing sequence $\{i_k\} \subset \mathbb{N}$ tending to infinity such that

$$\mathcal{D}_f(p, w_{i_k}) \leq \mathcal{D}_f(p, w_{i_k+1}) \text{ and } \mathcal{D}_f(p, w_k) \leq \mathcal{D}_f(p, w_{i_k+1}),\ \forall k \in \mathbb{N}.$$

In Lemma 13 we obtain

$$\mathcal{D}_f(p, w_{i_k+1}) \leq \alpha_{i_k}\mathcal{D}_f(p, w_1) + (1-\alpha_{i_k})\mathcal{D}_f(p, w_{i_k}).$$

Therefore, the fact $\lim_{i\to\infty} \alpha_i = 0$ implies

$$\mathcal{D}_f(p, w_{i_k+1}) - \mathcal{D}_f(p, w_{i_k}) \to 0\ (k \to \infty).$$

Due to boundedness of $\{w_{i_k}\}$, there is a subsequence of $\{w_{i_k}\}$ still denoted by $\{w_{i_k}\}$ and $w_{i_k} \rightharpoonup q$. As stated in **Case 1**, we have

$$\lim_{k\to\infty} \|w_{i_k+1} - w_{i_k}\| = 0, \quad \limsup_{k\to\infty} \langle \nabla f(w_1) - \nabla f(p), w_{i_k+1} - p\rangle \leq 0.$$

According to (29),

$$\begin{aligned}\mathcal{D}_f(p, w_{i_k+1}) &\leq (1-\alpha_{i_k})\mathcal{D}_f(p, w_{i_k}) + \alpha_{i_k}\langle \nabla f(w_1) - \nabla f(p), w_{i_k+1} - p\rangle \\ &\leq (1-\alpha_{i_k})\mathcal{D}_f(p, w_{i_k+1}) + \alpha_{i_k}\langle \nabla f(w_1) - \nabla f(p), w_{i_k+1} - p\rangle.\end{aligned}$$

The fact $\alpha_{i_k} > 0$ shows that

$$\mathcal{D}_f(p, w_k) \leq \mathcal{D}_f(p, w_{i_k+1}) \leq \langle \nabla f(w_1) - \nabla f(p), w_{i_k+1} - p\rangle.$$

Namely,

$$\limsup_{k\to\infty} \mathcal{D}_f(p, w_k) \leq \limsup_{k\to\infty} \langle \nabla f(w_1) - \nabla f(p), w_{i_k+1} - p\rangle \leq 0.$$

Hence, $w_k \to p$ $(k \to \infty)$, which is the desired result. □

Next we apply the above result to Hilbert space. Let $f(s) = \frac{1}{2}\|s\|^2$, we have $\mathcal{D}_f(s,t) := \dfrac{1}{2}\|s-t\|^2,\ \forall s, t \in \mathcal{C}$, then

$$Proj_{\mathcal{C}}^f\big(\nabla f^*(\nabla f(x_i) - \tau \mathscr{F} x_i)\big) =: \mathscr{P}_{\mathcal{C}}(x_i - \tau \mathscr{A} x_i),$$

where $\mathcal{C}$ is nonempty closed convex subset of Hilbert space $\mathscr{H}$, $\mathscr{P}_{\mathcal{C}}$ is the metric projection of $\mathscr{H}$ onto $\mathcal{C}$. In this case, if $\theta_i = 0$, we can obtain the following result.

Corollary 1 *Assume that the feasible set $\mathcal{C}$ is a nonempty closed convex subset of a real Hilbert space $\mathscr{H}$, the solution set $\mathcal{S}$ is nonempty and the operator $\mathscr{A} : \mathscr{H} \to \mathscr{H}$ is pseudomonotone, L-Lipschitz continuous and satisfies $\{q_i\} \subset \mathcal{C}, q_i \rightharpoonup q \Rightarrow \|\mathscr{A} q\| \leq \liminf_{i\to\infty} \|\mathscr{A} q_i\|$. Suppose $\tau \in (0, \frac{1}{L})$, $\xi \in (0, \tau]$, $T_i := \{x \in \mathscr{H} : \langle x_i - \tau\mathscr{A} x_i - t_i, x - t_i\rangle \leq 0\}$, the sequence $\{\alpha_i\} \subset (0,1)$ satisfies $\lim_{i\to\infty} \alpha_i = 0$ and $\sum_{i=1}^{\infty} \alpha_i = \infty$. Then the sequence $\{x_i\}$ produced by the following algorithm strongly converges to an element $p = \mathscr{P}_{\mathcal{S}}(x_1)$.*

