Fixed Point Theory, 25(2024), No. 1, 309-332
DOI: 10.24193/fpt-ro.2024.1.19
http://www.math.ubbcluj.ro/~nodeacj/sfptcj.html

EFFICIENT EXTRAGRADIENT METHODS FOR BILEVEL PSEUDOMONOTONE VARIATIONAL INEQUALITIES WITH NON-LIPSCHITZ OPERATORS AND THEIR APPLICATIONS

X. QIN*,**, A. PETRUŞEL***, B. TAN**** AND J.C. YAO*****

*Department of Mathematics, Yibin University, Sichuan, Yibin, China

**Center for Converging Humanities, Kyung Hee University, Seoul, Korea
E-mail: qxlxajh@163.com

***Department of Mathematics, Babeş-Bolyai University, Cluj-Napoca, Romania
E-mail: petrusel@math.ubbcluj.ro

****School of Mathematics and Statistics, Southwest University, Chongqing, China
E-mail: bingtan72@gmail.com

*****Research Center for Interneural Computing, China Medical University Hospital, China Medical University, Taichung, Taiwan
E-mail: yaojc@mail.cmu.edu.tw (Corresponding author)

Abstract. In this paper, four extragradient-type algorithms are presented for solving bilevel variational inequalities of pseudomonotone and non-Lipschitz continuous operators in real Hilbert spaces. The proposed iterative schemes employ two Armijo-type linesearch methods making them work adaptively. Strong convergence theorems of the suggested algorithms are established under some mild conditions. Finally, some numerical experiments and applications are given to verify the advantages and efficiency of the proposed algorithms over some previously known ones.
Key Words and Phrases: Bilevel variational inequality problem, inertial extragradient method, Armjio stepsize, pseudomonotone mapping, non-Lipschitz operator, fixed point.
2020 Mathematics Subject Classification: 65K15, 90C30, 47H10.

1. Introduction

Throughout the paper, 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 $\|\cdot\|$. Let us first review the classical variational inequality problem (shortly VIP), which is described as follows:

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

where $M : C \to \mathcal{H}$ is an operator. One denotes by $\mathrm{VI}(C, M)$ the set of all solutions of (VIP). It is known that variational inequalities play a significant role in applied science and optimization theory. They provide a general and useful framework for solving engineering problems, data sciences, and other fields. Therefore, numerical methods for studying variational inequalities have attracted numerous interest among researchers due to the wide application of VIPs. In this paper, we focus on the special case of VIPs with variational inequality constraints. That is, we want to find the solution to the following bilevel variational inequality problem (shortly BVIP), which reads as follows:

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

Bilevel variational inequality problems cover a number of nonlinear optimization problems, such as fixed point problems, quasi-variational inequality problems, complementary problems, saddle problems, and minimum norm problems. Bilevel optimization problems are hierarchical optimization problems in which the feasible region of the upper-level problem is restricted by the solution set of the lower-level problem. BVIP is a bilevel optimization problem where both the upper-level problem and the lower-level problem can be written as variational inequalities. For more details on the theory, algorithms, and applications of the bilevel optimization problem, we refer the reader to the recent monograph [7].

A basic method for solving VIPs is the projection one. The earliest and simplest numerical method to solve (VIP) is the projected gradient method (shortly, PGM), which generates an iterative sequence $\{x_n\}$ from x_1 in the following way: $x_{n+1} = P_C(x_n - \chi_n M x_n)$, where χ_n represents a set of appropriate parameters and $P_C : \mathcal{H} \to C$ denotes the metric (nearest point) projection from $\mathcal{H}$ onto C, characterized by $P_C(x) := \arg\min\{\|x - y\| : y \in C\}$ and $P_C(x) \in C$ for all $x \in \mathcal{H}$. It is worth noting that the convergence condition of PGM is especially strong, that is, the operator M is required to be strongly monotone and Lipschitz continuous, which limits the implementation of such methods in practical applications. In order to weaken the strong monotonicity of operator M, Korpelevich [12] proposed a two-step iterative scheme (now called the extragradient method, shortly EGM) to solve monotone variational inequality problems. However, the disadvantage of the EGM is that the projection onto the feasible set needs to be calculated twice in each iteration. It is noted that computing projection is equivalent to solving an optimization problem, which may not be easy to solve when the feasible set has a complex structure. Therefore, the EGM not only weakens the operator but also increases computational consumption. Next, we introduce three feasible methods to improve the computational efficiency of the EGM. The first is the Tseng's extragradient method (shortly TEGM) proposed by Tseng [26], which replaces the second step of EGM with an explicit calculation step. Another feasible scheme is to convert the projection of the second step of EGM onto the feasible set into the projection onto the half-space (note that the projection onto a half-space can be explicitly calculated). This method is now referred to as the subgradient extragradient method (shortly SEGM), which was introduced by Censor, Gibali and Reich [5]. The third method of using only one projection for the feasible set is the projection and contraction method (shortly PCM) suggested by He [9]. The first step of PCM is the same as the first step of EGM, but the calculation of the second step is updated by some previous information without involving any projection process. Therefore, these three methods (TEGM, SEGM, and PCM) only need to calculate the projection onto the feasible set once in each iteration, which greatly improves the computational efficiency of EGM.

Note that the extragradient-type methods (EGM, TEGM, SEGM, and PCM) mentioned above need to calculate the projection onto the feasible set at least once in each iteration. Is there a way to avoid calculating projections and solve variational inequalities? Indeed, Yamada [28] introduced a new iterative scheme, which is now stated the hybrid steepest descent method and is read as follows: $x_{n+1} = (I - \sigma\alpha_{n+1}F)Ux_n$, $\forall n \geq 1$, where I is the identity mapping, F is an η-strongly monotone and k-Lipschitz continuous mapping, U is a nonexpansive mapping, $\sigma \in (0, 2\eta/k^2)$ and $\{\alpha_n\}$ is a sequence that satisfies some restrictions. He proved that the iterative sequence generated by this method converges strongly to a point x^*, which is the unique solution of the variational inequality problem over the fixed point set, that is, find $x^* \in \mathrm{Fix}(U)$ such that $\langle Fx^*, y - x^*\rangle \geq 0$, $\forall y \in \mathrm{Fix}(U)$, where $\mathrm{Fix}(U) = \{x \in \mathcal{H} : Ux = x\}$ denotes the fixed point set of U. On the other hand, the inertial idea has been studied by many researchers as a technique to accelerate the convergence speed of algorithms. The main feature of the inertial method is that the next iteration depends on the combination of the previous two iterations. This small change can significantly improve the computational efficiency of the algorithms without inertial terms.

Recently, a large number of numerical algorithms have been proposed for solving BVIPs (see, e.g., [13, 11, 1, 10, 14, 21, 20]). A common characteristic enjoyed by these algorithms is that the operator M is required to be Lipschitz continuous. However, this condition may be difficult to be satisfied in real applications because there exist classes of mappings that do not satisfy Lipschitz continuity (such as uniformly continuous mappings). Recently, some new numerical methods with Armijo-type linesearch rules have been proposed to solve non-Lipschitz continuous variational inequality problems; see, e.g., [19, 4, 22, 23, 17]. It should be emphasized that the Armijo-type step size criterion suggested by Cai et al. [4] is different from the other ones [22, 23, 17]. Moreover, it is worth noting that the operator M in the algorithms proposed in [13, 10] are monotone, while the operator M in [11, 1, 21, 20] are pseudomonotone. It is known that pseudomonotone mappings contain monotone mappings. *A natural question is how to modify the extragradient algorithms so that they can solve the bilevel pseudomonotone variational inequality problem containing a non-Lipschitz continuous mapping.* To answer this question, in this paper we propose several modified extragradient methods with Armijo-type stepsizes for solving bilevel pseudomonotone variational inequalities. Strong convergence theorems of the suggested algorithms are established under some suitable and weaker conditions. Some numerical experiments are presented to verify the advantages and efficiency of the proposed algorithms.

The organizational structure of this paper is as follows. We review some basic definitions and lemmas that need to be used in Section 2. Section 3 states the suggested iterative schemes and analyzes their convergence properties. In Section 4, we perform some numerical examples to demonstrate the advantages of the proposed algorithms in comparison with some related ones. Finally, we conclude the paper with a brief summary in Section 5, the last section.

