RACSAM (2021) 115:174
https://doi.org/10.1007/s13398-021-01116-1

ORIGINAL PAPER

An accelerated extragradient algorithm for bilevel pseudomonotone variational inequality problems with application to optimal control problems

Bing Tan[1] · **Songxiao Li**[1] · **Xiaolong Qin**[2]

Received: 5 March 2021 / Accepted: 13 August 2021

Abstract
In this paper, an inertial extragradient algorithm with a new non-monotonic stepsize is proposed to solve the bilevel pseudomonotone variational inequality problem in real Hilbert spaces. The advantages of the suggested iterative algorithm are that only one projection onto the feasible set needs to be performed in each iteration and the prior knowledge of the Lipschitz constant of the mapping involved does not require to be known. The strong convergence theorem of the suggested algorithm is established under some suitable conditions. Numerical experiments are reported to illustrate the advantages and efficiency of the presented algorithm over the existing related ones.

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

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

1 Introduction

Variational inequality problems are important mathematical tools in many fields and have been rapidly developed in theories, algorithms and applications; see, e.g., [1–6]. Let us review the classical variational inequality problem (in short, VIP), which reads as follows:

$$\text{find } y^* \in C \text{ such that } \langle Ay^*, z - y^* \rangle \geq 0, \quad \forall z \in C, \tag{VIP}$$

✉ Songxiao Li
jyulsx@163.com

Bing Tan
bingtan72@gmail.com

Xiaolong Qin
qxlxajh@163.com

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

2 Department of Mathematics, Zhejiang Normal University, Zhejiang 321004, China

Published online: 23 August 2021

where C is a nonempty closed convex subset of a real Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$ and induced norm $\|\cdot\|$, and $A:\mathcal{H}\to\mathcal{H}$ is a mapping. We use $\mathrm{VI}(C,A)$ to represent the solution set of the (VIP). Many researchers have proposed various types of numerical methods to solve the (VIP); see, e.g., [7–15] and the references therein. In this paper, we focus on finding a solution to the following variational inequality problem with a variational inequality constraint defined on a nonempty closed convex subset C. The problem is now referred to as the bilevel variational inequality problem (in short, BVIP) and is described as follows:

$$\text{find } x^*\in \mathrm{VI}(C,A) \text{ such that } \left\langle Fx^*, y-x^*\right\rangle \geq 0, \quad \forall y\in \mathrm{VI}(C,A), \tag{BVIP}$$

where $F:\mathcal{H}\to\mathcal{H}$ is a mapping. It is known that the (BVIP) involves various types of mathematical applications with some constraints, including, such as equilibrium problems, variational inequalities, bilevel linear programs, minimum-norm solution problems and bilevel optimization problems. The purpose of this paper is to develop an efficient numerical method to solve the (BVIP) in an infinite-dimensional Hilbert space.

Another problem related to (BVIP) is the following one introduced by Yamada [16]:

$$\text{find } x^*\in \mathrm{Fix}(T) \text{ such that } \left\langle Fx^*, y-x^*\right\rangle \geq 0, \quad \forall y\in \mathrm{Fix}(T), \tag{1.1}$$

where $F:\mathcal{H}\to\mathcal{H}$ is a L_F-Lipschitz continuous and ψ-strongly monotone mapping (see the definition in Sect. 2), $T:C\to C$ is a nonexpansive mapping and its fixed point set denoted by $\mathrm{Fix}(T)$. Note that problem (1.1) can be converted to the (BVIP) by setting $T(x)=P_C(x-\chi Ax)$ in (1.1). It is easy to verify $\mathrm{VI}(C,A)=\mathrm{Fix}(T)$ according to the property of the projection. In order to solve problem (1.1), Yamada [16] presented a new iterative scheme, which is now called the hybrid steepest descent method (shortly, HSDM) and its form is as follows: $x_{n+1}=T\left(I-\chi_{n+1}\mu F\right)(x_n)$, where $\{\chi_n\}$ and μ are suitable parameters that satisfy some restrictions. The iterative sequence generated by the method (HSDM) converges strongly to the unique solution of problem (1.1). Therefore, we can use the method (HSDM) to solve the (BVIP). However, the convergence of the method needs to satisfy the assumption that T is nonexpansive, which will further impose strict restrictions on the mapping A (e.g., the mapping A requires to meet inverse strongly monotonicity) and affect the method used.

Recall that among the various numerical methods for solving the monotone (VIP), the extragradient method (EGM) based on projection-type introduced by Korpelevich [17] is very popular among researchers. The EGM is a two-step iterative scheme that requires two projections to be performed on the feasible set C in each iteration. Note that calculating the projection onto a closed convex set C is equivalent to solving a minimum distance optimization problem. This may affect the computational efficiency of the method used when C has a complex structure. To overcome this shortcoming, many scholars have made great efforts and achieved some important results; see, for example, the projection and contraction method (PCM) proposed by He [18], the Tseng's extragradient method (TEGM) introduced by Tseng [19], the subgradient extragradient method (SEGM) suggested by Censor et al. [20] and the modified subgradient extragradient method (MSEGM) presented by Malitsky and Semenov [21]. A common feature of these methods is that they only need to execute one projection on the feasible set in each iteration and this change significantly improves the computational efficiency of the EGM. It is known that the extragradient method for variational inequality problem given in [22] was extended in [21] to equilibrium problems for pseudomonotone and Lipschitz-type continuous bifunctions, replacing the naturally two projections considered in [21] by two optimization programs. With hybrid extragradient methods, the authors introduced in [23] a new iterative process for approximating a common element of the set of solutions of an equilibrium problem and a common zero of a finite family of monotone operators in Hadamard spaces. Two new extragradient variants for the classical equilibrium problem in real Hilbert spaces are introduced in [24], getting strong convergence with weak convergence assumptions. Recently, some iterative schemes based on the methods PCM, TEGM, SEGM, MSEGM and HSDM have been offered to solve the (BVIP) in real Hilert spaces; see, e.g., [25–30] and the references therein.

It is worth noting that the algorithm stated by Thong and Hieu [25, Algorithm 3.1] requires to know the prior information of the Lipschitz constant of the mapping A. In other words, the update of the step size of the algorithm needs the Lipschitz constant as an input parameter, which will affect the implementation of such algorithm without knowing the Lipschitz constant. To overcome this difficulty, the self-adaptive methods that do not necessitate to know the Lipschitz constant of the mapping in advance are very valuable. The algorithm proposed by Ceng [26, Algorithm 3.1] uses an Armijo-type linesearch criteria to update the iteration stepsize. It is known that an approach with a linesearch will require many additional computations and further reduces the computational efficiency of the method used. Recently, many methods with a simple stepsize have been proposed for solving the (BVIP); see, e.g., [27, Algorithm 1], [28, Algorithms 1 and 2], [29, Algorithm 1], and [30, Algorithms 3.1 and 3.2]. A common characteristic of these methods is that they can automatically update the stepsize in each iteration by using some previously known information to perform a simple calculation. However, it should be emphasized that the stepsize of these methods does not increase, which may affect the execution efficiency of these algorithms since they rely on the selection of the initial stepsize.

On the other hand, the idea of inertial has been studied by many scholars as one of the important tools to accelerate the convergence speed of the algorithms. They have constructed a large number of fast numerical algorithms to solve variational inequalities, split feasibility problems, fixed point problems, and inclusion problems; see, e.g., [6,9,14,15,26,30–35] and the references therein. The computational efficiency of these inertial-type methods has been demonstrated in many numerical experiments and applications.

Inspired and motivated by the above research results, in this paper, we introduce a new inertial self-adaptive extragradient method to solve the bilevel variational inequality problem in a real Hilbert space. The advantages of the suggested iterative scheme are that (1) the operator A involved is pseudomonotone rather than monotone; (2) the projection onto the feasible set needs to be evaluated only once in each iteration; (3) the method uses a new non-monotonic step size so that it can work without knowing the Lipschitz constant of the mapping; and (4) the algorithm embeds inertial terms making it can to accelerate the convergence speed of the algorithm without inertial terms. The strong convergence theorem of the stated iterative scheme is established under some suitable assumptions imposed on the operators and parameters. Finally, some numerical experiments and applications are provided to demonstrate the advantages and efficiency of the offered algorithm over the previously known ones in [29,30].

