



Two modified inertial projection algorithms for bilevel pseudomonotone variational inequalities with applications to optimal control problems

Bing Tan¹ · Xiaolong Qin² · Jen-Chih Yao^{3,4}

Received: 6 August 2020 / Accepted: 24 February 2021 / Published online: 16 April 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

In this paper, we investigate two new algorithms for solving bilevel pseudomonotone variational inequality problems in real Hilbert spaces. The advantages of our algorithms are that they only need to calculate one projection on the feasible set in each iteration, and do not require the prior information of the Lipschitz constant of the cost operator. Furthermore, two new algorithms are derived to solve variational inequality problems. We establish the strong convergence of the proposed algorithms under some suitable conditions imposed on parameters. Finally, several numerical results and applications in optimal control problems are reported to illustrate the efficiency and advantages of the proposed algorithms.

Keywords Bilevel variational inequality problem · Inertial extragradient method · Projection and contraction method · Hybrid steepest descent method · Pseudomonotone mapping

Mathematics Subject Classification (2010) 47H05 · 47J20 · 47J25 · 65Y05 · 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

Let C be a closed and convex nonempty subset in a real Hilbert space \mathcal{H} with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Let us first review some nonlinear mappings in convex analysis. For any $x, y \in \mathcal{H}$, the mapping $M : \mathcal{H} \rightarrow \mathcal{H}$ is said to be (i) L -Lipschitz continuous with $L > 0$ if $\|Mx - My\| \leq L\|x - y\|$ (if $L = 1$, then M is called nonexpansive); (ii) γ -inverse strongly monotone (or γ -cocoercive) if there exists $\gamma > 0$ such that $\langle Mx - My, x - y \rangle \geq \gamma\|Mx - My\|^2$; (iii) α -strongly monotone if there exists $\alpha > 0$ such that $\langle Mx - My, x - y \rangle \geq \alpha\|x - y\|^2$; (iv) monotone if $\langle Mx - My, x - y \rangle \geq 0$; (v) pseudomonotone if $\langle Mx, y - x \rangle \geq 0 \implies \langle My, y - x \rangle \geq 0$; (vi) sequentially weakly continuous if for each sequence $\{x_n\}$ converges weakly to x implies $\{Mx_n\}$ converges weakly to Mx . Through the above definitions, it is easy to see that (iii) \implies (iv) \implies (v). The main purpose of this paper is to devote several efficient numerical methods for solving the bilevel variational inequality problem (shortly, **BVIP**) involving a pseudomonotone mapping in real Hilbert spaces. Let $M : \mathcal{H} \rightarrow \mathcal{H}$ and $S : \mathcal{H} \rightarrow \mathcal{H}$ be two single-valued mappings. Recall that the variational inequality problem (**VIP**) for the mapping M on C is described as follows:

$$\text{find } y^* \in C \text{ such that } \langle My^*, z - y^* \rangle \geq 0, \quad \forall z \in C. \quad (\text{VIP})$$

We denote Ω the set of all solutions of the (**VIP**). Recall that the (**BVIP**) is formed as follows:

$$\text{find } x^\dagger \in \Omega \text{ such that } \langle Sx^\dagger, y - x^\dagger \rangle \geq 0, \quad \forall y \in \Omega. \quad (\text{BVIP})$$

Since the bilevel variational inequality problems include a number of problems, such as, quasi-variational inequality problems, complementary problems, and so on. It is therefore necessary to develop some fast and efficient numerical approaches to solve the bilevel variational inequalities. Some recent work on solution methods for (**BVIP**) can be found in [1–4]. There are also a number of methods dealing with approximation solution of (**VIP**); see, e.g., [5–8]. The simplest of these algorithms is the projected-gradient method, which, starting from any $x_0 \in C$, iteratively updates x_{n+1} according to the formula

$$x_{n+1} = P_C (x_n - \vartheta Mx_n), \quad (1.1)$$

where M is a nonlinear mapping, ϑ is a positive fixed step size and P_C denotes the metric projection onto C (see definition in Section 2). The projected-gradient method is based on the observation that $x^\dagger \in C$ is a solution of (**VIP**) if and only if

$$x^\dagger = P_C (x^\dagger - \vartheta Mx^\dagger). \quad (1.2)$$

This projected-gradient method (1.1) can be easily implemented because it only needs to calculate the function value and the projection onto C once in each iteration. However, the projected-gradient method requires a restrictive hypothesis on M for the convergence, that is, M is strongly monotone and Lipschitz continuous. To relax the strong assumptions required by the projected-gradient method and thus broaden

the class of the problems that we can solve, the extragradient method was proposed. Because of (1.2), $x^\dagger \in C$ is a solution of (VIP) if and only if

$$x^\dagger = P_C \left(x^\dagger - \vartheta M \left(P_C \left(x^\dagger - \vartheta M x^\dagger \right) \right) \right).$$

The basic idea of this method is to update x_{n+1} according to the double projection formula

$$x_{n+1} = P_C(x_n - \vartheta M(P_C(x_n - \vartheta M x_n))).$$

The extragradient method (EGM) was first proposed by Korpelevich [9], as shown below. Taking the initial value $x_0 \in C$, we generate a succession $\{x_n\}$ such that

$$\begin{cases} y_n = P_C(x_n - \vartheta M x_n), \\ x_{n+1} = P_C(x_n - \vartheta M y_n). \end{cases} \quad (\text{EGM})$$

It is known that the convergence of the extragradient method is proved under the following hypothesis: $\Omega \neq \emptyset$, mapping M is L -Lipschitz continuous monotone and fixed step size $\vartheta \in (0, 1/L)$. However, we note that the (EGM) needs to perform two projection calculations on the feasible set C in each iteration, which may seriously affect the computational performance, especially when C is a general closed convex set. To overcome this disadvantage, Censor, Gibali and Reich [10] introduced the subgradient extragradient method (SEGM), which can be seen as a modification of the (EGM). They replaced the second projection onto C with a projection onto a half-space. More precisely, their algorithm is expressed as follows:

$$\begin{cases} y_n = P_C(x_n - \vartheta M x_n), \\ T_n = \{z \in \mathcal{H} \mid \langle x_n - \vartheta M x_n - y_n, z - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \vartheta M y_n), \end{cases} \quad (\text{SEGM})$$

where mapping M is L -Lipschitz continuous monotone and fixed step size ϑ is in $(0, 1/L)$. They confirmed that the (SEGM) is weakly convergent in a Hilbert space. It is worth noting that the projection onto a half-space T_n can be calculated by an explicit formula. This greatly improves the computational performance of the (EGM).

Recently, Dong, Jiang and Gibali [11] proposed a modified subgradient extragradient method (MSEGM) for solving the (VIP) by improving the step size in the second step of the (SEGM). This method was inspired by the subgradient extragradient method and the projection and contraction method [12, 13]. Their primary example demonstrates the numerical performance and advantages of this new method compared with some existing approaches. Indeed, the (MSEGM) is of the form:

$$\begin{cases} y_n = P_C(x_n - \vartheta_n M x_n), \\ T_n = \{z \in \mathcal{H} \mid \langle x_n - \vartheta_n M x_n - y_n, z - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \theta \vartheta_n \chi_n M y_n), \end{cases} \quad (\text{MSEGM})$$

where $\theta \in (0, 2)$ and

$$\chi_n := \frac{\langle x_n - y_n, c_n \rangle}{\|c_n\|^2}, \quad c_n := x_n - y_n - \vartheta_n(M x_n - M y_n), \quad (1.3)$$

and $\vartheta_n := \delta \zeta^{m_n}$ ($\delta > 0$, $\zeta \in (0, 1)$) and m_n is the smallest nonnegative integer such that

$$\vartheta_n \|Mx_n - My_n\| \leq \phi \|x_n - y_n\|, \quad \phi \in (0, 1).$$

They proved that the iterative sequence formed by the (MSEGM) converges weakly to a solution of the (VIP) under some approximate conditions. Note that the (MSEGM) uses an Armijo-type line search rule to update the step size in each iteration, so it does not require the prior knowledge of the Lipschitz constant of the mapping.

Next, we introduce a problem related to the (BVIP). Yamada [14] studied the problem of finding a solution of the variational inequality problem over the fixed point set of nonexpansive mappings. More precisely, this problem is stated as follows:

$$\text{find } x^\dagger \in \text{Fix}(T) \text{ such that } \langle Sx^\dagger, y - x^\dagger \rangle \geq 0, \quad \forall y \in \text{Fix}(T), \quad (1.4)$$

where $T : C \rightarrow C$ is a nonexpansive mapping, and $\text{Fix}(T) = \{x \in \mathcal{H} : Tx = x\}$ represents its fixed point set. Yamada introduced the hybrid steepest descent method for solving problem (1.4), which read as follows:

$$x_{n+1} = Tx_n - \delta_{n+1} \vartheta S(Tx_n), \quad (1.5)$$

where mapping S is γ -inverse strongly monotone and L -Lipschitz continuous, fixed step size ϑ is in $(0, 2\gamma/L^2)$ and $\{\delta_n\}$ is a suitable sequence that satisfies some conditions. He proved that the iterative sequence $\{x_n\}$ formed by (1.5) converges to a solution of problem (1.4) in norm. Recently, many scholars have used this method to solve various optimization problems, such as split feasibility problems and variational inequalities; see, e.g., [4, 15, 16]. Let $\delta > 0$. By setting $Tx = P_C(x - \delta Mx)$, we see that $x \in \text{Fix}(T)$ iff $x \in \Omega$. Thus, the (BVIP) becomes problem (1.4) and we can use iterative algorithm (1.5) to solve the (BVIP). However, the convergence of the hybrid steepest descent method requires that mapping M is inversely strong monotone, and this strict assumption may affect the efficiency of the used algorithm. Furthermore, it can be seen from the iterative algorithm (1.5) that the constants L and γ of the mapping S must be known.

Recently, Thong and Hieu [17] combined the modified subgradient extragradient method (MSEGM) with the hybrid steepest descent method (1.5), and introduced a strongly convergent modified subgradient extragradient method for solving bilevel pseudomonotone variational inequality problems in real Hilbert spaces. Their algorithm is illustrated as follows:

$$\begin{cases} y_n = P_C(x_n - \vartheta Mx_n), \\ T_n = \{x \in \mathcal{H} \mid \langle x_n - \vartheta Mx_n - y_n, x - y_n \rangle \leq 0\}, \\ z_n = P_{T_n}(x_n - \theta \vartheta \chi_n My_n), \\ x_{n+1} = z_n - \varphi_n \gamma Sz_n, \end{cases} \quad (1.6)$$

where mapping M is L_M -Lipschitz continuous pseudomonotone on \mathcal{H} , sequentially weakly continuous on C , mapping S is L_S -Lipschitz continuous and α -strongly monotone on \mathcal{H} , $\{\chi_n\}$ is defined in (1.3), the fixed step size ϑ is in $(0, 1/L_M)$, $\theta \in (0, 2)$, $\gamma \in (0, 2\alpha/L_S^2)$ and $\{\varphi_n\}$ is a real sequence in $(0, 1)$ that satisfies $\lim_{n \rightarrow \infty} \varphi_n = 0$ and $\sum_{n=1}^{\infty} \varphi_n = \infty$. Then, the sequence $\{x_n\}$ devised by (1.6)

converges to the unique solution of the (BVIP) in norm. Moreover, their numerical experiments showed that the new algorithm has a better performance than the related one. It should be mentioned that Algorithm (1.6) uses a fixed step size and thus the Lipschitz constant of mapping M must be known.

