J. Appl. Numer. Optim. 4 (2022), No. 3, pp. 425-444
Available online at http://jano.biemdas.com
https://doi.org/10.23952/jano.4.2022.3.08

REVISITING INERTIAL SUBGRADIENT EXTRAGRADIENT ALGORITHMS FOR SOLVING BILEVEL VARIATIONAL INEQUALITY PROBLEMS

BING TAN[1,2], SONGXIAO LI[1,*], SUN YOUNG CHO[3]

[1] *Institute of Fundamental and Frontier Sciences,*
University of Electronic Science and Technology of China, Chengdu 611731, China
[2] *Department of Mathematics, University of British Columbia, Kelowna, BC, V1V 1V7, Canada*
[3] *Research Center for Interneural Computing, China Medical University, Taichung 40447, Taiwan*

Abstract. In this paper, four modified subgradient extragradient algorithms are proposed for solving bilevel pseudomonotone variational inequality problems in real Hilbert spaces. The proposed algorithms can work adaptively without the prior knowledge of the Lipschitz constant of the pseudomonotone mapping. Strong convergence theorems for the suggested algorithms are established under suitable and mild conditions. Finally, some numerical experiments and applications are performed to verify the efficiency of the proposed algorithms with respect to some previously known ones.
Keywords. Bilevel variational inequality problem; Inertial method; Subgradient extragradient method; Optimal control; Adaptive stepsize.

1. INTRODUCTION

Recall that the classical variational inequality problem 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 C is a nonempty, closed, and convex subset of a real Hilbert space $\mathscr{H}$, which is endowed with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$, and $M : C \to \mathscr{H}$ is an operator. It is known that the variational inequality model provides a general and useful framework for solving many problems in economics, engineering, data sciences, optimal control, mathematical programming, and other fields (see, e.g., [1–3]). Recently, many scholars proposed a large number of numerical algorithms to solve the variational inequality problem and its extensions in infinite-dimensional spaces; see, e.g., [4–12] and the references therein. Our focus in this paper is to investigate several numerical algorithms to solve the following bilevel variational inequality problem:

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

where $\mathrm{VI}(C, M)$ denotes the set of all solutions of (VIP) and $F : C \to \mathscr{H}$ is an operator.

*Corresponding author.
Email addresses: bingtan72@gmail.com (B. Tan), jyulsx@163.com (S. Li), ooly61@hotmail.com (S.Y. Cho).
Received October 10, 2022; Accepted November 7, 2022.

The bilevel variational inequality problem covers a number of nonlinear optimization problems, such as fixed point problems, quasi-variational inequality problems, complementary problems, saddle problems, and minimum norm problems. Therefore, it is useful to explore effective numerical algorithms for solving the (BVIP). Recently, various iterative algorithms were proposed for solving the (BVIP); see, e.g., [13–18]. Here, we state the drawbacks of the existing algorithms in the literature for solving (BVIP), which is the main motivation for this paper. Note that the algorithms introduced in [13–15] require two computations of the projection on the feasible set in each iteration. This may affect the computational efficiency of these algorithms if the projection on the feasible set is difficult to compute. Moreover, it is worth noting that the operator M in the algorithms proposed in [13, 15] are monotone, while the operator M in [14, 16–18] are pseudomonotone. On the other hand, the algorithms presented in [13, 14] require the prior information of the Lipschitz constant of the operator M, while the algorithms given in [15–18] can work adaptively since they use some adaptive step size criterion. However, the algorithms introduced in [15–18] generate a non-increasing sequence of step sizes, which may affect the execution efficiency of such adaptive algorithms. To overcome this difficulty, Liu and Yang [19] offered several methods with non-monotonic step size sequences for solving variational inequality problems. Furthermore, a common characteristic enjoyed by these algorithms suggested in [13–18] 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 some mappings that do not satisfy Lipschitz continuity (such as uniformly continuous mappings). Very recently, Cai, Dong, and Peng [20] introduced an iterative scheme with a new Armijo-type step size rule to solve pseudomonotone and non-Lipschitz continuous variational inequality problems in real Hilbert spaces. In the past decades, inertial techniques are widely used by scholars in algorithms to speed up the convergence of algorithms. They proposed many inertial algorithms for solving variational inequalities, bilevel variational inequalities, equilibrium problems, fixed point problems, split feasibility problems, and others; see, e.g., [10, 11, 18, 21, 22].