This paper is organized as follows. Some basic definitions and lemmas that need to be used in the proof are given in Sect. 2. Section 3 deals with the stated iterative algorithm and analyzes its convergence. In Sect. 4, several computational tests appearing in finite- and infinite-dimensional spaces are provided to illustrate the efficiency and performance of the suggested algorithm. Finally, the paper concludes with a brief summary in Sect. 5, the last section.

2 Preliminaries

The weak convergence and strong convergence of $\{x_n\}_{n=1}^{\infty}$ to x are represented by $x_n \rightharpoonup x$ and $x_n \to x$, respectively. For any $x, y \in \mathcal{H}$, the operator $T : \mathcal{H} \to \mathcal{H}$ is said to be (i) *L-Lipschitz continuous* with $L > 0$ if $\|Tx - Ty\| \le L\|x - y\|$ (if $L = 1$, then T is called *nonexpansive*); (ii) *ψ-strongly monotone* if there exists $\psi > 0$ such that $\langle Tx - Ty, x - y\rangle \ge \psi\|x - y\|^2$; (iii) *monotone* if $\langle Tx - Ty, x - y\rangle \ge 0$; (iv) *pseudomonotone* if $\langle Tx, y - x\rangle \ge 0 \Longrightarrow \langle Ty, y - x\rangle \ge 0$; (v) *sequentially weakly continuous* if for each sequence $\{x_n\}$ converges weakly to x implies $\{Tx_n\}$ converges weakly to Tx. For each $x, y \in \mathcal{H}$, we have

$$\|x + y\|^2 \le \|x\|^2 + 2\langle y, x + y\rangle. \tag{2.1}$$

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

$$\langle x - P_C(x), y - P_C(x)\rangle \le 0, \quad \forall x \in \mathcal{H}, y \in C. \tag{2.2}$$

The following lemmas are important for the convergence analysis of the proposed algorithm.

Lemma 2.1 [16] *Let $\gamma > 0$ and $\sigma \in (0, 1]$. Let $F : \mathcal{H} \to \mathcal{H}$ be a ψ-strongly monotone and L-Lipschitz continuous mapping with $0 < \psi \le L$. Associating with a nonexpansive mapping $T : \mathcal{H} \to \mathcal{H}$, define a mapping $T^\gamma : \mathcal{H} \to \mathcal{H}$ by $T^\gamma x = (I - \sigma\gamma F)(Tx), \forall x \in \mathcal{H}$. Then, T^γ is a contraction provided $\gamma < \frac{2\psi}{L^2}$, that is,*

$$\|T^\gamma x - T^\gamma y\| \le (1 - \sigma\eta)\|x - y\|, \quad \forall x, y \in \mathcal{H},$$

where $\eta = 1 - \sqrt{1 - \gamma\left(2\psi - \gamma L^2\right)} \in (0, 1)$.

Lemma 2.2 [36] *Let $\{a_n\}$ be a sequence of nonnegative real numbers, $\{\sigma_n\}$ be a sequence of real numbers in $(0, 1)$ with $\sum_{n=1}^{\infty} \sigma_n = \infty$, and $\{b_n\}$ be a sequence of real numbers. Assume that*

$$a_{n+1} \le (1 - \sigma_n)\, a_n + \sigma_n b_n, \quad \forall n \ge 1.$$

If $\limsup_{k\to\infty} b_{n_k} \le 0$ for every subsequence $\left\{a_{n_k}\right\}$ of $\{a_n\}$ satisfying $\liminf_{k\to\infty} \left(a_{n_k+1} - a_{n_k}\right) \ge 0$, then $\lim_{n\to\infty} a_n = 0$.

3 Main results

In this section, we present a new inertial extragradient method for solving the (BVIP) and analyze its convergence. First, we suppose that the proposed method satisfy the following conditions.

- (C1) The feasible set C is a nonempty, convex and closed set.
- (C2) The solution set of the (VIP) is nonempty, that is, $\mathrm{VI}(C, A) \ne \emptyset$.
- (C3) The mapping $A : \mathcal{H} \to \mathcal{H}$ is L_A-Lipschitz continuous and pseudomonotone on $\mathcal{H}$, and sequentially weakly continuous on C.
- (C4) The mapping $F : \mathcal{H} \to \mathcal{H}$ is L_F-Lipschitz continuous and ψ-strongly monotone on $\mathcal{H}$ such that $L_F \ge \psi$. In addition, we denote by p the unique solution of the (BVIP).

(C5) Assume $\{\xi_n\}$ and $\{\epsilon_n\}$ are two non-negative positive sequences such that $\sum_{n=1}^{\infty}\xi_n < \infty$ and $\lim_{n\to\infty}\frac{\epsilon_n}{\sigma_n} = 0$, where $\{\sigma_n\} \subset (0,1)$ satisfies $\sum_{n=1}^{\infty}\sigma_n = \infty$ and $\lim_{n\to\infty}\sigma_n = 0$.

The single projection algorithm proposed in this paper for solving the (BVIP) is stated in Algorithm 3.1 below.

Algorithm 3.1 Inertial projection and contraction extragradient algorithm for (BVIP)

Initialization: Take $\theta > 0$, $\chi_1 > 0$, $\mu \in (0,1)$, $\phi \in (0,2)$ and $\gamma \in \left(0, \frac{2\psi}{L_F^2}\right)$. Let $x_0, x_1 \in \mathcal{H}$.
Iterative Steps: Take the iterates x_{n-1} and x_n $(n \geq 1)$. Calculate x_{n+1} as follows:
Step 1. Compute $u_n = x_n + \theta_n\left(x_n - x_{n-1}\right)$, where

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

Step 2. Compute $y_n = P_C\left(u_n - \chi_n A u_n\right)$.
Step 3. Compute $z_n = u_n - \phi\delta_n d_n$, where

$$d_n := u_n - y_n - \chi_n\left(Au_n - Ay_n\right), \quad \delta_n := \begin{cases} \dfrac{\langle u_n - y_n, d_n\rangle}{\|d_n\|^2}, & \text{if } d_n \neq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.2}$$

Step 4. Compute $x_{n+1} = z_n - \sigma_n\gamma F z_n$, and update the step size χ_{n+1} by

$$\chi_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\|u_n - y_n\|}{\|Au_n - Ay_n\|}, \chi_n + \xi_n\right\}, & \text{if } Au_n - Ay_n \neq 0; \\ \chi_n + \xi_n, & \text{otherwise.} \end{cases} \tag{3.3}$$

Remark 3.1 It follows from (3.1) that $\lim_{n\to\infty}\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| = 0$. Indeed, we have $\theta_n\|x_n - x_{n-1}\| \leq \epsilon_n$ for all $n \geq 1$, which together with $\lim_{n\to\infty}\frac{\epsilon_n}{\sigma_n} = 0$ implies that

$$\lim_{n\to\infty}\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| \leq \lim_{n\to\infty}\frac{\epsilon_n}{\sigma_n} = 0.$$

The following lemmas are quite helpful to analyze the convergence of our algorithm.

Lemma 3.1 [37] *Suppose that Condition (C3) holds. Then the sequence $\{\chi_n\}$ generated by (3.3) is well defined and* $\lim_{n\to\infty}\chi_n = \chi$ *and* $\chi \in \left[\min\{\frac{\mu}{L_A}, \chi_1\}, \chi_1 + \sum_{n=1}^{\infty}\xi_n\right]$.

Proof We can easily prove the lemma by means of [37, Lemma 3.1]. We omit the proof here to avoid repetitive expressions. □

Remark 3.2 The idea of the step size χ_n defined in (3.3) is derived from [37]. It is worth noting that the step size χ_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 χ_1. On the other hand, because of $\sum_{n=1}^{\infty}\xi_n < +\infty$, which implies that $\lim_{n\to\infty}\xi_n = 0$. Thus, the step size χ_n may not increase when n is large enough. If $\xi_n = 0$, then the step size χ_n in Algorithm 3.1 is similar to the approaches in [12,14,15,29,30].