In recent years, the development of fast iterative algorithms has attracted enormous interest, in particular, the inertial technology, which is based on discrete versions of a second-order dissipative dynamic system. Many researchers have constructed various fast iterative algorithms by using the inertial technology; see, e.g., [18–21] 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 algorithms. Recently, Dong et al. [22] introduced an inertial projection and contraction method (IPCM) to solve the monotone (VIP). For any initial points $x_0, x_1 \in \mathcal{H}$, the iterative sequence $\{x_{n+1}\}$ is devised as follows:

$$\begin{cases} u_n = x_n + \tau_n(x_n - x_{n-1}), \\ y_n = P_C(u_n - \vartheta M u_n), \\ x_{n+1} = u_n - \theta \chi_n c_n, \end{cases} \quad (\text{IPCM})$$

where $\theta \in (0, 2)$, $\vartheta > 0$ and

$$\chi_n := \frac{\langle u_n - y_n, c_n \rangle}{\|c_n\|^2}, \quad c_n := u_n - y_n - \vartheta(M u_n - M y_n).$$

They proved that the (IPCM) achieves the weak convergence in a Hilbert space under appropriate assumptions. Moreover, the stated Algorithm (IPCM) shown the advantages and efficiency over other algorithms through some computational tests.

Motivated and inspired by the above work, in this paper, we introduce two new self-adaptive iterative algorithms for solving bilevel pseudomonotone variational inequality problems in real Hilbert spaces. Our algorithms do not require the prior knowledge of the Lipschitz constant of the potential mapping, and only need to calculate one projection on the feasible set in each iteration. Under certain suitable conditions, we prove that the iterative sequences generated by our algorithms converge strongly to a solution of (BVIP). Based on this, we derive two new strongly convergent methods to solve pseudomonotone (VIP). Finally, some computational tests are presented to support the theoretical results of our new iterative schemes.

The present paper is built up as follows. Some essential definitions and technical lemmas, that need to be used, are given in the next section. Section 3 describes the algorithms and analyzes their convergence. In Section 4, some numerical examples are presented to illustrate the behavior of our algorithms and compare them with the related one. In Section 5, we apply the derived methods to solve optimal control problems. Finally, a brief summary is given in Section 6, the last section.

2 Preliminaries

Let C be a closed and convex nonempty subset of a real Hilbert space \mathcal{H} . The weak convergence and strong convergence of $\{x_n\}_{n=1}^\infty$ to x are represented by $x_n \rightharpoonup x$ and

$x_n \rightarrow x$, respectively. For each $x, y \in \mathcal{H}$, we have

$$\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle. \quad (2.1)$$

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

$$\begin{aligned} \|P_C(x) - P_C(y)\|^2 &\leq \langle P_C(x) - P_C(y), x - y \rangle, \quad \forall x \in \mathcal{H}, y \in \mathcal{H}. \quad (2.2) \\ \langle x - P_C(x), y - P_C(x) \rangle &\leq 0, \quad \forall x \in \mathcal{H}, y \in C. \quad (2.3) \end{aligned}$$

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

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

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

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

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

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

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

The following lemmas play important roles in our proofs.

Lemma 2.1 ([23]) *Let C be a closed and convex nonempty subset of a real Hilbert space \mathcal{H} and let $M : C \rightarrow \mathcal{H}$ be a continuous and pseudomonotone mapping. Then, x^\dagger is a solution of the (VIP) if and only if $\langle Mx, x - x^\dagger \rangle \geq 0, \forall x \in C$.*

Lemma 2.2 ([14]) *Suppose that mapping $S : \mathcal{H} \rightarrow \mathcal{H}$ is L_S -Lipschitz continuous and α -strongly monotone with $0 < \alpha \leq L_S$. Define the mapping $T^\gamma : \mathcal{H} \rightarrow \mathcal{H}$ by $T^\gamma x = (I - \varphi\gamma S)(Tx), \forall x \in \mathcal{H}$, where $T : \mathcal{H} \rightarrow \mathcal{H}$ is a nonexpansive mapping, $\gamma > 0$ and $\varphi \in (0, 1]$. Then, T^γ is a contraction provided that $\gamma < \frac{2\alpha}{L_S^2}$, that is,*

$$\|T^\gamma x - T^\gamma y\| \leq (1 - \varphi\chi)\|x - y\|, \quad \forall x, y \in \mathcal{H},$$

where $\chi = 1 - \sqrt{1 - \gamma(2\alpha - \gamma L_S^2)} \in (0, 1)$.

Lemma 2.3 ([24]) *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*

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

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

3 Main results

In this section, we introduce two new self-adaptive iterative methods for solving the (BVIP) and analyze their convergence. The algorithms are inspired by the inertial method, the hybrid steepest descent method (1.5), the modified subgradient extragradient algorithm (1.6) and the inertial projection and contraction method (IPCM). Furthermore, our iterative schemes embed an Armijo-type step size criterion so that they can work without the prior knowledge of the Lipschitz constant of the involved mapping. Before starting to state the main results, we assume that our algorithms satisfy the following assumptions.

- (C1) The feasible set C is a nonempty, closed and convex subset of \mathcal{H} .
- (C2) The solution set of the (VIP) is nonempty, that is, $\Omega \neq \emptyset$.
- (C3) The mapping $M : \mathcal{H} \rightarrow \mathcal{H}$ is L_M -Lipschitz continuous and pseudomonotone on \mathcal{H} , and sequentially weakly continuous on C .
- (C4) The mapping $S : \mathcal{H} \rightarrow \mathcal{H}$ is L_S -Lipschitz continuous and α -strongly monotone on \mathcal{H} such that $\alpha \leq L_S$.
- (C5) Suppose that the positive sequence $\{\varepsilon_n\}$ satisfies $\lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\varphi_n} = 0$, where $\{\varphi_n\} \subset (0, 1)$ such that $\lim_{n \rightarrow \infty} \varphi_n = 0$ and $\sum_{n=1}^{\infty} \varphi_n = \infty$.

3.1 The modified inertial subgradient extragradient algorithm

In this subsection, we propose a new self-adaptive iterative scheme that performs only one projection onto the feasible set. Now, we state the suggested Algorithm 3.1 as follows.

Remark 3.1 We make the following observations for Algorithm 3.1.

- (i) It follows from (3.1) and Assumption (C5) that

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

Indeed, we obtain $\tau_n \|x_n - x_{n-1}\| \leq \varepsilon_n, \forall n$, which together with $\lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\varphi_n} = 0$ yields

$$\lim_{n \rightarrow \infty} \frac{\tau_n}{\varphi_n} \|x_n - x_{n-1}\| \leq \lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\varphi_n} = 0.$$

- (ii) It is worth noting that the definition of χ_n in our Algorithm 3.1 is different from that in (IPCM). Combining (3.2) and (3.3), one sees that

$$\begin{aligned} \frac{\langle u_n - y_n, c_n \rangle}{\|c_n\|^2} &= \frac{\|u_n - y_n\|^2 - \vartheta_n \langle Mu_n - My_n, u_n - y_n \rangle}{\|c_n\|^2} \\ &\geq \frac{\|u_n - y_n\|^2 - \vartheta_n \|Mu_n - My_n\| \|u_n - y_n\|}{\|c_n\|^2} \\ &\geq \frac{(1 - \phi) \|u_n - y_n\|^2}{\|c_n\|^2}. \end{aligned}$$

Algorithm 3.1 Modified inertial subgradient extragradient algorithm for (BVIP).

Initialization: Take $\tau > 0$, $\delta > 0$, $\zeta \in (0, 1)$, $\phi \in (0, 1)$, $\theta \in (0, 2)$ and $\gamma \in \left(0, \frac{2\alpha}{L_S^2}\right)$. 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. Give the iterates x_{n-1} and x_n ($n \geq 1$). Set

$$u_n = x_n + \tau_n(x_n - x_{n-1}),$$

where

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

Step 2. Calculate

$$y_n = P_C(u_n - \vartheta_n M u_n),$$

where ϑ_n is chosen to be the largest $\vartheta \in \{\delta, \delta\zeta, \delta\zeta^2, \dots\}$ satisfying

$$\vartheta \|M u_n - M y_n\| \leq \phi \|u_n - y_n\|. \quad (3.2)$$

If $u_n = y_n$ then stop and y_n is a solution of the (VIP). Otherwise, go to **Step 3**.

Step 3. Calculate

$$z_n = P_{T_n}(u_n - \theta \vartheta_n \chi_n M y_n),$$

where $T_n := \{x \in \mathcal{H} \mid \langle u_n - \vartheta_n M u_n - y_n, x - y_n \rangle \leq 0\}$, and χ_n is defined as follows

$$\chi_n := (1 - \phi) \frac{\|u_n - y_n\|^2}{\|c_n\|^2}, \quad c_n := u_n - y_n - \vartheta_n(M u_n - M y_n). \quad (3.3)$$

Step 4. Calculate

$$x_{n+1} = z_n - \phi_n \gamma S z_n.$$

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

- (iii) It is well known that if $S : \mathcal{H} \rightarrow \mathcal{H}$ is L -Lipschitz continuous and α -strongly monotone on \mathcal{H} and if Ω is a nonempty, closed and convex subset of \mathcal{H} , then the (BVIP) has a unique solution (see, e.g., [25]).

Next, we give some lemmas, which are very useful to prove the convergence of our algorithms.