$$\begin{cases} t_i = \mathscr{P}_{\mathcal{C}}(x_i - \tau\mathscr{A} x_i), \\ s_i = \mathscr{P}_{T_i}(x_i - \xi\mathscr{A} t_i), \\ x_{i+1} = \alpha_i x_1 + (1-\alpha_i)s_i. \end{cases}$$

4 Numerical experiments

In this section, we perform two numerical examples to show the behaviors of Algorithm 3, and compare them with other algorithms. In both experiments the parameters are chosen as $\alpha_i = \frac{1}{i+1}$ and

$$\theta_i = \begin{cases} \min\left\{\frac{\epsilon_i}{\|w_i - w_{i-1}\|}, \theta\right\}, & \text{if } w_i \neq w_{i-1}, \\ \theta, & \text{otherwise,} \end{cases} \tag{38}$$

where $\epsilon_i = \frac{1}{(i+2)^{1.1}}$.

Example 1 Let $\mathscr{B} = L^2([0,1])$ endowed with norm $\|x\|_{\mathscr{B}} = (\int_0^1 |x(t)|^2 dt)^{\frac{1}{2}}$ and inner product $\langle x, y\rangle_{\mathscr{B}} = \int_0^1 x(t)y(t)dt$ for all $x, y \in \mathscr{B}$. Consider $\mathcal{C} := \{x \in \mathscr{B} : \|x\| \leq 2\}$. Let $g : \mathcal{C} \to \mathbb{R}$ be defined by

$$g(s) := \frac{1}{1 + \|s\|^2}.$$

Note that g is L_g-Lipchitz continuous with $L_g = \frac{16}{25}$ and $\frac{1}{5} \leq g(s) \leq 1, \quad \forall s \in \mathcal{C}$. The Volterra integral operator $V : \mathscr{B} \to \mathscr{B}$ is given by

$$V(s)(t) := \int_0^t s(x)dx, \quad \forall s \in L^2([0,1]), t \in [0,1].$$

Then V is bounded linear monotone operator. Define $\mathscr{F} : \mathcal{C} \to \mathscr{B}$ by

$$\mathscr{F}(s)(t) := g(s)V(s)(t), \quad \forall s \in \mathcal{C}, t \in [0,1].$$

We see that $\mathscr{F}$ is not monotone on $\mathcal{C}$ but pseudomonotone and L-Lipschitz continuous with $L = 82/\pi$.

We compare our Algorithm 3 (shortly, **OUR**) with Algorithm 3.5 in Cai et al. (2018) (shortly, **CGIS**) and Algorithm 3 in Xie et al. (2023) (shortly, **XCD**). The parameters are set as follows:

Our Algorithm 3: $\tau = 0.99/L, \xi = 0.9/L, \theta = 0.63$ and $\alpha_i = \frac{1}{i+1}$;

Algorithm 3 in Xie et al. (2023): $\mu = 0.9, \ \tau_1 = 2$ and $\alpha_i = \frac{1}{i+1}$;

Algorithm 3.5 in Cai et al. (2018): $a = 0.0001, \lambda_i = a + \frac{i(\frac{1}{L}-1)}{i+1}$ and $\alpha_i = \frac{1}{i+1}$.

Let $x_0 = x_1 = e^{2t} + \sin(3t)$, we observe from Fig. 1 that our Algorithm 3 is better than Algorithm 3.5 in Cai et al. (2018) and Algorithm 3 in Xie et al. (2023).