2. Preliminaries

Let C be a nonempty, closed, and convex subset of a real Hilbert space $\mathcal{H}$. The weak convergence and strong convergence of $\{x_n\}_{n=1}^{\infty}$ to x are represented by $x_n \rightharpoonup x$ and $x_n \to x$, respectively. For each $x, y \in \mathcal{H}$, we have the following inequality

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

It is known that P_C is nonexpansive and has the following basic properties, for all $x \in \mathcal{H}$ and $y \in C$, $\|P_C(x)-y\|^2 \leq \|x-y\|^2 - \|x-P_C(x)\|^2$, $\langle x-P_C(x), y-P_C(x)\rangle \leq 0$, and $\|P_C(x)-P_C(y)\|^2 \leq \langle P_C(x)-P_C(y), x-y\rangle$

Recall that a mapping $M : \mathcal{H} \to \mathcal{H}$ is said to be:

(i) *L-Lipschitz continuous* with $L > 0$ if $\|Mx - My\| \leq L\|x-y\|$ for all $x, y \in \mathcal{H}$. (If $L \in (0,1)$ then mapping M is called *contraction*. In particular, mapping M is called *nonexpansive* when $L = 1$.)
(ii) *uniformly continuous on $\mathcal{H}$ if for any $\varepsilon > 0$, there exists $\delta > 0$ such that if $x, y \in \mathcal{H}$ and $\|x-y\| < \delta$, then $\|Mx - My\| < \varepsilon$.*
(iii) *α-strongly monotone* if there exists a constant $\alpha > 0$ such that $\langle Mx - My, x - y\rangle \geq \alpha\|x-y\|^2$ for all $x, y \in \mathcal{H}$.
(iv) *monotone* if $\langle Mx - My, x-y\rangle \geq 0$ for all $x, y \in \mathcal{H}$.
(v) *pseudomonotone* if $\langle Mx, y-x\rangle \geq 0 \Rightarrow \langle My, y-x\rangle \geq 0$ for all $x, y \in \mathcal{H}$.
(vi) *sequentially weakly continuous* if for each sequence $\{x_n\}$ converges weakly to x implies $\{Mx_n\}$ converges weakly to Mx.

The following lemmas play an important role in our proofs.

Lemma 2.1 ([28])**.** *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, $\|T^\gamma x - T^\gamma y\| \leq (1-\alpha\eta)\|x-y\|$ for all $x, y \in \mathcal{H}$, where $\eta = 1 - \sqrt{1-\gamma(2\beta - \gamma L^2)} \in (0,1)$.*

Lemma 2.2 ([18])**.** *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 $p_{n+1} \leq \alpha_n q_n + (1-\alpha_n)p_n$ for all $n \geq 1$. If $\limsup_{k\to\infty} q_{n_k} \leq 0$ for every subsequence $\{p_{n_k}\}$ of $\{p_n\}$ satisfying $\liminf_{k\to\infty}(p_{n_k+1} - p_{n_k}) \geq 0$, then $\lim_{n\to\infty} p_n = 0$.*

3. Main results

In this section, we present four efficient algorithms to solve the bilevel pseudomonotone variational inequality problem with a non-Lipschitz continuous operator. The following several conditions are assumed to be satisfied before introducing our algorithms.

(C1) The feasible set C is a nonempty, closed, and convex subset of a real Hilbert space $\mathcal{H}$.
(C2) The solution set of the problem (VIP) is nonempty, that is, $\mathrm{VI}(C, M) \neq \emptyset$.

(C3) The operator $M : \mathcal{H} \to \mathcal{H}$ is pseudomonotone, uniformly continuous on $\mathcal{H}$, and satisfies the following condition

$$\text{whenever } \{x_n\} \subset C,\ x_n \rightharpoonup z,\ \text{one has } \|Mz\| \leq \liminf_{n\to\infty} \|Mx_n\|. \tag{con1}$$

(C4) The mapping $F : \mathcal{H} \to \mathcal{H}$ is L_F-Lipschitz continuous and β-strongly monotone on $\mathcal{H}$.

(C5) Let $\{\epsilon_n\}$ be a positive sequence 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 $\sum_{n=1}^{\infty} \alpha_n = \infty$.

3.1. **First modified subgradient extragradient algorithm.** Now we are ready to state the first iterative scheme with a new Armijo-type stepsize criterion, which is based on the subgradient extragradient method, the inertial method, and the hybrid steepest descent method. The Algorithm 3.1 is formulated as follows.

Algorithm 3.1 First modified inertial subgradient extragradient method for solving (BVIP).

Initialization: Take $\theta > 0,\ \delta > 0,\ \ell \in (0,1),\ \mu \in (0,1),\ \gamma \in (0, 2\beta/L_F^2)$ and 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

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

Step 2. Compute $y_n = P_C(w_n - \chi_n M w_n)$.

Step 3. Compute $z_n = P_{T_n}(w_n - \chi_n M y_n)$, where the half-space T_n is defined by

$$T_n := \{x \in \mathcal{H} \mid \langle w_n - \chi_n M w_n - y_n, x - y_n\rangle \leq 0\},$$

and $\chi_n := \delta \ell^{m_n}$ and m_n is the smallest nonnegative integer m satisfying

$$\delta \ell^m \langle M y_n - M w_n, y_n - z_n\rangle \leq \frac{\mu}{2}\left[\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right]. \tag{Ar-1}$$

Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$.

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

Remark 3.1. Before considering the convergence of our algorithm, let us emphasize that the following observations from Algorithm 3.1.

(i) It should be noted that the Armijo criterion (Ar-1) is derived from the recent article by Cai et al. [4] and it is not of the same type as the one that already exists in the literature [19, 22, 23, 17]. On the other hand, the proposed Algorithm 3.1 uses a different type of step size than Algorithm 3.1 of Tan et al. [20], which leads to different convergence conditions for them. Specifically, the proposed Algorithm 3.1 only requires that the operator M is uniformly continuous while Algorithm 3.1 of Tan et al. [20] requires that it is Lipschitz continuous. It is known that the uniform continuity of the operator is weaker than the Lipschitz continuity, so the proposed Algorithm 3.1 has a wider application.

(ii) When mapping M is monotone, it is not necessary to impose Condition (con1) (see [6]).

(iii) We note here that inertial calculation criterion (In-Cri) is easy to implement since the term $\|x_n - x_{n-1}\|$ is known before calculating θ_n. Moreover, it follows from (In-Cri) and the assumptions on $\{\alpha_n\}$ that $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$. Indeed, we obtain $\theta_n\|x_n - x_{n-1}\| \leq \epsilon_n$ for all $n \geq 1$, which together with $\lim_{n\to\infty} \frac{\epsilon_n}{\alpha_n} = 0$ implies that $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \leq \lim_{n\to\infty} \frac{\epsilon_n}{\alpha_n} = 0$.

(iv) Note that the condition (con1) is used by many recent work on pseudomonotone variational inequalities; see, e.g., [17, 25]. It is easy to check that Condition (con1) is weaker than the sequential weak continuity of the mapping M (see [25, Remark 3.2]).

The following lemmas play an important role in the convergence analysis of Algorithm 3.1.