Inspired and motivated by the above works, we introduce four new adaptive modified subgradient extragradient methods for approximating the solutions of bilevel variational inequality problems in real Hilbert spaces. The operators M involved in our algorithms are Lipschitz continuous and pseudomonotone (the Lipschitz constant does not need to be known) and non-Lipschitz continuous and pseudomonotone. In addition, we use two new non-monotonic step size criteria that allow the proposed algorithms to work adaptively. The strong convergence of the iterative sequences generated by the proposed methods is established under suitable conditions. Some numerical experiments and applications are provided to verify the computational efficiency of the proposed algorithms.

This paper is organized as follows. Some basic definitions and lemmas that need to be used are stated in Sect. 2. The proposed algorithms and their convergence analysis are shown in Sect. 3. In Sect. 4, some numerical experiments are performed to demonstrate the computational efficiency of our algorithms. Finally, we conclude this paper with a brief summary in Sect. 5, the last section.

2. Preliminaries

In the whole paper, we use the symbol $x_n \to x$ (resp., $x_n \rightharpoonup x$) to represent the strong convergence (resp., weak convergence) of the sequence $\{x_n\}$ to x, and use $P_C : \mathscr{H} \to C$ to denote the metric projection from $\mathscr{H}$ onto C, i.e., $P_C(x) := \arg\min\{\|x-y\|, y \in C\}$. It is known that P_C has the following basic properties:

$$\langle x - P_C(x), y - P_C(x)\rangle \le 0, \quad \forall x \in \mathscr{H}, y \in C. \tag{2.1}$$

$$\|P_C(x) - y\|^2 \le \|x-y\|^2 - \|x - P_C(x)\|^2, \quad \forall x \in \mathscr{H}, y \in C. \tag{2.2}$$

$$\|P_C(x) - P_C(y)\|^2 \le \langle P_C(x) - P_C(y), x-y\rangle, \quad \forall x \in \mathscr{H}, y \in \mathscr{H}. \tag{2.3}$$

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

(i) *L-Lipschitz continuous* with $L > 0$ if $\|Mx - My\| \le L\|x-y\|, \forall x, y \in \mathscr{H}$ (If $L \in (0,1)$ then the mapping M is called a *contraction*).

(ii) *α-strongly monotone* if there exists a constant $\alpha > 0$ such that $\langle Mx - My, x-y\rangle \ge \alpha\|x-y\|^2, \forall x, y \in \mathscr{H}$.

(iii) *monotone* if $\langle Mx - My, x-y\rangle \ge 0, \forall x, y \in \mathscr{H}$.

(iv) *pseudomonotone* if $\langle Mx, y-x\rangle \ge 0 \Rightarrow \langle My, y-x\rangle \ge 0, \forall x, y \in \mathscr{H}$.

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

The following lemmas are crucial in the convergence analysis of our algorithms.

Lemma 2.1 ([16, 23])**.** *Let $\gamma > 0$ and $\alpha \in (0,1]$. Let $F : \mathscr{H} \to \mathscr{H}$ be a ζ-strongly monotone and L-Lipschitz continuous mapping (noting that $0 \le \zeta \le L$). Associating with a nonexpansive mapping $T : \mathscr{H} \to \mathscr{H}$, define a mapping $T^\gamma : \mathscr{H} \to \mathscr{H}$ by $T^\gamma x = (I - \alpha\gamma F)(Tx), \forall x \in \mathscr{H}$. Then, T^γ is contraction provided $\gamma \in \left(0, \frac{2\zeta}{L^2}\right)$, that is,*

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

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

Lemma 2.2 ([24])**.** *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*

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

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

3. Main results

In this section, we present four modified subgradient extragradient methods for solving the problem (BVIP). The following conditions are assumed to be satisfied in our algorithms.

(C1) The feasible set C is a nonempty, closed, and convex subset of a real Hilbert space $\mathscr{H}$.

(C2) The solution set of the problem (VIP) is nonempty, that is, $\mathrm{VI}(C, M) \ne \emptyset$.

(C3) The mapping $F : \mathscr{H} \to \mathscr{H}$ is L_F-Lipschitz continuous and ζ_F-strongly monotone on $\mathscr{H}$.

(C4) Let $\{\varepsilon_n\}$ be a positive sequence such that $\lim_{n\to\infty} \frac{\varepsilon_n}{\alpha_n} = 0$, where $\{\alpha_n\} \subset (0,1)$ satisfies $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=1}^{\infty} \alpha_n = \infty$.