Example 2 Compare the performance of Algorithm 3 at different Bregman distances. Define the feasible set $\mathcal{C}$ by

$$\mathcal{C} = \left\{x \in \mathbb{R}^m : x_i \geq 0, \sum_{i=1}^m x_i = 1\right\}$$

and define $f : \mathbb{R}^m \to \mathbb{R}$ by

$$(i) f(x) = \sum_{i=1}^m x_i \log(x_i), \quad (ii) f(x) = \frac{1}{2}\|x\|^2.$$

We see that f is strongly convex with $\beta = 1$. Due to $\mathscr{B} = \mathbb{R}^m$, then $\nabla f^*(x) = (\nabla f)^{-1}(x)$. Therefore, for each f,

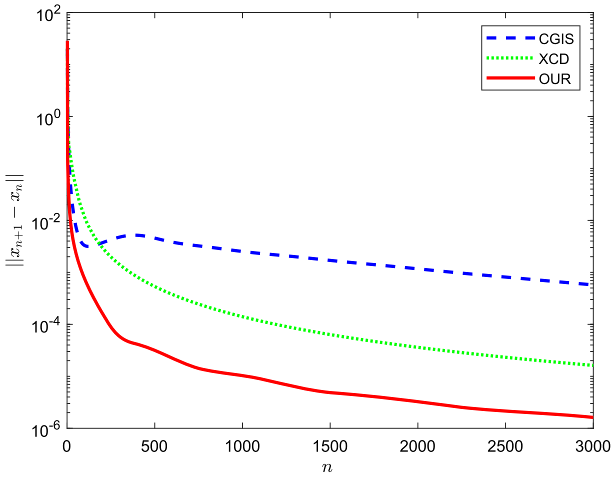

Fig. 1 The value of error versus the iteration numbers for Example 1

(i) $\nabla f(x) = (1 + \log(x_1), \ldots, 1 + \log(x_m))^T$,
$(\nabla f)^{-1}(x) = (\exp(x_1 - 1), \ldots, \exp(x_m - 1))^T$,
(ii) $\nabla f(x) = x$ and $(\nabla f)^{-1}(x) = x$.

The corresponding Bregman distances are given by

(i) $\mathcal{D}_f(x, y) = \sum_{i=1}^{m}(x_i \log(\frac{x_i}{y_i}) + y_i - x_i)$ which is the Kullback–Leibler distance (shortly, **KLD**),
(ii) $\mathcal{D}_f(x, y) = \frac{1}{2}\|x - y\|^2$ which is the squared Euclidean distance (shortly, **SED**).

Next we define the operator $\mathscr{F}$ by $\mathscr{F}(x) = \max(x, 0)$. Clearly, $\mathscr{F}$ is monotone and $\mathcal{S} = 0$. We compare the performance of Algorithm 3 using SED and KLD. The initial points are generated randomly for $m = 20, 50, 80, 120$, and $D_n = \|x_{n+1} - x_n\| < 10^{-4}$ is used as stopping criterion. The computation results are shown in Fig. 2 and Table 1. It can be seen that the convergence of the algorithm is different for different Bregman distances. Different functions correspond to different Bregman distances. Therefore, we use the Bregman projection method to obtain better convergence behavior.

Example 3 Let the operator $\mathscr{F}(x) := Mx + q$, where

$$M = BB^T + G + D,$$

and B is an $m \times m$ matrix, G is an $m \times m$ skew-symmetric matrix, D is an $m \times m$ diagonal matrix, whose diagonal entries are non-negative (so M is positive semidefinite), q is a vector in $\mathbb{R}^m$. The feasible set $\mathcal{C} \subset \mathbb{R}^m$ is a closed and convex subset defined by $\mathcal{C} := \{x \in \mathbb{R}^m : Qx \leq b\}$, where Q is an $k \times m$ matrix and b is a non-negative vector. It is clear that $\mathscr{F}$ is monotone and L-Lipschitz continuous with $L = \|M\|$. Let $q = 0$. Then, the solution set is $\{0\}$.

In this example, the initial values x_0 and x_1 are both set to $(1, 1, \ldots, 1)$, $k = 20$, $m = 10$, $\alpha_i = \frac{1}{i+1}$ and $\theta = 0.5$. $\|x_{n+1} - x_n\|$ represents the error of the n-th step iteration. For the two key parameters τ and ξ in Algorithm 3, we fix one parameter and change the remaining one. Finally, we get the following two cases, which are respectively shown in Figs. 3 and 4.