Lemma 3.4 Suppose that Assumptions (C1)–(C3) hold. The Armijo-like search rule (3.2) is well defined and

$$\min \left\{ \delta, \frac{\phi\zeta}{L_M} \right\} \leq \vartheta_n \leq \delta.$$

Proof Since M is L_M -Lipschitz continuous, one has

$$\|M u_n - M(P_C(u_n - \vartheta M u_n))\| \leq L_M \|u_n - P_C(u_n - \vartheta M u_n)\|,$$

which is equivalent to

$$\frac{\phi}{L_M} \|Mu_n - M(P_C(u_n - \vartheta Mu_n))\| \leq \phi \|u_n - P_C(u_n - \vartheta Mu_n)\|.$$

This implies that (3.2) holds for all $\vartheta \leq \frac{\phi}{L_M}$. Thus, ϑ_n is well defined. It is easy to see that $\vartheta_n \leq \delta$. If $\vartheta_n = \delta$, then this lemma is proved; otherwise, if $\vartheta_n < \delta$, by the search rule (3.2), we know that $\frac{\vartheta_n}{\zeta} > \vartheta_n$ must violate inequality (3.2), that is,

$$\frac{\vartheta_n}{\zeta} \left\| Mu_n - M \left(P_C \left(u_n - \frac{\vartheta_n}{\zeta} Mu_n \right) \right) \right\| > \phi \left\| u_n - P_C \left(u_n - \frac{\vartheta_n}{\zeta} Mu_n \right) \right\|,$$

which, combining with the fact that the mapping M is L_M -Lipschitz continuous, gets $\vartheta_n > \frac{\phi\zeta}{L_M}$. This completes the proof of Lemma 3.4. \square

Lemma 3.5 *If $y_n = u_n$ or $c_n = 0$ in Algorithm 3.1, then $y_n \in \Omega$.*

Proof From the fact that mapping M is L_M -Lipschitz continuous and (3.2), we get

$$\begin{aligned} \|c_n\| &= \|u_n - y_n - \vartheta_n(Mu_n - My_n)\| \\ &\geq \|u_n - y_n\| - \vartheta_n \|Mu_n - My_n\| \\ &\geq (1 - \phi) \|u_n - y_n\|. \end{aligned}$$

It can be easily proved that $\|c_n\| \leq (1 + \phi) \|u_n - y_n\|$. Therefore,

$$(1 - \phi) \|u_n - y_n\| \leq \|c_n\| \leq (1 + \phi) \|u_n - y_n\|,$$

and thus $u_n = y_n$ iff $c_n = 0$. Hence, if $u_n = y_n$ or $c_n = 0$, then we obtain $y_n = P_C(y_n - \vartheta_n My_n)$. In view of (1.2), we get $y_n \in \Omega$. That is the desired conclusion. \square

Lemma 3.6 *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 $x^\dagger \in \Omega$,*

$$\|z_n - x^\dagger\|^2 \leq \|u_n - x^\dagger\|^2 - \|u_n - z_n - \theta \chi_n c_n\|^2 - \theta(2 - \theta) \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_n - y_n\|^2.$$

Proof From $x^\dagger \in \Omega \subset C \subset T_n$ and the property of projection (2.2), we get

$$\begin{aligned} 2\|z_n - x^\dagger\|^2 &\leq 2\left\langle z_n - x^\dagger, u_n - \theta \vartheta_n \chi_n My_n - x^\dagger \right\rangle \\ &= \|z_n - x^\dagger\|^2 + \|u_n - \theta \vartheta_n \chi_n My_n - x^\dagger\|^2 - \|z_n - u_n + \theta \vartheta_n \chi_n My_n\|^2 \\ &= \|z_n - x^\dagger\|^2 + \|u_n - x^\dagger\|^2 + \theta^2 \vartheta_n^2 \chi_n^2 \|My_n\|^2 - 2\left\langle u_n - x^\dagger, \theta \vartheta_n \chi_n My_n \right\rangle \\ &\quad - \|z_n - u_n\|^2 - \theta^2 \vartheta_n^2 \chi_n^2 \|My_n\|^2 - 2\langle z_n - u_n, \theta \vartheta_n \chi_n My_n \rangle \\ &= \|z_n - x^\dagger\|^2 + \|u_n - x^\dagger\|^2 - \|z_n - u_n\|^2 - 2\left\langle z_n - x^\dagger, \theta \vartheta_n \chi_n My_n \right\rangle, \end{aligned}$$

which implies that

$$\|z_n - x^\dagger\|^2 \leq \|u_n - x^\dagger\|^2 - \|z_n - u_n\|^2 - 2\theta \vartheta_n \chi_n \left\langle z_n - x^\dagger, My_n \right\rangle. \quad (3.4)$$

Combining the pseudomonotonicity of mapping M , $y_n \in C$ and $x^\dagger \in \Omega$, we can show from Lemma 2.1 that $\langle My_n, y_n - x^\dagger \rangle \geq 0$, which means that $\langle My_n, z_n - x^\dagger \rangle \geq \langle My_n, z_n - y_n \rangle$. Hence,

$$-2\theta\vartheta_n\chi_n\langle My_n, z_n - x^\dagger \rangle \leq -2\theta\vartheta_n\chi_n\langle My_n, z_n - y_n \rangle. \quad (3.5)$$

Since $z_n \in T_n$, one obtains $\langle u_n - \vartheta_n Mu_n - y_n, z_n - y_n \rangle \leq 0$. This shows that

$$\langle u_n - y_n - \vartheta_n(Mu_n - My_n), z_n - y_n \rangle \leq \vartheta_n\langle My_n, z_n - y_n \rangle. \quad (3.6)$$

Using (3.5), (3.6) and the definition of c_n , we obtain

$$\begin{aligned} -2\theta\vartheta_n\chi_n\langle My_n, z_n - x^\dagger \rangle &\leq -2\theta\chi_n\langle c_n, z_n - y_n \rangle \\ &= -2\theta\chi_n\langle c_n, u_n - y_n \rangle + 2\theta\chi_n\langle c_n, u_n - z_n \rangle. \end{aligned} \quad (3.7)$$

Now, we estimate $-2\theta\chi_n\langle c_n, u_n - y_n \rangle$ and $2\theta\chi_n\langle c_n, u_n - z_n \rangle$. From the definitions of χ_n and c_n and (3.2), we have

$$\begin{aligned} \langle c_n, u_n - y_n \rangle &= \|u_n - y_n\|^2 - \vartheta_n\langle Mu_n - My_n, u_n - y_n \rangle \\ &\geq \|u_n - y_n\|^2 - \vartheta_n\|Mu_n - My_n\|\|u_n - y_n\| \\ &\geq \|u_n - y_n\|^2 - \phi\|u_n - y_n\|^2 \\ &= \chi_n\|c_n\|^2, \end{aligned}$$

which indicates that

$$-2\theta\chi_n\langle c_n, u_n - y_n \rangle \leq -2\theta\chi_n^2\|c_n\|^2. \quad (3.8)$$

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

$$2\theta\chi_n\langle c_n, u_n - z_n \rangle = \|u_n - z_n\|^2 + \theta^2\chi_n^2\|c_n\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2. \quad (3.9)$$

It follows from Lemma 3.5 that $\|c_n\| \leq (1 + \phi)\|u_n - y_n\|$, which combining the definition of χ_n yields that

$$\chi_n^2\|c_n\|^2 = (1 - \phi)^2 \frac{\|u_n - y_n\|^4}{\|c_n\|^2} \geq \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_n - y_n\|^2. \quad (3.10)$$

Combining (3.4), (3.7), (3.8), (3.9) and (3.10), we conclude that

$$\|z_n - x^\dagger\|^2 \leq \|u_n - x^\dagger\|^2 - \|u_n - z_n - \theta\chi_n c_n\|^2 - \theta(2 - \theta) \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_n - y_n\|^2.$$

This completes the proof. \square

Lemma 3.7 [26, 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 $\lim_{k \rightarrow \infty} \|u_{n_k} - y_{n_k}\| = 0$, then $z \in \Omega$.*

Now, we are in a position to prove the convergence of the suggested Algorithm 3.1.

Theorem 3.1 *Suppose that Assumptions (C1)–(C5) hold. Then, the sequence $\{x_n\}$ defined by Algorithm 3.1 converges to the unique solution of the (BVIP) in norm.*

Proof We divide the proof into four statements.

Claim 1 The sequence $\{x_n\}$ is bounded. Indeed, thanks to Lemma 3.6 and $\theta \in (0, 2)$, one has

$$\|z_n - x^\dagger\| \leq \|u_n - x^\dagger\|, \quad \forall n \geq 1. \quad (3.11)$$

From the definition of u_n , one sees that

$$\begin{aligned} \|u_n - x^\dagger\| &\leq \|x_n - x^\dagger\| + \tau_n \|x_n - x_{n-1}\| \\ &= \|x_n - x^\dagger\| + \varphi_n \cdot \frac{\tau_n}{\varphi_n} \|x_n - x_{n-1}\|. \end{aligned} \quad (3.12)$$

It follows from Remark 3.1 (i) that $\frac{\tau_n}{\varphi_n} \|x_n - x_{n-1}\| \rightarrow 0$ as $n \rightarrow \infty$. Thus, there is a constant $Q_1 > 0$ such that

$$\frac{\tau_n}{\varphi_n} \|x_n - x_{n-1}\| \leq Q_1, \quad \forall n \geq 1. \quad (3.13)$$

Combining (3.11), (3.12) and (3.13), one obtains

$$\|z_n - x^\dagger\| \leq \|u_n - x^\dagger\| \leq \|x_n - x^\dagger\| + \varphi_n Q_1, \quad \forall n \geq 1. \quad (3.14)$$

Using the definition of x_{n+1} , Lemma 2.2 and (3.14), one concludes

$$\begin{aligned} \|x_{n+1} - x^\dagger\| &= \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger - \varphi_n \gamma Sx^\dagger\| \\ &\leq \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger\| + \varphi_n \gamma \|Sx^\dagger\| \\ &\leq (1 - \varphi_n \chi) \|z_n - x^\dagger\| + \varphi_n \gamma \|Sx^\dagger\| \\ &\leq (1 - \varphi_n \chi) \|x_n - x^\dagger\| + \varphi_n \chi \cdot \frac{Q_1}{\chi} + \varphi_n \chi \cdot \frac{\gamma}{\chi} \|Sx^\dagger\| \quad (3.15) \\ &\leq \max \left\{ \|x_n - x^\dagger\|, \frac{Q_1 + \gamma \|Sx^\dagger\|}{\chi} \right\} \\ &\leq \cdots \leq \max \left\{ \|x_1 - x^\dagger\|, \frac{Q_1 + \gamma \|Sx^\dagger\|}{\chi} \right\}, \end{aligned}$$

where $\chi = 1 - \sqrt{1 - \gamma(2\alpha - \gamma L_S^2)} \in (0, 1)$. This shows that $\{x_n\}$ is bounded. We assert that $\{u_n\}$, $\{y_n\}$, $\{z_n\}$ and $\{Sz_n\}$ are also bounded sequences.

Claim 2

$$\begin{aligned} &\theta(2 - \theta) \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_n - y_n\|^2 + \|u_n - z_n - \theta \chi_n c_n\|^2 \\ &\leq \|x_n - x^\dagger\|^2 - \|x_{n+1} - x^\dagger\|^2 + \varphi_n Q_4 \end{aligned}$$

for some $Q_4 > 0$. Indeed, it follows from (3.14) that

$$\begin{aligned} \|u_n - x^\dagger\|^2 &\leq \left(\|x_n - x^\dagger\| + \varphi_n Q_1 \right)^2 \\ &= \|x_n - x^\dagger\|^2 + \varphi_n \left(2Q_1 \|x_n - x^\dagger\| + \varphi_n Q_1^2 \right) \quad (3.16) \\ &\leq \|x_n - x^\dagger\|^2 + \varphi_n Q_2 \end{aligned}$$

for some $Q_2 > 0$. From (2.1), (3.16), Lemmas 2.2 and 3.6, we obtain

$$\begin{aligned}
 \|x_{n+1} - x^\dagger\|^2 &= \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger - \varphi_n \gamma Sx^\dagger\|^2 \\
 &\leq \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger\|^2 - 2\varphi_n \gamma \langle Sx^\dagger, x_{n+1} - x^\dagger \rangle \\
 &\leq (1 - \varphi_n \chi)^2 \|z_n - x^\dagger\|^2 + 2\varphi_n \gamma \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle \\
 &\leq \|z_n - x^\dagger\|^2 + \varphi_n Q_3 \\
 &\leq \|x_n - x^\dagger\|^2 + \varphi_n Q_4 - \|u_n - z_n - \theta \chi_n c_n\|^2 - \theta(2 - \theta) \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_n - y_n\|^2,
 \end{aligned} \tag{3.17}$$

