Communications in Nonlinear Science and Numerical Simulation 107 (2022) 106160

Contents lists available at ScienceDirect

Communications in Nonlinear Science and Numerical Simulation

journal homepage: www.elsevier.com/locate/cnsns

Two adaptive modified subgradient extragradient methods for bilevel pseudomonotone variational inequalities with applications

Bing Tan [a], Sun Young Cho [b,*]

[a] Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, China
[b] Department of Human Health Care, Gyeongsang National University, Jinjiu, Republic of Korea

ARTICLE INFO

Article history:
Received 19 August 2021
Received in revised form 3 November 2021
Accepted 23 November 2021
Available online 7 December 2021

MSC:
47J20
47J25
47J30
68W10
65K15

Keywords:
Bilevel variational inequality
Inertial method
Extragradient method
Adaptive stepsize
Pseudomonotone operator

ABSTRACT

We consider the bilevel variational inequality problem with a pseudomonotone operator in real Hilbert spaces and investigate two modified subgradient extragradient methods with inertial terms. Our first scheme requires the operator to be Lipschitz continuous (the Lipschitz constant does not need to be known) while the second one only requires it to be uniformly continuous. The proposed methods employ two adaptive stepsizes making them work without the prior knowledge of the Lipschitz constant of the mapping. The strong convergence properties of the iterative sequences generated by the proposed algorithms are obtained under mild conditions. Some numerical tests and applications are given to demonstrate the advantages and efficiency of the stated schemes over previously known ones.

1. Introduction and preliminaries

Throughout the paper, we always assume that C is a nonempty, closed, and convex subset of a real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. We first review the definition of some nonlinear mappings that are relevant to this paper.

Definition 1.1. A mapping $M : \mathcal{H} \to \mathcal{H}$ is said to be

(i) *L-Lipschitz continuous* with $L > 0$ if $\|Mx - My\| \leq L\|x - y\|, \ \forall x, y \in \mathcal{H}$ (If $L \in (0, 1)$ then mapping M is called *contraction*. In particular, when $L = 1$, mapping M is called *nonexpansive*).
(ii) *α-strongly monotone* if there exists a constant $\alpha > 0$ such that $\langle Mx - My, x - y\rangle \geq \alpha \|x - y\|^2, \ \forall x, y \in \mathcal{H}$.
(iii) *monotone* if $\langle Mx - My, x - y\rangle \geq 0, \ \forall x, y \in \mathcal{H}$.
(iv) *pseudomonotone* if $\langle Mx, y - x\rangle \geq 0 \Longrightarrow \langle My, y - x\rangle \geq 0, \ \forall x, y \in \mathcal{H}$.

* Corresponding author.
E-mail addresses: bingtan72@gmail.com (B. Tan), chosy@gnu.ac.kr (S.Y. Cho).

https://doi.org/10.1016/j.cnsns.2021.106160

(v) *sequentially weakly continuous* if for each sequence $\{x_n\}$ converging weakly to x, one has the weak convergence of $\{Mx_n\}$ to Mx.

Note that the relation (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (iv) holds but its converse is generally incorrect (see, e.g., Example 3.2 in our Section 3).

In this paper, we focus on the bilevel variational inequality problem (shortly, BVIP). BVIPs contain a number of nonlinear optimization problems and have many potential applications (see [1] for more details). Let $F : C \to \mathcal{H}$ and $M : C \to \mathcal{H}$ be two operators. Recall that the BVIP is described as follows:

$$\text{find } x^* \in \Omega \text{ such that } \left\langle Fx^*, y - x^* \right\rangle \geq 0, \quad \forall y \in \Omega \,, \tag{BVIP}$$

where Ω denotes the set of all solutions of the following variational inequality problem (shortly, VIP):

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

It is known that VIPs provide a general and useful framework for solving engineering problems, image processing, data sciences, electronics, and other fields; see, e.g., [2–9] and the references therein. Thus, numerical methods for studying variational inequalities have attracted numerous interests among researchers. For some recent advances in variational inequalities, we recommend the readers to refer to [10].

Next we state some algorithms for solving the (VIP) and the (BVIP), and these motivate us to develop several new efficient iterative schemes. One of the most popular projection-type methods for solving variational inequality problems is the extragradient method (for short, EGM) proposed by Korpelevich [11]. It is known that the EGM obtains weak convergence in infinite-dimensional Hilbert spaces if the operator M involved is monotone and Lipschitz continuous. Note that the EGM is a two-step iterative scheme and requires the projection onto the feasible set to be computed twice in each iteration, which increases the computational burden of the method especially when the projection onto the feasible set is difficult to evaluate. To overcome this drawback, a large number of variants of the EGM have been introduced to solve variational inequalities in finite- and infinite-dimensional spaces; see, for example, [12–15]. One of the methods to be highlighted is the subgradient extragradient method (for short, SEGM) proposed by Censor, Gibali, and Reich [13]. The SEGM replaces the projection onto the feasible set in the second step of the EGM with the projection onto a special half-space. This modification improves the computational efficiency of the EGM due to the fact that the projection onto the half-space can be computed explicitly. Moreover, the weak convergence of the SEGM is established in an infinite-dimensional Hilbert space. Recently, Dong, Jiang, and Gibali [16] introduced a modified subgradient extragradient method (shortly, MSEGM) inspired by the SEGM and the projection contraction method (shortly, PCM) [14] to approximate the solution of the (VIP). The basic idea of the MSEGM is to improve the stepsize in the second step of the SEGM. They provided primary numerical experiments to demonstrate the computational efficiency of the MSEGM compared to the SEGM and the PCM.

Another issue of interest in the computational efficiency of the algorithm is the step size. A common feature enjoyed by the EGM, the SEGM, and some of their variant forms is that the update of the stepsize requires the prior information of the Lipschitz constant of the mapping involved. However, the Lipschitz constant is not readily available for practical applications or estimating a suitable range requires more computational burden. Recently, some adaptive methods have been offered to solve the variational inequality problem when the Lipschitz constant is unknown; see, e.g., [17–19] and the references therein. However, the schemes proposed in [17–19] generate a non-increasing sequence of stepsizes, which will further affect the computational efficiency of the algorithms used. Recently, Liu and Yang [20] introduced a new stepsize criterion that generates a non-monotonic sequence of stepsizes. Their numerical experiments demonstrate the performance of the algorithms with this new stepsize. On the other hand, there are some mappings in real-world problems that do not satisfy the Lipschitz continuity condition, which will lead to the failure of those algorithms that require the operator to be Lipschitz continuous. To overcome this shortcoming, some methods with Armijo-type stepsizes are proposed for solving monotone and uniformly continuous VIPs (see, e.g., [21,22]) and pseudomonotone and uniformly continuous VIPs (see, e.g., [23–26]).

In recent years, inertial terms have attracted the interest and research of scholars as a technique to accelerate the convergence speed of algorithms. A common feature of inertial-type algorithms is that the next iteration depends on the combination of the previous two iterations (see [27,28] for more details). This small change greatly improves the computational efficiency of inertial-type algorithms. Recently, many researchers have constructed a large number of inertial-type algorithms to solve variational inequality problems, fixed point problems, equilibrium problems, split feasibility problems, and other optimization problems; see, e.g., [29–32] and the references therein. The computational efficiency of these inertial-type algorithms was demonstrated by a number of computational tests and applications.

We next state some algorithms and difficulties in the literature for solving the BVIP, which will lead us to the motivation of our research in this paper. Yamada [33] introduces a new iterative algorithm (now known as the hybrid steepest descent method) to solve the bilevel problem described as a variational inequality problem restricted by a fixed point problem. Note that the method does not contain any projection step and obtains strong convergence in infinite-dimensional Hilbert spaces under some suitable conditions. It is known that the variational inequality problem and the fixed point problem are interconvertible and thus we can use the hybrid steepest descent method to approximate the solutions of BVIPs. Recently, a number of numerical algorithms that based on the hybrid steepest descent method have been proposed for solving the bilevel monotone variational inequalities (see, e.g., [34,35]) and the bilevel pseudomonotone variational inequalities (see, e.g., [36–39]). Very recently, Thong and Hieu [37] introduced a new modified subgradient extragradient method stimulated by the MSEGM of Dong et al. [16] for solving the BVIP involving a pseudomonotone operator. However, the step size in their scheme is bounded by the inverse of the Lipschitz constant of the mapping, which means that the Lipschitz constant of the mapping must be entered into the iterative process as a known parameter. To overcome this difficulty, some adaptive algorithms for solving BVIPs have been proposed by scholars (see, e.g., [35,38,39]). These methods can work without the prior knowledge of the Lipschitz constant of the mapping. Another thing we need to point out is that the operator M of these algorithms is required to satisfy Lipschitz continuity. This condition is very strong will cause that they will not be available in some cases.

A natural question is how to modify the algorithm suggested by Thong and Hieu [37] so that they can work adaptively and obtain a faster convergence speed. To answer this question, in this paper we present two modified inertial extragradient-type methods to solve the bilevel pseudomonotone variational inequality problem. Our first scheme requires that the operator M be Lipschitz continuous while the second one only requires that it be uniformly continuous. The strong convergence of the proposed methods is established under some mild conditions imposed on the parameters. Some numerical experiments and applications are given to verify the theoretical results. The algorithms suggested in this paper improve some known results in the literature for solving bilevel variational inequalities [34–39] and variational inequalities [13,21,22].

To end this section, we review the following two lemmas that will be used in the convergence analysis of the algorithms.