1. $\tau = 0.99/L,\ 0.5/L,\ 0.01/L$ and $\xi = 0.01/L$.

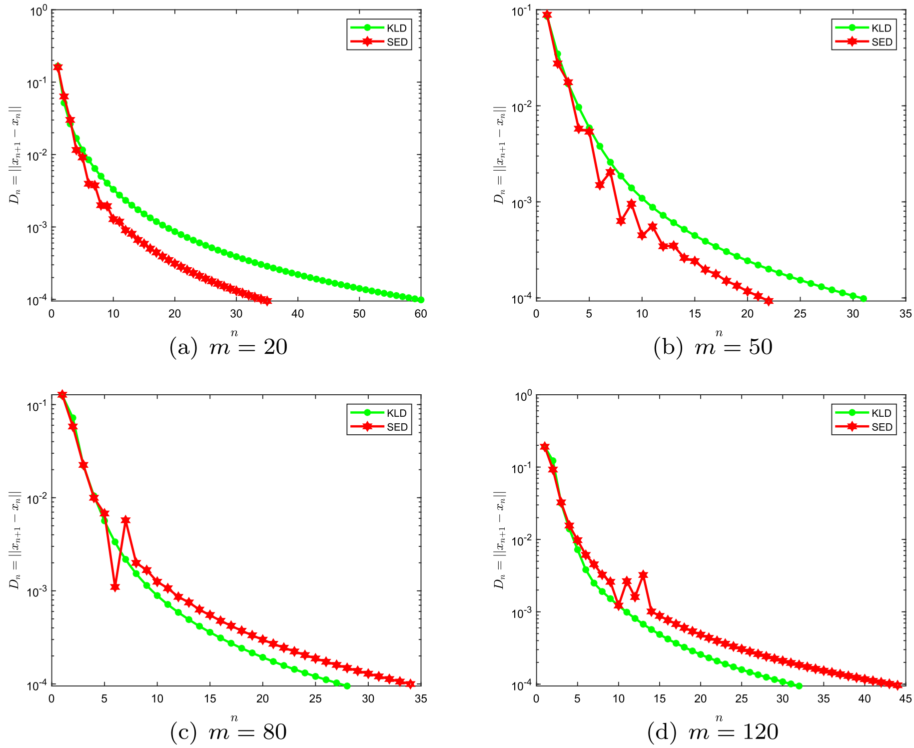

Fig. 2 Numerical behavior of all algorithms with different m in Example 2

Table 1 The number of termination iterations and execution time of our algorithm with different distances in Example 2

Bregman distance	$m = 20$		$m = 50$		$m = 80$		$m = 120$	
	Iter.	Times	Iter.	Times	Iter.	Times	Iter.	Times
KLD	60	0.0625	31	0.0469	28	0.0625	32	0.1094
SED	35	0.4531	22	0.3750	34	1.4062	44	3.4219

2. $\xi = 0.01/L,\ 0.1/L,\ 0.2/L$ and $\tau = 0.2/L$.

As shown in Figs. 3 and 4, each parameter will affect the convergence rate of our proposed algorithm, and the result of this effect is not monotonically increasing or decreasing with the parameter value. In particular, the values of τ and ξ are often consistent in other known results, but our algorithm breaks this restriction. We can choose that τ and ξ do not necessarily have to be the same, resulting in a better convergence behavior.