Lemma 3.1. *Suppose that Conditions (C1)–(C3) hold. Then the Armijo-like criteria* (Ar-1) *is well defined.*

Proof. The proof is trivial, and we omit the details here. □

Lemma 3.2. *Suppose that Conditions (C1)–(C3) hold. Let $\{w_n\}$ and $\{y_n\}$ be two sequences formulated by Algorithm 3.1. 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 $z \in \mathrm{VI}(C, M)$.*

Proof. The proof of this lemma follows the proof of Lemma 3.2 in [4], so it is omitted. □

Lemma 3.3. *Let $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ be three sequences generated by Algorithm 3.1 and $p \in \mathrm{VI}(C, M)$. Then $\|z_n - p\|^2 \leq \|w_n - p\|^2 - (1-\mu)(\|y_n - w_n\|^2 + \|z_n - y_n\|^2)$.*

Proof. By the definition of z_n and the property of projection, one sees that

$$\begin{aligned} 2\|z_n - p\|^2 &\leq 2\langle z_n - p, w_n - \chi_n M y_n - p\rangle \\ &= \|z_n - p\|^2 + \|w_n - \chi_n M y_n - p\|^2 - \|z_n - w_n + \chi_n M y_n\|^2 \\ &= \|z_n - p\|^2 + \|w_n - p\|^2 + \chi_n^2\|My_n\|^2 - 2\langle w_n - p, \chi_n M y_n\rangle \\ &\quad - \|z_n - w_n\|^2 - \chi_n^2\|My_n\|^2 - 2\langle z_n - w_n, \chi_n M y_n\rangle \\ &= \|z_n - p\|^2 + \|w_n - p\|^2 - \|z_n - w_n\|^2 - 2\langle z_n - p, \chi_n M y_n\rangle . \end{aligned}$$

This implies that

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \|z_n - w_n\|^2 - 2\langle z_n - p, \chi_n M y_n\rangle . \tag{3.1}$$

Since p is the solution of (VIP), we have $\langle Mp, x - p\rangle \geq 0$ for all $x \in C$. By the pseudomonotonicity of M, we obtain $\langle Mx, x - p\rangle \geq 0$ for all $x \in C$. Taking $x = y_n \in C$, one infers that $\langle My_n, p - y_n\rangle \leq 0$. Consequently,

$$\langle My_n, p - z_n\rangle = \langle My_n, p - y_n\rangle + \langle My_n, y_n - z_n\rangle \leq \langle My_n, y_n - z_n\rangle . \tag{3.2}$$

Combining (3.1) and (3.2), one obtains

$$\begin{aligned}\|z_n - p\|^2 &\le \|w_n - p\|^2 - \|z_n - w_n\|^2 + 2\chi_n \langle My_n, y_n - z_n\rangle \\ &= \|w_n - p\|^2 - \|z_n - y_n\|^2 - \|y_n - w_n\|^2 \\ &\quad - 2\langle z_n - y_n, y_n - w_n\rangle + 2\chi_n \langle My_n, y_n - z_n\rangle \\ &= \|w_n - p\|^2 - \|z_n - y_n\|^2 - \|y_n - w_n\|^2 \\ &\quad + 2\langle z_n - y_n, w_n - \chi_n My_n - y_n\rangle .\end{aligned} \tag{3.3}$$

According to $z_n \in T_n$ and the definition of T_n, one obtains

$$\begin{aligned}&2\langle w_n - \chi_n My_n - y_n, z_n - y_n\rangle \\ &= 2\langle w_n - \chi_n Mw_n - y_n, z_n - y_n\rangle + 2\chi_n \langle Mw_n - My_n, z_n - y_n\rangle \\ &\le 2\chi_n \langle Mw_n - My_n, z_n - y_n\rangle .\end{aligned} \tag{3.4}$$

Combining (Ar-1), (3.3), and (3.4), we have

$$\begin{aligned}\|z_n - p\|^2 &\le \|w_n - p\|^2 - \|w_n - y_n\|^2 - \|y_n - z_n\|^2 + 2\chi_n \langle My_n - Mw_n, y_n - z_n\rangle \\ &\le \|w_n - p\|^2 - \|w_n - y_n\|^2 - \|y_n - z_n\|^2 + \mu\left[\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right] \\ &= \|w_n - p\|^2 - (1-\mu)(\|y_n - w_n\|^2 + \|z_n - y_n\|^2).\end{aligned}$$

This completes the proof of the lemma. □

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

Proof. We divide the proof into four claims.
Claim 1. The sequence $\{x_n\}$ is bounded. It follows from Lemma 3.3 that

$$\|z_n - p\| \le \|w_n - p\|, \quad \forall n \ge 1. \tag{3.5}$$

By the definition of w_n, one has

$$\|w_n - p\| \le \alpha_n \cdot \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| + \|x_n - p\|. \tag{3.6}$$

According to Remark 3.1 we have $\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \to 0$ as $n \to \infty$. Therefore, there exists a constant $Q_1 > 0$ such that

$$\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \le Q_1, \quad \forall n \ge 1. \tag{3.7}$$

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

$$\|z_n - p\| \le \|w_n - p\| \le \|x_n - p\| + \alpha_n Q_1, \quad \forall n \ge 1. \tag{3.8}$$

Using Lemma 2.1 and (3.5), one has

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

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 $\{w_n\}$ and $\{z_n\}$ are also bounded.
Claim 2.

$$(1-\mu)\big(\|y_n-w_n\|^2+\|z_n-y_n\|^2\big)\le\|x_n-p\|^2-\|x_{n+1}-p\|^2+\alpha_nQ_4$$

for some $Q_4>0$. Indeed, using (2.1), one has

$$\begin{aligned}\|x_{n+1}-p\|^2 &= \|\left(I-\alpha_n\gamma F\right)z_n-\left(I-\alpha_n\gamma F\right)p-\alpha_n\gamma Fp\|^2 \\ &\le \left(1-\alpha_n\eta\right)^2\|z_n-p\|^2+2\alpha_n\gamma\left\langle Fp,p-x_{n+1}\right\rangle \\ &\le \|z_n-p\|^2+\alpha_nQ_2\end{aligned} \tag{3.9}$$

for some $Q_2>0$. In the light of Lemma 3.3, we obtain

$$\|x_{n+1}-p\|^2\le\|w_n-p\|^2-(1-\mu)\big(\|y_n-w_n\|^2+\|z_n-y_n\|^2\big)+\alpha_nQ_2. \tag{3.10}$$

It follows from (3.8) that

$$\begin{aligned}\|w_n-p\|^2 &\le \left(\|x_n-p\|+\alpha_nQ_1\right)^2 \\ &= \|x_n-p\|^2+\alpha_n\left(2Q_1\|x_n-p\|+\alpha_nQ_1^2\right) \\ &\le \|x_n-p\|^2+\alpha_nQ_3\end{aligned} \tag{3.11}$$

for some $Q_3>0$. Combining (3.10) and (3.11), we obtain

$$(1-\mu)\big(\|y_n-w_n\|^2+\|z_n-y_n\|^2\big)\le\|x_n-p\|^2-\|x_{n+1}-p\|^2+\alpha_nQ_4,$$

where $Q_4:=Q_2+Q_3$.
Claim 3.

$$\|x_{n+1}-p\|^2\le\left(1-\alpha_n\eta\right)\|x_n-p\|^2+\alpha_n\eta\Big[\frac{2\gamma}{\eta}\left\langle Fp,p-x_{n+1}\right\rangle+\frac{3Q\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|\Big]$$

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

$$\|w_n-p\|^2\le\|x_n-p\|^2+2\theta_n\|x_n-p\|\|x_n-x_{n-1}\|+\theta_n^2\|x_n-x_{n-1}\|^2. \tag{3.12}$$

Combining (3.5) and (3.9), we deduce

$$\|x_{n+1}-p\|^2\le\left(1-\alpha_n\eta\right)\|w_n-p\|^2+2\alpha_n\gamma\left\langle Fp,p-x_{n+1}\right\rangle. \tag{3.13}$$

Substituting (3.12) into (3.13), we obtain

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

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

Claim 4. The sequence $\{\|x_n-p\|\}$ converges to zero. By Lemma 2.2, it needs to show that $\limsup_{k\to\infty}\langle Fp, p-x_{n_k+1}\rangle \le 0$ for every subsequence $\{\|x_{n_k}-p\|\}$ of $\{\|x_n-p\|\}$ satisfying $\liminf_{k\to\infty}(\|x_{n_k+1}-p\|-\|x_{n_k}-p\|)\ge 0$.

