Journal of Global Optimization (2022) 82:523–557
https://doi.org/10.1007/s10898-021-01095-y

Strong convergence of inertial projection and contraction methods for pseudomonotone variational inequalities with applications to optimal control problems

Bing Tan[1] · Xiaolong Qin[2] · Jen-Chih Yao[3,4]

Received: 5 February 2021 / Accepted: 15 September 2021 / Published online: 6 October 2021

Abstract

This paper investigates some inertial projection and contraction methods for solving pseudomonotone variational inequality problems in real Hilbert spaces. The algorithms use a new non-monotonic step size so that they can work without the prior knowledge of the Lipschitz constant of the operator. Strong convergence theorems of the suggested algorithms are obtained under some suitable conditions. Some numerical experiments in finite- and infinite-dimensional spaces and applications in optimal control problems are implemented to demonstrate the performance of the suggested schemes and we also compare them with several related results.

Keywords Variational inequality problem · Projection and contraction method · Subgradient extragradient method · Inertial method · Pseudomonotone mapping · Optimal control problem

Mathematics Subject Classification 47H05 · 47H09 · 49J15 · 47J20 · 65K15

✉ Xiaolong Qin
qxlxajh@163.com

Bing Tan
bingtan72@gmail.com

Jen-Chih Yao
yaojc@mail.cmu.edu.tw

1 Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China

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

3 Research Center for Interneural Computing, China Medical University Hospital, China Medical University, Taichung 40447, Taiwan

4 Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 80424, Taiwan

1 Introduction

The theory of variational inequalities has become an important tool of pure and applied sciences. It builds a unified and general framework for many problems. It is known that numerous problems in society and science can be represented via the model of variational inequalities that plays an essential role in both nonlinear optimization theory and practical applications; see, for example, [1,8,19,26,27]. Recently, variational inequalities attracted considerable attention from many researchers who are interested not only in obtaining theoretical results but also in numerical approaches to solve such problems approximately. Let us first review the classical variational inequality problem (shortly, VIP) whose form is as follows:

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

where $\mathcal{C}$ is a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$, and $M : \mathcal{H} \to \mathcal{H}$ is a nonlinear mapping. The solution set of (VIP) is denoted by $\mathrm{VI}(\mathcal{C}, M)$ in this paper.

The earliest and easiest way to solve (VIP) is the projected gradient method. However, the convergence of this method needs to ensure the strong monotonicity of the cost operator, which limits the implementation of the method. Recently, a large number of approaches have been exploited to avoid the hypothesis of the strong monotonicity of mapping M; see, e.g., [6,9,20,35]. In this paper, we mainly consider the methods based on projections. The extragradient method (EGM) was proposed by Korpelevich [20] to solve the monotone VIP. In view of the EGM, we see that it needs to calculate two projections on the feasible set $\mathcal{C}$ in each iteration, which may severely affect the computational performance of the algorithm when $\mathcal{C}$ has a complex structure. There are three noteworthy methods in the literature to overcome the shortcomings of the EGM. The first is the Tseng's extragradient method (TEGM) suggested by Tseng [35]. This method is a two-step iterative scheme and the second step does not involve any projection steps. The second is the subgradient extragradient method (SEGM) proposed by Censor, Gibali and Reich [6]. The SEGM replaces the second projection of the EGM with the projection on a half-space. It is known that the projection on a half-space can be calculated by an explicit formula. The third approach is the projection and contraction method (PCM) considered in [17,33], which is a two-step iterative method and only calculates one projection per iteration. It is worth noting that TEGM, SEGM and PCM only calculate the projection on the feasible set $\mathcal{C}$ once in each iteration, which greatly improves the computational efficiency of such algorithms.

Note that the algorithms mentioned above all achieve weak convergence in infinite-dimensional Hilbert spaces. However, the applications appearing in medical imaging and machine learning tell us that the strong convergence is preferable to the weak convergence in an infinite-dimensional space. In 2017, Yekini and Olaniyi [28] suggested a viscosity-type modification of the SEGM with the adoption of the Armijo-like step size rule. Under some standard and mild conditions, they established the strong convergence of the suggested iterative scheme. In addition, inspired by the Mann-type method, the SEGM and the TEGM, Thong and Hieu [36] proposed two new self-adaptive iterative methods that use some previous results to automatically update the step size. Under some suitable assumptions, they obtained the strong convergence of the iterative sequence generated by their algorithms. Recently, Dong, Jiang and Gibali [11] stated a modified subgradient extragradient method that combines the SEGM and the PCM. They proved that the iterative sequence formed by their method converges weakly to a solution of the (VIP) under some conditions. Their example illustrates the numerical achievements and benefits of this new method compared with the SEGM and the PCM. Excited by the work of Dong et al. [11], Thong and Gibali [37], and Gibali et al. [15] obtained some strongly convergent methods to solve the monotone VIP by combining the Mann method and the viscosity method.

In recent years, the development of fast iterative algorithms has aroused great interest from the researchers working in the signal and imaging. They built various fast numerical algorithms by employing the inertial technology, see, for instance, [12,16,29,30,39,40] and the references therein. One of the common features of these algorithms is that the next iteration depends on the combination of the previous two iterations. Note that this minor change greatly improves the performance of the algorithm used. Recently, Dong et al. [13] introduced an inertial projection and contraction method (IPCM) by connecting the inertial method and the PCM to solve the monotone VIP. They confirmed that their iterative procedure achieved the weak convergence in Hilbert spaces under suitable assumptions. Moreover, the IPCM has shown benefits and performance over other algorithms through some computational tests. By associating the IPCM with the Mann method and the viscosity method, respectively, Thong et al. [38] and Cholamjiak et al. [7] established the strong convergence theorems of the proposed iterative schemes.

Note that the above mentioned methods [11,13,15,28,36–38] were obtained under the premise that the operator is monotone. However, many problems do not satisfy the fact that the potential operator is monotone. Since the class of pseudomonotone mappings includes the class of monotone mappings, this paper focuses on the pseudomonotone VIP. There are some numerical methods based on the SEGM and the PCM in the literature [7,31,32,42] that can solve the pseudomonotone VIP. It should be pointed out that the methods suggested in [32,42] achieve the weak convergence only. Furthermore, the algorithms proposed in [7,31,38] need to know the prior information of the Lipschitz constant of the mapping, which limits the realization of such algorithms when the Lipschitz constant of the mapping associated with the problem is unknown. It is known that the prior information of the Lipschitz constant of the operator in practical problems is not easy to obtain. A natural problem that arises is how to update the iteration step size when we do not know the prior information of the Lipschitz constant of the cost operator related to the problem. The methods suggested in [11,15,28,37] used the Armijo-like criteria to update the iteration step size. The disadvantage of this criterion is that, in each iteration, in order to determine an approximate step size, the value of operator M needs to be evaluated multiple times, which will greatly increase the execution time of the algorithm used. Recently, Shehu and Iyiola [32] studied the IPCM with a simple step size to solve the VIP and established the strong convergence of the proposed iterative method in infinite-dimensional Hilbert spaces. However, the iteration step size in [32] will not increase from iteration to iteration, which will affect the calculation performance of the algorithm. Very recently, Yang [43] introduced two self-adaptive iterative schemes with a non-monotonic step size to solve the equilibrium problem and established the weak convergence of the algorithms in real Hilbert spaces. Their numerical experiments show that the suggested algorithms have competitive advantages over other ones.

In this paper, motivated and stimulated by the results mentioned above, we introduce several inertial projection and contraction methods to find the solution of the pseudomonotone VIP in real Hilbert spaces. These methods use a new non-monotonic step size criterion so that they do not need to know the Lipschitz constant of the pseudomonotone mapping. Furthermore, they embedded inertial terms to accelerate the convergence speed of the algorithms. Under reasonable assumptions on the parameters, the strong convergence theorems of the suggested algorithms are obtained. Our iterative schemes improve and extend some previously known results in [7,11,13,15,28,31,32,36–38,42].

The rest of this paper is organized as follows. In Sect. 2, we recall some definitions and lemmas that need to be used. Section 3 is devoted to describing the algorithms and analyzing their convergence. In Sect. 4, we present some computational tests to show the efficiency of the suggested approaches over several existing ones. In Sect. 5, the proposed methods are investigated to solve the optimal control problems. Finally, the paper ends with a brief remark in Sect. 6, the last section.

2 Preliminaries

Let $\mathcal{C}$ 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) $\|\alpha x + (1-\alpha)y\|^2 = \alpha\|x\|^2 + (1-\alpha)\|y\|^2 - \alpha(1-\alpha)\|x-y\|^2, \alpha \in \mathbb{R}$.
(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 $\mathcal{C}$, denoted by $P_{\mathcal{C}}(x)$, such that $P_{\mathcal{C}}(x) := \operatorname{argmin}\{\|x-y\|,\ y \in \mathcal{C}\}$. $P_{\mathcal{C}}$ is called the metric projection of $\mathcal{H}$ onto $\mathcal{C}$. It is known that $P_{\mathcal{C}}$ has the following basic properties:

- $\langle x - P_{\mathcal{C}}(x), y - P_{\mathcal{C}}(x)\rangle \leq 0,\ \forall y \in \mathcal{C}$.
- $\|P_{\mathcal{C}}(x) - P_{\mathcal{C}}(y)\|^2 \leq \langle P_{\mathcal{C}}(x) - P_{\mathcal{C}}(y), x-y\rangle,\ \forall y \in \mathcal{H}$.

We give some projection calculation formulas that need to be used in numerical experiments. For more calculations on projections on specific sets, we refer to [2].

(1) The projection of x onto a half-space $H_{u,v} = \{x \in \mathbb{R}^n : \langle u, x\rangle \leq v\}$ 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 $\mathrm{Box}[a,b] = \{x \in \mathbb{R}^n : a \leq x \leq b\}$ is computed by

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

(3) The projection of x onto the intersection of a hyperplane and a box $\mathcal{C} = H_{u,v} \cap \mathrm{Box}[a,b] = \left\{x \in \mathbb{R}^n : u^\mathsf{T} x = v, a \leq x \leq b\right\}$ is computed by

$$P_{\mathcal{C}}(x) = P_{\mathrm{Box}[a,b]}(x - \mu^* u)\,,$$

where μ^* is a solution of the equation $\varphi(\mu) = u^\mathsf{T} P_{\mathrm{Box}[a,b]}(x - \mu u) - v$.

(4) The projection of x onto a ball $B[p,q] = \{x \in \mathbb{R}^n : \|x-p\| \leq q\}$ is computed by

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

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

(1) *η-strongly monotone* with $\eta > 0$ if $\langle Mx - My, x-y\rangle \geq \eta\|x-y\|^2$.
(2) *L-Lipschitz continuous* with $L > 0$ if $\|Mx - My\| \leq L\|x-y\|$. If $L \in (0,1)$ then mapping M is called *contraction*. In particular, when $L = 1$, mapping M is called *nonexpansive*.
(3) *monotone* if $\langle Mx - My, x-y\rangle \geq 0$.
(4) *pseudomonotone* if $\langle Mx, y-x\rangle \geq 0 \Longrightarrow \langle My, y-x\rangle \geq 0$.

(5) *sequentially weakly continuous* if for each sequence $\{x_n\}$ converges weakly to x implies $\{Mx_n\}$ converges weakly to Mx.

According to the above definitions, we see that the following fact holds: (1) $\Longrightarrow$ (3) $\Longrightarrow$ (4). But the inverse operation is usually not true as can be seen in the following example.

Example 2.1 Let $M : (0, \infty) \to (0, \infty)$ be a mapping defined by $Mx = \frac{a}{a+x}$ with $a > 0$. It is easy to check that M is pseudomonotone but not monotone.

The following lemmas will be used frequently in the proof of our main results.

Lemma 2.1 ([10]) *Assume that $\mathcal{C}$ is a closed and convex subset of a real Hilbert space $\mathcal{H}$. Let operator $M : \mathcal{C} \to \mathcal{H}$ be continuous and pseudomonotone. Then, x^* is a solution of (VIP) if and only if $\langle Mx, x - x^* \rangle \geq 0$, $\forall x \in \mathcal{C}$.*

Lemma 2.2 ([34]) *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 $\sum_{n=1}^{\infty} \sigma_n = \infty$. Suppose that $p_{n+1} \leq \sigma_n q_n + (1 - \sigma_n) p_n$, $\forall n \geq 1$. If $\limsup_{k\to\infty} q_{n_k} \leq 0$ for every subsequence $\{p_{n_k}\}$ of $\{p_n\}$ satisfying $\liminf_{k\to\infty} \left(p_{n_k+1} - p_{n_k}\right) \geq 0$, then $\lim_{n\to\infty} p_n = 0$.*

3 Main results

In this section, we introduce six new iterative schemes based on the SEGM and the IPCM to solve the pseudomonotone VIP in a real Hilbert space. These algorithms guarantee the strong convergence with the aid of the Mann-type method and the viscosity-type method. The advantage of our approaches is that we do not need to know the Lipschitz constant of the pseudomonotone mapping in advance. In order to analyze the convergence of the algorithms, the mapping and parameters involved in our methods need to meet the following assumptions.

(1) The feasible set $\mathcal{C}$ is a nonempty, closed and convex subset of $\mathcal{H}$;
(2) The solution set of the (VIP) is nonempty, that is $\mathrm{VI}(\mathcal{C}, M) \neq \emptyset$;
(3) The mapping $M : \mathcal{H} \to \mathcal{H}$ is L-Lipschitz continuous and pseudomonotone on $\mathcal{H}$, and sequentially weakly continuous on $\mathcal{C}$;
(4) Let $\{\epsilon_n\}$ and $\{\xi_n\}$ be two nonnegative sequences such that $\lim_{n\to\infty} \frac{\epsilon_n}{\sigma_n} = 0$ and $\sum_{n=1}^{\infty} \xi_n < +\infty$, where $\{\sigma_n\} \subset (0, 1)$ satisfies $\lim_{n\to\infty} \sigma_n = 0$ and $\sum_{n=1}^{\infty} \sigma_n = \infty$. Let $\{\varphi_n\} \subset (a, b) \subset (0, 1 - \sigma_n)$ for some $a > 0, b > 0$.

Remark 3.1 We note here that the assumption (C4) is easily satisfied by, for example, taking $\sigma_n = 1/(n+1)$, $\epsilon_n = 1/(n+1)^2$, $\varphi_n = 0.5(1 - \sigma_n)$ and $\xi_n = 1/(n+1)^{1.1}$. Moreover, it is not necessary to impose the sequential weak continuity when mapping M is monotone, see [14].

