J. Nonlinear Var. Anal. 8 (2024), No. 3, pp. 396-417
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.8.2024.3.04

A BREGMAN PROJECTION ALGORITHM WITH SELF ADAPTIVE STEP SIZES FOR SPLIT VARIATIONAL INEQUALITY PROBLEMS INVOLVING NON-LIPSCHITZ OPERATORS

LIYA LIU[1], SUN YOUNG CHO[2,*]

[1] *School of Mathematics and Statistics, Southwest University, Chongqing 400715, China*
[2] *Department of Human Health Care, Gyeongsang National University, Jinju, Korea*

Abstract. The purpose of this paper is to investigate a Bregman projection algorithm for solving the split variational inequality problem governed by pseudomonotone and not necessarily Lipschitz continuous operators in real Hilbert spaces. The proposed algorithm is motivated by the ideas of the Halpern method, the CQ method, and Tseng's extragradient method. The step size sequences are determined by employing Armijo line search techniques. The strong convergence theorem is established without the prior knowledge of the operator norm and the Lipschitz continuous assumption on the operators involved. Some numerical experiments with graphical illustrations are presented to demonstrate the effectiveness and the performance of our proposed algorithm in comparison with some existing ones.

Keywords. Bregman projection; Line search rule; Pseudomonotone operator; Split variational inequality problem; Tseng's extragradient method.

1. Introduction

Let H be a real Hilbert space endowed with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$. Let C be a nonempty, convex, and closed subset of space H, and let $F: H \to H$ be a mapping. Recall that the variational inequality problem is formulated as finding a vector $x \in C$ such that

$$\langle Fx, y-x\rangle \geq 0,\ \forall x \in C.$$

The solution set of the variational inequality problem is denoted by $S(C,F)$. The notion of the variational inequality problem was first introduced by Stampacchia [1]. This problem has been intensively and widely studied since it provides a fundamental framework for solving several problems in engineering, mechanics, finance, data sciences, and economics, and so on; see, e.g., [2–5].

Recall that the split feasibility problem is to find a vector $x^* \in C$ such that $Ax^* \in Q$, where C and Q are nonempty, convex, and closed sets in real Hilbert spaces H_1 and H_2, respectively, and A is a bounded and linear operator from H_1 to H_2. The split feasibility problem was first introduced by Censor and Elfying [6]. It is known that this model plays a key role in the inverse problems arising in intensity-modulated radiation therapy and treatment planning; see [7].

*Corresponding author.
E-mail address: liya42@swu.edu.cn (L. Liu), ooly61@hotmail.com (S.Y. Cho).
Received 19 December 2023; Accepted 5 March 2024; Published 5 April 2024.

As an important generalization of the variational inequality problem and the split feasibility problem, Censor et al. [8] introduced a split variational inequality problem (shortly, SVI). Let H_1 and H_2 be two real Hilbert spaces, and let $A : H_1 \to H_2$ be a bounded and linear operator. Let C and Q be two nonempty, convex, and closed sets in H_1 and H_2, respectively. Given two operators $F_1 : H_1 \to H_1$ and $F_2 : H_2 \to H_2$, the SVI is to find a solution x^* of the variational inequality problem in space H_1 so that the image $y = A(x^*)$, under a given bounded and linear operator A, is a solution to another variational inequality problem in space H_2. More specifically, the SVI can be formulated as finding an element $x^* \in C$ that solves

$$\langle F_1(x^*), x - x^* \rangle \geq 0, \ \forall x \in C, \tag{1.1}$$

such that $y^* = A(x^*) \in Q$ solves

$$\langle F_2(y^*), y - y^* \rangle \geq 0, \ \forall y \in Q. \tag{1.2}$$

Throughout the paper, we use Γ to denote the solution set of the SVI above, that is, $\Gamma := \{x^* \in S(C, F_1) | Ax^* \in S(Q, F_2))\}$. The SVI, which can be viewed as a combination of the variational inequality problem and the split feasibility problem, provides a unified model for treating a wide range of mathematical problems, including linear inverse problems, split zero problems, split minimization problems, and variational inclusion problems; see, e.g., [8–10].

We next recall some known algorithms for solving variational inequality problems. One of the efficient methods is the extragradient method proposed by Korpelevich [11]. Notice that the extragradient method may be costly computed, since it requires two orthogonal projections onto a given feasible set per iteration. To obtain better implementable and more efficient algorithms for solving variational inequality problems, one important task is to minimize the number of projections onto the feasible set involved per iteration. Motivated by this research trend, Censor et al. [12] proposed the so-called subgradient extragradient method, which replaces the second projection onto the feasible set in the extragradient method with the one onto a specific constructible half-space. Another method proposed by Tseng [13], is called Tseng's extragradient method, wherein only one projection is required at each iteration. The saving of one projection step allows this method to be more potentially efficient.

Notice that $S(C, F_1)$ and $S(Q, F_2)$ are the solution sets of (1.1) and (1.2), respectively. It is obvious that the SVI can be transformed into a fixed point problem, that is, $x^* \in \Gamma$ is equivalent to $x^* = P_C(I - F_1)(x^* + \tau A^*(P_Q(I - F_2) - I)Ax^*),\ \tau > 0$. Censor et al. [8] were inspired to extend the well-established CQ algorithm proposed by Byrne [7] to solve the SVI. Starting with an arbitrary initial $x_1 \in H_1$, the sequence $\{x_n\}_{n=2}^{\infty}$ is generated by

$$x_{n+1} = P_C(I - \lambda F_1)(x_n + \tau A^*(P_Q(I - \lambda F_2) - I)Ax_n), \tag{1.3}$$

where P_C and P_Q are metric (nearest point) projections onto sets C and Q, respectively, $F_1 : H_1 \to H_1$ and $F_2 : H_2 \to H_2$ are inverse-strongly monotone mapping with constants η_1 and η_2, respectively, $\lambda \in [0, 2\min\{\eta_1, \eta_2\}]$, $A : H_1 \to H_2$ is a bounded and linear operator with its adjoint operator $A^* : H_2 \to H_1$, and $\tau \in (0, 1/L)$ with L being the spectral radius of the operator A^*A. They proved that the generated sequence $\{x_n\}$ converges weakly to a solution of the SVI provided that $\Gamma \neq \emptyset$. The convergence theorem requires to calculate the spectral norm of the operator A. The implementation of this method depends on the knowledge of the bounded linear operator norm.

It also should be noted that strong convergence results are much more desirable than weak convergence results in infinite dimensional Hilbert spaces. Some methods, such as the Halpern method, the viscosity approximation method, and the hybrid projection method, can be employed to guarantee the strong convergence result. For solving the SVI, Thuy et al. [10] modified Algorithm (1.3) by using the ideas of Tseng's extragradient method and the hybrid steepest descent method, that is, for all $x_1, x_2 \in H_1$,

$$\begin{cases} w_n = x_n + \alpha_n(x_n - x_{n-1}), \text{ where } \alpha_n = \begin{cases} \min\{\frac{\theta_n}{\|x_n - x_{n-1}\|}, \theta\}, & \text{if } x_n \neq x_{n-1}, \\ 0, & \text{otherwise}, \end{cases} \\ y_n = \beta_n w_n + (1-\beta_n)P_C(I-\gamma F_1)w_n, \\ u_n = P_Q(I-\gamma F_2)A(y_n), \\ z_n = y_n - \lambda_n A^*(A(y_n)-u_n), \text{ where } \lambda_n = \begin{cases} \frac{\rho_n\|A(y_n)-u_n\|^2}{\|A^*(A(y_n)-u_n)\|^2}, & \text{if } \|A^*(A(y_n)-u_n)\| \neq 0, \\ 0, & \text{otherwise}, \end{cases} \\ x_{n+1} = z_n - \delta_n G(z_n), \end{cases} \tag{1.4}$$

where $\{\beta_n\} \subset (0,1), \{\rho_n\} \subset (0,1), \{\delta_n\} \subset (0,1)$, $\lim_{n\to\infty} \delta_n = 0$, $\sum_{n=1}^{\infty} \delta_n = \infty$, $\lim_{n\to\infty} \frac{\alpha_n}{\delta_n} = 0$, $\gamma \in (0, 2\min\{\eta_1, \eta_2\}]$, $F_1 : H_1 \to H_1$ and $F_2 : H_2 \to H_2$ are two inverse-strongly monotone with constants η_1 and η_2, $G : H_1 \to H_1$ is a strongly monotone and Lipschitz continuous operator, and $A : H_1 \to H_2$ is a linear and bounded operator with its adjoint operator $A^* : H_2 \to H_1$. They proved that $\{x_n\}$ converges strongly to a point $\tilde{x} \in \Gamma$ such that $\tilde{x} = P_\Gamma(I-G)\tilde{x}$ in a real Hilbert space. The main feature of the algorithm is that the step sizes do not depend on the norm of the operator A. However, the weakness of (1.3) and (1.4) is that the convergence requires operators involved to be inverse-strongly monotone, which is a restrictive assumption.

In order to relax the monotonicity of the operators involved, authors proposed various modified projection-based methods for solving the SVI. Since the SVI can be transformed into an equivalent constrained variational inequality in a product space, Censor et al. [8] employed the subgradient extragradient method to solve the SVI. The convergence of their method requires that their operators F_1 and F_2 are monotone and Lipschitz continuous. He et al. [14] proposed a relaxed projection algorithm for solving the SVI. Their method is consisted of a prediction step with two projections and a correction step. In a dimensional Euclidean space, the global convergence was established under the condition that F_1 and F_2 are monotone and Lipschitz continuous. Izuchukwu et al. [9] proposed a modified projection and contraction method for solving the SVI with the monotone and Lipschitz continuous assumption. To relax the monotonicity of the operators involved to pseudomonotonicity, Huy et al. [15] proposed a modified subgradient extragradient method of the Halpern type for solving the SVI. Their method requires two projections onto feasible sets and two projections onto half-spaces at each iteration and proved the strong convergence result under a mild assumption.

The step-size sequences play an essential role in the computational efficiency of algorithms. The update of step sizes often depends on the prior information of either the Lipschitz constant or the norm of the operators involved. This usually slows down the convergence rate of algorithms. However, the coefficient or the norm of given operators may not be known or may be difficult to estimate in many practical cases, which often affects the implementation and the convergence of algorithms. To overcome this drawback, the construction of self-adaptive step sizes has aroused numerous interest among researchers; see [10, 17]. Recently, some self-adaptive step-size techniques were employed to generate non-increasing sequences of step-sizes; see, e.g., [15–19]. In our work, the step sizes will be selected self-adaptively by using Armijo line-search techniques without requiring the knowledge of the Lipschitz constants and the operator norm.

It is worth noting that the methods mentioned above are based on norm distances and metric projections. In recent years, Bregman projection algorithms for solving optimization problems have become an important and interesting topic, due to the fact that the Bregman distance is a useful substitute for the norm distance. It is known that the Bregman distance can be regarded as an elegant and effective technique for solving problems arising in the nonlinear analysis and optimization theory. On the other hand, the applications of the Bregman distance associated with various choices of functions gives us an alternative way in selecting different projections. From the numerical point of view, the Bregman projection method is implementable. Then it seems to be more flexible to use the Bregman distance instead of the usual norm distance. Due to the wide applications of Bregman distances, methods for solving variational inequality problems with Bregman projections can be found in [20–24]. For instance, Gibali [23] proposed a nice extension of the subgradient extragradient method by using Bregman distance techniques; Sunthrayuth et al. [24] modified and improved Tseng's extragradient algorithm with the Bregman projection; Jolaoso et al. [22] proposed two relaxed inertial Halpern-type algorithms of Bregman distances; and Hieu and Reich [21] proposed two Bregman extragradient-like methods.