5 Conclusions

In this paper, we introduced an improved subgradient extragradient method with Bregman distance for solving pseudomonotone variational inequality problems. We generalized the

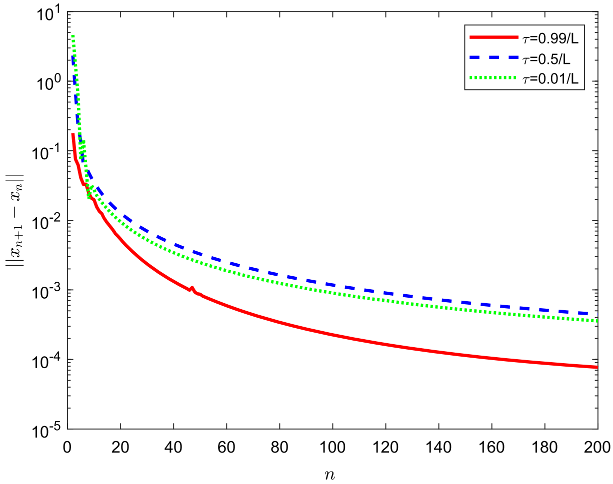

Fig. 3 The value of error versus the iteration numbers with different value on τ for Example 3 with $k = 20, m = 10, \xi = 0.01/L$

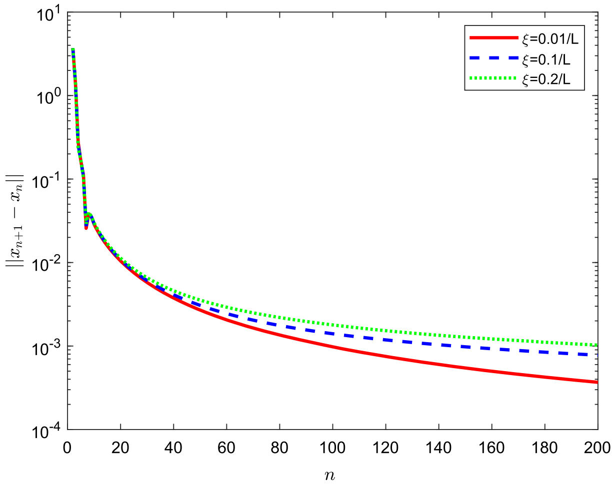

Fig. 4 The value of error versus the iteration numbers for Example 3 with $k = 20, m = 10, \tau = 0.2/L$

classical subgradient extragradient method to reflexive Banach spaces and effectively added an inertia term to speed up the convergence process of the algorithm. In addition, we set two different step size parameters, which help us to choose more appropriate values and thus increase the convergence rate. Under reasonable assumptions, we proved that the sequence generated by the proposed algorithm strongly converges to a solution of variational inequality problems. Finally, we presented numerical experiments to verify that our algorithm is effective in improving the iteration efficiency.

Acknowledgements This work was supported by the NSF of China (Grant no. 12171062).

Data availability statement Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Abass HA, Godwin GC, Narain OK, Darvish V (2022) Inertial extragradient method for solving variational inequality and fixed point problems of a Bregman demigeneralized mapping in a reflexive Banach spaces. Numer Funct Anal Optim 43:933–960

Abubakar J, Kumam P, Rehman HU (2022) Self-adaptive inertial subgradient extragradient scheme for pseudomonotone variational inequality problem. Int J Nonlinear Sci Numer Simul 23:77–96

Alber YI (1996) Metric and generalized projection operators in Banach spaces: properties and applications. Theory and applications of nonlinear operators of accretive and monotone type, Lecture Notes in Pure and Appl. Math., vol 178. Dekker, New York, pp 15–50

Barbagallo A, Di Vincenzo R (2015) Evolutionary variational inequality with long-term memory and applications to economic networks. Optim. Methods Softw. 30:253–275

Butnariu D, Iusem AN (2000) Totally convex functions for fixed points computation and infinite dimensional optimization. Applied Optimization, vol 40. Kluwer Academic Publishers, Dordrecht

Cai G, Gibali A, Iyiola OS, Shehu Y (2018) A new double-projection method for solving variational inequalities in Banach spaces. J Optim Theory Appl 178:219–239

Ceng LC, Petrusel A, Yao JC (2014) Composite viscosity approximation methods for equilibrium problem, variational inequality and common fixed points. J Nonlinear Convex Anal 15:219–240

Censor Y, Lent A (1981) An iterative row-action method for interval convex programming. J Optim Theory Appl 34:321–353

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

Dong QL, Lu YY, Yang J (2016) The extragradient algorithm with inertial effects for solving the variational inequality. Optimization 65:2217–2226