3.1 The Mann-type inertial subgradient extragradient algorithm

The first algorithm is based on the IPCM, the SEGM and the Mann-type method, and its details are described below.

Remark 3.2 It follows from (3.1) and Assumption (C4) that

$$\lim_{n\to\infty} \frac{\tau_n}{\sigma_n} \|x_n - x_{n-1}\| = 0\,.$$

Indeed, we obtain $\tau_n \|x_n - x_{n-1}\| \le \epsilon_n$, $\forall n \ge 1$, which together with $\lim_{n\to\infty} \frac{\epsilon_n}{\sigma_n} = 0$ implies that

$$\lim_{n\to\infty} \frac{\tau_n}{\sigma_n} \|x_n - x_{n-1}\| \le \lim_{n\to\infty} \frac{\epsilon_n}{\sigma_n} = 0\,.$$

The following basic lemmas are very helpful in analyzing the convergence of the algorithms.

Lemma 3.1 *Suppose that Assumption (C3) holds. The sequence* $\{\vartheta_n\}$ *generated by* (3.2) *is well defined and* $\lim_{n\to\infty} \vartheta_n = \vartheta$ *and* $\vartheta \in \left[\min\left\{\frac{\mu}{L}, \vartheta_1\right\}, \vartheta_1 + \Xi\right]$, *where* $\Xi = \sum_{n=1}^{\infty} \xi_n$.

Proof From the fact that mapping M is L-Lipschitz continuous, one has

$$\frac{\mu \|u_n - y_n\|}{\|Mu_n - My_n\|} \ge \frac{\mu \|u_n - y_n\|}{L\|u_n - y_n\|} = \frac{\mu}{L}\,, \text{ if } Mu_n \ne My_n\,.$$

Thus, $\vartheta_n \ge \min\left\{\frac{\mu}{L}, \vartheta_1\right\}$. It follows from the definition of ϑ_{n+1} that $\vartheta_{n+1} \le \vartheta_1 + \Xi$. Consequently, the sequence $\{\vartheta_n\}$ defined in (3.2) is bounded and $\vartheta_n \in \left[\min\left\{\frac{\mu}{L}, \vartheta_1\right\}, \vartheta_1 + \Xi\right]$. For simplicity, we define

$$(\vartheta_{n+1} - \vartheta_n)^+ = \max\{0, \vartheta_{n+1} - \vartheta_n\}$$

and

$$(\vartheta_{n+1} - \vartheta_n)^- = \max\{0, -(\vartheta_{n+1} - \vartheta_n)\}\,.$$

By the definition of $\{\vartheta_n\}$, one obtains $\sum_{n=1}^{\infty} (\vartheta_{n+1} - \vartheta_n)^+ \le \sum_{n=1}^{\infty} \xi_n < +\infty$, which implies that the series $\sum_{n=1}^{\infty} (\vartheta_{n+1} - \vartheta_n)^+$ is convergent.

Next, we show the convergence of $\sum_{n=1}^{\infty} (\vartheta_{n+1} - \vartheta_n)^-$. Suppose that $\sum_{n=1}^{\infty} (\vartheta_{n+1} - \vartheta_n)^- = +\infty$. Note that $\vartheta_{n+1} - \vartheta_n = (\vartheta_{n+1} - \vartheta_n)^+ - (\vartheta_{n+1} - \vartheta_n)^-$. Therefore,

$$\vartheta_{k+1} - \vartheta_1 = \sum_{n=1}^{k} (\vartheta_{n+1} - \vartheta_n) = \sum_{n=1}^{k} (\vartheta_{n+1} - \vartheta_n)^+ - \sum_{n=1}^{k} (\vartheta_{n+1} - \vartheta_n)^-\,.$$

Taking $k \to +\infty$ in the above equation, we obtain $\lim_{k\to+\infty} \vartheta_k \to -\infty$, which is a contradiction. Hence, we deduce that $\lim_{n\to\infty} \vartheta_n = \vartheta$ and $\vartheta \in \left[\min\left\{\frac{\mu}{L}, \vartheta_1\right\}, \vartheta_1 + \Xi\right]$. □

Remark 3.3 The idea of the step size ϑ_n defined in (3.2) is derived from [23]. It is worth noting that the step size ϑ_n generated in Algorithm 3.1 is allowed to increase when the iteration increases. Therefore, the use of this type of step size reduces the dependence on the initial step size ϑ_1. On the other hand, because of $\sum_{n=1}^{\infty} \xi_n < +\infty$, which implies that $\lim_{n\to\infty} \xi_n = 0$. Consequently, the step size ϑ_n may not increase when n is large enough. If $\xi_n = 0$, then the step size ϑ_n in Algorithm 3.1 is similar to the approaches in [32,36].

Lemma 3.2 *If* $y_n = u_n$ *or* $c_n = 0$ *in Algorithm* 3.1, *then* $y_n \in \mathrm{VI}(\mathcal{C}, M)$.

Proof From the L-Lipschitz continuity of the mapping M and (3.2), we obtain

$$\begin{aligned}\|c_n\| &\ge \|u_n - y_n\| - \vartheta_n \|Mu_n - My_n\| \\ &\ge \|u_n - y_n\| - \frac{\mu\vartheta_n}{\vartheta_{n+1}} \|u_n - y_n\| \\ &= \left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right) \|u_n - y_n\|\,.\end{aligned}$$

Algorithm 3.1 The Mann-type inertial subgradient extragradient algorithm

Initialization: Take $\tau > 0$, $\mu \in (0,1)$, $\vartheta_1 > 0$, $\theta \in (0,2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:
Step 1. Given the iterates x_{n-1} and x_n ($n \geq 1$). Set $u_n = x_n + \tau_n(x_n - x_{n-1})$, where

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

Step 2. Compute $y_n = P_{\mathcal{C}}(u_n - \vartheta_n M u_n)$, where step size ϑ_{n+1} is updated by

$$\vartheta_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\|u_n - y_n\|}{\|Mu_n - My_n\|}, \vartheta_n + \xi_n\right\}, & \text{if } Mu_n \neq My_n; \\ \vartheta_n + \xi_n, & \text{otherwise.} \end{cases} \tag{3.2}$$

If $u_n = y_n$, then stop and y_n is a solution of (VIP). Otherwise, go to **Step 3**.
Step 3. Compute $z_n = P_{T_n}(u_n - \theta\vartheta_n\chi_n M y_n)$, where the half-space T_n is defined by $T_n := \{x \in \mathcal{H} \mid \langle u_n - \vartheta_n M u_n - y_n, x - y_n\rangle \leq 0\}$, and

$$\chi_n = \frac{\langle u_n - y_n, c_n\rangle}{\|c_n\|^2}, \quad c_n = u_n - y_n - \vartheta_n(Mu_n - My_n). \tag{3.3}$$

Step 4. Compute $x_{n+1} = (1 - \sigma_n - \varphi_n)u_n + \varphi_n z_n$.
Set $n := n+1$ and go to **Step 1**.

It can be easily proved that $\|c_n\| \leq \left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)\|u_n - y_n\|$. Therefore, we have

$$\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)\|u_n - y_n\| \leq \|c_n\| \leq \left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)\|u_n - y_n\|,$$

and thus $u_n = y_n$ iff $c_n = 0$. Hence, if $u_n = y_n$ or $c_n = 0$, then $y_n = P_{\mathcal{C}}(y_n - \vartheta_n M y_n)$. This implies that $y_n \in \mathrm{VI}(\mathcal{C}, M)$. The proof is completed. □

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

Proof From the property of projection and $y_n = P_{\mathcal{C}}(u_n - \vartheta_n M u_n)$, we have

$$\langle u_{n_k} - \vartheta_{n_k} M u_{n_k} - y_{n_k}, x - y_{n_k}\rangle \leq 0, \quad \forall x \in \mathcal{C},$$

which can be written as follows

$$\frac{1}{\vartheta_{n_k}}\langle u_{n_k} - y_{n_k}, x - y_{n_k}\rangle \leq \langle M u_{n_k}, x - y_{n_k}\rangle, \quad \forall x \in \mathcal{C}.$$

Through a direct calculation, we obtain

$$\frac{1}{\vartheta_{n_k}}\langle u_{n_k} - y_{n_k}, x - y_{n_k}\rangle + \langle M u_{n_k}, y_{n_k} - u_{n_k}\rangle \leq \langle M u_{n_k}, x - u_{n_k}\rangle, \quad \forall x \in \mathcal{C}. \tag{3.4}$$

We obtain that $\{u_{n_k}\}$ is bounded since $\{u_{n_k}\}$ is converges weakly to $z \in \mathcal{H}$. Combining the Lipschitz continuity of mapping M and $\{u_{n_k}\}$ is bounded, we have $\{Mu_{n_k}\}$ is bounded. It follows form $\|u_{n_k} - y_{n_k}\| \to 0$ that $\{\|u_{n_k} - y_{n_k}\|\}$ is bounded, which together with the boundedness of $\{u_{n_k}\}$ and the inequality $\|y_{n_k}\| \leq \|u_{n_k}\| + \|y_{n_k} - u_{n_k}\|$, yields that $\{y_{n_k}\}$ is also bounded. One concludes from (3.4) that

$$\liminf_{k\to\infty}\langle M u_{n_k}, x - u_{n_k}\rangle \geq 0, \quad \forall x \in \mathcal{C}. \tag{3.5}$$

Moreover, one has

$$\langle My_{n_k}, x - y_{n_k}\rangle = \langle My_{n_k} - Mu_{n_k}, x - u_{n_k}\rangle + \langle Mu_{n_k}, x - u_{n_k}\rangle + \langle My_{n_k}, u_{n_k} - y_{n_k}\rangle . \tag{3.6}$$

Since $\lim_{k\to\infty} \|u_{n_k} - y_{n_k}\| = 0$ and the mapping M is Lipschitz continuous, we obtain that

$$\lim_{k\to\infty} \|Mu_{n_k} - My_{n_k}\| = 0.$$

This together with (3.5) and (3.6) yields that $\liminf_{k\to\infty}\langle My_{n_k}, x - y_{n_k}\rangle \geq 0$.

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

$$\langle My_{n_j}, x - y_{n_j}\rangle + \zeta_k \geq 0, \quad \forall j \geq N_k . \tag{3.7}$$

It can be easily seen that the sequence $\{N_k\}$ is increasing because $\{\zeta_k\}$ is decreasing. Moreover, for any k, from $\{y_{N_k}\} \subset \mathcal{C}$, we can assume $My_{N_k} \neq 0$ (otherwise, y_{N_k} is a solution) and set $s_{N_k} = My_{N_k}/\|My_{N_k}\|^2$. Then, we obtain $\langle My_{N_k}, s_{N_k}\rangle = 1, \forall k$. Now, we can deduce from (3.7) that $\langle My_{N_k}, x + \zeta_k s_{N_k} - y_{N_k}\rangle \geq 0, \forall k$. According to the fact that M is pseudomonotone on $\mathcal{H}$, we can show that

$$\langle M\left(x + \zeta_k s_{N_k}\right), x + \zeta_k s_{N_k} - y_{N_k}\rangle \geq 0\,,$$

which further yields that

$$\langle Mx, x - y_{N_k}\rangle \geq \langle Mx - M\left(x + \zeta_k s_{N_k}\right), x + \zeta_k s_{N_k} - y_{N_k}\rangle - \zeta_k\langle Mx, s_{N_k}\rangle . \tag{3.8}$$

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

$$0 \leq \limsup_{k\to\infty} \|\zeta_k s_{N_k}\| = \limsup_{k\to\infty} \left(\frac{\zeta_k}{\|My_{n_k}\|}\right) \leq \frac{\limsup_{k\to\infty} \zeta_k}{\liminf_{k\to\infty} \|My_{n_k}\|} = 0\,.$$

That is, $\lim_{k\to\infty} \zeta_k s_{N_k} = 0$. Thus, from the fact that M is Lipschitz continuous, the sequences $\{y_{N_k}\}$ and $\{u_{N_k}\}$ are bounded and $\lim_{k\to\infty} \zeta_k s_{N_k} = 0$, we can conclude from (3.8) that $\liminf_{k\to\infty}\langle Mx, x - y_{N_k}\rangle \geq 0$. Therefore,

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

Consequently, we observe that $z \in \mathrm{VI}(\mathcal{C}, M)$ by means of Lemma 2.1. This completes the proof. □

Lemma 3.4 *Suppose that Assumptions (C1)–(C3) hold. Let $\{z_n\}$, $\{y_n\}$ and $\{u_n\}$ be three sequences created by Algorithm 3.1. Then, for all $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$,*

$$\|z_n - x^\dagger\|^2 \leq \|u_n - x^\dagger\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2 - \theta(2-\theta)\frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2\,.$$

Proof From $x^\dagger \in \mathrm{VI}(\mathcal{C}, M) \subset \mathcal{C} \subset T_n$ and the property of projection, we obtain

$$\begin{aligned}2\|z_n - x^\dagger\|^2 &= 2\|P_{T_n}(u_n - \theta\vartheta_n\chi_n M y_n) - P_{T_n}(x^\dagger)\|^2 \le 2\langle z_n - x^\dagger, u_n - \theta\vartheta_n\chi_n M y_n - x^\dagger\rangle \\ &= \|z_n - x^\dagger\|^2 + \|u_n - \theta\vartheta_n\chi_n M y_n - x^\dagger\|^2 - \|z_n - u_n + \theta\vartheta_n\chi_n M y_n\|^2 \\ &= \|z_n - x^\dagger\|^2 + \|u_n - x^\dagger\|^2 + \theta^2\vartheta_n^2\chi_n^2\|M y_n\|^2 - 2\langle u_n - x^\dagger, \theta\vartheta_n\chi_n M y_n\rangle \\ &\quad - \|z_n - u_n\|^2 - \theta^2\vartheta_n^2\chi_n^2\|M y_n\|^2 - 2\langle z_n - u_n, \theta\vartheta_n\chi_n M y_n\rangle \\ &= \|z_n - x^\dagger\|^2 + \|u_n - x^\dagger\|^2 - \|z_n - u_n\|^2 - 2\langle z_n - x^\dagger, \theta\vartheta_n\chi_n M y_n\rangle\,,\end{aligned}$$

which implies that

$$\|z_n - x^\dagger\|^2 \le \|u_n - x^\dagger\|^2 - \|z_n - u_n\|^2 - 2\theta\vartheta_n\chi_n\langle z_n - x^\dagger, M y_n\rangle\,. \tag{3.9}$$