Motivated by the results above and the ongoing research interest in this direction, our interest in this paper is to develop Tseng's extragradient algorithm in the sense of the Bregman divergence for solving the SVI. The strong convergence analysis is investigated when the involved operators are pseudomonotone, uniformly continuous, and not necessarily Lipschitz continuous. Numerical examples are presented to demonstrate the efficiency of the proposed method in comparison with some existing ones. Our proposed method has key features as follows.

(i) The proposed method is independent on the norm of the bounded and linear operator nor on the Lipschitz constants of the involved operators, unlike the ones in [17–19].
(ii) The strong convergence result requires the underlying operators to be pseudomonotone, which is much more weaker than the monotonicity. This allows our method to be applied to a wider class of nonlinear operators in comparison with the methods in [8, 9, 14].
(iii) The algorithm based on norm distances for solving the SVI is extended to the more general framework of Bregman distances in Hilbert spaces; see [18].
(iv) Our method converges strongly to a solution of the SVI, which is an important factor to consider in an infinite dimensional space.

The rest of the paper is organized as follows. In Section 2, we recall some essential definitions and technical lemmas. Section 3 states the algorithm and analyzes its strong convergence. In Section 4, numerical implementations and comparisons are present to support our theoretical findings. Finally, the paper is concluded with a brief summary in Section 5.

2. Preliminaries

Throughout this paper, we denote by H a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $\mathbb{N}$ and $\mathbb{R}$ denote the set of all positive integers and the set of all real numbers, respectively. For a sequence $\{x_n\} \subset H$, the strong convergence and the weak convergence of

$\{x_n\}$ to $x \in H$ are denoted by $x_n \to x$ and $x_n \rightharpoonup x$, as $n \to \infty$, respectively. The weak limit set of the sequence $\{x_n\}$ is denoted by $\omega_w(x_n) = \{x \in H | x_{n_k} \rightharpoonup x \text{ as } k \to \infty, \text{for some subsequence } \{x_{n_k}\} \text{ of } \{x_n\}\}$. We now recall some definitions and lemmas.

An operator $F : H \to H$ is said to be

(i) η-inverse strongly monotone if there exists a constant $\eta > 0$ such that $\langle Fx - Fy, x - y\rangle \geq \eta \|Fx - Fy\|^2$ for all $x, y \in H$;
(ii) monotone if $\langle Fx - Fy, x - y\rangle \geq 0$ for all $x, y \in H$;
(iii) pseudomonotone if $\langle Fx, y - x\rangle \geq 0 \Rightarrow \langle Fy, y - x\rangle \geq 0$ for all $x, y \in H$;
(iv) L-Lipschitz continuous if there exists a constant $L > 0$ such that $\|Fx - Fy\| \leq L\|x - y\|$ for all $x, y \in H$.

Let H_1 and H_2 be two Hilbert spaces, and let $A : H_1 \to H_2$ be a bounded and linear operator. An operator $A^* : H_2 \to H_1$ is called the adjoint operator of A if $\langle A(x), y\rangle = \langle x, A^*(y)\rangle$ for all $x \in H_1$ and $y \in H_2$.

Let $f : H \to \mathbb{R}$ be a convex and differentiable function with a nonempty domain $domf = \{x \in H | f(x) < \infty\}$. The subdifferential set of f at x is defined by $\partial f(x) = \{u \in H | f(y) \geq f(x) + \langle u, y - x\rangle, \forall y \in H\}$. When f is differentiable, then the element in $\partial f(x)$ is the gradient of f at x, denoted by $\nabla f(x)$. The interior of the domain of f is denoted by $int(domf)$. We say that f is Gâteaux differentiable at $x \in int(domf)$ if the limit

$$\lim_{t \to 0} \frac{f(x + ty) - f(x)}{t} = \langle \nabla f(x), y\rangle \tag{2.1}$$

exists for every $y \in H$. Moreover, f is said to be Gâteaux diferentiable if it is Gâteaux diferentiable at every $x \in int(domf)$. We say that f is uniformly Fréchet differentiable on a subset C of a real Hilbert space H, if limit (2.1) is attained uniformly for any $y \in H$ with $\|y\| = 1$ and $x \in C$. Furthermore, ∇f is uniformly continuous on a bounded subset C of a real Hilbert space H, if f is uniformly Fréchet differentiable and bounded on C; see [25].

The Bregman distance $D_f : domf \times int(domf) \to [0, \infty)$ corresponding to a strictly convex and differentiable function f with its gradient ∇f is defined by [26]

$$D_f(x, y) = f(x) - f(y) - \langle \nabla f(y), x - y\rangle, \ \forall x \in domf, \ y \in int(domf).$$

The Bregman distance is not a usual metric (see [23, 27] for examples). However, it can be characterized by the following properties (see [28])

(i) (The two point identity) for any $x, y \in int(domf)$,

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

(ii) (The three point identity) for any $x \in domf$ and $y, z \in int(domf)$,

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

We recall that a function $f : H \to \mathbb{R}$ is said to be strongly convex with a constant $\sigma > 0$ if

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - \frac{\sigma}{2} t(1 - t)\|x - y\|^2, \ \forall x, y \in domf, \forall t \in [0, 1].$$

The Bregman distance D_f corresponding to the σ-strongly convex function f can be characterized by the inequality of $D_f(x, y) \geq \frac{\sigma}{2}\|x - y\|^2$, $\forall x \in domf$, $y \in int(domf)$; see [22].

The Bregman projection associated with f of $x \in int(domf)$ onto a nonempty, convex and closed set $C \subset int(domf)$ [26] is the unique point in C, Π_C^f, defined by

$$\Pi_C^f(x) := \arg\min\{D_f(y,x) : y \in C\}.$$

The Fenchel conjugate function of f is the convex function $f^* : H \to (-\infty, +\infty]$ defined by

$$f^*(x^*) := \sup_{x \in H}\{\langle x^*, x\rangle - f(x)\}.$$

A function $f : H \to \mathbb{R}$ is called Legendre [23, 29] if it satisfies

(i) $int(domf) \neq \emptyset$ and the subdifferential ∂f is single-valued on its domain;
(ii) $int(domf^*) \neq \emptyset$ and ∂f^* is single-valued on its domain.

Let $V_f : domf^* \times domf \to [0, +\infty)$ associated with a Legendre function f be defined by

$$V_f(\eta, x) := f(x) - \langle \eta, x\rangle + f^*(\eta), \ \forall \eta \in domf^*, x \in domf.$$

Some properties of the function V_f can be summarized as follows (see [30])

(i) $V_f(x,y) = D_f(x, \nabla f^*(y)), \ \forall x \in domf, \ y \in int(domf)$;
(ii) $V_f(x,\eta) + \langle \zeta, \nabla f^*(\eta) - x\rangle \leq V_f(x, \eta + \zeta), \ \forall \eta, \zeta \in domf^*, \ x \in domf$;
(iii) V_f is nonnegative and convex in the second variable.

Since V_f is convex in the second variable, then, for $N \in \mathbb{N}$,

$$D_f\left(x, \nabla f^*\left(\sum_{i=1}^N \lambda_i \nabla f(y_i)\right)\right) \leq \sum_{i=1}^N \lambda_i D_f(x, y_i), \ \forall x \in domf,$$

where $\{y_i\}_{i=1}^N \subset H$ and $\{\lambda_i\}_{i=1}^N \subset [0,1]$ with $\sum_{i=1}^N \lambda_i = 1$.

Lemma 2.1. *[31] The Bregman projection $\Pi_C^f(x)$ has the following properties, for each $x \in H$,*

(i) $\langle \nabla f(\Pi_C^f(x)) - \nabla f(x), y - \Pi_C^f(x)\rangle \geq 0, \ \forall y \in C$;
(ii) $D_f(y, \Pi_C^f(x)) + D_f(\Pi_C^f(x), x) \leq D_f(y,x), \ \forall y \in C$.

Lemma 2.2. *[24] Let $f : H \to \mathbb{R}$ be strongly convex, Fréchet differentiable, and bounded on bounded subsets of H. Let $\{x_n\}$ and $\{y_n\}$ be two sequences in H. If $\lim_{n\to\infty} D_g(x_n, y_n) = 0$, then $\lim_{n\to\infty}\|x_n - y_n\| = 0$.*

Lemma 2.3. *[32] Let H_1 and H_2 be two real Hilbert spaces. Suppose that $f : H_1 \to H_2$ is uniformly continuous on a bounded subset C of H_1. Then $f(C) := \{f(x)|x \in C\}$ is bounded.*

Lemma 2.4. *[22, 33] For all $x \in H$ and $\mu \geq \nu > 0$, the following inequalities hold*

$$\left\|\frac{x - \Pi_C^f \nabla f^*(\nabla f(x) - \mu F(x))}{\mu}\right\| \leq \left\|\frac{x - \Pi_C^f \nabla f^*(\nabla f(x) - \nu F(x))}{\nu}\right\|.$$

Lemma 2.5. *[34] Let C be a nonempty, convex, and closed subset of a real Hilbert space H, and let $F : C \to H$ be a pseudomonotone and continuous operator. Then $\tilde{x} \in S(C,F)$ if and only if $\langle Fz, z - \tilde{x}\rangle \geq 0, \ \forall z \in C$.*

Lemma 2.6. *[35] Let $\{a_n\}$ be a nonnegative real sequence with $a_{n+1} \leq \lambda_n \beta_n + (1-\lambda_n)a_n$ for all $n \geq 1$, where $\{\lambda_n\}$ is a sequence in $(0,1)$ with the condition $\sum_{n=1}^\infty \lambda_n = \infty$, and $\{\beta_n\}$ is a sequence with the condition $\limsup_{n\to\infty}\beta_n \leq 0$. Then $\lim_{n\to\infty} a_n = 0$.*

Lemma 2.7. *[36] Let $\{a_n\}$ be a nonnegative real sequence such that there exists a subsequence $\{n_i\}$ of $\{n\}$ such that $a_{n_i} < a_{n_i+1}$, $\forall i \in \mathbb{N}$. Then there exists an increasing sequence $\{\varphi(m)\} \subset \mathbb{N}$ such that $\lim_{m\to\infty} \varphi(m) = \infty$ and the following properties are satisfied by all (sufficiently large) numbers $m \in \mathbb{N}$: $a_{\varphi(m)} \leq a_{\varphi(m)+1}$ and $a_m \leq a_{\varphi(m)+1}$. In fact, $\varphi(m) = \max\{j \leq m | a_j \leq a_{j+1}\}$.*

3. The Algorithm and Convergence Analysis

In this section, we propose a Bregman projection algorithm for solving the SVI. In order to establish the strong convergence of the algorithm, we make the following standard assumptions.

Assumption 3.1. ($A1$) C is a nonempty, convex, and closed subset of a real Hilbert space H_1; Q is a nonempty, convex, and closed subset of a real Hilbert space H_2.

($A2$) The function $f : H_1 \to \mathbb{R}$ is σ-strongly convex, Legendre, which is bounded and uniformly Fréchet differentiable on bounded subsets of H_1.

($A3$) The function $g : H_2 \to \mathbb{R}$ is ς-strongly convex, Legendre, which is bounded and uniformly Fréchet differentiable on bounded subsets of H_2.