Hieu DV, Cho YJ, Xiao Y, Kumam P (2020) Relaxed extragradient algorithm for solving pseudomonotone variational inequalities in Hilbert spaces. Optimization 69:2279–304

Jitpeera T, Kumam P (2010) An extragradient type method for a system of equilibrium problems, variational inequality problems and fixed points of finitely many nonexpansive mappings. J Nonlinear Anal Optim 1:71–91

Jolaoso LO, Shehu Y (2022) Single Bregman projection method for solving variational inequalities in reflexive Banach spaces. Appl Anal 101:4807–4828

Jolaoso LO, Qin X, Shehu Y, Yao JC (2021) Improved subgradient extragradient methods with self-adaptive stepsizes for variational inequalities in Hilbert spaces. J Nonlinear Convex Anal 22:1591–1614

Jolaoso LO, Oyewole OK, Aremu KO (2022) A Bregman subgradient extragradient method with self-adaptive technique for solving variational inequalities in reflexive Banach spaces. Optimization 71:3835–3860

Korpelevich GM (1976) The extragradient method for finding saddle points and other problems. Ekon Mat Metod 12:747–756

Kraikaew R, Saejung S (2014) Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces. J Optim Theory Appl 163:399–412

Lions JL (1977) Numerical methods for variational inequalities-applications in physics and in control theory. Inf Process 77:917–924

Maingé PE (2008) Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization. Set-Val Anal 16:899–912

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

Mashreghi J, Nasri M (2010) Forcing strong convergence of Korpelevich's method in Banach spaces with its applications in game theory. Nonlinear Anal 72:2086–2099

Naraghirad E, Yao JC (2013) Bregman weak relatively nonexpansive mappings in Banach spaces. Fixed Point Theory Appl 141:43

Oyewole OK, Jolaoso LO, Aremu KO, Olayiwola MO (2022) Inertial self-adaptive Bregman projection method for finite family of variational inequality problems in reflexive Banach spaces. Comput Appl Math 41:22

Phelps RR (1993) Convex functions, monotone operators and differentiability. Lecture Notes in Mathematics, vol 1364, 2nd edn. Springer, Berlin

Popov LD (1980) A modification of the Arrow-Hurwitz method of search for saddle points. Mat Zamet 28:777–784

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

Reich S, Tuyen TM, Sunthrayuth P, Cholamjiak P (2021) Two new inertial algorithms for solving variational inequalities in reflexive Banach spaces. Numer Funct Anal Optim 42:1954–1984

Tan B, Liu L, Qin X (2022) Strong convergence of inertial extragradient algorithms for solving variational inequalities and fixed point problems. Fixed Point Theory 23:707–727

Thong DV, Hieu DV (2018) Modified subgradient extragradient method for variational inequality problems. Numer Algorithms 79:597–610

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

Xie Z, Cai G, Li X, Dong QL (2021) Strong convergence of the modified inertial extragradient method with line-search process for solving variational inequality problems in Hilbert spaces. J Sci Comput 88:19

Xie Z, Cai G, Dong QL (2023) Strong convergence of Bregman projection method for solving variational inequality problems in reflexive Banach spaces. Numer Algorithms 93:269–294

Xu HK (2002) Iterative algorithms for nonlinear operators. J Lond Math Soc 66:240–256

Yang J, Liu H, Li G (2020) Convergence of a subgradient extragradient algorithm for solving monotone variational inequalities. Numer Algorithms 84:389–405

Yao Y, Liou YC, Yao JC (2011) New relaxed hybrid-extragradient method for fixed point problems, a general system of variational inequality problems and generalized mixed equilibrium problems. Optimization 60:395–412

Yao Y, Postolache M, Yao JC (2020) Strong convergence of an extragradient algorithm for variational inequality and fixed point problems. Politehn Univ Buchar Sci Bull Ser A Appl Math Phys. 82:3–12

Yao Y, Iyiola OS, Shehu Y (2022) Subgradient extragradient method with double inertial steps for variational inequalities. J Sci Comput 90:29

Zâlinescu C (2002) Convex analysis in general vector spaces. World Scientific Publishing, Singapore