Analysis and Mathematical Physics (2022) 12:26
https://doi.org/10.1007/s13324-021-00638-6

Modified inertial projection and contraction algorithms for solving variational inequality problems with non-Lipschitz continuous operators

Bing Tan[1] · **Xiaolong Qin**[2]

Received: 6 August 2021 / Revised: 14 December 2021 / Accepted: 16 December 2021

Abstract

In this paper, we present four modified inertial projection and contraction methods to solve the variational inequality problem with a pseudo-monotone and non-Lipschitz continuous operator in real Hilbert spaces. Strong convergence theorems of the proposed algorithms are established without the prior knowledge of the Lipschitz constant of the operator. Several numerical experiments and the applications to optimal control problems are provided to verify the advantages and efficiency of the proposed algorithms.

Keywords Variational inequality · Projection and contraction method · Subgradient extragradient method · Inertial method · Pseudomonotone mapping · Uniformly continuous

Mathematics Subject Classification 47J20 · 47J25 · 47J30 · 68W10 · 65K15

1 Introduction

Our goal in this paper is to develop some adaptive strongly convergent iterative algorithms to solve variational inequality problems in infinite-dimensional Hilbert spaces.

✉ Xiaolong Qin
qxlxajh@163.com

Bing Tan
bingtan72@gmail.com

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

2 Department of Mathematics, Zhejiang Normal University, Zhejiang, China

Published online: 31 December 2021

Recall that the variational inequality problem is formed as follows:

$$\text{find } x^* \in C \text{ such that } \langle Ax^*, x - x^* \rangle \geq 0, \quad \forall x \in C. \tag{VIP}$$

where C is a nonempty, closed, and convex subset of a real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, and $A : \mathcal{H} \to \mathcal{H}$ is a nonlinear operator. We denote the solution set of (VIP) by $\mathrm{VI}(C, A)$ and assume that the set is nonempty throughout the paper. The variational inequality problem has attracted extensive research as one of the important problems in applied mathematics, and it can be interconverted with fixed point problems and split feasibility problems. The variational inequality can be used as a model for solving many practical problems, such as optimal control problems, image processing problems, signal recovery problems, and so on; see, e.g., [1–4].

We next recall some known solution methods in the literature for solving variational inequality problems, which motivate us to develop new iterative algorithms. Recently, the extragradient method (shortly, EGM) proposed by Korpelevich [5] has received a lot of attention due to its simple form and easy implementation in numerical experiments. More precisely, the EGM is a two-step iterative scheme and produces the following iterative process:

$$\begin{cases} s_n = P_C(x_n - \lambda A x_n), \\ x_{n+1} = P_C(x_n - \lambda A s_n), \end{cases} \tag{EGM}$$

where P_C is the metric projection of $\mathcal{H}$ onto C defined by $P_C(x) := \arg\min_{y \in C}\{\|x - y\|\}$ for a point $x \in \mathcal{H}$, mapping A is monotone and L-Lipschitz continuous, and fixed step size λ is in $(0, 1/L)$. It is noted that the drawback of the EGM is the computational effort due to the fact that it needs to compute the projection on the feasible set twice in each iteration. It is known that computing the projection is equivalent to an optimal distance problem. The corresponding optimization problem may not be easy to solve when the form of the feasible set is complex. Therefore, how to improve the computational efficiency of extragradient type methods has become a hot topic of research. In the past decades, a large number of methods that require computing the projection on the feasible set only once have been proposed to improve the computational efficiency of the EGMs, such as the Tseng's extragradient method introduced by Tseng [6], the subgradient extragradient method (shortly, SEGM) suggested by Censor, Gibali and Reich [7–9], and the projection and contraction method (shortly, PCM) presented by He [10]. We focus on the SEGM and the PCM in this paper. Recall that the SEGM is described as follows:

$$\begin{cases} s_n = P_C(x_n - \lambda A x_n), \\ x_{n+1} = P_{H_n}(x_n - \lambda A s_n), \end{cases} \tag{SEGM}$$

where H_n is defined by

$$H_n := \{x \in \mathcal{H} \mid \langle x_n - \lambda A x_n - s_n, x - s_n \rangle \leq 0\}, \tag{1.1}$$

and the fixed step size $\lambda \in (0, 1/L)$. It is worth noting that the projection onto a half-space H_n can be calculated by an explicit formula. Thus, the SEGM actually only needs to compute the projection on the feasible set once in each iteration, which greatly improves the computational efficiency of the EGM.

It is important to note that the EGM uses the same step size in both projection steps in each iteration, and the SEGM also has this observation. The appropriate choice of step sizes is very significant for the convergence speed of the algorithm. Indeed, there is a method that uses two different step sizes in each iteration, called the projection and contraction method [10]. More precisely, the PCM is expressed in the following form:

$$\begin{cases} s_n = P_C(x_n - \lambda_n A x_n), \\ x_{n+1} = x_n - \tau \lambda_n \sigma_n d_n, \end{cases} \tag{PCM}$$

where $\tau \in (0, 2)$, $\lambda_n \in (0, 1/L)$ or $\{\lambda_n\}$ is selected self-adaptively, and

$$\sigma_n := \frac{\langle x_n - s_n, d_n \rangle}{\|d_n\|^2}, \quad d_n := x_n - s_n - \lambda_n (A x_n - A s_n). \tag{1.2}$$

The numerical experiments given in [11] show that the (PCM) is approximately half the computational effort of the (EGM). Recently, inspired by the (SEGM) and the (PCM), Dong et al. [12] introduced a modified subgradient extragradient method (shortly, MSEGM) using two different step sizes in each iteration to solve the variational inequality problem. Their method is stated as follows:

$$\begin{cases} s_n = P_C(x_n - \lambda_n A x_n), \\ x_{n+1} = P_{H_n}(x_n - \tau \lambda_n \sigma_n A s_n), \end{cases} \tag{MSEGM}$$

where $\tau \in (0, 2)$, H_n, and σ_n are defined in (1.1) and (1.2), respectively. The step size λ_n is chosen to be the largest $\lambda \in \{\gamma, \gamma\ell, \gamma\ell^2, \ldots\}$ such that

$$\lambda \|A x_n - A s_n\| \le \phi \|x_n - s_n\|, \quad \gamma > 0, \ell \in (0, 1), \phi \in (0, 1). \tag{1.3}$$

This step size search procedure (1.3) is known as the Armijo step size criterion. Notice that Algorithm (MSEGM) can work without the prior knowledge of the Lipschitz constant of the mapping by applying the step size update criterion (1.3). The computational efficiency of the proposed algorithm with respect to some previously known schemes was verified by some numerical experiments.

Notice that Algorithm (MSEGM) is only weakly convergent in an infinite-dimensional Hilbert space. Many practical applications occurring in machine learning tell us that strong convergence is preferable to weak convergence in infinite-dimensional spaces. Recently, combining the (MSEGM), the Mann-type method, and the viscosity-type method, Thong and Gibali [13] proposed two new strongly convergent algorithms for solving the (VIP) with a monotone operator in real Hilbert spaces. More precisely, their algorithms are expressed as follows:

$$\begin{cases} s_n = P_C\left(x_n - \lambda_n A x_n\right), \\ t_n = P_{H_n}\left(x_n - \tau \lambda_n \sigma_n A s_n\right), \\ x_{n+1} = \left(1 - \zeta_n - \sigma_n\right) x_n + \sigma_n t_n, \end{cases} \tag{TG Alg. 3.1}$$

and

$$\begin{cases} s_n = P_C\left(x_n - \lambda_n A x_n\right), \\ t_n = P_{H_n}\left(x_n - \tau \lambda_n \sigma_n A s_n\right), \\ x_{n+1} = \zeta_n f\left(x_n\right) + \left(1 - \zeta_n\right) t_n, \end{cases} \tag{TG Alg. 3.2}$$

where $\tau \in (0, 2)$, H_n is defined in (1.1), λ_n is generated by (1.3), and σ_n is defined by

$$\sigma_n := (1 - \phi)\frac{\|x_n - s_n\|^2}{\|d_n\|^2}, \quad d_n := x_n - s_n - \lambda_n\left(A x_n - A s_n\right). \tag{1.4}$$

In addition, Gibali et al. [14] provided two iterative schemes to solve the monotone variational inequality problem in real Hilbert spaces. The form of their algorithms are as follows:

$$\begin{cases} s_n = P_C\left(x_n - \lambda_n A x_n\right), \\ t_n = x_n - \tau \sigma_n d_n, \\ x_{n+1} = \left(1 - \zeta_n - \sigma_n\right) x_n + \sigma_n t_n, \end{cases} \tag{GTT Alg. 3.1}$$

and

$$\begin{cases} s_n = P_C\left(x_n - \lambda_n A x_n\right), \\ t_n = x_n - \tau \sigma_n d_n, \\ x_{n+1} = \zeta_n f\left(x_n\right) + \left(1 - \zeta_n\right) t_n, \end{cases} \tag{GTT Alg. 3.2}$$

where $\tau \in (0, 2)$, λ_n is generated by (1.3), σ_n, and d_n are defined in (1.4).

Notice that the strong convergence of the algorithms presented in [13,14] is guaranteed in the case that mapping A is monotonic. It is known that the class of pseudo-monotone mappings includes the class of monotone mappings. Due to the wide range of applications of pseudo-monotone mappings in practice, there has been a great interest in the pseudo-monotone variational inequality problem and a large number of iterative algorithms have been proposed to solve it; see, for example, [15–18] and the references therein. On the other hand, the inertial idea has been studied by many researchers as a technique to accelerate the convergence speed of algorithms. A common feature of inertial-type methods is that the next iteration depends on the combination of the previous two (or more) iterations. Notice that this small change can improve the convergence speed of algorithms without inertial. Recently, a large number of numerical methods have been constructed to solve variational inequalities, split feasibilities, fixed point problems, and other optimization problems; see, e.g., [17–23] and the references therein.

Inspired and motivated by the above works, in this paper, we propose four new inertial extragradient methods with Armijo step sizes to solve the variational inequality problem involving a pseudo-monotone operator in real Hilbert spaces. Strong convergence theorems of the proposed algorithms are established under the condition that the Lipschitz continuity of the operators is not required. Some numerical experiments are given to verify the computational efficiency of the proposed algorithms.

The remainder of this paper is organized as follows. Some important definitions and lemmas are reviewed in Sect. 2. Section 3 introduces four new adaptive inertial extragradient methods and proves their strong convergence. We report some numerical experiments and applications to optimal control problems in Sect. 4 and conclude the paper in Sect. 5.