($A4$) The operator $F_1 : H_1 \to H_1$ is pseudomonotone and uniformly continuous, which satisfies whenever $\{p_n\} \subset H_1, p_n \rightharpoonup p$ as $n \to \infty$, one has $\|F_1(p)\| \leqslant \liminf_{n\to\infty} \|F_1(p_n)\|$.

($A5$) The operator $F_2 : H_2 \to H_2$ is pseudomonotone and uniformly continuous, which satisfies whenever $\{q_n\} \subset H_2, q_n \rightharpoonup q$ as $n \to \infty$, one has $\|F_2(q)\| \leqslant \liminf_{n\to\infty} \|F_2(q_n)\|$.

($A6$) The operator $A : H_1 \to H_2$ is a bounded linear operator with $\|A\| \neq 0$ and an adjoint $A^* : H_2 \to H_1$.

($A7$) The real sequence $\{\lambda_n\} \subset (0,1)$ satisfies that $\sum_{n=1}^{\infty} \lambda_n = \infty$ and $\lim_{n\to\infty} \lambda_n = 0$.

($A8$) The solution set $\Gamma := \{x \in S(C,F_1) | Ax \in S(Q,F_2)\}$ is nonempty.

The algorithm is of the following form

Algorithm 1

Step 1: Let $x_1 \in H_1$ be arbitrary. Take $\kappa, \iota, \chi, \rho, \gamma \in (0,1)$, $\mu \in (0,\varsigma)$, $\upsilon \in (0,\sigma)$, $\varpi \in (0,\infty)$, $\eta \in (1,\infty)$, $\alpha \in (0,\chi/\eta)$ and $\{\lambda_n\}$ is a real sequence given by ($A7$). Set $n = 1$.

Step 2: Compute

$$u_n = \Pi_Q^g \nabla g^*(\nabla g(A(x_n)) - \mu_n F_2(A(x_n))),$$

where $\mu_n = \kappa \iota^{k_n}$ with k_n being the smallest nonnegative integer k satisfying

$$\kappa \iota^k \|F_2(u_n) - F_2(A(x_n))\| \leq \mu \|u_n - A(x_n)\|. \tag{3.1}$$

Step 3: Compute

$$t_n = \nabla g^*(\nabla g(u_n) - \mu_n(F_2(u_n) - F_2(A(x_n)))).$$

Step 4: Compute

$$y_n = \nabla f^*(\nabla f(x_n) + \alpha_n A^*(\nabla g(t_n) - \nabla g(A(x_n)))),$$

where α_n is defined by $\alpha_n = \varpi \eta^{l_n}$ with l_n being the smallest integer l satisfying

$$\alpha D_f(y_n, x_n) \leq \alpha_n D_g(A(y_n), A(x_n)). \tag{3.2}$$

Step 5: Compute

$$v_n = \Pi_C^f \nabla f^*(\nabla f(y_n) - \upsilon_n F_1(y_n)),$$

where $\upsilon_n = \rho\gamma^{m_n}$ with m_n being the smallest nonnegative integer m satisfying

$$\rho\gamma^m \|F_1(v_n) - F_1(y_n)\| \le \upsilon \|v_n - y_n\|. \tag{3.3}$$

If $u_n = A(x_n)$ and $v_n = y_n$, then stop. Otherwise, go to ***Step 6***.

Step 6: Compute

$$s_n = \nabla f^*(\nabla f(v_n) - \upsilon_n(F_1(v_n) - F_1(y_n))).$$

Step 7: Compute

$$x_{n+1} = \nabla f^*(\lambda_n \nabla f(x_1) + (1-\lambda_n)\nabla f(s_n)).$$

Set $n := n+1$ and return back to ***Step 2***.

Remark 3.1. (i) The choice of step-size sequences $\{\mu_n\}$ and $\{\nu_n\}$ are determined by adopting linesearch techniques without assuming that F_1 and F_2 are Lipschitz continuous.

(ii) The nonmonotonic step-size sequence $\{\alpha_n\}$ is given by utilizing a self-adaptive step size technique with known parameters α, ϖ, and η, avoiding the use of the operator norm $\|A\|$.

Before investigating the convergence of our algorithm, we require some lemmas. The following one is concerned with the well-definedness of step size sequences $\{\mu_n\}$ and $\{\upsilon_n\}$.

Lemma 3.1. *(3.1) and (3.3), the Armijo-line search rules, are well-defined.*

Proof. The proof is similar to that of [22, Lemma 3.3]. □

Lemma 3.2. *The self-adaptive step size rule (3.2) is well-defined.*

Proof. If $D_g(A(y_n), A(x_n)) = 0$, then $D_f(y_n, x_n) = 0$. In this case, $l_n = 0$. If $D_g(A(y_n), A(x_n)) \neq 0$, we assume that the contrary of (3.2) holds for any integer l, that is, $\varpi\eta^l D_g(A(y_n), A(x_n)) < \alpha D_f(y_n, x_n)$. In this case, it follows that $A(y_n) \neq A(x_n)$. Since A is linear, we obtain that $y_n \neq x_n$. It further implies that $D_f(y_n, x_n) > 0$. By considering that A is bounded and $\eta > 1$, one finds that $\lim_{l\to+\infty} \varpi\eta^l D_g(A(y_n), A(x_n)) = +\infty$. From above, we obtain that

$$+\infty = \lim_{l\to+\infty} \varpi\eta^l D_g(A(y_n), A(x_n)) < \lim_{l\to+\infty} \alpha D_f(y_n, x_n) = \alpha D_f(y_n, x_n).$$

Thus we obtain a contradiction. The proof is completed. □

We precede the proof of the following lemma, which is crucial to the convergence theorem.

Lemma 3.3. *Let Assumption 3.1* (A1)*-(*A8*) be satisfied. Let $\{x_n\}$ be the sequence generated by Algorithm 1. Then, for any $x \in \Gamma$,*

$$\begin{aligned} D_f(x, x_{n+1}) \le& \lambda_n D_f(x, x_1) + (1-\lambda_n) D_f(x, s_n) \\ \le& \lambda_n D_f(x, x_1) + (1-\lambda_n)\Big[D_f(x, x_n) - (1-\alpha) D_f(y_n, x_n) \\ & - \left(1 - \frac{\upsilon}{\sigma}\right) D_f(v_n, y_n) - \left(1 - \frac{\upsilon}{\sigma}\right) D_f(s_n, v_n)\Big], \ \forall n \in \mathbb{N}. \end{aligned} \tag{3.4}$$

Proof. Let $x \in \Gamma$. Then $x \in S(C, F_1)$ and $A(x) \in S(Q, F_2)$. By considering the definition of $\{u_n\}$, one has that $u_n \in Q$. This together with $A(x) \in S(Q, F_2)$ implies that $\langle F_2(A(x)), u_n - A(x)\rangle \ge 0$. The pseudomonotonicity of F_2 gives that

$$\langle F_2(u_n), u_n - A(x)\rangle \ge 0. \tag{3.5}$$

Denote $z_n = A(x_n)$. By the definition of $\{u_n\}$ and Lemma 2.1 (i), one finds that

$$\langle \nabla g(u_n) - \nabla g(\nabla g^*(\nabla g(z_n) - \mu_n F_2(z_n))), A(x) - u_n\rangle \geq 0. \tag{3.6}$$

By combining (3.5) with (3.6), one sees that

$$\langle \nabla g(z_n) - \nabla g(u_n), A(x) - u_n\rangle \leq \mu_n \langle F_2(z_n) - F_2(u_n), A(x) - u_n\rangle. \tag{3.7}$$

By using the three point identity of the Bregman distance and (3.7), one finds that

$$\begin{aligned}
&D_g(A(x),t_n)\\
&= D_g(A(x),A(x_n)) - D_g(u_n,A(x_n)) + \langle \nabla g(A(x_n)) - \nabla g(u_n), A(x) - u_n\rangle - D_g(t_n,u_n)\\
&\quad + \mu_n\langle F_2(u_n) - F_2(A(x_n)), A(x) - t_n\rangle\\
&\leq D_g(A(x),A(x_n)) - D_g(u_n,A(x_n)) - D_g(t_n,u_n) + \mu_n\langle F_2(A(x_n)) - F_2(u_n), t_n - u_n\rangle.
\end{aligned} \tag{3.8}$$

In view of (3.1), one obtains that

$$\begin{aligned}
\mu_n\langle F_2(A(x_n)) - F_2(u_n), t_n - u_n\rangle &\leq \mu_n \|F_2(A(x_n)) - F_2(u_n)\|\|t_n - u_n\|\\
&\leq \frac{\mu}{2}(\|A(x_n) - u_n\|^2 + \|t_n - u_n\|^2).
\end{aligned} \tag{3.9}$$

By substituting (3.9) into (3.8) and using the relation $D_g(x,y) \geq \frac{\varsigma}{2}\|x-y\|^2$ $(\forall x \in dom(g), y \in int(dom(g)))$, one concludes that

$$\begin{aligned}
D_g(A(x),t_n) \leq& D_g(A(x),A(x_n)) - D_g(u_n,A(x_n)) - D_g(t_n,u_n) + \frac{\mu}{2}(\|A(x_n) - u_n\|^2 + \|t_n - u_n\|^2)\\
\leq& D_g(A(x),A(x_n)) - \left(1 - \frac{\mu}{\varsigma}\right) D_g(u_n,A(x_n)) - \left(1 - \frac{\mu}{\varsigma}\right) D_g(t_n,u_n).
\end{aligned} \tag{3.10}$$

Therefore, (3.10) implies that

$$D_g(A(x),t_n) \leq D_g(A(x),A(x_n)). \tag{3.11}$$

By the definition of $\{v_n\}$ and Lemma 2.1 (i), one has that

$$\langle \nabla f(v_n) - \nabla f(\nabla f^*(\nabla f(y_n) - \upsilon_n F_1(y_n))), x - v_n\rangle \geq 0. \tag{3.12}$$

Since $x \in S(C,F_1)$ and $v_n \in C$, one find that $\langle F_1(x), v_n - x\rangle \geq 0$. The pseudomonotonicity of F_1 yields that $\langle F_1(v_n), v_n - x\rangle \geq 0$, which together with (3.12) yields that

$$\upsilon_n\langle F_1(v_n) - F_1(y_n), x - v_n\rangle \leq \langle \nabla f(v_n) - \nabla f(y_n), x - v_n\rangle. \tag{3.13}$$

Using the three point identity of the Bregman distance and the definition of $\{s_n\}$ yields that

$$\begin{aligned}
D_f(x,s_n) =& D_f(x,v_n) - D_f(s_n,v_n) + \langle \nabla f(v_n) - \nabla f(s_n), x - s_n\rangle\\
=& D_f(x,v_n) - D_f(s_n,v_n) + \upsilon_n\langle (F_1(v_n) - F_1(y_n)), x - s_n\rangle.
\end{aligned} \tag{3.14}$$

Again, by using the three point identity of the Bregman distance, one sees that

$$D_f(x,v_n) = D_f(x,y_n) - D_f(v_n,y_n) + \langle \nabla f(y_n) - \nabla f(v_n), x - v_n\rangle. \tag{3.15}$$

It follows from (3.13), (3.14), and (3.15) that

$$\begin{aligned}
D_f(x,s_n) \leq& D_f(x,y_n) - D_f(v_n,y_n) + \upsilon_n\langle F_1(y_n) - F_2(v_n), x - v_n\rangle - D_f(s_n,v_n)\\
&+ \upsilon_n\langle (F_1(v_n) - F_1(y_n)), x - s_n\rangle\\
=& D_f(x,y_n) - D_f(v_n,y_n) - D_f(s_n,v_n) + \upsilon_n\langle F_1(y_n) - F_2(v_n), s_n - v_n\rangle.
\end{aligned} \tag{3.16}$$