Lemma 1.1 ([33]). *Let $\gamma > 0$ and $\alpha \in (0, 1]$. Let $F : \mathcal{H} \to \mathcal{H}$ be a β-strongly monotone and L-Lipschitz continuous mapping. Associating with a nonexpansive mapping $T : \mathcal{H} \to \mathcal{H}$, define a mapping $T^\gamma : \mathcal{H} \to \mathcal{H}$ by $T^\gamma x = (I - \alpha\gamma F)(Tx), \forall x \in \mathcal{H}$. Then, T^γ is a contraction provided $\gamma < \frac{2\beta}{L^2}$, that is,*

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

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

Lemma 1.2 ([40]). *Let $\{p_n\}$ be a positive sequence, $\{q_n\}$ be a sequence of real numbers, and $\{\alpha_n\}$ be a sequence in $(0, 1)$ such that $\sum_{n=1}^{\infty} \alpha_n = \infty$. Assume that*

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

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

2. The algorithms and their convergence analysis

In this section, we introduce two adaptive modified subgradient extragradient methods for finding the solutions of the bilevel pseudomonotone variational inequality problem (BVIP) in real Hilbert spaces. In the sequel, we use the notation $x_n \to x$ (resp., $x_n \rightharpoonup x$) to denote the strong convergence (resp., weak convergence) of the sequence $\{x_n\}$ to x and use $P_C : \mathcal{H} \to C$ to represent the metric projection from $\mathcal{H}$ onto C, defined as $P_C(x) := \arg\min\{\|x - y\|, \ y \in C\}$.

2.1. Algorithm for Lipschitz continuous operators

In this subsection, we present an adaptive algorithm for solving the (BVIP) with a pseudomonotone and Lipschitz continuous operator. Suppose that the following assumptions (A1)–(A5) hold for our Algorithm 2.1.

(A1) The feasible set C is a nonempty, closed, and convex subset of a real Hilbert space $\mathcal{H}$.
(A2) The solution set of the problem (VIP) is nonempty, that is, $\Omega \neq \emptyset$.
(A3) The mapping $F : \mathcal{H} \to \mathcal{H}$ is L_F-Lipschitz continuous and β-strongly monotone on $\mathcal{H}$.
(A4) The operator $M : \mathcal{H} \to \mathcal{H}$ is pseudomonotone, L_M-Lipschitz continuous on $\mathcal{H}$ and the operator $M : \mathcal{H} \to \mathcal{H}$ satisfies the following condition

$$\text{whenever } \{x_n\} \subset C, \ x_n \rightharpoonup z, \ \text{one has } \|Mz\| \leq \liminf_{n\to\infty} \|Mx_n\| . \tag{2.1}$$

(A5) Assume $\{\xi_n\}$ and $\{\epsilon_n\}$ are two non-negative positive sequences such that $\sum_{n=1}^{\infty} \xi_n < \infty$ and $\lim_{n\to\infty} \frac{\epsilon_n}{\alpha_n} = 0$, where $\{\alpha_n\} \subset (0, 1)$ satisfies $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$.

We now state the first scheme in Algorithm 2.1.

Algorithm 2.1 Adaptive-type inertial modified subgradient extragradient method for (BVIP).

Initialization: Take $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0, 1)$, $\delta \in (0, 2/\mu)$, $\tau \in (\delta/2, 1/\mu)$, $\gamma \in (0, 2\beta/L_F^2)$. Select three sequences $\{\xi_n\}$, $\{\epsilon_n\}$, and $\{\alpha_n\}$ to satisfy Assumption (A5). Let $x_0, x_1 \in \mathcal{H}$ be arbitrary.
Iterative Steps: Given the iterates x_{n-1} and x_n $(n \geq 1)$, calculate x_{n+1} as follows:
Step 1. Compute $w_n = x_n + \theta_n(x_n - x_{n-1})$, where

$$\theta_n = \begin{cases} \min\left\{\dfrac{\epsilon_n}{\|x_n - x_{n-1}\|}, \theta\right\}, & \text{if } x_n \neq x_{n-1}; \\ \theta, & \text{otherwise.} \end{cases} \tag{2.2}$$

Step 2. Compute $y_n = P_C(w_n - \tau\lambda_n M w_n)$, where the next step size λ_{n+1} is updated by

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\,\|w_n - y_n\|}{\|Mw_n - My_n\|}, \lambda_n + \xi_n\right\}, & \text{if } Mw_n \neq My_n; \\ \lambda_n + \xi_n, & \text{otherwise.} \end{cases} \tag{2.3}$$

Step 3. Compute $z_n = P_{T_n}(w_n - \delta\lambda_n\chi_n My_n)$, where the half-space T_n and χ_n are defined by

$$\begin{aligned} &T_n := \{x \in \mathcal{H} \mid \langle w_n - \tau\lambda_n M w_n - y_n, x - y_n\rangle \leq 0\}\,, \\ &\chi_n := \frac{\langle w_n - y_n, c_n\rangle}{\|c_n\|^2}, \quad c_n := w_n - y_n - \tau\lambda_n(Mw_n - My_n)\,. \end{aligned} \tag{2.4}$$

Step 4. Compute $x_{n+1} = z_n - \alpha_n\gamma F z_n$.
Set $n := n + 1$ and go to **Step 1**.

Remark 2.1. We have the following observations for the hypotheses and the Algorithm 2.1.

- Note that the condition (2.1) is used by many recent works on pseudomonotone variational inequalities (see, e.g., [25,41]). It is easy to check that Condition (2.1) is weaker than the sequential weak continuity of the mapping M (see [41, Remark 3.2]). Moreover, it is not necessary to impose Condition (2.1) if mapping M is monotone (see [42]).
- The idea of the step size λ_n defined in (2.3) is derived from [20]. It is worth noting that the step size λ_n generated in Algorithm 2.1 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 λ_1. On the other hand, because of $\sum_{n=1}^{\infty}\xi_n < +\infty$, which implies that $\lim_{n\to\infty}\xi_n = 0$. Consequently, the step size λ_n may not increase when n is large enough. If $\xi_n = 0$, then the step size λ_n in Algorithm 2.1 is similar to the approaches in [17–19].
- Notice that there is an explicit formula to calculate the projection on the half-space T_n (see, e.g., [43]). Thus, the proposed Algorithm 2.1 needs to compute the projection on the feasible set C only once in each iteration.
- The step size used in Algorithm 2.1 to compute z_n is $\delta\lambda_n\chi_n$, which is larger than the step size λ_n used in the subgradient extragradient method proposed by Censor et al. [13] to compute z_n. The method involving this new step size was introduced by Dong et al. in [16]. On the other hand, it is important to emphasize that the step size used to compute y_n in the second step of Algorithm 2.1 is $\tau\lambda_n$, where $\tau \in (\delta/2, 1/\mu)$. This small change can improve the convergence speed of the algorithm with $\tau = 1$ (see Section 3).

The following lemmas are important in the convergence analysis of Algorithm 2.1.

Lemma 2.1. *Suppose that Assumption (A4) holds and the sequence $\{\lambda_n\}$ is generated by (2.3). Then $\lim_{n\to\infty}\lambda_n = \lambda$ and $\lambda \in \left[\min\{\mu/L_M, \lambda_1\}, \lambda_1 + \sum_{n=1}^{\infty}\xi_n\right]$.*

Proof. The proof of this lemma follows as that of Lemma 3.1 in [20] and thus it is omitted. □

Lemma 2.2. *Suppose that Assumptions (A1), (A2), and (A4) hold. Let $\{w_n\}$ and $\{y_n\}$ be two sequences formulated by Algorithm 2.1. If there exists a subsequence $\{w_{n_k}\}$ of $\{w_n\}$ such that $\{w_{n_k}\}$ converges weakly to $z \in \mathcal{H}$ and $\lim_{k\to\infty}\|w_{n_k} - y_{n_k}\| = 0$, then $z \in \Omega$.*

Proof. The proof follows that of Lemma 3.8 in [38]. So it is omitted. □

Lemma 2.3. *If $y_n = w_n$ or $c_n = 0$ in Algorithm 2.1, then $y_n \in \Omega$.*

Proof. From the definition of c_n and (2.3), one has

$$\begin{aligned}\|c_n\| &\geq \|w_n - y_n\| - \tau\lambda_n\|Mw_n - My_n\| \\ &\geq \|w_n - y_n\| - \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\|w_n - y_n\| = \left(1 - \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|\,.\end{aligned}$$

It can be easily proved that $\|c_n\| \leq \left(1 + \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|$. Therefore

$$\left(1 - \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\| \leq \|c_n\| \leq \left(1 + \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|,$$

and thus $w_n = y_n$ if and only if $c_n = 0$. Hence, if $w_n = y_n$ or $c_n = 0$, then $y_n = P_C(y_n - \tau\lambda_n My_n)$. This implies that $y_n \in \Omega$. The proof is completed. □

Lemma 2.4. *Suppose that Assumptions (A1), (A2), and (A4) hold. Let $\{z_n\}$, $\{y_n\}$ and $\{w_n\}$ be three sequences created by Algorithm 2.1. Then, for all $p \in \Omega$,*

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \|w_n - z_n - \frac{\delta}{\tau}\chi_n c_n\|^2 - \frac{\delta}{\tau^2}(2\tau - \delta)\frac{\left(1 - \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}{\left(1 + \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}\|w_n - y_n\|^2.$$