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

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

By Claim 2 and the assumption on $\{\alpha_n\}$, one obtains

$$\begin{aligned}&\limsup_{k\to\infty}\left[(1-\mu)\left(\|y_{n_k}-w_{n_k}\|^2+\|z_{n_k}-y_{n_k}\|^2\right)\right] \\ &\le \limsup_{k\to\infty}\left[\alpha_{n_k}Q_4+\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2\right] \\ &\le \limsup_{k\to\infty}\alpha_{n_k}Q_4+\limsup_{k\to\infty}\left[\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2\right] \\ &= -\liminf_{k\to\infty}\left[\|x_{n_k+1}-p\|^2-\|x_{n_k}-p\|^2\right]\le 0,\end{aligned}$$

which implies that $\lim_{k\to\infty}\|y_{n_k}-w_{n_k}\| = \lim_{k\to\infty}\|z_{n_k}-y_{n_k}\| = 0$. Therefore, we obtain

$$\lim_{k\to\infty}\|z_{n_k}-w_{n_k}\| = 0. \tag{3.14}$$

Moreover, we can show that

$$\|x_{n_k+1}-z_{n_k}\| = \alpha_{n_k}\gamma\|Fz_{n_k}\| \to 0 \text{ as } n\to\infty, \tag{3.15}$$

and

$$\|x_{n_k}-w_{n_k}\| = \alpha_{n_k}\cdot\frac{\theta_{n_k}}{\alpha_{n_k}}\|x_{n_k}-x_{n_k-1}\| \to 0 \text{ as } n\to\infty. \tag{3.16}$$

Combining (3.14), (3.15), and (3.16), we obtain

$$\|x_{n_k+1}-x_{n_k}\| \le \|x_{n_k+1}-z_{n_k}\|+\|z_{n_k}-w_{n_k}\|+\|w_{n_k}-x_{n_k}\| \to 0 \text{ as } n\to\infty. \tag{3.17}$$

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 $z\in\mathcal{H}$. Moreover,

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

By (3.16), we obtain $w_{n_k}\rightharpoonup z$ as $k\to\infty$. This together with $\lim_{k\to\infty}\|w_{n_k}-y_{n_k}\|=0$ and Lemma 3.2 yields $z\in \mathrm{VI}(C,M)$. From the assumption that p is the unique solution of the (BVIP), we deduce

$$\limsup_{k\to\infty} \langle Fp, p - x_{n_k}\rangle = \langle Fp, p - z\rangle \le 0. \tag{3.18}$$

Using (3.17) and (3.18), we obtain

$$\limsup_{k\to\infty} \langle Fp, p - x_{n_k+1}\rangle = \limsup_{k\to\infty} \langle Fp, p - x_{n_k}\rangle \le 0. \tag{3.19}$$

From $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$ and (3.19), we have

$$\limsup_{k\to\infty} \left[\frac{2\gamma}{\eta}\langle Fp, p - x_{n_k+1}\rangle + \frac{3Q\theta_{n_k}}{\alpha_{n_k}\eta}\|x_{n_k} - x_{n_k-1}\|\right] \le 0. \tag{3.20}$$

Combining Claim 3, Condition (C5), and (3.20), in the light of Lemma 2.2, one concludes 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 3.1. Set $F(x) = x - f(x)$ in Algorithm 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 obtain an inertial subgradient extragradient algorithm with a new Armijo-type step size for solving (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 $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, and $\mu \in (0,1)$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and the iterative sequence $\{x_n\}$ be generated by the following*

$$\boxed{\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \chi_n M w_n), \\ z_n = P_{T_n}(w_n - \chi_n M y_n), \\ T_n := \{x \in \mathcal{H} \mid \langle w_n - \chi_n M w_n - y_n, x - y_n\rangle \le 0\}, \\ x_{n+1} = (1-\alpha_n) z_n + \alpha_n f(z_n), \end{cases}} \tag{3.21}$$

where θ_n and χ_n are defined in (In-Cri) *and* (Ar-1)*, respectively. Then the iterative sequence $\{x_n\}$ formed by* (3.21) *converges to p in norm, where $p = P_{\mathrm{VI}(C,M)}(f(p))$.*

3.2. **Second modified subgradient extragradient algorithm.** In this subsection, we present another version of Algorithm 3.1. Our second iterative scheme is shown in Algorithm 3.2. The only difference between this method and Algorithm 3.1 is that their iteration step size are updated in two different ways.

Remark 3.2. Following the proof method of Lemma 3.1 in [22], we can obtain that the Armijo-type criterion (Ar-2) is well defined. Let x_n be a sequence generated by Algorithm 3.2, then Lemma 3.2 still holds by replacing x_n in the proof process in Lemma 3.3 of [22] with w_n.

Algorithm 3.2 Second modified inertial subgradient extragradient method for solving (BVIP).

Initialization: Take $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, $\gamma \in (0, 2\beta/L_F^2)$ and 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 θ_n is defined in (In-Cri).
Step 2. Compute $y_n = P_C(w_n - \chi_n M w_n)$, where $\chi_n := \delta \ell^{m_n}$ and m_n is the smallest nonnegative integer m satisfying

$$\delta \ell^m \langle M w_n - M y_n, w_n - y_n \rangle \leq \mu \|w_n - y_n\|^2. \tag{Ar-2}$$

Step 3. Compute $z_n = P_{T_n}(w_n - \chi_n M y_n)$, where the half-space T_n is defined by

$$T_n := \{x \in \mathcal{H} \mid \langle w_n - \chi_n M w_n - y_n, x - y_n \rangle \leq 0\}.$$

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

Lemma 3.4. *Assume that Conditions (C1)–(C3) hold. Let $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ be three sequences generated by Algorithm 3.2. Then, for all $p \in \mathrm{VI}(C, M)$,*

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - (1-\mu)(\|y_n - w_n\|^2 + \|z_n - y_n\|^2).$$

Proof. Using the definition of χ_n in (Ar-2), one obtains $\chi_n \|M w_n - M y_n\| \leq \mu \|w_n - y_n\|$. Moreover, we obtain

$$\begin{aligned} 2\chi_n \langle M w_n - M y_n, z_n - y_n \rangle &\leq 2\chi_n \|M w_n - M y_n\| \|y_n - z_n\| \\ &\leq 2\mu \|w_n - y_n\| \|y_n - z_n\| \\ &\leq \mu \|w_n - y_n\|^2 + \mu \|y_n - z_n\|^2 . \end{aligned}$$

Applying the same statements as (3.1)–(3.4) in the proof of Lemma 3.3, we can obtain the desired conclusion. This completes the proof of the lemma. □

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

Proof. The proof of this result follows almost in the same way as that of Theorem 3.1 but we apply Lemma 3.4 in place of Lemma 3.3. □

3.3. **Modified Tseng's extragradient algorithm.** In this subsection, we introduce a new iterative algorithm for solving (BVIP), which is based on the Tseng's extragradient method, the inertial method, and the hybrid steepest descent method. More precisely, the method is displayed in Algorithm 3.3.

Lemma 3.5. *Assume that Conditions (C1)–(C3) hold. Let $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ be three sequences generated by Algorithm 3.3. Then, for all $p \in \mathrm{VI}(C, M)$,*

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - (1-\mu^2) \|y_n - w_n\|^2 .$$

Proof. The proof of this lemma is very similar to the proof of Lemma 3.2 in [24]. So we omit the details. □

Algorithm 3.3 Modified inertial Tseng's extragradient method for solving (BVIP).

Initialization: Take $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, $\gamma \in (0, 2\beta/L_F^2)$ and 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 θ_n is defined in (In-Cri).
Step 2. Compute $y_n = P_C(w_n - \chi_n M w_n)$, where χ_n is defined in (Ar-2).
Step 3. Compute $z_n = y_n - \chi_n (M y_n - M w_n)$.
Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$.
Set $n := n+1$ and go to *Step 1.*

Theorem 3.3. *Assume that Conditions (C1)–(C5) hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.3 converges to the unique solution of the* (BVIP) *in norm.*

Proof. As shown in Claim 1 of the proof of Theorem 3.1, we obtain that the sequences $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ are bounded. From (3.9), (3.11), and Lemma 3.5, we have

$$(1-\mu^2)\|y_n - w_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n Q_4. \tag{Eq1}$$

In addition, we can obtain the same conclusion as Claim 3 in Theorem 3.1. Finally, we show that $\{\|x_n - p\|\}$ converges to zero. For this purpose, one assumes that $\{\|x_{n_k} - p\|\}$ is a subsequence of $\{\|x_n - p\|\}$ such that $\liminf_{k\to\infty} (\|x_{n_k+1} - p\| - \|x_{n_k} - p\|) \geq 0$. By (Eq1) and the assumption on $\{\alpha_n\}$, one obtains

$$\limsup_{k\to\infty} \left[(1-\mu^2)\|y_{n_k} - w_{n_k}\|^2\right] \leq \limsup_{k\to\infty} \left[\alpha_{n_k} Q_4 + \|x_{n_k} - p\|^2 - \|x_{n_k+1} - p\|^2\right] \leq 0,$$

which implies that $\lim_{k\to\infty} \|y_{n_k} - w_{n_k}\| = 0$. From the definition of z_n and (Ar-2), we have $\|z_n - y_n\| \leq \mu \|w_n - y_n\|$. Thus we obtain $\lim_{k\to\infty} \|z_{n_k} - y_{n_k}\| = 0$. Following the same statements as (3.14)–(3.20) in Theorem 3.1, we conclude that $x_n \to p$ as $n \to \infty$. The proof is completed. □

3.4. **Modified projection and contraction algorithm.** Finally, inspired by the inertial method, the projection and contraction method, and the hybrid steepest descent method, the last iterative scheme for solving (BVIP) is given. The concrete expression of Algorithm 3.4 is shown below.