3.1. **The first type of subgradient extragradient methods.** In this subsection, we introduce two new numerical algorithms for solving the (BVIP). First, a new iterative scheme with a non-monotonic step size criterion is given in Algorithm 3.1. We assume that the operator M in the suggested Algorithm 3.1 satisfies the following condition (C5).

(C5) The operator $M : \mathscr{H} \to \mathscr{H}$ is pseudomonotone, L_M-Lipschitz continuous, and sequentially weakly continuous on C. Choose a nonnegative real sequence $\{\xi_n\}$ such that $\sum_{n=1}^{\infty} \xi_n < +\infty$.

The Algorithm 3.1 is formulated as follows.

Algorithm 3.1

Initialization: Take $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0,1)$, $\beta \in (0, 2/(1+\mu))$, and $\gamma \in (0, 2\zeta_F / L_F^2)$. Select two sequences $\{\varepsilon_n\}$ and $\{\alpha_n\}$ to satisfy Condition (C4), and choose a sequence $\{\xi_n\}$ to satisfy Condition (C5). Let $x_0, x_1 \in \mathscr{H}$ 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 $u_n = x_n + \theta_n(x_n - x_{n-1})$, where

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

Step 2. Compute $y_n = P_C(u_n - \lambda_n M u_n)$. If $u_n = y_n$ or $My_n = 0$, then stop and y_n is a solution of (VIP); Otherwise, go to **Step 3**.

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

$$T_n = \{x \in \mathscr{H} : \langle u_n - \lambda_n M u_n - y_n, x - y_n \rangle \leq 0\}.$$

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

$$\lambda_{n+1} = \begin{cases} \min\left\{ \mu \dfrac{\|u_n - y_n\|^2 + \|z_n - y_n\|^2}{2\langle M u_n - M y_n, z_n - y_n \rangle}, \lambda_n + \xi_n \right\}, & \text{if } \langle M u_n - M y_n, z_n - y_n \rangle > 0; \\ \lambda_n + \xi_n, & \text{otherwise.} \end{cases} \tag{3.2}$$

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

Remark 3.1. We note here that the inertial calculation criterion (3.1) is easy to implement since the term $\|x_n - x_{n-1}\|$ is known before calculating θ_n. Moreover, it follows from (3.1) 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 $\theta_n \|x_n - x_{n-1}\| \leq \varepsilon_n$ for all $n \geq 1$, which together with $\lim_{n\to\infty} \frac{\varepsilon_n}{\alpha_n} = 0$ implies that $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n} \|x_n - x_{n-1}\| \leq \lim_{n\to\infty} \frac{\varepsilon_n}{\alpha_n} = 0$.

Remark 3.2. In Algorithm 3.1, if $u_n = y_n$ or $My_n = 0$, then stop and y_n is a solution of (VIP); this can be easily verified according to the property of the projection (2.1). On the other hand, combining the fact that $\mathrm{VI}(C, M)$ is a closed and convex set and Conditions (C2) and (C3) yields a unique solution for (BVIP) (see, e.g., [15, 25]).

The following lemmas are useful in the convergence analysis of Algorithm 3.1.

Lemma 3.1. *Suppose that Condition (C5) holds. Then the sequence* $\{\lambda_n\}$ *generated by* (3.2) *is well defined and* $\lim_{n\to\infty}\lambda_n=\lambda$ *and* $\lambda\in[\min\{\mu/L_M,\lambda_1\},\lambda_1+\Xi]$, *where* $\Xi=\sum_{n=1}^{\infty}\xi_n$.

Proof. The proof of this lemma is very similar to Lemma 3.1 in [19]. Therefore we omit the details. □

Lemma 3.2. *Assume that Condition (C5) holds. Let* $\{z_n\}$ *be a sequence generated by Algorithm 3.1. Then, for all* $p\in\mathrm{VI}(C,M)$,

$$\|z_n-p\|^2\le\|u_n-p\|^2-\beta_n^*\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right),$$

where $\beta_n^*=2-\beta-\frac{\beta\mu\lambda_n}{\lambda_{n+1}}$ *if* $\beta\in[1,2/(1+\mu))$ *and* $\beta_n^*=\beta-\frac{\beta\mu\lambda_n}{\lambda_{n+1}}$ *if* $\beta\in(0,1)$.