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 $\{x_n\}$ to x are represented by $x_n \rightharpoonup x$ and $x_n \to x$, 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) $\|\tau x + \beta y + \lambda z\|^2 = \tau\|x\|^2 + \beta\|y\|^2 + \lambda\|z\|^2 - \tau\beta\|x - y\|^2 - \tau\lambda\|x - z\|^2 - \beta\lambda\|y - z\|^2$, where $\tau, \beta, \lambda \in [0, 1]$ with $\tau + \beta + \lambda = 1$.

For any $x, y \in \mathcal{H}$, a mapping $A : \mathcal{H} \to \mathcal{H}$ is said to be:

1. *L-Lipschitz continuous* with $L > 0$ if $\|Ax - Ay\| \leq L\|x - y\|$ (If $L \in (0, 1)$, then mapping A is called *contraction*).
2. *monotone* if $\langle Ax - Ay, x - y\rangle \geq 0$.
3. *pseudo-monotone* if $\langle Ax, y - x\rangle \geq 0 \Longrightarrow \langle Ay, y - x\rangle \geq 0$.
4. *sequentially weakly continuous* if for each sequence $\{x_n\}$ converges weakly to x implies that $\{Ax_n\}$ converges weakly to Ax.

It is known that P_C has the following basic properties:

$$\langle x - P_C(x), y - P_C(x)\rangle \leq 0,\ \forall x \in \mathcal{H}, \forall y \in C, \tag{2.1}$$

and

$$\|P_C(x) - P_C(y)\|^2 \leq \langle P_C(x) - P_C(y), x - y\rangle,\ \forall x, y \in \mathcal{H}. \tag{2.2}$$

We give some explicit formulas to calculate projections on special feasible sets.

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

$$P_{H_{u,v}}(x) = x - \max\left\{\frac{\langle u, x\rangle - v}{\|u\|^2}, 0\right\} u.$$

2. The projection of x onto a box $\mathrm{Box}[a, b] = \{x : a \leq x \leq b\}$ is given by

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

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

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

Lemma 2.1 ([24]) *Let $\{x_n\}$ be a positive sequence, $\{q_n\}$ be a sequence of real numbers, and $\{\zeta_n\}$ be a sequence in $(0, 1)$ such that $\sum_{n=1}^{\infty} \zeta_n = \infty$. Suppose that*

$$x_{n+1} \leq (1 - \zeta_n) x_n + \zeta_n q_n, \quad \forall n \geq 1.$$

If $\limsup_{k\to\infty} q_{n_k} \leq 0$ for any subsequence $\{x_{n_k}\}$ of $\{x_n\}$ satisfying $\liminf_{k\to\infty} \left(x_{n_k+1} - x_{n_k}\right) \geq 0$, then $\lim_{n\to\infty} x_n = 0$.

3 Main results

In this section, we introduce four new modified inertial projection and contraction methods with Armijo line search step sizes to solve pseudo-monotone variational inequalities in real Hilbert spaces. We first assume that the following conditions hold in order to analyze the convergence of the proposed algorithms.

(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, $\mathrm{VI}(C, A) \neq \emptyset$.
(C3) The mapping $A : \mathcal{H} \to \mathcal{H}$ is pseudo-monotone, uniformly continuous on $\mathcal{H}$, and sequentially weakly continuous on C.

3.1 The first Mann-type projection algorithm

Based on the inertial method, the modified subgradient extragradient method (MSEGM), and the Mann-type method, our first scheme is stated in Algorithm 3.1.

Algorithm 3.1

Initialization: Take $\delta > 0$, $\gamma > 0$, $\ell \in (0,1)$, $\phi \in (0,1)$, $\tau \in (0, 2/\phi)$, and $\beta \in (\tau/2, 1/\phi)$. Let $x_0, x_1 \in \mathcal{H}$.
Iterative Steps: Given the iterates x_{n-1} and x_n ($n \geq 1$). Calculate the next iteration point x_{n+1} as follows:
Step 1. Compute $q_n = x_n + \delta_n(x_n - x_{n-1})$, where

$$\delta_n = \begin{cases} \min\left\{\dfrac{\epsilon_n}{\|x_n - x_{n-1}\|}, \delta\right\}, & \text{if } x_n \neq x_{n-1}; \\ \delta, & \text{otherwise.} \end{cases} \tag{3.1}$$

Step 2. Compute $s_n = P_C(q_n - \beta\lambda_n A q_n)$, where the step size λ_n is chosen to be the largest $\lambda \in \left\{\gamma, \gamma\ell, \gamma\ell^2, \ldots\right\}$ satisfying

$$\lambda\|Aq_n - As_n\| \leq \phi\|q_n - s_n\|. \tag{3.2}$$

If $q_n = s_n$, then stop and s_n is a solution of (VIP). Otherwise, go to **Step 3**.
Step 3. Compute $t_n = P_{H_n}(q_n - \tau\lambda_n\theta_n A s_n)$, where

$$H_n := \{x \in \mathcal{H} \mid \langle q_n - \beta\lambda_n A q_n - s_n, x - s_n\rangle \leq 0\},$$

and

$$\theta_n = (1 - \beta\phi)\frac{\|q_n - s_n\|^2}{\|d_n\|^2}, \quad d_n = q_n - s_n - \beta\lambda_n(Aq_n - As_n). \tag{3.3}$$

Step 4. Compute $x_{n+1} = (1 - \zeta_n - \sigma_n)q_n + \sigma_n t_n$.
Set $n := n + 1$ and go to **Step 1**.

Remark 3.1 Notice that Algorithm 3.1 is different from the extragradient-type methods in the literature (e.g., [12–14]) when calculating the values of s_n and t_n. Specifically, we use two different step sizes in the calculation of s_n and t_n. Our algorithm has a faster convergence speed and accuracy when choosing a suitable value of the parameter β (see Sect. 4). In addition, the range of τ in our algorithm is $\tau \in (0, 2/\phi)$ with $\phi \in (0,1)$, while it is $\tau \in (0,2)$ in [12–14].

We can obtain the following conclusions of Lemmas 3.1 and 3.2 by a simple modification of Lemmas 3.1 and 3.3 in [25], respectively. To avoid repetitive expressions, we omit their proofs here.

Lemma 3.1 ([25]) *Suppose that Conditions (C1)–(C3) hold. The Armijo-like criteria* (3.2) *is well defined.*

Lemma 3.2 ([25]) *Suppose that Conditions (C1)–(C3) hold. Let $\{q_n\}$ and $\{s_n\}$ be two sequences created by Algorithm 3.1. If there exists a subsequence $\{q_{n_k}\}$ of $\{q_n\}$ such that $\{q_{n_k}\}$ converges weakly to $z \in \mathcal{H}$ and $\lim_{k\to\infty}\|q_{n_k} - s_{n_k}\| = 0$, then $z \in$* $\mathrm{VI}(C, A)$.

The following lemma is very helpful in analyzing the convergence of Algorithm 3.1.

Lemma 3.3 *Suppose that Conditions (C1)–(C3) hold. Let $\{q_n\}$, $\{s_n\}$, and $\{t_n\}$ be three sequences generated by Algorithm 3.1. Then, for all $x^* \in \mathrm{VI}(C, A)$,*

$$\|t_n - x^*\|^2 \le \|q_n - x^*\|^2 - \|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2 - \frac{\tau}{\beta^2}(2\beta - \tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2.$$

Proof From $x^* \in \mathrm{VI}(C, A) \subset C \subset H_n$ and (2.2), we have

$$\begin{aligned}
2\|t_n - x^*\|^2 &= 2\|P_{H_n}(q_n - \tau\lambda_n\theta_n As_n) - P_{H_n}(x^*)\|^2\\
&\le 2\langle t_n - x^*, q_n - \tau\lambda_n\theta_n As_n - x^*\rangle\\
&= \|t_n - x^*\|^2 + \|q_n - \tau\lambda_n\theta_n As_n - x^*\|^2 - \|t_n - q_n + \tau\lambda_n\theta_n As_n\|^2\\
&= \|t_n - x^*\|^2 + \|q_n - x^*\|^2 + \tau^2\lambda_n^2\theta_n^2\|As_n\|^2 - 2\langle q_n - x^*, \tau\lambda_n\theta_n As_n\rangle\\
&\quad - \|t_n - q_n\|^2 - \tau^2\lambda_n^2\theta_n^2\|As_n\|^2 - 2\langle t_n - q_n, \tau\lambda_n\theta_n As_n\rangle\\
&= \|t_n - x^*\|^2 + \|q_n - x^*\|^2 - \|t_n - q_n\|^2 - 2\langle t_n - x^*, \tau\lambda_n\theta_n As_n\rangle,
\end{aligned}$$

which yields that

$$\|t_n - x^*\|^2 \le \|q_n - x^*\|^2 - \|t_n - q_n\|^2 - 2\tau\lambda_n\theta_n\langle t_n - x^*, As_n\rangle. \tag{3.1}$$

Combining $s_n \in C$ and $x^* \in \mathrm{VI}(C, A)$, it follows that $\langle Ax^*, s_n - x^*\rangle \ge 0$. We obtain $\langle As_n, s_n - x^*\rangle \ge 0$ according to the pseudo-monotonicity of A. This implies that $\langle As_n, t_n - x^*\rangle \ge \langle As_n, t_n - s_n\rangle$. Hence,

$$-2\tau\lambda_n\theta_n\langle As_n, t_n - x^*\rangle \le -2\tau\lambda_n\theta_n\langle As_n, t_n - s_n\rangle. \tag{3.2}$$

By using $t_n \in H_n$, one has $\langle q_n - \beta\lambda_n Aq_n - s_n, t_n - s_n\rangle \le 0$. This means that

$$\langle \underbrace{q_n - s_n - \beta\lambda_n(Aq_n - As_n)}_{d_n}, t_n - s_n\rangle \le \beta\lambda_n\langle As_n, t_n - s_n\rangle. \tag{3.3}$$

Combining (3.2), (3.3), and the definition of d_n, we obtain

$$\begin{aligned}
-2\tau\lambda_n\theta_n\langle As_n, t_n - x^*\rangle &\le -2\frac{\tau}{\beta}\theta_n\langle d_n, t_n - s_n\rangle\\
&= -2\frac{\tau}{\beta}\theta_n\langle d_n, q_n - s_n\rangle + 2\frac{\tau}{\beta}\theta_n\langle d_n, q_n - t_n\rangle.
\end{aligned} \tag{3.4}$$

From the definitions of θ_n, d_n, and (3.2), we obtain

$$\begin{aligned}
\langle d_n, q_n - s_n\rangle &\ge \|q_n - s_n\|^2 - \beta\lambda_n\|Aq_n - As_n\|\|q_n - s_n\|\\
&\ge \|q_n - s_n\|^2 - \beta\phi\|q_n - s_n\|^2\\
&= (1-\beta\phi)\|q_n - s_n\|^2 = \theta_n\|d_n\|^2,
\end{aligned}$$

which indicates that

$$-2\frac{\tau}{\beta}\theta_n\langle d_n, q_n - s_n\rangle \le -2\frac{\tau}{\beta}\theta_n^2\|d_n\|^2. \tag{3.5}$$