Lemma 3.6. *Assume that Conditions (C1)–(C3) hold. Let $\{w_n\}$, $\{y_n\}$, and $\{z_n\}$ be three sequences generated by Algorithm 3.4. Then*

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \frac{(2-\phi)}{\phi}\|z_n - w_n\|^2, \quad \forall p \in \mathrm{VI}(C, M),$$

and

$$\|w_n - y_n\|^2 \leq \frac{(1+\mu)^2}{[(1-\mu)\phi]^2}\|z_n - w_n\|^2.$$

Proof. The two conclusions of this lemma are easily obtained by a simple modification of Lemma 3.2 of Dong et al. [8]. Thus we omit the details. □

Algorithm 3.4 Modified inertial projection and contraction method for solving (BVIP).

Initialization: Take $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, $\phi \in (0,2)$, $\gamma \in (0, 2\beta/L_F^2)$ and 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 θ_n is defined in (In-Cri).

Step 2. Compute $y_n = P_C(w_n - \chi_n M w_n)$, where χ_n is defined in (Ar-2).

Step 3. Compute $z_n = w_n - \phi\delta_n d_n$, where d_n and δ_n are defined by

$$d_n := w_n - y_n - \chi_n \left(M w_n - M y_n\right), \quad \delta_n := \begin{cases} \langle w_n - y_n,\, d_n\rangle / \|d_n\|^2, & \text{if } d_n \neq 0; \\ 0, & \text{otherwize.} \end{cases}$$

Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$.

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

Theorem 3.4. *Assume that Conditions (C1)–(C5) hold. Then the sequence $\{x_n\}$ generated by Algorithm 3.4 converges to the unique solution of the* (BVIP) *in norm.*

Proof. The proof of the theorem is similar to the proof of Theorem 3.3. We leave it to the reader to verify it. □

Similar to Corollary 3.1, we have the following conclusions for Algorithms 3.2–3.4.

Corollary 3.2. *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 $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, and $\phi \in (0,2)$. Let $x_0, x_1 \in \mathcal{H}$ be two arbitrary initial points and the iterative sequence $\{x_n\}$ be generated by*

$$\boxed{\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \chi_n M w_n), \\ z_n = P_{T_n}(w_n - \chi_n M y_n), \\ T_n = \left\{x \in \mathcal{H} \mid \langle w_n - \chi_n M w_n - y_n, x - y_n\rangle \leq 0\right\}, \\ x_{n+1} = (1 - \alpha_n) z_n + \alpha_n f(z_n), \end{cases}} \tag{3.22}$$

$$\boxed{\begin{cases} w_n = x_n + \theta_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \chi_n M w_n), \\ z_n = y_n - \chi_n \left(M y_n - M w_n\right), \\ x_{n+1} = (1 - \alpha_n) z_n + \alpha_n f(z_n), \end{cases}} \tag{3.23}$$

$$
\begin{cases}
w_n = x_n + \theta_n(x_n - x_{n-1}), \\
y_n = P_C(w_n - \chi_n M w_n), \\
z_n = w_n - \phi \delta_n d_n, \\
d_n = w_n - y_n - \chi_n \left(M w_n - M y_n\right), \\
\delta_n = \begin{cases} \langle w_n - y_n,\, d_n\rangle / \|d_n\|^2, & \text{if } d_n \neq 0; \\ 0, & \text{otherwize.} \end{cases} \\
x_{n+1} = (1-\alpha_n) z_n + \alpha_n f(z_n),
\end{cases}
\tag{3.24}
$$

where θ_n and χ_n are defined in (In-Cri) *and* (Ar-2)*, respectively. Then the iterative sequence $\{x_n\}$ formed by Algorithm* [(3.22), (3.23), (3.24)] *converges to p in norm, where $p = P_{\mathrm{VI}(C,M)}(f(p))$.*

Remark 3.3. We make the following observations for the suggested algorithms.

(i) The four algorithms obtained in this paper can solve the bilevel pseudomonotone variational inequality problem, while the algorithms suggested in [13, 10] can only solve the bilevel monotone variational inequality problem. On the other hand, we replace the Lipschitz continuity of mapping M in the literature [13, 11, 1, 10, 21, 20] with the uniform continuity of mapping M in the suggested algorithms. Therefore, our proposed approaches have a wider range of applications.

(ii) It should be emphasized that the Armijo-type criterion (Ar-2) in Algorithm 3.2 does not use the information of sequence $\{z_n\}$ when updating the step size in each iteration, while the Armijo-type criterion (Ar-1) in Algorithm 3.1 uses the information of sequence $\{z_n\}$, which makes Algorithm 3.1 converge faster than Algorithm 3.2 (see the numerical examples in Section 4).

(iii) Our Algorithms (3.21)–(3.24) improve many numerical methods in the literature [19, 4, 22, 23, 17] for solving variational inequality problems due to the fact that the mapping M involved in the proposed algorithms is pseudomonotone and uniformly continuous. Moreover, if we set $\theta_n = 0$ and $\phi = 1$ in our Algorithm (3.24), then it degenerates to the Algorithm 3 proposed by Thong et al. [22]. Indeed, they compute the value of z_n through the projection of x_n on the half-space C_n, where C_n is defined as $C_n := \{x \in \mathcal{H} : \langle d_n, x - y_n\rangle \leq 0\}$ and $d_n = x_n - y_n - \chi_n \left(M x_n - M y_n\right)$. It is known that the projection on the half-space can be computed explicitly. Therefore

$$
z_n = P_{C_n}\left(x_n\right) = x_n - \frac{\langle d_n, x_n - y_n\rangle}{\|d_n\|^2} d_n,
$$

which is equivalent to the calculation of z_n in Algorithm (3.24) (by setting $\theta_n = 0$ and $\phi = 1$).

(iv) Our algorithms are embedded with inertial terms making them converge faster than the algorithms without inertial (see Section 4).

4. Numerical experiments and applications

In this section, we provide some computational tests to demonstrate the numerical behavior of the proposed Algorithms 3.1–3.4, and also to compare them with the Algorithm 1 introduced by Thong et al. [21] and the Algorithm 3.2 suggested by Tan, Liu and Qin [20]. Notice that the algorithms mentioned in [21, 20] do not require the priori information about the Lipschitz constant of mapping M. All the programs were implemented in MATLAB 2018a on a Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz computer with RAM 8.00 GB. The parameters of all algorithms are set as follows.

- In the proposed Algorithms 3.1–3.4, we set $\theta = 0.4$, $\epsilon_n = 100/(n+1)^2$, $\delta = 2$, $\ell = 0.5$, $\mu = 0.1$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\beta/L_F^2$. Pick $\phi = 1.5$ for Algorithm 3.4.
- In the Algorithm 1 introduced by Thong et al. [21], we choose $\mu = 0.1$, $\chi_1 = 0.6$, $\phi = 1.5$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\beta/L_F^2$.
- In the Algorithm 3.2 suggested by Tan et al. [20], we take $\theta = 0.4$, $\epsilon_n = 100/(n+1)^2$, $\mu = 0.1$, $\chi_1 = 0.6$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\beta/L_F^2$.