By $y_n \in \mathcal{C}$ and $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$, one has $\langle M x^\dagger, y_n - x^\dagger\rangle \ge 0$, which combined with the pseudo-monotonicity of M yields $\langle M y_n, y_n - x^\dagger\rangle \ge 0$. This means that $\langle M y_n, z_n - x^\dagger\rangle \ge \langle M y_n, z_n - y_n\rangle$. Hence,

$$-2\theta\vartheta_n\chi_n\langle M y_n, z_n - x^\dagger\rangle \le -2\theta\vartheta_n\chi_n\langle M y_n, z_n - y_n\rangle\,. \tag{3.10}$$

Since $z_n \in T_n$, we have $\langle u_n - \vartheta_n M u_n - y_n, z_n - y_n\rangle \le 0$. This shows that

$$\langle u_n - y_n - \vartheta_n(M u_n - M y_n), z_n - y_n\rangle \le \vartheta_n\langle M y_n, z_n - y_n\rangle\,. \tag{3.11}$$

Using (3.10), (3.11), and the definitions of c_n and χ_n, we obtain

$$\begin{aligned}-2\theta\vartheta_n\chi_n\langle M y_n, z_n - x^\dagger\rangle &\le -2\theta\chi_n\langle c_n, z_n - y_n\rangle \\ &= -2\theta\chi_n\langle c_n, u_n - y_n\rangle + 2\theta\chi_n\langle c_n, u_n - z_n\rangle \\ &= -2\theta\chi_n^2\|c_n\|^2 + 2\theta\chi_n\langle c_n, u_n - z_n\rangle\,.\end{aligned} \tag{3.12}$$

Now, we estimate $2\theta\chi_n\langle c_n, u_n - z_n\rangle$. According to the basic inequality $2ab = a^2 + b^2 - (a-b)^2$, we also have

$$2\theta\chi_n\langle c_n, u_n - z_n\rangle = \|u_n - z_n\|^2 + \theta^2\chi_n^2\|c_n\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2\,. \tag{3.13}$$

It follows from (3.2) that $\|M u_n - M y_n\| \le (\mu/\vartheta_{n+1})\|u_n - y_n\|$, $\forall n \ge 1$, which combining with the definition of χ_n yields that

$$\begin{aligned}\chi_n\|c_n\|^2 = \langle c_n, u_n - y_n\rangle &\ge \|u_n - y_n\|^2 - \vartheta_n\,\|M u_n - M y_n\|\,\|u_n - y_n\| \\ &\ge \left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)\|u_n - y_n\|^2\,.\end{aligned}$$

This together with the fact that $\|c_n\| \le (1 + \mu\vartheta_n/\vartheta_{n+1})\|u_n - y_n\|$ implies

$$\chi_n^2\|c_n\|^2 \ge \left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2\frac{\|u_n - y_n\|^4}{\|c_n\|^2} \ge \frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2\,. \tag{3.14}$$

Combining (3.9), (3.12), (3.13) and (3.14), we conclude that

$$\|z_n - x^\dagger\|^2 \le \|u_n - x^\dagger\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2 - \theta(2-\theta)\frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2\,.$$

This completes the proof. □

Theorem 3.1 *Suppose that Assumptions (C1)–(C4) hold. Then the sequence $\{x_n\}$ formed by Algorithm 3.1 converges to $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ in norm, where $\|x^\dagger\| = \min\{\|z\| : z \in \mathrm{VI}(\mathcal{C}, M)\}$.*

Proof For convenience, we divide the proof into four claims.
Claim 1. The sequence $\{x_n\}$ is bounded. Indeed, thanks to Lemma 3.4, one has

$$\|z_n - x^\dagger\| \le \|u_n - x^\dagger\|, \quad \forall n \ge 1 . \tag{3.15}$$

From the definition of u_n, one sees that

$$\|u_n - x^\dagger\| \le \|x_n - x^\dagger\| + \sigma_n \cdot \frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| . \tag{3.16}$$

According to Remark 3.2, we have $\frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| \to 0$ as $n \to \infty$. Therefore, there exists a constant $Q_1 > 0$ such that

$$\frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| \le Q_1, \quad \forall n \ge 1 ,$$

which together with (3.15) and (3.16) implies that

$$\|z_n - x^\dagger\| \le \|u_n - x^\dagger\| \le \|x_n - x^\dagger\| + \sigma_n Q_1, \quad \forall n \ge 1 . \tag{3.17}$$

By the definition of x_{n+1}, one obtains

$$\|x_{n+1} - x^\dagger\| \le \|(1 - \sigma_n - \varphi_n)(u_n - x^\dagger) + \varphi_n(z_n - x^\dagger)\| + \sigma_n\|x^\dagger\| . \tag{3.18}$$

It follows from (3.15) that

$$\begin{aligned}
&\|(1 - \sigma_n - \varphi_n)(u_n - x^\dagger) + \varphi_n(z_n - x^\dagger)\|^2 \\
&\quad \le (1 - \sigma_n - \varphi_n)^2\|u_n - x^\dagger\|^2 + 2(1 - \sigma_n - \varphi_n)\varphi_n\|z_n - x^\dagger\|\|u_n - x^\dagger\| + \varphi_n^2\|z_n - x^\dagger\|^2 \\
&\quad \le (1 - \sigma_n - \varphi_n)^2\|u_n - x^\dagger\|^2 + 2(1 - \sigma_n - \varphi_n)\varphi_n\|u_n - x^\dagger\|^2 + \varphi_n^2\|u_n - x^\dagger\|^2 \\
&\quad = (1 - \sigma_n)^2\|u_n - x^\dagger\|^2 ,
\end{aligned}$$

which yields

$$\|(1 - \sigma_n - \varphi_n)(u_n - x^\dagger) + \varphi_n(z_n - x^\dagger)\| \le (1 - \sigma_n)\|u_n - x^\dagger\| . \tag{3.19}$$

Using (3.17), (3.18) and (3.19), we deduce that

$$\begin{aligned}
\|x_{n+1} - x^\dagger\| &\le (1 - \sigma_n)\|u_n - x^\dagger\| + \sigma_n\|x^\dagger\| \\
&\le (1 - \sigma_n)\|x_n - x^\dagger\| + \sigma_n(\|x^\dagger\| + Q_1) \\
&\le \max\{\|x_n - x^\dagger\|, \|x^\dagger\| + Q_1\} \\
&\le \cdots \le \max\{\|x_0 - x^\dagger\|, \|x^\dagger\| + Q_1\} .
\end{aligned}$$

That is, the sequence $\{x_n\}$ is bounded. So the sequences $\{u_n\}$, $\{y_n\}$ and $\{z_n\}$ are also bounded.
Claim 2.

$$\begin{aligned}
&\varphi_n\|u_n - z_n - \theta\chi_n c_n\|^2 + \varphi_n\theta(2 - \theta)\frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2 \\
&\le \|x_n - x^\dagger\|^2 - \|x_{n+1} - x^\dagger\|^2 + \sigma_n(\|x^\dagger\|^2 + Q_2)
\end{aligned}$$

for some $Q_2 > 0$. Indeed, from (3.17), one sees that

$$\begin{aligned}\|u_n - x^\dagger\|^2 &\le (\|x_n - x^\dagger\| + \sigma_n Q_1)^2 \\ &= \|x_n - x^\dagger\|^2 + \sigma_n(2Q_1\|x_n - x^\dagger\| + \sigma_n Q_1^2) \\ &\le \|x_n - x^\dagger\|^2 + \sigma_n Q_2 \end{aligned} \tag{3.20}$$

for some $Q_2 > 0$. By the definition of x_{n+1} and Assumption (C4), we find that

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 &= \|(1-\sigma_n-\varphi_n)(u_n - x^\dagger) + \varphi_n(z_n - x^\dagger) + \sigma_n(-x^\dagger)\|^2 \\ &= (1-\sigma_n-\varphi_n)\|u_n - x^\dagger\|^2 + \varphi_n\|z_n - x^\dagger\|^2 + \sigma_n\|x^\dagger\|^2 \\ &\quad - \varphi_n(1-\sigma_n-\varphi_n)\|u_n - z_n\|^2 - \sigma_n(1-\sigma_n-\varphi_n)\|u_n\|^2 - \sigma_n\varphi_n\|z_n\|^2 \\ &\le (1-\sigma_n-\varphi_n)\|u_n - x^\dagger\|^2 + \varphi_n\|z_n - x^\dagger\|^2 + \sigma_n\|x^\dagger\|^2 . \end{aligned} \tag{3.21}$$

Thus, using Lemma 3.4, (3.20) and (3.21), we obtain

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 &\le (1-\sigma_n-\varphi_n)\|u_n - x^\dagger\|^2 + \varphi_n\|u_n - x^\dagger\|^2 - \varphi_n\|u_n - z_n - \theta\chi_n c_n\|^2 \\ &\quad - \varphi_n\theta(2-\theta)\frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2 + \sigma_n\|x^\dagger\|^2 \\ &\le \|x_n - x^\dagger\|^2 - \varphi_n\|u_n - z_n - \theta\chi_n c_n\|^2 \\ &\quad - \varphi_n\theta(2-\theta)\frac{\left(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n - y_n\|^2 + \sigma_n(\|x^\dagger\|^2 + Q_2) . \end{aligned}$$

The desired result can be achieved by a simple conversion.
Claim 3.

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 \le (1-\sigma_n)\|x_n - x^\dagger\|^2 + \sigma_n\big[&2\varphi_n\|u_n - z_n\|\|x_{n+1} - x^\dagger\| \\ &+ 2\langle x^\dagger, x^\dagger - x_{n+1}\rangle + \frac{3Q\tau_n}{\sigma_n}\|x_n - x_{n-1}\|\big]\end{aligned}$$

for some $Q > 0$. Indeed, from the definition of u_n, we can write

$$\begin{aligned}\|u_n - x^\dagger\|^2 &\le \|x_n - x^\dagger\|^2 + 2\tau_n\|x_n - x^\dagger\|\|x_n - x_{n-1}\| + \tau_n^2\|x_n - x_{n-1}\|^2 \\ &\le \|x_n - x^\dagger\|^2 + 3Q\tau_n\|x_n - x_{n-1}\| , \end{aligned} \tag{3.22}$$

where $Q := \sup_{n\in\mathbb{N}}\{\|x_n - x^\dagger\|, \tau\|x_n - x_{n-1}\|\} > 0$. Set $t_n = (1-\varphi_n)u_n + \varphi_n z_n$, it follows from (3.15) that

$$\begin{aligned}\|t_n - x^\dagger\| &= \|(1-\varphi_n)(u_n - x^\dagger) + \varphi_n(z_n - x^\dagger)\| \\ &\le (1-\varphi_n)\|u_n - x^\dagger\| + \varphi_n\|u_n - x^\dagger\| = \|u_n - x^\dagger\| . \end{aligned} \tag{3.23}$$

From (3.22) and (3.23), we obtain

$$\begin{aligned}\|x_{n+1}-x^\dagger\|^2 &= \|(1-\varphi_n)u_n+\varphi_n z_n-\sigma_n u_n-x^\dagger\|^2\\ &= \|(1-\sigma_n)(t_n-x^\dagger)-\sigma_n(u_n-t_n)-\sigma_n x^\dagger\|^2\\ &\le (1-\sigma_n)^2\|t_n-x^\dagger\|^2-2\sigma_n\langle u_n-t_n+x^\dagger, x_{n+1}-x^\dagger\rangle\\ &= (1-\sigma_n)^2\|t_n-x^\dagger\|^2+2\sigma_n\langle u_n-t_n, x^\dagger-x_{n+1}\rangle+2\sigma_n\langle x^\dagger, x^\dagger-x_{n+1}\rangle\\ &\le (1-\sigma_n)\|t_n-x^\dagger\|^2+2\sigma_n\|u_n-t_n\|\|x_{n+1}-x^\dagger\|+2\sigma_n\langle x^\dagger, x^\dagger-x_{n+1}\rangle\\ &\le (1-\sigma_n)\|x_n-x^\dagger\|^2+\sigma_n\big[2\varphi_n\|u_n-z_n\|\|x_{n+1}-x^\dagger\|\\ &\quad+2\langle x^\dagger, x^\dagger-x_{n+1}\rangle+\frac{3Q\tau_n}{\sigma_n}\|x_n-x_{n-1}\|\big].\end{aligned}$$

Claim 4. The sequence $\{\|x_n-x^\dagger\|^2\}$ converges to zero. From now on, we always assume that $\{\|x_{n_k}-x^\dagger\|\}$ is a subsequence of $\{\|x_n-x^\dagger\|\}$ such that $\liminf_{k\to\infty}(\|x_{n_k+1}-x^\dagger\|-\|x_{n_k}-x^\dagger\|)\ge 0$. Then,

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

By Claim 2 and Assumption (C4), we observe that

$$\begin{aligned}&\varphi_{n_k}\theta(2-\theta)\frac{\left(1-\frac{\mu\vartheta_{n_k}}{\vartheta_{n_k+1}}\right)^2}{\left(1+\frac{\mu\vartheta_{n_k}}{\vartheta_{n_k+1}}\right)^2}\|u_{n_k}-y_{n_k}\|^2+\varphi_{n_k}\|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|^2\\ &\quad\le\limsup_{k\to\infty}\left[\|x_{n_k}-x^\dagger\|^2-\|x_{n_k+1}-x^\dagger\|^2\right]+\limsup_{k\to\infty}\sigma_{n_k}\left(\|x^\dagger\|^2+Q_2\right)\\ &\quad=-\liminf_{k\to\infty}\left[\|x_{n_k+1}-x^\dagger\|^2-\|x_{n_k}-x^\dagger\|^2\right]\le 0,\end{aligned}$$

which implies that

$$\lim_{k\to\infty}\|y_{n_k}-u_{n_k}\|=0 \ \text{ and } \ \lim_{k\to\infty}\|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|=0.$$

From the definition of χ_{n_k}, we obtain

$$\begin{aligned}\|u_{n_k}-z_{n_k}\| &\le \|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|+\theta\chi_{n_k}\|c_{n_k}\|\\ &= \|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|+\theta\frac{\langle u_{n_k}-y_{n_k}, c_{n_k}\rangle}{\|c_{n_k}\|}\\ &\le \|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|+\theta\|u_{n_k}-y_{n_k}\|.\end{aligned}$$

Hence, we obtain that $\lim_{k\to\infty}\|z_{n_k}-u_{n_k}\|=0$. This together with the boundedness of $\{x_n\}$ yields that

$$\lim_{k\to\infty}\varphi_{n_k}\|u_{n_k}-z_{n_k}\|\|x_{n_k+1}-x^\dagger\|=0. \tag{3.24}$$