According to the basic inequality $2ab = a^2 + b^2 - (a-b)^2$, we also have

$$2\frac{\tau}{\beta}\theta_n\langle d_n, q_n - t_n\rangle = \|q_n - t_n\|^2 + \frac{\tau^2}{\beta^2}\theta_n^2\|d_n\|^2 - \|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2. \tag{3.6}$$

It follows from the definition of d_n and (3.2) that

$$\begin{aligned}\|d_n\| &\le \|q_n - s_n\| + \beta\lambda_n\|Aq_n - As_n\|\\ &\le \|q_n - s_n\| + \beta\phi\|q_n - s_n\|\\ &= (1+\beta\phi)\|q_n - s_n\|,\end{aligned}$$

which combining the definition of θ_n yields that

$$\theta_n^2\|d_n\|^2 = (1-\beta\phi)^2\frac{\|q_n - s_n\|^4}{\|d_n\|^2} \ge \frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2. \tag{3.7}$$

Combining (3.1), (3.4), (3.5), (3.6), (3.7), and $\beta \in (\tau/2, 1/\phi)$, we conclude that

$$\begin{aligned}\|t_n - x^*\|^2 &\le \|q_n - x^*\|^2 - \|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2\\ &-\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2.\end{aligned}$$

This completes the proof. □

Theorem 3.1 *Suppose that Conditions (C1)–(C4) hold. Then the sequence* $\{x_n\}$ *generated by Algorithm 3.1 converges strongly to* $x^* \in \mathrm{VI}(C, A)$*, where* $\|x^*\| = \min\{\|z\| : z \in \mathrm{VI}(C, A)\}$.

(C4) *Let* $\{\epsilon_n\}$ *be a positive sequence such that* $\lim_{n\to\infty}\frac{\epsilon_n}{\zeta_n} = 0$*, where* $\{\zeta_n\} \subset (0,1)$ *satisfies* $\lim_{n\to\infty}\zeta_n = 0$ *and* $\sum_{n=1}^{\infty}\zeta_n = \infty$*. Let* $\{\sigma_n\} \subset (a, b) \subset (0, 1-\zeta_n)$ *for some* $a > 0, b > 0$.

Proof First, we show that the sequence $\{x_n\}$ is bounded. Indeed, thanks to Lemma 3.3, $\tau \in (0, 2/\phi)$, and $\beta \in (\tau/2, 1/\phi)$, one has

$$\|t_n - x^*\| \le \|q_n - x^*\|, \quad \forall n \ge 1. \tag{3.8}$$

By the definition of q_n, one has

$$\|q_n - x^*\| \le \|x_n - x^*\| + \zeta_n \cdot \frac{\delta_n}{\zeta_n}\|x_n - x_{n-1}\|. \tag{3.9}$$

From (3.1), we obtain $\delta_n \|x_n - x_{n-1}\| \le \epsilon_n, \forall n \ge 1$, which together with $\lim_{n\to\infty} \frac{\epsilon_n}{\zeta_n} = 0$ implies that

$$\lim_{n\to\infty} \frac{\delta_n}{\zeta_n} \|x_n - x_{n-1}\| \le \lim_{n\to\infty} \frac{\epsilon_n}{\zeta_n} = 0. \tag{3.10}$$

According to (3.10), we obtain that $\frac{\delta_n}{\zeta_n} \|x_n - x_{n-1}\| \to 0$ as $n \to \infty$. Therefore, there exists a constant $W_1 > 0$ such that

$$\frac{\delta_n}{\zeta_n} \|x_n - x_{n-1}\| \le W_1, \quad \forall n \ge 1,$$

which combining with (3.8) and (3.9) yields that

$$\|t_n - x^*\| \le \|q_n - x^*\| \le \|x_n - x^*\| + \zeta_n W_1. \tag{3.11}$$

From the definition of x_{n+1}, one sees that

$$\|x_{n+1} - x^*\| \le \|(1 - \zeta_n - \sigma_n)(q_n - x^*) + \sigma_n(t_n - x^*)\| + \zeta_n \|x^*\|. \tag{3.12}$$

It follows from (3.8) that

$$\begin{aligned}
&\|(1 - \zeta_n - \sigma_n)(q_n - x^*) + \sigma_n(t_n - x^*)\|^2 \\
&\quad \le (1 - \zeta_n - \sigma_n)^2 \|q_n - x^*\|^2 + \sigma_n^2 \|t_n - x^*\|^2 \\
&\qquad + 2(1 - \zeta_n - \sigma_n)\sigma_n \|t_n - x^*\| \|q_n - x^*\| \\
&\quad \le (1 - \zeta_n - \sigma_n)^2 \|q_n - x^*\|^2 + \sigma_n^2 \|q_n - x^*\|^2 \\
&\qquad + 2(1 - \zeta_n - \sigma_n)\sigma_n \|q_n - x^*\|^2 \\
&\quad = (1 - \zeta_n)^2 \|q_n - x^*\|^2,
\end{aligned}$$

which implies that

$$\|(1 - \zeta_n - \sigma_n)(q_n - x^*) + \sigma_n(t_n - x^*)\| \le (1 - \zeta_n)\|q_n - x^*\|. \tag{3.13}$$

From (3.11), (3.12), and (3.13), we conclude that

$$\begin{aligned}
\|x_{n+1} - x^*\| &\le (1 - \zeta_n)\|q_n - x^*\| + \zeta_n \|x^*\| \\
&\le (1 - \zeta_n)\|x_n - x^*\| + \zeta_n (\|x^*\| + W_1) \\
&\le \max\{\|x_n - x^*\|, \|x^*\| + W_1\} \\
&\le \cdots \le \max\{\|x_1 - x^*\|, \|x^*\| + W_1\}.
\end{aligned}$$

That is, $\{x_n\}$ is bounded, and so are $\{q_n\}$, $\{s_n\}$, and $\{t_n\}$.

From (3.11), one sees that

$$\begin{aligned}\|q_n - x^*\|^2 &\le (\|x_n - x^*\| + \zeta_n W_1)^2 \\ &= \|x_n - x^*\|^2 + \zeta_n(2W_1\|x_n - x^*\| + \zeta_n W_1^2) \\ &\le \|x_n - x^*\|^2 + \zeta_n W_2\end{aligned} \tag{3.14}$$

for some $W_2 > 0$. By the definition of x_{n+1} and Condition (C4), we have

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &= \|(1-\zeta_n-\sigma_n)(q_n - x^*) + \sigma_n(t_n - x^*) + \zeta_n(-x^*)\|^2 \\ &= (1-\zeta_n-\sigma_n)\|q_n - x^*\|^2 + \sigma_n\|t_n - x^*\|^2 + \zeta_n\|x^*\|^2 \\ &\quad - \sigma_n(1-\zeta_n-\sigma_n)\|q_n - t_n\|^2 - \zeta_n\sigma_n\|t_n\|^2 \\ &\quad - \zeta_n(1-\zeta_n-\sigma_n)\|q_n\|^2 \\ &\le (1-\zeta_n-\sigma_n)\|q_n - x^*\|^2 + \sigma_n\|t_n - x^*\|^2 + \zeta_n\|x^*\|^2.\end{aligned} \tag{3.15}$$

From Lemma 3.3, (3.14), and (3.15), we obtain

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\le (1-\zeta_n-\sigma_n)\|q_n - x^*\|^2 + \sigma_n\|q_n - x^*\|^2 - \sigma_n\|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2 \\ &\quad - \sigma_n\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2 + \zeta_n\|x^*\|^2 \\ &\le \|x_n - x^*\|^2 - \sigma_n\|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2 + \zeta_n(\|x^*\|^2 + W_2) \\ &\quad - \sigma_n\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2.\end{aligned}$$

Thus we have

$$\begin{aligned}&\sigma_n\|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2 + \sigma_n\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n - s_n\|^2 \\ &\le \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + \zeta_n(\|x^*\|^2 + W_2).\end{aligned} \tag{Eq1}$$

From the definition of q_n, one sees that

$$\begin{aligned}\|q_n - x^*\|^2 &\le \|x_n - x^*\|^2 + 2\delta_n\|x_n - x^*\|\|x_n - x_{n-1}\| + \delta_n^2\|x_n - x_{n-1}\|^2 \\ &\le \|x_n - x^*\|^2 + 3W\delta_n\|x_n - x_{n-1}\|,\end{aligned} \tag{3.16}$$

where $W := \sup_{n\in\mathbb{N}}\{\|x_n - x^*\|, \delta\|x_n - x_{n-1}\|\} > 0$. Set $b_n = (1-\sigma_n)q_n + \sigma_n t_n$. It follows from (3.8) that

$$\|b_n - x^*\| \le (1-\sigma_n)\|q_n - x^*\| + \sigma_n\|t_n - x^*\| \le \|q_n - x^*\|. \tag{3.17}$$

From (3.16) and (3.17), we have

$$
\begin{aligned}
&\|x_{n+1} - x^*\|^2 \qquad\qquad \text{(Eq2)}\\
&\quad = \|(1-\zeta_n)(b_n - x^*) - \zeta_n(q_n - b_n) - \zeta_n x^*\|^2\\
&\quad \le (1-\zeta_n)^2\|b_n - x^*\|^2 - 2\zeta_n\langle q_n - b_n + x^*, x_{n+1} - x^*\rangle\\
&\quad = (1-\zeta_n)^2\|b_n - x^*\|^2 + 2\zeta_n\langle q_n - b_n, x^* - x_{n+1}\rangle + 2\zeta_n\langle x^*, x^* - x_{n+1}\rangle\\
&\quad \le (1-\zeta_n)\|b_n - x^*\|^2 + 2\zeta_n\|q_n - b_n\|\|x_{n+1} - x^*\| + 2\zeta_n\langle x^*, x^* - x_{n+1}\rangle\\
&\quad \le (1-\zeta_n)\|x_n - x^*\|^2 + \zeta_n\Big[2\sigma_n\|q_n - t_n\|\|x_{n+1} - x^*\|\\
&\qquad + 2\langle x^*, x^* - x_{n+1}\rangle + \frac{3W\delta_n}{\zeta_n}\|x_n - x_{n-1}\|\Big].
\end{aligned}
$$