Lemma 3.2 *Assume that Conditions (C1)–(C3) hold. Let* $\{y_n\}$ *and* $\{z_n\}$ *be two sequences formed by Algorithm 3.1. Then,*

$$\|z_n - p\|^2 \le \|u_n - p\|^2 - \frac{2-\phi}{\phi}\|u_n - z_n\|^2, \quad \forall p \in \mathrm{VI}(C, A),$$

and

$$\|u_n - y_n\|^2 \le \frac{\left(1 + \frac{\mu\chi_n}{\chi_{n+1}}\right)^2}{\left[\left(1 - \frac{\mu\chi_n}{\chi_{n+1}}\right)\phi\right]^2}\|u_n - z_n\|^2.$$

Proof If there exists $n_1 \ge 1$ such that $d_{n_1} = 0$, then $z_{n_1} = u_{n_1}$ and the first inequality holds. Now, we consider $d_n \ne 0$ for each $n \ge 1$. By the definition of z_n, one obtains

$$\begin{aligned} \|z_n - p\|^2 &= \|u_n - \phi\delta_n d_n - p\|^2 \\ &= \|u_n - p\|^2 - 2\phi\delta_n\langle u_n - p, d_n\rangle + \phi^2\delta_n^2\|d_n\|^2. \end{aligned} \tag{3.4}$$

According to the definition of d_n, one sees that

$$\begin{aligned} \langle u_n - p, d_n\rangle &= \langle u_n - y_n, d_n\rangle + \langle y_n - p, d_n\rangle \\ &= \langle u_n - y_n, d_n\rangle + \langle y_n - p, u_n - y_n - \chi_n(Au_n - Ay_n)\rangle. \end{aligned} \tag{3.5}$$

From $y_n = P_C(u_n - \chi_n Au_n)$ and the property of projection (2.2), we have

$$\langle u_n - y_n - \chi_n Au_n, y_n - p\rangle \ge 0. \tag{3.6}$$

Using $p \in \mathrm{VI}(C, A)$ and $y_n \in C$, we obtain that $\langle Ap, y_n - p\rangle \ge 0$, which combining with the pseudomonotonicity of A yields that

$$\chi_n\langle Ay_n, y_n - p\rangle \ge 0. \tag{3.7}$$

Using (3.5), (3.6) and (3.7), we obtain

$$\langle u_n - p, d_n\rangle \ge \langle u_n - y_n, d_n\rangle. \tag{3.8}$$

It follows from the definition of z_n that $z_n - u_n = \phi\delta_n d_n$. Combining (3.4), (3.8) and the definition of δ_n, we conclude that

$$\begin{aligned} \|z_n - p\|^2 &\le \|u_n - p\|^2 - 2\phi\delta_n\langle u_n - y_n, d_n\rangle + \phi^2\delta_n^2\|d_n\|^2 \\ &= \|u_n - p\|^2 - 2\phi\delta_n^2\|d_n\|^2 + \phi^2\delta_n^2\|d_n\|^2 \\ &= \|u_n - p\|^2 - \frac{2-\phi}{\phi}\|\phi\delta_n d_n\|^2 \\ &= \|u_n - p\|^2 - \frac{2-\phi}{\phi}\|u_n - z_n\|^2. \end{aligned}$$

Therefore, the first inequality is proved. According to the definition of (3.3), one has $\|Au_n - Ay_n\| \le (\mu/\chi_{n+1})\|u_n - y_n\|$, $\forall n \ge 1$, which together with the definition of δ_n implies that

$$\begin{aligned} \delta_n\|d_n\|^2 = \langle d_n, u_n - y_n\rangle \;&\ge \|u_n - y_n\|^2 - \chi_n\,\|Au_n - Ay_n\|\,\|u_n - y_n\| \\ &\ge \left(1 - \frac{\mu\chi_n}{\chi_{n+1}}\right)\|u_n - y_n\|^2. \end{aligned} \tag{3.9}$$

Using the definition of d_n and (3.3), we have

$$\begin{aligned}\|d_n\| &\leq \|u_n - y_n\| + \chi_n \|Au_n - Ay_n\| \\ &\leq \|u_n - y_n\| + \frac{\mu\chi_n}{\chi_{n+1}}\|u_n - y_n\| \\ &= \left(1 + \frac{\mu\chi_n}{\chi_{n+1}}\right)\|u_n - y_n\|. \end{aligned} \tag{3.10}$$

Combining (3.9) and (3.10), one obtains

$$\delta_n^2\|d_n\|^2 \geq \left(1 - \frac{\mu\chi_n}{\chi_{n+1}}\right)^2 \frac{\|u_n - y_n\|^4}{\|d_n\|^2} \geq \frac{\left(1 - \frac{\mu\chi_n}{\chi_{n+1}}\right)^2}{\left(1 + \frac{\mu\chi_n}{\chi_{n+1}}\right)^2}\|u_n - y_n\|^2. \tag{3.11}$$

By the definition of z_n and (3.11), we obtain

$$\|z_n - u_n\|^2 = \phi^2\delta_n^2\|d_n\|^2 \geq \phi^2 \frac{\left(1 - \frac{\mu\chi_n}{\chi_{n+1}}\right)^2}{\left(1 + \frac{\mu\chi_n}{\chi_{n+1}}\right)^2}\|u_n - y_n\|^2.$$

Thus we get

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

The proof of the lemma is now complete. $\square$

Lemma 3.3 [38, Lemma 3.3] *Assume that Conditions (C1)–(C3) hold. Let $\{u_n\}$ be a sequence generated by Algorithm 3.1. If there exists a subsequence $\{u_{n_k}\}$ converges weakly to $z \in \mathcal{H}$ and $\lim_{k\to\infty}\|u_{n_k} - y_{n_k}\| = 0$, then $z \in \mathrm{VI}(C, A)$.*

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

Proof It is known that if Conditions (C2) and (C4) hold, then the (BVIP) has a unique solution (see, e.g., [39]). We divide the proof into four steps.

Claim 1. The sequence $\{x_n\}$ is bounded. From Lemma 3.2 and $\phi \in (0, 2)$, one sees that

$$\|z_n - p\| \leq \|u_n - p\|, \quad \forall n \geq 1. \tag{3.12}$$

By the definition of u_n, we can write

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

According to Remark 3.1, we have $\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| \to 0$ as $n \to \infty$. Therefore, there exists a constant $M_1 > 0$ such that

$$\frac{\theta_n}{\sigma_n}\|x_n - x_{n-1}\| \leq M_1, \quad \forall n \geq 1. \tag{3.14}$$

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

$$\|z_n - p\| \leq \|u_n - p\| \leq \|x_n - p\| + \sigma_n M_1, \quad \forall n \geq 1. \tag{3.15}$$

Using Lemma 2.1 and (3.15), it follows that

$$\begin{aligned}\|x_{n+1}-p\| &= \|(I-\sigma_n\gamma F)z_n-(I-\sigma_n\gamma F)p-\sigma_n\gamma Fp\| \\ &\le (1-\sigma_n\eta)\|z_n-p\|+\sigma_n\gamma\|Fp\| \\ &\le (1-\sigma_n\eta)\|x_n-p\|+\sigma_n\eta\cdot\frac{M_1}{\eta}+\sigma_n\eta\cdot\frac{\gamma}{\eta}\|Fp\| \\ &\le \max\left\{\frac{M_1+\gamma\|Fp\|}{\eta},\|x_n-p\|\right\} \\ &\le \cdots\le \max\left\{\frac{M_1+\gamma\|Fp\|}{\eta},\|x_1-p\|\right\}. \end{aligned} \tag{3.16}$$

where $\eta=1-\sqrt{1-\gamma\left(2\psi-\gamma L_F^2\right)}\in(0,1)$. That is, the sequence $\{x_n\}$ is bounded. We get that the sequences $\{u_n\}$, $\{y_n\}$ and $\{z_n\}$ are also bounded.