Proof. From the definition of z_n and the property of projection (2.2), we have

$$\begin{aligned}\|z_n-p\|^2&=\|P_{T_n}(u_n-\beta\lambda_nMy_n)-p\|^2\\&\le\|u_n-\beta\lambda_nMy_n-p\|^2-\|u_n-\beta\lambda_nMy_n-z_n\|^2\\&=\|u_n-p\|^2+(\beta\lambda_n)^2\|My_n\|^2-2\langle u_n-p,\beta\lambda_nMy_n\rangle-\|u_n-z_n\|^2\\&\quad-(\beta\lambda_n)^2\|My_n\|^2+2\langle u_n-z_n,\beta\lambda_nMy_n\rangle\\&=\|u_n-p\|^2-\|u_n-z_n\|^2-2\langle\beta\lambda_nMy_n,z_n-p\rangle\\&=\|u_n-p\|^2-\|u_n-z_n\|^2-2\langle\beta\lambda_nMy_n,z_n-y_n\rangle-2\langle\beta\lambda_nMy_n,y_n-p\rangle.\end{aligned}\tag{3.3}$$

Since $p\in\mathrm{VI}(C,M)$ and $y_n\in C$, we obtain $\langle Mp,y_n-p\rangle\ge0$. By the pseudomonotonicity of the mapping M, we obtain $\langle My_n,y_n-p\rangle\ge0$. Thus, inequality (3.3) reduces to

$$\|z_n-p\|^2\le\|u_n-p\|^2-\|u_n-z_n\|^2-2\langle\beta\lambda_nMy_n,z_n-y_n\rangle.\tag{3.4}$$

Now we estimate $2\langle\beta\lambda_nMy_n,z_n-y_n\rangle$. Note that

$$-\|u_n-z_n\|^2=-\|u_n-y_n\|^2-\|y_n-z_n\|^2+2\langle u_n-y_n,z_n-y_n\rangle.\tag{3.5}$$

In addition,

$$\begin{aligned}\langle u_n-y_n,z_n-y_n\rangle&=\langle u_n-y_n-\lambda_nMu_n+\lambda_nMu_n-\lambda_nMy_n+\lambda_nMy_n,z_n-y_n\rangle\\&=\langle u_n-\lambda_nMu_n-y_n,z_n-y_n\rangle+\lambda_n\langle Mu_n-My_n,z_n-y_n\rangle\\&\quad+\langle\lambda_nMy_n,z_n-y_n\rangle.\end{aligned}\tag{3.6}$$

Since $z_n\in T_n$, one has

$$\langle u_n-\lambda_nMu_n-y_n,z_n-y_n\rangle\le0.\tag{3.7}$$

According to the definition of λ_{n+1}, we deduce

$$\langle Mu_n-My_n,z_n-y_n\rangle\le\frac{\mu}{2\lambda_{n+1}}\|u_n-y_n\|^2+\frac{\mu}{2\lambda_{n+1}}\|z_n-y_n\|^2.\tag{3.8}$$

Substituting (3.6), (3.7), and (3.8) into (3.5), we obtain

$$-\|u_n-z_n\|^2\le-\left(1-\frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right)+2\langle\lambda_nMy_n,z_n-y_n\rangle,$$

which implies that

$$-2\langle\beta\lambda_n My_n, z_n-y_n\rangle \le -\beta\left(1-\frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right)+\beta\|u_n-z_n\|^2. \tag{3.9}$$

Combining (3.4) and (3.9), we conclude that

$$\begin{aligned}\|z_n-p\|^2 \le \|u_n-p\|^2 &-\beta\left(1-\frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right)\\ &-(1-\beta)\|u_n-z_n\|^2.\end{aligned} \tag{3.10}$$

Note that

$$\|u_n-z_n\|^2 \le 2\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right),$$

which yields that

$$-(1-\beta)\|u_n-z_n\|^2 \le -2(1-\beta)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right), \quad \forall\beta\ge 1.$$

This together with (3.10) implies

$$\|z_n-p\|^2 \le \|u_n-p\|^2-\left(2-\beta-\frac{\beta\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right), \quad \forall\beta\ge 1.$$

On the other hand, if $\beta\in(0,1)$, then we obtain

$$\|z_n-p\|^2 \le \|u_n-p\|^2-\beta\left(1-\frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right), \quad \forall\beta\in(0,1).$$

This completes the proof. □

Remark 3.3. It follows from Lemmas 3.1 and 3.2 that

$$\lim_{n\to\infty}\beta_n^* = \begin{cases} 2-\beta-\beta\mu, & \text{if } \beta\in[1,2/(1+\mu));\\ \beta-\beta\mu, & \text{if } \beta\in(0,1).\end{cases}$$

Thus, there exists a constant n_0 such that $\beta_n^*>0$ for all $n\ge n_0$ in Lemma 3.2 always holds.