Finally, we show that $\{\|x_n - x^*\|\}$ converges to zero. Throughout this paper, we always assume that $\{\|x_{n_k} - x^*\|\}$ is a subsequence of $\{\|x_n - x^*\|\}$ such that $\liminf_{k\to\infty}(\|x_{n_k+1} - x^*\| - \|x_{n_k} - x^*\|) \ge 0$. Then,

$$
\begin{aligned}
&\liminf_{k\to\infty}\left(\|x_{n_k+1} - x^*\|^2 - \|x_{n_k} - x^*\|^2\right)\\
&\quad = \liminf_{k\to\infty}\left[(\|x_{n_k+1} - x^*\| - \|x_{n_k} - x^*\|)(\|x_{n_k+1} - x^*\| + \|x_{n_k} - x^*\|)\right] \ge 0.
\end{aligned}
$$

Combining (Eq1), Condition (C4), $\tau \in (0, 2/\phi)$, and $\beta \in (\tau/2, 1/\phi)$, we have

$$
\begin{aligned}
&\limsup_{k\to\infty}\left\{\sigma_{n_k}\frac{\tau}{\beta^2}(2\beta - \tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_{n_k} - s_{n_k}\|^2 + \sigma_{n_k}\|q_{n_k} - t_{n_k} - \frac{\tau}{\beta}\theta_{n_k}d_{n_k}\|^2\right\}\\
&\quad \le \limsup_{k\to\infty}\left[\|x_{n_k} - x^*\|^2 - \|x_{n_k+1} - x^*\|^2\right] + \limsup_{k\to\infty}\zeta_{n_k}(\|x^*\|^2 + W_2)\\
&\quad = -\liminf_{k\to\infty}\left[\|x_{n_k+1} - x^*\|^2 - \|x_{n_k} - x^*\|^2\right] \le 0,
\end{aligned}
$$

which yields that

$$
\lim_{k\to\infty}\|s_{n_k} - q_{n_k}\| = 0 \quad \text{and} \quad \lim_{k\to\infty}\|q_{n_k} - t_{n_k} - \frac{\tau}{\beta}\theta_{n_k}d_{n_k}\| = 0.
$$

From $\|d_{n_k}\| \ge (1-\beta\phi)\|q_{n_k} - s_{n_k}\|$ and the definition of θ_{n_k}, we obtain

$$
\begin{aligned}
\|q_{n_k} - t_{n_k}\| &\le \|q_{n_k} - t_{n_k} - \frac{\tau}{\beta}\theta_{n_k}d_{n_k}\| + \frac{\tau}{\beta}\theta_{n_k}\|d_{n_k}\|\\
&= \|q_{n_k} - t_{n_k} - \frac{\tau}{\beta}\theta_{n_k}d_{n_k}\| + \frac{\tau}{\beta}(1-\beta\phi)\frac{\|q_{n_k} - s_{n_k}\|^2}{\|d_{n_k}\|}\\
&\le \|q_{n_k} - t_{n_k} - \frac{\tau}{\beta}\theta_{n_k}d_{n_k}\| + \frac{\tau}{\beta}\|q_{n_k} - s_{n_k}\|.
\end{aligned}
$$

Hence, we have that $\lim_{k\to\infty}\|t_{n_k}-q_{n_k}\|=0$. This together with the boundedness of $\{x_n\}$ means that

$$\lim_{k\to\infty}\sigma_{n_k}\|q_{n_k}-t_{n_k}\|\|x_{n_k+1}-x^*\|=0. \tag{3.18}$$

Combining (3.10) and Condition (C4), we have

$$\|x_{n_k+1}-q_{n_k}\|=\zeta_{n_k}\|q_{n_k}\|+\sigma_{n_k}\|t_{n_k}-q_{n_k}\|\to 0 \text{ as } n\to\infty,$$

and

$$\|x_{n_k}-q_{n_k}\|=\zeta_{n_k}\cdot\frac{\delta_{n_k}}{\zeta_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0 \text{ as } n\to\infty.$$

From the above facts, we conclude that

$$\|x_{n_k+1}-x_{n_k}\|\leq\|x_{n_k+1}-q_{n_k}\|+\|q_{n_k}-x_{n_k}\|\to 0 \text{ as } n\to\infty. \tag{3.19}$$

Since the sequence $\{x_{n_k}\}$ is bounded, there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$ such that $x_{n_{k_j}}\rightharpoonup z$ when $j\to\infty$. Furthermore,

$$\limsup_{k\to\infty}\langle x^*,x^*-x_{n_k}\rangle=\lim_{j\to\infty}\langle x^*,x^*-x_{n_{k_j}}\rangle=\langle x^*,x^*-z\rangle. \tag{3.20}$$

We obtain that $q_{n_k}\rightharpoonup z$ since $\|x_{n_k}-q_{n_k}\|\to 0$. This together with $\lim_{k\to\infty}\|q_{n_k}-s_{n_k}\|=0$ and Lemma 3.2 yields that $z\in \mathrm{VI}(C,A)$. By using (2.1), (3.20), and the definition of x^*, we obtain

$$\limsup_{k\to\infty}\langle x^*,x^*-x_{n_k}\rangle=\langle x^*,x^*-z\rangle\leq 0. \tag{3.21}$$

From (3.19) and (3.21), we have

$$\limsup_{k\to\infty}\langle x^*,x^*-x_{n_k+1}\rangle\leq\limsup_{k\to\infty}\langle x^*,x^*-x_{n_k}\rangle\leq 0. \tag{3.22}$$

Combining (3.10), (3.18), (3.22), (Eq2), and Lemma 2.1, we conclude that $x_n\to x^*$ as $n\to\infty$. The proof is completed. □

Remark 3.2 Notice that the convergence of the proposed Algorithm 3.1 is proved in the case where operator A is uniformly continuous rather than Lipschitz continuous. This cannot be achieved by many methods in the literature that use self-adaptive step sizes (see, e.g., [15,18]).

3.2 The first viscosity-type projection algorithm

In this subsection, we present a viscosity-type inertial modified subgradient extragradient method for solving (VIP). The details of this scheme are described in Algorithm 3.2.

Algorithm 3.2

Initialization: Take $\delta > 0$, $\gamma > 0$, $\ell \in (0,1)$, $\phi \in (0,1)$, $\tau \in (0, 2/\phi)$, and $\beta \in (\tau/2, 1/\phi)$. Let $x_0, x_1 \in \mathcal{H}$.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} q_n = x_n + \delta_n \left(x_n - x_{n-1}\right), \\ s_n = P_C\left(q_n - \beta\lambda_n A q_n\right), \\ t_n = P_{H_n}(q_n - \tau\lambda_n\theta_n A s_n), \\ H_n := \{x \in \mathcal{H} \mid \langle q_n - \beta\lambda_n A q_n - s_n, x - s_n\rangle \leq 0\}, \\ x_{n+1} = \zeta_n f(t_n) + (1-\zeta_n) t_n, \end{cases}$$

where $\{\delta_n\}$, $\{\lambda_n\}$, and $\{\theta_n\}$ are defined in (3.1), (3.2), and (3.3), respectively.

Theorem 3.2 *Suppose that Conditions (C1)–(C3) and (C5) hold. Then the sequence* $\{x_n\}$ *created by Algorithm 3.2 converges strongly to* $x^* \in \mathrm{VI}(C, A)$*, where* $x^* = P_{\mathrm{VI}(C,A)} \circ f(x^*)$.

(C5) *Let* $f : \mathcal{H} \to \mathcal{H}$ *be a* κ*-contraction mapping with* $\kappa \in [0, 1)$ *and let* $\{\epsilon_n\}$ *be a positive sequence such that* $\lim_{n\to\infty} \frac{\epsilon_n}{\zeta_n} = 0$*, where* $\{\zeta_n\} \subset (0, 1)$ *satisfies* $\lim_{n\to\infty} \zeta_n = 0$ *and* $\sum_{n=1}^{\infty} \zeta_n = \infty$.

Proof By the definition of x_{n+1} and (3.11), we have

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \zeta_n \|f(t_n) - f(x^*)\| + \zeta_n \|f(x^*) - x^*\| + (1-\zeta_n)\|t_n - x^*\| \\ &\leq (1 - (1-\kappa)\zeta_n)\|x_n - x^*\| + (1-\kappa)\zeta_n \frac{W_1 + \|f(x^*) - x^*\|}{1-\kappa} \\ &\leq \max\left\{\|x_1 - x^*\|, \frac{W_1 + \|f(x^*) - x^*\|}{1-\kappa}\right\}. \end{aligned}$$

This means that $\{x_n\}$ is bounded. Hence, $\{q_n\}$, $\{s_n\}$, $\{t_n\}$, and $\{f(t_n)\}$ are also bounded. From Lemma 3.3 and (3.14), we obtain

$$\begin{aligned} \|x_{n+1} - x^*\|^2 &\leq \zeta_n (\|t_n - x^*\| + \|f(x^*) - x^*\|)^2 + (1-\zeta_n)\|t_n - x^*\|^2 \\ &= \zeta_n \|t_n - x^*\|^2 + (1-\zeta_n)\|t_n - x^*\|^2 \\ &\quad + \zeta_n \left(2\|t_n - x^*\|\|f(x^*) - x^*\| + \|f(x^*) - x^*\|^2\right) \\ &\leq \|t_n - x^*\|^2 + \zeta_n W_3 \\ &\leq \|x_n - x^*\|^2 - \|q_n - t_n - \frac{\tau}{\beta}\theta_n d_n\|^2 \end{aligned}$$

$$-\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n-s_n\|^2+\zeta_n W_4,$$

where $W_3 := \max\{2\|t_n-x^*\|\|f(x^*)-x^*\|+\|f(x^*)-x^*\|^2\}$ and $W_4 := W_2+W_3$. Thus,

$$\begin{aligned}&\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_n-s_n\|^2+\|q_n-t_n-\frac{\tau}{\beta}\theta_n d_n\|^2\\&\quad\le\|x_n-x^*\|^2-\|x_{n+1}-x^*\|^2+\zeta_n W_4.\end{aligned}\tag{Eq3}$$

By using (3.8) and (3.16), we have

$$\begin{aligned}&\|x_{n+1}-x^*\|^2\\&\quad=\|\zeta_n(f(t_n)-f(x^*))+(1-\zeta_n)(t_n-x^*)+\zeta_n(f(x^*)-x^*)\|^2\\&\quad\le\zeta_n\kappa\|t_n-x^*\|^2+(1-\zeta_n)\|t_n-x^*\|^2+2\zeta_n\langle f(x^*)-x^*,x_{n+1}-x^*\rangle\\&\quad\le(1-(1-\kappa)\zeta_n)\|x_n-x^*\|^2+(1-\kappa)\zeta_n\cdot\Big[\frac{3W}{1-\kappa}\cdot\frac{\delta_n}{\zeta_n}\|x_n-x_{n-1}\|\\&\qquad+\frac{2}{1-\kappa}\langle f(x^*)-x^*,x_{n+1}-x^*\rangle\Big].\end{aligned}\tag{Eq4}$$

