

INERTIAL EXTRAGRADIENT METHODS FOR SOLVING PSEUDOMONOTONE VARIATIONAL INEQUALITIES WITH NON-LIPSCHITZ MAPPINGS AND THEIR OPTIMIZATION APPLICATIONS

BING TAN¹, SUN YOUNG CHO^{2,*}

¹*Institute of Fundamental and Frontier Sciences,
University of Electronic Science and Technology of China, Chengdu, China*

²*Gyeongsang National University, Jinju-Si, Korea*

Abstract. In this paper, four extragradient-type algorithms with inertial terms are presented for solving the variational inequality problem with a pseudomonotone and non-Lipschitz continuous operator in real Hilbert spaces. Strong convergence theorems of the suggested methods are established under some suitable conditions imposed on the parameters. Finally, several computational tests and applications in optimal control problems are given to illustrate the efficiency and advantages of the proposed iterative schemes over some known ones.

Keywords. Variational inequality; Inertial extragradient method; Viscosity method; Pseudomonotone operator; Non-Lipschitz mapping.

1. INTRODUCTION

This paper deals with several fast and efficient numerical methods for solving variational inequality problems in a real Hilbert space \mathcal{H} . Let C be a nonempty closed convex subset in \mathcal{H} with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Let $A : \mathcal{H} \rightarrow \mathcal{H}$ be an operator. The variational inequality problem (shortly, VIP) for A on C is to find a point $x^* \in C$ such that

$$\langle Ax^*, x - x^* \rangle \geq 0, \quad \forall x \in C. \quad (\text{VIP})$$

Let $\text{VI}(C, A)$ be the solution set of (VIP), which we assume to be nonempty, i.e., $\text{VI}(C, A) \neq \emptyset$. Variational inequality is an essential tool for studying many fields of mathematics and applied science (such as physics, regional, social, engineering and other issues); see, for example, [1, 2, 3, 4]. The theories and methods of variational inequalities have been implemented in numerous areas of science and have proven to be successful and creative. The theory has been shown to provide an easy, common, and consistent structure for dealing with possible issues. In the past few decades, researchers have been very interested in developing effective and robust numerical approaches for solving variational inequality problems. In particular, there has been great interest in projection-based methods and their variants. To see various projection-type methods, one refers to [5, 6, 7, 8, 9, 10, 11, 12] and the references therein.

*Corresponding author.

Email addresses: bingtan72@gmail.com (B. Tan), chosy@gnu.ac.kr (S.Y. Cho).

Received August 4, 2020; Accepted October 9, 2020.

The simplest and earliest projection-type method for solving (VIP) is the projected gradient method: $x_{n+1} = P_C(x_n - \lambda_n Ax_n)$. Note that only one projection onto the feasible set is performed. However, the convergence of this method requires a slightly strong hypothesis that the operator is strongly monotone. To avoid this strong hypothesis, Korpelevich [5] introduced the extragradient method (EGM) to solve saddle point problems in Euclidean spaces. Indeed, the extragradient method is of the form:

$$\begin{cases} y_n = P_C(x_n - \lambda_n Ax_n), \\ x_{n+1} = P_C(x_n - \lambda_n Ay_n), \end{cases} \quad (1.1)$$

where operator A is monotone and L -Lipschitz continuous, P_C denotes the metric projection from \mathcal{H} onto C and $\lambda_n \in (0, 1/L)$. It is known that the sequence $\{x_n\}$ generated by process (1.1) converges to an element in $\text{VI}(C, A)$ when the solution set is nonempty.

Note that the EGM will fail when the Lipschitz constant of the mapping A is unknown or the mapping A is not Lipschitz continuous. To overcome these difficulties, Iusem [6] proposed a new modified extragradient method, which generates a sequence $\{x_n\}$ through the following scheme:

$$\begin{cases} y_n = P_C(x_n - \lambda_n Ax_n), \\ x_{n+1} = P_C(x_n - \beta_n Ay_n), \end{cases} \quad (1.2)$$

where mapping A is continuous and monotone, sequence $\{\lambda_n\}$ is updated by some backtracking search method (also called Armijo criteria) and $\beta_n := \langle Ay_n, x_n - y_n \rangle / \|Ay_n\|^2$. Let us briefly explain the method. Define $\partial H_n = \{x \in \mathbb{R}^m \mid \langle Ay_n, y_n - x \rangle = 0\}$ as a hyperplane. We claim that all solutions x^* of (VIP) lie on one side of ∂H_n . Indeed, since A is monotone (or pseudomonotone), for all $x^* \in C$ and $y_n \in C$, we obtain $\langle Ax^*, y_n - x^* \rangle \geq 0$ and thus $\langle Ay_n, y_n - x^* \rangle \geq 0$. If x_n is on the other side, i.e., $\langle Ay_n, y_n - x_n \rangle < 0$, then ∂H_n separates x_n and x^* . Note that $x_n - \beta_n Ay_n$ is the orthogonal projection of x_n onto ∂H_n . Thus, the iteration point x_{n+1} obtained by the second equation of (1.2) is x_n is first projected onto ∂H_n and then onto the feasible set C .

It should be noted that the method in [6] may not be efficient because it uses Armijo-type line search criteria. In fact, most backtracking methods have the following feature: in order to find out whether a certain ‘‘candidate’’ step size λ_n meets the required inequality, it may be necessary to compute $P_C(x_n - \lambda_n Ax_n)$ many times. This implies that if the backtracking search at iterative n needs m_n steps, then finding y_n requires us to compute m_n projections on the feasible set C . This will further lead to a lot of calculation work since the estimation of the projection is usually not easy. To overcome this shortcoming, an algorithm similar to [6] was proposed by Iusem and Svaiter [8], who used a new search strategy to determine β_n indirectly, which allows the method to compute only two orthogonal projections on the feasible set C in each iteration. They proved that the sequence generated by the method converges to the solution of (VIP) when $\text{VI}(C, A) \neq \emptyset$ and A is a continuous monotone operator. However, the disadvantage of this method is that the convergence speed is very slow. From the computational point of view, this method is not efficient and practical. A few years later, Solodov and Svaiter [9] modified the method of Iusem-Svaiter [8] with a more efficient iterative scheme. The Solodov-Svaiter method is divided into two steps. First, a backtracking search method is used to find a suitable hyperplane that separates the current iteration from the solution of the problem; in the second step, the next iteration is determined as the projection of the current iteration at the intersection of the feasible set and the half-space containing the solution set. More precisely, their iterative scheme is described in Algorithm 1.1.

Algorithm 1.1 The Solodov-Svaiter method for solving (VIP)

Initialization: Take $\ell \in (0, 1)$ and $\mu \in (0, 1)$. Let $x_1 \in \mathbb{R}^m$ be arbitrary.

Iterative Steps: Given the current iterate x_n , calculate x_{n+1} as follows:

Step 1. Compute $y_n = P_C(x_n - Ax_n)$. Set $r(x_n) = x_n - y_n$. If $r(x_n) = 0$, then stop; $x_n \in \text{VI}(C, A)$. Otherwise, go to Step 2.

Step 2. Compute $t_n = x_n - \tau_n r(x_n)$, where $\tau_n = \ell^{m_n}$ and m_n is the smallest non-negative integer m such that $\langle A(x_n - \ell^m r(x_n)), r(x_n) \rangle \geq \mu \|r(x_n)\|^2$.

Step 3. Compute $x_{n+1} = P_{C \cap H_n}(x_n)$, where $H_n = \{x \in \mathbb{R}^m : \langle At_n, x - t_n \rangle \leq 0\}$. Set $n = n + 1$ and go to Step 1.

The geometric interpretation of the methods in [8] and [9] is shown in Fig. 1. As can be seen from Fig. 1: Iusem-Svaiter method is projected first onto the separating hyperplane ∂H_n and then onto C ; the second projection step of the Solodov-Svaiter method is onto the intersection $C \cap H_n$. We can observe that the iteration point of Solodov-Svaiter method is closer to the solution x^* of the problem than the iteration point calculated by Iusem-Svaiter method. Note that the Iusem-Svaiter method may be difficult to converge when x^* is close to the hyperplane ∂H_n . Furthermore, it should be pointed out that adding one more linear constraint to perform a projection onto $C \cap H_n$ will not increase the cost compared to projecting onto C .

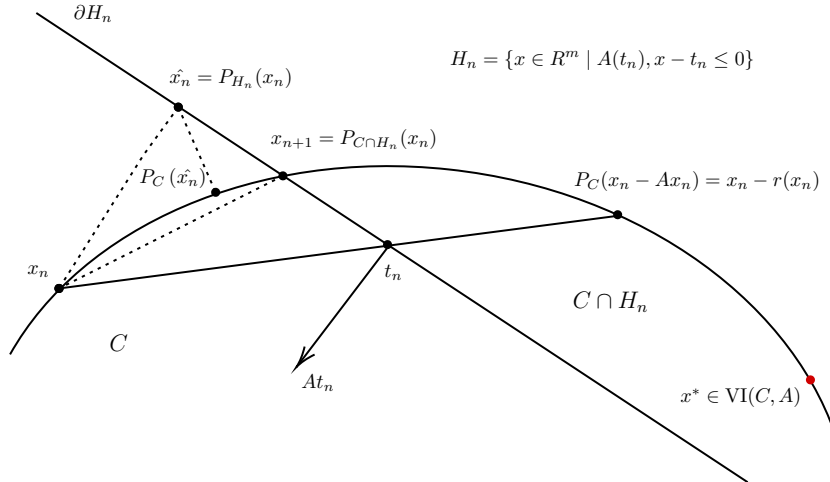


FIGURE 1. Comparison between Iusem-Svaiter method and Solodov-Svaiter method

It should be highlighted that Algorithm 1.1 needs to compute two projections onto the feasible set in each iteration. It is significant to reduce the number of projections on the feasible set C in some cases that the convex set C has a complex expression. This is a fairly common situation, although the operators of interest in many applications are given by closed formulas that are easy to calculate, except for orthogonal projections on convex sets. In the case of spheres, boxes or hyperplanes, the calculation of the projection is in itself is an optimization problem that must be solved approximately by means of some extra numerical schemes. Therefore, many authors have constructed different methods to reduce the number of projections in each iteration, see, e.g., [10, 11, 12, 13] and the references therein. Recently, based on the idea of [9], Vuong and Shehu [14] combined the Halpern method and the Solodov-Svaiter method to propose a new

iterative scheme. Compared with (1.2) and Algorithm 1.1, the advantage of their method is that only one projection on the feasible set needs to be computed in each iteration and strong convergence can be obtained. Indeed, the algorithm is described as follows:

Algorithm 1.2 The Vuong-Shehu method for solving (VIP)

Initialization: Take $\{\alpha_n\} \subset (0, 1)$, $\ell \in (0, 1)$ and $\mu \in (0, 1)$. Let $x_1 \in C$ be arbitrary.

Iterative Steps: Given the current iterate x_n , calculate x_{n+1} as follows:

Step 1. Compute $y_n = P_C(x_n - Ax_n)$. Set $r(x_n) = x_n - y_n$. If $r(x_n) = 0$, then stop; $x_n \in \text{VI}(C, A)$. Otherwise, go to Step 2.

Step 2. Compute $t_n = x_n - \tau_n r(x_n)$, where $\tau_n = \ell^{m_n}$ and m_n is the smallest non-negative integer m such that $\langle A(x_n - \ell^m r(x_n)), r(x_n) \rangle \geq \mu/2 \|r(x_n)\|^2$.

Step 3. Compute $x_{n+1} = \alpha_n x_1 + (1 - \alpha_n) P_{H_n}(x_n)$, where $H_n = \{x \in C : \langle At_n, x - t_n \rangle \leq 0\}$. Set $n = n + 1$ and go to Step 1.

They proved that the iterative sequence $\{x_n\}$ formed by Algorithm 1.2 converges an element of $\text{VI}(C, A)$ in norm, provided that mapping A is pseudomonotone, uniformly continuous and sequentially weakly continuous on bounded subsets of C and the sequence $\{\alpha_n\}$ satisfies $\lim_{n \rightarrow \infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$. It should be pointed out that Algorithm 1.2 degenerates to Algorithm 1.1 when $\alpha_n = 0$ for any n . Notice that Algorithm 1.2 neither requires that the mapping A be Lipschitz continuous nor that the Lipschitz constant be known. In fact, mapping A is uniformly continuous and sequentially weakly continuous, which is a weaker condition than Lipschitz continuity. For more recent research, see, e.g., [15, 16, 17, 18, 19, 20] and the references therein.

Recently, many scholars have developed various types of inertial algorithms by employing inertial extrapolation techniques. The inertial method is based on the discrete version of the second-order dissipative dynamical system originally proposed by Polyak [21]. The main feature of the inertial-type methods is that they use the previously known sequence information to generate the next iteration point. More precisely, the procedure requires two iteration steps and the second iteration step is implemented by the previous two iterations. Note that this small change can greatly accelerate the convergence speed of the iterative algorithm. In recent years, many researchers have studied inertial-type algorithms extensively and have successfully solved many problems, see, e.g., [22, 23, 24, 25, 26, 27] and the references therein.