The relation $D_f(x,y) \geq \frac{\sigma}{2}\|x-y\|^2$ $(\forall x \in dom f, y \in int(dom f))$ and (3.3) yield that

$$\begin{aligned}\upsilon_n\langle F_1(y_n)-F_2(v_n), s_n-v_n\rangle &\leq \upsilon\|y_n-v_n\|\|s_n-v_n\|\\ &\leq \frac{\upsilon}{2}(\|y_n-v_n\|^2+\|s_n-v_n\|^2)\\ &\leq \frac{\upsilon}{\sigma}(D_f(v_n,y_n)+D_f(s_n,v_n)).\end{aligned} \tag{3.17}$$

After combining (3.16) with (3.17), one concludes that

$$D_f(x,s_n) \leq D_f(x,y_n) - \left(1-\frac{\upsilon}{\sigma}\right)D_f(v_n,y_n) - \left(1-\frac{\upsilon}{\sigma}\right)D_f(s_n,v_n). \tag{3.18}$$

By using the three point identity of the Bregman distance, one infers that

$$\begin{aligned}\langle A^*(\nabla g(t_n)-\nabla g(A(x_n))), x-x_n\rangle &= \langle \nabla g(t_n)-\nabla g(A(x_n)), A(x)-A(x_n)\rangle\\ &= D_g(A(x),A(x_n)) - D_g(A(x),t_n) + D_g(A(x_n),t_n),\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}&\langle A^*(\nabla g(t_n)-\nabla g(A(x_n))), x_n-y_n\rangle\\ &= \langle \nabla g(t_n)-\nabla g(A(x_n)), A(x_n)-A(y_n)\rangle\\ &= -[(D_g(A(y_n),A(x_n)) - D_g(A(y_n),t_n) + D_g(A(x_n),t_n))].\end{aligned} \tag{3.20}$$

By combining (3.19) with (3.20), one obtains that

$$\begin{aligned}&\langle A^*(\nabla g(t_n)-\nabla g(A(x_n))), x-y_n\rangle\\ &= D_g(A(x),A(x_n)) - D_g(A(x),t_n) + D_g(A(x_n),t_n)\\ &\quad - [(D_g(A(y_n),A(x_n)) - D_g(A(y_n),t_n) + D_g(A(x_n),t_n))]\\ &\geq D_g(A(x),A(x_n)) - D_g(A(x),t_n) + D_g(A(y_n),t_n)) - D_g(A(y_n),A(x_n))\\ &\geq D_g(A(x),A(x_n)) - D_g(A(x),t_n) - D_g(A(y_n),A(x_n)).\end{aligned} \tag{3.21}$$

By recalling that $\alpha \in (0, \chi/\eta)$ and invoking (3.2), one finds that

$$\alpha_n D_g(A(y_n),A(x_n)) \leq \frac{\chi}{\alpha}\alpha_n\eta^{-1}D_g(A(y_n),A(x_n)) \leq \chi D_f(y_n,x_n). \tag{3.22}$$

By using the three point identity of the Bregman distance, (3.11), (3.21), and (3.22), one sees that

$$\begin{aligned}D_f(x,y_n) &= D_f(x,x_n) - D_f(y_n,x_n) - \alpha_n\langle A^*(\nabla g(t_n)-\nabla g(A(x_n))), x-y_n\rangle\\ &\leq D_f(x,x_n) - D_f(y_n,x_n) - \alpha_n[D_g(A(x),A(x_n)) - D_g(A(x),t_n)] + \alpha_n D_g(A(y_n),A(x_n))\\ &\leq D_f(x,x_n) - (1-\chi)D_f(y_n,x_n) - \alpha_n[D_g(A(x),A(x_n)) - D_g(A(x),t_n)]\\ &\leq D_f(x,x_n) - (1-\chi)D_f(y_n,x_n).\end{aligned} \tag{3.23}$$

Combining (3.18) with (3.23) shows that

$$\begin{aligned}D_f(x,x_{n+1}) &\leq \lambda_n D_f(x,x_1) + (1-\lambda_n)D_f(x,s_n)\\ &\leq \lambda_n D_f(x,x_1) + (1-\lambda_n)\left[D_f(x,x_n) - (1-\chi)D_f(y_n,x_n)\right.\\ &\quad \left. - \left(1-\frac{\upsilon}{\sigma}\right)D_f(v_n,y_n) - \left(1-\frac{\upsilon}{\sigma}\right)D_f(s_n,v_n)\right].\end{aligned}$$

This completes the proof of this lemma. □

Now, we are in a position to prove the strong convergence theorem of Algorithm 1.

Theorem 3.1. *Let Assumption 3.1* (A1)*-(*A8*) be satisfied. Then the sequence* $\{x_n\}$ *generated by Algorithm 1 converges strongly to* $\hat{x} = \Pi_{\Gamma}^{f}(x_1)$.

Proof. By using (3.4) in Lemma 3.3, and invoking that $\alpha \in (0,1)$ and $\upsilon \in (0,\sigma)$, we obtain that

$$\begin{aligned} D_f(x,x_{n+1}) \leq &\lambda_n D_f(x,x_1) + (1-\lambda_n)D_f(x,x_n) \leq \max\{D_f(x,x_1),D_f(x,x_n))\} \\ \leq &\max\{D_f(x,x_1),D_f(x,x_{n-1}))\} \leq \cdots \leq D_f(x,x_1). \end{aligned}$$

Hence, by induction we obtain that $D_f(x,x_n) \leq D_f(x,x_1)$ for all $n \in \mathbb{N}$. Therefore, we conclude that $\{D_f(x,x_n)\}$ is a bounded sequence, which implies that $\{x_n\}$ is bounded too. By using (3.23) in Lemma 3.3, one finds that $\{y_n\}$ is also bounded. Let $\hat{x} = \Pi_{\Gamma}^{f}(x_1)$. By replacing x by $\hat{x}$ in (3.4), one obtains that

$$\begin{aligned} &(1-\lambda_n)[(1-\chi)D_f(y_n,x_n) + \left(1-\frac{\upsilon}{\sigma}\right)D_f(v_n,y_n) + \left(1-\frac{\upsilon}{\sigma}\right)D_f(s_n,v_n)] \\ \leq &\lambda_n D_f(\hat{x},x_1) + (1-\lambda_n)[D_f(\hat{x},x_n) - D_f(\hat{x},x_{n+1})] - \lambda_n D_f(\hat{x},x_{n+1}) \\ \leq &\lambda_n D_f(\hat{x},x_1) + (1-\lambda_n)[D_f(\hat{x},x_n) - D_f(\hat{x},x_{n+1})]. \end{aligned} \tag{3.24}$$

Now, one considers the following two possible cases to prove $\lim_{n\to\infty} D_f(\hat{x},x_n) = 0$.

Case 1. There exists $N \in \mathbb{N}$ such that $D_f(\hat{x},x_{n+1}) \leq D_f(\hat{x},x_n)$ for all $n \geq N$, which gives that $\{D_f(\hat{x},x_n)\}$ is convergent and

$$\lim_{n\to\infty}(D_f(\hat{x},x_n) - D_f(\hat{x},x_{n+1})) = 0. \tag{3.25}$$

By putting together (3.24), (3.25), and the condition of $\chi \in (0,1)$, one concludes that

$$\lim_{n\to\infty} D_f(y_n,x_n) = \lim_{n\to\infty} D_f(v_n,y_n) = \lim_{n\to\infty} D_f(s_n,v_n) = 0. \tag{3.26}$$

It follows that

$$\lim_{n\to\infty} \|\nabla f(y_n) - \nabla f(x_n)\| = 0. \tag{3.27}$$

Based on (3.23), it holds that

$$\alpha_n[D_g(A(x),A(x_n)) - D_g(A(x),t_n)] \leq D_f(x,x_n) - D_f(x,y_n) - (1-\chi)D_f(y_n,x_n). \tag{3.28}$$

By using the three point identity of the Bregman distance, we find that

$$D_f(x,y_n) - D_f(x,x_n) = -D_f(y_n,x_n) + \langle \nabla f(x_n) - \nabla f(y_n), x - y_n\rangle. \tag{3.29}$$

We consider two possible cases in proving that

$$\lim_{n\to\infty}(D_g(A(x),A(x_n)) - D_g(A(x),t_n)) = 0. \tag{3.30}$$

(i) *First, we consider the case that* $D_g(A(y_n),A(x_n)) \neq 0$. In this case, we see that $A(y_n) \neq A(x_n)$. By invoking that A is a linear operator, we further find that $y_n \neq x_n$. Thus $D_f(y_n,x_n) \neq 0$. Moreover, (3.2) yields that $0 < \frac{1}{\alpha_n} \leq \frac{D_g(A(y_n),A(x_n))}{\alpha D_f(y_n,x_n)}$, when $D_g(A(y_n),A(x_n)) \neq 0$. By recalling Assumption 3.1 ($A2$), we obtain that ∇f is Lipschitz continuous with $\frac{1}{\sigma}$, the constant. This together with the relation $D_f(x,y) \geq \frac{\sigma}{2}\|x-y\|^2$ $(\forall x \in dom f, y \in int(dom f))$, (3.11), (3.28), and

(3.29), yields that

$$\begin{aligned}0 \leq & D_g(A(x),A(x_n)) - D_g(A(x),t_n)\\ \leq & \frac{D_g(A(y_n),A(x_n))}{\alpha D_f(y_n,x_n)}\left[D_f(x,x_n) - D_f(x,y_n) - (1-\chi)D_f(y_n,x_n)\right]\\ = & \frac{D_g(A(y_n),A(x_n))}{\alpha D_f(y_n,x_n)}[D_f(y_n,x_n) - \langle \nabla f(x_n) - \nabla f(y_n), x-y_n\rangle] - \frac{1-\chi}{\alpha}D_g(A(y_n),A(x_n))\\ \leq & \frac{D_g(A(y_n),A(x_n))}{\alpha} + \frac{D_g(A(y_n),A(x_n))}{\alpha D_f(y_n,x_n)}\|\nabla f(x_n) - \nabla f(y_n)\|\|x-y_n\| - \frac{1-\chi}{\alpha}D_g(A(y_n),A(x_n))\\ \leq & \frac{D_g(A(y_n),A(x_n))}{\alpha} + \frac{2D_g(A(y_n),A(x_n))}{\alpha\sigma^2}\|x-y_n\| - \frac{1-\chi}{\alpha}D_g(A(y_n),A(x_n)).\end{aligned} \tag{3.31}$$

It follows from (3.26) that $\lim_{n\to\infty}\|y_n - x_n\| = 0$. Since A is linear, we find that $\lim_{n\to\infty}\|A(y_n) - A(x_n)\| = 0$. Thus we further obtain that $\lim_{n\to\infty} D_g(A(y_n),A(x_n)) = 0$. This together with (3.11) in Lemma 3.3, (3.31), and the boundedness of $\{y_n\}$ yields that (3.30) holds.