Lemma 3.3. *Suppose that Condition (C5) holds. Let $\{u_n\}$ and $\{y_n\}$ be two sequences formulated by Algorithm 3.1. If there exists a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that $\{u_{n_k}\}$ converges weakly to $z\in\mathscr{H}$ and $\lim_{k\to\infty}\|u_{n_k}-y_{n_k}\|=0$, then $z\in \mathrm{VI}(C,M)$.*

Proof. The proof of this lemma follows that of Lemma 3.3 in [26] and thus it is omitted. □

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

Proof. First, we show that the sequence $\{x_n\}$ is bounded. It follows from Lemma 3.2 and Remark 3.3 that

$$\|z_n-p\| \le \|u_n-p\|, \quad \forall n\ge n_0. \tag{3.11}$$

By the definition of u_n, one has

$$\|u_n-p\| \le \alpha_n\cdot\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|+\|x_n-p\|. \tag{3.12}$$

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

$$\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \le Q_1, \quad \forall n \ge 1. \tag{3.13}$$

Combining (3.11), (3.12), and (3.13), we obtain

$$\|z_n - p\| \le \|u_n - p\| \le \|x_n - p\| + \alpha_n Q_1, \quad \forall n \ge n_0. \tag{3.14}$$

Using Lemma 2.1 and (3.11), it follows that

$$\begin{aligned}
\|x_{n+1} - p\| &= \|(I - \alpha_n\gamma F)z_n - (I - \alpha_n\gamma F)p - \alpha_n\gamma F p\| \\
&\le (1 - \alpha_n\eta)\|z_n - p\| + \alpha_n\gamma\|Fp\| \\
&\le (1 - \alpha_n\eta)\|x_n - p\| + \alpha_n\eta \cdot \left(\frac{Q_1}{\eta} + \frac{\gamma}{\eta}\|Fp\|\right) \\
&\le \max\left\{\frac{Q_1 + \gamma\|Fp\|}{\eta}, \|x_n - p\|\right\}, \ \forall n \ge n_0 \\
&\le \cdots \le \max\left\{\frac{Q_1 + \gamma\|Fp\|}{\eta}, \|x_{n_0} - p\|\right\},
\end{aligned}$$

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

Using Lemma 2.1 and the inequality $\|x+y\|^2 \le \|x\|^2 + 2\langle y, x+y\rangle, \forall x,y \in \mathscr{H}$, one has

$$\begin{aligned}
\|x_{n+1} - p\|^2 &= \|(I - \alpha_n\gamma F)z_n - (I - \alpha_n\gamma F)p - \alpha_n\gamma F p\|^2 \\
&\le (1 - \alpha_n\eta)^2\|z_n - p\|^2 + 2\alpha_n\gamma\langle Fp, p - x_{n+1}\rangle \\
&\le \|z_n - p\|^2 + \alpha_n Q_2
\end{aligned} \tag{3.15}$$

for some $Q_2 > 0$. In the light of Lemma 3.2, we obtain

$$\|x_{n+1} - p\|^2 \le \|u_n - p\|^2 - \beta_n^*\left(\|y_n - u_n\|^2 + \|z_n - y_n\|^2\right) + \alpha_n Q_2. \tag{3.16}$$

It follows from (3.14) that

$$\begin{aligned}
\|u_n - p\|^2 &\le (\|x_n - p\| + \alpha_n Q_1)^2 \\
&= \|x_n - p\|^2 + \alpha_n\left(2Q_1\|x_n - p\| + \alpha_n Q_1^2\right) \\
&\le \|x_n - p\|^2 + \alpha_n Q_3
\end{aligned} \tag{3.17}$$

for some $Q_3 := \sup_{n\in\mathbb{N}}\{2Q_1\|x_n - p\| + \alpha_n Q_1^2\} > 0$. Combining (3.16) and (3.17), we deduce

$$\beta_n^*\left(\|y_n - u_n\|^2 + \|z_n - y_n\|^2\right) \le \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n Q_4, \tag{3.18}$$

where $Q_4 := Q_2 + Q_3$.