Motivated and inspired by the above work, and through continuous research in these aspects, in this paper, we introduce four new extragradient-type algorithms for solving variational inequality problems in real Hilbert spaces. Strong convergence theorems of these algorithms are obtained under some mild hypotheses. Finally, we give several numerical examples to support the theoretical results. Moreover, an application of the optimal control problem is implemented to demonstrate the performance of our algorithms. Our methods improve and unify some known results in the literature. Simply put, our contributions in this paper are the following:

- Extend the results in the literature [6, 7, 8, 9] from finite-dimensional spaces to infinite-dimensional Hilbert spaces.
- Several different backtracking search methods are applied in our algorithms, which allow the suggested schemes can work without knowing the prior knowledge of the Lipschitz constant of the mapping. Moreover, our Algorithms 3.2–3.4 only need to calculate the projection on the feasible set in each iteration.

- The operator involved in the variational inequality is uniformly continuous pseudomonotone. This extends many recent results (see, e.g., [8, 9, 15, 22, 23]) on variational inequality where the involved operator has to be uniformly continuous monotone.
- Add an inertial term to the algorithms proposed in [17, 18, 19, 20], which accelerate the convergence speed of the suggested algorithms (see Section 4 and Section 5).
- The strong convergence results of the suggested schemes are obtained under reasonable assumptions. The viscosity method is used in our algorithms, which gives them a faster convergence speed than the Halpern-type algorithms proposed in [14, Algorithm 3.3], [16, Algorithm 4.3] and [17, Algorithm 3.11] (see Section 4 and Section 5).
- Some examples occurring in finite- and infinite-dimensional spaces are presented to illustrate the efficiency of our algorithms over some existing results in the literature. In addition, our algorithms are also devised to solve optimal control problems.

The paper is built up as follows. Some essential definitions and technical lemmas that need to be used are given in the next section. In Section 3, we propose four algorithms and analyze their convergence. Some numerical experiments to verify our theoretical results are presented in Section 4. In Section 5, the proposed methods are investigated to solve optimal control problems. Finally, the paper ends with a brief remark in Section 6, the last section.

2. PRELIMINARIES

Let C be a nonempty closed and convex subset of a real Hilbert space \mathcal{H} . The weak convergence and strong convergence of $\{x_n\}_{n=1}^{\infty}$ to x are represented by $x_n \rightharpoonup x$ and $x_n \rightarrow x$, respectively. For each $x, y, z \in \mathcal{H}$, we have the following inequalities.

- (1) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle$;
- (2) $\|\alpha x + (1 - \alpha)y\|^2 = \alpha\|x\|^2 + (1 - \alpha)\|y\|^2 - \alpha(1 - \alpha)\|x - y\|^2$, $\alpha \in \mathbb{R}$;
- (3) $\|\alpha x + \beta y + \gamma z\|^2 = \alpha\|x\|^2 + \beta\|y\|^2 + \gamma\|z\|^2 - \alpha\beta\|x - y\|^2 - \alpha\gamma\|x - z\|^2 - \beta\gamma\|y - z\|^2$, where $\alpha, \beta, \gamma \in [0, 1]$ with $\alpha + \beta + \gamma = 1$.

For every point $x \in \mathcal{H}$, there exists a unique nearest point in C , denoted by $P_C(x)$ such that $P_C(x) = \operatorname{argmin}\{\|x - y\|, y \in C\}$. P_C is called the metric projection of \mathcal{H} onto C . It is known that P_C is nonexpansive and has the following basic properties:

- (i) $\langle x - P_C(x), y - P_C(x) \rangle \leq 0$, $\forall y \in C$;
- (ii) $\|P_C(x) - P_C(y)\|^2 \leq \langle P_C(x) - P_C(y), x - y \rangle$, $\forall y \in \mathcal{H}$;
- (iii) $\|P_C(x) - y\|^2 \leq \|x - y\|^2 - \|x - P_C(x)\|^2$, $\forall y \in C$.

A mapping $A : \mathcal{H} \rightarrow \mathcal{H}$ is said to be:

- (1) *L-Lipschitz continuous* with $L > 0$ if

$$\|Ax - Ay\| \leq L\|x - y\|, \quad \forall x, y \in \mathcal{H}.$$

If $L \in (0, 1)$ then mapping A is called *contraction*. In particular, when $L = 1$, mapping A is called *nonexpansive*.

- (2) *monotone* if

$$\langle Ax - Ay, x - y \rangle \geq 0, \quad \forall x, y \in \mathcal{H}.$$

- (3) *pseudomonotone* if

$$\langle Ax, y - x \rangle \geq 0 \implies \langle Ay, y - x \rangle \geq 0, \quad \forall x, y \in \mathcal{H}.$$

- (4) *sequentially weakly continuous* if for each sequence $\{x_n\}$ converges weakly to x implies $\{Ax_n\}$ converges weakly to Ax .

We give some projection calculation formulas that need to be used in numerical experiments. For more projection calculations on specific sets, see [28].

- (1) The projection of x onto a half-space $H_{u,v} = \{x : \langle u, x \rangle \leq v\}$ is computed by

$$P_{H_{u,v}}(x) = x - \max\{[\langle u, x \rangle - v] / \|u\|^2, 0\}u.$$

- (2) The projection of x onto a box $\text{Box}[a, b] = \{x : a \leq x \leq b\}$ is computed by

$$P_{\text{Box}[a,b]}(x)_i = \min\{b_i, \max\{x_i, a_i\}\}.$$

- (3) The projection of x onto a ball $B[p, q] = \{x : \|x - p\| \leq q\}$ is computed by

$$P_{B[p,q]}(x) = p + \frac{q}{\max\{\|x - p\|, q\}}(x - p).$$

The following lemmas will be used in the convergence analysis of our algorithms.

Lemma 2.1 ([29]). *For $x \in \mathcal{H}$ and $\alpha \geq \beta > 0$ the following inequality holds.*

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

Lemma 2.2 ([30]). *Let \mathcal{H}_1 and \mathcal{H}_2 be two real Hilbert spaces. Suppose $A : \mathcal{H}_1 \rightarrow \mathcal{H}_2$ is uniformly continuous on bounded subsets of \mathcal{H}_1 and M is a bounded subset of \mathcal{H}_1 . Then, $A(M)$ is bounded.*

Lemma 2.3 ([31]). *Assume that C is a convex and closed nonempty subset of a real Hilbert space \mathcal{H} . Let h be a real-valued function on \mathcal{H} and define $K = \{x \in C : h(x) \leq 0\}$. If K is nonempty and h is θ -Lipschitz continuous on C , then*

$$\text{dist}(x, K) \geq \theta^{-1} \max\{h(x), 0\}, \quad \forall x \in C,$$

where $\text{dist}(x, K)$ denotes the distance function from x to K .

Lemma 2.4 ([32]). *Assume that C is a closed and convex subset of a real Hilbert space \mathcal{H} . Let operator $A : C \rightarrow \mathcal{H}$ be continuous and pseudomonotone. Then, x^* is a solution of (VIP) if and only if $\langle Ax, x - x^* \rangle \geq 0, \forall x \in C$.*

Lemma 2.5 ([33]). *Let $\{p_n\}$ be a positive sequence, $\{q_n\}$ be a sequence of real numbers, and $\{\sigma_n\}$ be a sequence in $(0, 1)$ such that $\sum_{n=1}^{\infty} \sigma_n = \infty$. Assume that*

$$p_{n+1} \leq (1 - \sigma_n) p_n + \sigma_n q_n, \quad \forall n \geq 1.$$

If $\limsup_{k \rightarrow \infty} q_{n_k} \leq 0$ for every subsequence $\{p_{n_k}\}$ of $\{p_n\}$ satisfying $\liminf_{k \rightarrow \infty} (p_{n_{k+1}} - p_{n_k}) \geq 0$, then $\lim_{n \rightarrow \infty} p_n = 0$.

3. MAIN RESULTS

In this section, we introduce four new iterative schemes for solving the variational inequality problem involving a uniformly continuous and pseudomonotone operator in a real Hilbert space. These algorithms guarantee strong convergence through the viscosity-type method. In order to analyze the convergence properties of the algorithms, the mapping and parameters involved in our methods need to satisfy the following assumptions.

- (C1) The feasible set C is a nonempty, closed and convex subset of \mathcal{H} .
- (C2) The solution set of the (VIP) is nonempty, that is $\text{VI}(C, A) \neq \emptyset$.
- (C3) The operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is pseudomonotone, uniformly continuous and sequentially weakly continuous on bounded subsets of C . In finite-dimensional spaces, it suffices to assume that operator $A : \mathcal{H} \rightarrow \mathcal{H}$ is continuous pseudomonotone on \mathcal{H} .
- (C4) The mapping $f : C \rightarrow C$ is ρ -contractive with constant $\rho \in [0, 1)$.
- (C5) Let $\{\varepsilon_n\}$ be a positive sequence such that $\lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\alpha_n} = 0$, where $\{\alpha_n\} \subset (0, 1)$ satisfies $\lim_{n \rightarrow \infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$.

3.1. The Algorithm 3.1. Our first algorithm is based on Algorithm 1.2 and the Algorithm 3.1.1 presented by Thong and Gibali [17]. The suggested scheme combines inertial method and viscosity technique to solve (VIP), which enhances the convergence speed of the method proposed in [17] (see our numerical experiments). The Algorithm 3.1 is described as follows.

Algorithm 3.1 The modified viscosity-type extragradient algorithm

Initialization: Take $\gamma > 0$, $\ell \in (0, 1)$, $\mu \in (0, 1)$. Let $x_0, x_1 \in C$ be arbitrary.

Iterative Steps: Calculate x_{n+1} as follows:

Step 1. Given the iterates x_{n-1} and x_n ($n \geq 1$). Set $w_n = x_n + \theta_n(x_n - x_{n-1})$, where

$$\theta_n = \begin{cases} \min \left\{ \frac{\varepsilon_n}{\|x_n - x_{n-1}\|}, \theta \right\}, & \text{if } x_n \neq x_{n-1}; \\ \theta, & \text{otherwise.} \end{cases} \quad (3.1)$$

Step 2. Compute $y_n = P_C(w_n - \lambda_n A w_n)$, where $\lambda_n := \gamma \ell^{m_n}$ and m_n is the smallest non-negative integer m satisfying

$$\gamma \ell^m \langle A w_n - A y_n, w_n - y_n \rangle \leq \mu \|w_n - y_n\|^2. \quad (3.2)$$

If $w_n = y_n$ or $A y_n = 0$ then stop and y_n is a solution of (VIP). Otherwise, go to Step 2.

Step 3. Compute $z_n = P_C(w_n - \beta_n A y_n)$, where

$$\beta_n = \frac{1 - \mu \|w_n - y_n\|^2}{\gamma \|A y_n\|^2}.$$

Step 4. Compute $x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n) z_n$.

Set $n = n + 1$ and go to Step 1.

Remark 3.1. The following comments are important and basic.

- (i) We note here that the inertia calculation criterion (3.1) is easy to implement since the term $\|x_n - x_{n-1}\|$ is known before calculating θ_n . Moreover, it follows from (3.1) and Assumption (C5) that

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

Indeed, we obtain $\theta_n \|x_n - x_{n-1}\| \leq \varepsilon_n, \forall n \geq 1$, which together with $\lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\alpha_n} = 0$ implies that $\lim_{n \rightarrow \infty} \frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \leq \lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\alpha_n} = 0$.

(ii) We prove that if $w_n = y_n$ or $Ay_n = 0$ then $y_n \in \text{VI}(C, A)$. Indeed, from $0 < \lambda_n \leq \gamma$ and Lemma 2.1, we have

$$0 = \frac{\|w_n - y_n\|}{\lambda_n} = \frac{\|w_n - P_C(w_n - \lambda_n Aw_n)\|}{\lambda_n} \geq \frac{\|w_n - P_C(w_n - \gamma Aw_n)\|}{\gamma},$$

which indicates that w_n is a solution of (VIP), so $y_n \in \text{VI}(C, A)$. On the other hand, since $y_n \in C$, one sees that if $Ay_n = 0$ then $y_n = P_C(y_n - \gamma Ay_n)$, that is $y_n \in \text{VI}(C, A)$.

Lemma 3.1. *Suppose that Assumptions (C1)–(C3) hold. The Armijo-like criteria (3.2) is well defined. Moreover, we get $\lambda_n \leq \gamma$.*

Proof. If $w_n \in \text{VI}(C, A)$ then $w_n = P_C(w_n - \gamma Aw_n)$, which implies that $w_n = y_n$ and thus $m_n = 0$. If $w_n \notin \text{VI}(C, A)$, we assume that the opposite of (3.2) holds, that is,

$$\gamma \ell^m \langle Aw_n - AP_C(w_n - \gamma \ell^m Aw_n), w_n - P_C(w_n - \gamma \ell^m Aw_n) \rangle > \mu \|w_n - P_C(w_n - \gamma \ell^m Aw_n)\|^2.$$

Therefore, we get

$$\|Aw_n - AP_C(w_n - \gamma \ell^m Aw_n)\| > \mu \frac{\|w_n - P_C(w_n - \gamma \ell^m Aw_n)\|}{\gamma \ell^m}. \quad (3.3)$$

We study two cases of w_n . First, suppose that $w_n \in C$. Since A and P_C are continuous, we obtain

$$\lim_{m \rightarrow \infty} \|w_n - P_C(w_n - \gamma \ell^m Aw_n)\| = 0.$$

From the fact that A is uniformly continuous, one has

$$\lim_{m \rightarrow \infty} \|Aw_n - AP_C(w_n - \gamma \ell^m Aw_n)\| = 0,$$

which combining with (3.3) yields

$$\lim_{m \rightarrow \infty} \frac{\|w_n - P_C(w_n - \gamma \ell^m Aw_n)\|}{\gamma \ell^m} = 0. \quad (3.4)$$

Let $z_m = P_C(w_n - \gamma \ell^m Aw_n)$. According to the characteristics of projection, one obtains

$$\langle z_m - w_n + \gamma \ell^m Aw_n, x - z_m \rangle \geq 0, \quad \forall x \in C,$$

which means that

$$\langle (z_m - w_n)/\gamma \ell^m, x - z_m \rangle + \langle Aw_n, x - z_m \rangle \geq 0, \quad \forall x \in C.$$

This together with (3.4), we get that $\langle Aw_n, x - w_n \rangle \geq 0, \forall x \in C$. This shows that $w_n \in \text{VI}(C, A)$, which contradicts the hypothesis.