(ii) *Second, we consider another case that* $D_g(A(y_n),A(x_n)) = 0$. In this case, (3.2) yields that $\alpha_n = \varpi$. Moreover, we obtain that $A(y_n) = A(x_n)$. Since A is linear, we further obtain that $D_f(y_n,x_n) = 0$. By using (3.28) and (3.29), we have

$$\begin{aligned}& D_g(A(x),A(x_n)) - D_g(A(x),t_n)\\ \leq & \frac{1}{\varpi}[D_f(y_n,x_n) - \langle\nabla f(x_n) - \nabla f(y_n), x-y_n\rangle - (1-\chi)D_f(y_n,x_n)]\\ \leq & \frac{1}{\varpi}[D_f(y_n,x_n) + \|\nabla f(x_n) - \nabla f(y_n)\|\|x-y_n\| - (1-\chi)D_f(y_n,x_n)].\end{aligned} \tag{3.32}$$

By using (3.11), (3.26), (3.27), (3.32), and the boundedness of $\{y_n\}$, we prove (3.30). In view of $\mu < \varsigma$, we find from (3.10) and (3.30) that $\lim_{n\to\infty} D_g(u_n, A(x_n)) = 0$, which together with (3.26) indicates that

$$\lim_{n\to\infty}\|y_n - x_n\| = \lim_{n\to\infty}\|v_n - y_n\| = \lim_{n\to\infty}\|s_n - v_n\| = \lim_{n\to\infty}\|u_n - A(x_n)\| = 0. \tag{3.33}$$

The boundedness of $\{x_n\}$ yields that $\{s_n\}$ is bounded, which together with Assumption 3.1 (*A2*) implies that $\{D_f(s_n,x_1)\}$ is bounded. Furthermore, we obtain that

$$D_f(s_n,x_{n+1}) \leq \lambda_n D_f(s_n,x_1) + (1-\lambda_n)D_f(s_n,s_n) = \lambda_n D_f(s_n,x_1). \tag{3.34}$$

Due to the facts that $\lim_{n\to\infty}\lambda_n = 0$ and $\{s_n\}$ is bounded, we have that $\lim_{n\to\infty} D_f(s_n,x_{n+1}) = 0$. By using the relation $D_f(x,y) \geq \frac{\sigma}{2}\|x-y\|^2$ $(\forall x \in dom f, y \in int(dom f))$, we obtain that $\lim_{n\to\infty}\|s_n - x_{n+1}\| = 0$, which together with (3.33) yields $\lim_{n\to\infty}\|x_n - x_{n+1}\| = 0$. Since $\{x_n\}$ is bounded, there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \rightharpoonup \tilde{x}$ as $k\to\infty$, which together with (3.33) yields that $\{y_n\}$ is bounded and $y_{n_k} \rightharpoonup \tilde{x}$ as $k\to\infty$. Assumption 3.1 (*A6*) implies that A is sequential weakly continuous. Thus we obtain that $A(x_{n_k}) \rightharpoonup A(\tilde{x})$ as $k\to\infty$.

Now, we show that $\tilde{x} \in S(C,F_1)$. According to (3.12), we find that

$$\langle F_1(y_{n_k}), x - y_{n_k}\rangle \geq -\frac{1}{\upsilon_n}\langle\nabla f(v_{n_k}) - \nabla f(y_{n_k}), x - v_{n_k}\rangle - \langle F_1(y_{n_k}), y_{n_k} - v_{n_k}\rangle. \tag{3.35}$$

We consider two possible cases in proving that

$$\liminf_{k\to\infty}\langle F_1(y_{n_k}), z - y_{n_k}\rangle \geq 0,\ \forall z \in C. \tag{3.36}$$

(i) *Suppose that* $\liminf_{k\to\infty} \upsilon_{n_k} > 0$. Since F_1 is uniformly continuous and $\{y_n\}$ is bounded, we see that Lemma 2.3 indicates that $\{F_1(y_{n_k})\}$ is bounded. Note that Assumption 3.1 (*A2*) implies that ∇f is uniformly continuous. Thus it follows from (3.33) that

$$\lim_{k\to\infty} \|\nabla f(v_{n_k}) - \nabla f(y_{n_k})\| = 0. \tag{3.37}$$

By taking the limit inferior as $k \to \infty$ in (3.35), we prove (3.36).

(ii) *Suppose that* $\liminf_{k\to\infty} \upsilon_{n_k} = 0$. Denote $p_{n_k} = \Pi_C^f \nabla f^*(\nabla f(y_{n_k}) - \upsilon_{n_k}\gamma^{-1}F_1(y_{n_k}))$, for all $k \geq 1$. Since $\gamma \in (0,1)$, we have $\upsilon_{n_k}\gamma^{-1} > \upsilon_{n_k}$. According to Lemma 2.4, we obtain that $\gamma^{-1}\|y_{n_k} - p_{n_k}\| \leq \|y_{n_k} - v_{n_k}\|$, which together with (3.33) yields that $\lim_{k\to\infty} \|y_{n_k} - p_{n_k}\| = 0$. Since F_1 is uniformly continuous, we find that $\lim_{k\to\infty} \|F_1(y_{n_k}) - F_1(p_{n_k})\| = 0$. By using (3.3), we obtain that $\upsilon_{n_k}\gamma^{-1}\|F_1(p_{n_k}) - F_1(y_{n_k})\| > \upsilon\|p_{n_k} - y_{n_k}\|$. Thus $\lim_{k\to\infty} \frac{\|p_{n_k} - y_{n_k}\|}{\upsilon_{n_k}\gamma^{-1}} = 0$. Thus

$$\lim_{k\to\infty} \frac{\|\nabla f(p_{n_k}) - \nabla f(y_{n_k})\|}{\upsilon_{n_k}\gamma^{-1}} = 0. \tag{3.38}$$

The definition of $\{p_{n_k}\}$ and Lemma 2.1 (i) indicate that

$$\langle \nabla f(p_{n_k}) - \nabla f(\nabla f^*(\nabla f(y_{n_k}) - \upsilon_{n_k}\gamma^{-1}F_1(y_{n_k}))), z - p_{n_k}\rangle \geq 0, \ \forall z \in C, \tag{3.39}$$

which in turn implies that

$$\langle F_1(y_{n_k}), z - y_{n_k}\rangle + \langle F_1(y_{n_k}), y_{n_k} - p_{n_k}\rangle + \left\langle \frac{\nabla f(p_{n_k}) - \nabla f(y_{n_k})}{\upsilon_{n_k}\gamma^{-1}}, z - p_{n_k}\right\rangle \geq 0, \ \forall z \in C. \tag{3.40}$$

Taking the limit inferior as $k \to \infty$ in (3.40), we obtain (3.36) by (3.38). Let $\{\chi_k\}$ be a strictly decreasing sequence of positive numbers such that $\chi_k \to 0$ as $k \to \infty$. For each $k \geq 1$, it follows from (3.36) that the smallest positive integer, denoted by $\phi(k)$, such that

$$\langle F_1(y_{n_i}), z - y_{n_i}\rangle + \chi_k \geq 0, \ \forall z \in C, \ \forall i \geq \phi(k) \tag{3.41}$$

exists. For each $k \geq 1$, suppose that $F_1(y_{n_{\phi(k)}}) \neq 0$ (otherwise, $y_{n_{\phi(k)}} \in S(C, F_1)$). By setting $q_k = \frac{F_1(y_{n_{\phi(k)}})}{\|F_1(y_{n_{\phi(k)}})\|^2}$, we see that $\langle F_1(y_{n_{\phi(k)}}), q_{n_{\phi(k)}}\rangle = 1$, which together with (3.41) yields that $\langle F_1(y_{n_{\phi(k)}}), z - y_{n_{\phi(k)}} + \chi_k q_k\rangle \geq 0$ for all $z \in C$. Since F_1 is pseudomonotone, then

$$\langle F_1(z + \chi_k q_k), z - y_{n_{\phi(k)}} + \chi_k q_k\rangle \geq 0, \ \forall z \in C. \tag{3.42}$$

In view of (3.42), we obtain that

$$\begin{aligned} &\langle F_1(z), z - y_{n_{\phi(k)}}\rangle \\ =&\langle F_1(z) - F_1(z + \chi_k q_k), z - y_{n_{\phi(k)}}\rangle + \langle F_1(z + \chi_k q_k), z + \chi_k q_k - y_{n_{\phi(k)}}\rangle - \langle F_1(z + \chi_k q_k), \chi_k q_k\rangle \\ \geq&\langle F_1(z) - F_1(z + \chi_k q_k), z - y_{n_{\phi(k)}}\rangle - \langle F_1(z + \chi_k q_k), \chi_k q_k\rangle. \end{aligned} \tag{3.43}$$

Now we demonstrate that $\lim_{k\to\infty} \chi_k q_k = 0$. By considering that $y_{n_{\phi(k)}} \rightharpoonup \tilde{x}$ as $k \to \infty$, it follows from Assumption 3.1 (*A4*) that $\|F_1(\tilde{x})\| \leq \liminf_{k\to\infty} \|F_1(y_{n_{\phi(k)}})\|$. We assume that $F_1(\tilde{x}) \neq 0$ (otherwise, $\tilde{x} \in S(C, F_1)$). Since $\chi_k \to 0$ as $k \to \infty$, we find that

$$0 \leq \limsup_{k\to\infty} \|\chi_k q_k\| = \limsup_{k\to\infty} \frac{\chi_k}{\|F_1(y_{n_{\phi(k)}})\|} \leq \frac{\limsup_{k\to\infty} \chi_k}{\liminf_{k\to\infty} \|F_1(y_{n_{\phi(k)}})\|} = 0.$$

Thus $\lim_{k\to\infty}\|\chi_k q_k\|=0$, which together with (3.42) and Assumption 3.1 (*A4*) implies that $\liminf_{k\to\infty}\langle F_1(z), z-y_{n_{\phi(k)}}\rangle \geq 0$ for all $z\in C$. It follows that

$$\langle F_1(z), z-\tilde{x}\rangle = \lim_{k\to\infty}\langle F_1(z), z-y_{n_{\phi(k)}}\rangle = \liminf_{k\to\infty}\langle F_1(z), z-y_{n_{\phi(k)}}\rangle \geq 0,\ \forall z\in C.$$

By using Lemma 2.5, we prove that $\tilde{x}\in S(C,F_1)$.

By following the similar proof used in obtaining $\tilde{x}\in S(C,F_1)$, we seet that $A\tilde{x}\in S(Q,F_2)$. Since $\hat{x}=\Pi_{\Gamma}^{f}(x_1)$, Lemma 2.1 (i) indicates that

$$\limsup_{n\to\infty}\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{n+1}-\hat{x}\rangle = \langle \nabla f(x_1)-\nabla f(\hat{x}), \tilde{x}-\hat{x}\rangle \leq 0. \tag{3.44}$$

By combining (3.16) with (3.23), we observe that $D_f(\hat{x},s_n)\leq D_f(\hat{x},x_n)$, which together with the property of $V_f(\cdot,\cdot)$ yields that

$$\begin{aligned}
D_f(\hat{x},x_{n+1}) =& V_f(\hat{x},\lambda_n\nabla f(x_1)+(1-\lambda_n)\nabla f(s_n))\\
\leq& V_f(\hat{x},\lambda_n\nabla f(x_1)+(1-\lambda_n)\nabla f(s_n)-\lambda_n(\nabla f(x_1)-\nabla f(\hat{x})))\\
&+\langle \lambda_n(\nabla f(x_1)-\nabla f(\hat{x})), \nabla f^*(\lambda_n\nabla f(x_1)+(1-\lambda_n)\nabla f(s_n))-\hat{x}\rangle\\
=& V_f(\hat{x},\lambda_n\nabla f(\hat{x})+(1-\lambda_n)\nabla f(s_n))+\langle \lambda_n(\nabla f(x_1)-\nabla f(\hat{x})), x_{n+1}-\hat{x}\rangle\\
=& D_f(\hat{x},\nabla f^*(\lambda_n\nabla f(\hat{x})+(1-\lambda_n)\nabla f(s_n)))+\langle \lambda_n(\nabla f(x_1)-\nabla f(\hat{x})), x_{n+1}-\hat{x}\rangle\\
\leq& (1-\lambda_n)D_f(\hat{x},s_n)+\lambda_n\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{n+1}-\hat{x}\rangle\\
\leq& (1-\lambda_n)D_f(\hat{x},x_n)+\lambda_n\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{n+1}-\hat{x}\rangle.
\end{aligned} \tag{3.45}$$

Using Lemma 2.6, (3.44), and (3.45), we obtain that $\lim_{n\to\infty}D_f(\hat{x},x_n)=0$, which together with the relation $D_f(x,y)\geq \frac{\sigma}{2}\|x-y\|^2$ $(\forall x\in dom f, y\in int(dom f))$ yields that $\lim_{n\to\infty}\|\hat{x}-x_n\|=0$. Then $\lim_{n\to\infty}x_n=\hat{x}$.