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

Claim 3

$$\|x_{n+1} - x^\dagger\|^2 \leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + \varphi_n \chi \left[\frac{2\gamma}{\chi} \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle + \frac{3Q\tau_n}{\varphi_n \chi} \|x_n - x_{n-1}\| \right]$$

for some $Q > 0$. Indeed, we have

$$\begin{aligned}
 \|u_n - x^\dagger\|^2 &= \|x_n + \tau_n(x_n - x_{n-1}) - x^\dagger\|^2 \\
 &= \|x_n - x^\dagger\|^2 + 2\tau_n \langle x_n - x^\dagger, x_n - x_{n-1} \rangle + \tau_n^2 \|x_n - x_{n-1}\|^2 \\
 &\leq \|x_n - x^\dagger\|^2 + 2\tau_n \|x_n - x^\dagger\| \|x_n - x_{n-1}\| + \tau_n^2 \|x_n - x_{n-1}\|^2.
 \end{aligned} \tag{3.18}$$

Using (2.1), (3.1) and (3.11), one has

$$\begin{aligned}
 \|x_{n+1} - x^\dagger\|^2 &= \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger - \varphi_n \gamma Sx^\dagger\|^2 \\
 &\leq \|(I - \varphi_n \gamma S)z_n - (I - \varphi_n \gamma S)x^\dagger\|^2 - 2\varphi_n \gamma \langle Sx^\dagger, x_{n+1} - x^\dagger \rangle \\
 &\leq (1 - \varphi_n \chi)^2 \|z_n - x^\dagger\|^2 + 2\varphi_n \gamma \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle \\
 &\leq (1 - \varphi_n \chi) \|u_n - x^\dagger\|^2 + 2\varphi_n \gamma \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle.
 \end{aligned} \tag{3.19}$$

Substituting (3.18) into (3.19), we obtain

$$\begin{aligned}
 \|x_{n+1} - x^\dagger\|^2 &\leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + 2\varphi_n \gamma \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle \\
 &\quad + \tau_n \|x_n - x_{n-1}\| \left(2\|x_n - x^\dagger\| + \tau \|x_n - x_{n-1}\| \right) \\
 &\leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + \varphi_n \chi \left[\frac{2\gamma}{\chi} \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle + \frac{3Q\tau_n}{\varphi_n \chi} \|x_n - x_{n-1}\| \right],
 \end{aligned}$$

where $Q := \sup_{n \in \mathbb{N}} \{\|x_n - x^\dagger\|, \tau \|x_n - x_{n-1}\|\} > 0$ and $\chi \in (0, 1)$ is defined in Claim 2.

Claim 4 The sequence $\{\|x_n - x^\dagger\|^2\}$ converges to zero. From Lemma 2.3 and Remark 3.1 (i), it remains to show that $\limsup_{k \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_k+1} \rangle \leq 0$ for every subsequence $\{\|x_{n_k} - x^\dagger\|\}$ of $\{\|x_n - x^\dagger\|\}$ satisfying

$$\liminf_{k \rightarrow \infty} \left(\|x_{n_k+1} - x^\dagger\| - \|x_{n_k} - x^\dagger\| \right) \geq 0.$$

For this purpose, we assume that $\{\|x_{n_k} - x^\dagger\|\}$ is a subsequence of $\{\|x_n - x^\dagger\|\}$ such that $\liminf_{k \rightarrow \infty} (\|x_{n_{k+1}} - x^\dagger\| - \|x_{n_k} - x^\dagger\|) \geq 0$. Then,

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

It follows from Claim 2 and Assumption (C5) that

$$\begin{aligned} & \theta(2 - \theta) \frac{(1 - \phi)^2}{(1 + \phi)^2} \|u_{n_k} - y_{n_k}\|^2 + \|u_{n_k} - z_{n_k} - \theta \chi_{n_k} c_{n_k}\|^2 \\ & \leq \limsup_{k \rightarrow \infty} \left[\|x_{n_k} - x^\dagger\|^2 - \|x_{n_{k+1}} - x^\dagger\|^2 + \varphi_{n_k} Q_4 \right] \\ & \leq \limsup_{k \rightarrow \infty} \left[\|x_{n_k} - x^\dagger\|^2 - \|x_{n_{k+1}} - x^\dagger\|^2 \right] + \limsup_{k \rightarrow \infty} \varphi_{n_k} Q_4 \\ & = -\liminf_{k \rightarrow \infty} \left[\|x_{n_{k+1}} - x^\dagger\|^2 - \|x_{n_k} - x^\dagger\|^2 \right] \\ & \leq 0, \end{aligned}$$

which indicates that

$$\lim_{k \rightarrow \infty} \|y_{n_k} - u_{n_k}\| = 0 \text{ and } \lim_{k \rightarrow \infty} \|u_{n_k} - z_{n_k} - \theta \chi_{n_k} c_{n_k}\| = 0. \quad (3.20)$$

From $\|c_{n_k}\| \geq (1 - \phi)\|u_{n_k} - y_{n_k}\|$ and the definition of χ_{n_k} , we have

$$\begin{aligned} \|u_{n_k} - z_{n_k}\| & \leq \|u_{n_k} - z_{n_k} - \theta \chi_{n_k} c_{n_k}\| + \theta \chi_{n_k} \|c_{n_k}\| \\ & = \|u_{n_k} - z_{n_k} - \theta \chi_{n_k} c_{n_k}\| + \theta(1 - \phi) \frac{\|u_{n_k} - y_{n_k}\|^2}{\|c_{n_k}\|} \\ & \leq \|u_{n_k} - z_{n_k} - \theta \chi_{n_k} c_{n_k}\| + \theta \|u_{n_k} - y_{n_k}\|. \end{aligned} \quad (3.21)$$

Combining (3.20) and (3.21), we get

$$\lim_{k \rightarrow \infty} \|z_{n_k} - u_{n_k}\| = 0. \quad (3.22)$$

Moreover, we have

$$\|x_{n_{k+1}} - z_{n_k}\| = \varphi_{n_k} \gamma \|S z_{n_k}\| \rightarrow 0, \text{ as } n \rightarrow \infty, \quad (3.23)$$

and

$$\|x_{n_k} - u_{n_k}\| = \varphi_{n_k} \cdot \frac{\tau_{n_k}}{\varphi_{n_k}} \|x_{n_k} - x_{n_{k-1}}\| \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (3.24)$$

From (3.22), (3.23) and (3.24), we obtain

$$\|x_{n_{k+1}} - x_{n_k}\| \leq \|x_{n_{k+1}} - z_{n_k}\| + \|z_{n_k} - u_{n_k}\| + \|u_{n_k} - x_{n_k}\| \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (3.25)$$

Since $\{x_{n_k}\}$ is bounded, it follows that there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$, which converges weakly to some $z \in \mathcal{H}$, such that

$$\limsup_{k \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_k} \rangle = \lim_{j \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_{k_j}} \rangle = \langle Sx^\dagger, x^\dagger - z \rangle. \quad (3.26)$$

By using (3.24), we get that $u_{n_k} \rightharpoonup z$. This together with $\lim_{k \rightarrow \infty} \|u_{n_k} - y_{n_k}\| = 0$ and Lemma 3.7 concludes that $z \in \Omega$. From (3.26) and x^\dagger is the unique solution of the (BVIP), we get

$$\limsup_{k \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_k} \rangle = \langle Sx^\dagger, x^\dagger - z \rangle \leq 0. \quad (3.27)$$

Using (3.25) and (3.27), we obtain

$$\limsup_{k \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_{k+1}} \rangle = \limsup_{k \rightarrow \infty} \langle Sx^\dagger, x^\dagger - x_{n_k} \rangle \leq 0. \quad (3.28)$$

Therefore, combining (3.28), Remark 3.1 (i) and Claim 3, in the light of Lemma 2.3, we conclude that $\lim_{n \rightarrow \infty} \|x_n - x^\dagger\| = 0$, i.e., $x_n \rightarrow x^\dagger$. We have thus proved the theorem. \square

Now, we give a special case of Algorithm 3.1. Set $S(x) = x - f(x)$ in Theorem 3.1, where mapping $f : \mathcal{H} \rightarrow \mathcal{H}$ is a ρ -contraction. It can be easily verified that mapping $S : \mathcal{H} \rightarrow \mathcal{H}$ is $(1 + \rho)$ -Lipschitz continuous and $(1 - \rho)$ -strongly monotone. In this situation, by picking $\gamma = 1$, we get a new modified inertial subgradient extragradient algorithm for solving (VIP). More specifically, we get the following results.

Corollary 3.1 *Suppose that mapping $M : \mathcal{H} \rightarrow \mathcal{H}$ is L_M -Lipschitz continuous pseudomonotone on \mathcal{H} and sequentially weakly continuous on C , and mapping $f : \mathcal{H} \rightarrow \mathcal{H}$ is a ρ -contraction with $\rho \in [0, \sqrt{5} - 2)$. Take $\tau > 0$, $\delta > 0$, $\zeta \in (0, 1)$, $\phi \in (0, 1)$ and $\theta \in (0, 2)$. Assume that the positive sequence $\{\epsilon_n\}$ satisfies $\lim_{n \rightarrow \infty} \frac{\epsilon_n}{\varphi_n} = 0$, where $\{\varphi_n\} \subset (0, 1)$ such that $\lim_{n \rightarrow \infty} \varphi_n = 0$ and $\sum_{n=1}^{\infty} \varphi_n = \infty$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and iterative sequence $\{x_n\}$ be generated by the following*

$$\begin{cases} u_n = x_n + \tau_n(x_n - x_{n-1}), \\ y_n = P_C(u_n - \vartheta_n M u_n), \\ T_n := \{x \in \mathcal{H} \mid \langle u_n - \vartheta_n M u_n - y_n, x - y_n \rangle \leq 0\}, \\ z_n = P_{T_n}(u_n - \theta \vartheta_n \chi_n M y_n), \\ x_{n+1} = (1 - \varphi_n)z_n + \varphi_n f(z_n), \end{cases} \quad (3.29)$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively. Then the iterative sequence $\{x_n\}$ formed by (3.29) converges to $x^\dagger \in \Omega$ in norm, where $x^\dagger = P_\Omega(f(x^\dagger))$.

Remark 3.2 Our Corollary 3.1 improves and generalizes some recent results in the literature [11, 17, 27] based on the following observations. (1) Our Algorithm (3.29) is strongly convergent while the algorithm (MSEG) introduced by Dong et al. [11] only obtains weak convergence in an infinite-dimensional Hilbert space. (2) The Algorithm (40) suggested by Thong and Hieu in [17] is a fixed step algorithm, but the proposed iterative scheme (3.29) is self-adaptive, i.e., it means that our algorithm can

work without knowing the prior information of the Lipschitz constant of the mapping. (3) When the inertial parameter $\tau_n = 0$ in (3.29), then the stated Algorithm (3.29) is similar to Thong and Gibali's Algorithm 3.2 [27]. Notice that the mapping contained in our Algorithm (3.29) is pseudomonotone, whereas the corresponding mappings in Dong et al.'s Algorithm (MSEG) and Thong and Gibali's Algorithm 3.2 [27] are monotone.