On the other hand, if $w_n \notin C$, then we obtain

$$\lim_{m \rightarrow \infty} \|w_n - P_C(w_n - \gamma \ell^m Aw_n)\| = \|w_n - P_C(w_n)\| > 0,$$

and

$$\lim_{m \rightarrow \infty} \gamma \ell^m \|Aw_n - AP_C(w_n - \gamma \ell^m Aw_n)\| = 0.$$

Combining these with (3.3), we get an opposite. The proof is completed. \square

Remark 3.2. It is worth noting that we did not use the pseudo-monotonicity of mapping A in the proof of Lemma 3.1.

Lemma 3.2. *Suppose that Assumptions (C1)–(C3) hold. Let sequences $\{w_n\}$ and $\{y_n\}$ be created by Algorithm 3.1. Then, we obtain*

$$\langle Ay_n, w_n - p \rangle \geq \frac{1 - \mu}{\gamma} \|w_n - y_n\|^2, \quad \forall p \in \text{VI}(C, A).$$

Proof. Recall the projection property $\|x - P_C(y)\|^2 \leq \langle x - y, x - P_C(y) \rangle, \forall x \in C, y \in \mathcal{H}$. By setting $x = w_n$ and $y = w_n - \lambda_n A w_n$, we have

$$\|w_n - P_C(w_n - \lambda_n A w_n)\|^2 \leq \lambda_n \langle A w_n, w_n - P_C(w_n - \lambda_n A w_n) \rangle,$$

which means that $\langle A w_n, w_n - y_n \rangle \geq 1/\lambda_n \|w_n - y_n\|^2$. Since $y_n \in C$ and $p \in \text{VI}(C, A)$, we get that $\langle A p, y_n - p \rangle \geq 0$. From the pseudomonotonicity of mapping A , one sees that $\langle A y_n, y_n - p \rangle \geq 0$. Thus, using (3.2), we deduce that

$$\begin{aligned} \langle A y_n, w_n - p \rangle &= \langle A y_n, w_n - y_n \rangle + \langle A y_n, y_n - p \rangle \geq \langle A y_n, w_n - y_n \rangle \\ &= \langle A w_n, w_n - y_n \rangle - \langle A w_n - A y_n, w_n - y_n \rangle \\ &\geq \frac{1}{\lambda_n} \|w_n - y_n\|^2 - \frac{\mu}{\lambda_n} \|w_n - y_n\|^2 \\ &= \frac{1 - \mu}{\lambda_n} \|w_n - y_n\|^2 \geq \frac{1 - \mu}{\gamma} \|w_n - y_n\|^2. \end{aligned}$$

This completes the proof of the lemma. \square

Remark 3.3. By Lemma 3.2, one can see that if $A y_n = 0$ then $w_n = y_n$, that is, $y_n \in \text{VI}(C, A)$.

Lemma 3.3. *Suppose that Assumptions (C1)–(C3) hold. Let $\{w_n\}$ and $\{y_n\}$ be two sequences formulated by Algorithm 3.1. If there exists a subsequence $\{w_{n_k}\}$ of $\{w_n\}$ such that $\{w_{n_k}\}$ converges weakly to $z \in \mathcal{H}$ and $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$, then $z \in \text{VI}(C, A)$.*

Proof. From $\{y_n\} \subset C$, $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$ and $w_{n_k} \rightharpoonup z$, we have $z \in C$. According to the property of projection and the definition of y_{n_k} , we obtain

$$\langle w_{n_k} - \lambda_{n_k} A w_{n_k} - y_{n_k}, x - y_{n_k} \rangle \leq 0, \quad \forall x \in C,$$

which can be written as

$$\lambda_{n_k}^{-1} \langle w_{n_k} - y_{n_k}, x - y_{n_k} \rangle + \langle A w_{n_k}, y_{n_k} - w_{n_k} \rangle \leq \langle A w_{n_k}, x - w_{n_k} \rangle, \quad \forall x \in C. \quad (3.5)$$

Now, we prove that $\liminf_{k \rightarrow \infty} \langle A w_{n_k}, x - w_{n_k} \rangle \geq 0$ by considering two possible situations of λ_{n_k} . First, we assume that $\liminf_{k \rightarrow \infty} \lambda_{n_k} > 0$. Since sequence $\{w_{n_k}\}$ is bounded and mapping A is uniformly continuous, in the light of Lemma 2.2, one gets $\{A w_{n_k}\}$ is bounded. Combining $\|w_{n_k} - y_{n_k}\| \rightarrow 0$ and (3.5), we have $\liminf_{k \rightarrow \infty} \langle A w_{n_k}, x - w_{n_k} \rangle \geq 0$. Next, one supposes that $\liminf_{k \rightarrow \infty} \lambda_{n_k} = 0$. Setting $s_{n_k} = P_C(w_{n_k} - \lambda_{n_k} \ell^{-1} A w_{n_k})$, by means of Lemma 2.1, one obtains $\ell \|w_{n_k} - s_{n_k}\| \leq \|w_{n_k} - y_{n_k}\| \rightarrow 0$. Thus, $s_{n_k} \rightarrow z \in C$, which means that sequence $\{s_{n_k}\}$ is bounded. This together with the uniform continuity of mapping A , we get

$$\lim_{k \rightarrow \infty} \|A w_{n_k} - A s_{n_k}\| \rightarrow 0. \quad (3.6)$$

Combining (3.2) and the Cauchy-Schwartz inequality ($\langle a, b \rangle \leq \|a\| \|b\|$), we have

$$\lambda_{n_k} \ell^{-1} \|A w_{n_k} - A P_C(w_{n_k} - \lambda_{n_k} \ell^{-1} A w_{n_k})\| > \mu \|w_{n_k} - P_C(w_{n_k} - \lambda_{n_k} \ell^{-1} A w_{n_k})\|.$$

Therefore, we get $\frac{1}{\mu} \|Aw_{n_k} - As_{n_k}\| > \frac{\|w_{n_k} - s_{n_k}\|}{\lambda_{n_k} \ell^{-1}}$. This combining with (3.6) gives

$$\lim_{k \rightarrow \infty} \frac{\|w_{n_k} - s_{n_k}\|}{\lambda_{n_k} \ell^{-1}} = 0.$$

Moreover, according to the definition of s_{n_k} and the property of projection, we obtain

$$\langle w_{n_k} - \lambda_{n_k} \ell^{-1} Aw_{n_k} - s_{n_k}, x - s_{n_k} \rangle \leq 0, \quad \forall x \in C,$$

which yields

$$\frac{1}{\lambda_{n_k} \ell^{-1}} \langle w_{n_k} - s_{n_k}, x - s_{n_k} \rangle + \langle Aw_{n_k}, s_{n_k} - w_{n_k} \rangle \leq \langle Aw_{n_k}, x - w_{n_k} \rangle, \quad \forall x \in C.$$

Taking the limits on the left and right sides of the above inequality, one has

$$\liminf_{k \rightarrow \infty} \langle Aw_{n_k}, x - w_{n_k} \rangle \geq 0. \quad (3.7)$$

Hence, we achieved the desired result.

Now, we show that $z \in \text{VI}(C, A)$. Indeed, one sees that

$$\langle Ay_{n_k}, x - y_{n_k} \rangle = \langle Ay_{n_k} - Aw_{n_k}, x - w_{n_k} \rangle + \langle Aw_{n_k}, x - w_{n_k} \rangle + \langle Ay_{n_k}, w_{n_k} - y_{n_k} \rangle. \quad (3.8)$$

Since $\|w_{n_k} - y_{n_k}\| \rightarrow 0$ and mapping A is uniformly continuous, we get $\lim_{k \rightarrow \infty} \|Aw_{n_k} - Ay_{n_k}\| = 0$. This together with (3.7) and (3.8) yields that $\liminf_{k \rightarrow \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle \geq 0$.

Next, we select a positive number decreasing sequence $\{\zeta_k\}$ such that $\zeta_k \rightarrow 0$ as $k \rightarrow \infty$. For any k , we represent the smallest positive integer with N_k such that

$$\langle Ay_{n_j}, x - y_{n_j} \rangle + \zeta_k \geq 0, \quad \forall j \geq N_k. \quad (3.9)$$

It can be easily seen that the sequence $\{N_k\}$ is increasing because $\{\zeta_k\}$ is decreasing. Moreover, for any k , from $\{y_{N_k}\} \subset C$, we can assume $Ay_{N_k} \neq 0$ (otherwise, y_{N_k} is a solution) and set $u_{N_k} = Ay_{N_k} / \|Ay_{N_k}\|^2$. Then, we get $\langle Ay_{N_k}, u_{N_k} \rangle = 1, \forall k$. Now, we can deduce from (3.9) that $\langle Ay_{N_k}, x + \zeta_k u_{N_k} - y_{N_k} \rangle \geq 0, \forall k$. According to the fact that A is pseudomonotone on \mathcal{H} , we can show that

$$\langle A(x + \zeta_k u_{N_k}), x + \zeta_k u_{N_k} - y_{N_k} \rangle \geq 0,$$

which further yields that

$$\langle Ax, x - y_{N_k} \rangle \geq \langle Ax - A(x + \zeta_k u_{N_k}), x + \zeta_k u_{N_k} - y_{N_k} \rangle - \zeta_k \langle Ax, u_{N_k} \rangle. \quad (3.10)$$

Now, we prove that $\lim_{k \rightarrow \infty} \zeta_k u_{N_k} = 0$. We get that $y_{n_k} \rightharpoonup z$ since $w_{n_k} \rightharpoonup z$ and $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$. From $\{y_n\} \subset C$, we have $z \in C$. In view of A is sequentially weakly continuous on C , one has $Ay_{n_k} \rightharpoonup Az$. One assumes that $Az \neq 0$ (otherwise, z is a solution). According to the fact that norm mapping is sequentially weakly lower semicontinuous, we obtain $0 < \|Az\| \leq \liminf_{k \rightarrow \infty} \|Ay_{n_k}\|$. Using $\{y_{N_k}\} \subset \{y_{n_k}\}$ and $\zeta_k \rightarrow 0$ as $k \rightarrow \infty$, we have

$$0 \leq \limsup_{k \rightarrow \infty} \|\zeta_k u_{N_k}\| = \limsup_{k \rightarrow \infty} \left(\frac{\zeta_k}{\|Ay_{n_k}\|} \right) \leq \frac{\limsup_{k \rightarrow \infty} \zeta_k}{\liminf_{k \rightarrow \infty} \|Ay_{n_k}\|} = 0.$$

That is, $\lim_{k \rightarrow \infty} \zeta_k u_{N_k} = 0$. Thus, from the facts that A is uniformly continuous, sequences $\{y_{N_k}\}$ and $\{u_{N_k}\}$ are bounded and $\lim_{k \rightarrow \infty} \zeta_k u_{N_k} = 0$, we can conclude from (3.10) that $\liminf_{k \rightarrow \infty} \langle Ax, x - y_{N_k} \rangle \geq 0$. Therefore,

$$\langle Ax, x - z \rangle = \lim_{k \rightarrow \infty} \langle Ax, x - y_{N_k} \rangle = \liminf_{k \rightarrow \infty} \langle Ax, x - y_{N_k} \rangle \geq 0, \quad \forall x \in C.$$

Consequently, we observe that $z \in \text{VI}(C, A)$ by Lemma 2.4. This completes the proof. \square

Remark 3.4. If mapping A is monotone, then A does not need to satisfy the sequential weak continuity. Moreover, if the step size is fixed and not updated through (3.2), then Lemma 3.3 is obviously true.

Theorem 3.1. *Suppose that Assumptions (C1)–(C5) hold. Then the iterative sequence $\{x_n\}$ formulated by Algorithm 3.1 converges to $p \in \text{VI}(C, A)$ in norm, where $p = P_{\text{VI}(C, A)}(f(p))$.*

Proof. We divide our proof in four steps.