From the definition of u_n, one sees that

$$\|u_n - p\|^2 \le \|x_n - p\|^2 + 2\theta_n\|x_n - p\|\|x_n - x_{n-1}\| + \theta_n^2\|x_n - x_{n-1}\|^2. \tag{3.19}$$

Combining (3.11) and (3.15), we obtain

$$\|x_{n+1} - p\|^2 \le (1 - \alpha_n\eta)\|u_n - p\|^2 + 2\alpha_n\gamma\langle Fp, p - x_{n+1}\rangle. \tag{3.20}$$

Substituting (3.19) into (3.20), we obtain

$$\begin{aligned}\|x_{n+1}-p\|^2 &\leq (1-\alpha_n\eta)\|x_n-p\|^2+2\alpha_n\gamma\langle Fp, p-x_{n+1}\rangle \\ &\quad +\theta_n\|x_n-x_{n-1}\|(2\|x_n-p\|+\theta\|x_n-x_{n-1}\|) \\ &\leq (1-\alpha_n\eta)\|x_n-p\|^2+\alpha_n\eta\left[\frac{2\gamma}{\eta}\langle Fp, p-x_{n+1}\rangle+\frac{3Q\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|\right],\end{aligned} \tag{3.21}$$

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

Finally, we need to prove that the sequence $\{\|x_n-p\|^2\}$ converges to zero. By Lemma 2.2 and (3.21), one assumes that $\{\|x_{n_k}-p\|^2\}$ is a subsequence of $\{\|x_n-p\|^2\}$ satisfying

$$\liminf_{k\to\infty}\left(\|x_{n_k+1}-p\|^2-\|x_{n_k}-p\|^2\right)\geq 0.$$

From (3.18), Remark 3.3, and the assumption on $\{\alpha_n\}$, one infers that

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

which implies that

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

Therefore, we obtain

$$\lim_{k\to\infty}\|z_{n_k}-u_{n_k}\|=0. \tag{3.22}$$

Moreover, we can show that

$$\|x_{n_k+1}-z_{n_k}\|=\alpha_{n_k}\gamma\|Fz_{n_k}\|\to 0 \text{ as } n\to\infty, \tag{3.23}$$

and

$$\|x_{n_k}-u_{n_k}\|=\alpha_{n_k}\cdot\frac{\theta_{n_k}}{\alpha_{n_k}}\|x_{n_k}-x_{n_k-1}\|\to 0 \text{ as } n\to\infty. \tag{3.24}$$

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

$$\|x_{n_k+1}-x_{n_k}\|\leq\|x_{n_k+1}-z_{n_k}\|+\|z_{n_k}-u_{n_k}\|+\|u_{n_k}-x_{n_k}\|\to 0 \text{ as } n\to\infty. \tag{3.25}$$

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

$$\limsup_{k\to\infty}\langle Fp, p-x_{n_k}\rangle=\langle Fp, p-z\rangle\leq 0. \tag{3.26}$$

Using (3.25) and (3.26), we obtain

$$\limsup_{k\to\infty}\langle Fp, p-x_{n_k+1}\rangle\leq\limsup_{k\to\infty}\langle Fp, p-x_{n_k}\rangle\leq 0. \tag{3.27}$$

From $\lim_{n\to\infty}\frac{\theta_n}{\alpha_n}\|x_n-x_{n-1}\|=0$ and (3.27), we deduce

$$\limsup_{k\to\infty}\left[\frac{2\gamma}{\eta}\left\langle Fp,p-x_{n_k+1}\right\rangle+\frac{3Q\theta_{n_k}}{\alpha_{n_k}\eta}\|x_{n_k}-x_{n_k-1}\|\right]\le 0. \tag{3.28}$$

Combining $\sum_{n=1}^{\infty}\alpha_n\eta=\infty$, (3.21), and (3.28), 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. □

Next, we present an iterative scheme (see Algorithm 3.2 below) for solving the (BVIP) with a pseudomonotone and non-Lipschitz continuous operator. In our Algorithm 3.2, we replace the condition (C5) in Algorithm 3.1 with the following condition (C6).

(C6) The operator $M:\mathscr{H}\to\mathscr{H}$ is pseudomonotone, uniformly continuous on $\mathscr{H}$, and sequentially weakly continuous on C.

Now we are ready to describe the proposed Algorithm 3.2.

Algorithm 3.2

Initialization: Take $\theta>0$, $\delta>0$, $\ell\in(0,1)$, $\mu\in(0,1)$, $\beta\in(0,2/(1+\mu))$, and $\gamma\in(0,2\zeta_F/L_F^2)$. Select two sequences $\{\epsilon_n\}$ and $\{\alpha_n\}$ to satisfy Condition (C4). Let $x_0,x_1\in\mathscr{H}$ 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 $u_n=x_n+\theta_n(x_n-x_{n-1})$, where θ_n is defined in (3.1).