Moreover, using Remark 3.2 and Assumption (C4), we have

$$\begin{aligned}\|x_{n_k+1}-u_{n_k}\| &= \sigma_{n_k}\|u_{n_k}\|+\varphi_{n_k}\|z_{n_k}-u_{n_k}\|\to 0,\\ \|x_{n_k}-u_{n_k}\| &= \sigma_{n_k}\cdot\frac{\tau_{n_k}}{\sigma_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0.\end{aligned}$$

From the above facts, we conclude that

$$\|x_{n_k+1} - x_{n_k}\| \leq \|x_{n_k+1} - u_{n_k}\| + \|u_{n_k} - x_{n_k}\| \rightarrow 0\,. \tag{3.25}$$

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

$$\limsup_{k\to\infty} \langle x^\dagger, x^\dagger - x_{n_k}\rangle = \lim_{j\to\infty} \langle x^\dagger, x^\dagger - x_{n_{k_j}}\rangle = \langle x^\dagger, x^\dagger - z\rangle\,. \tag{3.26}$$

We obtain that $u_{n_k} \rightharpoonup z$ since $\|x_{n_k} - u_{n_k}\| \to 0$. This together with $\lim_{k\to\infty} \|u_{n_k} - y_{n_k}\| = 0$ and Lemma 3.3 implies that $z \in \mathrm{VI}(\mathcal{C}, M)$. From the definition of $x^\dagger$ and (3.26), we obtain

$$\limsup_{k\to\infty} \langle x^\dagger, x^\dagger - x_{n_k}\rangle = \langle x^\dagger, x^\dagger - z\rangle \leq 0\,. \tag{3.27}$$

Combining (3.25) and (3.27), we find that

$$\limsup_{k\to\infty} \langle x^\dagger, x^\dagger - x_{n_k+1}\rangle \leq \limsup_{k\to\infty} \langle x^\dagger, x^\dagger - x_{n_k}\rangle \leq 0\,. \tag{3.28}$$

Thus, combining Assumption (C4), Remark 3.2, (3.24), (3.28) and Claim 3, in the light of Lemma 2.2, we conclude that $x_n \to x^\dagger$ as $n \to \infty$, which is the desired result. □

3.2 The first viscosity-type inertial subgradient extragradient algorithm

In this subsection, we introduce a viscosity-type inertial projection and contraction subgradient extragradient method. First, we use the following Assumption (C5) to replace the Assumption (C4) described in Sect. 3.

(C5) Let $f : \mathcal{H} \to \mathcal{H}$ be a κ-contraction mapping with $\kappa \in [0, 1)$. Let $\{\epsilon_n\}$ and $\{\xi_n\}$ be two nonnegative sequences such that $\lim_{n\to\infty} \frac{\epsilon_n}{\sigma_n} = 0$ and $\sum_{n=1}^{\infty} \xi_n < +\infty$, where $\{\sigma_n\} \subset (0, 1)$ satisfies $\lim_{n\to\infty} \sigma_n = 0$ and $\sum_{n=1}^{\infty} \sigma_n = \infty$.

The Algorithm 3.2 is of the following form.

Algorithm 3.2 The first viscosity-type inertial subgradient extragradient algorithm

Initialization: Take $\tau > 0$, $\mu \in (0, 1)$, $\vartheta_1 > 0$, $\theta \in (0, 2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n \left(x_n - x_{n-1}\right) , \\ y_n = P_{\mathcal{C}} \left(u_n - \vartheta_n M u_n\right) , \\ z_n = P_{T_n} (u_n - \theta \vartheta_n \chi_n M y_n) , \\ T_n = \{x \in \mathcal{H} \mid \langle u_n - \vartheta_n M u_n - y_n, x - y_n\rangle \leq 0\} , \\ x_{n+1} = \sigma_n f(x_n) + (1 - \sigma_n) z_n , \end{cases}$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively.

Theorem 3.2 *Suppose that Assumptions (C1)–(C3) and (C5) hold. Then the sequence $\{x_n\}$ constructed by Algorithm 3.2 converges to $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ in norm, where $x^\dagger = P_{\mathrm{VI}(\mathcal{C},M)} \circ f(x^\dagger)$.*

Proof As in the proof of Theorem 3.1, we also prove it in four steps.
Claim 1. The sequence $\{x_n\}$ is bounded. Using the definition of x_{n+1} and (3.17), we obtain

$$\begin{aligned}
\|x_{n+1}-x^\dagger\| &= \|\sigma_n(f(x_n)-x^\dagger)+(1-\sigma_n)(z_n-x^\dagger)\| \\
&\le \sigma_n\|f(x_n)-f(x^\dagger)\|+\sigma_n\|f(x^\dagger)-x^\dagger\|+(1-\sigma_n)\|z_n-x^\dagger\| \\
&\le \sigma_n\kappa\|x_n-x^\dagger\|+\sigma_n\|f(x^\dagger)-x^\dagger\|+(1-\sigma_n)\|z_n-x^\dagger\| \\
&\le (1-(1-\kappa)\sigma_n)\|x_n-x^\dagger\|+(1-\kappa)\sigma_n\frac{Q_1+\|f(x^\dagger)-x^\dagger\|}{1-\kappa} \\
&\le \max\left\{\|x_n-x^\dagger\|,\frac{Q_1+\|f(x^\dagger)-x^\dagger\|}{1-\kappa}\right\} \\
&\le \cdots \le \max\left\{\|x_0-x^\dagger\|,\frac{Q_1+\|f(x^\dagger)-x^\dagger\|}{1-\kappa}\right\}.
\end{aligned}$$

This implies that the sequence $\{x_n\}$ is bounded. We also get that the sequences $\{u_n\}$, $\{z_n\}$ and $\{f(x_n)\}$ are bounded.
Claim 2.

$$\begin{aligned}
&(1-\sigma_n)\theta(2-\theta)\frac{\left(1-\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1+\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n-y_n\|^2+(1-\sigma_n)\|u_n-z_n-\theta\chi_n c_n\|^2 \\
&\quad\le \|x_n-x^\dagger\|^2-\|x_{n+1}-x^\dagger\|^2+\sigma_n Q_4
\end{aligned}$$

for some $Q_4>0$. Combining Lemma 3.4 and (3.20), we see that

$$\begin{aligned}
\|x_{n+1}-x^\dagger\|^2 &\le \sigma_n\big(\|f(x_n)-f(x^\dagger)\|+\|f(x^\dagger)-x^\dagger\|\big)^2+(1-\sigma_n)\|z_n-x^\dagger\|^2 \\
&\le \sigma_n\big(\|x_n-x^\dagger\|+\|f(x^\dagger)-x^\dagger\|\big)^2+(1-\sigma_n)\|z_n-x^\dagger\|^2 \\
&= \sigma_n\|x_n-x^\dagger\|^2+(1-\sigma_n)\|z_n-x^\dagger\|^2 \\
&\quad+\sigma_n\big(2\|x_n-x^\dagger\|\cdot\|f(x^\dagger)-x^\dagger\|+\|f(x^\dagger)-x^\dagger\|^2\big) \\
&\le \sigma_n\|x_n-x^\dagger\|^2+(1-\sigma_n)\|z_n-x^\dagger\|^2+\sigma_n Q_3 \\
&\le \|x_n-x^\dagger\|^2-(1-\sigma_n)\|u_n-z_n-\theta\chi_n c_n\|^2 \\
&\quad-(1-\sigma_n)\theta(2-\theta)\frac{\left(1-\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1+\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n-y_n\|^2+\sigma_n Q_4\,,
\end{aligned}$$

where $Q_4:=Q_2+Q_3$. The desired result can be delivered by a simple conversion.
Claim 3.

$$\begin{aligned}
\|x_{n+1}-x^\dagger\|^2 &\le (1-(1-\kappa)\sigma_n)\|x_n-x^\dagger\|^2+(1-\kappa)\sigma_n\cdot\Big[\frac{3Q}{1-\kappa}\cdot\frac{\tau_n}{\sigma_n}\|x_n-x_{n-1}\| \\
&\quad+\frac{2}{1-\kappa}\langle f(x^\dagger)-x^\dagger,x_{n+1}-x^\dagger\rangle\Big].
\end{aligned}$$

Using (3.15) and (3.22), we obtain

$$\begin{aligned}\|x_{n+1}-x^\dagger\|^2 &= \|\sigma_n(f(x_n)-f(x^\dagger))+(1-\sigma_n)(z_n-x^\dagger)+\sigma_n(f(x^\dagger)-x^\dagger)\|^2\\ &\le \|\sigma_n(f(x_n)-f(x^\dagger))+(1-\sigma_n)(z_n-x^\dagger)\|^2+2\sigma_n\langle f(x^\dagger)-x^\dagger, x_{n+1}-x^\dagger\rangle\\ &\le \sigma_n\|f(x_n)-f(x^\dagger)\|^2+(1-\sigma_n)\|z_n-x^\dagger\|^2+2\sigma_n\langle f(x^\dagger)-x^\dagger, x_{n+1}-x^\dagger\rangle\\ &\le \sigma_n\kappa\|x_n-x^\dagger\|^2+(1-\sigma_n)\|u_n-x^\dagger\|^2+2\sigma_n\langle f(x^\dagger)-x^\dagger, x_{n+1}-x^\dagger\rangle\\ &\le (1-(1-\kappa)\sigma_n)\|x_n-x^\dagger\|^2+(1-\kappa)\sigma_n\cdot\Big[\frac{3Q}{1-\kappa}\cdot\frac{\tau_n}{\sigma_n}\|x_n-x_{n-1}\|\\ &\quad+\frac{2}{1-\kappa}\langle f(x^\dagger)-x^\dagger, x_{n+1}-x^\dagger\rangle\Big].\end{aligned}$$

Claim 4. The sequence $\{\|x_n-x^\dagger\|^2\}$ converges to zero. By Claim 2 and Assumption (C5), we observe that

$$\begin{aligned}&(1-\sigma_{n_k})\theta(2-\theta)\frac{\left(1-\frac{\mu\vartheta_{n_k}}{\vartheta_{n_k+1}}\right)^2}{\left(1+\frac{\mu\vartheta_{n_k}}{\vartheta_{n_k+1}}\right)^2}\|u_{n_k}-y_{n_k}\|^2+(1-\sigma_{n_k})\|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|^2\\ &\le \limsup_{k\to\infty}\left[\|x_{n_k}-x^\dagger\|^2-\|x_{n_k+1}-x^\dagger\|^2+\sigma_{n_k}Q_4\right]\le 0\,,\end{aligned}$$

which implies that

$$\lim_{k\to\infty}\|y_{n_k}-u_{n_k}\|=0 \ \text{ and } \lim_{k\to\infty}\|u_{n_k}-z_{n_k}-\theta\chi_{n_k}c_{n_k}\|=0\,.$$

As stated in Claim 4 of Theorem 3.1, it is easy to see that $\lim_{k\to\infty}\|z_{n_k}-u_{n_k}\|=0$. Moreover, using Remark 3.2 and Assumption (C5), we have

$$\begin{aligned}\|x_{n_k+1}-z_{n_k}\| &= \sigma_{n_k}\|z_{n_k}-f(x_{n_k})\|\to 0\,,\\ \|x_{n_k}-u_{n_k}\| &= \sigma_{n_k}\cdot\frac{\tau_{n_k}}{\sigma_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0\,.\end{aligned}$$

It follows that

$$\|x_{n_k+1}-x_{n_k}\|\le\|x_{n_k+1}-z_{n_k}\|+\|z_{n_k}-u_{n_k}\|+\|u_{n_k}-x_{n_k}\|\to 0\,. \tag{3.29}$$

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

$$\limsup_{k\to\infty}\langle f(x^\dagger)-x^\dagger, x_{n_k}-x^\dagger\rangle=\lim_{j\to\infty}\langle f(x^\dagger)-x^\dagger, x_{n_{k_j}}-x^\dagger\rangle=\langle f(x^\dagger)-x^\dagger, z-x^\dagger\rangle\,. \tag{3.30}$$

We obtain that $u_{n_k}\rightharpoonup z$ since $\|x_{n_k}-u_{n_k}\|\to 0$. This, together with $\lim_{k\to\infty}\|u_{n_k}-y_{n_k}\|=0$ and Lemma 3.3, obtains that $z\in \mathrm{VI}(\mathcal{C}, M)$. From the definition of $x^\dagger$ and (3.30), we obtain

$$\limsup_{k\to\infty}\langle f(x^\dagger)-x^\dagger, x_{n_k}-x^\dagger\rangle=\langle f(x^\dagger)-x^\dagger, z-x^\dagger\rangle\le 0\,. \tag{3.31}$$

Combining (3.29) and (3.31), we obtain

$$\limsup_{k\to\infty}\langle f(x^\dagger)-x^\dagger, x_{n_k+1}-x^\dagger\rangle \le \limsup_{k\to\infty}\langle f(x^\dagger)-x^\dagger, x_{n_k}-x^\dagger\rangle \le 0\,. \tag{3.32}$$

Thus, from Assumption (C5), Remark 3.2, (3.32) and Claim 3, in the light of Lemma 2.2, we conclude that $x_n \to x^\dagger$ as $n \to \infty$. The proof of the Theorem 3.2 is now complete. □

3.3 The second viscosity-type inertial subgradient extragradient algorithm

In this subsection, we introduce another viscosity-type iterative scheme that is different from Algorithm 3.2. The details of this scheme are described in Algorithm 3.3.

Algorithm 3.3 The second viscosity-type inertial subgradient extragradient algorithm

Initialization: Take $\tau > 0$, $\mu \in (0,1)$, $\vartheta_1 > 0$, $\theta \in (0,2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n\left(x_n - x_{n-1}\right)\,,\\ y_n = P_{\mathcal{C}}\left(u_n - \vartheta_n M u_n\right)\,,\\ z_n = P_{T_n}\left(u_n - \theta\vartheta_n\chi_n M y_n\right),\\ T_n = \{x \in \mathcal{H} \mid \langle u_n - \vartheta_n M u_n - y_n, x - y_n\rangle \le 0\}\,,\\ x_{n+1} = \sigma_n f\left(z_n\right) + \left(1-\sigma_n\right) z_n\,, \end{cases}$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively.

Theorem 3.3 *Suppose that Assumptions (C1)–(C3) and (C5) hold. Then the sequence* $\{x_n\}$ *collected by Algorithm* 3.3 *converges to* $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ *in norm, where* $x^\dagger = P_{\mathrm{VI}(\mathcal{C},M)} \circ f(x^\dagger)$.