Claim 1. The sequence $\{x_n\}$ is bounded. Combining P_C is nonexpansive, Lemma 3.2 and the definition of β_n , we get

$$\begin{aligned}
 \|z_n - p\|^2 &= \|P_C(w_n - \beta_n A y_n) - P_C(p)\|^2 \leq \|w_n - \beta_n A y_n - p\|^2 \\
 &= \|w_n - p\|^2 - 2\beta_n \langle A y_n, w_n - p \rangle + \beta_n^2 \|A y_n\|^2 \\
 &\leq \|w_n - p\|^2 - 2\beta_n \frac{1 - \mu}{\gamma} \|w_n - y_n\|^2 + \beta_n^2 \|A y_n\|^2 \\
 &\leq \|w_n - p\|^2 - 2 \frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|A y_n\|^2} + \frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|A y_n\|^2} \\
 &= \|w_n - p\|^2 - \frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|A y_n\|^2}, \quad \forall p \in \text{VI}(C, A).
 \end{aligned} \tag{3.11}$$

This indicates that

$$\|z_n - p\| \leq \|w_n - p\|. \tag{3.12}$$

By the definition of w_n , one sees that

$$\|w_n - p\| \leq \|x_n - p\| + \alpha_n \cdot \frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\|. \tag{3.13}$$

From Remark 3.1 (i), one gets $\frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \rightarrow 0$. Thus, there is a constant $Q_1 > 0$ that satisfies

$$\frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \leq Q_1, \quad \forall n \geq 1. \tag{3.14}$$

Using (3.12), (3.13) and (3.14), we obtain

$$\|z_n - p\| \leq \|w_n - p\| \leq \|x_n - p\| + \alpha_n Q_1, \quad \forall n \geq 1. \tag{3.15}$$

From the definition of x_{n+1} and (3.15), we have

$$\begin{aligned}
 \|x_{n+1} - p\| &= \|\alpha_n (f(x_n) - p) + (1 - \alpha_n)(z_n - p)\| \\
 &\leq \alpha_n \|f(x_n) - f(p)\| + \alpha_n \|f(p) - p\| + (1 - \alpha_n) \|z_n - p\| \\
 &\leq [1 - \alpha_n(1 - \rho)] \|x_n - p\| + \alpha_n(1 - \rho) \cdot (\|f(p) - p\| + Q_1) / (1 - \rho) \\
 &\leq \max \{ \|x_n - p\|, (\|f(p) - p\| + Q_1) / (1 - \rho) \} \\
 &\leq \dots \leq \max \{ \|x_0 - p\|, (\|f(p) - p\| + Q_1) / (1 - \rho) \}.
 \end{aligned} \tag{3.16}$$

That is, $\{x_n\}$ is bounded. We have $\{w_n\}$, $\{y_n\}$, $\{z_n\}$, $\{A y_n\}$ and $\{f(x_n)\}$ are also bounded.

Claim 2.

$$\frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|A y_n\|^2} \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3)$$

for some $Q_2, Q_3 > 0$. Indeed, it follows from (3.15) that

$$\begin{aligned} \|w_n - p\|^2 &\leq \|x_n - p\|^2 + \alpha_n(2Q_1\|x_n - p\| + \alpha_n Q_1^2) \\ &\leq \|x_n - p\|^2 + \alpha_n Q_2 \end{aligned} \quad (3.17)$$

for some $Q_2 > 0$. Combining (3.11) and (3.17), we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|(1 - \alpha_n)(z_n - p) + \alpha_n(f(x_n) - p)\|^2 \\ &\leq (1 - \alpha_n)\|z_n - p\|^2 + 2\alpha_n\langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|x_n - p\|^2 - \frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|Ay_n\|^2} + 2\alpha_n\langle f(x_n) - p, x_{n+1} - p \rangle + \alpha_n Q_2. \end{aligned}$$

Therefore, we obtain

$$\begin{aligned} \frac{(1 - \mu)^2}{\gamma^2} \frac{\|w_n - y_n\|^4}{\|Ay_n\|^2} &\leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + 2\alpha_n\|f(x_n) - p\|\|x_{n+1} - p\| + \alpha_n Q_2 \\ &\leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3), \end{aligned}$$

where $Q_3 = \sup_{n \in \mathbb{N}} \{2\|f(x_n) - p\|\|x_{n+1} - p\|\}$.

Claim 3.

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - (1 - \rho)\alpha_n)\|x_n - p\|^2 + (1 - \rho)\alpha_n \cdot \left[\frac{3Q}{1 - \rho} \cdot \frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \right. \\ &\quad \left. + \frac{2}{1 - \rho} \langle f(p) - p, x_{n+1} - p \rangle \right]. \end{aligned}$$

Using the definition of w_n , we can show that

$$\begin{aligned} \|w_n - p\|^2 &\leq \|x_n - p\|^2 + 2\theta_n\|x_n - p\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2 \\ &\leq \|x_n - p\|^2 + 3Q\theta_n\|x_n - x_{n-1}\|, \end{aligned} \quad (3.18)$$

where $Q = \sup_{n \in \mathbb{N}} \{\|x_n - p\|, \theta\|x_n - x_{n-1}\|\} > 0$. Using (3.12) and (3.18), we get

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|\alpha_n f(z_n) + (1 - \alpha_n)z_n - p\|^2 \\ &= \|\alpha_n(f(z_n) - f(p)) + (1 - \alpha_n)(z_n - p) + \alpha_n(f(p) - p)\|^2 \\ &\leq \|\alpha_n(f(z_n) - f(p)) + (1 - \alpha_n)(z_n - p)\|^2 + 2\alpha_n\langle f(p) - p, x_{n+1} - p \rangle \\ &\leq \alpha_n\|f(z_n) - f(p)\|^2 + (1 - \alpha_n)\|z_n - p\|^2 + 2\alpha_n\langle f(p) - p, x_{n+1} - p \rangle \\ &\leq \alpha_n \rho^2 \|z_n - p\|^2 + (1 - \alpha_n)\|z_n - p\|^2 + 2\alpha_n\langle f(p) - p, x_{n+1} - p \rangle \quad (\text{TC}) \\ &\leq (1 - (1 - \rho)\alpha_n)\|z_n - p\|^2 + 2\alpha_n\langle f(p) - p, x_{n+1} - p \rangle \\ &\leq (1 - (1 - \rho)\alpha_n)\|x_n - p\|^2 + (1 - \rho)\alpha_n \cdot \left[\frac{3Q}{1 - \rho} \cdot \frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \right. \\ &\quad \left. + \frac{2}{1 - \rho} \langle f(p) - p, x_{n+1} - p \rangle \right]. \end{aligned}$$

Claim 4. The sequence $\{\|x_n - p\|^2\}$ converges to zero. From Lemma 2.5 and Remark 3.1 (i), it remains to show that $\limsup_{k \rightarrow \infty} \langle f(p) - p, x_{n_k+1} - p \rangle \leq 0$ for any subsequence $\{\|x_{n_k} - p\|\}$ of $\{\|x_n - p\|\}$ satisfies $\liminf_{k \rightarrow \infty} (\|x_{n_k+1} - p\| - \|x_{n_k} - p\|) \geq 0$.

For this purpose, we assume that $\{\|x_{n_k} - p\|\}$ is a subsequence of $\{\|x_n - p\|\}$ such that

$$\liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - p\| - \|x_{n_k} - p\|) \geq 0. \quad (3.19)$$

Then, we have

$$\begin{aligned} & \liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - p\|^2 - \|x_{n_k} - p\|^2) \\ &= \liminf_{k \rightarrow \infty} [(\|x_{n_{k+1}} - p\| - \|x_{n_k} - p\|)(\|x_{n_{k+1}} - p\| + \|x_{n_k} - p\|)] \geq 0. \end{aligned}$$

It follows from Claim 2 and Assumption (C5) that

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \frac{(1 - \mu)^2 \|w_{n_k} - y_{n_k}\|^4}{\gamma^2 \|Ay_{n_k}\|^2} \\ & \leq \limsup_{k \rightarrow \infty} [\|x_{n_k} - p\|^2 - \|x_{n_{k+1}} - p\|^2] + \limsup_{k \rightarrow \infty} \alpha_{n_k} (Q_2 + Q_3) \\ & = -\liminf_{k \rightarrow \infty} [\|x_{n_{k+1}} - p\|^2 - \|x_{n_k} - p\|^2] \leq 0, \end{aligned}$$

which yields that $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$. According to the definition of z_n and β_n , one has

$$\begin{aligned} \|z_{n_k} - w_{n_k}\| &= \|P_C(w_{n_k} - \beta_{n_k} Ay_{n_k}) - P_C(w_{n_k})\| \leq \|w_{n_k} - \beta_{n_k} Ay_{n_k} - w_{n_k}\| \\ &= \beta_{n_k} \|Ay_{n_k}\| = \frac{1 - \mu}{\gamma} \frac{\|w_{n_k} - y_{n_k}\|^2}{\|Ay_{n_k}\|}. \end{aligned}$$

Hence, we get $\lim_{k \rightarrow \infty} \|z_{n_k} - w_{n_k}\| = 0$. Moreover, using Remark 3.1 (i) and Assumption (C5), we have

$$\|x_{n_k} - w_{n_k}\| = \alpha_{n_k} \cdot \frac{\theta_{n_k}}{\alpha_{n_k}} \|x_{n_k} - x_{n_{k-1}}\| \rightarrow 0,$$

and

$$\|x_{n_{k+1}} - z_{n_k}\| = \alpha_{n_k} \|z_{n_k} - f(z_{n_k})\| \rightarrow 0.$$

Therefore, we conclude that

$$\|x_{n_{k+1}} - x_{n_k}\| \leq \|x_{n_{k+1}} - z_{n_k}\| + \|z_{n_k} - w_{n_k}\| + \|w_{n_k} - x_{n_k}\| \rightarrow 0. \quad (3.20)$$

From the sequence $\{x_{n_k}\}$ is bounded, there is a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$ that satisfies $x_{n_{k_j}} \rightarrow q$. Furthermore,

$$\limsup_{k \rightarrow \infty} \langle f(p) - p, x_{n_k} - p \rangle = \lim_{j \rightarrow \infty} \langle f(p) - p, x_{n_{k_j}} - p \rangle = \langle f(p) - p, q - p \rangle. \quad (3.21)$$

We get $w_{n_k} \rightarrow q$ since $\|x_{n_k} - w_{n_k}\| \rightarrow 0$. This together with $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$ and Lemma 3.3 yields that $q \in \text{VI}(C, A)$. By the definition of $p = P_{\text{VI}(C, A)}(f(p))$ and (3.21), we infer that

$$\limsup_{k \rightarrow \infty} \langle f(p) - p, x_{n_k} - p \rangle = \langle f(p) - p, q - p \rangle \leq 0. \quad (3.22)$$

Combining (3.20) and (3.22), we get

$$\limsup_{k \rightarrow \infty} \langle f(p) - p, x_{n_{k+1}} - p \rangle \leq \limsup_{k \rightarrow \infty} \langle f(p) - p, x_{n_k} - p \rangle \leq 0. \quad (3.23)$$

From Remark 3.1 (i), (3.23), Claim 3 and Lemma 2.5, we conclude that $x_n \rightarrow p$ as $n \rightarrow \infty$. The proof of Theorem 3.1 is now complete. \square

3.2. The Algorithm 3.2. The second iterative scheme is stated in Algorithm 3.2. This method is derived from the algorithm 3 suggested by Thong, Shehu and Iyiola [18]. Our contribution is to add an inertial term to speed up the convergence speed of [18, Algorithm 3].

Algorithm 3.2 The modified inertial projection-type algorithm

Initialization: Take $\gamma > 0, \ell \in (0, 1), \mu \in (0, 1)$. Let $x_0, x_1 \in C$ be arbitrary.

Iterative Steps: Calculate x_{n+1} as follows:

$$\begin{cases} w_n = x_n + \theta_n (x_n - x_{n-1}), \\ y_n = P_C (w_n - \lambda_n A w_n), \\ x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n) P_{H_n} (w_n). \end{cases}$$

where $H_n = \{x \in \mathcal{H} : h_n(x) \leq 0\}$, $h_n(x) = \langle w_n - y_n - \lambda_n (A w_n - A y_n), x - y_n \rangle$, $\{\theta_n\}$ and $\{\lambda_n\}$ are defined in (3.1) and (3.2), respectively.

The following lemma is very helpful for the convergence analysis of the algorithm.

Lemma 3.4. *Suppose that Assumptions (C1)–(C3) hold. Let x^* be a solution of (VIP). Then $h_n(x^*) \leq 0$ and $h_n(w_n) \geq (1 - \mu)\|w_n - y_n\|^2$.*

Proof. From $x^* \in \text{VI}(C, A)$ and $y_n \in C$, it follows from Lemma 2.4 that $\langle A y_n, x^* - y_n \rangle \leq 0$. Combining the definition of y_n and the property of projection, we have

$$h_n(x^*) = \langle w_n - y_n - \lambda_n A w_n, x^* - y_n \rangle + \lambda_n \langle A y_n, x^* - y_n \rangle \leq 0.$$

Using (3.2) we can show that

$$\begin{aligned} h_n(w_n) &= \langle w_n - y_n - \lambda_n (A w_n - A y_n), w_n - y_n \rangle \\ &= \|w_n - y_n\|^2 - \lambda_n \langle A w_n - A y_n, w_n - y_n \rangle \\ &\geq \|w_n - y_n\|^2 - \mu \|w_n - y_n\|^2 = (1 - \mu) \|w_n - y_n\|^2. \end{aligned}$$

The proof of this lemma is now complete. \square

Remark 3.5. From Remark 3.1 (ii), if $w_n = y_n$ then $y_n \in \text{VI}(C, A)$. Otherwise, it follows from Lemma 3.4 that $h_n(w_n) > 0$, which indicates that $w_n \notin H_n$. According to the calculation formula of projection onto a half-space, we get

$$P_{H_n}(w_n) = w_n - \frac{\langle v_n, w_n - y_n \rangle}{\|v_n\|^2} v_n, \text{ where } v_n = w_n - y_n - \lambda_n (A w_n - A y_n).$$

Theorem 3.2. *Suppose that Assumptions (C1)–(C5) hold. Then the sequence $\{x_n\}$ formed by Algorithm 3.2 converges to $p \in \text{VI}(C, A)$ in norm, where $p = P_{\text{VI}(C, A)}(f(p))$.*

Proof. We divide our proof in four steps.