Step 2. Compute $y_n=P_C(u_n-\lambda_n Mu_n)$. If $u_n=y_n$ or $My_n=0$, then stop and y_n is a solution of (VIP). Otherwise, go to **Step 3**.

Step 3. Compute $z_n=P_{T_n}(u_n-\beta\lambda_n My_n)$, where the half space T_n is defined by

$$T_n=\{x\in\mathscr{H}:\langle u_n-\lambda_n Mu_n-y_n,x-y_n\rangle\le 0\},$$

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

$$\delta\ell^m\langle My_n-Mu_n,y_n-z_n\rangle\le\frac{\mu}{2}\left[\|u_n-y_n\|^2+\|y_n-z_n\|^2\right]. \tag{3.29}$$

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

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

Remark 3.4. Suppose that Condition (C6) holds. Let $\{u_n\}$ and $\{y_n\}$ be two sequences generated by Algorithm 3.2. Following the proof of Lemma 3.1 in [27], we can obtain that the Armijo-type criterion (3.29) is well defined. Moreover, Lemma 3.3 still holds by replacing x_n in the proof process in Lemma 3.2 of [20] with u_n.

Lemma 3.4. *Assume that Condition (C6) hold. Let $\{z_n\}$ be a sequence generated by Algorithm 3.2. Then,*

$$\|z_n-p\|^2\le\|u_n-p\|^2-\beta^{**}\left(\|u_n-y_n\|^2+\|z_n-y_n\|^2\right),\quad\forall p\in\mathrm{VI}(C,M),$$

where $\beta^{**}=2-\beta-\beta\mu$ if $\beta\in[1,2/(1+\mu))$ and $\beta^{**}=\beta-\beta\mu$ if $\beta\in(0,1)$.

Proof. The conclusion is easily obtained by following the proof of Lemma 3.2. Therefore we omit the details. □

Theorem 3.2. *Assume that Conditions (C1)–(C4) and (C6) holds. Then the sequence* $\{x_n\}$ *generated by Algorithm 3.2 converges strongly to the unique solution of the* (BVIP)*.*

Proof. With the aid of Lemma 3.4, we conclude from Theorem 3.1 the desired conclusion immediately. □

3.2. **The second type of subgradient extragradient methods.** In this section, we introduce two new modified subgradient extragradient algorithms to solve the (BVIP). Now, we present another version of the proposed Algorithm 3.1. The scheme is shown in Algorithm 3.3 below.

Algorithm 3.3

Initialization: Take $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0,1)$, $\beta \in (1/(2-\mu), 1/\mu)$ and $\gamma \in (0, 2\zeta_F/L_F^2)$. Select two sequences $\{\varepsilon_n\}$, and $\{\alpha_n\}$ to satisfy Condition (C4), and choose a sequence $\{\xi_n\}$ to satisfy Condition (C5). Let $x_0, x_1 \in \mathscr{H}$ 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 $u_n = x_n + \theta_n(x_n - x_{n-1})$, where θ_n is defined in (3.1).
Step 2. Compute $y_n = P_C(u_n - \beta\lambda_n M u_n)$. If $u_n = y_n$ or $My_n = 0$, then stop and y_n is a solution of (VIP). Otherwise, go to **Step 3**.
Step 3. Compute $z_n = P_{H_n}(u_n - \lambda_n M y_n)$, where the half space H_n is defined by

$$H_n = \{x \in \mathscr{H} : \langle u_n - \beta\lambda_n M u_n - y_n, x - y_n\rangle \leq 0\}.$$

Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$. Update the next step size λ_{n+1} by (3.2).
Set $n := n+1$ and go to **Step 1**.

Lemma 3.5. *Assume that Condition (C5) holds. Let* $\{z_n\}$ *be a sequence generated by Algorithm 3.3. Then, for all* $p \in \mathrm{VI}(C,M)$*,*

$$\|z_n - p\|^2 \leq \|u_n - p\|^2 - \beta_n^\dagger \left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right),$$

where $\beta_n^\dagger = 2 - \frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}$ *if* $\beta \in (1/(2-\mu), 1]$ *and* $\beta_n^\dagger = \frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}$ *if* $\beta \in (1, 1/\mu)$*.*

Proof. From (3.3) and (3.4), we obtain

$$\|z_n - p\|^2 \leq \|u_n - p\|^2 - \|u_n - z_n\|^2 - 2\langle \lambda_n M y_n, z_n - y_n\rangle. \tag{3.30}$$