Proof The proof of this theorem is very similar to Theorem 3.2. We also separate it into four steps.
Claim 1. The sequence $\{x_n\}$ is bounded. Using the definition of x_{n+1} and (3.17), we obtain

$$\begin{aligned} \|x_{n+1}-x^\dagger\| &\le \sigma_n\|f(z_n)-f(x^\dagger)\| + \sigma_n\|f(x^\dagger)-x^\dagger\| + (1-\sigma_n)\|z_n-x^\dagger\| \\ &\le (1-(1-\kappa)\sigma_n)\|x_n-x^\dagger\| + (1-\kappa)\sigma_n\frac{Q_1+\|f(x^\dagger)-x^\dagger\|}{1-\kappa} \\ &\le \max\left\{\|x_0-x^\dagger\|, \frac{Q_1+\|f(x^\dagger)-x^\dagger\|}{1-\kappa}\right\}. \end{aligned}$$

This indicates that the sequence $\{x_n\}$ is bounded. We also get that the sequences $\{u_n\}$, $\{z_n\}$ and $\{f(z_n)\}$ are bounded.
Claim 2.

$$\begin{aligned} &\theta(2-\theta)\frac{\left(1-\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1+\frac{\mu\vartheta_n}{\vartheta_{n+1}}\right)^2}\|u_n-y_n\|^2 + \|u_n-z_n-\theta\chi_n c_n\|^2 \\ &\le \|x_n-x^\dagger\|^2 - \|x_{n+1}-x^\dagger\|^2 + \sigma_n Q_6 \end{aligned}$$

for some $Q_6 > 0$. Combining Lemma 3.4 and (3.20), we find that

$$\begin{aligned}
\|x_{n+1} - x^\dagger\|^2 &\le \sigma_n(\|z_n - x^\dagger\| + \|f(x^\dagger) - x^\dagger\|)^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 \\
&= \sigma_n\|z_n - x^\dagger\|^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 \\
&\quad + \sigma_n\big(2\|z_n - x^\dagger\| \cdot \|f(x^\dagger) - x^\dagger\| + \|f(x^\dagger) - x^\dagger\|^2\big) \\
&\le \|z_n - x^\dagger\|^2 + \sigma_n Q_5 \\
&\le \|x_n - x^\dagger\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2 \\
&\quad - \theta(2-\theta)\frac{\big(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}}\big)^2}{\big(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\big)^2}\|u_n - y_n\|^2 + \sigma_n Q_6\,,
\end{aligned}$$

where $Q_6 := Q_2 + Q_5$. The desired result can be delivered by a simple conversion.
Claim 3.

$$\begin{aligned}
\|x_{n+1} - x^\dagger\|^2 \le (1-(1-\kappa)\sigma_n)\|x_n - x^\dagger\|^2 + (1-\kappa)\sigma_n \cdot \Big[\frac{3Q}{1-\kappa}\cdot\frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| \\
+ \frac{2}{1-\kappa}\langle f(x^\dagger) - x^\dagger, x_{n+1} - x^\dagger\rangle\Big].
\end{aligned}$$

Using (3.15) and (3.22), we have

$$\begin{aligned}
\|x_{n+1} - x^\dagger\|^2 &= \|\sigma_n(f(z_n) - f(x^\dagger)) + (1-\sigma_n)(z_n - x^\dagger) + \sigma_n(f(x^\dagger) - x^\dagger)\|^2 \\
&\le \sigma_n\kappa\|z_n - x^\dagger\|^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 + 2\sigma_n\langle f(x^\dagger) - x^\dagger, x_{n+1} - x^\dagger\rangle \\
&\le (1-(1-\kappa)\sigma_n)\|x_n - x^\dagger\|^2 + (1-\kappa)\sigma_n \cdot \Big[\frac{3Q}{1-\kappa}\cdot\frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| \\
&\quad + \frac{2}{1-\kappa}\langle f(x^\dagger) - x^\dagger, x_{n+1} - x^\dagger\rangle\Big].
\end{aligned}$$

Claim 4. The sequence $\{\|x_n - x^\dagger\|^2\}$ converges to zero. This conclusion can be easily obtained by inferences similar to Claim 4 of Theorem 3.2. This completes the proof. □

In the next part, we will introduce three new simple numerical methods that only need to calculate the projection once in each iteration.

3.4 The modified Mann-type inertial projection and contraction algorithm

Our first modified iterative process is stated in Algorithm 3.4. Compared with Algorithm 3.1, the calculation of the iterative sequence $\{z_n\}$ replaces the projection on the half-space with a display formula.

The following lemma plays an important role in studying the convergence of the algorithms.

Lemma 3.5 *Suppose that Assumption (C3) holds. Let $\{z_n\}$ and $\{u_n\}$ be two sequences produced by Algorithm 3.4. Then, for all $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$,*

$$\|z_n - x^\dagger\|^2 \le \|u_n - x^\dagger\|^2 - \frac{2-\theta}{\theta}\|u_n - z_n\|^2\,,$$

and

$$\|u_n - y_n\|^2 \le \frac{\big(1 + \frac{\mu\vartheta_n}{\vartheta_{n+1}}\big)^2}{\big[(1 - \frac{\mu\vartheta_n}{\vartheta_{n+1}})\theta\big]^2}\|u_n - z_n\|^2\,.$$

Algorithm 3.4 The modified Mann-type inertial projection and contraction algorithm

Initialization: Take $\tau > 0$, $\mu \in (0,1)$, $\vartheta_1 > 0$, $\theta \in (0,2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n \left(x_n - x_{n-1}\right) , \\ y_n = P_{\mathcal{C}} \left(u_n - \vartheta_n M u_n\right) , \\ z_n = u_n - \theta \chi_n c_n , \\ x_{n+1} = (1 - \sigma_n - \varphi_n) u_n + \varphi_n z_n , \end{cases}$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively.

Proof By using the definition of z_n, one obtains

$$\begin{aligned} \|z_n - x^\dagger\|^2 &= \|u_n - \theta \chi_n c_n - x^\dagger\|^2 \\ &= \|u_n - x^\dagger\|^2 - 2\theta \chi_n \langle u_n - x^\dagger, c_n\rangle + \theta^2 \chi_n^2 \|c_n\|^2 . \end{aligned} \tag{3.33}$$

According to the definition of c_n, one sees that

$$\begin{aligned} \langle u_n - x^\dagger, c_n\rangle &= \langle u_n - y_n, c_n\rangle + \langle y_n - x^\dagger, c_n\rangle \\ &= \langle u_n - y_n, c_n\rangle + \langle y_n - x^\dagger, u_n - y_n - \vartheta_n (M u_n - M y_n)\rangle . \end{aligned} \tag{3.34}$$

From $y_n = P_{\mathcal{C}}(u_n - \vartheta_n M u_n)$ and the property of projection, we have

$$\langle u_n - y_n - \vartheta_n M u_n, y_n - x^\dagger\rangle \geq 0 . \tag{3.35}$$

Using $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ and $y_n \in \mathcal{C}$, we obtain that $\langle M x^\dagger, y_n - x^\dagger\rangle \geq 0$, which combined with the pseudomonotonicity of M yields that

$$\vartheta_n \langle M y_n, y_n - x^\dagger\rangle \geq 0 . \tag{3.36}$$

By using (3.34), (3.35) and (3.36), we obtain

$$\langle u_n - x^\dagger, c_n\rangle \geq \langle u_n - y_n, c_n\rangle . \tag{3.37}$$

It follows from the definition of z_n that $z_n - u_n = \theta \chi_n c_n$. From the definition of χ_n, one gets $\langle u_n - y_n, c_n\rangle = \chi_n \|c_n\|^2$. Combining (3.33) and (3.37), we conclude that

$$\begin{aligned} \|z_n - x^\dagger\|^2 &\leq \|u_n - x^\dagger\|^2 - 2\theta \chi_n \langle u_n - y_n, c_n\rangle + \theta^2 \chi_n^2 \|c_n\|^2 \\ &= \|u_n - x^\dagger\|^2 - 2\theta \chi_n^2 \|c_n\|^2 + \theta^2 \chi_n^2 \|c_n\|^2 \\ &= \|u_n - x^\dagger\|^2 - \frac{2-\theta}{\theta} \|\theta \chi_n c_n\|^2 \\ &= \|u_n - x^\dagger\|^2 - \frac{2-\theta}{\theta} \|u_n - z_n\|^2 . \end{aligned}$$

On the other hand, by the definition of z_n and (3.14), we have

$$\|z_n - u_n\|^2 = \theta^2 \chi_n^2 \|c_n\|^2 \geq \theta^2 \frac{\left(1 - \frac{\mu \vartheta_n}{\vartheta_{n+1}}\right)^2}{\left(1 + \frac{\mu \vartheta_n}{\vartheta_{n+1}}\right)^2} \|u_n - y_n\|^2 .$$

Thus, we obtain

$$\|u_n - y_n\|^2 \leq \frac{\left(1 + \frac{\mu \vartheta_n}{\vartheta_{n+1}}\right)^2}{\left[(1 - \frac{\mu \vartheta_n}{\vartheta_{n+1}})\theta\right]^2} \|u_n - z_n\|^2 .$$

The proof of the lemma is now complete. □

Theorem 3.4 *Suppose that Assumptions (C1)–(C4) hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.4 converges to $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ in norm, where $\|x^\dagger\| = \min\{\|z\| : z \in \mathrm{VI}(\mathcal{C}, M)\}$.*

Proof As we did before, we state the proof into four steps.
Claim 1. The sequence $\{x_n\}$ is bounded. Indeed, thanks to Lemma 3.5 and $\theta \in (0, 2)$, we have

$$\|z_n - x^\dagger\| \leq \|u_n - x^\dagger\|, \quad \forall n \geq 1\,. \tag{3.38}$$

Using the same facts as stated in Claim 1 of Theorem 3.1, we obtain that the sequences $\{x_n\}$, $\{u_n\}$, $\{y_n\}$ and $\{z_n\}$are bounded.
Claim 2.

$$\varphi_n \frac{2-\theta}{\theta}\|u_n - z_n\|^2 \leq \|x_n - x^\dagger\|^2 - \|x_{n+1} - x^\dagger\|^2 + \sigma_n\big(\|x^\dagger\|^2 + Q_2\big)\,.$$

By using Lemma 3.5, (3.20) and (3.21), we obtain

$$\begin{aligned}
\|x_{n+1} - x^\dagger\|^2 &\leq (1-\sigma_n-\varphi_n)\|u_n - x^\dagger\|^2 + \varphi_n\|z_n - x^\dagger\|^2 + \sigma_n\|x^\dagger\|^2 \\
&\leq (1-\sigma_n-\varphi_n)\|u_n - x^\dagger\|^2 + \varphi_n\|u_n - x^\dagger\|^2 - \varphi_n\frac{2-\theta}{\theta}\|u_n - z_n\|^2 + \sigma_n\|x^\dagger\|^2 \\
&\leq \|x_n - x^\dagger\|^2 - \varphi_n\frac{2-\theta}{\theta}\|u_n - z_n\|^2 + \sigma_n\big(\|x^\dagger\|^2 + Q_2\big)\,.
\end{aligned}$$

The desired result can be achieved by a simple conversion.
Claim 3.

$$\begin{aligned}
\|x_{n+1} - x^\dagger\|^2 \leq (1-\sigma_n)\|x_n - x^\dagger\|^2 + \sigma_n\Big[&2\varphi_n\|u_n - z_n\|\|x_{n+1} - x^\dagger\| \\
&+ 2\langle x^\dagger, x^\dagger - x_{n+1}\rangle + \frac{3Q\tau_n}{\sigma_n}\|x_n - x_{n-1}\|\Big]\,.
\end{aligned}$$

The result can be obtained by using the same facts as declared in Claim 3 of Theorem 3.1.
Claim 4. The sequence $\{\|x_n - x^\dagger\|^2\}$ converges to zero. By Claim 2 and Assumption (C4), we have

$$\begin{aligned}
\varphi_{n_k}\frac{2-\theta}{\theta}\|u_{n_k} - z_{n_k}\|^2 &\leq \limsup_{k\to\infty}\left[\|x_{n_k} - x^\dagger\|^2 - \|x_{n_k+1} - x^\dagger\|^2 + \sigma_{n_k}(\|x^\dagger\|^2 + Q_2)\right] \\
&\leq 0\,,
\end{aligned}$$

which implies that $\lim_{k\to\infty}\|z_{n_k} - u_{n_k}\| = 0$. In view of Lemma 3.5, we observe that $\lim_{k\to\infty}\|y_{n_k} - u_{n_k}\| = 0$. As asserted in Claim 4 of Theorem 3.1, we can obtain the same result as (3.24)–(3.28). Therefore, we obtain that $x_n \to x^\dagger$ as $n \to \infty$. This completes the proof. □

3.5 The first modified viscosity-type inertial projection and contraction algorithm

By replacing the calculation process of the iterative sequence $\{z_n\}$ in Algorithm 3.2, we obtain the following Algorithm 3.5.

Theorem 3.5 *Suppose that Assumptions (C1)–(C3) and (C5) hold. Then the sequence $\{x_n\}$ designed by Algorithm 3.5 converges to $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ in norm, where $x^\dagger = P_{\mathrm{VI}(\mathcal{C},M)} \circ f(x^\dagger)$.*

Algorithm 3.5 The first modified viscosity-type inertial projection and contraction algorithm

Initialization: Take $\tau > 0$, $\mu \in (0, 1)$, $\vartheta_1 > 0$, $\theta \in (0, 2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n \left(x_n - x_{n-1}\right) , \\ y_n = P_{\mathcal{C}} \left(u_n - \vartheta_n M u_n\right) , \\ z_n = u_n - \theta \chi_n c_n , \\ x_{n+1} = \sigma_n f\left(x_n\right) + \left(1 - \sigma_n\right) z_n , \end{cases}$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively.