3.2 The new modified inertial projection and contraction algorithm

In this subsection, we introduce a new modified inertial projection and contraction algorithm for solving (BVIP). The iterative procedure only involves the calculation of one projection, and it can work without the prior information of the Lipschitz constant of the mapping. The new Algorithm 3.2 is stated as follows.

Algorithm 3.2 Modified inertial projection and contraction algorithm for (BVIP).

Initialization: Take $\tau > 0$, $\delta > 0$, $\zeta \in (0, 1)$, $\phi \in (0, 1)$, $\theta \in (0, 2)$ and $\gamma \in \left(0, \frac{2\alpha}{L_S^2}\right)$. Let $x_0, x_1 \in \mathcal{H}$ be two initial points.

Iterative Steps: Calculate the next iteration point x_{n+1} as follows:

$$\begin{cases} u_n = x_n + \tau_n(x_n - x_{n-1}), \\ y_n = P_C(u_n - \vartheta_n M u_n), \\ z_n = u_n - \theta \chi_n c_n, \\ x_{n+1} = z_n - \varphi_n \gamma S z_n, \end{cases}$$

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

The following lemma plays a vital role in the convergence analysis of the algorithm.

Lemma 3.8 *Suppose that Assumptions (C1)–(C3) hold. Let $\{z_n\}$, $\{y_n\}$ and $\{u_n\}$ be three sequences generated by Algorithm 3.2. Then, for all $x^\dagger \in \Omega$,*

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

and

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

Proof By using of the definition of z_n , one sees that

$$\begin{aligned} \|z_n - x^\dagger\|^2 &= \|u_n - \theta \chi_n c_n - x^\dagger\|^2 \\ &= \|u_n - x^\dagger\|^2 - 2\theta \chi_n \langle u_n - x^\dagger, c_n \rangle + \theta^2 \chi_n^2 \|c_n\|^2. \end{aligned} \quad (3.30)$$

Combining (3.2) and (3.3), one obtains

$$\begin{aligned}
 \langle u_n - x^\dagger, c_n \rangle &= \langle u_n - y_n, c_n \rangle + \langle y_n - x^\dagger, c_n \rangle \\
 &= \|u_n - y_n\|^2 - \vartheta_n \langle u_n - y_n, Mu_n - My_n \rangle + \langle y_n - x^\dagger, c_n \rangle \\
 &\geq \|u_n - y_n\|^2 - \vartheta_n \|u_n - y_n\| \|Mu_n - My_n\| + \langle y_n - x^\dagger, c_n \rangle \quad (3.31) \\
 &\geq (1 - \phi) \|u_n - y_n\|^2 + \langle y_n - x^\dagger, u_n - y_n - \vartheta_n (Mu_n - My_n) \rangle.
 \end{aligned}$$

From $y_n = P_C(u_n - \vartheta_n Mu_n)$ and the property of projection (2.3), one has

$$\langle u_n - y_n - \vartheta_n Mu_n, y_n - x^\dagger \rangle \geq 0. \quad (3.32)$$

Using $x^\dagger \in \Omega$ and $y_n \in C$, one gets $\langle Mx^\dagger, y_n - x^\dagger \rangle \geq 0$. With the aid of the pseudomonotonicity of M , one obtains

$$\langle My_n, y_n - x^\dagger \rangle \geq 0. \quad (3.33)$$

It follows from (3.3) that $(1 - \phi) \|u_n - y_n\|^2 = \chi_n \|c_n\|^2$. This together with (3.31), (3.32) and (3.33) implies that

$$\langle u_n - x^\dagger, c_n \rangle \geq (1 - \phi) \|u_n - y_n\|^2 = \chi_n \|c_n\|^2. \quad (3.34)$$

Combining the definition of z_n , (3.30) and (3.34), one concludes that

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

On the other hand, by using of the definition of z_n and (3.3), one sees that

$$\begin{aligned}
 \|u_n - y_n\|^2 &= \frac{\chi_n}{1 - \phi} \|c_n\|^2 = \frac{1}{\chi_n(1 - \phi)} \|\chi_n c_n\|^2 \\
 &= \frac{1}{\chi_n(1 - \phi)\theta^2} \|u_n - z_n\|^2. \quad (3.35)
 \end{aligned}$$

From $\|c_n\| \leq (1 + \phi) \|u_n - y_n\|$ and the definition of χ_n , one obtains

$$\chi_n = (1 - \phi) \frac{\|u_n - y_n\|^2}{\|c_n\|^2} \geq \frac{1 - \phi}{(1 + \phi)^2}. \quad (3.36)$$

It implies from (3.35) and (3.36) that

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

This completes the proof of Lemma 3.8. \square

Theorem 3.2 Suppose that Assumptions (C1)–(C5) hold. Then the sequence $\{x_n\}$ formed by Algorithm 3.2 converges to the unique solution of the (BVIP) in norm.

Proof The proof of the theorem is very similar to the proof of Theorem 3.1, we will simplify some derivation steps.

Claim 1 The sequence $\{x_n\}$ is bounded. Indeed, thanks to Lemma 3.8, we have

$$\|z_n - x^\dagger\| \leq \|u_n - x^\dagger\|, \quad \forall n \geq 1. \quad (3.37)$$

From (3.12), (3.13) and (3.37), we obtain

$$\|z_n - x^\dagger\| \leq \|u_n - x^\dagger\| \leq \|x_n - x^\dagger\| + \varphi_n Q_1, \quad \forall n \geq 1. \quad (3.38)$$

Using (3.15) and (3.38), we get

$$\begin{aligned} \|x_{n+1} - x^\dagger\| &\leq (1 - \varphi_n \chi) \|z_n - x^\dagger\| + \varphi_n \gamma \|Sx^\dagger\| \\ &\leq \max \left\{ \|x_n - x^\dagger\|, \frac{Q_1 + \gamma \|Sx^\dagger\|}{\chi} \right\} \\ &\leq \cdots \leq \max \left\{ \|x_1 - x^\dagger\|, \frac{Q_1 + \gamma \|Sx^\dagger\|}{\chi} \right\}. \end{aligned}$$

Thus, sequence $\{x_n\}$ is bounded. Hence, we assert $\{u_n\}$, $\{y_n\}$, $\{z_n\}$ and $\{Sz_n\}$ are also bounded sequences.

Claim 2

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

From (3.16), (3.17) and Lemma 3.8, we can immediately get

$$\begin{aligned} \|x_{n+1} - x^\dagger\|^2 &\leq \|z_n - x^\dagger\|^2 + \varphi_n Q_3 \\ &\leq \|x_n - x^\dagger\|^2 + \varphi_n Q_4 - \frac{2 - \theta}{\theta} \|u_n - z_n\|^2, \end{aligned}$$

where Q_4 is defined in the Claim 2 of Theorem 3.1.

Claim 3

$$\|x_{n+1} - x^\dagger\|^2 \leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + \varphi_n \chi \left[\frac{2\gamma}{\chi} \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle + \frac{3Q\tau_n}{\varphi_n \chi} \|x_n - x_{n-1}\| \right].$$

Combining (3.18), (3.19), (3.37), one obtains

$$\begin{aligned} \|x_{n+1} - x^\dagger\|^2 &\leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + 2\varphi_n \gamma \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle \\ &\quad + \tau_n \|x_n - x_{n-1}\| (2\|x_n - x^\dagger\| + \tau \|x_n - x_{n-1}\|) \\ &\leq (1 - \varphi_n \chi) \|x_n - x^\dagger\|^2 + \varphi_n \chi \left[\frac{2\gamma}{\chi} \langle Sx^\dagger, x^\dagger - x_{n+1} \rangle + \frac{3Q\tau_n}{\varphi_n \chi} \|x_n - x_{n-1}\| \right], \end{aligned}$$

where $Q := \sup_{n \in \mathbb{N}} \{\|x_n - x^\dagger\|, \tau \|x_n - x_{n-1}\|\} > 0$.

Claim 4 $\{\|x_n - x^\dagger\|^2\}$ converges to zero. As proved in Claim 4 of Theorem 3.1, from Claim 2 and $\theta \in (0, 2)$, one has

$$\frac{2 - \theta}{\theta} \|u_{n_k} - z_{n_k}\|^2 \leq \limsup_{k \rightarrow \infty} \left[\|x_{n_k} - x^\dagger\|^2 - \|x_{n_k+1} - x^\dagger\|^2 + \varphi_{n_k} Q_4 \right] \leq 0.$$

Thus, we get $\lim_{k \rightarrow \infty} \|z_{n_k} - u_{n_k}\| = 0$. This together with Lemma 3.8 gives that $\lim_{k \rightarrow \infty} \|y_{n_k} - u_{n_k}\| = 0$. The rest of the proof can refer to the Claim 4 of Theorem 3.1. We leave it to readers to verify. The proof of the theorem is now complete. \square

By setting $S(x) = x - f(x)$ in Theorem 3.2 and choosing $\gamma = 1$, we have the following result.

Corollary 3.2 *Suppose that mapping $M : \mathcal{H} \rightarrow \mathcal{H}$ is L_M -Lipschitz continuous pseudomonotone on \mathcal{H} and sequentially weakly continuous on C , and mapping $f : \mathcal{H} \rightarrow \mathcal{H}$ is ρ -contraction with $\rho \in [0, \sqrt{5} - 2)$. Take $\tau > 0$, $\delta > 0$, $\zeta \in (0, 1)$, $\phi \in (0, 1)$ and $\theta \in (0, 2)$. Assume that the positive sequence $\{\varepsilon_n\}$ satisfies $\lim_{n \rightarrow \infty} \frac{\varepsilon_n}{\varphi_n} = 0$, where $\{\varphi_n\} \subset (0, 1)$ such that $\lim_{n \rightarrow \infty} \varphi_n = 0$ and $\sum_{n=1}^{\infty} \varphi_n = \infty$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and iterative sequence $\{x_n\}$ be created by the following*

$$\begin{cases} u_n = x_n + \tau_n(x_n - x_{n-1}), \\ y_n = P_C(u_n - \vartheta_n M u_n), \\ z_n = u_n - \theta \chi_n c_n, \\ x_{n+1} = (1 - \varphi_n)z_n + \varphi_n f(z_n), \end{cases} \quad (3.39)$$

where $\{\tau_n\}$, $\{\vartheta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively. Then the iterative sequence $\{x_n\}$ generated by (3.39) converges to $x^\dagger \in \Omega$ in norm, where $x^\dagger = P_\Omega(f(x^\dagger))$.

Remark 3.3 Compared with some recent approaches presented in [22, 28–30], the stated Algorithm (3.39) has the following advantages. (i) If the inertial parameter $\tau_n = 0$ in Algorithm (3.39), then it is similar to the viscosity projection-type algorithm 3.2 proposed by Gibali, Thong and Tuan [28]. (ii) Note that the suggested Algorithm (3.39) can update the step size adaptively by applying the Armijo-type step size criterion, while Algorithm 1 introduced by Thong, Vinh and Cho [29] and Algorithm 4.12 proposed by Gibali and Shehu [30] are both fixed step size iterative schemes. (iii) Our Algorithm (3.39) obtains strong convergence in an infinite-dimensional Hilbert space, whereas the Algorithm (IPCM) of Dong et al. [22] can only achieve weak convergence. (iv) The methods offered in [22, 28–30] are used to solve monotone variational inequality problems, while the recommended Algorithm (3.39) can solve pseudomonotone variational inequalities. It is known that the class of pseudomonotone mappings covers the class of monotone mappings, and thus our algorithm is more applicable. Therefore, the stated Corollary 3.2 develops and summarizes some recent results in the literature.