Claim 1. The sequence $\{x_n\}$ is bounded. Let $z_n = P_{H_n}(w_n)$. From Lemma 2.3 and the property of projection, we get

$$\begin{aligned} \|z_n - p\|^2 &\leq \|w_n - p\|^2 - \|w_n - P_{H_n}(w_n)\|^2 \\ &= \|w_n - p\|^2 - \text{dist}^2(w_n, H_n), \end{aligned} \tag{3.24}$$

which means that $\|z_n - p\| \leq \|w_n - p\|$. Using the same argument as (3.13)–(3.16), we get $\{x_n\}$ is bounded. So $\{w_n\}$, $\{y_n\}$, $\{A w_n\}$, $\{A y_n\}$ and $\{f(x_n)\}$ are also bounded.

Claim 2.

$$[M^{-1}(1-\mu)\|w_n - y_n\|^2]^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3),$$

and

$$\|z_n - w_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3).$$

Indeed, there exists $M > 0$ such that $\|w_n - y_n - \lambda_n(Aw_n - Ay_n)\| \leq M, \forall n$. For any $u, v \in \mathcal{H}$, we show that

$$\begin{aligned} \|h_n(u) - h_n(v)\| &= \|\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), u - v \rangle\| \\ &\leq \|w_n - y_n - \lambda_n(Aw_n - Ay_n)\| \|u - v\| \\ &\leq M\|u - v\|. \end{aligned}$$

This shows that $h_n(x)$ is M -Lipschitz continuous on \mathcal{H} . From Lemma 2.3 and Lemma 3.4, we observe that

$$\text{dist}(w_n, H_n) \geq M^{-1}h_n(w_n) \geq M^{-1}(1-\mu)\|w_n - y_n\|^2,$$

which together with (3.24) implies

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - [M^{-1}(1-\mu)\|w_n - y_n\|^2]^2. \quad (3.25)$$

By the definition of x_{n+1} , (3.17) and (3.25), we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|\alpha_n(f(x_n) - p) + (1 - \alpha_n)(z_n - p)\|^2 \\ &\leq (1 - \alpha_n)\|z_n - p\|^2 + 2\alpha_n\langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|x_n - p\|^2 - [M^{-1}(1-\mu)\|w_n - y_n\|^2]^2 + \alpha_n(Q_2 + Q_3). \end{aligned}$$

where Q_2 and Q_3 are defined in Claim 2 of Theorem 3.1. Moreover, it follows from (3.17) and (3.24) that

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq \|z_n - p\|^2 + 2\alpha_n\langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|x_n - p\|^2 - \|z_n - w_n\|^2 + \alpha_n(Q_2 + Q_3). \end{aligned}$$

The two required results can be obtained through some simple transformations.

Claim 3. We can get the same conclusion as stated in Claim 3 of Theorem 3.1 by using (TC).

Claim 4. The sequence $\{\|x_n - p\|^2\}$ converges to zero. This proof is very similar to Claim 4 in Theorem 3.1. Let $\{\|x_{n_k} - p\|\}$ be a subsequence of $\{\|x_n - p\|\}$ such that (3.19) holds. It follows from Claim 2 and Assumption (C5) that

$$\limsup_{k \rightarrow \infty} [M^{-1}(1-\mu)\|w_{n_k} - y_{n_k}\|^2]^2 \leq 0 \quad \text{and} \quad \limsup_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\|^2 \leq 0.$$

Thus, we get $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$ and $\lim_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\| = 0$. Using the same arguments as (3.20)–(3.22), we get that (3.23) holds. Therefore, from Remark 3.1 (i), Claim 3 and Lemma 2.5, we conclude that $x_n \rightarrow p$ as $n \rightarrow \infty$. The proof of Theorem 3.2 is now complete. \square

3.3. The Algorithm 3. The third method is explained in Algorithm 3.3. This approach improves the results of Vuong and Shehu [14] and Thong et al. [19].

Algorithm 3.3 The first modified inertial extragradient-viscosity algorithm

Initialization: Take $\ell \in (0, 1)$, $\mu > 0$, $\lambda \in (0, 1/\mu)$. Let $x_0, x_1 \in C$ be arbitrary.

Iterative Steps: Calculate x_{n+1} as follows:

$$\begin{cases} w_n = x_n + \theta_n (x_n - x_{n-1}), \\ y_n = P_C (w_n - \lambda A w_n), \\ t_n = w_n - \tau_n r_\lambda (w_n), \quad r_\lambda (w_n) = w_n - y_n, \\ x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n) P_{H_n} (w_n). \end{cases}$$

where $\{\theta_n\}$ is defined in (3.1), $H_n = \{x \in C : h_n(x) \leq 0\}$, $h_n(x) = \langle A t_n, x - t_n \rangle$, $\tau_n = \ell^{m_n}$ and m_n is the smallest non-negative integer m satisfying

$$\langle A w_n - A (w_n - \ell^m r_\lambda (w_n)), r_\lambda (w_n) \rangle \leq \mu \|r_\lambda (w_n)\|^2. \quad (3.26)$$

Before analyzing the convergence of the algorithm, we introduce the following lemmas.

Lemma 3.5. *Suppose that Assumptions (C1)–(C3) hold. The Armijo-line search rule (3.26) is well defined.*

Proof. Since mapping A is uniformly continuous on C and $\ell \in (0, 1)$, we get

$$\lim_{m \rightarrow \infty} \langle A w_n - A (w_n - \ell^m r_\lambda (w_n)), r_\lambda (w_n) \rangle = 0.$$

Furthermore, it can be easily seen that $\|r_\lambda (w_n)\| > 0$ (otherwise, y_n is a solution of (VIP)). Thus, there exists a non-negative integer m_n satisfying (3.26). \square

Lemma 3.6. *Suppose that Assumptions (C1)–(C3) hold. Let x^* be a solution of (VIP). Then $h_n(x^*) \leq 0$ and $h_n(w_n) \geq \tau_n (\lambda^{-1} - \mu) \|r_\lambda (w_n)\|^2$. In particular, if $r_\lambda (w_n) \neq 0$ then $h_n(w_n) > 0$.*

Proof. From $x^* \in \text{VI}(C, A)$ and $t_n \in C$, it follows from Lemma 2.4 that $h_n(x^*) = \langle A t_n, x^* - t_n \rangle \leq 0$. Using the definition of h_n , one sees that

$$h_n(w_n) = \langle A t_n, w_n - t_n \rangle = \langle A t_n, \tau_n r_\lambda (w_n) \rangle = \tau_n \langle A t_n, r_\lambda (w_n) \rangle. \quad (3.27)$$

By the projection property $\|x - P_C(y)\|^2 \leq \langle x - y, x - P_C(y) \rangle$, $\forall x \in C, y \in \mathcal{H}$. Taking $x = w_n$ and $y = w_n - \lambda A w_n$, we get

$$\|w_n - P_C(w_n - \lambda A w_n)\|^2 \leq \lambda \langle A w_n, w_n - P_C(w_n - \lambda A w_n) \rangle,$$

which yields that $\langle A w_n, r_\lambda (w_n) \rangle \geq \lambda^{-1} \|r_\lambda (w_n)\|^2$. From (3.26), one has

$$\begin{aligned} \langle A t_n, r_\lambda (w_n) \rangle &\geq \langle A w_n, r_\lambda (w_n) \rangle - \mu \|r_\lambda (w_n)\|^2 \\ &\geq (\lambda^{-1} - \mu) \|r_\lambda (w_n)\|^2. \end{aligned} \quad (3.28)$$

Combining (3.27) and (3.28), we observe that $h_n(w_n) \geq \tau_n (\lambda^{-1} - \mu) \|r_\lambda (w_n)\|^2$. \square

Lemma 3.7. *Suppose that Assumptions (C1)–(C3) hold. Let sequence $\{w_n\}$ be created by Algorithm 3.3. If $\lim_{n \rightarrow \infty} \tau_n \|r_\lambda (w_n)\|^2 = 0$ then $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$.*

Proof. We show that $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$ by consider two cases of τ_n . First, we assume that $\liminf_{n \rightarrow \infty} \tau_n > 0$. Thus, there exists a positive number τ such that $\tau_n \geq \tau > 0, \forall k \in \mathbb{N}$. Moreover, one sees that

$$\|w_n - y_n\|^2 = \frac{1}{\tau_n} \tau_n \|w_n - y_n\|^2 \leq \frac{1}{\tau} \cdot \tau_n \|w_n - y_n\|^2.$$

Therefore, we get $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$. On the other hand, we assume $\liminf_{n \rightarrow \infty} \tau_n = 0$. In this situation, we suppose that $\{n_k\}$ is a subsequence of $\{n\}$ such that

$$\lim_{k \rightarrow \infty} \tau_{n_k} = 0 \quad \text{and} \quad \lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = a > 0. \quad (3.29)$$

Let $y_k = w_{n_k} - \frac{1}{\ell} \tau_{n_k} (w_{n_k} - y_{n_k})$. It follows that

$$\lim_{k \rightarrow \infty} \|y_k - w_{n_k}\|^2 = \lim_{k \rightarrow \infty} \frac{1}{\ell^2} \tau_{n_k} \cdot \tau_{n_k} \|w_{n_k} - y_{n_k}\|^2 = 0,$$

which together with the fact that mapping A is uniformly continuous gives $\lim_{k \rightarrow \infty} \|Aw_{n_k} - Ay_k\| = 0$. From the definition of y_k and (3.26), we obtain

$$\langle Aw_{n_k} - Ay_k, w_{n_k} - y_{n_k} \rangle > \mu \|w_{n_k} - y_{n_k}\|^2, \quad (3.30)$$

which further yields that $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$. This contradicts the Hypothesis (3.29). Thus, we conclude that $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$. The proof is completed. \square

Theorem 3.3. *Suppose that Assumptions (C1)–(C5) hold. Then the sequence $\{x_n\}$ produced by Algorithm 3.3 converges to $p \in \text{VI}(C, A)$ in norm, where $p = P_{\text{VI}(C, A)}(f(p))$.*

Proof. We divide our proof in four steps.

Claim 1. The sequence $\{x_n\}$ is bounded. Indeed, put $z_n = P_{H_n}(w_n)$, the formula (3.24) also holds. Thus, we get that $\|z_n - p\| \leq \|w_n - p\|$. Using the same facts as (3.13)–(3.16), we have $\{x_n\}$ is bounded. So $\{w_n\}$, $\{y_n\}$, $\{t_n\}$ and $\{f(x_n)\}$ are also bounded.

Claim 2.

$$\|z_n - w_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3),$$

and

$$[L^{-1} \tau_n (\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2]^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3).$$

Combining (3.17) and (3.24), we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \alpha_n) \|z_n - p\|^2 + 2\alpha_n \langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|w_n - p\|^2 - \|z_n - w_n\|^2 + 2\alpha_n \langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|x_n - p\|^2 - \|z_n - w_n\|^2 + \alpha_n(Q_2 + Q_3), \end{aligned}$$

where Q_2 and Q_3 are defined in Claim 2 of Theorem 3.1. The first inequality can be obtained by a simple conversion.

Since $\{At_n\}$ is bounded, there is $L > 0$ such that $\|At_n\| \leq L, \forall n$. For any $u, v \in \mathcal{H}$, we get

$$\|h_n(u) - h_n(v)\| = \|\langle At_n, u - v \rangle\| \leq \|At_n\| \|u - v\| \leq L \|u - v\|,$$

which means that $h_n(x)$ is L -Lipschitz continuous on \mathcal{H} . From Lemma 2.3 and Lemma 3.6, we find that

$$\text{dist}(w_n, H_n) \geq L^{-1} h_n(w_n) \geq L^{-1} \tau_n (\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2.$$

This together with (3.24) gives

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - [L^{-1}\tau_n(\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2]^2.$$

Therefore, we get

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \alpha_n) \|z_n - p\|^2 + 2\alpha_n \langle f(x_n) - p, x_{n+1} - p \rangle \\ &\leq \|x_n - p\|^2 - [L^{-1}\tau_n(\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2]^2 + \alpha_n(Q_2 + Q_3), \end{aligned}$$

A simple transformation of the above equation can get the second inequality.

Claim 3. We can get the same conclusion as stated in Claim 3 of Theorem 3.1 by using (TC).

Claim 4. The sequence $\{\|x_n - p\|^2\}$ converges to zero. This proof is very similar to Claim 4 in Theorem 3.1. Let $\{\|x_{n_k} - p\|\}$ be a subsequence of $\{\|x_n - p\|\}$ such that (3.19) holds. It follows from Claim 2 and Assumption (C5) that

$$\limsup_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\|^2 \leq 0 \text{ and } \limsup_{k \rightarrow \infty} [L^{-1}\tau_{n_k}(\lambda^{-1} - \mu) \|r_\lambda(w_{n_k})\|^2]^2 \leq 0,$$

Therefore, we get $\lim_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\| = 0$ and $\lim_{k \rightarrow \infty} \tau_{n_k} \|r_\lambda(w_{n_k})\|^2 = 0$. It follows from Lemma 3.7 that $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$. Using the same arguments as (3.20)–(3.22), we get that (3.23) holds. Therefore, from Remark 3.1 (i), Claim 3 and Lemma 2.5, we conclude that $x_n \rightarrow p$ as $n \rightarrow \infty$. The proof of Theorem 3.3 is now complete. \square

3.4. The Algorithm 3.4. Finally, we propose the last iterative scheme with a new separating hyperplane. This idea comes from the recent paper by Reich et al. [20], and the method can be seen as a direct modification of Algorithm 3.3.