Claim 2.

$$\frac{2-\phi}{\phi}\|u_n-z_n\|^2\le\|x_n-p\|^2-\|x_{n+1}-p\|^2+\sigma_n M_4,\quad \forall n\ge 1$$

for some $M_4>0$. By (2.1) and Lemma 2.1, one has

$$\begin{aligned}\|x_{n+1}-p\|^2 &= \|(I-\sigma_n\gamma F)z_n-(I-\sigma_n\gamma F)p-\sigma_n\gamma Fp\|^2 \\ &\le \|(I-\sigma_n\gamma F)z_n-(I-\sigma_n\gamma F)p\|^2-2\sigma_n\gamma\langle Fp,x_{n+1}-p\rangle \\ &\le (1-\sigma_n\eta)^2\|z_n-p\|^2+2\sigma_n\gamma\langle Fp,p-x_{n+1}\rangle \\ &\le \|z_n-p\|^2+\sigma_n M_2 \end{aligned} \tag{3.17}$$

for some $M_2>0$. It follows from (3.15) that

$$\begin{aligned}\|u_n-p\|^2 &\le (\|x_n-p\|+\sigma_n M_1)^2 \\ &= \|x_n-p\|^2+\sigma_n\left(2M_1\|x_n-p\|+\sigma_n M_1^2\right) \\ &\le \|x_n-p\|^2+\sigma_n M_3 \end{aligned} \tag{3.18}$$

for some $M_3>0$. Combining Lemma 3.2, (3.17) and (3.18), we have

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le \|u_n-p\|^2-\frac{2-\phi}{\phi}\|u_n-z_n\|^2+\sigma_n M_2 \\ &\le \|x_n-p\|^2+\sigma_n M_3-\frac{2-\phi}{\phi}\|u_n-z_n\|^2+\sigma_n M_2, \end{aligned}$$

which yields

$$\frac{2-\phi}{\phi}\|u_n-z_n\|^2\le\|x_n-p\|^2-\|x_{n+1}-p\|^2+\sigma_n M_4,$$

where $M_4:=M_2+M_3$. That is the desired conclusion.

Claim 3.

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le (1-\sigma_n\eta)\|x_n-p\|^2 \\ &\quad+\sigma_n\eta\left[\frac{2\gamma}{\eta}\langle Fp,p-x_{n+1}\rangle+\frac{3M\theta_n}{\sigma_n\eta}\|x_n-x_{n-1}\|\right],\quad \forall n\ge 1\end{aligned}$$

for some $M > 0$. From the definition of u_n, one sees that

$$\begin{aligned}\|u_n - p\|^2 &= \|x_n + \theta_n (x_n - x_{n-1}) - p\|^2\\ &\leq \|x_n - p\|^2 + 2\theta_n \|x_n - p\| \|x_n - x_{n-1}\| + \theta_n^2 \|x_n - x_{n-1}\|^2\\ &\leq \|x_n - p\|^2 + 3M\theta_n \|x_n - x_{n-1}\|, \end{aligned} \tag{3.19}$$

where $M := \sup_{n\in\mathbb{N}} \{\|x_n - p\|, \theta \|x_n - x_{n-1}\|\} > 0$. Combining (3.12), (3.17) and (3.19), we obtain

$$\begin{aligned}\|x_{n+1} - p\|^2 &\leq (1 - \sigma_n \eta)^2 \|z_n - p\|^2 + 2\sigma_n \gamma \langle Fp, p - x_{n+1}\rangle\\ &\leq (1 - \sigma_n \eta) \|u_n - p\|^2 + 2\sigma_n \gamma \langle Fp, p - x_{n+1}\rangle\\ &\leq (1 - \sigma_n \eta) \|x_n - p\|^2 + 3M\theta_n \|x_n - x_{n-1}\| + 2\sigma_n \gamma \langle Fp, p - x_{n+1}\rangle\\ &= (1 - \sigma_n \eta) \|x_n - p\|^2 + \sigma_n \eta \left[\frac{2\gamma}{\eta} \langle Fp, p - x_{n+1}\rangle + \frac{3M\theta_n}{\sigma_n \eta} \|x_n - x_{n-1}\|\right],\\ &\quad \forall n \geq 1.\end{aligned}$$

Claim 4. The sequence $\{\|x_n - p\|\}$ converges to zero. Indeed, by Lemma 2.2, it suffices to show that $\limsup_{k\to\infty} \langle Fp, p - x_{n_k+1}\rangle \leq 0$ for every subsequence $\{\|x_{n_k} - p\|\}$ of $\{\|x_n - p\|\}$ satisfying

$$\liminf_{k\to\infty} (\|x_{n_k+1} - p\| - \|x_{n_k} - p\|) \geq 0.$$

For this purpose, one assumes that $\{\|x_{n_k} - p\|\}$ is a subsequence of $\{\|x_n - p\|\}$ such that $\liminf_{k\to\infty} (\|x_{n_k+1} - p\| - \|x_{n_k} - p\|) \geq 0$. Then,

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

By Claim 2 and $\lim_{k\to\infty} \sigma_{n_k} = 0$, one has

$$\begin{aligned}\limsup_{k\to\infty} \left[\frac{2-\phi}{\phi} \|u_{n_k} - z_{n_k}\|^2\right] &\leq \limsup_{k\to\infty} [\sigma_{n_k} M_4 + \|x_{n_k} - p\|^2 - \|x_{n_k+1} - p\|^2]\\ &\leq \limsup_{k\to\infty} \sigma_{n_k} M_4 + \limsup_{k\to\infty} [\|x_{n_k} - p\|^2 - \|x_{n_k+1} - p\|^2]\\ &= -\liminf_{k\to\infty} [\|x_{n_k+1} - p\|^2 - \|x_{n_k} - p\|^2] \leq 0,\end{aligned}$$

which together with $\phi \in (0, 2)$ implies that

$$\lim_{k\to\infty} \|z_{n_k} - u_{n_k}\| = 0. \tag{3.20}$$

Moreover, we can show that

$$\|x_{n_k+1} - z_{n_k}\| = \sigma_{n_k} \gamma \|F z_{n_k}\| \to 0 \quad \text{as } n \to \infty, \tag{3.21}$$

and

$$\|x_{n_k} - u_{n_k}\| = \sigma_{n_k} \cdot \frac{\theta_{n_k}}{\sigma_{n_k}} \|x_{n_k} - x_{n_k-1}\| \to 0 \quad \text{as } n \to \infty. \tag{3.22}$$

Combining (3.20), (3.21) and (3.22), we obtain

$$\|x_{n_k+1} - x_{n_k}\| \leq \|x_{n_k+1} - z_{n_k}\| + \|z_{n_k} - u_{n_k}\| + \|u_{n_k} - x_{n_k}\| \to 0 \quad \text{as } n \to \infty. \tag{3.23}$$

From Lemma 3.2, one sees that

$$\|u_{n_k} - y_{n_k}\|^2 \le \frac{\left(1 + \frac{\mu\chi_{n_k}}{\chi_{n_k+1}}\right)^2}{\left[\left(1 - \frac{\mu\chi_{n_k}}{\chi_{n_k+1}}\right)\phi\right]^2}\|u_{n_k} - z_{n_k}\|^2.$$

which together with (3.20) implies that $\lim_{k\to\infty}\|u_{n_k} - y_{n_k}\| = 0$. Since the sequence $\{x_{n_k}\}$ is bounded, there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$, which converges weakly to some $z \in \mathcal{H}$, and moreover

$$\limsup_{k\to\infty}\left\langle Fp, p - x_{n_k}\right\rangle = \lim_{j\to\infty}\left\langle Fp, p - x_{n_{k_j}}\right\rangle = \langle Fp.p - z\rangle. \tag{3.24}$$