Proof The proof of this theorem is very similar to Theorem 3.2. We divide it into four steps.
Claim 1. The sequence $\{x_n\}$ is bounded. Using the same arguments as declared in Claim 1 of Theorem 3.2, we obtain that the sequences $\{x_n\}$, $\{u_n\}$, $\{y_n\}$, $\{z_n\}$ and $\{f(x_n)\}$ are bounded.
Claim 2.

$$(1 - \sigma_n)\frac{2-\theta}{\theta}\|u_n - z_n\|^2 \leq \|x_n - x^\dagger\|^2 - \|x_{n+1} - x^\dagger\|^2 + \sigma_n Q_4 .$$

In view of Lemma 3.5 and (3.20), we have

$$\begin{aligned} \|x_{n+1} - x^\dagger\|^2 &= \|\sigma_n(f(x_n) - x^\dagger) + (1-\sigma_n)(z_n - x^\dagger)\|^2 \\ &\leq \sigma_n\|f(x_n) - x^\dagger\|^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 \\ &= \sigma_n\big(\|f(x_n) - f(x^\dagger)\| + \|f(x^\dagger) - x^\dagger\|\big)^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 \\ &\leq \sigma_n\|x_n - x^\dagger\|^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 \\ &\quad + \sigma_n\big(2\|x_n - x^\dagger\| \cdot \|f(x^\dagger) - x^\dagger\| + \|f(x^\dagger) - x^\dagger\|^2\big) \\ &\leq \sigma_n\|x_n - x^\dagger\|^2 + (1-\sigma_n)\|z_n - x^\dagger\|^2 + \sigma_n Q_3 \\ &\leq \|x_n - x^\dagger\|^2 - (1-\sigma_n)\frac{2-\theta}{\theta}\|u_n - z_n\|^2 + \sigma_n Q_4 , \end{aligned}$$

where $Q_4 := Q_2 + Q_3$. The desired result can be gained through a simple conversion.
Claim 3.

$$\begin{aligned} \|x_{n+1} - x^\dagger\|^2 \leq (1 - (1-\kappa)\sigma_n)\|x_n - x^\dagger\|^2 + (1-\kappa)\sigma_n \cdot \Big[\frac{3Q}{1-\kappa} \cdot \frac{\tau_n}{\sigma_n}\|x_n - x_{n-1}\| \\ + \frac{2}{1-\kappa}\langle f(x^\dagger) - x^\dagger, x_{n+1} - x^\dagger\rangle\Big]. \end{aligned}$$

The result can be achieved by using the same facts as stated in Claim 3 of Theorem 3.2.
Claim 4. The sequence $\{\|x_n - x^\dagger\|^2\}$ converges to zero. By Claim 2 Assumption (C5), one has

$$(1 - \sigma_{n_k})\frac{2-\theta}{\theta}\|u_{n_k} - z_{n_k}\|^2 \leq 0 ,$$

which indicates that $\lim_{k\to\infty}\|z_{n_k} - u_{n_k}\| = 0$. This together with Lemma 3.5 finds that $\lim_{k\to\infty}\|y_{n_k} - u_{n_k}\| = 0$. As stated in Claim 4 of Theorem 3.2, we can get the same facts as (3.29)–(3.32). Therefore, we obtain $x_n \to x^\dagger$ as $n \to \infty$. The proof is completed. □

3.6 The second modified viscosity-type inertial projection and contraction algorithm

Our last iteration scheme is depicted in Algorithm 3.6.

Algorithm 3.6 The second modified viscosity-type inertial projection and contraction algorithm

Initialization: Take $\tau > 0$, $\mu \in (0, 1)$, $\vartheta_1 > 0$, $\theta \in (0, 2)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.
Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n \left(x_n - x_{n-1}\right) , \\ y_n = P_{\mathcal{C}} \left(u_n - \vartheta_n M u_n\right) , \\ z_n = u_n - \theta \chi_n c_n , \\ x_{n+1} = \sigma_n f\left(z_n\right) + \left(1 - \sigma_n\right) z_n , \end{cases}$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively.

Theorem 3.6 *Suppose that Assumptions (C1)–(C3) and (C5) hold. Then the sequence* $\{x_n\}$ *determined by Algorithm 3.6 converges to* $x^\dagger \in \mathrm{VI}(\mathcal{C}, M)$ *in norm, where* $x^\dagger = P_{\mathrm{VI}(\mathcal{C},M)} \circ f(x^\dagger)$.

Proof Combining the proofs of Theorems 3.3 and 3.5, we can easily get the desired conclusion. This part is left to the reader to verify. □

Remark 3.4 As shown in Table 1, we compare the suggested iterative schemes with some known algorithms in the literature.

The six algorithms obtained in this paper directly improve some recent results in [7,11,13,15,28,31,32,36–38,42] based on the following observations:

(i) We investigated and confirmed the strong convergence of the proposed algorithms, while the algorithms suggested by Dong et al. [11,13], Shehu and Iyiola [32] and

Table 1 Compare the proposed iterative schemes with some known algorithms

Algorithms	Operator	Inertial	Stepsize	Convergence
Our Algorithms 3.1–3.6	Pseudomonotone	Yes	Non-monotonic	Strong
[7, Algorithm 3.1]	Pseudomonotone	Yes	Fixed	Strong
[11, Algorithm 3.1]	Monotone	No	Armijo-like	Weak
[13, Algorithm 3.1]	Monotone	Yes	Fixed	Weak
[15, Algorithms 3.1 and 3.2]	Monotone	No	Armijo-like	Strong
[28, Algorithm 3.1]	Monotone	No	Armijo-like	Strong
[31, Algorithm 4.3]	Pseudomonotone	No	Fixed	Strong
[32, Algorithm 2]	Pseudomonotone	Yes	Self-adaptive	Weak
[36, Algorithms 3.1 and 3.2]	Monotone	No	Self-adaptive	Strong
[37, Algorithms 3.1 and 3.2]	Monotone	No	Armijo-like	Strong
[38, Algorithm 1]	Monotone	Yes	Fixed	Strong
[42, Algorithm 3.1]	Pseudomonotone	Yes	Self-adaptive	Weak

Yang [42] converge weakly to a solution of VIP in infinite-dimensional Hilbert spaces. Many problems in applied sciences indicate that strong convergence is better than weak convergence in an infinite-dimensional space.

(ii) The algorithms proposed by Cholamjiak et al. [7], Dong et al. [13], Shehu et al. [31] and Thong et al. [38] use a fixed step size in each iteration, which indicates that the Lipschitz constant of the cost mapping must need to be received in advance. In practical large-scale nonlinear optimization problems, the Lipschitz constant is not easy to obtain or requires more calculation to estimate. On the other hand, the algorithms suggested in [11,15,28,37] require more execution time because they use an Armijo-like step size criteria. Furthermore, the algorithms proposed in [32,36,42] use a non-increasing step size and may depend on the choice of initial step size, which will affect the efficiency of the algorithms. However, our algorithms use a new non-monotonic step size criteria to automatically update the iteration step size, which makes them more intelligent in applications.

(iii) The inertial method is used by many scholars as an acceleration technique. Many works of literature show that this technique can speed up the convergence speed of the methods used. The iterative schemes devised in this paper combined inertial terms, which also accelerate the convergence speed of the algorithms without inertial terms in [15,37] (see the numerical experiments in Sects. 5 and 6).

(iv) The algorithms presented in [11,13,15,28,36–38] will not be available when the mapping is not monotone. It is known that there are some mappings that are not monotone, such as pseudomonotone mappings. Therefore, the algorithms proposed in this study to solve pseudomonotone VIP are more useful. In addition, the algorithm proposed by Shehu et al. [31] combines the Halpern-type method to ensure strong convergence, which makes its convergence speed very slow, while our suggested algorithms have a faster convergence speed because they use the viscosity-type and the Mann-type method.

As mentioned above, the approaches established in this paper have competitive advantages over some known results in the literature and are more desirable in practical applications.

4 Numerical experiments

In this section, we provide some numerical experiments to illustrate the effectiveness of our proposed iterative schemes and compare them with some known strong convergent algorithms, including the Algorithm 3.1 suggested by Shehu and Iyiola [28] (shortly, SI Alg. 3.1), the Algorithms 3.1 and 3.2 introduced by Thong and Gibali [37] (shortly, TG Algs. 3.1 and 3.2), the Algorithms 3.1 and 3.2 proposed by Gibali et al. [15] (shortly, GTT Algs. 3.1 and 3.2) and the Algorithms 4.3 presented by Shehu et al. [31] (shortly, SDJ Alg. 4.3). We use the FOM Solver [3] to effectively calculate the projections onto $\mathcal{C}$ and T_n. All the programs are implemented in MATLAB 2018a on a personal computer.

Our parameters are set as follows. For all algorithms, we set $\sigma_n = 1/(n+1)$, $\varphi_n = 0.5(1-\sigma_n)$ and $f(x) = 0.1x$. For our proposed algorithms, we take $\mu = 0.4$, $\vartheta_1 = 0.5$, $\xi_n = 1/(n+1)^{1.1}$ and $\theta = 1.5$. In [37, Algorithms 3.1 and 3.2] and [15, Algorithms 3.1 and 3.2], we pick $\delta = \zeta = 0.5$, $\phi = 0.4$ and $\theta = 1.5$. Adopt inertial parameters $\tau = 0.4$ and $\epsilon_n = 100/(n+1)^2$ in our algorithms. For [28, Algorithm 3.1], we choose $\zeta = 0.5$ and $\phi = 0.4$. Take fixed step size $\vartheta_n = 0.5/L$ and $\theta = 1.5$ in Shehu et al.'s [31, Algorithms

Table 2 Numerical results for Example 4.1

Algorithms	$\varsigma_n = 10^{-2}$		$\varsigma_n = 10^{-3}$		$\varsigma_n = 10^{-4}$		$\varsigma_n = 10^{-5}$	
	Iter.	CPU	Iter.	CPU	Iter.	CPU	Iter.	CPU
Our Alg. 3.1	69	0.0034	130	0.0071	413	0.0227	1309	0.0842
Our Alg. 3.2	16	0.0010	69	0.0033	237	0.0131	629	0.0388
Our Alg. 3.3	35	0.0025	69	0.0034	237	0.0130	629	0.0490
Our Alg. 3.4	58	0.0026	137	0.0064	433	0.0257	1373	0.0784
Our Alg. 3.5	14	0.0007	116	0.0053	272	0.0126	712	0.0450
Our Alg. 3.6	46	0.0021	117	0.0052	271	0.0126	711	0.0519
TG Alg. 3.1	47	0.0099	290	0.0590	572	0.1216	1808	0.4421
TG Alg. 3.2	22	0.0047	155	0.0310	256	0.0547	811	0.2201
GTT Alg. 3.1	48	0.0105	309	0.0715	590	0.1266	1870	0.4805
GTT Alg. 3.2	24	0.0050	134	0.0331	266	0.0651	842	0.2077
SI Alg. 3.1	24	0.0059	71	0.0358	271	0.0809	854	0.2541

4.3]. We use $D_n = \|x_n - x^*\|$ to measure the n-th iteration error of all algorithms, where x^* represents the solution to our problems.

Example 4.1 Our first test example is the nonlinear complementarity problem (NCP) considered by many researchers. Recall that the NCP is described as follows:

$$\text{find } x^* \in \mathcal{C} \text{ such that } x^* \geq 0,\ \ Mx^* \geq 0 \text{ and } \langle x^*, Mx^* \rangle = 0\,.$$

In fact, NCP is a special case when the constraint of the (VIP) is non-negative, that is, the feasible set of NCP is $\mathcal{C} = \mathbb{R}^n_+$. Assume that the mapping $M : \mathbb{R}^4 \to \mathbb{R}^4$ is given by

$$Mx = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_1 + x_2^2 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix}.$$

The feasible set $\mathcal{C}$ is defined by $\mathcal{C} = \{x \in \mathbb{R}^4_+ \mid x_1 + x_2 + x_3 + x_4 = 4\}$. Since the Lipschitz constant of the mapping M is unknown, Shehu et al. [31, Algorithms 4.3] will not participate in the comparison of this example. Moreover, we do not know the exact solution of the problem, so we use $E_n = \|u_n - y_n\|^2$ to measure the error of the n-th iteration and $E_n < \varsigma_n$ is used as the stopping criterion. According to Lemma 3.2, we obtain that $y_n \in \mathrm{VI}(\mathcal{C}, M)$ when $E_n = 0$. The initial values $x_0, x_1 \in \mathbb{R}^4_+$ are randomly generated by MATLAB. Table 2 shows the number of iterations and CPU times in seconds required by all algorithms under different stopping criteria. An approximate solution to this problem is $x^* = (1.2404, 0, 0, 2.7553)^\mathsf{T}$ by using the proposed algorithms.

Example 4.2 Consider the form of linear operator $M : \mathbb{R}^m \to \mathbb{R}^m$ ($m = 10, 30, 60, 100$) as follows: $M(x) = Gx + g$, 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 $\mathcal{C}$ is a box constraint with the form $\mathcal{C} = [-2, 5]^m$. It is easy to see that M 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 $g = \mathbf{0}$. Then

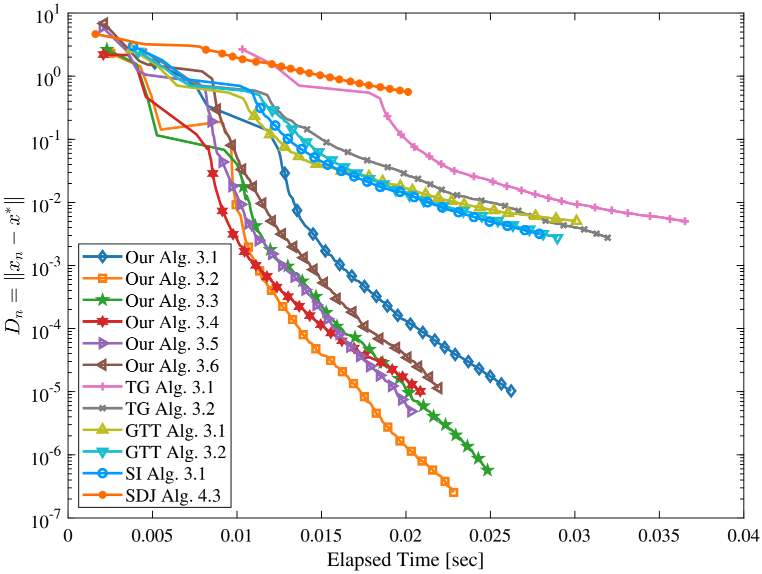


Fig. 1 Example 4.2 for $m = 10$

the solution set is $x^* = \{\mathbf{0}\}$. The maximum number of iterations 200 as a common stopping criterion and the initial values $x_0 = x_1$ are randomly generated by *5rand*$(m, 1)$ in MATLAB. The numerical results are shown in Fig. 1 and Table 3, where "CPU" in Table 3 indicates the execution time in seconds for all algorithms.