4 Numerical examples

In this section, we implement some examples that appear in finite- and infinite-dimensional spaces to show the numerical performance of the proposed Algorithms 3.1 and 3.2, and also to compare them with the Algorithm (1.6) suggested

by Thong and Hieu [17]. We use the FOM Solver [31] to effectively calculate the projections onto C and T_n . All the programs are implemented in MATLAB 2018a on a personal computer. In our experiments, we consider only variational inequalities governed by pseudomonotone operators which are not monotone.

Example 1 Let an operator $M : \mathbb{R}^m \rightarrow \mathbb{R}^m$ ($m = 5, 10, 30, 50$) be given by

$$M(x) = \frac{1}{\|x\|^2 + 1} \arg \min_{y \in \mathbb{R}^m} \left\{ \frac{\|y\|^4}{4} + \frac{1}{2} \|x - y\|^2 \right\}.$$

We emphasize that the operator M is not monotone. However, the operator M is Lipschitz continuous pseudomonotone (see [32]). The mapping S is defined by $S(x) = \frac{1}{2}x$. It can be easily verified that mapping S is $\frac{1}{2}$ -Lipschitz continuous and $\frac{1}{2}$ -strongly monotone. In this example, we choose the feasible set is a box constraint $C = [1, 3]^m$. Moreover, we estimate that $L_M \approx 1.4$ by using MATLAB. Our parameters are set as follows. In all the algorithms, we set $\theta = 1.5$, $\varphi_n = \frac{0.1}{n+1}$ and $\gamma = \frac{\alpha}{L_S}$. Take $\tau = \{0.1, 0.8\}$, $\varepsilon_n = \frac{1}{(n+1)^2}$, $\delta = \zeta = 0.9$, $\phi = 0.6$ in our Algorithms 3.1 and 3.2. For Algorithm (1.6), we choose the fixed step size $\vartheta = \frac{0.5}{L_M}$. Since we do not know the exact solution to this problem, we use $D_n = \|x_n - x_{n-1}\|$ to study the numerical behavior of all the algorithms. Take initial values $x_0 = x_1$ are randomly generated by $\text{rand}(m, 1)$ in MATLAB. The maximum iteration 50 as a common stopping criterion. For the four different dimensions of the operator M , the numerical results of all the algorithms are presented in Fig. 1.

Example 2 Consider a mapping $S : \mathbb{R}^5 \rightarrow \mathbb{R}^5$ of the form $S(x) = Bx + q$, where $B \in \mathbb{R}^{5 \times 5}$ is a positive-definite and symmetric matrix and $q \in \mathbb{R}^5$ with their entries generated randomly in $(-2, 2)$. It is clear that S is L_S -Lipschitz continuous and α -strongly monotone with $L_S = \max\{\text{eig}(B)\}$ and $\alpha = \min\{\text{eig}(B)\}$, where $\text{eig}(B)$ represents all eigenvalues of B . Taking the feasible set $C = \{x \in \mathbb{R}^5 : 1 \leq x_i \leq 3, i = 1, 2, \dots, 5\}$, we consider the following quadratic fractional programming problem

$$\min_{x \in C} f(x) = \frac{x^\top Qx + a^\top x + a_0}{b^\top x + b_0},$$

with

$$Q = \begin{bmatrix} 5 & -1 & 2 & 0 & 2 \\ -1 & 6 & -1 & 3 & 0 \\ 2 & -1 & 3 & 0 & 1 \\ 0 & 3 & 0 & 5 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{bmatrix}, \quad a = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \\ 1 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \quad a_0 = -2, \quad b_0 = 20.$$

By a straightforward computation, we get

$$M(x) := \nabla f(x) = \frac{(b^\top x + b_0)(2Qx + a) - b(x^\top Qx + a^\top x + a_0)}{(b^\top x + b_0)^2}.$$

The operator M is Lipschitz continuous on C with the constant $L = \max\{\|\nabla M(x)\| : x \in C\}$, see [33]. We get that $L_M \approx 149$ by using MATLAB. It is

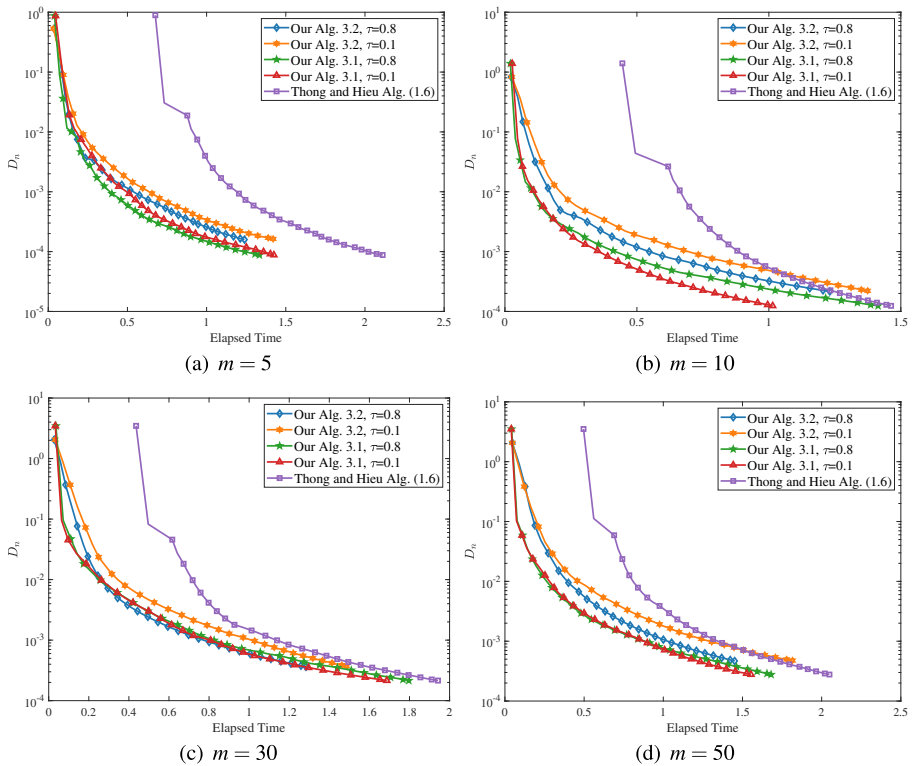


Fig. 1 Numerical results of all the algorithms in Example 1

not sure that whether M is monotone or not? However, M is pseudomonotone because f is pseudo-convex. The exact solution of the problem is $x^* = (1, 1, 1, 1, 1)^T$. In this example, the Lipschitz constant is very large. If the parameters are selected as in Example 1, our algorithms will oscillate due to the large step size. Therefore, we adjust the parameters of the Armijo-type criterion in Algorithms 3.1 and 3.2 to $\delta = 0.003$, $\zeta = 0.9$ and $\phi = 0.1$. Our other parameters are the same as in Example 1. The maximum iteration of 500 as the stopping criterion. Figure 2 shows the numerical performance of $D_n = \|x_n - x^\dagger\|$ of all the algorithms under four different initial values $x_0 = x_1$, which are randomly created by $k \times \text{rand}(5, 1)$ in MATLAB.

Example 3 Finally, we focus on an example that appears in an infinite-dimensional Hilbert space $\mathcal{H} = L^2[0, 1]$ with inner product $\langle x, y \rangle = \int_0^1 x(t)y(t) dt$ and induced norm $\|x\| = \left(\int_0^1 |x(t)|^2 dt \right)^{1/2}$ for all $x, y \in \mathcal{H}$. Let r, R be two positive real numbers such that $R/(k+1) < r/k < r < R$ for some $k > 1$. Take the feasible set $C = \{x \in \mathcal{H} : \|x\| \leq r\}$ and the operator $M : \mathcal{H} \rightarrow \mathcal{H}$ given by

$$M(x) = (R - \|x\|)x, \quad \forall x \in \mathcal{H}.$$

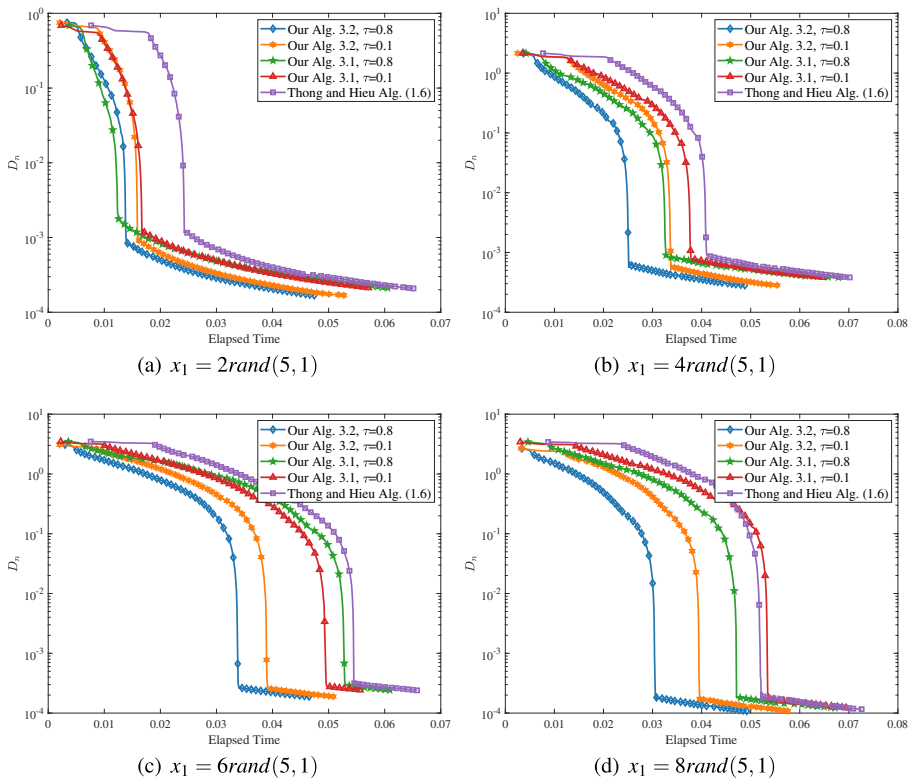


Fig. 2 Numerical results of all the algorithms in Example 2

Note that M is not monotone. Taking a particular pair $(\tilde{x}, \tilde{y}) = (\tilde{x}, k\tilde{x})$, one picks $\tilde{x} \in C$ to satisfy $R/(k+1) < \|\tilde{x}\| < r/k$. One can see that $k\|\tilde{x}\| \in C$. By a simple operation, one gets

$$\langle M(\tilde{x}) - M(\tilde{y}), \tilde{x} - \tilde{y} \rangle = (1-k)^2 \|\tilde{x}\|^2 (R - (1+k)\|\tilde{x}\|) < 0.$$