By (3.22), we get $u_{n_k} \rightharpoonup z$ as $k \to \infty$. This together with $\lim_{k\to\infty}\|u_{n_k} - y_{n_k}\| = 0$ and Lemma 3.3 yields that $z \in \mathrm{VI}(C, A)$. From (3.24) and the assumption that p is the unique solution of the (BVIP), we have

$$\limsup_{k\to\infty}\left\langle Fp, p - x_{n_k}\right\rangle = \langle Fp, p - z\rangle \le 0. \tag{3.25}$$

Combining (3.23) and (3.25), we obtain

$$\begin{aligned}\limsup_{k\to\infty}\left\langle Fp, p - x_{n_k+1}\right\rangle &= \limsup_{k\to\infty}\left\langle Fp, p - x_{n_k}\right\rangle = \langle Fp, p - z\rangle \\ &\le 0. \end{aligned} \tag{3.26}$$

From Remark 3.1 and (3.26), one has

$$\limsup_{k\to\infty}\left[\frac{2\gamma}{\eta}\left\langle Fp, p - x_{n_k+1}\right\rangle + \frac{3M\theta_{n_k}}{\sigma_{n_k}\eta}\|x_{n_k} - x_{n_k-1}\|\right] \le 0. \tag{3.27}$$

Hence, combining Claim 3, Condition (C5) and (3.27), in the light of Lemma 2.2, one concludes that $\lim_{n\to\infty}\|x_n - p\| = 0$, i.e., $x_n \to p$ as $n \to \infty$. We have thus proved the theorem. □

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

Corollary 3.1 *Suppose that Conditions (C1)–(C3) and (C5) holds. Let mapping $f : \mathcal{H} \to \mathcal{H}$ be ρ-contraction with $\rho \in [0, \sqrt{5} - 2)$. Take $\theta > 0$, $\chi_1 > 0$, $\mu \in (0, 1)$, $\phi \in (0, 2)$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and the iterative sequence $\{x_n\}$ be generated by the following*

$$\begin{cases} u_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(u_n - \chi_n A u_n), \\ z_n = u_n - \phi\delta_n d_n, \\ x_{n+1} = (1 - \sigma_n)z_n + \sigma_n f(z_n), \end{cases} \tag{3.28}$$

where $\{\theta_n\}$, $\{\delta_n\}$ and $\{\chi_n\}$ are defined in (3.1), (3.2) and (3.3), respectively. Then the iterative sequence $\{x_n\}$ formed by Algorithm (3.28) converges to p in norm, where $p = P_{\mathrm{VI}(C,A)}(f(p))$.

Remark 3.3 Our Algorithm (3.28) improves many of the results in the literature for solving variational inequalities based on the following observation.

- The proposed Algorithm (3.28) obtains a strong convergence theorem in an infinite-dimensional Hilbert space, while the methods presented in [8,9,14,17–21,31] can only obtain weak convergence theorems.
- Notice that our Algorithm (3.28) only needs to compute the projection on the feasible set once in each iteration, which improves some results in the literature (see, e.g., [17,24]) that require computing the projection twice.
- It is known that the choice of step size has an important impact on the convergence speed of the algorithm used. The suggested iterative scheme (3.28) uses a non-monotonic step size criterion, which makes it better than the algorithms in [12,14,29,30] that use non-increasing step sizes, the algorithm in [8] that use Armijo-type step sizes, and the algorithms in [9,11] that use fixed step sizes.
- Our Algorithm (3.28) is designed to solve pseudo-monotone variational inequality problems, which improves the methods proposed in the literature (see, e.g., [8,9,11,12,15]) for solving monotone variational inequalities. In addition, our Algorithm (3.28) also adds inertial effects, which can accelerate the convergence speed of the algorithm without inertial terms.

In summary, the Algorithm (3.28) proposed in this paper is useful and efficient.

4 Numerical examples

In this section, we provide some computational tests to demonstrate the numerical behavior of the proposed Algorithm 3.1, and also to compare it with the Algorithm 1 introduced by Thong et al. [29] and the Algorithm 3.2 suggested by Tan, Liu and Qin [30]. The parameters of all algorithms are set as follows.

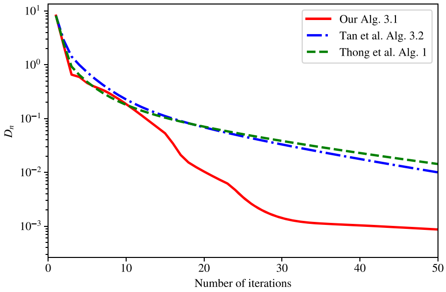
- In the proposed Algorithm 3.1, we set $\theta = 0.2$, $\epsilon_n = \frac{100}{(n+1)^2}$, $\mu = 0.4$, $\chi_1 = 0.6$, $\xi_n = \frac{1}{(n+1)^{1.1}}$, $\phi = 1$, $\sigma_n = \frac{1}{n+1}$ and $\gamma = \frac{1.7\psi}{L_F^2}$.
- In the Algorithm 1 introduced by Thong et al. [29], we choose $\mu = 0.4$, $\chi_1 = 0.6$, $\phi = 1$, $\sigma_n = \frac{1}{n+1}$ and $\gamma = \frac{1.7\psi}{L_F^2}$.
- In the Algorithm 3.2 suggested by Tan et al. [30], we take $\theta = 0.2$, $\epsilon_n = \frac{100}{(n+1)^2}$, $\mu = 0.4$, $\chi_1 = 0.6$, $\sigma_n = \frac{1}{n+1}$ and $\gamma = \frac{1.7\psi}{L_F^2}$.

Notice that these algorithms mentioned above do not require the prior information about the Lipschitz constant of the mapping. In addition, we apply the offered Algorithm (3.28) to solve optimal control problems and compare it with some strongly convergent algorithms in the literature. All the programs were implemented in MATLAB 2018a on a Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz computer with RAM 8.00 GB.

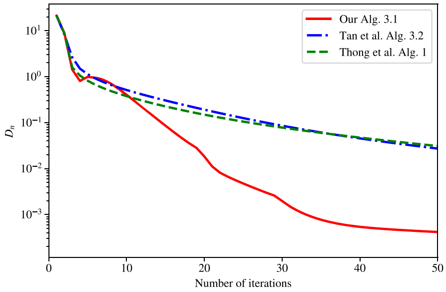
Example 4.1 Consider a mapping $F : \mathbb{R}^m \to \mathbb{R}^m$ $(m = 5)$ of the form $F(x) = Mx + q$, where

$$M = BB^{\mathsf{T}} + D + K,$$

and B is a $m \times m$ matrix with their entries being generated in $(0, 1)$, D is a $m \times m$ skew-symmetric matrix with their entries being generated in $(-1, 1)$, K is a $m \times m$ diagonal matrix,



(a) $x_0 = x_1 = 10rand(5,1)$



(b) $x_0 = x_1 = 20rand(5,1)$

Fig. 1 The behavior of all algorithms with different initial values in Example 4.1

whose diagonal entries are positive in (0, 1) (so M is positive semidefinite), $q \in \mathbb{R}^m$ is a vector with entries being generated in (0, 1). It is clear that F is L_F-Lipschitz continuous and ψ-strongly monotone with $L_F = \max\{\mathrm{eig}(M)\}$ and $\psi = \min\{\mathrm{eig}(M)\}$, where $\mathrm{eig}(M)$ represents all eigenvalues of M. Next, we consider the following fractional programming problem:

$$\begin{aligned} &\min f(x) = \frac{x^\mathsf{T} Qx + a^\mathsf{T} x + a_0}{b^\mathsf{T} x + b_0}, \\ &\text{subject to } x \in C := \left\{x \in \mathbb{R}^5 : b^\mathsf{T} x + b_0 > 0\right\}, \end{aligned}$$

where

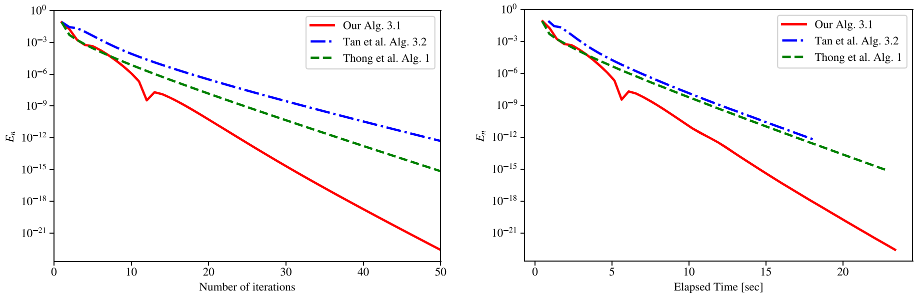
$$Q = \begin{bmatrix} 5 & -1 & 2 & 0 & 2 \\ -1 & 6 & -1 & 3 & 0 \\ 2 & -1 & 3 & 0 & 1 \\ 0 & 3 & 0 & 5 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{bmatrix}, \ a = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \\ 1 \end{bmatrix}, \ b = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \ a_0 = -2, b_0 = 20,$$