Proof. From the property of projection $\|P_C(x) - P_C(y)\|^2 \leq \langle P_C(x) - P_C(y), x - y\rangle$, $\forall x, y \in \mathcal{H}$ and $p \in \Omega \subset C \subset T_n$, we obtain

$$\begin{aligned}2\|z_n - p\|^2 &= 2\|P_{T_n}(w_n - \delta\lambda_n\chi_n My_n) - P_{T_n}(p)\|^2 \\ &\leq 2\langle z_n - p, w_n - \delta\lambda_n\chi_n My_n - p\rangle \\ &= \|z_n - p\|^2 + \|w_n - \delta\lambda_n\chi_n My_n - p\|^2 - \|z_n - w_n + \delta\lambda_n\chi_n My_n\|^2 \\ &= \|z_n - p\|^2 + \|w_n - p\|^2 + \delta^2\lambda_n^2\chi_n^2\|My_n\|^2 - 2\langle w_n - p, \delta\lambda_n\chi_n My_n\rangle \\ &\quad - \|z_n - w_n\|^2 - \delta^2\lambda_n^2\chi_n^2\|My_n\|^2 - 2\langle z_n - w_n, \delta\lambda_n\chi_n My_n\rangle \\ &= \|z_n - p\|^2 + \|w_n - p\|^2 - \|z_n - w_n\|^2 - 2\langle z_n - p, \delta\lambda_n\chi_n My_n\rangle\,,\end{aligned}$$

which implies that

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \|z_n - w_n\|^2 - 2\delta\lambda_n\chi_n\langle z_n - p, My_n\rangle\,. \tag{2.5}$$

By using $y_n \in C$ and $p \in \Omega$, in the light of (VIP), one has $\langle Mp, y_n - p\rangle \geq 0$. This together with the pseudo-monotonicity of the mapping M yields $\langle My_n, y_n - p\rangle \geq 0$, which means that $\langle My_n, z_n - p\rangle \geq \langle My_n, z_n - y_n\rangle$. Hence,

$$-2\delta\lambda_n\chi_n\langle My_n, z_n - p\rangle \leq -2\delta\lambda_n\chi_n\langle My_n, z_n - y_n\rangle\,. \tag{2.6}$$

From $z_n \in T_n$ and the definition of T_n, we have $\langle w_n - \tau\lambda_n Mw_n - y_n, z_n - y_n\rangle \leq 0$. This shows that

$$\langle w_n - y_n - \tau\lambda_n(Mw_n - My_n), z_n - y_n\rangle \leq \tau\lambda_n\langle My_n, z_n - y_n\rangle\,. \tag{2.7}$$

By using (2.6), (2.7), and the definitions of c_n and χ_n, we obtain

$$\begin{aligned}-2\delta\lambda_n\chi_n\langle My_n, z_n - p\rangle &\leq -2\frac{\delta}{\tau}\chi_n\langle c_n, z_n - y_n\rangle \\ &= -2\frac{\delta}{\tau}\chi_n\langle c_n, w_n - y_n\rangle + 2\frac{\delta}{\tau}\chi_n\langle c_n, w_n - z_n\rangle \\ &= -2\frac{\delta}{\tau}\chi_n^2\|c_n\|^2 + 2\frac{\delta}{\tau}\chi_n\langle c_n, w_n - z_n\rangle\,.\end{aligned} \tag{2.8}$$

Now, we estimate $2\frac{\delta}{\tau}\chi_n\langle c_n, w_n - z_n\rangle$. According to the formula $2ab = a^2 + b^2 - (a-b)^2$, we deduce

$$2\frac{\delta}{\tau}\chi_n\langle c_n, w_n - z_n\rangle = \|w_n - z_n\|^2 + \frac{\delta^2}{\tau^2}\chi_n^2\|c_n\|^2 - \|w_n - z_n - \frac{\delta}{\tau}\chi_n c_n\|^2\,. \tag{2.9}$$

It follows from (2.3) that $\|Mw_n - My_n\| \leq (\mu/\lambda_{n+1})\|w_n - y_n\|$, $\forall n \geq 1$, which combining with the definition of χ_n yields that

$$\begin{aligned}\chi_n\|c_n\|^2 = \langle c_n, w_n - y_n\rangle &\geq \|w_n - y_n\|^2 - \tau\lambda_n\,\|Mw_n - My_n\|\,\|w_n - y_n\| \\ &\geq \left(1 - \frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|^2\,.\end{aligned}$$

This together with the fact that $\|c_n\| \leq (1+\tau\mu\lambda_n/\lambda_{n+1})\|w_n - y_n\|$ implies

$$\chi_n^2\|c_n\|^2 \geq \left(1-\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2 \frac{\|w_n-y_n\|^4}{\|c_n\|^2} \geq \frac{\left(1-\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}{\left(1+\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}\|w_n-y_n\|^2 . \tag{2.10}$$

Combining (2.5), (2.8), (2.9), and (2.10), we conclude that

$$\|z_n-p\|^2 \leq \|w_n-p\|^2 - \|w_n-z_n-\frac{\delta}{\tau}\chi_n c_n\|^2 - \frac{\delta}{\tau^2}(2\tau-\delta)\frac{\left(1-\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}{\left(1+\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}\|w_n-y_n\|^2.$$

This completes the proof. □

Theorem 2.1. *Assume that Assumptions (A1)–(A5) hold. Then the sequence $\{x_n\}$ generated by Algorithm 2.1 converges to the unique solution of the (BVIP) in norm.*

Proof. We first show that the sequence $\{x_n\}$ is bounded. Let $p \in \Omega$. From Lemma 2.4, $\delta \in (0, 2/\mu)$, and $\tau \in (\delta/2, 1/\mu)$, we have

$$\|z_n-p\| \leq \|w_n-p\|, \quad \forall n \geq 1.$$

It follows from (2.2) and the assumptions on $\{\alpha_n\}$ that $\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\| \to 0$ as $n \to \infty$. Therefore, there exists a constant $Q_1 > 0$ such that $\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\| \leq Q_1$, $\forall n \geq 1$. By the definition of w_n, one has $\|w_n-p\| \leq \alpha_n \cdot Q_1 + \|x_n-p\|$. Thus

$$\|z_n-p\| \leq \|w_n-p\| \leq \|x_n-p\| + \alpha_n Q_1, \quad \forall n \geq 1 . \tag{2.11}$$

From Lemma 1.1 and (2.11), it follows that

$$\begin{aligned}
\|x_{n+1}-p\| &= \|(I-\alpha_n\gamma F)z_n - (I-\alpha_n\gamma F)p - \alpha_n\gamma Fp\| \\
&\leq (1-\alpha_n\eta)\|z_n-p\| + \alpha_n\gamma\|Fp\| \\
&\leq (1-\alpha_n\eta)\|x_n-p\| + \alpha_n\eta\cdot\frac{Q_1}{\eta} + \alpha_n\eta\cdot\frac{\gamma}{\eta}\|Fp\| \\
&\leq \max\left\{\frac{Q_1+\gamma\|Fp\|}{\eta}, \|x_n-p\|\right\} \\
&\leq \cdots \leq \max\left\{\frac{Q_1+\gamma\|Fp\|}{\eta}, \|x_1-p\|\right\} ,
\end{aligned}$$

where $\eta = 1-\sqrt{1-\gamma\left(2\beta-\gamma L_F^2\right)} \in (0,1]$. This implies that the sequence $\{x_n\}$ is bounded. We obtain that the sequences $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ are also bounded.

Combining the inequality $\|x+y\|^2 \leq \|x\|^2 + 2\langle y, x+y\rangle$, $\forall x,y \in \mathcal{H}$, (2.11), and Lemma 2.4, we have

$$\begin{aligned}
\|x_{n+1}-p\|^2 &= \|(I-\alpha_n\gamma F)z_n - (I-\alpha_n\gamma F)p - \alpha_n\gamma Fp\|^2 \\
&\leq (1-\alpha_n\eta)^2\|z_n-p\|^2 + 2\alpha_n\gamma\langle Fp, p-x_{n+1}\rangle \\
&\leq \|z_n-p\|^2 + \alpha_n Q_2 \\
&\leq \|x_n-p\|^2 + \alpha_n\left(2Q_1\|x_n-p\| + \alpha_n Q_1^2\right) + \alpha_n Q_2 \\
&\quad - \|w_n-z_n-\frac{\delta}{\tau}\chi_n c_n\|^2 - \frac{\delta}{\tau^2}(2\tau-\delta)\frac{\left(1-\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}{\left(1+\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}\|w_n-y_n\|^2
\end{aligned}$$

for some $Q_2 > 0$. Thus

$$\begin{aligned}
&\|w_n-z_n-\frac{\delta}{\tau}\chi_n c_n\|^2 + \frac{\delta}{\tau^2}(2\tau-\delta)\frac{\left(1-\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}{\left(1+\frac{\tau\mu\lambda_n}{\lambda_{n+1}}\right)^2}\|w_n-y_n\|^2 \\
&\leq \|x_n-p\|^2 - \|x_{n+1}-p\|^2 + \alpha_n Q_3 ,
\end{aligned} \tag{2.12}$$

where $Q_3 := \sup_{n\in\mathbb{N}}\left\{2Q_1\|x_n-p\| + \alpha_n Q_1^2 + Q_2\right\} > 0$.

From the definition of w_n and (2.11), we obtain

$$\begin{aligned}\|x_{n+1}-p\|^2 &\leq (1-\alpha_n\eta)\|w_n-p\|^2+2\alpha_n\gamma\langle Fp, p-x_{n+1}\rangle \\ &\leq (1-\alpha_n\eta)\|x_n-p\|^2+2\alpha_n\gamma\langle Fp, p-x_{n+1}\rangle \\ &\quad +\theta_n\|x_n-x_{n-1}\|\left(2\|x_n-p\|+\theta\|x_n-x_{n-1}\|\right) \\ &\leq \alpha_n\eta\Big[\frac{2\gamma}{\eta}\langle Fp, p-x_{n+1}\rangle+\frac{3Q\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|\Big] \\ &\quad +(1-\alpha_n\eta)\|x_n-p\|^2,\end{aligned} \tag{2.13}$$

where $Q:=\sup_{n\in\mathbb{N}}\{\|x_n-p\|,\theta\|x_n-x_{n-1}\|\}>0$.