Algorithm 3.4 The second modified inertial extragradient-viscosity algorithm

Initialization: Take $\ell \in (0, 1)$, $\mu > 0$, $\lambda \in (0, 1/\mu)$. Let $x_0, x_1 \in C$ be arbitrary.

Iterative Steps: Calculate x_{n+1} as follows:

$$\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \lambda A w_n), \\ t_n = w_n - \tau_n r_\lambda(w_n), \quad r_\lambda(w_n) = w_n - y_n, \\ x_{n+1} = \alpha_n f(x_n) + (1 - \alpha_n) P_{H_n}(w_n). \end{cases}$$

where $\{\theta_n\}$ is defined in (3.1), $H_n = \{x \in C : h_n(x) \leq 0\}$, $h_n(x) = \langle A t_n, x - w_n \rangle + \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2$, $\tau_n = \ell^{m_n}$ and m_n is the smallest non-negative integer m satisfying

$$\langle A w_n - A(w_n - \ell^m r_\lambda(w_n)), r_\lambda(w_n) \rangle \leq \mu/2 \|r_\lambda(w_n)\|^2. \quad (3.31)$$

We start the convergence analysis of the algorithm by proving the following lemma.

Lemma 3.8. Suppose that Assumptions (C1)–(C3) hold. Let x^* be a solution of (VIP). Then $h_n(x^*) \leq 0$ and $h_n(w_n) = \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2$. In particular, if $r_\lambda(w_n) \neq 0$ then $h_n(w_n) > 0$.

Proof. It follows from Lemma 3.6 that $\langle A t_n, x^* - t_n \rangle \leq 0$ and $\langle A w_n, r_\lambda(w_n) \rangle \geq \lambda^{-1} \|r_\lambda(w_n)\|^2$. From (3.31), one has

$$\langle A t_n, r_\lambda(w_n) \rangle \geq \langle A w_n, r_\lambda(w_n) \rangle - \frac{\mu}{2} \|r_\lambda(w_n)\|^2 \geq \left(\frac{1}{\lambda} - \frac{\mu}{2} \right) \|r_\lambda(w_n)\|^2,$$

which together with the definition of $h_n(x)$ and t_n yields that

$$\begin{aligned}
 h_n(x^*) &= -\langle At_n, w_n - t_n \rangle + \langle At_n, x^* - t_n \rangle + \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2 \\
 &\leq -\tau_n \langle At_n, r_\lambda(w_n) \rangle + \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2 \\
 &\leq -\frac{\tau_n}{2} (2\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2 + \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2 \\
 &= -\frac{\tau_n}{2} (\lambda^{-1} - \mu) \|r_\lambda(w_n)\|^2 \leq 0.
 \end{aligned}$$

On the other hand, from the definition of $h_n(x)$, it can be easily seen that $h_n(w_n) = \frac{\tau_n}{2\lambda} \|r_\lambda(w_n)\|^2$. This completes the proof. \square

Theorem 3.4. *Suppose that Assumptions (C1)–(C5) hold. Then the sequence $\{x_n\}$ provided by Algorithm 3.4 converges to $p \in \text{VI}(C, A)$ in norm, where $p = P_{\text{VI}(C, A)}(f(p))$.*

Proof. We divide our proof in four steps.

Claim 1. The sequence $\{x_n\}$ is bounded. Take $z_n = P_{H_n}(w_n)$, then the formula (3.24) also holds. Thus, we get that $\|z_n - p\| \leq \|w_n - p\|$. Using the same facts as (3.13)–(3.16), we get that the sequence $\{x_n\}$ is bounded. So the sequences $\{w_n\}$, $\{y_n\}$, $\{t_n\}$ and $\{f(x_n)\}$ are also bounded.

Claim 2.

$$\|z_n - w_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3),$$

and

$$\left[\frac{\tau_n}{2\lambda L} \|r_\lambda(w_n)\|^2 \right]^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n(Q_2 + Q_3).$$

As stated in Claim 2 of Theorem 3.3, we can easily obtain the first inequality. Since $\{At_n\}$ is bounded, there is $L > 0$ such that $\|At_n\| \leq L, \forall n$. For any $u, v \in \mathcal{H}$, we get

$$\|h_n(u) - h_n(v)\| = \|\langle At_n, u - v \rangle\| \leq \|At_n\| \|u - v\| \leq L \|u - v\|,$$

which indicates that $h_n(x)$ is L -Lipschitz continuous on \mathcal{H} . From Lemma 2.3 and Lemma 3.8, we observe that

$$\text{dist}(w_n, H_n) \geq L^{-1} h_n(w_n) = \frac{\tau_n}{2\lambda L} \|r_\lambda(w_n)\|^2.$$

This together with (3.17), (3.24) and the definition of x_{n+1} yields

$$\begin{aligned}
 \|x_{n+1} - p\|^2 &\leq (1 - \alpha_n) \|z_n - p\|^2 + 2\alpha_n \langle f(x_n) - p, x_{n+1} - p \rangle \\
 &\leq \|w_n - p\|^2 - \left[\frac{\tau_n}{2\lambda L} \|r_\lambda(w_n)\|^2 \right]^2 + \alpha_n Q_3 \\
 &\leq \|x_n - p\|^2 - \left[\frac{\tau_n}{2\lambda L} \|r_\lambda(w_n)\|^2 \right]^2 + \alpha_n(Q_2 + Q_3),
 \end{aligned}$$

where Q_2 and Q_3 are defined in Claim 2 of Theorem 3.1. A simple transformation of the above equation can get the second inequality.

Claim 3. We can get the same conclusion as stated in Claim 3 of Theorem 3.1 by using (TC).

Claim 4. The sequence $\{\|x_n - p\|^2\}$ converges to zero. This proof is very similar to Claim 4 in Theorem 3.1. Let $\{\|x_{n_k} - p\|\}$ be a subsequence of $\{\|x_n - p\|\}$ such that (3.19) holds. It follows from Claim 2 and Assumption (C5) that

$$\limsup_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\|^2 \leq 0 \quad \text{and} \quad \limsup_{k \rightarrow \infty} \left[\frac{\tau_{n_k}}{2\lambda L} \|r_\lambda(w_{n_k})\|^2 \right]^2 \leq 0.$$

Therefore, we get $\lim_{k \rightarrow \infty} \|w_{n_k} - z_{n_k}\| = 0$ and $\lim_{k \rightarrow \infty} \tau_{n_k} \|r_\lambda(w_{n_k})\|^2 = 0$. By Lemma 3.7, one has $\lim_{k \rightarrow \infty} \|w_{n_k} - y_{n_k}\| = 0$. Using the same arguments as (3.20)–(3.22), we get that (3.23) holds. Therefore, from Remark 3.1 (i), Claim 3 and Lemma 2.5, we conclude that $x_n \rightarrow p$ as $n \rightarrow \infty$. The proof of Theorem 3.4 is now complete. \square

Remark 3.6. Note that the proposed algorithms 3.1–3.4 have obtained strong convergence theorems in real Hilbert spaces. They can solve non-Lipschitz continuous and pseudomonotone variational inequalities. These two properties give them an advantage and applicability compared with algorithms involving monotone and Lipschitz continuous mappings. It should be point out that Algorithms 3.1 and 3.2 need to calculate multiple projections onto the feasible set in each iteration, while Algorithms 3.3 and 3.4 only require to compute one projection onto the feasible set per iteration. Furthermore, our algorithms combine inertial technology and viscosity method making them converge faster than the algorithms proposed in the literature [14, 16, 17, 18, 19, 20, 34] (see Section 4 and Section 5).

4. NUMERICAL EXAMPLES

In this section, we perform some numerical tests to show the computational efficiency of the proposed iterative schemes and compare them with several previously known strongly convergent algorithms, including the Algorithm 3.3 suggested by Vuong and Shehu [14] (shortly, VS Alg. 3.3), the Algorithm 4.3 proposed by Shehu et al. [16] (SDJ Alg. 4.3), the Algorithm 3.11 introduced by Thong and Gibali [17] (TG Alg. 3.11), the Algorithm 3 offered by Thong et al. [18] (TSI1 Alg. 3), the Algorithm 3.3 presented by Thong et al. [19] (TSI2 Alg. 3.3), the Algorithm 4 given by Reich et al. [20] (RTDLD Alg. 4) and the Algorithm 3.2 provided by Gibali et al. [34] (GTT Alg. 3.2). We use the FOM Solver [35] to solve the projections involved in all algorithms. All the programs were implemented in MATLAB 2018a on a Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz computer with RAM 8.00 GB. The parameters of all algorithms are set as follows.

- For all algorithms, we set $\alpha_n = 1/(n+1)$ and $f(x) = 0.1x$.
- For the suggested Algorithms 3.1–3.4, TG Alg. 3.11, TSI1 Alg. 3, TSI2 Alg. 3.3, RTDLD Alg. 4, we select $\gamma = 1.5$, $\ell = 0.5$, $\mu = 0.4$. Adopt inertial parameters $\tau = 0.4$ and $\varepsilon_n = 100/(n+1)^2$ in our proposed algorithms. In addition, we set $\lambda = 0.5/\mu$ in TSI2 Alg. 3.3, RTDLD Alg. 4 and our Algorithms 3.3 and 3.4.
- Pick $\gamma = 0.5$ and $\sigma = 0.4$ in VS Alg 3.3. For GTT Alg. 3.2, we choose $\lambda = 1.5$, $\ell = 0.5$, $\mu = 0.4$, $\gamma = 1.5$. Take the fixed step size $\lambda_n = 0.5/L$ and $\gamma = 1.5$ in SDJ Alg. 4.3.

Example 4.1. Consider the form of linear operator $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ as follows: $A(x) = Gx + g$, where $g \in \mathbb{R}^m$ and $G = BB^\top + S + E$, matrix $B \in \mathbb{R}^{m \times m}$, matrix $S \in \mathbb{R}^{m \times m}$ is skew-symmetric, and matrix $E \in \mathbb{R}^{m \times m}$ is diagonal matrix whose diagonal terms are non-negative (hence G is positive symmetric definite). We choose the feasible set C is a box constraint with the form $C = [-2, 5]^m$. It is easy to see that A is Lipschitz continuous monotone and its Lipschitz constant $L = \|G\|$. In this numerical example, all entries of B, E are generated randomly in $[0, 2]$ and S is generated randomly in $[-2, 2]$. Let $q = \mathbf{0}$. Then the solution set is $x^* = \{\mathbf{0}\}$. The maximum number of iterations of 200 as a common stopping criterion and the initial values $x_0 = x_1$ are randomly generated by $rand(m, 1)$ in MATLAB. We use $D_n = \|x_n - x^*\|$ to measure the n -th iteration error of all algorithms. The numerical results of all algorithms with four different dimensions are shown in Fig. 2.

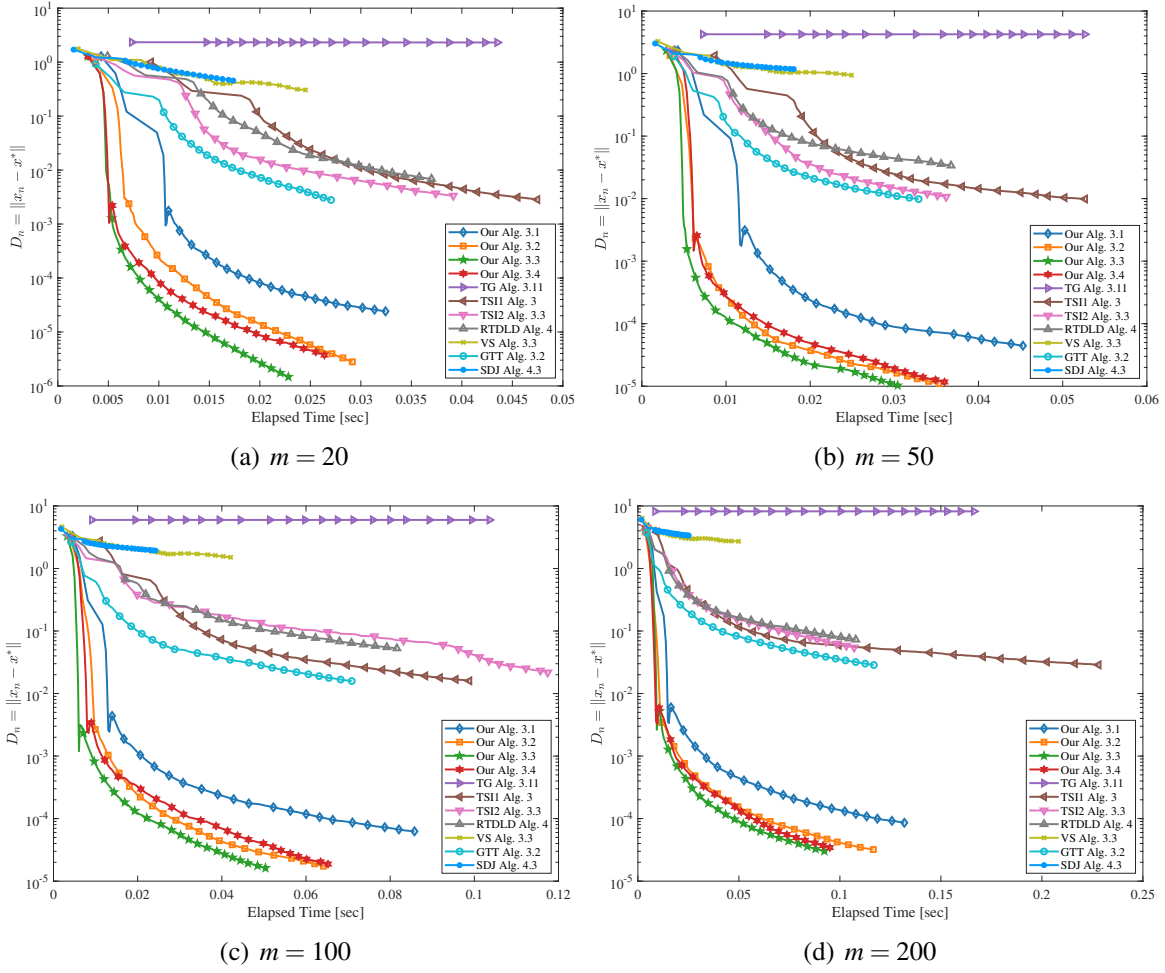


FIGURE 2. Numerical results for Example 4.1

Example 4.2. In this example, we consider a test that occurs in an infinite-dimensional space. Let $\mathcal{H} = L^2([0, 1])$. Its inner product and norm are defined as $\langle x, y \rangle = \int_0^1 x(t)y(t) dt$ and $\|x\| = (\int_0^1 |x(t)|^2 dt)^{1/2}$, respectively. Assume that the feasible set is a ball and its form is $C = \{x \in \mathcal{H} : \|x\| \leq 2\}$. Define a mapping $h : C \rightarrow \mathbb{R}$ by $h(m) = 1/(1 + \|m\|^2)$. It is easy to verify that the mapping h is bounded ($h(m) \in [0.2, 1]$) and L_h -Lipschitz continuous with $L_h = 16/25$. Recall that the Volterra integration operator $V : \mathcal{H} \rightarrow \mathcal{H}$ is given by