Case 2. There exists a subsequence $\{D_f(\hat{x},x_{n_m})\}$ of $\{D_f(\hat{x},x_n)\}$ such that $D_f(\hat{x},x_{n_m})\leq D_f(\hat{x},x_{n_m+1})$ for all $m\in\mathbb{N}$. By applying Lemma 2.7, we see that there exists an increasing sequence $\{\varphi(m)\}\subset\mathbb{N}$ such that $\lim_{m\to\infty}\varphi(m)=\infty$ and the following inequalities hold, for any $m\in\mathbb{N}$, $D_f(\hat{x},x_{\varphi(m)})\leq D_f(\hat{x},x_{\varphi(m)+1})$ and $D_f(\hat{x},x_m)\leq D_f(\hat{x},x_{\varphi(m)+1})$. In view of (3.24), we obtain that

$$\begin{aligned}
&(1-\lambda_{\varphi(m)})[(1-\chi)D_f(y_{\varphi(m)},x_{\varphi(m)})+\left(1-\frac{\upsilon}{\sigma}\right)D_f(v_{\varphi(m)},y_{\varphi(m)})+\left(1-\frac{\upsilon}{\sigma}\right)D_f(s_{\varphi(m)},v_{\varphi(m)})]\\
&\leq \lambda_{\varphi(m)}D_f(\hat{x},x_1)+(1-\lambda_{\varphi(m)})[D_f(\hat{x},x_{\varphi(m)})-D_f(\hat{x},x_{\varphi(m)+1})].
\end{aligned} \tag{3.46}$$

It follows from (3.45) that

$$D_f(\hat{x},x_{\varphi(m)+1}) \leq (1-\lambda_{\varphi(m)})D_f(\hat{x},x_{\varphi(m)})+\lambda_{\varphi(m)}\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{\varphi(m)+1}-\hat{x}\rangle. \tag{3.47}$$

According to (3.46), Assumption 3.1 (*A7*), $\alpha\in(0,1)$, and $\upsilon\in(0,\sigma)$, we find that

$$\lim_{m\to\infty}D_f(y_{\varphi(m)},x_{\varphi(m)})=\lim_{m\to\infty}D_f(v_{\varphi(m)},y_{\varphi(m)})=\lim_{m\to\infty}D_f(s_{\varphi(m)},v_{\varphi(m)})=0.$$

By repeating the same arguments as in the proof of ***Case 1***, we conclude that

$$\limsup_{m\to\infty}\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{\varphi(m)}-\hat{x}\rangle \leq 0. \tag{3.48}$$

By (3.47), we arrive at

$$D_f(\hat{x},x_{\varphi(m)+1}) \leq (1-\lambda_{\varphi(m)})D_f(\hat{x},x_{\varphi(m)+1})+\lambda_{\varphi(m)}\langle \nabla f(x_1)-\nabla f(\hat{x}), x_{\varphi(m)+1}-\hat{x}\rangle.$$

It follows that

$$D_f(\hat{x},x_m) \le D_f(\hat{x},x_{\varphi(m)+1}) \le \langle \nabla f(x_1) - \nabla f(\hat{x}), x_{\varphi(m)+1} - \hat{x} \rangle. \tag{3.49}$$

In view of (3.48) and (3.49), we conclude that $\limsup_{m\to\infty} D_f(\hat{x},x_m) = 0$, which together with the relation $D_f(x,y) \ge \frac{\sigma}{2}\|x-y\|^2$ ($\forall x \in dom f, y \in int(dom f)$) yields that $\lim_{m\to\infty} \|\hat{x} - x_m\| = 0$. Then $\lim_{m\to\infty} x_m = \hat{x}$. From ***Case 1*** and ***Case 2***, we conclude that $\{x_n\}$ converges strongly to $\hat{x} = \Pi_\Gamma^f(x_1)$, which completes the proof. □

Remark 3.2. By setting $f(x) = \frac{1}{2}\|x\|^2$ for all $x \in H_1$ and $g(y) = \frac{1}{2}\|y\|^2$ for all $y \in H_2$, we obtain a special case of Algorithm 1. Given parameters $\chi \in (0,1)$, $\varpi \in (0,\infty)$, $\eta \in (1,\infty)$, $\alpha \in (0,\chi/\eta)$, and $x_1 \in H_1$, the iterative sequence $\{x_n\}$ is generated by the following:

$$\begin{cases} u_n = P_Q(A(x_n) - \mu_n F_2(A(x_n))), \text{where } \mu_n \text{ is defined by (3.1)}, \\ t_n = u_n - \mu_n(F_2(u_n) - F_2(A(x_n))), \\ y_n = x_n + \alpha_n A^*(t_n - A(x_n)), \text{where } \alpha_n \text{ is defined by } \alpha_n = \varpi\eta^{l_n} \text{ with} \\ l_n \text{ being the smallest integer } l \text{ satisfying } \alpha\|y_n - x_n\| \le \alpha_n\|A(y_n) - A(x_n)\|, \\ v_n = P_C(y_n - \upsilon_n F_1(y_n)), \\ s_n = v_n - \upsilon_n(F_1(v_n) - F_1(y_n)), \text{where } \upsilon_n \text{ is defined by (3.3)}, \\ x_{n+1} = \lambda_n x_1 + (1-\lambda_n)s_n. \end{cases}$$

Remark 3.3. By setting $F_1(x) = 0$ for all $x \in H_1$ in (1.1) and $F_2(y) = 0$ for all $y \in H_2$ in (1.2), we see that the SVI is reduced to a split feasibility problem. By setting $g(y) = \frac{1}{2}\|y\|^2$ for all $y \in H_2$, then Bregman projection Π_Q^g is reduced to the metric projection P_Q. In this situation, by picking $\lambda_n = 0$ for all $n \ge 1$, Algorithm 1 is reduced to the method proposed in [37, Algorithm 3.1].

4. Numerical Examples

In this section, we present two test examples to illustrate the advantage and the efficiency of our proposed algorithm in comparison with the methods proposed in [15, 19]. In the following experiments, we define $Err_n := \frac{1}{2}(\|x_n - P_C(x_n - F_1(x_n))\|^2 + \|A(x_n) - P_Q(A(x_n) - F_2(A(x_n)))\|^2)$ for all $n \in \mathbb{N}$. We use the stopping criterion $Err_n < \varepsilon$ for the iterative process, where ε is the predetermined error. If $Err_n = 0$, then $x_n \in \Gamma$. All codes are written in Python 3.9 on a PC Desktop Intel(R) Core(TM) i5-11300H @ 3.10 GHz(8 CPUs), 3.1 GHz, RAM 16 384 MB.

We first consider an example in finite dimensional spaces.

Example 4.1. Let $H_1 = \mathbb{R}^4$ and $H_2 = \mathbb{R}^2$. Let $A(x) = (x_1 + x_3 + x_4, x_2 + x_3 - x_4)^T$ for all $x = (x_1,x_2,x_3,x_4)^T \in \mathbb{R}^4$. Hence, A is a bounded linear operator from $\mathbb{R}^4$ into $\mathbb{R}^2$. Let $B(y) = (y_1, y_2, y_1 + y_2, y_1 - y_2)^T$ for all $y = (y_1,y_2)^T \in \mathbb{R}^2$. Since $\langle A(x), y\rangle = \langle x, B(y)\rangle$ for all $x = (x_1,x_2,x_3,x_4)^T \in \mathbb{R}^4$ and $y = (y_1,y_2)^T \in \mathbb{R}^2$. Thus $B = A^*$ is an adjoint operator of A. Furthermore, we find that $\|A\| = \|A^*\| = \sqrt{3}$. Let $C = \{(x_1,x_2,x_3,x_4)^T \in \mathbb{R}^4 | x_1 - x_2 - x_3 + 2x_4 \ge -1\}$ and $Q = \{(x_1,x_2)^T \in \mathbb{R}^2 | 2x_1 - 3x_2 \ge -4\}$. Define an operator $F_1 : \mathbb{R}^4 \to \mathbb{R}^4$ by $F_1(x) = (\sin\|x\| + 2)p_0$ for all $x \in \mathbb{R}^4$, where $p_0 = (1,-1,-1,2)^T \in \mathbb{R}^4$. Define another operator $F_2 : \mathbb{R}^2 \to \mathbb{R}^2$ by $F_2(y) = (\sin\|y\| + 3)q_0$, $\forall y \in \mathbb{R}^2$, where $q_0 = (2,-3)^T \in \mathbb{R}^2$. It is noted that F_1 satisfies the condition $(A4)$ in Assumption 3.1 on $\mathbb{R}^4$. Moreover, F_2 satisfies the condition $(A5)$ in Assumption 3.1 on $\mathbb{R}^2$; see [18]. It is not difficult to check that F_1 is $\sqrt{7}$-Lipschitz continuous and F_2 is $\sqrt{13}$-Lipschitz continuous. Let $f(x) = \frac{1}{2}\|x\|^2, \forall x \in \mathbb{R}^4$ and $g(y) = \frac{1}{2}\|y\|^2, \forall y \in \mathbb{R}^2$. It is clear that f and g are strongly convex with modulus 1.

In this experiment, we give a numerical comparison of Algorithm 1 with Method-MPC of Ogwo et al. [19, Algorithm 1] and Method-MHSE of Huy et al. [15, Algorithm 1]. We test these methods by using the following parameters:

(i) For Algorithm 1, parameters $\kappa, \iota, \varpi, \chi, \rho, \gamma, \mu$, and υ are generated randomly in $(0,1)$. $\eta \in (1,\infty)$ is generated randomly in $(1,10)$. α is generated randomly in $(0,\chi/\eta)$. We define $\lambda_n := \frac{1}{n+1}$ for all $n \in \mathbb{N}$.

(ii) For Method-MPC, we select $\alpha := 4$ and $a := 10^{-4}$ and define $\delta_n := \frac{1}{n+1}$ and $\tau_n := \frac{1}{(n+1)^2}$ for all $n \in \mathbb{N}$. We randomly choose λ in $(0,1/L_2)$, μ in $(0,1/L_1)$, γ_1 in $(0,2)$, and γ_2 in $(0,2)$.

(iii) For Method-MHSE, we randomly generate μ_0, λ_0, μ, and λ in $(0,1)$. For each $n \in \mathbb{N}$, we randomly choose δ_n in $(0.1,0.2)$. We define $\alpha_n := \frac{1}{n+1}, \forall n \in \mathbb{N}$.