Finally we show that the sequence $\{\|x_n-p\|^2\}$ converges to zero. We set

$$p_n=\|x_n-p\|^2,\quad q_n=\frac{2\gamma}{\eta}\langle Fp, p-x_{n+1}\rangle+\frac{3Q\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|.$$

Then the last inequality in (2.13) can be written as $p_{n+1}\leq(1-\alpha_n\eta)p_n+\alpha_n\eta q_n$, $\forall n\geq 1$. Note that the sequence $\{\alpha_n\eta\}$ is in $(0,1)$ and $\sum_{n=1}^{\infty}\alpha_n\eta=\infty$. By Lemma 1.2, it remains to show that $\limsup_{k\to\infty}q_{n_k}\leq 0$ for every subsequence $\{p_{n_k}\}$ of $\{p_n\}$ satisfying $\liminf_{k\to\infty}\left(p_{n_{k+1}}-p_{n_k}\right)\geq 0$. For this purpose, we assume $\{p_{n_k}\}$ is a subsequence of $\{p_n\}$ such that $\liminf_{k\to\infty}\left(p_{n_{k+1}}-p_{n_k}\right)\geq 0$. Combining (2.12), the assumption on $\{\alpha_n\}$, $\delta\in(0,2/\mu)$, and $\tau\in(\delta/2,1/\mu)$, one obtains

$$\begin{aligned}&\limsup_{k\to\infty}\left\{\|w_{n_k}-z_{n_k}-\frac{\delta}{\tau}\chi_{n_k}c_{n_k}\|^2+\frac{\delta}{\tau^2}(2\tau-\delta)\frac{\left(1-\frac{\tau\mu\lambda_{n_k}}{\lambda_{n_k+1}}\right)^2}{\left(1+\frac{\tau\mu\lambda_{n_k}}{\lambda_{n_k+1}}\right)^2}\|w_{n_k}-y_{n_k}\|^2\right\} \\ &\leq \limsup_{k\to\infty}\alpha_{n_k}Q_3+\limsup_{k\to\infty}\left(p_{n_k}-p_{n_k+1}\right) \\ &\leq -\liminf_{k\to\infty}\left(p_{n_k+1}-p_{n_k}\right)\leq 0,\end{aligned}$$

which implies that

$$\lim_{k\to\infty}\|y_{n_k}-w_{n_k}\|=0\quad\text{and}\quad\lim_{k\to\infty}\|w_{n_k}-z_{n_k}-\frac{\delta}{\tau}\chi_{n_k}c_{n_k}\|=0.$$

From the definition of χ_n, we obtain

$$\begin{aligned}\|w_{n_k}-z_{n_k}\| &\leq \|w_{n_k}-z_{n_k}-\frac{\delta}{\tau}\chi_{n_k}c_{n_k}\|+\frac{\delta}{\tau}\chi_{n_k}\|c_{n_k}\| \\ &= \|w_{n_k}-z_{n_k}-\frac{\delta}{\tau}\chi_{n_k}c_{n_k}\|+\frac{\delta}{\tau}\frac{\langle w_{n_k}-y_{n_k},c_{n_k}\rangle}{\|c_{n_k}\|} \\ &\leq \|w_{n_k}-z_{n_k}-\frac{\delta}{\tau}\chi_{n_k}c_{n_k}\|+\frac{\delta}{\tau}\|w_{n_k}-y_{n_k}\|.\end{aligned}$$

Hence, we have that $\lim_{k\to\infty}\|z_{n_k}-w_{n_k}\|=0$. Moreover, we can show that

$$\|x_{n_k+1}-z_{n_k}\|=\alpha_{n_k}\gamma\|Fz_{n_k}\|\to 0\text{ as }n\to\infty,$$

and

$$\|x_{n_k}-w_{n_k}\|=\alpha_{n_k}\cdot\frac{\theta_{n_k}}{\alpha_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0\text{ as }n\to\infty.$$

Thus

$$\|x_{n_k+1}-x_{n_k}\|\leq\|x_{n_k+1}-z_{n_k}\|+\|z_{n_k}-w_{n_k}\|+\|w_{n_k}-x_{n_k}\|\to 0\text{ as }n\to\infty.$$

Since the sequence $\{x_{n_k}\}$ is bounded, there exists a subsequence $\{x_{n_{k_j}}\}$ of $\{x_{n_k}\}$ such that $x_{n_{k_j}}\rightharpoonup z\in\mathcal{H}$ as $j\to\infty$ and

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

By using $\lim_{k\to\infty}\|x_{n_k}-w_{n_k}\|=0$, we obtain $w_{n_{k_j}}\rightharpoonup z$ as $j\to\infty$. This combining with $\lim_{k\to\infty}\|w_{n_k}-y_{n_k}\|=0$ and Lemma 2.2 yields that $z\in\Omega$. From the assumption that p is the unique solution of the (BVIP) and $\lim_{k\to\infty}\|x_{n_k+1}-x_{n_k}\|=0$, we deduce

$$\limsup_{k\to\infty}\langle Fp, p-x_{n_k+1}\rangle\leq\limsup_{k\to\infty}\langle Fp, p-x_{n_k}\rangle=\langle Fp, p-z\rangle\leq 0,$$

which together with $\lim_{n\to\infty}\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|=0$ yields that $\limsup_{k\to\infty}q_{n_k}\leq 0$. Therefore, we conclude that $\lim_{n\to\infty}\|x_n-p\|=0$. That is, $x_n\to p$ as $n\to\infty$. This completes the proof. □

Now, we give a special case of Theorem 2.1. Set $F(x) = x - f(x)$ in Algorithm 2.1, 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 obtain a new inertial modified subgradient extragradient algorithm for solving the (VIP). More specifically, we have the following result.

Corollary 2.1. *Assume that Assumptions (A1), (A2), (A4), and (A5) hold. Let mapping $f : \mathcal{H} \to \mathcal{H}$ be ρ-contraction with $\rho \in [0, \sqrt{5}-2)$. Take $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0,1)$, $\delta \in (0, 2/\mu)$ and $\tau \in (\delta/2, 1/\mu)$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and the iterative sequence $\{x_n\}$ be generated by*

$$\begin{cases} w_n = x_n + \theta_n (x_n - x_{n-1}), \quad y_n = P_C(w_n - \tau\lambda_n M w_n), \\ x_{n+1} = \alpha_n f(z_n) + (1-\alpha_n) z_n, \quad z_n = P_{T_n}(w_n - \delta\lambda_n \chi_n M y_n), \\ \text{where } \theta_n, \lambda_n, \chi_n \text{ and } T_n \text{ are defined in (2.2), (2.3), and (2.4)}. \end{cases} \tag{2.14}$$

Then the iterative sequence $\{x_n\}$ formed by (2.14) converges to $p \in \Omega$ in norm, where $p = P_\Omega(f(p))$.

2.2. Algorithm for uniformly continuous operators

In this subsection, we propose an Armijo-type iterative scheme to approximate the solution of the (BVIP) with a pseudomonotone and non-Lipschitz continuous operator. We replace Assumptions (A4) and (A5) in Algorithm 2.1 with the following Assumptions (A6) and (A7), respectively.

(A6) The operator $M : \mathcal{H} \to \mathcal{H}$ is pseudomonotone, uniformly continuous on $\mathcal{H}$ and satisfies the Assumption (2.1).
(A7) Suppose $\{\epsilon_n\}$ is a non-negative positive sequences such that $\lim_{n\to\infty} \frac{\epsilon_n}{\alpha_n} = 0$, where $\{\alpha_n\} \subset (0,1)$ satisfies $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$.

The second scheme is stated in Algorithm 2.2.

Algorithm 2.2 Armijo-type inertial modified subgradient extragradient method for (BVIP).

Initialization: Take $\theta > 0$, $\sigma > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, $\delta \in (0, 2/\mu)$, $\tau \in (\delta/2, 1/\mu)$, $\gamma \in (0, 2\beta/L_F^2)$. Select two sequences $\{\epsilon_n\}$ and $\{\alpha_n\}$ to satisfy Assumption (A7). Let $x_0, x_1 \in \mathcal{H}$ be arbitrary.
Iterative Steps: Given the iterates x_{n-1} and x_n $(n \ge 1)$, calculate x_{n+1} as follows:
Step 1. Compute $w_n = x_n + \theta_n(x_n - x_{n-1})$, where θ_n is updated by (2.2).
Step 2. Compute $y_n = P_C(w_n - \tau\lambda_n M w_n)$, where $\lambda_n := \sigma\ell^{m_n}$ and m_n is the smallest nonnegative integer m satisfying

$$\sigma\ell^m \langle M w_n - M y_n, w_n - y_n\rangle \le \mu \|w_n - y_n\|^2 . \tag{Amj}$$

Step 3. Compute $z_n = P_{T_n}(w_n - \delta\lambda_n\chi_n M y_n)$, where the half-space T_n and χ_n are defined in (2.4).
Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$.
Set $n := n+1$ and go to **Step 1**.

We can obtain the following conclusions of Lemmas 2.5 and 2.6 by a simple modification of Lemmas 3.1 and 3.3 in [44], respectively. To avoid repetitive expressions, we omit their proofs here.

Lemma 2.5 ([44]). *Suppose that Assumption (A6) holds. Then the Armijo criteria (Amj) is well defined.*

Lemma 2.6 ([44]). *Suppose that Assumptions (A1), (A2), and (A6) hold. Let $\{w_n\}$ and $\{y_n\}$ be two sequences created by Algorithm 2.2. If there exists a subsequence $\{w_{n_k}\}$ of $\{w_n\}$ such that $\{w_{n_k}\}$ converges weakly to $z \in \mathcal{H}$ and $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$, then $z \in \Omega$.*

According to the proofs of Lemmas 2.3 and 2.4, we have the following Lemmas 2.7 and 2.8 without proof.