Finally, we show that $\{\|x_n-x^*\|\}$ converges to zero. From (Eq3), Condition (C5), $\tau\in(0,2/\phi)$, and $\beta\in(\tau/2,1/\phi)$, we have

$$\begin{aligned}&\limsup_{k\to\infty}\left\{\frac{\tau}{\beta^2}(2\beta-\tau)\frac{(1-\beta\phi)^2}{(1+\beta\phi)^2}\|q_{n_k}-s_{n_k}\|^2+\|q_{n_k}-t_{n_k}-\frac{\tau}{\beta}\theta_{n_k}d_{n_k}\|^2\right\}\\&\quad\le\limsup_{k\to\infty}\left[\|x_{n_k}-x^*\|^2-\|x_{n_k+1}-x^*\|^2+\zeta_{n_k}W_4\right]\le 0,\end{aligned}$$

which indicates that

$$\lim_{k\to\infty}\|s_{n_k}-q_{n_k}\|=0\ \text{ and }\ \lim_{k\to\infty}\|q_{n_k}-t_{n_k}-\frac{\tau}{\beta}\theta_{n_k}d_{n_k}\|=0.$$

As stated in Theorem 3.1, it is easy to see that $\lim_{k\to\infty}\|t_{n_k}-q_{n_k}\|=0$. From (3.10) and Condition (C5), we obtain

$$\|x_{n_k+1}-t_{n_k}\|=\zeta_{n_k}\|t_{n_k}-f(x_{n_k})\|\to 0 \text{ as } n\to\infty,$$

and

$$\|x_{n_k}-q_{n_k}\|=\zeta_{n_k}\cdot\frac{\delta_{n_k}}{\zeta_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0 \text{ as } n\to\infty.$$

Therefore,

$$\|x_{n_k+1}-x_{n_k}\|\le\|x_{n_k+1}-t_{n_k}\|+\|t_{n_k}-q_{n_k}\|+\|q_{n_k}-x_{n_k}\|\to 0 \text{ as } n\to\infty.$$

$$(3.23)$$

Since $\{x_{n_k}\}$ is bounded, there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$ such that $x_{n_{k_j}} \rightharpoonup z$ when $j \to \infty$. Moreover,

$$\begin{aligned}\limsup_{k\to\infty}\langle f(x^*) - x^*, x_{n_k} - x^*\rangle &= \lim_{j\to\infty}\langle f(x^*) - x^*, x_{n_{k_j}} - x^*\rangle \\ &= \langle f(x^*) - x^*, z - x^*\rangle.\end{aligned} \tag{3.24}$$

We have that $q_{n_k} \rightharpoonup z$ as $\|x_{n_k} - q_{n_k}\| \to 0$, which together with $\lim_{k\to\infty}\|q_{n_k} - s_{n_k}\| = 0$ and Lemma 3.2 implies that $z \in \mathrm{VI}(C, A)$. From (2.1), (3.24), and the definition of x^*, we have

$$\limsup_{k\to\infty}\langle f(x^*) - x^*, x_{n_k} - x^*\rangle = \langle f(x^*) - x^*, z - x^*\rangle \le 0. \tag{3.25}$$

By using (3.23) and (3.25), we obtain

$$\limsup_{k\to\infty}\langle f(x^*) - x^*, x_{n_k+1} - x^*\rangle \le \limsup_{k\to\infty}\langle f(x^*) - x^*, x_{n_k} - x^*\rangle \le 0. \tag{3.26}$$

From (3.10), (3.26), (Eq4), and Lemma 2.1, we conclude that $x_n \to x^*$ as $n \to \infty$. The proof of the Theorem 3.2 is now complete. □

3.3 The second Mann-type projection algorithm

Inspired by the algorithms of Gibali et al. [14], in this subsection, we introduce a new modified inertial projection and contraction algorithm to solve pseudo-monotone variational inequalities in infinite-dimensional Hilbert spaces. Now, we present the proposed scheme in Algorithm 3.3.

Algorithm 3.3

Initialization: Take $\delta > 0$, $\gamma > 0$, $\ell \in (0, 1)$, $\phi \in (0, 1)$, $\tau \in (0, 2)$, and $\beta \in (0, 1/\phi)$. Let $x_0, x_1 \in \mathcal{H}$.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} q_n = x_n + \delta_n\left(x_n - x_{n-1}\right), \\ s_n = P_C\left(q_n - \beta\lambda_n A q_n\right), \\ t_n = q_n - \tau\theta_n d_n, \\ x_{n+1} = (1 - \zeta_n - \sigma_n)q_n + \sigma_n t_n, \end{cases}$$

where $\{\delta_n\}$, $\{\lambda_n\}$, and $\{\theta_n\}$ are defined in (3.1), (3.2), and (3.3), respectively.

Remark 3.3 It is worth noting that our Algorithm 3.3 is different from the algorithms presented in [14] in the calculation of s_n, and our numerical experiments in the next section will show that our algorithm has a higher accuracy and faster convergence speed than the algorithms in [14] when choosing the appropriate value of β.

The following lemma plays a crucial role in studying the convergence of Algorithm 3.3.

Lemma 3.4 *Suppose that Conditions (C1)–(C3) hold. Let* $\{q_n\}$, $\{s_n\}$, *and* $\{t_n\}$ *be three sequences generated by Algorithm 3.3. Then,*

$$\|t_n - x^*\|^2 \leq \|q_n - x^*\|^2 - \frac{2-\tau}{\tau}\|q_n - t_n\|^2, \quad \forall x^* \in \mathrm{VI}(C, A),$$

and

$$\|q_n - s_n\|^2 \leq \left[\frac{1+\beta\phi}{(1-\beta\phi)\tau}\right]^2 \|q_n - t_n\|^2.$$

Proof From the definition of t_n, one sees that

$$\begin{aligned} \|t_n - x^*\|^2 &= \|q_n - \tau\theta_n d_n - x^*\|^2 \\ &= \|q_n - x^*\|^2 - 2\tau\theta_n\langle q_n - x^*, d_n\rangle + \tau^2\theta_n^2\|d_n\|^2. \end{aligned} \tag{3.27}$$

From (3.2) and (3.3), we have

$$\begin{aligned} \langle q_n - x^*, d_n\rangle &= \langle q_n - s_n, q_n - s_n - \beta\lambda_n(Aq_n - As_n)\rangle + \langle s_n - x^*, d_n\rangle \\ &\geq \|q_n - s_n\|^2 - \beta\lambda_n\|q_n - s_n\|\|Aq_n - As_n\| + \langle s_n - x^*, d_n\rangle \\ &\geq (1-\beta\phi)\|q_n - s_n\|^2 + \langle s_n - x^*, q_n - s_n - \beta\lambda_n(Aq_n - As_n)\rangle. \end{aligned} \tag{3.28}$$

Combining $s_n = P_C(q_n - \beta\lambda_n Aq_n)$ and (2.1), we obtain

$$\langle q_n - s_n - \beta\lambda_n Aq_n, s_n - x^*\rangle \geq 0. \tag{3.29}$$

Using $x^* \in \mathrm{VI}(C, A)$ and $s_n \in C$, we obtain that $\langle Ax^*, s_n - x^*\rangle \geq 0$. This together with the pseudo-monotonicity of mapping A implies that

$$\langle As_n, s_n - x^*\rangle \geq 0. \tag{3.30}$$

It follows from (3.3) that $(1-\beta\phi)\|q_n - s_n\|^2 = \theta_n\|d_n\|^2$. This combining with (3.28), (3.29), and (3.30) yields that

$$\langle q_n - x^*, d_n\rangle \geq (1-\beta\phi)\|q_n - s_n\|^2 = \theta_n\|d_n\|^2. \tag{3.31}$$

Combining (3.27) and (3.31), we conclude that

$$\begin{aligned} \|t_n - x^*\|^2 &\leq \|q_n - x^*\|^2 - 2\tau\theta_n^2\|d_n\|^2 + \tau^2\theta_n^2\|d_n\|^2 \\ &= \|q_n - x^*\|^2 - \frac{2-\tau}{\tau}\|q_n - t_n\|^2. \end{aligned}$$

From the definition of t_n and (3.3), we obtain

$$\|q_n - s_n\|^2 = \frac{1}{\theta_n(1-\beta\phi)}\|\theta_n d_n\|^2 = \frac{1}{\theta_n(1-\beta\phi)\tau^2}\|q_n - t_n\|^2. \tag{3.32}$$

Since $\|d_n\| \le (1+\beta\phi)\|q_n - s_n\|$, we have

$$\theta_n = (1-\beta\phi)\frac{\|q_n - s_n\|^2}{\|d_n\|^2} \ge \frac{1-\beta\phi}{(1+\beta\phi)^2}. \tag{3.33}$$

It follows from (3.32) and (3.33) that

$$\|q_n - s_n\|^2 \le \left[\frac{1+\beta\phi}{(1-\beta\phi)\tau}\right]^2 \|q_n - t_n\|^2.$$

The proof of the lemma is now complete. □

Theorem 3.3 *Suppose that Conditions (C1)–(C4) hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.3 converges strongly to $x^* \in \mathrm{VI}(C, A)$, where $\|x^*\| = \min\{\|z\| : z \in \mathrm{VI}(C, A)\}$.*

Proof Thanks to Lemma 3.4 and $\tau \in (0, 2)$, we obtain

$$\|t_n - x^*\| \le \|q_n - x^*\|, \quad \forall n \ge 1. \tag{3.34}$$

Using the same facts as stated in Theorem 3.1, we find that $\{x_n\}$, $\{q_n\}$, $\{s_n\}$, and $\{t_n\}$ are bounded. From Lemma 3.4, (3.14), and (3.15), we have

$$\begin{aligned}\|x_{n+1} - x^*\|^2 &\le (1-\zeta_n-\sigma_n)\|q_n - x^*\|^2 + \sigma_n\|q_n - x^*\|^2 \\ &\quad - \sigma_n\frac{2-\tau}{\tau}\|q_n - t_n\|^2 + \zeta_n\|x^*\|^2 \\ &\le \|x_n - x^*\|^2 - \sigma_n\frac{2-\tau}{\tau}\|q_n - t_n\|^2 + \zeta_n(\|x^*\|^2 + W_2).\end{aligned}$$

By a simple conversion, we assert that

$$\sigma_n\frac{2-\tau}{\tau}\|q_n - t_n\|^2 \le \|x_n - x^*\|^2 - \|x_{n+1} - x^*\|^2 + \zeta_n(\|x^*\|^2 + W_2). \tag{Eq5}$$

Moreover, we can obtain (Eq2) by using the same facts as declared in Theorem 3.1. Finally, we prove that $\{\|x_n - x^*\|\}$ converges to zero. From (Eq5) and Condition (C4), we have

$$\begin{aligned}&\limsup_{k\to\infty}\sigma_{n_k}\frac{2-\tau}{\tau}\|q_{n_k} - t_{n_k}\|^2 \\ &\quad\le \limsup_{k\to\infty}\left[\|x_{n_k} - x^*\|^2 - \|x_{n_k+1} - x^*\|^2 + \zeta_{n_k}(\|x^*\|^2 + W_2)\right] \\ &\quad\le 0,\end{aligned}$$

which means that $\lim_{k\to\infty}\|t_{n_k}-q_{n_k}\|=0$. In view of Lemma 3.4, we observe that $\lim_{k\to\infty}\|s_{n_k}-q_{n_k}\|=0$. As asserted in Theorem 3.1, we can obtain the same result as (3.18)–(3.22). Therefore, we have that $x_n\to x^*$ as $n\to\infty$. This completes the proof. □

3.4 The second viscosity-type projection algorithm

In this subsection, we present a viscosity version of Algorithm 3.3. Now, the last iterative scheme stated in this paper is shown in Algorithm 3.4 below.