Hence, the operator M is not monotone on C . Next one shows that M is pseudomonotone. Indeed, one assumes that $\langle M(x), y - x \rangle \geq 0$ for all $x, y \in C$, that is, $\langle (R - \|x\|)x, y - x \rangle \geq 0$. From $\|x\| < R$, one gets that $\langle x, y - x \rangle \geq 0$. Therefore,

$$\begin{aligned} \langle M(y), y - x \rangle &= \langle (R - \|y\|)y, y - x \rangle \\ &\geq (R - \|y\|)(\langle y, y - x \rangle - \langle x, y - x \rangle) \\ &= (R - \|y\|)\|y - x\|^2 \geq 0. \end{aligned}$$

Let $S : \mathcal{H} \rightarrow \mathcal{H}$ be an operator defined by $(Sx)(t) = \frac{1}{2}x(t)$, $t \in [0, 1]$. It is easy to see that S is $\frac{1}{2}$ -strongly monotone and $\frac{1}{2}$ -Lipschitz continuous. For the experiment, we take $R = 1.5$, $r = 1$, $k = 1.1$. We know that the solution to the problem is $x^\dagger(t) = 0$. Our parameters are the same as in Example 1. The maximum iteration of 50 as the stopping criterion. Figure 3 displays the numerical behavior

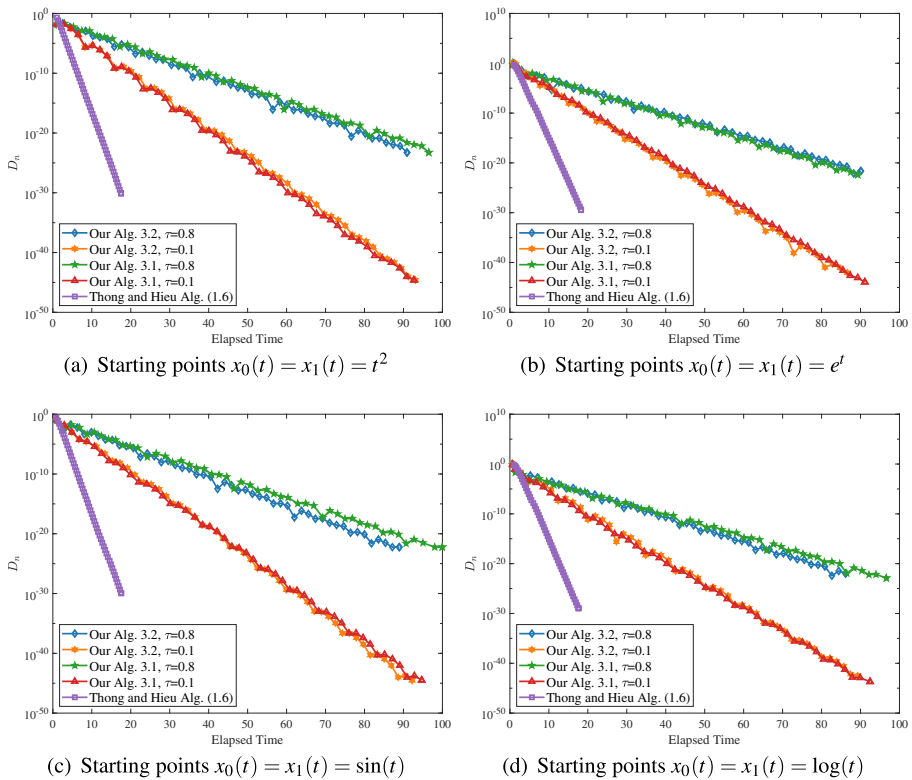


Fig. 3 Numerical results of all the algorithms in Example 3

of $D_n = \|x_n(t) - x^\dagger(t)\|$ generated by all the algorithms with four starting points $x_0(t) = x_1(t)$.

Remark 4.4 We have the following observations for Examples 1–3.

- From Figs. 1, 2, and 3, we can see that the proposed Algorithms 3.1 and 3.2 are more efficient and faster than the Algorithm (1.6) introduced by Thong and Hieu [17] under the appropriate parameters, and these results are independent of the selection of initial values and the size of dimensions. Moreover, our presented algorithms can work adaptively, while the fixed step iterative Algorithm (1.6) depends on the prior information of the Lipschitz constant of the involved mapping, which makes it invalid when the Lipschitz constant is unknown (see Examples 4–6). Therefore, the iterative schemes suggested in this paper are preferable in practical applications.
- From Example 3, it should be emphasized that the proposed Algorithms 3.1 and 3.2 can achieve higher accuracy than the Algorithm (1.6) under the same stopping criterion. However, they need to spend more running time in an infinite-dimensional space to achieve the same error accuracy, because they use an

Armijo-type rule to automatically update the step size and this update criterion requires that the value of operator M to be calculated multiple times in each iteration. It will be interesting to embed a new simple step size used recently in [3, 4, 19, 26] into the algorithms proposed in this paper, and this is also one of our future research topics.

5 Applications to optimal control problems

In this section, we use the derived algorithms (3.29) and (3.39) to solve the variational inequality that occurs in the optimal control problem. Assume that $L_2([0, T], \mathbb{R}^m)$ represents the square-integrable Hilbert space with inner product $\langle p, q \rangle = \int_0^T \langle p(t), q(t) \rangle dt$ and norm $\|p\|_2 = \sqrt{\langle p, p \rangle}$. The optimal control problem is described as follows:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Denoting $Gp(t) := W(t)^\top s(t)$, Khoroshilova [34] showed that Gp is the gradient of the objective function g . Therefore, system (5.5)–(5.7) is reduced to the variational inequality problem

$$\langle Gp^*, q - p^* \rangle \geq 0, \quad \forall q \in V. \quad (5.8)$$

Recently, there are many approaches to solve the optimal control problem, see, for example, [34–37]. Note that our algorithms (3.29) and (3.39) guarantee strong convergence and do not require the prior information of the Lipschitz constant of the mapping. Furthermore, the addition of inertial terms makes them converge faster.

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

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

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

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

and

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

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

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

It is well known that the Euler discretization has the error estimate $O(h)$ [38]. This indicates that the difference between the discretized solution $p^N(t)$ and the original solution $p^*(t)$ is proportional to the mesh size h . That is, there exists a constant $K > 0$ such that $\|p^N - p^*\| \leq Kh$.

Now, we provide some numerical examples to confirm the theoretical results of the derived algorithms (3.29) and (3.39). Our parameters are set as follows:

$$N = 100, \tau = 10^{-2}, \varepsilon_n = \frac{10^{-4}}{(n+1)^2}, \delta = \zeta = 0.5, \phi = 0.4, \theta = 1.5, \varphi_n = \frac{10^{-4}}{n+1}, f(x) = 0.1x.$$

The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-1, 1]$, and the stopping criterion is $\|p_{n+1} - p_n\| \leq 10^{-4}$.

Example 4 (Control of a harmonic oscillator, see [39])

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

The exact optimal control of Example 4 is known:

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

Algorithms (3.29) and (3.39) obtained an approximate solution after 107 and 109 iterations, respectively. They take 0.070847 s and 0.043855 s, respectively. Figure 4 shows the approximate optimal control and the corresponding trajectories of Algorithm (3.29).

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

Example 5 (Rocket car [35])

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

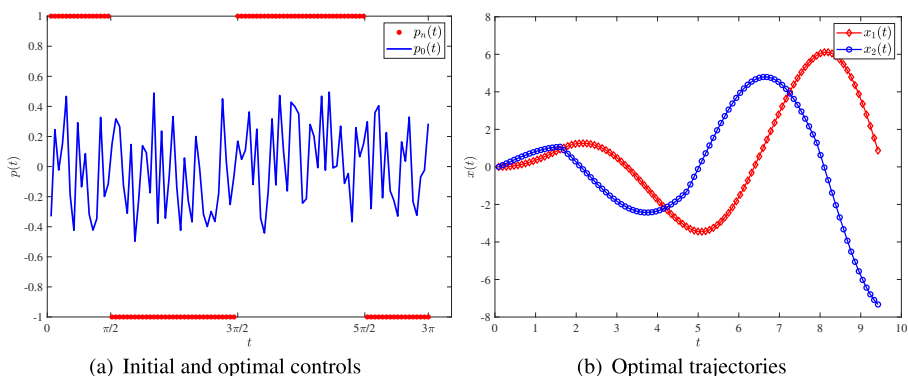


Fig. 4 Numerical results of Algorithm (3.29) in Example 4

The exact optimal control of Example 5 is

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

After 256 iterations, Algorithm (3.29) takes 0.14696 s to obtain an approximate solution. Moreover, Algorithm (3.39) takes 0.28532 s to achieve an approximate solution after 352 iterations. The approximate optimal control and the corresponding trajectories of Algorithm (3.39) are plotted in Fig. 5.

Example 6 (See [40])

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

The exact optimal control of Example 6 is

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

Algorithm (3.29) takes 0.098119 s to obtain an approximate solution after 256 iterations. In addition, Algorithm (3.39) takes 0.17002 s to reach an approximate solution after 346 iterations. Figure 6 gives the approximate optimal control and the corresponding trajectories of Algorithm (3.29).

To compare the execution efficiency of the suggested algorithms (3.29) and (3.39), we show the error estimates $\|p_{n+1} - p_n\|$ of the proposed algorithms for Examples 4–6 in Fig. 7.