It is easy to check that Q is symmetric and positive definite in $\mathbb{R}^5$ and hence f is pseudo-convex on $C = \left\{x \in \mathbb{R}^5 : b^\mathsf{T} x + b_0 > 0\right\}$. Let

$$A(x) := \nabla f(x) = \frac{\left(b^\mathsf{T} x + b_0\right)(2Qx + a) - b\left(x^\mathsf{T} Qx + a^\mathsf{T} x + a_0\right)}{\left(b^\mathsf{T} x + b_0\right)^2}.$$

It is known that A is pseudomonotone and Lipschitz continuous (see [40] for more details). Notice that estimating the Lipschitz constant for A is not easy.

Since we do not know the exact solution to the problem, we use $D_n = \|u_n - P_C(u_n - \chi_n A u_n)\|$ to measure the error of the n-th iteration for the proposed Algorithm 3.1 and the Algorithm 3.2 suggested by Tan et al. [30], and use $D_n = \|x_n - P_C(x_n - \chi_n A x_n)\|$ to measure the error of the n-th iteration for the Algorithm 1 introduced by Thong et al. [29]. By the property of the solution of the variational inequality problem (VIP), we know that if $D_n \to 0$ then u_n (or x_n) tends to the solution of the problem. The maximum number of iterations 200 as the stopping criterion for all algorithms. Numerical results of all algorithms with two different initial values are reported in Fig. 1.



(a) Comparison of the number of iterations **(b)** Comparison of the elapsed time

Fig. 2 The behavior of all algorithms at $x_0 = x_1 = \cos(t)$ in Example 4.2

Table 1 Numerical results of all algorithms with different initial values in Example 4.2

Algorithms	$x_0 = \cos(t)$		$x_0 = t^4$		$x_0 = e^t + t^2$		$x_0 = 2^t$	
	E_n	CPU (s)	E_n	CPU (s)	E_n	CPU (s)	E_n	CPU (s)
Our Alg. 3.1	2.59E−23	23.39	9.09E−24	23.23	1.23E−22	26.47	9.39E−24	23.58
Tan et al. Alg. 3.2	4.95E−13	18.26	1.87E−13	18.19	8.77E−13	19.91	7.69E−13	18.58
Thong et al. Alg. 1	6.96E−16	22.94	4.16E−16	23.07	1.26E−15	25.07	1.00E−15	23.37

Example 4.2 We consider an example that appears in the infinite-dimensional Hilbert space $\mathcal{H} = L^2[0, 1]$ with inner product

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

and induced norm

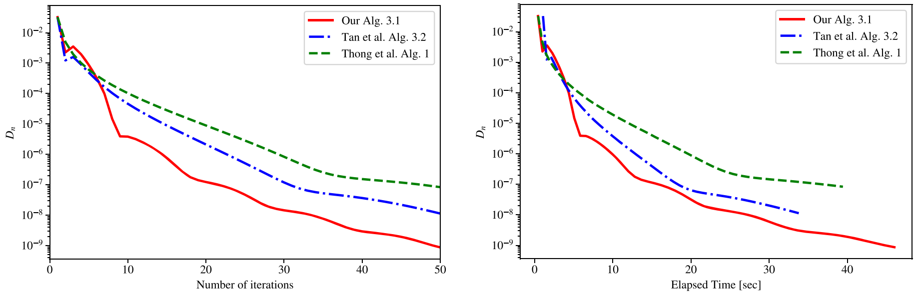
$$\|x\| = \left(\int_0^1 |x(t)|^2\mathrm{d}t\right)^{1/2}, \quad \forall x \in \mathcal{H}.$$

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 as follows $C := \{x \in \mathcal{H} : \|x\| \le r\}$. Let the operator $A : \mathcal{H} \to \mathcal{H}$ be given by

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

Note that operator A is pseudomonotone rather than monotone and Lipschitz continuous (see [41, Section 4]). Let $F : \mathcal{H} \to \mathcal{H}$ be an operator defined by $(Fx)(t) = 0.5x(t),\ t \in [0, 1]$. It is easy to see that F is 0.5-strongly monotone and 0.5-Lipschitz continuous. For the experiment, we choose $R = 1.5, r = 1, k = 1.1$. The solution of this problem is $x^*(t) = 0$. The maximum number of iterations 50 as the stopping criterion for all algorithms. Figure 2 and Table 1 show the numerical behaviors of $E_n = \|x_n(t) - x^*(t)\|$ generated by all algorithms under four different initial values $x_0(t) = x_1(t)$.

Example 4.3 Let $\mathcal{H} = L^2([0, 1])$ be an infinite-dimensional Hilbert space and have the same inner product and induced norm as in Example 4.2. Assume that the feasible set is given by



(a) Comparison of the number of iterations **(b)** Comparison of the elapsed time

Fig. 3 The behavior of all algorithms at $x_0 = x_1 = t^4$ in Example 4.3

Table 2 Numerical results of all algorithms with different initial values in Example 4.3

Algorithms	$x_0 = \cos(t)$		$x_0 = t^4$		$x_0 = e^t + t^2$		$x_0 = 2^t$	
	D_n	CPU (s)	D_n	CPU (s)	D_n	CPU (s)	D_n	CPU (s)
Our Alg. 3.1	8.56E−08	106.38	8.69E−10	46.02	4.03E−07	61.09	5.16E−07	61.03
Tan et al. Alg. 3.2	1.10E−06	82.91	1.13E−08	33.79	1.08E−06	48.94	1.07E−06	48.02
Thong et al. Alg. 1	5.40E−06	93.74	8.33E−08	39.44	4.98E−06	58.94	5.36E−06	54.29

$C = \{x \in \mathcal{H} : \|x\| \leq 2\}$. Define a mapping $h : C \to \mathbb{R}$ by

$$h(x) = \frac{1}{1 + \|x\|^2}.$$

Recall that the Volterra integration operator $V : \mathcal{H} \to \mathcal{H}$ is given by

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

Now, we define the mapping $A : C \to \mathcal{H}$ as follows:

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

Let $F : \mathcal{H} \to \mathcal{H}$ be an operator defined by $(Fx)(t) = 0.5x(t)$, $t \in [0, 1]$. Notice that the operator A is pseudo-monotone but not monotone and Lipschitz continuous (see [42, Example 4.2]). In addition, it is worth noting that the Lipschitz constant of A is not easy to estimate, and the solution of the variational inequality problem with A and C given above is unknown. We use $D_n = \|u_n(t) - P_C(u_n(t) - \chi_n A u_n(t))\|$ to measure the error of the n-th iteration for the proposed Algorithm 3.1 and the Algorithm 3.2 suggested by Tan et al. [30], and use $D_n = \|x_n(t) - P_C(x_n(t) - \chi_n A x_n(t))\|$ to measure the error of the nth iteration for the Algorithm 1 introduced by Thong et al. [29]. The numerical results of all algorithms with four starting points $x_0(t) = x_1(t)$ are shown in Fig. 3 and Table 2.

Remark 4.1 From Examples 4.1–4.3, we have the following observations:

(i) As shown in Figs. 1, 2, 3 and Tables 1 and 2, it can be seen that the proposed Algorithm 3.1 converges quickly, and it has a faster convergence speed and higher accuracy than the previously known ones in [29,30] under the same stopping conditions. These results are independent of the size of the dimension and the choice of initial values. Therefore, our Algorithm 3.1 is efficient and robust.