Lemma 2.7. *If $y_n = w_n$ or $c_n = 0$ in Algorithm 2.2, then $y_n \in \Omega$. Moreover, we have*

$$(1-\tau\mu)\|w_n - y_n\| \le \|c_n\| \le (1+\tau\mu)\|w_n - y_n\|.$$

Lemma 2.8. *Suppose that Assumptions (A1), (A2), and (A6) hold. Let $\{z_n\}$, $\{y_n\}$, and $\{w_n\}$ be three sequences created by Algorithm 2.2. Then, for all $p \in \Omega$,*

$$\|z_n - p\|^2 \le \|w_n - p\|^2 - \|w_n - z_n - \frac{\delta}{\tau}\chi_n c_n\|^2 - \frac{\delta}{\tau^2}(2\tau - \delta)\frac{(1-\tau\mu)^2}{(1+\tau\mu)^2}\|w_n - y_n\|^2.$$

Theorem 2.2. *Assume that Assumptions (A1)–(A3), (A6), and (A7) hold. Then the sequence $\{x_n\}$ generated by Algorithm 2.2 converges to the unique solution of the (BVIP) in norm.*

Proof. The proof is very similar to Theorem 2.1. We leave it to the reader to verify. □

Similar to Corollary 2.1, we have the following result for the special case of Theorem 2.2.

Corollary 2.2. *Assume that Assumptions (A1), (A2), (A6), and (A7) hold. Let mapping $f : \mathcal{H} \to \mathcal{H}$ be ρ-contraction with $\rho \in [0, \sqrt{5} - 2)$. Take $\theta > 0$, $\sigma > 0$, $\ell \in (0, 1)$, $\mu \in (0, 1)$, $\delta \in (0, 2/\mu)$ and $\tau \in (\delta/2, 1/\mu)$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and the iterative sequence $\{x_n\}$ be generated by*

$$\begin{cases} w_n = x_n + \theta_n (x_n - x_{n-1}), \quad y_n = P_C(w_n - \tau\lambda_n M w_n), \\ x_{n+1} = \alpha_n f(z_n) + (1 - \alpha_n) z_n, \quad z_n = P_{T_n}(w_n - \delta\lambda_n \chi_n M y_n), \\ \text{where } \theta_n, \lambda_n, \chi_n \text{ and } T_n \text{ are defined in (2.2), (Amj), and (2.4)}. \end{cases} \tag{2.15}$$

Then the iterative sequence $\{x_n\}$ formed by (2.15) converges to $p \in \Omega$ in norm, where $p = P_\Omega(f(p))$.

Remark 2.2. We discuss further contributions of this paper in the comments below.

(i) In the proposed Algorithms 2.1, 2.2, (2.14), and (2.15), y_n is computed as $y_n = P_C(w_n - \tau\lambda_n M w_n)$. Notice that the step size for computing y_n is $\tau\lambda_n$, which is different from the methods that already exist in the literature [13,16]. The proposed Algorithms (2.14) and (2.15) can be viewed as a modification of Algorithm 3.1 introduced by Dong et al. [16]. Our numerical experiments indicate that this modification can improve the convergence speed of some known algorithms (see Section 3).

(ii) Our Algorithms 2.1 and 2.2 improve some known numerical methods for solving the (BVIP) in the literature [34–39]. More precisely, we base on the following considerations: (1) the algorithms introduced in [34,36] require computing the projection onto the feasible set twice in each iteration, while our Algorithm 2.1 need to calculate it only once; (2) the methods stated in the literature [34,35] are used to solve bilevel monotone variational inequalities, while our algorithms can solve a wider range of bilevel pseudomonotone variational inequalities; (3) the update of the step size of the Algorithm 3.1 suggested by Thong and Hieu [37] requires the prior knowledge of the Lipschitz constant of the mapping, while our algorithms can adaptively update the step size without any prior information; (4) it should be emphasized that the proposed Algorithm 2.2 is designed to solve the (BVIP) with a non-Lipschitz continuous operator, which improves many algorithms in the literature [34–39] for solving the (BVIP) with a Lipschitz continuous operator; and (5) inertial effects are added to the proposed algorithms, which accelerates the convergence speed of the algorithm in [38] without inertial terms (see Section 3).

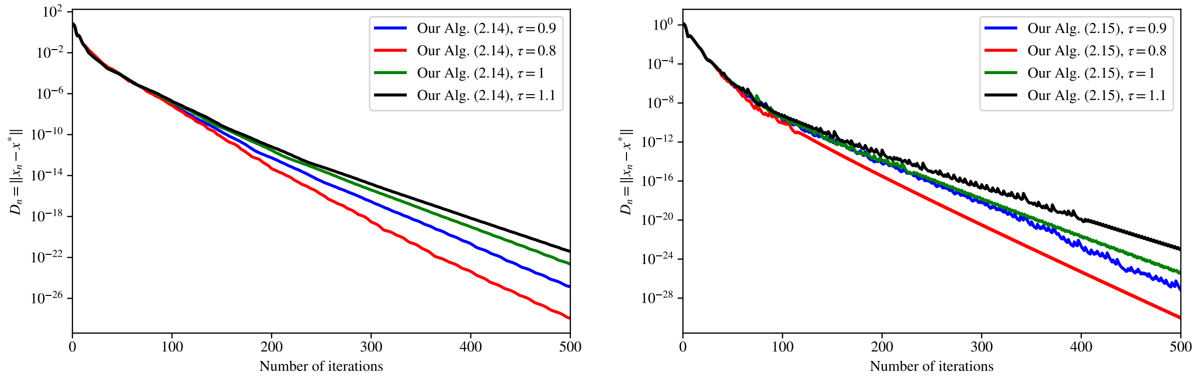
(iii) Our Algorithm (2.14) and Algorithm (2.15) can solve a wider range of variational inequalities with pseudomonotone and non-Lipschitz continuous operators, and thus they improve and unify many of the methods proposed in the literature (see, e.g., [13,16,21,22]) for solving variational inequalities. This is similar to the exposition in (ii) and thus we omit the details.

3. Numerical experiments and applications

In this section, we present some computational experiments to illustrate the numerical performance of the proposed algorithms over some existing ones. 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.

3.1. Theoretical examples

Example 3.1. Consider the linear operator $M : \mathbb{R}^m \to \mathbb{R}^m (m = 20)$ in the form $M(x) = Sx + q$, where $q \in \mathbb{R}^m$ and $S = NN^{\mathsf{T}} + Q + D$, N is a $m \times m$ matrix, Q is a $m \times m$ skew-symmetric matrix, and D is a $m \times m$ diagonal matrix with its diagonal entries being nonnegative (hence S is positive symmetric definite). The feasible set C is given by $C = \{x \in \mathbb{R}^m : -2 \le x_i \le 5,\ i = 1, \ldots, m\}$. It is clear that M is monotone and Lipschitz continuous with constant $L = \|S\|$. In this experiment, all entries of N, Q are generated randomly in $[-2, 2]$, D is generated randomly in $[0, 2]$ and $q = \mathbf{0}$. It is easy to check that the solution of the variational inequality problem is $x^* = \{\mathbf{0}\}$. The maximum number of iterations 500 is used as a stopping criterion and the function $D_n = \|x_n - x^*\|$ is used to measure the error of the n-th iteration step. We use the proposed Algorithms (2.14) and (2.15) to solve this problem. Take $\theta = 0.6$, $\epsilon_n = 100/(n+1)^2$, $\alpha_n = 1/(n+1)$, $\delta = 1.5$, and $f(x) = 0.1x$ for the presented algorithms. Choose $\lambda_1 = 1$, $\mu = 0.2$, and $\xi_n = 1/(n+1)^{1.1}$ for the suggested Algorithm (2.14). Select $\sigma = 2$, $\ell = 0.5$, and $\mu = 0.2$ for the proposed Algorithm (2.15). Fig. 1 shows the numerical performance of the proposed algorithms for different parameter τ.



(a) Our Algorithm (2.14) (b) Our Algorithm (2.15)

Fig. 1. Numerical results for Example 3.1.

Example 3.2. Let $\mathcal{H} = L^2([0,1])$ be an infinite-dimensional Hilbert space 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 as $C = \{x \in \mathcal{H} : \|x\| \le r\}$. The operator $M : \mathcal{H} \to \mathcal{H}$ is given by

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

Note that the operator M is pseudomonotone rather than monotone (see [45, 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 operator F is 0.5-strongly monotone and 0.5-Lipschitz continuous. We use the proposed Algorithms 2.1 and 2.2 to solve the (BVIP) with M, F and C given above, and compare them with two previously known strongly convergent algorithms, including the Algorithm 1 suggested by Thong et al. [38] (shortly, TLDCR Alg. 1) and the Algorithm 3.2 proposed by Tan et al. [39] (shortly, TLQ Alg. 3.2). The parameters of all algorithms are set as follows.

- In the proposed Algorithms 2.1 and 2.2, we set $\theta = 0.4$, $\epsilon_n = 1/(n+1)^2$, $\alpha_n = 1/(n+1)$, $\tau = 0.8$, $\delta = 1.5$, and $\gamma = 1.7\beta/L_F^2$. Pick $\lambda_1 = 0.5$, $\mu = 0.1$, and $\xi_n = 1/(n+1)^{1.1}$ for Algorithm 2.1. Adopt $\sigma = 2$, $\ell = 0.5$, and $\mu = 0.1$ for Algorithm 2.2.
- In the TLDCR Alg. 1 [38], we take $\mu = 0.1$, $\lambda_1 = 0.5$, $\delta = 1.5$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\beta/L_F^2$.
- In the TLQ Alg. 3.2 [39], we choose $\theta = 0.4$, $\epsilon_n = 1/(n+1)^2$, $\mu = 0.1$, $\lambda_1 = 0.5$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\beta/L_F^2$.