Example 4.3 We consider an example in the Hilbert space $\mathcal{H} = L^2([0, 1])$ associated with the inner product $\langle x, y \rangle := \int_0^1 x(t)y(t)\,\mathrm{d}t$ and the induced norm $\|x\| := (\int_0^1 |x(t)|^2\,\mathrm{d}t)^{1/2}$, $\forall x, y \in \mathcal{H}$. Let the feasible set be the unit ball $\mathcal{C} := \{x \in \mathcal{H} : \|x\| \leq 1\}$. Define an operator $M : \mathcal{C} \to \mathcal{H}$ by

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

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 M is monotone and L-Lipschitz continuous with $L = 2$ and $x^*(t) = \{\mathbf{0}\}$ is the solution of the corresponding variational inequality problem. Note that the projection on $\mathcal{C}$ is inherently explicit, that is,

$$P_{\mathcal{C}}(x) = \begin{cases} \frac{x}{\|x\|}, & \text{if } \|x\| > 1\,; \\ x, & \text{if } \|x\| \leq 1\,. \end{cases}$$

We choose the maximum number of iterations 50 as the common stopping criterion. The numerical behaviors of all the algorithms with four starting points $x_0(t) = x_1(t)$ are reported in Fig. 2 and Table 4, where "CPU" in Table 4 represents the execution time in seconds for all algorithms.

Table 3 Numerical results for Example 4.2

Algorithms	$m = 10$		$m = 30$		$m = 60$		$m = 100$	
	D_n	CPU	D_n	CPU	D_n	CPU	D_n	CPU
Our Alg. 3.1	1.03E−05	0.0262	1.17E−04	0.0245	4.75E−04	0.0241	6.52E−04	0.0308
Our Alg. 3.2	2.54E−07	0.0228	1.50E−05	0.0216	6.46E−05	0.0216	9.26E−05	0.0314
Our Alg. 3.3	5.69E−07	0.0248	3.97E−05	0.0235	2.77E−04	0.0237	2.88E−04	0.0331
Our Alg. 3.4	1.02E−05	0.0208	1.17E−04	0.0197	4.76E−04	0.0199	5.20E−04	0.0273
Our Alg. 3.5	4.89E−06	0.0203	9.06E−04	0.0191	4.56E−03	0.0198	6.35E−03	0.0240
Our Alg. 3.6	1.14E−05	0.0219	1.10E−03	0.0195	6.39E−03	0.0204	6.49E−03	0.0277
TG Alg. 3.1	5.01E−03	0.0365	1.55E−02	0.0419	4.32E−02	0.0416	5.86E−02	0.0877
TG Alg. 3.2	2.75E−03	0.0319	1.79E−02	0.0336	6.29E−02	0.0459	8.64E−02	0.0755
GTT Alg. 3.1	5.01E−03	0.0301	1.55E−02	0.0296	4.32E−02	0.0345	5.86E−02	0.0687
GTT Alg. 3.2	2.75E−03	0.0290	1.79E−02	0.0293	6.29E−02	0.0380	8.64E−02	0.0686
SI Alg. 3.1	3.18E−03	0.0279	1.89E−02	0.0326	6.45E−02	0.0423	8.82E−02	0.0749
SDJ Alg. 4.3	5.62E−01	0.0201	2.65E+00	0.0219	7.73E+00	0.0208	1.06E+01	0.0246

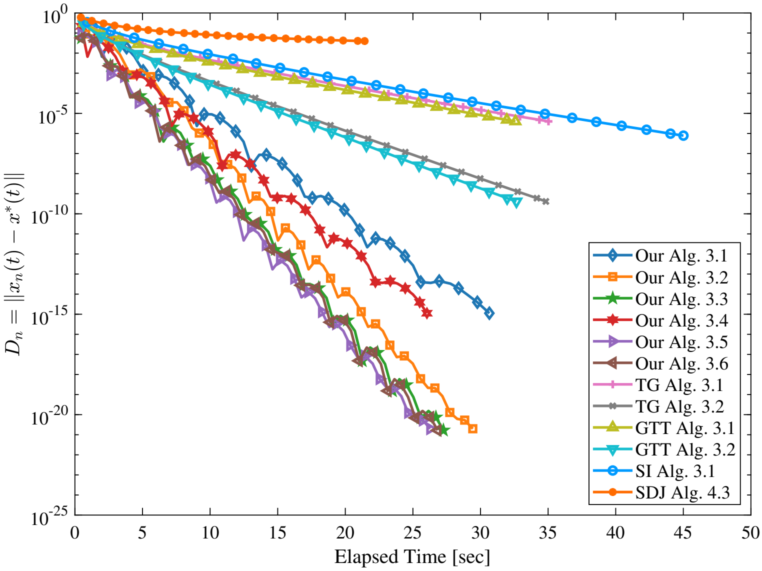


Fig. 2 Example 4.3 for starting points $x_1(t) = \sin(2t)$

Table 4 Numerical results for Example 4.3

Algorithms	$x_1(t) = t^2$		$x_1(t) = e^t$		$x_1(t) = \sin(2t)$		$x_1(t) = \log(t)$	
	D_n	CPU	D_n	CPU	D_n	CPU	D_n	CPU
Our Alg. 3.1	6.61E−16	29.15	2.08E−15	34.91	1.12E−15	30.65	2.65E−15	29.32
Our Alg. 3.2	1.12E−21	27.53	1.38E−18	34.20	1.99E−21	29.43	3.36E−21	28.53
Our Alg. 3.3	9.36E−22	27.27	2.70E−18	32.51	1.67E−21	27.26	3.49E−21	28.55
Our Alg. 3.4	6.61E−16	26.22	2.08E−15	32.27	1.12E−15	26.04	2.65E−15	27.61
Our Alg. 3.5	1.12E−21	26.34	1.41E−18	31.41	1.99E−21	26.27	3.36E−21	27.55
Our Alg. 3.6	9.36E−22	26.50	1.84E−18	31.63	1.67E−21	26.83	3.49E−21	27.85
TG Alg. 3.1	2.35E−06	35.85	7.84E−06	45.93	3.99E−06	35.03	8.06E−06	38.11
TG Alg. 3.2	2.45E−10	35.14	6.96E−10	44.46	4.13E−10	34.83	8.98E−10	37.97
GTT Alg. 3.1	2.35E−06	34.06	6.87E−06	47.08	3.99E−06	32.58	7.74E−06	37.78
GTT Alg. 3.2	2.45E−10	33.57	4.73E−10	46.13	4.13E−10	32.66	8.20E−10	37.83
SI Alg. 3.1	4.54E−07	46.23	1.11E−06	61.55	7.80E−07	45.01	1.24E−06	49.05
SDJ Alg. 4.3	2.32E−02	21.82	9.13E−02	28.52	3.99E−02	21.43	7.40E−02	23.83

Finally, we consider an example in 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 $\langle x, y\rangle = \sum_{i=1}^{\infty} x_i y_i$ and induced norm $\|x\| = \sqrt{\langle x, x\rangle}$ for any $x, y \in \mathcal{H}$.

Example 4.4 Let $C := \{x \in \mathcal{H} : |x_i| \le 1/i\}$ and $\beta > \alpha > \beta/2 > 0$. Define an operator $M : C \to \mathcal{H}$ by $Mx = \big(\beta - \|x\|\big)x$. It can be verified that mapping M is pseudomonotone on $\mathcal{H}$, $(\beta + 2\alpha)$-Lipschitz continuous on C and sequentially weakly continuous on C, but fails to be monotone on $\mathcal{H}$ (see [21, Example 3.1] for more details). In this example, we take $\beta = 5$, $\alpha = 3$ and $\mathcal{H} = \mathbb{R}^m$ for different values of m. Thus, the feasible set C is a box

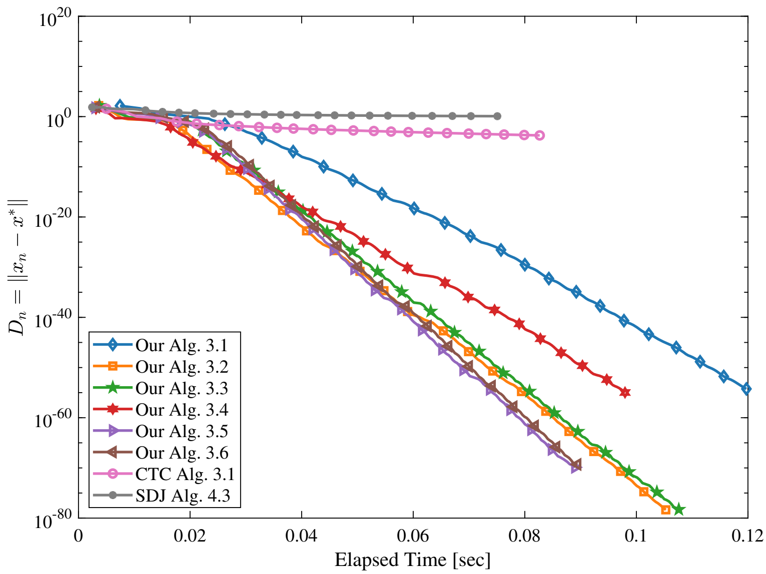


Fig. 3 Example 4.4 for $m = 10,000$

$C = \left\{x \in \mathbb{R}^m : -1/i \leq x_i \leq 1/i,\ i = 1, 2, \ldots, m\right\}$. It is easy to see that the solution of the (VIP) is $x^* = \{\mathbf{0}\}$. The stopping criterion and the choice of initial values are the same as in Example 4.2. We compare the suggested algorithms with the Algorithms 4.3 introduced by Shehu et al. [31] and the Algorithm 3.1 proposed by Cholamjiak et al. [7] (shortly, CTC Alg. 3.1). Take $\sigma_n = 1/(n+1)$, $\varphi_n = 0.5(1-\sigma_n)$, $\tau = 0.4$, $\epsilon_n = 100/(n+1)^2$, $\theta = 1.5$ and $\vartheta_n = 0.5/L$ for CTC Alg. 3.1. The numerical results of all algorithms with four different dimensions are reported in Fig. 3 and Table 5, where "CPU" in Table 5 denotes the execution time in seconds for all algorithms.

Remark 4.1 From Examples 4.1–4.4, we have the following observations.

(1) It can be seen from Examples 4.1–4.4 that the proposed algorithms are efficient and robust.
(2) From Tables 2, 3, 4, 5, and Figs. 1, 2, 3, it can be seen that the suggested algorithms are better than some existing ones [7,15,28,31,37] in terms of execution time and accuracy, and these results are independent of the size of the dimension and the selection of the initial values.
(3) It should be pointed out that the Algorithm 3.1 introduced by Cholamjiak et al. [7] and the Algorithm 4.3 proposed by Shehu et al. [31] use a fixed step size, that is, the update of step size needs to know the prior information of the Lipschitz constant of the operator M. However, our suggested algorithms can realize self-adaptive update iteration step size, which means that our algorithms are more intelligent and practical. Moreover, it is obvious from Figs. 1 and 2 that the proposed algorithms require less execution time and can obtain higher accuracy under the same stopping criterion than other compared algorithms. The following two reasons can explain this phenomenon: (1) the recommended algorithms have embedded inertial terms that can accelerate the algorithms without inertial terms; (2) the suggested methods use a non-monotonic iteration step size without any line search process, while some Armijo-like line search methods [15,28,37] require more execution

Table 5 Numerical results for Example 4.4

Algotithms	$m = 100$		$m = 1000$		$m = 10,000$		$m = 10,0000$	
	D_n	CPU	D_n	CPU	D_n	CPU	D_n	CPU
Our Alg. 3.1	2.72E−55	0.0339	5.47E−55	0.0321	5.58E−55	0.1197	2.96E−54	0.8197
Our Alg. 3.2	8.45E−79	0.0278	5.11E−79	0.0252	4.44E−79	0.1053	5.32E−79	0.6933
Our Alg. 3.3	1.45E−78	0.0256	6.89E−79	0.0243	5.15E−79	0.1076	6.73E−79	0.6757
Our Alg. 3.4	1.32E−55	0.0249	9.51E−56	0.0266	1.06E−55	0.0980	1.43E−55	0.6990
Our Alg. 3.5	1.91E−75	0.0213	6.96E−73	0.0245	1.33E−70	0.0890	5.91E−68	0.8542
Our Alg. 3.6	3.13E−74	0.0201	2.07E−72	0.0249	5.03E−70	0.0894	2.92E−67	0.6214
CTC Alg. 3.1	1.41E−04	0.0292	1.46E−04	0.0249	1.81E−04	0.0827	1.41E−04	0.5175
SDJ Alg. 4.3	6.82E−01	0.0178	8.17E−01	0.0200	1.22E+00	0.0751	3.13E+00	0.4690

time because they need to evaluate the value of operator M many times in order to find a suitable step size in each iteration.

(4) Notice that the operator M in Example 4.4 is pseudo-monotonic rather than monotonic. In this case the algorithms in the literature [15,28,37] that can only solve monotone variational inequalities will not be available. Therefore, the algorithms proposed in this paper for solving pseudomonotone variational inequalities have a broader range of applications.

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 $L_2\left([0,T],\mathbb{R}^m\right)$ represents the square-integrable Hilbert space with inner product $\langle p,q\rangle=\int_0^T\langle p(t),q(t)\rangle\,\mathrm{d}t$ and norm $\|p\|_2=\sqrt{\langle p,p\rangle}$. The optimal control problem is described as follows:

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

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

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

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

$$g(p)=\Phi(x(T)), \tag{5.3}$$

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:

$$\dot{x}(t)=\frac{\mathrm{d}}{\mathrm{d}t}x(t)=Q(t)x(t)+W(t)p(t),\quad 0\le t\le T,\quad x(0)=x_0, \tag{5.4}$$

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

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

$$\frac{\mathrm{d}}{\mathrm{d}t}s^*(t)=-Q(t)^{\mathsf{T}}s^*(t),\quad s^*(T)=\nabla\Phi\left(x^*(T)\right), \tag{5.6}$$

$$0\in W(t)^{\mathsf{T}}s^*(t)+N_V\left(p^*(t)\right), \tag{5.7}$$

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

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

Denoting $Gp(t):=W(t)^{\mathsf{T}}s(t)$, Khoroshilova [22] showed that Gp is the gradient of the objective function g. Therefore, system (5.5)–(5.7) is reduced to the variational inequality problem

$$\left\langle Gp^{*}, q - p^{*}\right\rangle \geq 0, \quad \forall q \in V\,. \tag{5.8}$$

Recently, there are many approaches to solve the optimal control problem, for example, see [18,22,24,41]. Note that our Algorithms 3.1–3.6 guarantee strong convergence and do not require the Lipschitz constant. Furthermore, 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 $p^N := (p_0, p_1, \ldots, p_{N-1})$ with its piece-wise constant extension:

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

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

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

and

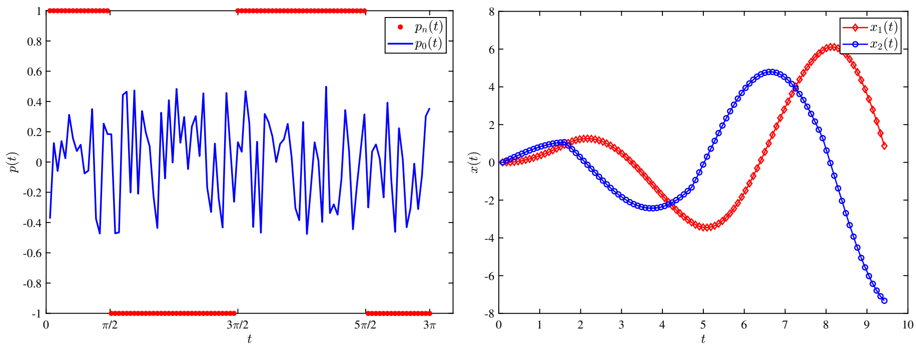
$$s^N(t) = s_i + \frac{t_i - t}{h}\left(s_{i-1} - s_i\right), \quad \forall t \in \left(t_{i-1}, t_i\right], \quad i = N, N-1, \ldots, 1\,.$$

We using 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

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

It is well known that the Euler discretization has the error estimate $O(h)$ [4]. 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 $\left\|p^N - p^*\right\| \leq Kh$.