(ii) It is noticed from Tables 1 and 2 that the proposed algorithm requires more time while achieving higher accuracy. The reason for this phenomenon is that the suggested Algorithm 3.1 needs to spend extra time to calculate the values of u_n, d_n and δ_n in an infinite-dimensional Hilbert space (see Examples 4.2 and 4.3).

(iii) In [40], Boţ, Csetnek and Vuong showed that the operator A in Example 4.1 is L_A-Lipschitz continuous with $L_A \approx 148.68$. In most cases, the prior knowledge of the Lipschitz constant for the problem under consideration is usually unknown, in which case the fixed stepsize algorithm proposed by Thong and Hieu [25, Algorithm 3.1] will not be available because it needs the Lipschitz constant as an input parameter. However, the offered method can work well since it does not require the prior information of the Lipschitz constant of the mapping.

(iv) It should be noted that the operators A in Examples 4.2 and 4.3 are pseudomonotone rather than monotone. The Algorithm 1 proposed by Hieu and Moudafi [27] for solving the bilevel monotone variational inequality problem will not be applicable in those cases.

Next, we use the proposed Algorithm (3.28) to solve the (VIP) that appears in optimal control problems. We recommend readers to refer to [3,30,43,44] for detailed description of the problem. We compare the suggested iterative scheme (3.28) with some strongly convergent algorithms in the literature. Two methods used to compare here are the Algorithm (31) (in short, TLDCR Alg. (31)) introduced by Thong et al. [29] and the Algorithm (3.39) (in short, TLQ Alg. (3.39)) proposed by Tan et al. [30]. The parameters of all algorithms are set as follows.

- In the proposed Algorithm (3.28), we set $N = 100$, $\theta = 0.01$, $\epsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\mu = 0.1$, $\chi_1 = 0.4$, $\xi_n = \frac{0.1}{(n+1)^{1.1}}$, $\phi = 1.5$, $\sigma_n = \frac{10^{-4}}{n+1}$ and $f(x) = 0.1x$.
- In the TLDCR Alg. (31), we choose $N = 100$, $\mu = 0.1$, $\chi_1 = 0.4$, $\phi = 1.5$ and $\sigma_n = \frac{10^{-4}}{n+1}$.
- In the TLQ Alg. (3.39), we take $N = 100$, $\theta = 0.01$, $\epsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\mu = 0.1$, $\chi_1 = 0.4$, $\sigma_n = \frac{10^{-4}}{n+1}$ and $f(x) = 0.1x$.

The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-1, 1]$. The stopping criterion is either $D_n = \|p_{n+1} - p_n\| \le 10^{-4}$, or maximum number of iterations which is set to 1000.

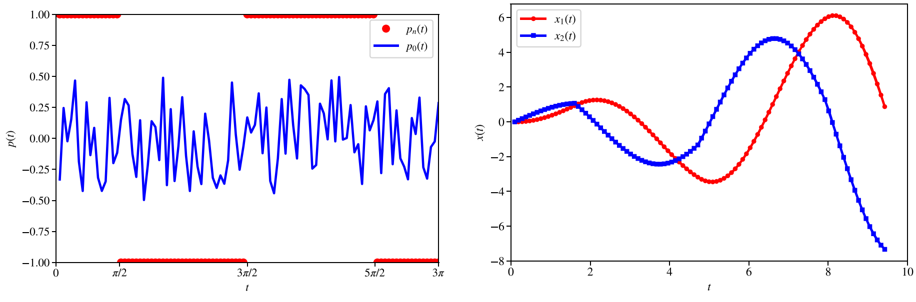
Example 4.4 (Control of a harmonic oscillator, see [45])

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

The exact optimal control of Example 4.4 is known:

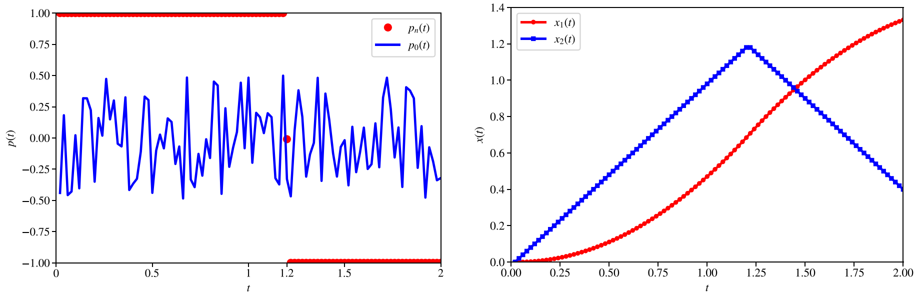
$$
p^*(t) = \begin{cases} 1, & \text{if } t \in [0, \pi/2) \cup (3\pi/2, 5\pi/2)\,; \\ -1, & \text{if } t \in (\pi/2, 3\pi/2) \cup (5\pi/2, 3\pi]. \end{cases}
$$

Figure 4 shows the approximate optimal control and the corresponding trajectories of the stated Algorithm (3.28).



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

Fig. 4 Numerical results of the proposed Algorithm (3.28) for Example 4.4



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

Fig. 5 Numerical results of the proposed Algorithm (3.28) for Example 4.5

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

Example 4.5 (see [46])

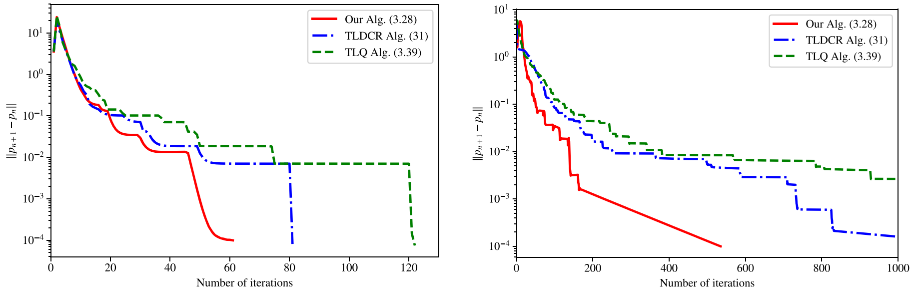
$$\begin{aligned}
\text{minimize} \quad & -x_1(2) + (x_2(2))^2\,, \\
\text{subject to} \quad & \dot{x}_1(t) = x_2(t), \\
& \dot{x}_2(t) = p(t), \quad \forall t \in [0,2], \\
& x_1(0) = 0, \quad x_2(0) = 0, \\
& p(t) \in [-1,1].
\end{aligned}$$

The exact optimal control of Example 4.5 is

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

The approximate optimal control and the corresponding trajectories of the suggested Algorithm (3.28) are plotted in Fig. 5.

Finally, we compare the offered Algorithm (3.28) with TLQ Alg. (3.39) and TLDCR Alg. (31) for Examples 4.4 and 4.5. Figure 6 presents the numerical behavior of the error estimate $\|p_{n+1} - p_n\|$ with respect to the number of iterations for all algorithms. In addition, the number of terminated iterations and the execution time of all algorithms are shown in Table 3.



(a) Example 4.4 **(b)** Example 4.5

Fig. 6 Error estimates of all algorithms for Examples 4.4 and 4.5

Table 3 Numerical results of all algorithms in Examples 4.4 and 4.5

Algorithms	Example 4.4			Example 4.5		
	Iter.	CPU (s)	D_n	Iter.	CPU (s)	D_n
Our Alg. (3.28)	60	0.029916	9.901E-05	534	0.18632	9.944E-05
TLDCR Alg. (31)	80	0.036433	8.096E-05	1000	0.33392	1.584E-04
TLQ Alg. (3.39)	121	0.078416	7.437E-05	1000	0.32060	2.647E-03

Remark 4.2 We draw the following observations from Examples 4.4 and 4.5.