Algorithm 3.4

Initialization: Take $\delta>0$, $\gamma>0$, $\ell\in(0,1)$, $\phi\in(0,1)$, $\tau\in(0,2)$, and $\beta\in(0,1/\phi)$. Let $x_0,x_1\in\mathcal{H}$.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} q_n=x_n+\delta_n\left(x_n-x_{n-1}\right),\\ s_n=P_C\left(q_n-\beta\lambda_n Aq_n\right),\\ t_n=q_n-\tau\theta_n d_n,\\ x_{n+1}=\zeta_n f\left(t_n\right)+\left(1-\zeta_n\right)t_n, \end{cases}$$

where $\{\delta_n\}$, $\{\lambda_n\}$, and $\{\theta_n\}$ are defined in (3.1), (3.2), and (3.3), respectively.

Theorem 3.4 *Suppose that Conditions (C1)–(C3) and (C5) hold. Then the sequence $\{x_n\}$ created by Algorithm 3.4 converges strongly to $x^*\in\mathrm{VI}(C,A)$, where $x^*=P_{\mathrm{VI}(C,A)}\circ f(x^*)$.*

Proof Using the same arguments as declared in Theorem 3.2, we conclude that $\{x_n\}$, $\{q_n\}$, $\{s_n\}$, $\{t_n\}$ and $\{f(t_n)\}$ are bounded. From Lemma 3.4, and (3.14), we obtain

$$\frac{2-\tau}{\tau}\|q_n-t_n\|^2\le\|x_n-x^*\|^2-\|x_{n+1}-x^*\|^2+\zeta_n W_4, \tag{Eq6}$$

where W_4 is defined in (Eq3). Moreover, we can obtain (Eq4) by using the same facts as stated in Theorem 3.2. Finally, we show that $\{\|x_n-x^*\|\}$ converges to zero. From (Eq6) and Condition (C5), we have

$$\limsup_{k\to\infty}\frac{2-\tau}{\tau}\|q_{n_k}-t_{n_k}\|^2\le 0,$$

which implies that $\lim_{k\to\infty}\|t_{n_k}-q_{n_k}\|=0$. This together with Lemma 3.4 yields that $\lim_{k\to\infty}\|s_{n_k}-q_{n_k}\|=0$. As stated in Theorem 3.2, we can obtain the same facts as (3.23)–(3.26). Therefore, we conclude that $x_n\to x^*$ as $n\to\infty$. The proof is completed. □

Remark 3.4 Our four algorithms have the following advantages over the algorithms presented in [12–19]. More precisely, our contributions in this paper are as follows.

(i) To improve the convergence speed and computational accuracy of the suggested algorithms, we add inertial terms to the proposed schemes and use two different step sizes in the computation of $\{s_n\}$ and $\{t_n\}$. The computational efficiency of our algorithms will be explained and shown in detail in Sect. 4.

(ii) The proposed algorithms can solve a wider range of pseudo-monotone variational inequalities, while the algorithms introduced in [12–14,19] are only applicable to monotone variational inequalities. Moreover, the variational inequality operators involved in our algorithms are non-Lipschitz continuous, which extends many of the results in [12–19] for solving Lipschitz continuous variational inequalities.

(iii) The fixed step size algorithms given in [16,17,19] require the prior information about the Lipschitz constant of the mapping to work, while our algorithms use an Armijo step size criterion that allows them to work adaptively without this information.

(iv) The iterative sequences generated by our algorithms obtain strong convergence in an infinite-dimensional Hilbert space, which improves on the weakly convergent algorithms obtained in [12,15,18,19].

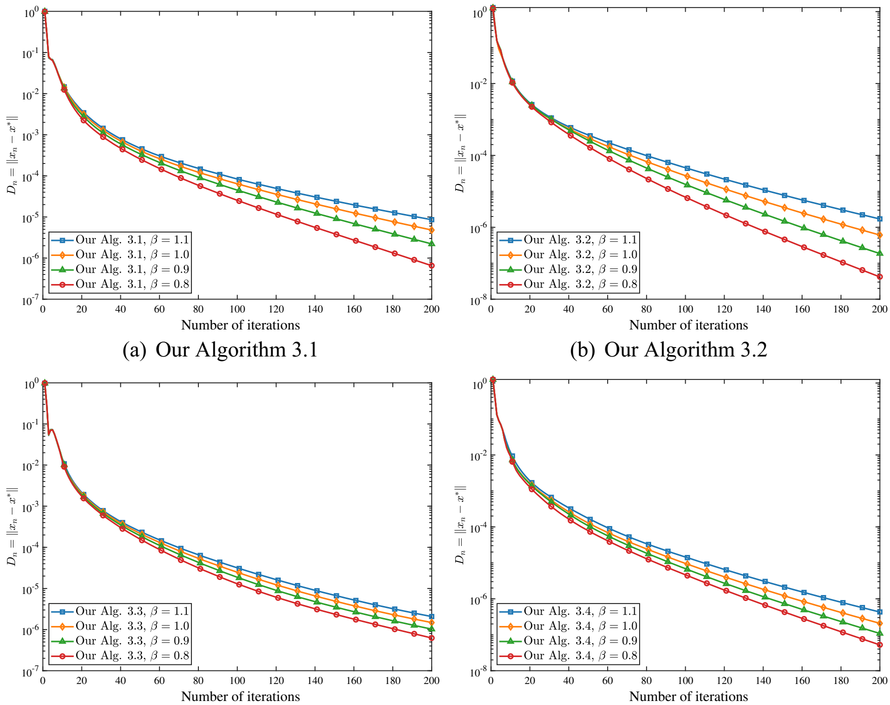
Thus, our four iterative schemes are intelligent, useful, and efficient in solving variational inequality problems.

4 Numerical experiments and applications

In this section, we provide some numerical experiments to demonstrate the advantages of the suggested methods and compare them with some known strongly convergent algorithms in [13,14], which including the algorithms 3.1 and 3.2 presented by Thong and Gibali [13] (shortly, TG Alg. 3.1 and TG Alg. 3.2), and the algorithms 3.1 and 3.2 introduced by Gibali et al. [14] (shortly, GTT Alg. 3.1 and GTT Alg. 3.2). All the programs are implemented in MATLAB 2018a on a personal computer.

4.1 Theoretical examples

Example 4.1 Consider the form of linear operator $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ($m = 20, 50, 100, 200$) as follows: $A(x) = Gx + g$, where $g \in \mathbb{R}^m$ and $G = BB^{\mathsf{T}} + S + E$, 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 and monotone, and its Lipschitz constant is $L = \|G\|$. In this numerical example, all entries of B and S are generated randomly in $[-2, 2]$, and E is generated randomly in $[0, 2]$. Let $g = \mathbf{0}$. Then the solution set is $x^* = \{\mathbf{0}\}$. The parameters of all algorithms are set as follows. Choose $\gamma = 2$, $\ell = 0.5$, $\phi = 0.6$, $\tau = 1.5$, $\zeta_n = 1/(n+1)$, $\sigma_n = 0.5(1-\zeta_n)$ and $f(x) = 0.1x$ for all algorithms. Take $\delta = 0.4$, $\epsilon_n = 100/(n+1)^2$ and $\beta = 0.8$ for the proposed algorithms. We use $D_n = \|x_n - x^*\|$ to measure the nth iteration error of all algorithms. The maximum number of iterations of 200 as a common stopping criterion and the initial values $x_0 = x_1$ are randomly generated by *5rand(m,1)* in MATLAB.



(a) Our Algorithm 3.1 (b) Our Algorithm 3.2

(c) Our Algorithm 3.3 (d) Our Algorithm 3.4

Fig. 1 The behavior of the proposed algorithms with different β in Example 4.1 ($m = 20$)

Figure 1 shows the convergence behavior of the proposed algorithms at $m = 20$ with different parameters β. Table 1 shows the numerical results of all algorithms with four dimensions.