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 is used as a common stopping criterion. The numerical behavior of $D_n = \|x_n(t) - x^*(t)\|$ of all algorithms with four starting points $x_0(t) = x_1(t)$ is shown in Fig. 2.

Example 3.3. Consider the Hilbert space $\mathcal{H} = l_2 := \{x = (x_1, x_2, \ldots, x_i, \ldots) \mid \sum_{i=1}^{\infty} |x_i|^2 < +\infty\}$ equipped with inner product

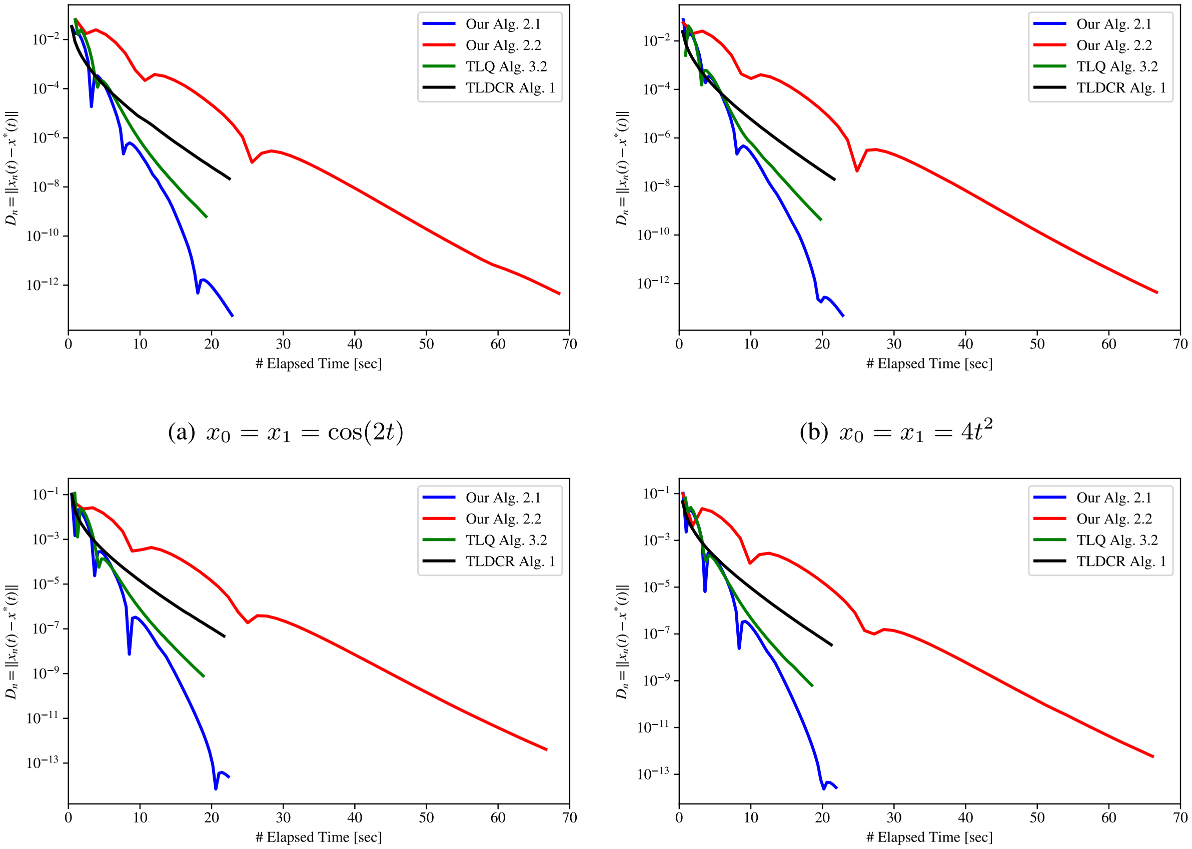
$$\langle x, y\rangle = \sum_{i=1}^{\infty} x_i y_i, \quad \forall x, y \in \mathcal{H},$$

and induced norm

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

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

$$Mx = (\|x\| + 1/(\|x\| + \varphi))\,x$$



(a) $x_0 = x_1 = \cos(2t)$ (b) $x_0 = x_1 = 4t^2$

(c) $x_0 = x_1 = \log(2t)$ (d) $x_0 = x_1 = 2^t$

Fig. 2. Numerical results for Example 3.2.

for some $\varphi > 0$. It can be verified that mapping M is pseudomonotone on $\mathcal{H}$, uniformly continuous and sequentially weakly continuous on C but not Lipschitz continuous on $\mathcal{H}$ (see [46]). In the following cases, we take $\varphi = 0.5$, $\mathcal{H} = \mathbb{R}^m$ for different values of m. In this case, the feasible set C is a box $C = \{x \in \mathbb{R}^m : -1/i \le x_i \le 1/i,\ i = 1, 2, \ldots, m\}$. We compare the proposed Algorithm (2.15) with several strongly convergent algorithms that can solve the (VIP) with uniformly continuous operators, including the Algorithm 4 proposed by Reich et al. [25] (shortly, RTDLD Alg. 4), the Algorithm 3.1 introduced by Cai et al. [26] (shortly, CDP Alg. 3.1) and the Algorithm 3 suggested by Thong et al. [47] (shortly, TSI Alg. 3). Take $\alpha_n = 1/(n+1)$, $f(x) = 0.1x$, $\sigma = 2$, $\ell = 0.5$, and $\mu = 0.6$ for all algorithms. Choose $\lambda = 0.5/\mu$ for RTDLD Alg. 4. Set $\theta = 0.4$, $\epsilon_n = 1/(n+1)^2$, $\tau = 0.8$, and $\delta = 1.5$ for the suggested Algorithm (2.15). The initial values $x_0 = x_1 = 5rand(m, 1)$ are randomly generated by MATLAB. The maximum number of iterations 200 is used as a common stopping criterion. The numerical performance of $D_n = \|x_n - x_{n-1}\|$ of all algorithms with four different dimensions is reported in Fig. 3.

Remark 3.1. We have the following observations for the results of Examples 3.1–3.3.

(i) The iterative methods proposed in this paper are efficient and robust. They have a better numerical performance than the algorithms presented in the literature [25,26,38,39,47] for the same stopping criterion, and these results are not significantly related to the choice of initial values and the size of the dimensions.
(ii) It can be seen from Fig. 1 that the values of the parameter τ have different effects on the proposed Algorithms (2.14) and (2.15). Specifically, the algorithms with $\tau = \{0.8, 0.9\}$ can accelerate the convergence speed of the algorithms with $\tau = 1$. Therefore, our schemes have a faster convergence speed when a suitable τ is chosen.
(iii) Note that the operator M in Example 3.2 is pseudomonotone rather than monotone and that the Lipschitz constant of the operator M is unknown. In these cases, the algorithms introduced in [34,35] for solving bilevel monotone variational inequalities and the algorithms offered in [36,37] that require the prior knowledge of the Lipschitz constant of the operator will be unavailable. On the other hand, as demonstrated in Example 3.3, the operator M is

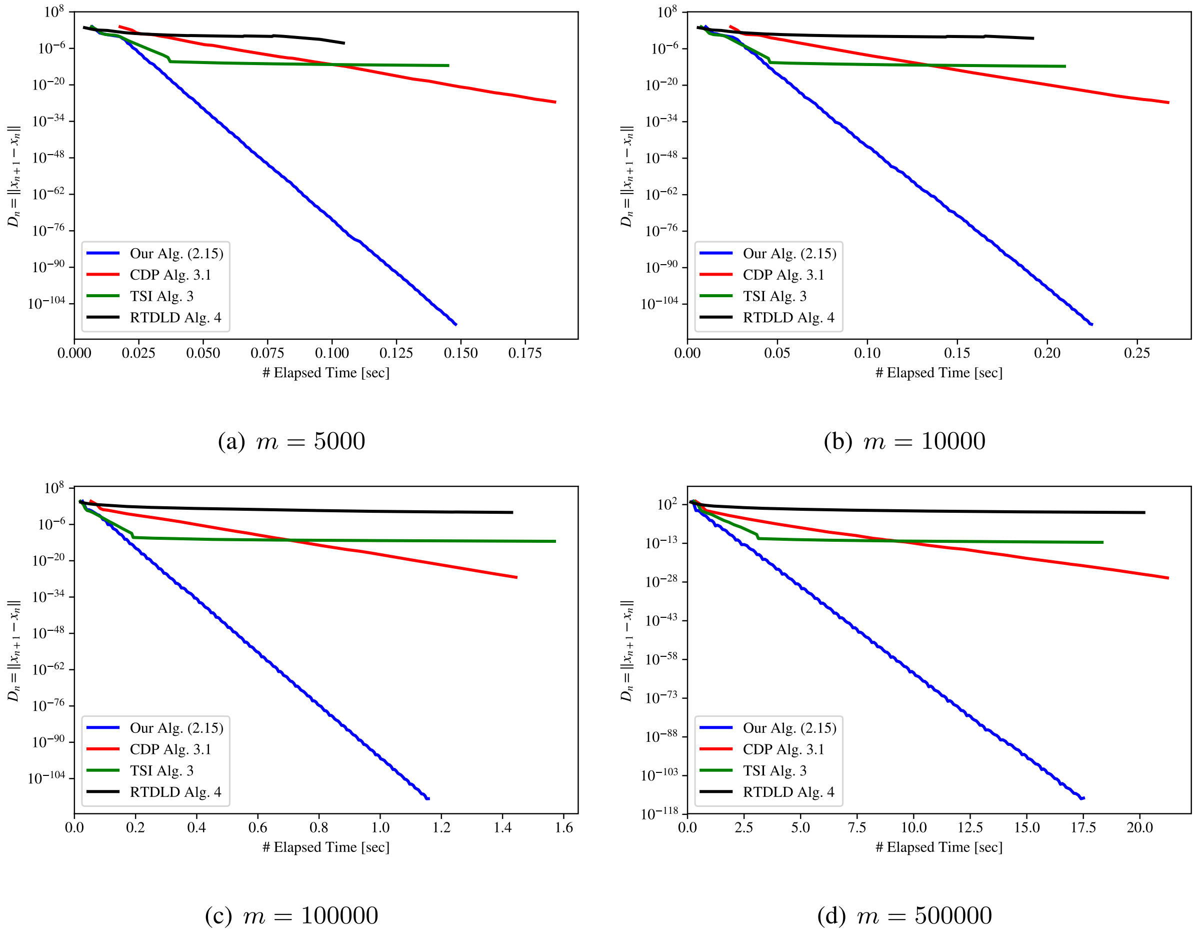


Fig. 3. Numerical results for Example 3.3.