For comparing Algorithm 1 with Method-MPC and Method-MHSE, we choose the same initials with the entries being randomly generated in the range of $(0,1)$. For all algorithms, we take $Err_n < \varepsilon = 5 \times 10^4$ as a common stopping criterion, which serves as the role of checking whether the algorithms converge to the solution of the SVI or not. The corresponding numerical results are reported in Figures 1, 2, and Table 1.

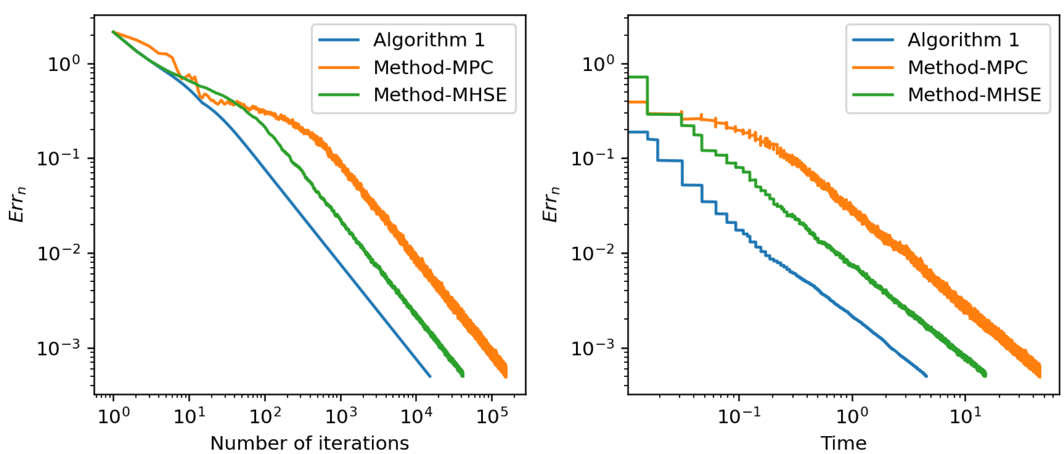

FIGURE 1. Numerical results for Algorithm 1, Method-MPC, and Method-MHSE.

TABLE 1. Comparison of the number of termination iterations and the execution time between Algorithm 1, Method-MPC and Method-MHSE.

Method	Algorithm 1	Method-MPC	Method-MHSE
Iterations	**15250**	153994	41628
CPU Time	**4.5494**	45.7233	15.0631

From the changing processes of the values of $\{Err_n\}$ in Figure 1, we observe that our algorithm has a faster convergence speed than Method-MPC and Method-MHSE in terms of both the number of iterations and the execution time in second elapses. This is due to the fact that our proposed algorithm uses Armijo-type stepsize criterions which finds appropriate step sizes per iteration and thus decreases the CPU time and the number of iterations of Algorithm 1. Furthermore, it can also be seen from Figure 1 that our proposed algorithm achieves a more stable and higher precision with the number of iterations. Besides, the convergence of $\{Err_n\}$ to 0 implies that the iterative sequences converge to the solution of the SVI. The numerical results summarized in Table 1 illustrate that Algorithm 1 requires fewer iterations and less execution time than Method-MPC and Method-MHSE to achieve the same error accuracy.

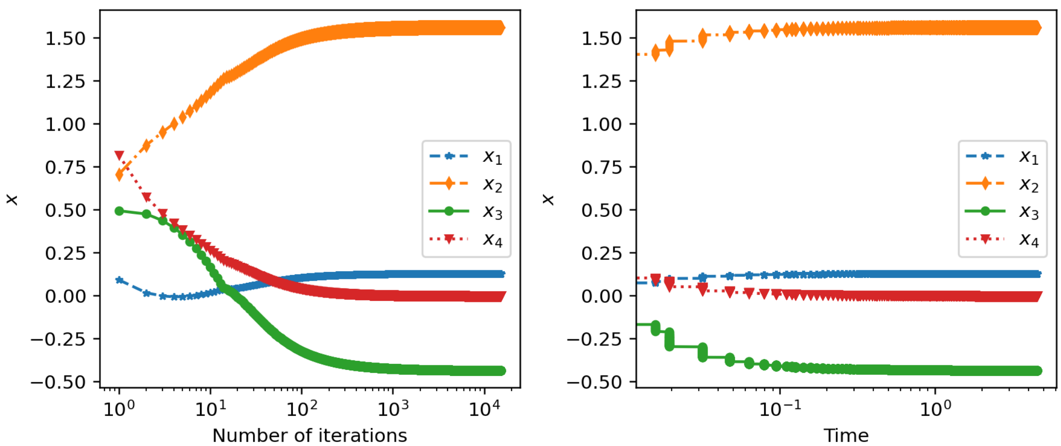

FIGURE 2. The behavior of elements of x of Example 4.1.

Figure 2 plots the changing processes of elements of x. Furthermore, the convergent point $x = (0.1284, 1.5612, -0.4358, -0.0014)^T$ is the solution of the SVI.

Example 4.2. Let $H_1 = H_2 = \ell_2(\mathbb{R})$ whose elements are square summable sequences, that is,

$$\ell_2(\mathbb{R}) = \{x = (x_1, x_2, \cdots, x_i, \cdots) | \sum_{i=1}^{\infty} |x_i|^2 < \infty\}.$$

The inner product $\langle \cdot, \cdot \rangle : \ell_2(\mathbb{R}) \times \ell_2(\mathbb{R}) \to \mathbb{R}$ and the norm $\|\cdot\| : \ell_2(\mathbb{R}) \to \mathbb{R}$ are respectively defined by

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i y_i, \ \forall x = (x_1, x_2, \cdots, x_i, \cdots), \ y = (y_1, y_2, \cdots, y_i, \cdots) \in \ell_2(\mathbb{R}),$$

and

$$\|x\| = \sqrt{\langle x, x \rangle}, \ \forall x = (x_1, x_2, \cdots, x_i, \cdots) \in \ell_2(\mathbb{R}).$$

We define $C = \{x \in \ell_2(\mathbb{R}) | \|x - y\| \le a\}$, where $y = (1, \frac{1}{2}, \cdots, \frac{1}{i}, \cdots)$ and $a = 3$. Define $Q = \{x \in \ell_2(\mathbb{R}) | \langle z, x \rangle \le b\}$, where $z = (1, 2, \cdots, i, \cdots)$ and $b = 4$. Therefore, C and Q are nonempty, convex, and closed subsets of $\ell_2(\mathbb{R})$. Based on [38], the explicit formulas for projections onto P_C and P_Q are defined by

$$P_C(x) = \begin{cases} \frac{x-y}{\|x-y\|} a + y, & \text{if } \|x - y\| > a, \\ x, & \text{otherwise}, \end{cases} \quad \text{and} \quad P_Q(x) = \begin{cases} \frac{b - \langle z, x \rangle}{\|z\|^2} z + x, & \text{if } \langle z, x \rangle > b, \\ x, & \text{otherwise}. \end{cases}$$

Let $A(x) = (0, x_1, \frac{x_2}{2}, \frac{x_3}{3}, \cdots)$ for all $x = (x_1, x_2, x_3, \cdots) \in \ell_2(\mathbb{R})$. Note that A is a bounded and linear operator on $\ell_2(\mathbb{R})$ with an adjoint operator $A^* y = (y_2, \frac{y_3}{2}, \frac{y_4}{3}, \cdots)$ for all $y = (y_1, y_2, y_3, \cdots) \in$

$\ell_2(\mathbb{R})$; see [19]. Define an operator $F_1 : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ by $F_1(x) = \left(\frac{1}{\|x\|+\theta}\|x\|\right)x$ for all $x \in \ell_2(\mathbb{R})$, where $\theta > 0$. Define another operator $F_2 : \ell_2(\mathbb{R}) \to \ell_2(\mathbb{R})$ by $F_2(x_1, x_2, x_3, \cdots) = (x_1 e^{-x_1^2}, 0, 0, \cdots)$ for all $x \in \ell_2(\mathbb{R})$. It is noted that F_1 and F_2 satisfy the conditions $(A4)$ and $(A5)$ in Assumption 3.1, respectively; see [19]. We define $f(x) = g(x) = \frac{1}{2}\|x\|^2$ for all $x \in \ell_2(\mathbb{R})$. Hence, f and g are strongly convex with modulus 1. It is clear that $\nabla f(x) = \nabla g(x) = x$ and $\nabla f^*(x) = \nabla g^*(x) = x$ for all $x \in \ell_2(\mathbb{R})$. Thus the corresponding Bregman distance is given by $D_f(x,y) = D_g(x,y) = \frac{1}{2}\|x-y\|^2$ for all $x, y \in \ell_2(\mathbb{R})$. In this experiment, we choose different starting points as follows

Case I: $x_1 = (0.5260, 0.8047, 0.7059, 0.8023, 0.6221, 0.1461, 0.5222, 0.3694, 0, \cdots, 0, \cdots)$;

Case II: $x_1 = (0.1711, 0.6004, 0.2114, 0.7504, 0, \cdots, 0, \cdots)$.

We take $Err_n < \varepsilon = 10^{-4}$ as the stopping criterion for the iterative process. It is noted that $\|x_n - x_{n+1}\|$ can be used to measure the error of the n-th iteration step of Algorithm 1. Numerical behaviors of our algorithm with different starting points are shown in Figures 3-6.

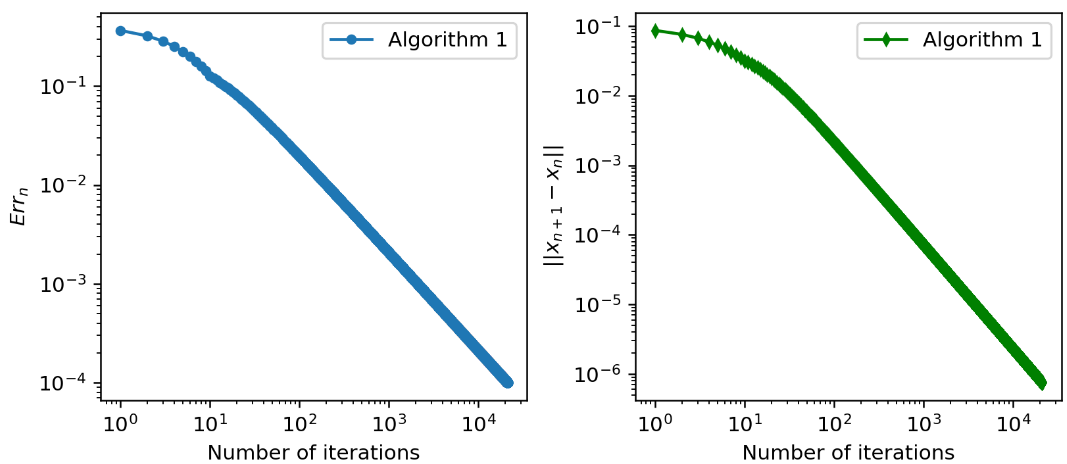

FIGURE 3. Numerical results for ***Case I*** of Example 4.2.

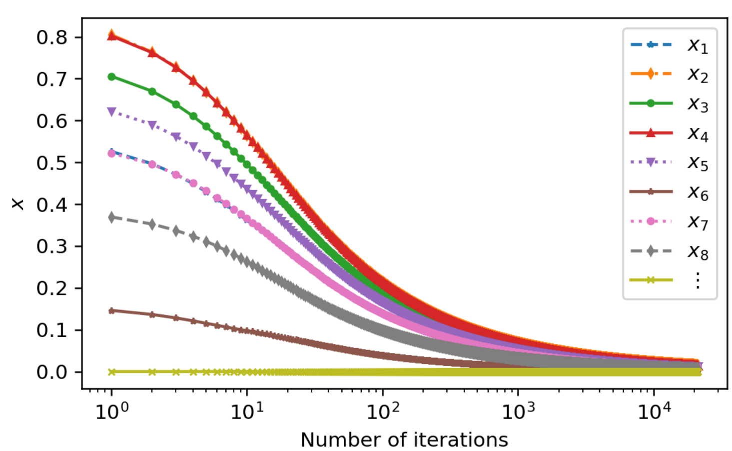

FIGURE 4. Behavior of entries of x with the number of iterations for ***Case I***.