Example 4.2 We consider an example in the Hilbert space $\mathcal{H} = L^2([0, 1])$ associated with inner product

$$\langle x, y\rangle := \int_0^1 x(t)y(t)\mathrm{d}t, \quad \forall x, y \in \mathcal{H}$$

and induced norm

$$\|x\| := \left(\int_0^1 |x(t)|^2 \mathrm{d}t\right)^{1/2}, \ \forall x \in \mathcal{H}.$$

Let the feasible set be the unit ball $C := \{x \in \mathcal{H} : \|x\| \leq 1\}$. Define an operator $A : C \to \mathcal{H}$ by

$$(Ax)(t) = \int_0^1 (x(t) - G(t, s)g(x(s)))\mathrm{d}s + h(t), \quad t \in [0, 1], \ x \in C,$$

Table 1 Numerical results of all algorithms for Example 4.1

Algorithms	$m = 20$		$m = 50$		$m = 100$		$m = 200$	
	D_n	CPU	D_n	CPU	D_n	CPU	D_n	CPU
Our Alg. 3.1	1.96E−06	0.0328	7.46E−05	0.0329	2.80E−04	0.0528	5.11E−04	0.0998
Our Alg. 3.2	2.02E−08	0.0314	8.89E−06	0.0356	4.82E−05	0.0606	1.94E−04	0.1002
Our Alg. 3.3	6.60E−06	0.0304	9.92E−05	0.0308	3.58E−04	0.0493	5.93E−04	0.0892
Our Alg. 3.4	1.54E−07	0.0287	2.37E−05	0.0301	1.04E−04	0.0475	2.36E−04	0.0899
TG Alg. 3.1	4.23E−03	0.0323	1.62E−02	0.0374	3.57E−02	0.0499	4.97E−02	0.1193
TG Alg. 3.2	9.48E−04	0.0317	1.09E−02	0.0508	3.90E−02	0.0475	6.33E−02	0.0990
GTT Alg. 3.1	4.23E−03	0.0277	1.62E−02	0.0329	3.56E−02	0.0506	4.97E−02	0.0924
GTT Alg. 3.2	9.48E−04	0.0267	1.09E−02	0.0316	3.89E−02	0.0532	6.33E−02	0.0994

where

$$G(t,s)=\frac{2ts\mathrm{e}^{t+s}}{e\sqrt{\mathrm{e}^2-1}},\quad g(x)=\cos x,\quad h(t)=\frac{2t\mathrm{e}^t}{e\sqrt{\mathrm{e}^2-1}}.$$

It is known that A is monotone and L-Lipschitz continuous with $L=2$ (see [26]), and $x^*(t)=\{\mathbf{0}\}$ is the solution of the corresponding variational inequality problem. The parameters of all algorithms are set as follows. Choose $\gamma=2$, $\ell=0.5$, $\phi=0.4$, $\tau=1.5$, $\zeta_n=1/(n+1)$, $\sigma_n=0.9(1-\zeta_n)$ and $f(x)=0.1x$ for all algorithms. Adopt $\delta=0.2$, $\epsilon_n=1/(n+1)^2$ and $\beta=0.8$ for the proposed algorithms. We use $D_n=\|x_n(t)-x^*(t)\|$ to measure the nth iteration error of all algorithms. We choose the maximum number of iterations of 20 as the common stopping criterion. Table 2 shows the numerical results of all algorithms with four starting points $x_0(t)=x_1(t)$.

Example 4.3 Consider the Hilbert space $\mathcal{H}=l_2:=\{x=(x_1,x_2,\ldots,x_i,\ldots)\mid \sum_{i=1}^{\infty}|x_i|^2<+\infty\}$ equipped with inner product

$$\langle x,y\rangle=\sum_{i=1}^{\infty}x_iy_i,\quad \forall x,y\in\mathcal{H}$$

and induced norm

$$\|x\|=\sqrt{\langle x,x\rangle},\quad \forall x\in\mathcal{H}.$$

Let $C:=\{x\in\mathcal{H}:|x_i|\le 1/i\}$. Define an operator $A:C\to\mathcal{H}$ by

$$Ax=\left(\|x\|+\frac{1}{\|x\|+\varphi}\right)x,\quad \varphi>0.$$

It can be verified that mapping A is pseudo-monotone on $\mathcal{H}$, uniformly continuous and sequentially weakly continuous on C but not Lipschitz continuous on $\mathcal{H}$ (see [27]). In this example, we take $\varphi=0.5$ and $\mathcal{H}=\mathbb{R}^m$ for different values of m. We compare the proposed algorithms with several strongly convergent ones that can solve the (VIP) with uniformly continuous operators, which including the Algorithm 4 proposed by Reich et al. [28] (shortly, RTDLD Alg. 4), the Algorithm 3 suggested by Thong et al. [29] (shortly, TSI Alg. 3), and the Algorithm 3.1 introduced by Cai et al. [30] (shortly, CDP Alg. 3.1). The parameters of all algorithms are set as follows.

- The parameters of the proposed algorithms are the same as those set in Example 4.1.
- In the RTDLD Alg. 4 [28], we take $\ell=0.5$, $\phi=0.4$, $\lambda=0.5/\phi$, $\zeta_n=1/(n+1)$, and $f(x)=0.1x$.
- In the TSI Alg. 3 [29] and the CDP Alg. 3.1 [30], we choose $\gamma=2$, $\ell=0.5$, $\phi=0.6$, $\zeta_n=1/(n+1)$, and $f(x)=0.1x$.

The initial values x_0 and x_1 are generated randomly by the function *rand(m,1)* in MATLAB. The maximum number of iterations 200 is used as a common stopping

Table 2 Numerical results of all algorithms for Example 4.2

Algorithms	$x_1(t) = t^2$		$x_1(t) = \mathrm{e}^t$		$x_1(t) = \sin(2t)$		$x_1(t) = \log(t)$	
	D_n	CPU	D_n	CPU	D_n	CPU	D_n	CPU
Our Alg. 3.1	3.39E−08	24.7869	2.97E−07	27.4160	7.80E−08	26.4355	5.57E−07	26.8071
Our Alg. 3.2	1.29E−08	23.9488	6.28E−08	26.8225	2.99E−08	25.3484	1.80E−07	26.2116
Our Alg. 3.3	1.56E−06	23.7991	6.06E−07	26.1013	2.52E−06	24.4954	6.96E−06	25.5490
Our Alg. 3.4	1.09E−06	23.8379	9.05E−07	25.9988	1.77E−06	24.6149	3.96E−06	25.5971
TG Alg. 3.1	3.32E−05	23.5227	9.60E−05	25.6065	5.59E−05	23.7105	1.20E−04	25.5283
TG Alg. 3.2	2.32E−05	23.3046	6.60E−05	25.4201	3.91E−05	24.2506	8.50E−05	25.5357
GTT Alg. 3.1	3.32E−05	22.1117	6.53E−05	25.2520	5.59E−05	22.3823	1.09E−04	25.2109
GTT Alg. 3.2	2.32E−05	21.8890	4.49E−05	25.0714	3.91E−05	22.5513	7.77E−05	25.1570

criterion. We measure the error of all algorithms at the nth iteration using $E_n = \|q_n - s_n\|$. By the property of projection (2.1), it is known that s_n can be seen as a solution of the problem (VIP) when $E_n \to 0$. The numerical results for all algorithms with four dimensions are reported in Table 3.

Remark 4.1 We have the following observations for Examples 4.1–4.3.

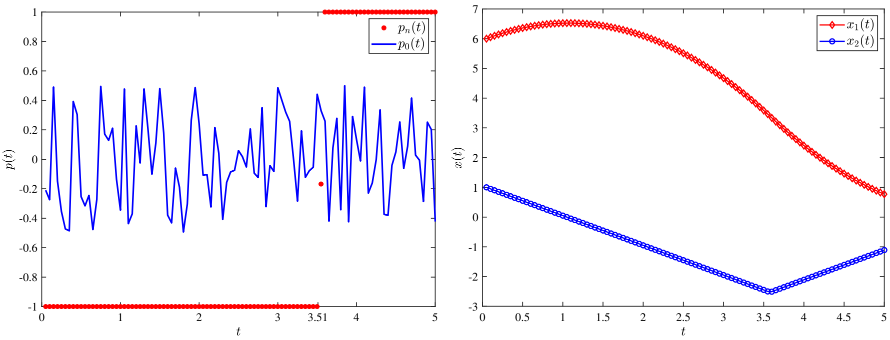
1. As can be seen in Fig. 1, different values of the parameter β have different effects on the proposed algorithms. When $\beta < 1$ can accelerate the convergence speed of the suggested methods when $\beta = 1$. Therefore, the schemes proposed in this paper are useful.
2. From Tables 1, 2 and 3, we know that our algorithms have a higher accuracy than some known methods in the literature [13,14,28–30] when performing the same stopping criterion, and this result is independent of the size of the dimension and the choice of the initial values. Therefore, our algorithms are efficient and robust.
3. It can be seen from Table 2 that our algorithms take more execution time in an infinite-dimensional space than the algorithms in [13,14], due to the fact that our algorithms need to compute the inertial parameters in each iteration. However, our algorithms can achieve a higher accuracy under the same stopping criterion.
4. In Example 4.3, the variational inequality operator A is pseudo-monotone and uniformly continuous rather than Lipschitz continuous. In this case, the algorithms used in [12–14] for solving monotone variational inequalities and the algorithms proposed in [15–18] for solving pseudo-monotone and Lipschitz continuous variational inequalities will not be available.
5. Our algorithms employ an Armijo step size criterion that allows them to work without the prior knowledge of the Lipschitz constant. The fixed step size algorithms proposed in [16,17,19] need to work with the knowledge of the Lipschitz constant. In other words, the fixed step size algorithms [16,17,19] will not work without the prior knowledge of the Lipschitz constant of the mapping. Therefore, the algorithms proposed in this paper are more useful than the fixed step size algorithms (e.g., [16,17,19]) in practical applications.

4.2 Application to optimal control problems

In this subsection, we use the proposed algorithms to solve the (VIP) that appears in optimal control problems. We recommend readers to refer to [1,31] for detailed description of the problem. Take $\gamma = 1, \ell = 0.5, \phi = 0.4, \tau = 1.5, \zeta_n = 10^{-4}/(n+1)$, $\sigma_n = 0.9(1 - \zeta_n)$, and $f(x) = 0.1x$ for all algorithms. Choose $\beta = 0.8$, $\delta = 0.01$, and $\epsilon_n = 10^{-4}/(n+1)^2$ for the proposed Algorithms 3.1 and 3.2. Select $\beta = 1.0$, $\delta = 0.01$, and $\epsilon_n = 10^{-4}/(n+1)^2$ for the proposed Algorithms 3.3 and 3.4. The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-1, 1]$ and the stopping criterion is $E_n = \|p_{n+1} - p_n\| \le 10^{-3}$.

Table 3 Numerical results of all algorithms for Example 4.3

Algorithms	$m = 100$		$m = 1000$		$m = 10000$		$m = 100000$	
	E_n	CPU	E_n	CPU	E_n	CPU	E_n	CPU
Our Alg. 3.1	4.05E−58	0.0243	1.21E−57	0.0338	2.19E−57	0.1854	2.36E−57	1.0939
Our Alg. 3.2	4.90E−86	0.0269	5.65E−86	0.0328	4.38E−86	0.1902	2.23E−86	1.1229
Our Alg. 3.3	3.88E−54	0.0235	7.93E−54	0.0306	5.91E−54	0.1553	7.53E−54	0.9775
Our Alg. 3.4	1.11E−72	0.0270	1.87E−72	0.0327	2.33E−72	0.1537	1.76E−72	0.9330
CDP Alg. 3.1	1.18E−27	0.0443	1.36E−27	0.0531	1.72E−27	0.2595	2.09E−27	1.4033
TSI Alg. 3	2.09E−11	0.0291	3.22E−11	0.0330	3.21E−11	0.1843	1.89E−11	1.1038
RTDLD Alg. 4	8.21E−11	0.0348	8.93E−11	0.0310	5.46E−09	0.1460	2.49E−03	0.9486



(a) Initial and optimal controls (b) Optimal trajectories

Fig. 2 Numerical results of the proposed Algorithm 3.1 for Example 4.4

Example 4.4 [Rocket car [31]]

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

The exact optimal control of Example 4.4 is $p^*(t) = -1$ if $t \in (0, 3.517]$ and $p^*(t) = 1$ if $t \in (3.517, 5]$. The approximate optimal control and the corresponding trajectories of Algorithm 3.1 are plotted in Fig. 2.