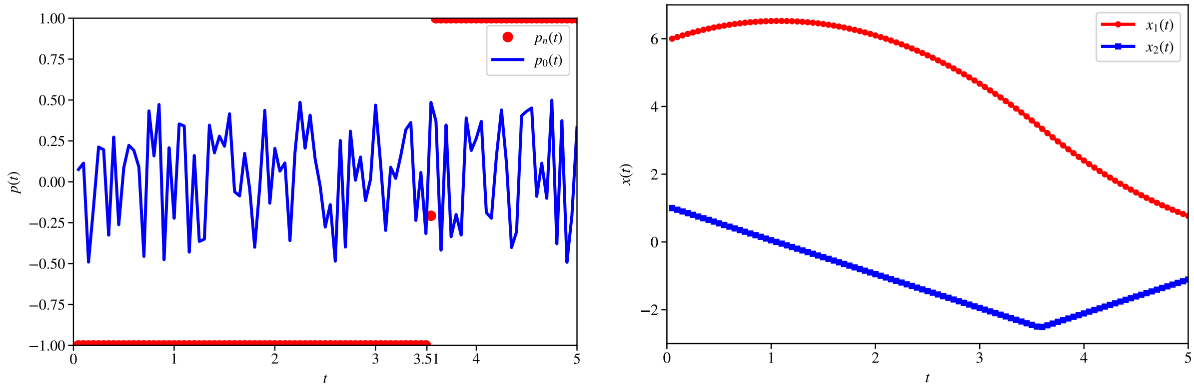
uniformly continuous without satisfying Lipschitz continuity. Many algorithms in the literature (see, e.g., [13,16]) that solve the variational inequality problem involving a Lipschitz continuous operator will fail in this case. However, the schemes stated in this paper can solve these problems well, and therefore they have a wider range of applications.

(iv) It can be seen from Fig. 2 that the proposed Algorithm 2.2 requires more execution time in an infinite-dimensional Hilbert space than the adaptive algorithm proposed in [38,39], because it uses an Armijo-type step size criterion that may require multiple computations of the projection onto the feasible set in each iteration. However, the proposed adaptive Algorithm 2.1 converges very fast due to the fact that it employs a new non-monotonic sequence of stepsizes.

3.2. Application to optimal control problems

Next, we use the proposed algorithms to solve the (VIP) that appears in optimal control problems. We recommend readers to refer to [23,48] for a detailed description of the problem. We compare the suggested Algorithm (2.14) and Algorithm (2.15) with some strongly convergent algorithms in the literature. Two methods used to compare here are the Algorithm (31) (shortly, TLDCR Alg. (31)) introduced by Thong et al. [38] and the Algorithm (3.39) (shortly, TLQ Alg. (3.39)) proposed by Tan et al. [39]. The parameters of all algorithms are set as follows.

- In the proposed Algorithms (2.14) and (2.15), we set $N = 100$, $\theta = 0.01$, $\epsilon_n = 10^{-4}/(n+1)^2$, $\delta = 1.5$, $\alpha_n = 10^{-4}/(n+1)$, $\tau = \{0.8, 1\}$, and $f(x) = 0.1x$. Pick $\lambda_1 = 0.5$, $\mu = 0.1$, and $\xi_n = 0.1/(n+1)^{1.1}$ for Algorithm (2.14). Adopt $\sigma = 2$, $\ell = 0.5$, and $\mu = 0.1$ for Algorithm (2.15).
- In the TLDCR Alg. (31) [38], we choose $N = 100$, $\mu = 0.1$, $\lambda_1 = 0.5$, $\delta = 1.5$, and $\alpha_n = 10^{-4}/(n+1)$.
- In the TLQ Alg. (3.39) [39], we take $N = 100$, $\theta = 0.01$, $\epsilon_n = 10^{-4}/(n+1)^2$, $\mu = 0.1$, $\lambda_1 = 0.5$, $\alpha_n = 10^{-4}/(n+1)$, and $f(x) = 0.1x$.



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

Fig. 4. Numerical results of the proposed Algorithm (2.14) for Example 3.4.

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

Example 3.4 (*Rocket car [48]*).

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

The exact optimal control of Example 3.4 is

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

The approximate optimal control and the corresponding trajectories of the suggested Algorithm (2.14) (with $\tau = 0.8$) are plotted in Fig. 4.

Example 3.5 (*see [49]*).

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

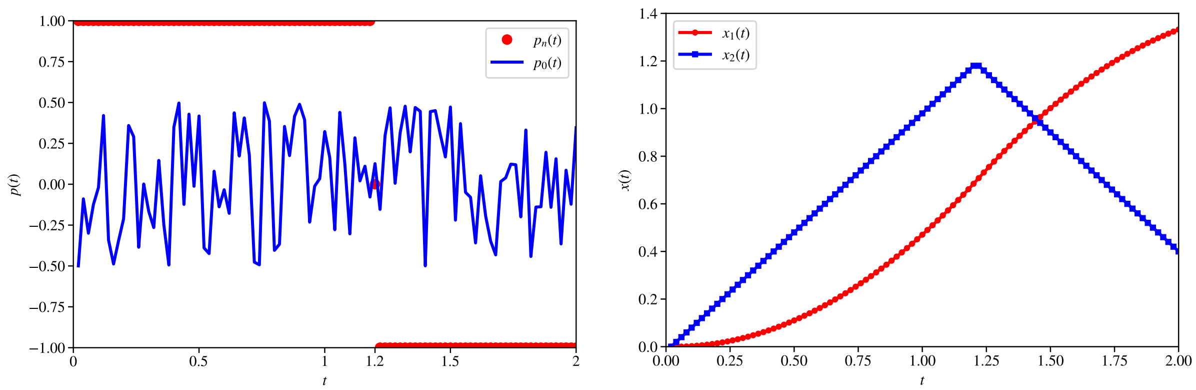
The exact optimal control of Example 3.5 is

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

Fig. 5 shows the approximate optimal control and the corresponding trajectories of the proposed Algorithm (2.15) (with $\tau = 0.8$).

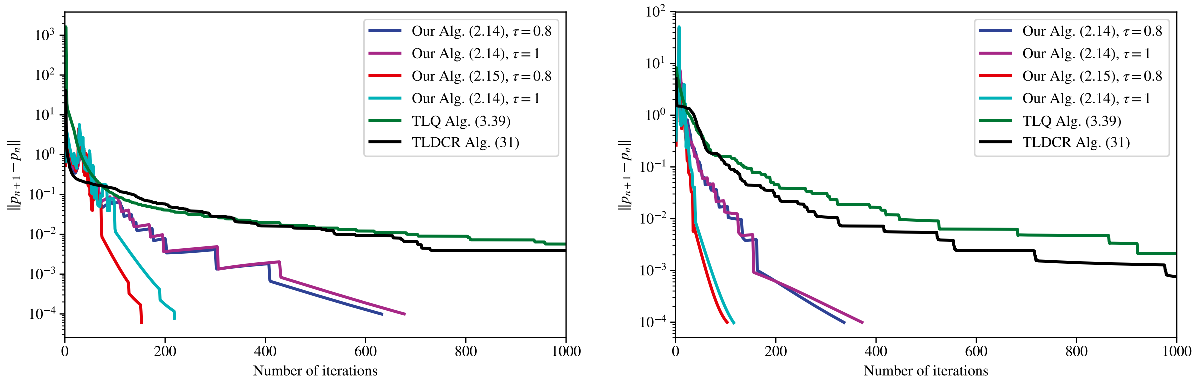
Finally, we compare the offered Algorithm (2.14) and Algorithm (2.15) with TLQ Alg. (3.39) and TLDCR Alg. (31) for Examples 3.4 and 3.5. Fig. 6 presents the numerical behavior of the error estimate $\|p_{n+1} - p_n\|$ with respect to the number of iterations for all algorithms. Moreover, the number of terminated iterations and the execution time of all algorithms are shown in Table 1.

Remark 3.2. The suggested Algorithms (2.14) and (2.15) can be applied to solve optimal control problems. As shown in Fig. 6 and Table 1, the proposed Algorithms (2.14) and (2.15) outperform the existing methods in the literature [38,39]. Moreover, when $\tau = 0.8$, the proposed algorithms converge faster than the algorithms with $\tau = 1.0$.



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

Fig. 5. Numerical results of the proposed Algorithm (2.15) for Example 3.5.



(a) Example 3.4 (b) Example 3.5

Fig. 6. Error estimates of all algorithms for Examples 3.4 and 3.5.