$$V(m)(t) = \int_0^t m(s) ds, \quad \forall t \in [0, 1], m \in \mathcal{H}.$$

Then V is bounded linear monotone (see [36, Exercise 20.12]) and its operator norm is $\|V\| = \frac{2}{\pi}$. Now, we define the mapping $A : C \rightarrow \mathcal{H}$ as follows:

$$A(m)(t) = h(m)V(m)(t), \quad \forall t \in [0, 1], m \in C.$$

Note that A is not monotone. For example, take $n = 1$ and $m = 2$, then $\langle An - Am, n - m \rangle = -\frac{1}{10} < 0$. In fact, A is pseudomonotone. Indeed, for all $m, n \in C$, assume that $\langle Am, n - m \rangle \geq 0$. Next we show that $\langle An, n - m \rangle \geq 0$. Note that $\langle Vm, n - m \rangle \geq 0$ (since $h(m) > 0.2$). Therefore,

we obtain

$$\begin{aligned} \langle An, n - m \rangle &= h(n) \langle V(n), n - m \rangle \geq h(n) [\langle V(n), n - m \rangle - \langle V(m), n - m \rangle] \\ &= h(n) \langle V(n) - V(m), n - m \rangle \geq 0. \end{aligned}$$

Hence, mapping A is pseudomonotone. Moreover, we have

$$\begin{aligned} \|Am - An\| &= \|h(m)V(m) - h(n)V(n)\| \\ &\leq \|h(m)V(m) - h(m)V(n)\| + \|h(m)V(n) - h(n)V(n)\| \\ &\leq |h(m)| \|V(m) - V(n)\| + \|V(n)\| \|h(m) - h(n)\| \\ &\leq (|h(m)| \|V\| + \|V\| \|n\| L_g) \|m - n\| \\ &\leq \frac{114}{25\pi} \|m - n\|, \quad \forall m, n \in C. \end{aligned}$$

Thus, mapping A is L -Lipschitz continuous with $L = 114/(25\pi)$. We choose the maximum number of iterations of 50 as a common stopping criterion and use $D_n = \|x_{n+1} - x_n\|$ to measure the error of the n -th step since we do not know the solution of the problem. Fig. 3 shows the numerical behaviors of all algorithms with four starting points.

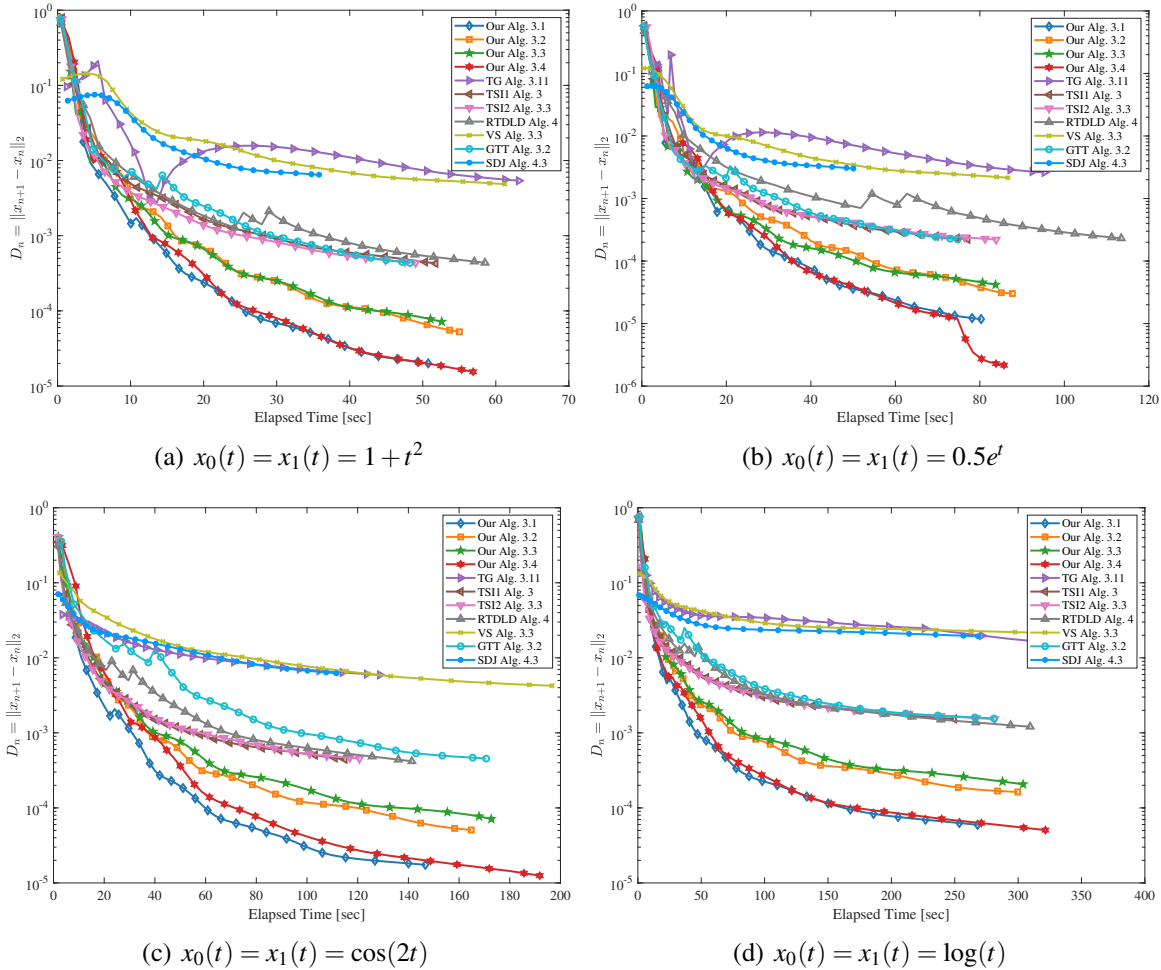


FIGURE 3. Numerical results for Example 4.2

Remark 4.1. As seen from Examples 4.1 and 4.2, the proposed schemes are easy to implement and efficient. Note that the choice of initial values and dimensions does not significantly affect on CPU time required to reach the same stopping criterion. Moreover, our algorithms embed inertial terms making them converge faster than some known algorithms in the literature [14, 16, 17, 18, 19, 20, 34]. It is worth noting that VS Alg. 3.3, TG Alg. 3.11 and SDJ Alg. 4.3 converge very slowly because they use the Halpern method to ensure strong convergence.

5. APPLICATIONS TO OPTIMAL CONTROL PROBLEMS

In this section, we use the proposed algorithms to solve the variational inequality that occurs in the optimal control problem. Assume that $L_2([0, T], \mathbb{R}^m)$ represents the square-integrable Hilbert space with inner product $\langle p, q \rangle = \int_0^T \langle p(t), q(t) \rangle dt$ and norm $\|p\|_2 = \sqrt{\langle p, p \rangle}$. The optimal control problem is described as follows:

$$p^*(t) \in \text{Argmin}\{g(p) \mid p \in V\}, \quad t \in [0, T], \quad (5.1)$$

where V represents a set of feasible controls composed of m piecewise continuous functions. Its form is expressed as follows:

$$V = \{p(t) \in L_2([0, T], \mathbb{R}^m) : p_i(t) \in [p_i^-, p_i^+], i = 1, 2, \dots, m\}. \quad (5.2)$$

In particular, the control $p(t)$ may be a piecewise constant function (bang-bang type). The terminal objective function has the form

$$g(p) = \Phi(x(T)), \quad (5.3)$$

where Φ is a convex and differentiable defined on the attainability set.

Assume that the trajectory $x(t) \in L_2([0, T])$ satisfies the constraints of the linear differential equation system:

$$\frac{d}{dt}x(t) = Q(t)x(t) + W(t)p(t), \quad 0 \leq t \leq T, \quad x(0) = x_0, \quad (5.4)$$

where $Q(t) \in \mathbb{R}^{n \times n}$, $W(t) \in \mathbb{R}^{n \times m}$ are given continuous matrices for every $t \in [0, T]$. By the solution of problem (5.1)–(5.4), we mean a control $p^*(t)$ and a corresponding (optimal) trajectory $x^*(t)$ such that its terminal value $x^*(T)$ minimizes objective function (5.3). From the Pontryagin maximum principle, there exists a function $s^* \in L_2([0, T])$ such that the triple (x^*, s^*, p^*) solves for a.e. $t \in [0, T]$ the system

$$\frac{d}{dt}x^*(t) = Q(t)x^*(t) + W(t)p^*(t), \quad x^*(0) = x_0, \quad (5.5)$$

$$\frac{d}{dt}s^*(t) = -Q(t)^\top s^*(t), \quad s^*(T) = \nabla \Phi(x^*(T)), \quad (5.6)$$

$$0 \in W(t)^\top s^*(t) + N_V(p^*(t)), \quad (5.7)$$

where $N_V(p)$ is the normal cone to V at p defined by

$$N_V(p) = \begin{cases} \emptyset, & \text{if } p \notin V; \\ \{t \in \mathcal{H} : \langle t, q - p \rangle \leq 0, \forall q \in V\}, & \text{if } p \in V. \end{cases}$$

Denoting $Gp(t) = W(t)^\top s(t)$, Khoroshilova [37] showed that Gp is the gradient of the objective function g . Therefore, system (5.5)–(5.7) is reduced to the variational inequality problem

$$\langle Gp^*, q - p^* \rangle \geq 0, \quad \forall q \in V.$$

Recently, there are many approaches to solve the optimal control problem (see, e.g., [14, 37, 38]). Note that our algorithms 3.1–3.4 guarantee strong convergence and do not require the Lipschitz constant. Moreover, the addition of inertial terms makes them converge faster.

For the convenience of numerical computation, we discretize the continuous functions. Given the mesh size $h = T/N$, where N is a natural number. We identify any discretized control $p^N = (p_0, p_1, \dots, p_{N-1})$ with its piece-wise constant extension:

$$p^N(t) = p_i, \quad \forall t \in [t_i, t_{i+1}), \quad t_i = ih, \quad i = 0, 1, \dots, N.$$

Furthermore, we identify the discretized state $x^N = (x_0, x_1, \dots, x_N)$ and co-state $s^N = (s_0, s_1, \dots, s_N)$. They have the form of piecewise linear interpolation:

$$x^N(t) = x_i + \frac{t - t_i}{h} (x_{i+1} - x_i), \quad \forall t \in [t_i, t_{i+1}), \quad i = 0, 1, \dots, N-1,$$

and

$$s^N(t) = s_i + \frac{t_i - t}{h} (s_{i-1} - s_i), \quad \forall t \in (t_{i-1}, t_i], \quad i = N, N-1, \dots, 1.$$

We use the classical Euler discretization method to solve the systems of ODEs (5.5) and (5.6). Thus, the Euler discretization of the original system (5.1)–(5.4) is given by

$$\begin{aligned} & \text{minimize} \quad \Phi_N(x^N, p^N) \\ & \text{subject to} \quad x_{i+1}^N = x_i^N + h [Q(t_i)x_i^N + W(t_i)p_i^N], \quad x^N(0) = x_0, \\ & \quad \quad \quad s_i^N = s_{i+1}^N + hQ(t_i)^\top s_{i+1}^N, \quad s(N) = \nabla\Phi(x_N), \\ & \quad \quad \quad p_i^N \in V. \end{aligned}$$

It is well known that the Euler discretization has the error estimate $O(h)$ [39]. This indicates that the difference between the discretized solution $p^N(t)$ and the original solution $p^*(t)$ is proportional to the mesh size h . That is, there exists a constant $K > 0$ such that $\|p^N - p^*\| \leq Kh$.