4.1. **Numerical examples of finite- and infinite-dimensions.**

Example 4.1. Consider a mapping $F : \mathbb{R}^m \to \mathbb{R}^m$ ($m = 5$) of the form $F(x) = Gx + q$, where $G = BB^\mathsf{T} + D + K$, 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, whose diagonal entries are positive in $(0, 1)$ (so G 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 β-strongly monotone with $L_F = \max\{\operatorname{eig}(G)\}$ and $\beta = \min\{\operatorname{eig}(G)\}$, where $\operatorname{eig}(G)$ represents all eigenvalues of G. Next, we consider the following fractional programming problem:

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

where

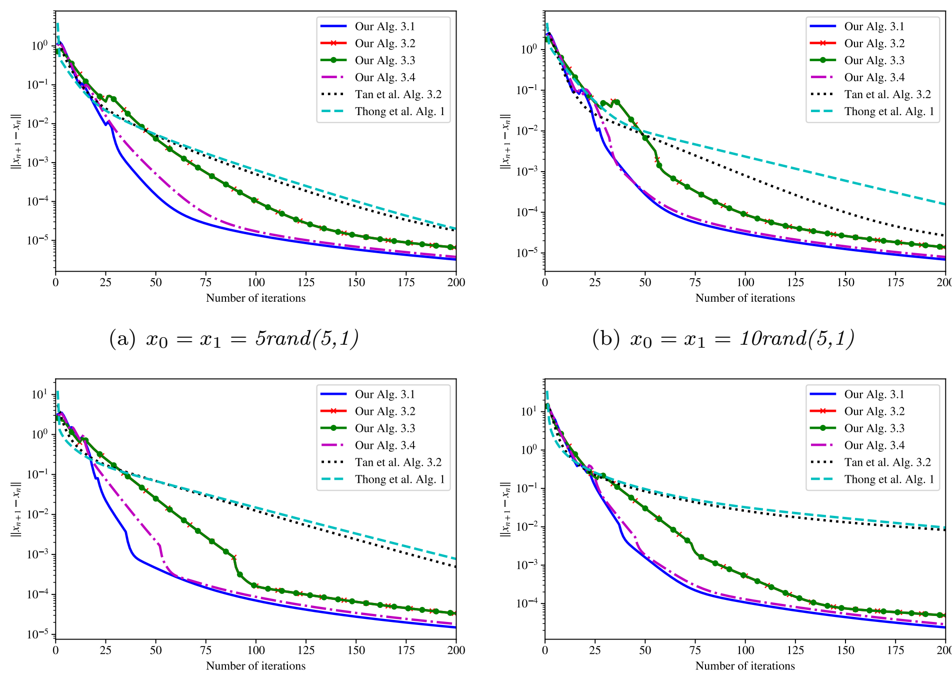
$$Q = \begin{bmatrix} 5 & -1 & 2 & 0 & 2 \\ -1 & 6 & -1 & 3 & 0 \\ 2 & -1 & 3 & 0 & 1 \\ 0 & 3 & 0 & 5 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{bmatrix}, \; a = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \\ 1 \end{bmatrix}, \; b = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, \; a_0 = -2, b_0 = 20.$$

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

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

It is known that the mapping M is pseudomonotone and Lipschitz continuous (see [2]).

We use $D_n = \|x_{n+1} - x_n\|$ to measure the error of the n-th iteration since we do not know the exact solution to the problem. The maximum number of iterations 200 is used as a common stopping criterion. Numerical results of all algorithms with four initial values are reported in Fig. 1.



(a) $x_0 = x_1 = \mathit{5rand(5,1)}$ (b) $x_0 = x_1 = \mathit{10rand(5,1)}$

(c) $x_0 = x_1 = \mathit{20rand(5,1)}$ (d) $x_0 = x_1 = \mathit{50rand(5,1)}$

FIGURE 1. Numerical results of all algorithms for Example 4.1

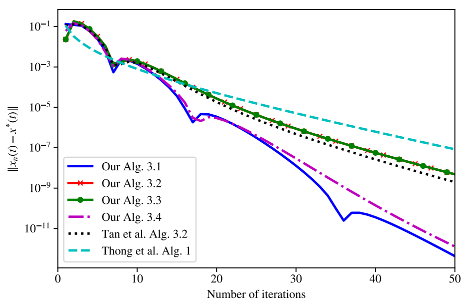
Example 4.2. We consider an example that appears in the infinite-dimensional Hilbert space $\mathcal{H} = L^2[0,1]$ with inner product

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

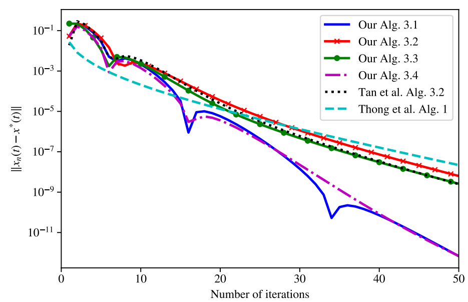
and induced norm

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

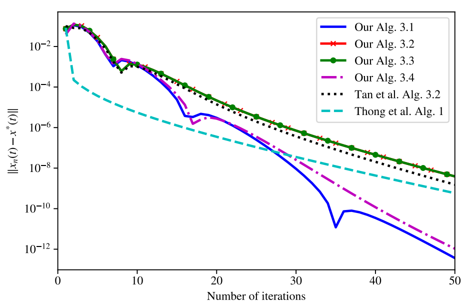
Let r, R be two positive real numbers such that $R/(k+1) < r/k < r < R$ for some $k > 1$. Take the feasible set as $C = \{x \in \mathcal{H} : \|x\| \leq r\}$. The operator $M : \mathcal{H} \to \mathcal{H}$ is given by $M(x) = (R - \|x\|)x$ for all $x \in \mathcal{H}$. Note that the operator M is pseudomonotone rather than monotone (see [20, Example 2]). 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 $R = 1.5$, $r = 1$ and $k = 1.1$. The solution of the (VIP) is $x^*(t) = 0$. The maximum number of iterations 50 is used as a common stopping criterion. Figure 2 shows the behaviors of $D_n = \|x_n(t) - x^*(t)\|$ generated by all algorithms under four different initial values $x_0(t) = x_1(t)$.



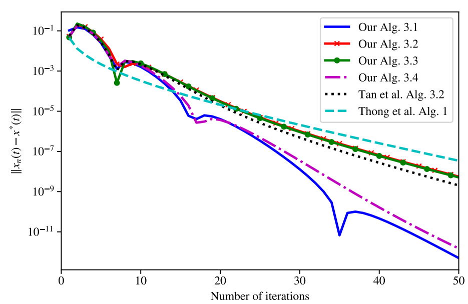
(a) $x_0 = x_1 = 2\cos(2t)$



(b) $x_0 = x_1 = 4t^2$



(c) $x_0 = x_1 = \log(3t)$



(d) $x_0 = x_1 = 2^t$

FIGURE 2. Numerical results of all algorithms for Example 4.2

Example 4.3. Consider the Hilbert space

$$\mathcal{H} = l_2 := \{x = (x_1, x_2, \ldots, x_i, \ldots) \mid \sum_{i=1}^{\infty} |x_i|^2 < +\infty\}$$

equipped with inner product

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

and induced norm $\|x\| = \sqrt{\langle x, x \rangle}$, $\forall x \in \mathcal{H}$. Let

$$C := \{x = (x_1, x_2, \ldots, x_i, \ldots) \in \mathcal{H} : |x_i| \leq 1/i,\ i = 1, 2, \ldots, n, \ldots\}.$$

Define an operator $M : C \to \mathcal{H}$ by

$$Mx = \left(\|x\| + \frac{1}{\|x\| + \varphi} \right) x$$

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 [23] for more details). In the following cases, we take $\varphi = 0.5$, and $\mathcal{H} = \mathbb{R}^m$ for different values of m. In this case, the feasible set C is a box

$$C = \left\{ x \in \mathbb{R}^m : \frac{-1}{i} \leq x_i \leq \frac{1}{i},\ i = 1, 2, \ldots, m \right\}.$$