Example 4.5 (See [32])

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

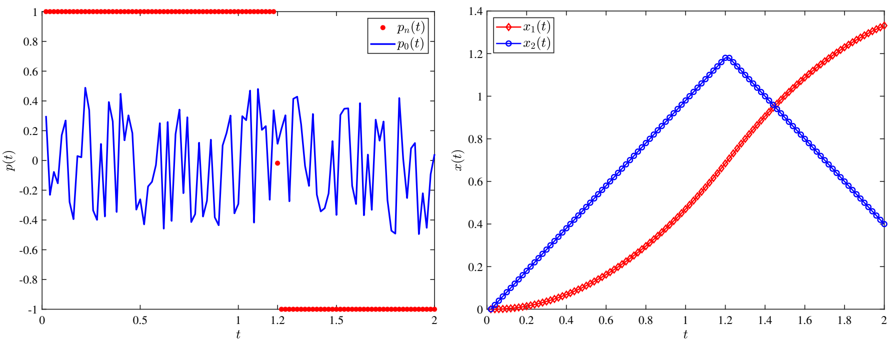
The exact optimal control of Example 4.5 is $p^*(t) = 1$ if $t \in [0, 1.2)$ and $p^*(t) = -1$ if $t \in (1.2, 2]$. Figure 3 gives the approximate optimal control and the corresponding trajectories of Algorithm 3.4.

Finally, the numerical results of all algorithms in Examples 4.4 and 4.5 are shown in Fig. 4 and Table 4.

Remark 4.2 From Figs. 2, 3, 4 and Table 4, we know that the suggested algorithms can be applied to solve optimal control problems and they perform well.

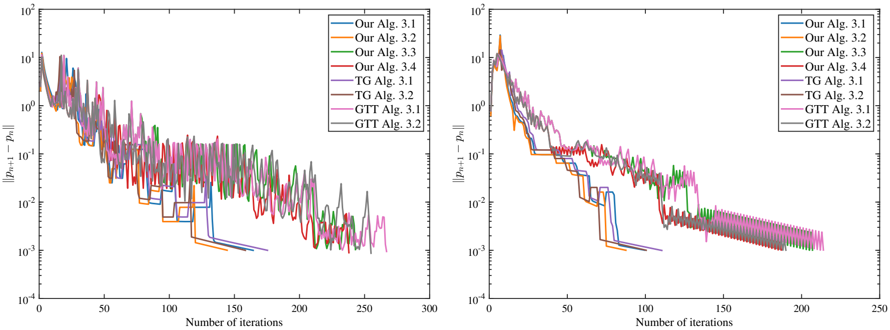
5 Conclusions

In this study, we proposed four accelerated inertial extragradient algorithms for solving variational inequalities in infinite-dimensional Hilbert spaces. Our methods are



(a) Initial and optimal controls (b) Optimal trajectories

Fig. 3 Numerical results of the proposed Algorithm 3.4 for Example 4.5



(a) Example 4.4 (b) Example 4.5

Fig. 4 Numerical behavior of all algorithms in Examples 4.4 and 4.5

Table 4 Numerical results of all algorithms for Examples 4.4 and 4.5

Algorithms	Example 4.4			Example 4.5		
	Iter.	CPU	E_n	Iter.	CPU	E_n
Our Alg. 3.1	164	0.1343	9.8832E−04	99	0.0518	9.9745E−04
Our Alg. 3.2	144	0.1065	9.9391E−04	87	0.0391	9.9366E−04
Our Alg. 3.3	242	0.2458	9.9761E−04	206	0.1204	9.6920E−04
Our Alg. 3.4	237	0.2360	8.5812E−04	186	0.1047	9.7685E−04
TG Alg. 3.1	175	0.1200	9.9061E−04	110	0.0445	9.8535E−04
TG Alg. 3.2	158	0.1099	9.8821E−04	100	0.0411	9.9252E−04
GTT Alg. 3.1	266	0.2617	9.1061E−04	213	0.1196	9.7499E−04
GTT Alg. 3.2	254	0.2454	8.3673E−04	189	0.1059	9.6885E−04

inspired by the subgradient extragradient method, the projection and contraction method, the inertial method, the Mann method, and the viscosity method. Notice that the variational inequality operators involved in our methods are pseudo-monotone and non-Lipschitz continuous. The proposed algorithms employ an Armijo step size criterion allowing them to work without the prior information of the Lipschitz constant of the mapping. The strong convergence of the sequences generated by the proposed iterative schemes is proved under some suitable conditions imposed on the parameters. Finally, some numerical tests and applications verified the advantages and performance of the proposed algorithms with respect to previously known schemes. The results obtained in this paper improved and generalized many algorithms proposed in the literature for solving variational inequalities.

Acknowledgements The authors would like to thank the Editor and the anonymous referee for their valuable comments, which greatly improved the quality of the initial manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Vuong, P.T., Shehu, Y.: Convergence of an extragradient-type method for variational inequality with applications to optimal control problems. Numer. Algorithms **81**, 269–291 (2019)
2. Ansari, Q.H., Islam, M., Yao, J.-C.: Nonsmooth variational inequalities on Hadamard manifolds. Appl. Anal. **99**, 340–358 (2020)
3. Bonacker, E., Gibali, A., Küfer, K.-H.: Accelerating two projection methods via perturbations with application to intensity-modulated radiation therapy. Appl. Math. Optim. **83**, 881–914 (2021)
4. Sahu, D.R., Yao, J.-C., Verma, M., Shukla, K.K.: Convergence rate analysis of proximal gradient methods with applications to composite minimization problems. Optimization **70**, 75–100 (2021)
5. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Èkonom. i Mat. Metody **12**, 747–756 (1976)
6. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)
7. Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalities in Hilbert space. J. Optim. Theory Appl. **148**, 318–335 (2011)
8. Censor, Y., Gibali, A., Reich, S.: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. Optim. Methods Softw. **26**, 827–845 (2011)
9. Censor, Y., Gibali, A., Reich, S.: Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. Optimization **61**, 1119–1132 (2012)
10. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35**, 69–76 (1997)
11. Cai, X., Gu, G., He, B.: On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators. Comput. Optim. Appl. **57**, 339–363 (2014)
12. Dong, Q.-L., Jiang, D., Gibali, A.: A modified subgradient extragradient method for solving the variational inequality problem. Numer. Algorithms **79**, 927–940 (2018)
13. Thong, D.V., Gibali, A.: Two strong convergence subgradient extragradient methods for solving variational inequalities in Hilbert spaces. Jpn. J. Ind. Appl. Math. **36**, 299–321 (2019)
14. Gibali, A., Thong, D.V., Tuan, P.A.: Two simple projection-type methods for solving variational inequalities. Anal. Math. Phys. **9**, 2203–2225 (2019)

15. Gibali, A., Thong, D.V.: A new low-cost double projection method for solving variational inequalities. Optim. Eng. **21**, 1613–1634 (2020)
16. Shehu, Y., Dong, Q.-L., Jiang, D.: Single projection method for pseudo-monotone variational inequality in Hilbert spaces. Optimization **68**, 385–409 (2019)
17. Cholamjiak, P., Thong, D.V., Cho, Y.J.: A novel inertial projection and contraction method for solving pseudomonotone variational inequality problems. Acta Appl. Math. **169**, 217–245 (2020)
18. Ogbuisi, F.U., Shehu, Y., Yao, J.-C.: An alternated inertial method for pseudomonotone variational inequalities in Hilbert spaces. Optim. Eng. (2021). https://doi.org/10.1007/s11081-021-09615-1
19. Dong, Q.-L., Cho, Y.J., Zhong, L.L., Rassias, T.M.: Inertial projection and contraction algorithms for variational inequalities. J. Global Optim. **70**, 687–704 (2018)
20. Shehu, Y., Iyiola, O.S.: Projection methods with alternating inertial steps for variational inequalities: weak and linear convergence. Appl. Numer. Math. **157**, 315–337 (2020)
21. Shehu, Y., Gibali, A.: New inertial relaxed method for solving split feasibilities. Optim. Lett. **15**, 2109–2126 (2021)
22. Liu, L., Cho, S.Y., Yao, J.-C.: Convergence analysis of an inertial Tseng's extragradient algorithm for solving pseudomonotone variational inequalities and applications. J. Nonlinear Var. Anal. **5**, 627–644 (2021)
23. Tan, B., Qin, X.: Self adaptive viscosity-type inertial extragradient algorithms for solving variational inequalities with applications. Math. Model. Anal. (2022). https://doi.org/10.3846/mma.2022.13846
24. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. **75**, 742–750 (2012)
25. Tan, B., Cho, S.Y.: Inertial extragradient methods for solving pseudomonotone variational inequalities with non-Lipschitz mappings and their optimization applications. Appl. Set-Valued Anal. Optim. **3**, 165–192 (2021)
26. Hieu, D.V., Anh, P.K., Muu, L.D.: Modified hybrid projection methods for finding common solutions to variational inequality problems. Comput. Optim. Appl. **66**, 75–96 (2017)
27. Thong, D.V., Shehu, Y., Iyiola, O.S.: A new iterative method for solving pseudomonotone variational inequalities with non-Lipschitz operators. Comput. Appl. Math. **39**, 108 (2020)
28. Reich, S., Thong, D.V., Dong, Q.-L., Li, X.H., Dung, V.T.: New algorithms and convergence theorems for solving variational inequalities with non-Lipschitz mappings. Numer. Algorithms **87**, 527–549 (2021)
29. Thong, D.V., Shehu, Y., Iyiola, O.S.: Weak and strong convergence theorems for solving pseudo-monotone variational inequalities with non-Lipschitz mappings. Numer. Algorithms **84**, 795–823 (2020)
30. Cai, G., Dong, Q.-L., Peng, Y.: Strong convergence theorems for solving variational inequality problems with pseudo-monotone and non-Lipschitz operators. J. Optim. Theory Appl. **188**, 447–472 (2021)
31. Preininger, J., Vuong, P.T.: On the convergence of the gradient projection method for convex optimal control problems with bang–bang solutions. Comput. Optim. Appl. **70**, 221–238 (2018)
32. Bressan, B., Piccoli, B.: Introduction to the Mathematical Theory of Control. AIMS Series on Applied Mathematics, San Francisco (2007)