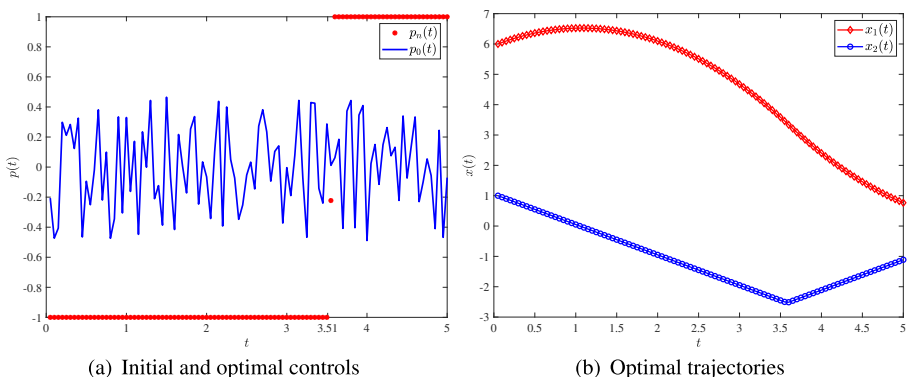


Fig. 5 Numerical results of Algorithm (3.39) in Example 5

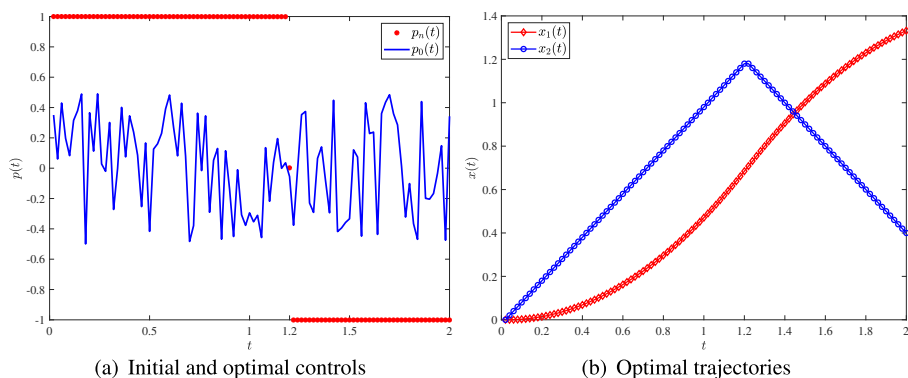


Fig. 6 Numerical results of Algorithm (3.29) in Example 6

Finally, we compare the proposed iterative schemes (3.29) and (3.39) with some strongly convergent algorithms in the literature. Four algorithms used to compare here are the Algorithm (3.39) (shortly, TLQ Alg. (3.39)) proposed by Tan, Liu and Qin [4], the Algorithm 3.2 (THR Alg. 3.2) suggested by Thong, Hieu and Rasiias [26], the Algorithm 3.2 (TG Alg. 3.2) introduced by Thong and Gibali [27] and the Algorithm 3.2 (GTT Alg. 3.2) presented by Gibali, Thong and Tuan [28]. The parameters are $N = 100$, $f(x) = 0.1x$, $\varphi_n = \frac{10^{-4}}{n+1}$ for all the algorithms; $\delta = 1$, $\zeta = 0.5$, $\phi = 0.4$, $\theta = 1.5$ for TG Alg. 3.2, GTT Alg. 3.2 and the suggested algorithms (3.29) and (3.39); $\tau = 10^{-2}$ and $\epsilon_n = \frac{10^{-4}}{(n+1)^2}$ for TLQ Alg. (3.39), THR Alg. 3.2 and the stated algorithms (3.29) and (3.39); $\mu = 0.1$ and $\lambda_1 = 0.4$ for TLQ Alg. (3.39) and THR Alg. 3.2. The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-1, 1]$. The stopping criterion is $\|p_{n+1} - p_n\| \leq 10^{-4}$ or reaching the maximum of 1000 iterations. Table 1 compares the number of iterations and execution time required for all the algorithms to reach the stopping criterion under Examples 4–6.

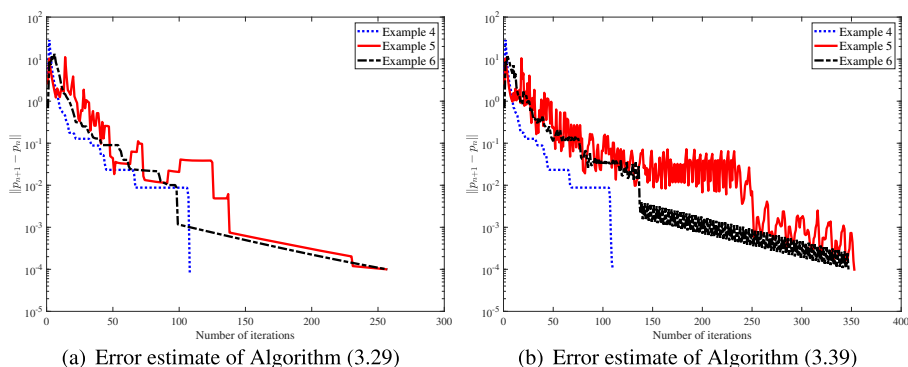


Fig. 7 Error estimates of the proposed algorithms in Examples 4–6

Table 1 Comparison of the number of iterations and execution time of all the algorithms in Examples 4–6

Algorithms	Example 4		Example 5		Example 6	
	Iter.	Time (s)	Iter.	Time (s)	Iter.	Time (s)
Our Alg. (3.29)	90	0.046053	91	0.0873	317	0.13714
Our Alg. (3.39)	101	0.051137	148	0.16716	392	0.21342
TLQ Alg. (3.39)	148	0.056682	1000	0.32776	1000	0.29783
THR Alg. 3.2	148	0.071484	1000	0.37165	1000	0.32059
TG Alg. 3.2	91	0.056392	91	0.12149	317	0.14152
GTT Alg. 3.2	101	0.058568	155	0.16431	456	0.24791

Remark 5.5 From Examples 4–6, we have the following observations.

- (i) From Figs. 4, 5, and 6, it can be seen that the derived algorithms (3.29) and (3.39) work well when the terminal function is linear or nonlinear. However, it is noticed from Fig. 7 that they perform better when the terminal function is linear than when it is nonlinear, that is, they require fewer iterations under the same stopping criterion and have a more stable behavior.
- (ii) As shown in Table 1, the proposed iterative schemes (3.29) and (3.39) have better performance than some known results in the literature, i.e., they need fewer iterations and execution time under the same stopping condition, and these results are independent of the form of the terminal function. Thus, our suggested algorithms are efficient and robust.
- (iii) It is noted that the presented algorithms use an Armijo-type step size criterion, which makes them work without knowing the prior knowledge of the Lipschitz constants of the mapping involved. Indeed, in practical applications, the prior information of the Lipschitz constant is not easy to obtain, and the fixed step size algorithms suggested in [10, 17, 29, 30] will fail in this case. Therefore, several self-adaptive methods proposed in this paper are more useful in reality.

6 Conclusions

In this paper, we proposed two new iterative methods for solving bilevel variational inequality problems in a real Hilbert space when the involved mapping is pseudomonotone and Lipschitz continuous but the Lipschitz constant is unknown. The advantage of the suggested algorithms is that only one projection onto the feasible set needs to be performed. Strong convergence theorems of the stated iterative schemes were proved without the prior knowledge of the Lipschitz constant of the involved mapping. Several numerical experiments were performed to demonstrate the efficiency of the proposed algorithms over the related one. Finally, the derived methods were applied to solve optimal control problems and compared them with existing algorithms.

Acknowledgements The authors are very grateful to the anonymous referees for their valuable and constructive comments, which greatly improved the readability and quality of the initial version of the paper.

References

1. Anh, P.N., Kim, J.K., Muu, L.D.: An extragradient algorithm for solving bilevel pseudomonotone variational inequalities. *J. Global Optim.* **52**, 627–639 (2012)
2. Hieu, D.V., Moudafi, A.: Regularization projection method for solving bilevel variational inequality problem. *Optim. Lett.* **15**, 205–229. <https://doi.org/10.1007/s11590-020-01580-5> (2021)
3. Thong, D.V., Triet, N.A., Li, X.H., Dong, Q.-L.: Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems. *Numer. Algorithms* **83**, 1123–1143 (2020)
4. Tan, B., Liu, L., Qin, X.: Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems. *Jpn. J. Ind. Appl. Math.* <https://doi.org/10.1007/s13160-020-00450-y> (2021)
5. Cho, S.Y.: A monotone Bregan projection algorithm for fixed point and equilibrium problems in a reflexive Banach space. *Filomat* **34**, 1487–1497 (2020)
6. Cho, S.Y.: Implicit extragradient-like method for fixed point problems and variational inclusion problems in a Banach space. *Symmetry* **12**, 998 (2020)
7. Shehu, Y., Iyiola, O.S., Li, X.H., Dong, Q.-L.: Convergence analysis of projection method for variational inequalities. *Comput. Appl. Math.* **38**, 161 (2019)
8. Tan, B., Xu, S., Li, S.: Inertial shrinking projection algorithms for solving hierarchical variational inequality problems. *J. Nonlinear Convex Anal.* **21**, 871–884 (2020)
9. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. *Ekon. Matematicheskie Metody* **12**, 747–756 (1976)
10. Censor, Y., Gibali, A., Reich, S.: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optim. Methods Softw.* **26**, 827–845 (2011)
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. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. *Appl. Math. Optim.* **35**, 69–76 (1997)
13. Sun, D.F.: A class of iterative methods for solving nonlinear projection equations. *J. Optim. Theory Appl.* **91**, 123–140 (1996)
14. Yamada, I.: The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. *Inherently Parallel Algorithm. Feasibility Optim. Appl.* **8**, 473–504 (2001)
15. Liu, L.: A hybrid steepest descent method for solving split feasibility problems involving nonexpansive mappings. *J. Nonlinear Convex Anal.* **20**, 471–488 (2019)
16. Slavakis, K., Yamada, I.: Fejér-monotone hybrid steepest descent method for affinely constrained and composite convex minimization tasks. *Optimization* **67**, 1963–2001 (2018)
17. Thong, D.V., Hieu, D.V.: A strong convergence of modified subgradient extragradient method for solving bilevel pseudomonotone variational inequality problems. *Optimization* **69**, 1313–1334 (2020)
18. Gibali, A., Hieu, D.V.: A new inertial double-projection method for solving variational inequalities. *J. Fixed Point Theory Appl.* **21**, 97 (2019)
19. 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)
20. Shehu, Y., Gibali, A.: New inertial relaxed method for solving split feasibilities. *Optim. Lett.* <https://doi.org/10.1007/s11590-020-01603-1> (2020)
21. Tan, B., Li, S.: Strong convergence of inertial Mann algorithms for solving hierarchical fixed point problems. *J. Nonlinear Var. Anal.* **4**, 337–355 (2020)
22. Dong, Q.-L., Cho, Y.J., Zhong, L.L., Rassias, T.H.M.: Inertial projection and contraction algorithms for variational inequalities. *J. Global Optim.* **70**, 687–704 (2018)
23. Cottle, R.W., Yao, J.C.: Pseudo-monotone complementarity problems in Hilbert space. *J. Optim. Theory Appl.* **75**, 281–295 (1992)

24. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. *Nonlinear Anal.* **75**, 742–750 (2012)
25. Muu, L.D., Quy, N.V.: On existence and solution methods for strongly pseudomonotone equilibrium problems. *J. Math.* **43**, 229–238 (2015)
26. Thong, D.V., Hieu, D.V., Rassias, T.M.: Self adaptive inertial subgradient extragradient algorithms for solving pseudomonotone variational inequality problems. *Optim. Lett.* **14**, 115–144 (2020)
27. 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)
28. Gibali, A., Thong, D.V., Tuan, P.A.: Two simple projection-type methods for solving variational inequalities. *Anal. Math. Phys.* **9**, 2203–2225 (2019)
29. 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 (2019)
30. Gibali, A., Shehu, Y.: An efficient iterative method for finding common fixed point and variational inequalities in Hilbert spaces. *Optimization* **68**, 13–32 (2019)
31. Beck, A., Guttman-Beck, N.: FOM—a MATLAB toolbox of first-order methods for solving convex optimization problems. *Optim. Methods Softw.* **34**, 172–193 (2019)
32. Hieu, D.V., Cho, Y.J., Xiao, Y.-b., Kumam, P.: Relaxed extragradient algorithm for solving pseudomonotone variational inequalities in Hilbert spaces. *Optimization* **69**, 2279–2304 (2020)
33. Boj, R.I., Csetnek, E.R., Vuong, P.T.: The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces. *Eur. J. Oper. Res.* **287**, 49–60 (2020)
34. Khoroshilova, E.V.: Extragradient-type method for optimal control problem with linear constraints and convex objective function. *Optim. Lett.* **7**, 1193–1214 (2013)
35. 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)
36. 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)
37. 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**, 112844 (2020)
38. 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)
39. 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)
40. Bressan, B., Piccoli, B.: Introduction to the Mathematical Theory of Control. Applied Mathematics, vol. 2. American Institute of Mathematical Sciences (AIMS), Springfield (2007)

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.