Table 1
Numerical results of all algorithms for Examples 3.4 and 3.5.

Algorithms	Example 3.4			Example 3.5		
	Iter.	Time (s)	D_n	Iter.	Time (s)	D_n
Our Alg. (2.14), $\tau = 0.8$	631	0.2453	9.94E−05	335	0.1137	9.89E−05
Our Alg. (2.14), $\tau = 1.0$	676	0.2513	1.00E−04	371	0.1246	9.93E−05
Our Alg. (2.15), $\tau = 0.8$	152	0.1473	6.02E−05	102	0.0434	9.93E−05
Our Alg. (2.15), $\tau = 1.0$	218	0.1662	7.79E−05	115	0.0502	9.77E−05
TLQ Alg. (3.39)	1000	0.3389	5.69E−03	1000	0.3541	2.12E−03
TLDCR Alg. (31)	1000	0.3251	3.87E−03	1000	0.3241	7.48E−04

4. Conclusions

In this paper, we introduced two new modified subgradient extragradient methods to approximate the solution of bilevel variational inequalities. The advantages of the proposed algorithms are that (1) only one projection onto the feasible set needs to be computed in each iteration; (2) the operator involved is pseudomonotone and Lipschitz continuous (or uniformly continuous); (3) the update of the step size does not require the prior knowledge of the Lipschitz constant of the mapping; and (4) the embedding of the inertial terms accelerates the convergence speed of the algorithms. Strong convergence theorems of the presented algorithms are established in the framework of real Hilbert spaces. Finally, the computational efficiency of our iterative schemes compared to the known methods in the literature is verified by some numerical tests and applications. The results obtained in this paper improved and extended many numerical methods in the literature for solving variational inequalities and bilevel variational inequalities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the editor and the anonymous reviewers for their insightful comments, which helped us to improve the readability and quality of the initial manuscript.

References

[1] Dempe S, Zemkoho A. Bilevel optimization. In: Springer optimization and its applications. Vol. 161. Cham: Springer; 2020, http://dx.doi.org/10.1007/978-3-030-52119-6_20.
[2] Goeleven D. Complementarity and variational inequalities in electronics. In: Mathematical analysis and its applications. London: Academic Press; 2017.
[3] Cubiotti P, Yao JC. On the Cauchy problem for a class of differential inclusions with applications. Appl Anal 2020;99:2543–54.
[4] An NT, Dong PD, Qin X. Robust feature selection via nonconvex sparsity-based methods. J Nonlinear Var Anal 2021;5:59–77.
[5] An NT, Nam NM, Qin X. Solving *k*-center problems involving sets based on optimization techniques. J Global Optim 2020;76:189–209.
[6] Sahu DR, Yao JC, Verma M, Shukla KK. Convergence rate analysis of proximal gradient methods with applications to composite minimization problems. Optimization 2021;70:75–100.
[7] Shehu Y, Vuong PT, Zemkoho A. An inertial extrapolation method for convex simple bilevel optimization. Optim Methods Softw 2021;36:1–19.
[8] Sofonea M, Xiao YB, Zeng SD. Generalized penalty method for history-dependent variational–hemivariational inequalities. Nonlinear Anal Real World Appl 2021;61:103329.
[9] Xiao YB, Sofonea M. Generalized penalty method for elliptic variational–hemivariational inequalities. Appl Math Optim 2021;83:789–812.
[10] Noor MA, Noor KI, Rassias MT. New trends in general variational inequalities. Acta Appl Math 2020;170:981–1064.
[11] Korpelevich GM. The extragradient method for finding saddle points and other problems. ÈKon I Mat Metody 1976;12:747–56.
[12] Tseng P. A modified forward–backward splitting method for maximal monotone mappings. SIAM J Control Optim 2000;38:431–46.
[13] Censor Y, Gibali A, Reich S. The subgradient extragradient method for solving variational inequalities in Hilbert space. J Optim Theory Appl 2011;148:318–35.
[14] He BS. A class of projection and contraction methods for monotone variational inequalities. Appl Math Optim 1997;35:69–76.
[15] Malitsky Y. Projected reflected gradient methods for monotone variational inequalities. SIAM J Optim 2015;25:502–20.
[16] Dong QL, Jiang D, Gibali A. A modified subgradient extragradient method for solving the variational inequality problem. Numer Algorithms 2018;79:927–40.
[17] Yang J, Liu H. Strong convergence result for solving monotone variational inequalities in Hilbert space. Numer Algorithms 2019;80:741–52.
[18] Hieu DV, Anh PK. Modified extragradient-like algorithms with new stepsizes for variational inequalities. Comput Optim Appl 2019;73:913–32.
[19] Shehu Y, Iyiola OS. Projection methods with alternating inertial steps for variational inequalities: Weak and linear convergence. Appl Numer Math 2020;157:315–37.
[20] Liu H, Yang J. Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput Optim Appl 2020;77:491–508.
[21] Shehu Y, Cholamjiak P. Iterative method with inertial for variational inequalities in Hilbert spaces. Calcolo 2019;56:4.
[22] Shehu Y, Iyiola OS, Li XH, Dong QL. Convergence analysis of projection method for variational inequalities. Comput Appl Math 2019;38:161.
[23] Vuong PT, Shehu Y. Convergence of an extragradient-type method for variational inequality with applications to optimal control problems. Numer Algorithms 2019;81:269–91.
[24] Shehu Y, Dong QL, Jiang D. Single projection method for pseudo-monotone variational inequality in Hilbert spaces. Optimization 2019;68:385–409.
[25] Reich S, Thong DV, Dong QL, Li XH, Dung VT. New algorithms and convergence theorems for solving variational inequalities with non-Lipschitz mappings. Numer Algorithms 2021;87:527–49.
[26] Cai G, Dong QL, Peng Y. Strong convergence theorems for solving variational inequality problems with pseudo-monotone and non-Lipschitz operators. J Optim Theory Appl 2021;188:447–72.
[27] Polyak B. Some methods of speeding up the convergence of iteration methods. USSR Comput Math Math Phys 1964;4:1–17.
[28] Alvarez F, Attouch H. An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal 2001;9:3–11.
[29] Gibali A, Hieu DV. A new inertial double-projection method for solving variational inequalities. J Fixed Point Theory Appl 2019;21:97.
[30] Qin X, Wang L, Yao JC. Inertial splitting method for maximal monotone mappings. J Nonlinear Convex Anal 2020;21:2325–33.
[31] Tan B, Li S. Strong convergence of inertial mann algorithms for solving hierarchical fixed point problems. J Nonlinear Var Anal 2020;4:337–55.
[32] Ceng LC, Shang M. Hybrid inertial subgradient extragradient methods for variational inequalities and fixed point problems involving asymptotically nonexpansive mappings. Optimization 2021;70:715–40.
[33] Yamada I. The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. In: Inherently parallel algorithms for feasibility and optimization and their applications. Amsterdam: Elsevier; 2001, p. 473–504.
[34] Maingé PE. A hybrid extragradient-viscosity method for monotone operators and fixed point problems. SIAM J Control Optim 2008;47:1499–515.
[35] Hieu DV, Moudai A. Regularization projection method for solving bilevel variational inequality problem. Optim Lett 2021;15:205–29.
[36] Anh PN, Kim JK, Muu LD. An extragradient algorithm for solving bilevel pseudomonotone variational inequalities. J Global Optim 2012;52:627–39.
[37] Thong DV, Hieu DV. A strong convergence of modified subgradient extragradient method for solving bilevel pseudomonotone variational inequality problems. Optimization 2020;69:1313–34.
[38] Thong DV, Li XH, Dong QL, Cho YJ, Rassias TM. A projection and contraction method with adaptive step sizes for solving bilevel pseudo-monotone variational inequality problems. Optimization 2020. http://dx.doi.org/10.1080/02331934.2020.1849206.
[39] Tan B, Liu L, Qin X. Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems. Jpn J Ind Appl Math 2021;38:519–43.
[40] Saejung S, Yotkaew P. Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal 2012;75:742–50.
[41] Thong DV, Yang J, Cho YJ, Rassias TM. Explicit extragradient-like method with adaptive stepsizes for pseudomonotone variational inequalities. Optim Lett 2021;15:2181–99.

[42] Denisov SV, Semenov VV, Chabak LM. Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators. Cybern Syst Anal 2015;51:757–65.
[43] Cegielski A. Iterative methods for fixed point problems in Hilbert spaces. In: Lecture notes in mathematics, vol. 2057, Berlin: Springer; 2012.
[44] Tan B, Cho SY. Inertial extragradient methods for solving pseudomonotone variational inequalities with non-Lipschitz mappings and their optimization applications. Appl Set-Valued Anal Optim 2021;3:165–92.
[45] Hieu DV, Cho YJ, Xiao YB, Kumam P. Modified extragradient method for pseudomonotone variational inequalities in infinite dimensional Hilbert spaces. Vietnam J Math 2021;49:1165–83.
[46] Thong DV, Shehu Y, Iyiola OS. A new iterative method for solving pseudomonotone variational inequalities with non-Lipschitz operators. Comput Appl Math 2020;39:108.
[47] Thong DV, Shehu Y, Iyiola OS. Weak and strong convergence theorems for solving pseudo-monotone variational inequalities with non-Lipschitz mappings. Numer Algorithms 2020;84:795–823.
[48] Preininger J, Vuong PT. On the convergence of the gradient projection method for convex optimal control problems with bang–bang solutions. Comput Optim Appl 2018;70:221–38.
[49] Bressan B, Piccoli B. Introduction to the mathematical theory of control. aims series on applied mathematics, 2. Springfield: American Institute of Mathematical Sciences (AIMS); 2007.