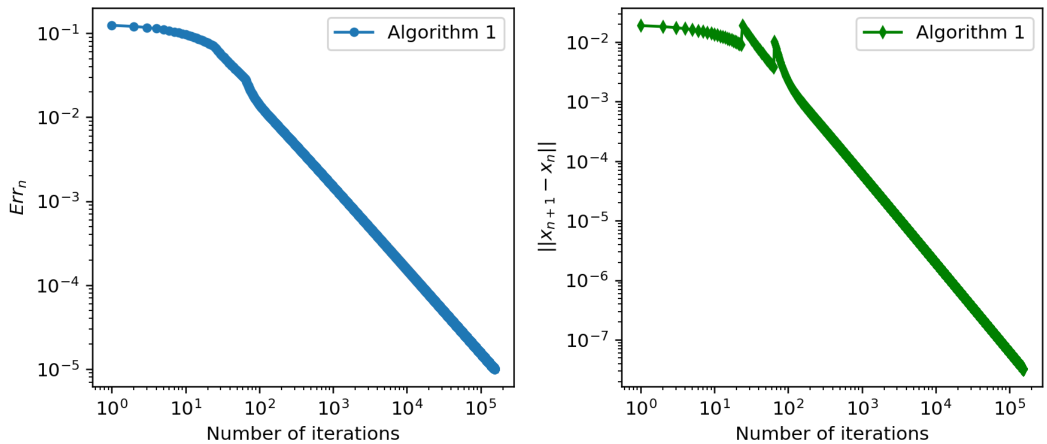

FIGURE 5. Numerical results for ***Case II*** of Example 4.2.

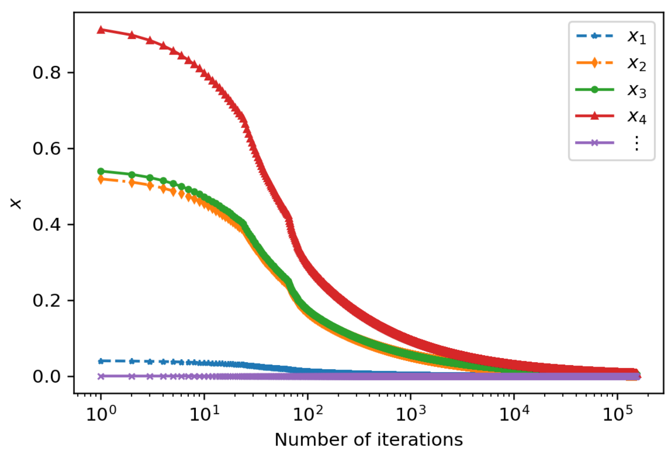

FIGURE 6. Behavior of entries of x with the number of iterations for ***Case II***.

In Figures 3 and 5, the left figure plots the changing process of the value of $\{Err_n\}$ with the number of iterations and the right figure plots the changing process of the value of $\{\|x_{n+1} - x_n\|\}$ with the number of iterations. From above, we observe that the convergence of $\{Err_n\}$ and $\{\|x_{n+1} - x_n\|\}$ to 0 implies that the iterative sequence converge to the solution of the SVI. Our proposed algorithm is efficient to implement for solving the SVI.

The numerical performance of x shown in Figure 4 implies that the convergence point in ***Case I*** is $(0.0098, 0.0150, 0.0131, 0.0149, 0.0116, 0.0027, 0.0097, 0.0068, 0, \cdots, 0, \cdots)$.

The numerical performance of x shown in Figure 6 implies that the convergence point in ***Case II*** is $(0.0003, 0.0043, 0.0045, 0.0077, 0, \cdots, 0, \cdots)$.

Remark 4.1. We have the following observations for Examples 4.1 and 4.2.

(i) The nonmonotonic step-size criterions utilized in our algorithm make it more efficient and more implementable than the method presented in [15] that uses a non-increasing step size criterion and the method presented in [19] that uses a fixed step-size criterion.

(ii) The observations have no significant relationship with the selection of initial points.

5. Conclusion

In this paper, the proposed algorithm for solving SVIs was motivated by the Halpern method, Tseng's extragradient method and the CQ algorithm. We employed different nonmonotonic step size criterions that allows the algorithm to work adaptively without the prior knowledge of the operator norm and operators' Lipschitz constants. The operators involved in the SVI are pseudomonotone and not necessarily Lipschitz continuous. The strong convergence theorem of the suggested method was established under mild conditions. Two numerical experiments were performed to verify theoretical results. Numerical results show the competitive advantage and the computational efficiency of the suggested algorithm over other methods. The obtained result improves and extends some related works in the literature.

Acknowledgements
This work was supported by National Natural Science Foundation of China (Grant No. 12201517), the Natural Science Foundation of Chongqing (Grant No. CSTB2022NSCQ-MSX0580) and the Fundamental Research Funds for the Central Universities (Grant No. SWU-KQ22014).

References

[1] G. Stampacchi, Formes bilineaires coercivites sur les ensembles convexes, C. R. Acad. Sci. Paris, 258 (1964), 4413-4416.
[2] W. Yata, M. Yamagishi, I. Yamada, A constrained LiGME model and its proximal splitting algorithm under overall convexity condition, J. Appl. Numer. Optim. 4 (2022), 245-271.
[3] G. Kotsalis, G. Lan, T. Li, Simple and optimal methods for stochastic variational inequalities, I: operator extrapolation, SIAM J. Optim. 32 (2022), 2041-2073.
[4] L.V. Nguyen, Q.H. Ansari, X. Qin, Weak sharpness and finite convergence for solutions of nonsmooth variational inequalities in Hilbert spaces, Appl. Math. Optim. 84 (2021), 807-828.
[5] T. Huang, et al., A CNN-based policy for optimizing continuous action control by learning state sequences, Neurocomputing, 468 (2022), 286-295.
[6] Y. Censor, T. Elfving, A multiprojection algorithm using Bregman projections in a product space, Numer. Algorithms 8 (1994), 221-239.
[7] C. Byrne, Iterative oblique projection onto convex sets and the split feasibility problem, Inverse Probl. 18 (2002), 441–453.
[8] Y. Censor, A. Gibali, S. Reich, Algorithms for the split variational inequality problem, Numer. Algorithms 59 (2012), 301-323.
[9] C. Izuchukwu, A. A. Mebawondu, O. T. Mewomo, A new method for solving split variational inequality problems without co-coerciveness, J. Fixed Point Theory Appl. 22 (2002), 98.
[10] N. T. T. Thuy, T. T. Tung, A self adaptive inertial algorithm for solving variational inequalities over the solution set of the split variational inequality problem, Optim. Lett. doi: 10.1007/s11590-023-02080-y.
[11] G. M. Korpelevich, The extragradient method for finding saddle points and other problems, Matecon 12 (1976), 747-756.
[12] Y. Censor, A. Gibali, S. Reich, The subgradient extragradient method for solving variational inequalities in Hilbert space, J. Optim. Theory Appl. 148 (2021), 318-335.

[13] P. Tseng, A modifed forward-backward splitting method for maximal monotone mappings, SIAM J Control Optim. 38 (2000), 431-446.
[14] H. He, C. Ling, H. K. Xu, A relaxed projection method for split variational inequalities, J. Optim. Theory Appl. 166 (2015), 213–233.
[15] P. V. Huy, N.D. Hien, T.V. Anh, A strongly convergent modified Halpern subgradient extragradient method for solving the split variational inequality problem, Vietna J. Math. 48 (2020), 187-204.
[16] L.C. Ceng, et al., A modified inertial subgradient extragradient method for solving pseudomonotone variational inequalities and common fixed point problems, Fixed Point Theory 21 (2020), 93-108.
[17] B. Tan, X. Qin, J.C. Yao, Strong convergence of self-adaptive inertial algorithms for solving split variational inclusion problems with applications, J. Sci. Computing 87 (2021), 20.
[18] P. V. Huy, et al., Modified Tseng's extragradient methods with self-adaptive step size for solving bilevel split variational inequality problems, Optimization 71 (2022), 1721-1748.
[19] G. N. Ogwo, C. Izuchukwu, O. T. Mewomo, Inertial methods for finding minimum-norm solutions of the split variational inequality problem beyond monotonicity, Numer. Algorithms 88 (2021), 1419-1456.
[20] T. O. Alakoya, V. A. Uzor, O. T. Mewomo, A new projection and contraction method for solving split monotone variational inclusion, pseudomonotone variational inequality, and common fixed point problems, Comput. Appl. Math. 42 (2023), 3.
[21] D. V. Hieu, S. Reich, Two Bregman projection methods for solving variational inequalities, Optimization 71 (2022), 1777-1802.
[22] L. O. Jolaoso, P. Sunthrayuth, P. Cholamjiak, et al., Analysis of two versions of relaxed inertial algorithms with Bregman divergences for solving variational inequalities, Comput. Appl. Math. 41 (2022), 300.
[23] A. Gibali, A new Bregman projection method for solving variational inequalities in Hilbert spaces, Pure Appl. Funct. Anal. 3 (2018), 403-415.
[24] P. Sunthrayuth, L. O. Jolaoso, P. Cholamjiak, New Bregman projection methods for solving pseudo-monotone variational inequality problem, J. Appl. Math. Comput. 68 (2022), 1565-1589.
[25] S. Reich, S. Sabach, A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces, J. Nonlinear Convex Anal. 10 (2009), 471-485.
[26] L. M. Bregman, The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, USSR Comput. Math. Math. Phys. 7 (1967), 200-217.
[27] H. H. Bauschke, et al., Bregman distances and Chebyshev sets, J. Approx. Theory 159 (2009), 3-25.
[28] Y. Censor, A. N. Iusem, S. A. Zenios, An interior point method with Bregman functions for the variational inequality problem with paramonotone operators, Math. Program. 81 (1998), 373-400.
[29] R. T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[30] V. Martín Márquez, S. Reich, S. Sabach, Iterative methods for approximating fixed points of Bregman nonexpansive operators, Discret. Contin. Dyn. Syst. Ser. S. 6 (2013), 1043–1063.

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

[32] A. N. Iusem, R. G. Otero, Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, Numer. Funct. Anal. Optim. 22 (2001), 609-640.

[33] S. V. Denisov, V. V. Semenov, L. M. Chabak, Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators, Cybern. Syst. Anal. 51 (2015), 757-765.

[34] R. W. Cottle, J. C. Yao, Pseudomonotone complementarity problems in Hilbert space, J. Optim. Theory Appl. 75 (1992), 281-295.

[35] H. K. Xu, Iterative algorithms for nonlinear operators, J. London Math. Soc. 66 (2002), 240-256.

[36] P. E. Maingé, Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization, Set-valued Anal. 16 (2008), 899-912.

[37] Y. Hao, J. Zhao, New bregman projection algorithms for solving the split feasibility problem, J. Nonlinear Var. Anal. 7 (2023), 971-984.

[38] C. C. Okeke, An improved inertial extragradient subgradient method for solving split variational inequality problems, Bol. Soc. Mat. Mex. 28 (2022), 16.