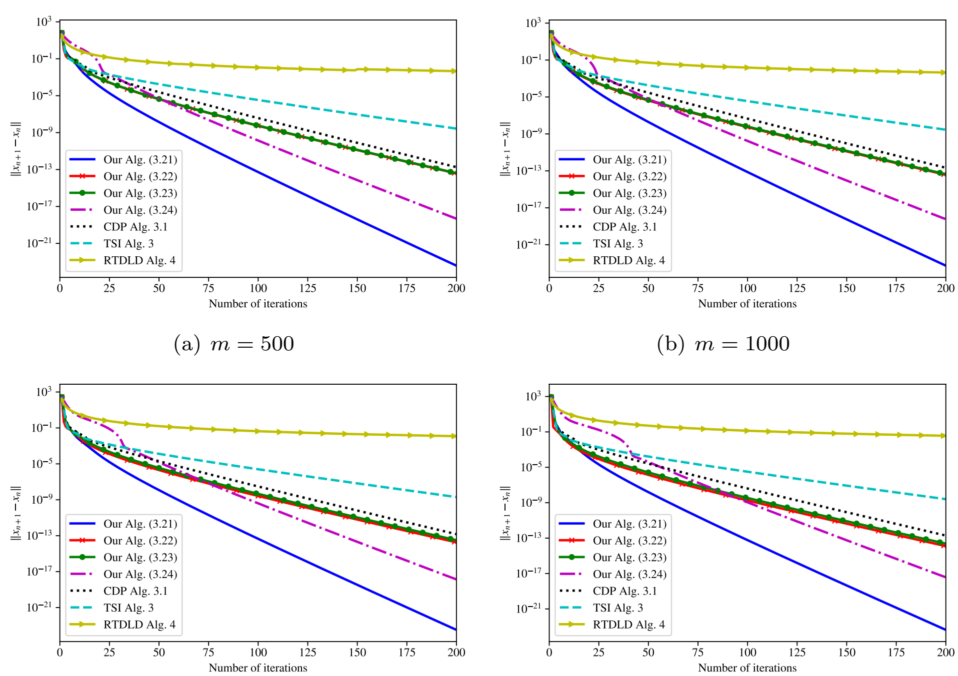
We compare the proposed Algorithms (3.21)–(3.24) with several previously known strongly convergent algorithms, including the Algorithm 3.1 introduced by Cai, Dong and Peng [4] (shortly, CDP Alg. 3.1), the Algorithm 3 suggested by Thong, Shehu and Iyiola [22] (shortly, TSI Alg. 3), and the Algorithm 4 proposed by Reich et al. [17] (shortly, RTDLD Alg. 4). Take

$$\alpha_n = 1/(n+1),\ f(x) = 0.1x, \delta = 2,\ \ell = 0.5,\ \text{and}\ \mu = 0.1$$

for all algorithms. Choose $\chi = 0.5/\mu$ for RTDLD Alg. 4. Set

$$\theta = 0.4 \text{ and } \epsilon_n = 100/(n+1)^2$$

for the suggested algorithms. The numerical performance of $D_n = \|x_{n+1} - x_n\|$ of all algorithms with four different dimensions is reported in Fig. 3.



(a) $m = 500$ (b) $m = 1000$

(c) $m = 10000$ (d) $m = 100000$

FIGURE 3. Numerical results of all algorithms for Example 4.3

Remark 4.1. From Example 4.1–4.3, we have the following observations.

(i) The proposed Algorithms 3.1–3.4 and Algorithms (3.21)–(3.24) are useful and converge quickly.
(ii) As shown in Figs. 1–3, the stated algorithms have higher accuracy than the previously known ones [21, 20, 4, 22, 17] under the same stopping conditions. These results are independent of the size of the dimension and the choice of initial values. Therefore, our suggested algorithms are efficient and robust.
(iii) Note that the operator M in Example 4.2 is pseudomonotone rather than monotone. The algorithms proposed in [13, 10] for solving the bilevel monotone variational inequality problem will not be applicable in this case. Moreover, the algorithms proposed in [21, 20] will not be available in Example 4.3 due to the fact that the mapping M involved in this example is uniformly continuous but not Lipschitz continuous.
(iv) It should be mentioned that Algorithm 3.2 and Algorithm 3.3 are equivalent in the following case. Indeed, if $(w_n - \chi_n M w_n) \in C$ always holds in the second step of Algorithm 3.2, then $y_n = w_n - \chi_n M w_n$ and thus $z_n = w_n - \chi_n M y_n$. On the other hand, if $(w_n - \chi_n M w_n) \in C$ always holds in Algorithm 3.3, then $z_n = w_n - \chi_n M y_n$. In conclusion, the proposed Algorithms 3.2 and 3.3 have the same numerical behavior in the case just described (see Figs. 1–3).

4.2. **Application to optimal control problems.** Next, we use the proposed algorithms to solve the variational inequality problem (VIP) that appears in optimal control problems. We recommend readers to refer to [16, 27] for detailed description of the problem. We compare the suggested iterative schemes (3.21)–(3.24) 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. [21] and the Algorithm (3.39) (in short, TLQ Alg. (3.39)) proposed by Tan, Liu and Qin [20]. The parameters of all algorithms are set as follows.

- In the proposed Algorithms (3.21)–(3.24), we set $N = 100$, $\theta = 0.01$, $\epsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\delta = 1$, $\ell = 0.5$, $\mu = 0.1$, $\alpha_n = \frac{10^{-4}}{n+1}$, and $f(x) = 0.1x$. Pick $\phi = 1.5$ for Algorithm (3.24).
- In the TLDCR Alg. (31), we choose $N = 100$, $\mu = 0.1$, $\chi_1 = 0.4$, $\phi = 1.5$, and $\alpha_n = \frac{10^{-4}}{n+1}$.
- In the TLQ Alg. (3.39), we take $N = 100$, $\theta = 0.01$, $\epsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\mu = 0.1$, $\chi_1 = 0.4$, $\alpha_n = \frac{10^{-4}}{n+1}$, 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 $D_n = \|p_{n+1} - p_n\| \leq 10^{-4}$, or maximum number of iterations which is set to 1000.

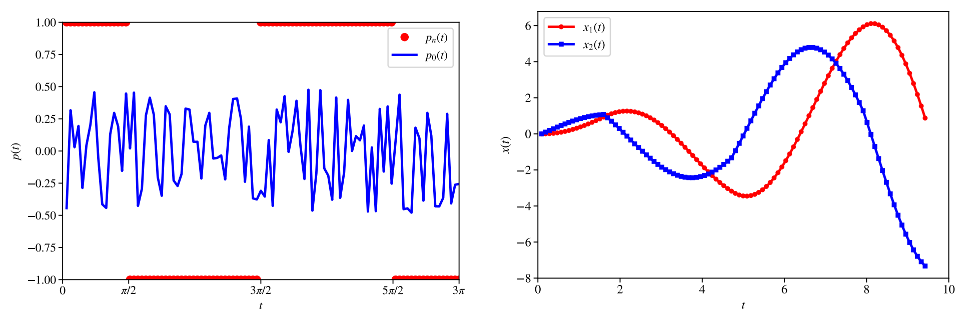
Example 4.4 (Control of a harmonic oscillator, see [15]).

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

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 (3.21).



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

FIGURE 4. Numerical results of the proposed Algorithm (3.21) for Example 4.4

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

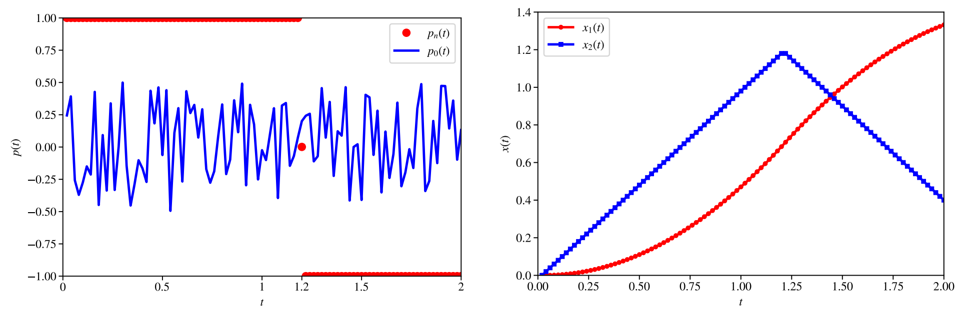
Example 4.5 (see [3]).