Next, we present several mathematical examples to illustrate the computational performance of all the algorithms. Our parameters are set as follows. For all algorithms, we set $\sigma_n = 10^{-4}/(n+1)$, $\varphi_n = 0.5(1-\sigma_n)$, $f(x) = 0.1x$ and $\theta = 1.5$. For the suggested Algorithms 3.1–3.6, we choose $\mu = 0.4$, $\vartheta_1 = 0.5$ and $\xi_n = 0.1/(n+1)^{1.1}$. For TG Algs. 3.1 and 3.2, GTT Algs. 3.1 and 3.2, and SI Alg. 3.1, we pick the three parameters of the Armijo-like step size as $\delta = 1$, $\zeta = 0.5$ and $\phi = 0.4$. Take inertial parameters $\tau = 0.01$ and $\epsilon_n = 10^{-4}/(n+1)^2$ in the proposed algorithms. 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 1000.



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

Fig. 4 Numerical results for Example 5.1

Example 5.1 (Control of a harmonic oscillator, see [25])

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

The exact optimal control of Example 5.1 is known:

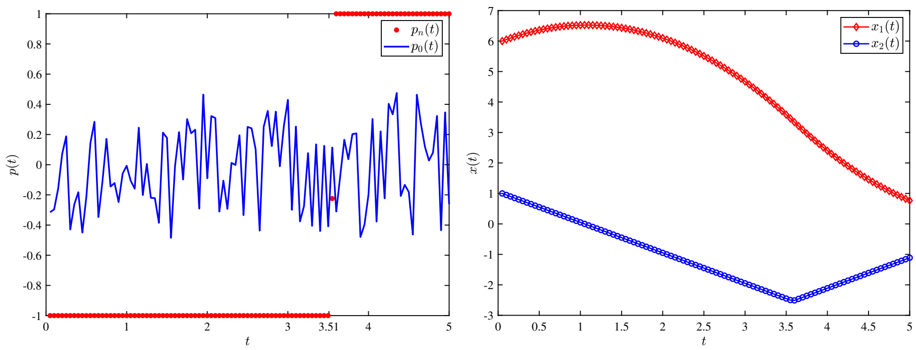
$$p^*(t) = \begin{cases} 1, & \text{if } \ t \in [0, \pi/2) \cup (3\pi/2, 5\pi/2)\,; \\ -1, & \text{if } \ t \in (\pi/2, 3\pi/2) \cup (5\pi/2, 3\pi]\,. \end{cases}$$

According to the previous analysis, we know that the gradient ∇g of the objective function g in the optimal control problem is the operator G in the corresponding variational inequality (5.8). Recall that $\nabla g : \mathbb{R}^n \to \mathbb{R}^n$ is called a monotone mapping if $(\nabla g(x) - \nabla g(y))^{\mathsf{T}}(x - y) \geq 0$. It is easy to see that the gradient of the objective function $g = x_2(3\pi)$ is $\nabla g = [0; 1]$. A simple calculation can verify that ∇g is monotonic. Therefore, Example 5.1 can be transformed into a monotone variational inequality problem (5.8). Consequently, we can use the proposed algorithms to solve the optimal control problem. Figure 4 shows the approximate optimal control and the corresponding trajectories of our Algorithm 3.1.

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

Example 5.2 (Rocket car [24])

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



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

Fig. 5 Numerical results for Example 5.2

The exact optimal control of Example 5.2 is

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

We can show that the gradient $\nabla g = [x_1; x_2]$ of the objective function in Example 5.2 is monotone. The approximate optimal control and the corresponding trajectories of our Algorithm 3.4 are plotted in Fig. 5.

Example 5.3 (See [5])

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

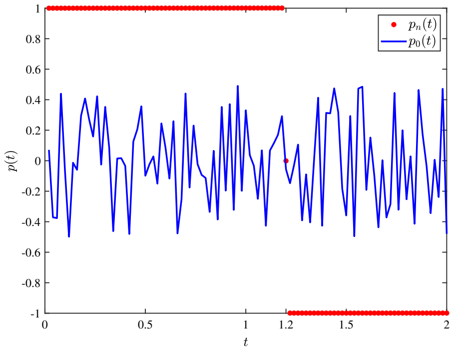
The exact optimal control of Example 5.3 is

$$p^*(t) = \begin{cases} 1, & \text{if } t \in [0, 1.2)\,; \\ -1, & \text{if } t \in (1.2, 2]\,. \end{cases}$$

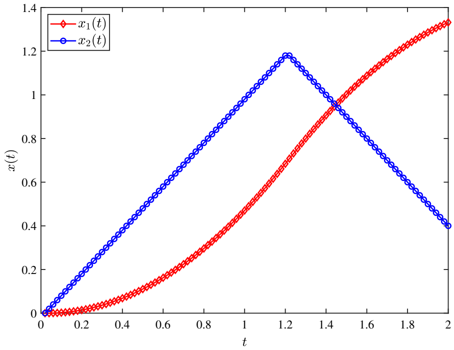
It is easy to check that the gradient of the objective function in Example 5.3 is monotone. Figure 6 gives the approximate optimal control and the corresponding trajectories of our Algorithm 3.6.

Finally, the numerical performance of all the algorithms in Examples 5.1–5.3 is shown in Table 6, where “CPU" in Table 6 stands for the execution time of all algorithms in seconds.

Remark 5.1 From Figs. 4, 5, 6 and Table 6, we know that the suggested algorithms can work well when the terminal function is linear or nonlinear. Moreover, the step size of Algorithm 4.3 [31] requires the prior information of the Lipschitz constant of the cost mapping, and our algorithms can automatically update the iteration step size.



(a) Initial and optimal controls



(b) Optimal trajectories

Fig. 6 Numerical results for Example 5.3

Table 6 Comparison of the number of iterations and execution time of all algorithms in Examples 5.1–5.3

Algorithms	Example 5.1		Example 5.2		Example 5.3	
	Iter.	CPU	Iter.	CPU	Iter.	CPU
Our Alg. 3.1	201	0.10152	583	0.21901	386	0.13496
Our Alg. 3.2	90	0.04435	286	0.12281	202	0.07001
Our Alg. 3.3	90	0.05974	286	0.11152	202	0.07033
Our Alg. 3.4	134	0.07025	1000	0.40773	846	0.25851
Our Alg. 3.5	60	0.04593	1000	0.32034	431	0.13518
Our Alg. 3.6	60	0.03608	1000	0.32069	431	0.13327
TG Alg. 3.1	202	0.10854	678	0.52515	476	0.21914
TG Alg. 3.2	91	0.06312	348	0.26483	238	0.12423
GTT Alg. 3.1	224	0.09730	1000	1.06640	966	0.59361
GTT Alg. 3.2	101	0.05076	445	0.46242	277	0.17807
SI Alg. 3.1	91	0.05042	289	0.29149	218	0.09755

6 Final remarks

This paper proposed six inertial projection and contraction methods to solve the pseudomonotone and Lipschitz continuous variational inequality problem in real Hilbert spaces. The advantage of our algorithms is that they do not require the prior knowledge of the Lipschitz constant of the cost operator. Note that the step size of the iterative schemes suggested in this paper is non-monotonic. The strong convergence theorems of the proposed schemes are obtained under some mild conditions. Finally, some numerical experiments and applications in optimal control problems are provided to illustrate the effectiveness and robustness of our algorithms over several existing related ones.

Acknowledgements The authors are very grateful to the editors and reviewers for their valuable and constructive comments that greatly improved the readability and quality of the initial version of the manuscript. This paper was supported by the National Natural Science Foundation of China under Grant No. 11401152.

References

1. An, N.T., Nam, N.M.: Convergence analysis of a proximal point algorithm for minimizing differences of functions. Optimization **66**, 129–147 (2017)
2. Beck, A.: First-Order Methods in Optimization. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia (2017)
3. Beck, A., Guttmann-Beck, N.: FOM-a MATLAB toolbox of first-order methods for solving convex optimization problems. Optim. Methods Softw. **34**, 172–193 (2019)
4. Bonnans, J.F., Festa, A.: Error estimates for the Euler discretization of an optimal control problem with first-order state constraints. SIAM J. Numer. Anal. **55**, 445–471 (2017)
5. Bressan, B., Piccoli, B.: Introduction to the Mathematical Theory of Control. American Institute of Mathematical Sciences, San Francisco (2007)
6. 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)
7. 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)
8. Cho, S.Y.: A monotone Bregan projection algorithm for fixed point and equilibrium problems in a reflexive Banach space. Filomat **34**, 1487–1497 (2020)
9. Cho, S.Y.: Implicit extragradient-like method for fixed point problems and variational inclusion problems in a Banach space. Symmetry-Basel **12**, 998 (2020)
10. Cottle, R.W., Yao, J.C.: Pseudo-monotone complementarity problems in Hilbert space. J. Optim. Theory Appl. **75**, 281–295 (1992)
11. Dong, Q.L., Jiang, D., Gibali, A.: A modified subgradient extragradient method for solving the variational inequality problem. Numer. Algorithms **79**, 927–940 (2018)
12. Dong, Q.L., Lu, Y.Y., Yang, J.: The extragradient algorithm with inertial effects for solving the variational inequality. Optimization **65**, 2217–2226 (2016)
13. 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)
14. Denisov, S.V., Semenov, V.V., Chabak, L.M.: Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators. Cybern. Syst. Anal. **51**, 757–765 (2015)
15. Gibali, A., Thong, D.V., Tuan, P.A.: Two simple projection-type methods for solving variational inequalities. Anal. Math. Phys. **9**, 2203–2225 (2019)
16. Gibali, A., Hieu, D.V.: A new inertial double-projection method for solving variational inequalities. J. Fixed Point Theory Appl. **21**, Article ID 97 (2019)
17. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35**, 69–76 (1997)
18. Hieu, D.V., Strodiot, J.J., Muu, L.D.: Strongly convergent algorithms by using new adaptive regularization parameter for equilibrium problems. J. Comput. Appl. Math. **376**, Article ID 112844 (2020)
19. Khan, A.A., Motreanu, D.: Existence theorems for elliptic and evolutionary variational and quasi-variational inequalities. J. Optim. Theory Appl. **167**, 1136–1161 (2015)
20. Korpelevich, G.M.: An extragradient method for finding saddle points and for other problems. Èkonom. i Mat. Metody **12**, 747–756 (1976)
21. Khanh, P.Q., Thong, D.V., Vinh, N.T.: Versions of the subgradient extragradient method for pseudomonotone variational inequalities. Acta Appl. Math. **170**, 319–345 (2020)
22. Khoroshilova, E.V.: Extragradient-type method for optimal control problem with linear constraints and convex objective function. Optim. Lett. **7**, 1193–1214 (2013)
23. Liu, H., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput. Optim. Appl. **77**, 491–508 (2020)
24. 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)
25. Pietrus, A., Scarinci, T., Veliov, V.M.: High order discrete approximations to Mayer's problems for linear systems. SIAM J. Control. Optim. **56**, 102–119 (2018)
26. Reich, S., Tuyen, T.M.: Two projection algorithms for solving the split common fixed point problem. J. Optim. Theory Appl. **186**, 148–168 (2020)
27. Shehu, Y., Gibali, A.: New inertial relaxed method for solving split feasibilities. Optim. Lett. **15**, 2109–2126 (2021)
28. Shehu, Y., Iyiola, O.S.: Strong convergence result for monotone variational inequalities. Numer. Algorithms **76**, 259–282 (2017)
29. Shehu, Y., Gibali, A., Sagratella, S.: Inertial projection-type methods for solving quasi-variational inequalities in real Hilbert spaces. J. Optim. Theory Appl. **184**, 877–894 (2020)

30. Shehu, Y., Li, X.H., Dong, Q.L.: An efficient projection-type method for monotone variational inequalities in Hilbert spaces. Numer. Algorithms **84**, 365–388 (2020)
31. Shehu, Y., Dong, Q.L., Jiang, D.: Single projection method for pseudo-monotone variational inequality in Hilbert spaces. Optimization **68**, 385–409 (2019)
32. 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)
33. Sun, D.F.: A class of iterative methods for solving nonlinear projection equations. J. Optim. Theory Appl. **91**, 123–140 (1996)
34. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. **75**, 742–750 (2012)
35. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control. Optim. **38**, 431–446 (2000)
36. Thong, D.V., Hieu, D.V.: Strong convergence of extragradient methods with a new step size for solving variational inequality problems. Comput. Appl. Math. **38**, Article ID 136 (2019)
37. 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)
38. Thong, D.V., Vinh, N.T., Cho, Y.J.: New strong convergence theorem of the inertial projection and contraction method for variational inequality problems. Numer. Algorithms **84**, 285–305 (2020)
39. Tan, B., Fan, J., Li, S.: Self-adaptive inertial extragradient algorithms for solving variational inequality problems. Comput. Appl. Math. **40**, Article ID 19 (2021)
40. Tan, B., Li, S.: Strong convergence of inertial Mann algorithms for solving hierarchical fixed point problems. J. Nonlinear Var. Anal. **4**, 337–355 (2020)
41. 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)
42. Yang, J.: Self-adaptive inertial subgradient extragradient algorithm for solving pseudomonotone variational inequalities. Appl. Anal. **100**, 1067–1078 (2021)
43. Yang, J.: The iterative methods for solving pseudomontone equilibrium problems. J. Sci. Comput. **84**, Article ID 50 (2020)