(i) The offered Algorithm (3.28) can be applied to solve optimal control problems, and it performs well when the terminal function is linear or nonlinear (cf. Figs. 4, 5).
(ii) As shown in Fig. 6 and Table 3, the proposed Algorithm (3.28) performs better when the terminal function is linear than when it is nonlinear, i.e., it requires less execution time and the number of termination iterations in the case where the terminal function is linear. Moreover, the proposed Algorithm (3.28) outperforms the existing methods in the literature [29,30], in other words, the presented Algorithm (3.28) converges faster than the others in [29,30] under the same stopping criterion.

5 Conclusions

In this paper, we presented a new inertial projection method to discover the solution of a bilevel pseudomonotone variational inequality problem in an infinite-dimensional Hilbert space. The suggested iterative scheme is constructed by the inertial method, the projection and contraction method and the hybrid steepest descent method. A new non-monotonic step size that does not contain any linesearch process is embedded into the proposed algorithm so that it can work well without knowing the prior knowledge of the Lipschitz constant of the mapping. Finally, the stated theoretical results are verified by several preliminary numerical experiments. The results obtained in this paper improved and generalized some relevant known algorithms in the field.

Acknowledgements The authors would like to thank the anonymous referees for their constructive comments, which improved the quality of earlier version of the manuscript.

Declarations

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

References

1. Sahu, D.R., Yao, J.C., Verma, M., Shukla, K.K.: Convergence rate analysis of proximal gradient methods with applications to composite minimization problems. Optimization **70**, 75–100 (2021)
2. Cuong, T.H., Yao, J.C., Yen, N.D.: Qualitative properties of the minimum sum-of-squares clustering problem. Optimization **69**, 2131–2154 (2020)
3. Hieu, D.V., Anh, P.K., Muu, L.D., Strodiot, J.J.: Iterative regularization methods with new stepsize rules for solving variational inclusions. J. Appl. Math. Comput. (2021). https://doi.org/10.1007/s12190-021-01534-9
4. Hieu, D.V., Reich, S., Anh, P.K., Ha, N.H.: A new proximal-like algorithm for solving split variational inclusion problems. Numer. Algorithms (2021). https://doi.org/10.1007/s11075-021-01135-4
5. Cho, S.Y.: A monotone Bregman projection algorithm for fixed point and equilibrium problems in a reflexive Banach space. Filomat **34**, 1487–1497 (2020)
6. Shehu, Y., Yao, J.C.: Rate of convergence for inertial iterative method for countable family of certain quasi-nonexpansive mappings. J. Nonlinear Convex Anal. **21**, 533–541 (2020)
7. Cho, S.Y.: A convergence theorem for generalized mixed equilibrium problems and multivalued asymptotically nonexpansive mappings. J. Nonlinear Convex Anal. **21**, 1017–1026 (2020)
8. Dong, Q.L., Jiang, D., Gibali, A.: A modified subgradient extragradient method for solving the variational inequality problem. Numer. Algorithms **79**, 927–940 (2018)
9. 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)
10. Hieu, D.V., Reich, S.: Two Bregman projection methods for solving variational inequalities. Optimization (2020). https://doi.org/10.1080/02331934.2020.1836634
11. Hieu, D.V., Muu, L.D., Duong, H.N., Thai, B.H.: Strong convergence of new iterative projection methods with regularization for solving monotone variational inequalities in Hilbert spaces. Math. Methods Appl. Sci. **43**, 9745–9765 (2020)
12. Hieu, D.V., Anh, P.K., Muu, L.D.: Strong convergence of subgradient extragradient method with regularization for solving variational inequalities. Optim. Eng. https://doi.org/10.1007/s11081-020-09540-9 (2020)
13. Hieu, D.V., Muu, L.D., Duong, H.N.: Iterative regularization methods for solving equilibrium problems. Int. J. Comput. Math (2021). https://doi.org/10.1080/00207160.2021.1906421
14. 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)
15. Tan, B., Fan, J., Li, S.: Self-adaptive inertial extragradient algorithms for solving variational inequality problems. Comput. Appl. Math. **40**, Article ID 19 (2021)
16. Yamada, I.: The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. Inherently Parallel Algorithms Feasibility Optim. Appl. **8**, 473–504 (2001)
17. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Èkonom. i Mat. Metody **12**, 747–756 (1976)
18. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35**, 69–76 (1997)
19. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)
20. 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)
21. Malitsky, Y.V., Semenov, V.V.: An extragradient algorithm for monotone variational inequalities. Cybern. Syst. Anal. **50**, 271–277 (2014)
22. Liu, Y., Kong, H.: The new extragradient method extended to equilibrium problems. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM **113**, 2113–2126 (2019)
23. Moharami, R., Eskandani, G.Z.: An extragradient algorithm for solving equilibrium problem and zero point problem in Hadamard spaces. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM **114**, Article ID 152 (2020)

24. Rehman, H.u., Gibali, A., Kumam, P., Sitthithakerngkiet, K.: Two new extragradient methods for solving equilibrium problems. Rev. R. Acad. Cienc. Exactas Fís. Nat. Ser. A Math. RACSAM **115**, Article ID 75 (2021)
25. 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)
26. Ceng, L.C.: Two inertial linesearch extragradient algorithms for the bilevel split pseudomonotone variational inequality with constraints. J. Appl. Numer. Optim. **2**, 213–233 (2020)
27. Hieu, D.V., Moudai, A.: Regularization projection method for solving bilevel variational inequality problem. Optim. Lett. (2020) **15**, 205–229 (2021)
28. 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)
29. Thong, D.V., Li, X.H., Dong, Q.L., Cho, Y.J., Rassias, T.M.: A projection and contraction method with adaptive step sizes for solving bilevel pseudo-monotone variational inequality problems. Optimization (2020). https://doi.org/10.1080/02331934.2020.1849206
30. Tan, B., Liu, L., Qin, X.: Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems. Jpn. J. Ind. Appl. Math. **38**, 519–543 (2021)
31. Dong, Q.L., Cho, Y.J., Zhong, L.L., Rassias, T.M.: Inertial projection and contraction algorithms for variational inequalities. J. Glob. Optim. **70**, 687–704 (2018)
32. Gibali, A., Shehu, Y.: An efficient iterative method for finding common fixed point and variational inequalities in Hilbert spaces. Optimization **68**, 13–32 (2019)
33. Shehu, Y., Gibali, A.: New inertial relaxed method for solving split feasibilities. Optim. Lett. **15**, 2109–2126 (2021)
34. Tan, B., Li, S.: Strong convergence of inertial Mann algorithms for solving hierarchical fixed point problems. J. Nonlinear Var. Anal. **4**, 337–355 (2020)
35. Tan, B., Cho, S.Y.: An inertial Mann-like algorithm for fixed points of nonexpansive mappings in Hilbert spaces. J. Appl. Numer. Optim. **2**, 335–351 (2020)
36. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. **75**, 742–750 (2012)
37. Liu, H., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput. Optim. Appl. **77**, 491–508 (2020)
38. 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)
39. Muu, L.D., Quy, N.V.: On existence and solution methods for strongly pseudomonotone equilibrium problems. Vietnam J. Math. **43**, 229–238 (2015)
40. Boţ, 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)
41. Hieu, D.V., Cho, Y.J., Xiao, Y.B., Kumam, P.: Modified extragradient method for pseudomonotone variational inequalities in infinite dimensional Hilbert spaces. Vietnam J. Math. (2020). https://doi.org/10.1007/s10013-020-00447-7
42. Tan, B., Cho, S.Y.: Inertial extragradient methods for solving pseudomonotone variational inequalities with non-Lipschitz mappings and their optimization applications. Appl. Set Valued Anal. Optim. **3**, 165–192 (2021)
43. 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)
44. 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)
45. Pietrus, A., Scarinci, T., Veliov, V.M.: High order discrete approximations to Mayers problems for linear systems. SIAM J. Control Optim. **56**, 102–119 (2018)
46. Bressan, B., Piccoli, B.: Introduction to the Mathematical Theory of Control. AIMS series on applied mathematics, San Francisco (2007)