Next, we present several examples to illustrate the numerical performance of all the algorithms (except the Algorithm 4.3 proposed by Shehu et al. [16] since we do not know the Lipschitz constant of the mapping related to the problems). For all algorithms, we set $\alpha_n = 10^{-4}/(n+1)$. Take inertial parameters $\tau = 0.4$ and $\varepsilon_n = 10^{-4}/(n+1)^2$ in the suggested iterative schemes. The remaining parameters are the same as those set in Section 4. The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-1, 1]$ and the stopping criterion is $\|p_{n+1} - p_n\| \leq 10^{-4}$ or the maximum number of iterations of 1000.

Example 5.1 (Control of a harmonic oscillator, see [40]).

$$\begin{aligned} & \text{minimize} \quad x_2(3\pi) \\ & \text{subject to} \quad \dot{x}_1(t) = x_2(t), \\ & \quad \quad \quad \dot{x}_2(t) = -x_1(t) + p(t), \quad \forall t \in [0, 3\pi], \\ & \quad \quad \quad x(0) = 0, \\ & \quad \quad \quad p(t) \in [-1, 1]. \end{aligned}$$

The exact optimal control of Example 5.1 is known:

$$p^*(t) = \begin{cases} 1, & \text{if } t \in [0, \pi/2) \cup (3\pi/2, 5\pi/2); \\ -1, & \text{if } t \in (\pi/2, 3\pi/2) \cup (5\pi/2, 3\pi]. \end{cases}$$

Fig. 4 shows the approximate optimal control and the corresponding trajectories of Algorithm 3.3.

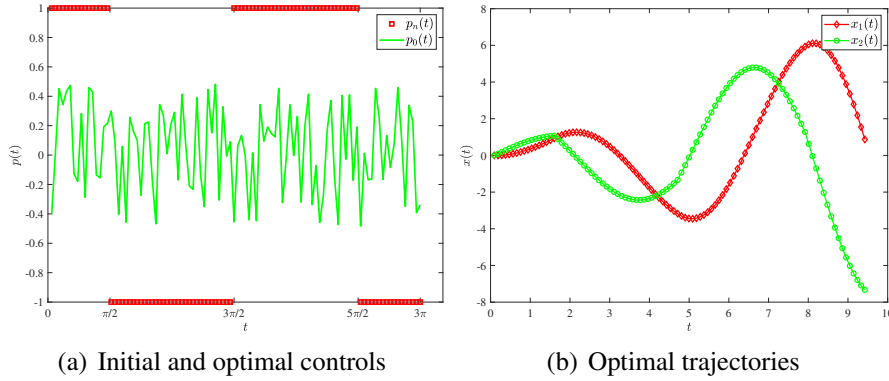


FIGURE 4. Numerical results for Example 5.1

We now consider an example in which the terminal function is not linear.

Example 5.2 (Rocket car [38]).

$$\begin{aligned} & \text{minimize } 0.5 \left((x_1(5))^2 + (x_2(5))^2 \right), \\ & \text{subject to } \dot{x}_1(t) = x_2(t), \\ & \quad \dot{x}_2(t) = p(t), \quad \forall t \in [0, 5], \\ & \quad x_1(0) = 6, \quad x_2(0) = 1, \\ & \quad p(t) \in [-1, 1]. \end{aligned}$$

The exact optimal control of Example 5.2 is

$$p^* = \begin{cases} 1 & \text{if } t \in (3.517, 5]; \\ -1 & \text{if } t \in (0, 3.517]. \end{cases}$$

Fig. 5 shows the approximate optimal control and the corresponding trajectories of Algorithm 3.2.

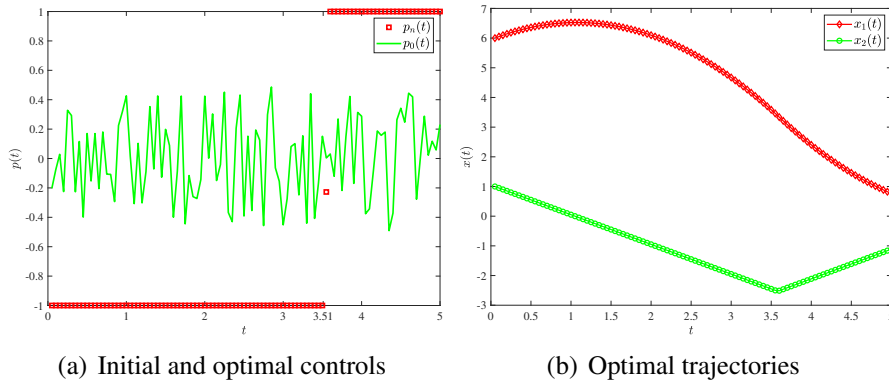


FIGURE 5. Numerical results for Example 5.2

Finally, the computational results of all algorithms in Examples 5.1 and 5.2 are shown in Table 1.

TABLE 1. Comparison of the number of iterations and execution time of all algorithms in Examples 5.1 and 5.2

Algorithms	Example 5.1		Example 5.2	
	Iter.	Time (s)	Iter.	Time (s)
Our Alg. 3.1	1000	0.33185	1000	1.4964
Our Alg. 3.2	90	0.032926	91	0.11501
Our Alg. 3.3	90	0.035086	1000	1.1794
Our Alg. 3.4	1000	0.41878	1000	1.5389
TG Alg. 3.11	1000	0.39974	1000	1.6707
TSI1 Alg. 3	91	0.046058	91	0.13369
TSI2 Alg. 3.3	91	0.044275	1000	1.2113
RTDLD Alg. 4	1000	0.43071	1000	1.6214
VS Alg. 3.3	100	0.043549	117	0.06323
GTT Alg. 3.2	101	0.042715	355	0.46617

Remark 5.1. From Fig. 4, Fig. 5 and Table 1, we know that the suggested algorithms can work well when the terminal function is linear or nonlinear. Moreover, the step size of the Algorithm 4.3 suggested by Shehu et al. [16] requires the prior information of the Lipschitz constant of the mapping, and our algorithms can automatically update the iteration step size.

6. FINAL REMARKS

This paper presented four extragradient projection-type algorithms to solve pseudomonotone and uniformly continuous variational inequalities in real Hilbert spaces. Strong convergence theorems of the suggested methods are confirmed under some suitable conditions. The advantage of our approaches is that no prior knowledge of the Lipschitz constant of the mapping is required. In addition, our algorithms add inertial terms to speed up the convergence speed of the algorithms without inertial terms. Numerical experiments in finite- and infinite-dimensional spaces show that the computational performance of the proposed iterative schemes is higher compared with previously known ones. Moreover, our methods are designed to solve optimal control problems. The schemes obtained in this paper improved and unified relevant results in the literature.

REFERENCES

- [1] N.T. An, N.M. Nam, X. Qin, Solving k -center problems involving sets based on optimization techniques, *J. Global Optim.* 76 (2020), 189–209.
- [2] A.A. Khan, D. Motreanu, Inverse problems for quasi-variational inequalities, *J. Global Optim.* 70 (2018), 401–411.
- [3] D.R. Sahu, J.C. Yao, M. Verma, K.K. Shukla, Convergence rate analysis of proximal gradient methods with applications to composite minimization problems, *Optimization* 70 (2021), 75–100.
- [4] Q.H. Ansari, M. Islam, J.C. Yao, Nonsmooth variational inequalities on Hadamard manifolds, *Appl. Anal.* 99 (2020), 340–358.

- [5] G.M. Korpelevich, An extragradient method for finding saddle points and for other problems, *Èkonom. i Mat. Metody* 12 (1976), 747–756.
- [6] A.N. Iusem, An iterative algorithm for the variational inequality problem, *Comput. Appl. Math.* 13 (1994), 103–114.
- [7] M.V. Solodov, P. Tseng, Modified projection-type methods for monotone variational inequalities, *SIAM J. Control Optim.* 34 (1996), 1814–1830.
- [8] A.N. Iusem, B.F. Svaiter, A variant of Korpelevich’s method for variational inequalities with a new search strategy, *Optimization* 42 (1997), 309–321.
- [9] M.V. Solodov, B.F. Svaiter, A new projection method for variational inequality problems, *SIAM J. Control Optim.* 37 (1999), 765–776.
- [10] B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. Optim.* 35 (1997), 69–76.
- [11] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.* 38 (2000), 431–446.
- [12] Y. Censor, A. Gibali, S. Reich, Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space, *Optim. Methods Softw.* 26 (2011), 827–845.
- [13] Y. Malitsky, Projected reflected gradient methods for monotone variational inequalities, *SIAM J. Optim.* 25 (2015), 502–520.
- [14] P.T. Vuong, Y. Shehu, Convergence of an extragradient-type method for variational inequality with applications to optimal control problems, *Numer. Algorithms* 81 (2019), 269–291.
- [15] Y. Shehu, P. Cholamjiak, Iterative method with inertial for variational inequalities in Hilbert spaces, *Calcolo* 56 (2019), Article ID 4.
- [16] Y. Shehu, Q.L. Dong, D. Jiang, Single projection method for pseudo-monotone variational inequality in Hilbert spaces, *Optimization* 68 (2019), 385–409.
- [17] D.V. Thong, A. Gibali, Extragradient methods for solving non-Lipschitzian pseudo-monotone variational inequalities, *J. Fixed Point Theory Appl.* 21 (2019), Article ID 20.
- [18] D.V. Thong, Y. Shehu, O.S. Iyiola, Weak and strong convergence theorems for solving pseudo-monotone variational inequalities with non-Lipschitz mappings, *Numer. Algorithms* 84 (2020), 795–823.
- [19] D.V. Thong, Y. Shehu, O.S. Iyiola, A new iterative method for solving pseudomonotone variational inequalities with non-Lipschitz operators, *Comput. Appl. Math.* 39 (2020), Article ID 108.
- [20] S. Reich, D.V. Thong, Q.L. Dong, X.H. Li, V.T. Dung, New algorithms and convergence theorems for solving variational inequalities with non-Lipschitz mappings, *Numer. Algorithms*. 2020. DOI: 10.1007/s11075-020-00977-8.
- [21] B.T. Polyak, Some methods of speeding up the convergence of iteration methods, *USSR Comput. Math. Math. Phys.* 4 (1964), 1–17.
- [22] Y. Shehu, O.S. Iyiola, X.H. Li, Q.L. Dong, Convergence analysis of projection method for variational inequalities, *Comput. Appl. Math.* 38 (2019), Article ID 161.
- [23] Y. Shehu, X.H. Li, Q.L. Dong, An efficient projection-type method for monotone variational inequalities in Hilbert spaces, *Numer. Algorithms* 84 (2020), 365–388.
- [24] L. Liu, X. Qin, Strong convergence of an extragradient-like algorithm involving pseudo-monotone mappings, *Numer. Algorithms* 83 (2020), 1577–1590.
- [25] B. Tan, S. Xu, S. Li, Inertial shrinking projection algorithms for solving hierarchical variational inequality problems, *J. Nonlinear Convex Anal.* 21 (2020), 871–884.
- [26] B. Tan, J. Fan, S. Li, Self-adaptive inertial extragradient algorithms for solving variational inequality problems, *Comput. Appl. Math.* 40 (2021), Article ID 19.
- [27] Z. Zhou, B. Tan, S. Li, A new accelerated self-adaptive stepsize algorithm with excellent stability for split common fixed point problems, *Comput. Appl. Math.* 39 (2020), Article ID 220.
- [28] A. Beck, *First-Order Methods in Optimization*, SIAM, Philadelphia, 2017.
- [29] 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.
- [30] A. Iusem, R.G. Otero, Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, *Numer. Funct. Anal. Optim.* 22 (2001), 609–640.

- [31] Y. He, A new double projection algorithm for variational inequalities, *J. Comput. Appl. Math.* 185 (2006), 166–173.
- [32] R.W. Cottle, J.C. Yao, Pseudo-monotone complementarity problems in Hilbert space, *J. Optim. Theory Appl.* 75 (1992), 281–295.
- [33] S. Saejung, P. Yotkaew, Approximation of zeros of inverse strongly monotone operators in Banach spaces, *Nonlinear Anal.* 75 (2012), 742–750.
- [34] A. Gibali, D.V. Thong, P.A. Tuan, Two simple projection-type methods for solving variational inequalities, *Anal. Math. Phys.* 9 (2019), 2203–2225.
- [35] A. Beck, N. Guttmann-Beck, FOM—a MATLAB toolbox of first-order methods for solving convex optimization problems, *Optim. Methods Softw.* 34 (2019), 172–193.
- [36] H.H. Bauschke, P.L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.
- [37] E.V. Khoroshilova, Extragradient-type method for optimal control problem with linear constraints and convex objective function, *Optim. Lett.* 7 (2013), 1193–1214.
- [38] J. Preininger, P.T. Vuong, On the convergence of the gradient projection method for convex optimal control problems with bang-bang solutions, *Comput. Optim. Appl.* 70 (2018), 221–238.
- [39] J.F. Bonnans, A. Festa, Error estimates for the Euler discretization of an optimal control problem with first-order state constraints, *SIAM J. Numer. Anal.* 55 (2017), 445–471.
- [40] A. Pietrus, T. Scarinci, V.M. Veliov, High order discrete approximations to Mayer’s problems for linear systems, *SIAM J. Control Optim.* 56 (2018), 102–119.