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

The exact optimal control of Example 4.5 is

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

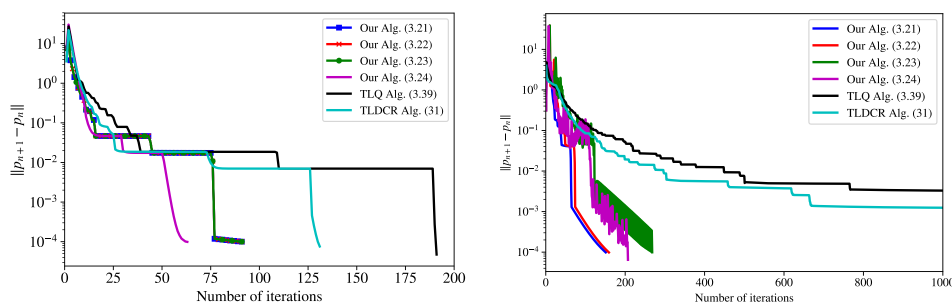
The approximate optimal control and the corresponding trajectories of the suggested Algorithm (3.21) are plotted in Fig. 5.



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

FIGURE 5. Numerical results of the proposed Algorithm (3.21) for Example 4.5

Finally, we compare the offered Algorithms (3.21)–(3.24) 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 1.



(a) Example 4.4 (b) Example 4.5

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

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

(i) The suggested Algorithms (3.21)–(3.24) can be applied to solve optimal control problems, and they perform well when the terminal function is linear or nonlinear.

(ii) As shown in Fig. 6 and Table 1, the proposed Algorithms (3.21)–(3.24) perform better when the terminal function is linear than when it is nonlinear. Moreover, the proposed Algorithms (3.21)–(3.24) outperform the existing methods in the literature [21, 20].

Table 1. Numerical results of all algorithms for Examples 4.4 and 4.5

Algorithms	Example 4.4			Example 4.5		
	Iter.	Time (s)	D_n	Iter.	Time (s)	D_n
Our Alg. (3.21)	90	0.0691	1.00E-04	150	0.0630	9.87E-05
Our Alg. (3.22)	90	0.0600	1.00E-04	158	0.0652	9.84E-05
Our Alg. (3.23)	90	0.0431	1.00E-04	267	0.1969	9.91E-05
Our Alg. (3.24)	62	0.0538	9.89E-05	206	0.1628	6.47E-05
TLQ Alg. (3.39)	190	0.1029	4.74E-05	1000	0.3127	3.29E-03
TLDCR Alg. (31)	130	0.0722	7.58E-05	1000	0.3134	1.24E-03

5. Conclusions

In this paper, we introduced four modified adaptive extragradient-type methods to solve bilevel variational inequality problems where the mapping involved is pseudomonotone and uniformly continuous. Our algorithms are inspired by the subgradient extragradient method, the Tseng's extragradient method, the projection and contraction method, and the hybrid steepest descent method. The strong convergence of the suggested algorithms is established under some suitable conditions imposed on the parameters. Finally, some numerical experiments are performed to verify the theoretical results. The algorithms presented in this paper improved and extended some known results for solving bilevel optimization problems and variational inequality problems.

References

[1] T.V. Anh, *A strongly convergent subgradient extragradient-halpern method for solving a class of bilevel pseudomonotone variational inequalities*, Vietnam J. Math., **45**(2017), 317-332.

[2] R.I. Boţ, E.R. Csetnek, P.T. Vuong, *The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces*, European J. Oper. Res., **287**(2020), 49-60.

[3] B. Bressan, B. Piccoli, *Introduction to the Mathematical Theory of Control*, San Francisco, American Institute of Mathematical Sciences, 2007.

[4] G. Cai, Q.L. Dong, Y. Peng, *Strong convergence theorems for solving variational inequality problems with pseudo-monotone and non-Lipschitz operators*, J. Optim. Theory Appl., **188**(2021), 447-472.

[5] Y. Censor, A. Gibali, S. Reich, *The subgradient extragradient method for solving variational inequalities in Hilbert space*, J. Optim. Theory Appl., **148**(2011), 318-335.

[6] S.V. Denisov, V.V. Semenov, L.M. Chabak, *Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators*, Cybernet. Systems Anal., **51**(2015), 757-765.

[7] S. Dempe, A. Zemkoho, *Bilevel Optimization*, Cham, Springer, 2020.

[8] Q.L. Dong, Y.J. Cho, L.L. Zhong, T.M. Rassias, *Inertial projection and contraction algorithms for variational inequalities*, J. Global Optim., **70**(2018), 687-704.

[9] B.S. He, *A class of projection and contraction methods for monotone variational inequalities*, Appl. Math. Optim., **35**(1997), 69-76.

[10] D.V. Hieu, A. Moudafi, *Regularization projection method for solving bilevel variational inequality problem*, Optim. Lett., **15**(2021), 205-229.

[11] K. Jantakarn, A. Kaewcharoen, *A Bregman hybrid extragradient method for solving pseudomonotone equilibrium and fixed point problems*, J. Nonlinear Funct. Anal., **2022**(2022), Article ID 6, pp. 1-17.

[12] G.M. Korpelevich, *The extragradient method for finding saddle points and other problems*, Ékonom. i Mat. Metody, **12**(1976), 747-756.

[13] P.E. Maingé, *A hybrid extragradient-viscosity method for monotone operators and fixed point problems*, SIAM J. Control Optim., **47**(2008), 1499-1515.

[14] F.U. Ogbuisi, Y. Shehu, *A new inertial relaxed Tseng extrgradient method for solving quasi-monotone bilevel variational inequality problems in Hilbert spaces*, J. Nonlinear Var. Anal., **7**(2023), 449-464.

[15] A. Pietrus, T. Scarinci, V.M. Veliov, *High order discrete approximations to Mayer's problems for linear systems*, SIAM J. Control Optim., **56**(2018), 102-119.

[16] J. Preininger, P.T. Vuong, *On the convergence of the gradient projection method for convex optimal control problems with bang-bang solutions*, Comput. Optim. Appl., **70**(2018), 221-238.

[17] S. Reich, D.V. Thong, Q.L. Dong, X.H. Li, V.T. Dung, *New algorithms and convergence theorems for solving variational inequalities with non-Lipschitz mappings*, Numer. Algorithms, **87**(2021), 527-549.

[18] S. Saejung, P. Yotkaew, *Approximation of zeros of inverse strongly monotone operators in Banach spaces*, Nonlinear Anal., **75**(2012), 742-750.

[19] Y. Shehu, X.H. Li, Q.L. Dong, *An efficient projection-type method for monotone variational inequalities in Hilbert spaces*, Numer. Algorithms., **84**(2020), 365-388.

[20] B. Tan, L. Liu, X. Qin, *Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems*, Jpn. J. Ind. Appl. Math., **38**(2021), 519-543.

[21] D.V. Thong, X.H. Li, Q.L. Dong, Y.J. Cho, T.M. Rassias, *A projection and contraction method with adaptive step sizes for solving bilevel pseudo-monotone variational inequality problems*, Optimization, **71**(2022), 2073-2096.

[22] D.V. Thong, Y. Shehu, O.S. Iyiola, *Weak and strong convergence theorems for solving pseudo-monotone variational inequalities with non-Lipschitz mappings*, Numer. Algorithms, **84**(2020), 795-823.

[23] D.V. Thong, Y. Shehu, O.S. Iyiola, *A new iterative method for solving pseudomonotone variational inequalities with non-Lipschitz operators*, Comput. Appl. Math., **39**(2020), Article ID 108.

[24] D.V. Thong, P.T. Vuong, *Modified Tseng's extragradient methods for solving pseudo-monotone variational inequalities*, Optimization, **68**(2019), 2207-2226.

[25] D.V. Thong, J. Yang, Y.J. Cho, T.M. Rassias, *Explicit extragradient-like method with adaptive stepsizes for pseudomonotone variational inequalities*, Optim. Lett., **15**(2021), 2181-2199.

[26] P. Tseng, *A modified forward-backward splitting method for maximal monotone mappings*, SIAM J. Control Optim., **38**(2000), 431-446.

[27] P.T. Vuong, Y. Shehu, *Convergence of an extragradient-type method for variational inequality with applications to optimal control problems*, Numer. Algorithms, **81**(2019), 269-291.

[28] I. Yamada, *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, pp. 473-504. Amsterdam: Elsevier, 2001.

Received: March 15, 2022; Accepted: July 7, 2022.