Now we estimate $2\langle \lambda_n M y_n, z_n - y_n\rangle$. Note that

$$-\|u_n - z_n\|^2 = -\|u_n - y_n\|^2 - \|y_n - z_n\|^2 + 2\langle u_n - y_n, z_n - y_n\rangle. \tag{3.31}$$

In addition,

$$\begin{aligned} &\langle u_n - y_n, z_n - y_n\rangle \\ &= \langle u_n - y_n - \beta\lambda_n M u_n + \beta\lambda_n M u_n - \beta\lambda_n M y_n + \beta\lambda_n M y_n, z_n - y_n\rangle \\ &= \langle u_n - \beta\lambda_n M u_n - y_n, z_n - y_n\rangle + \beta\lambda_n \langle M u_n - M y_n, z_n - y_n\rangle \\ &\quad + \langle \beta\lambda_n M y_n, z_n - y_n\rangle. \end{aligned} \tag{3.32}$$

Since $z_n \in H_n$, one obtains

$$\langle u_n - \beta\lambda_n M u_n - y_n, z_n - y_n\rangle \leq 0. \tag{3.33}$$

According to the definition of λ_{n+1}, we have

$$\langle Mu_n - My_n, z_n - y_n \rangle \le \frac{\mu}{2\lambda_{n+1}} \|u_n - y_n\|^2 + \frac{\mu}{2\lambda_{n+1}} \|z_n - y_n\|^2. \tag{3.34}$$

Substituting (3.32), (3.33), and (3.34) into (3.31), we obtain

$$-\|u_n - z_n\|^2 \le -\left(1 - \frac{\beta\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right) + 2\beta \langle \lambda_n M y_n, z_n - y_n \rangle,$$

which implies that

$$-2\langle \lambda_n M y_n, z_n - y_n \rangle \le -\left(\frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right) + \frac{1}{\beta}\|u_n - z_n\|^2. \tag{3.35}$$

Combining (3.30) and (3.35), we conclude that

$$\begin{aligned} \|z_n - p\|^2 \le \|u_n - p\|^2 &- \left(\frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right) \\ &- \left(1 - \frac{1}{\beta}\right)\|u_n - z_n\|^2. \end{aligned} \tag{3.36}$$

Note that

$$\|u_n - z_n\|^2 \le 2\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right),$$

which yields that

$$-\left(1 - \frac{1}{\beta}\right)\|u_n - z_n\|^2 \le -2\left(1 - \frac{1}{\beta}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right), \quad \forall \beta \in (0,1].$$

This together with (3.36) implies

$$\|z_n - p\|^2 \le \|u_n - p\|^2 - \left(2 - \frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right), \quad \forall \beta \in (0,1].$$

On the other hand, if $\beta > 1$, then

$$\|z_n - p\|^2 \le \|u_n - p\|^2 - \left(\frac{1}{\beta} - \frac{\mu\lambda_n}{\lambda_{n+1}}\right)\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right), \quad \forall \beta > 1.$$

This completes the proof. □

Remark 3.5. It follows from Lemmas 3.1 and 3.5 that

$$\lim_{n\to\infty} \beta_n^\dagger = \begin{cases} 2 - \frac{1}{\beta} - \mu, & \text{if } \beta \in (1/(2-\mu), 1]; \\ \frac{1}{\beta} - \mu, & \text{if } \beta \in (1, 1/\mu). \end{cases}$$

Thus, there exists a constant n_0 such that $\beta_n^\dagger > 0$ for all $n \ge n_0$ in Lemma 3.5 always holds.

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

Proof. The proof follows almost in the same way as that of Theorem 3.1 but we apply Lemma 3.5 in place of Lemma 3.2. □

Finally, the last iterative scheme proposed in this paper for solving (BVIP) is shown in Algorithm 3.4 below.

Algorithm 3.4

Initialization: Take $\theta > 0$, $\delta > 0$, $\ell \in (0,1)$, $\mu \in (0,1)$, $\beta \in (1/(2-\mu), 1/\mu)$ and $\gamma \in (0, 2\zeta_F/L_F^2)$. Select two sequences $\{\varepsilon_n\}$ and $\{\alpha_n\}$ to satisfy Condition (C4). Let $x_0, x_1 \in \mathscr{H}$ 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 $u_n = x_n + \theta_n(x_n - x_{n-1})$, where θ_n is defined in (3.1).
Step 2. Compute $y_n = P_C(u_n - \beta\lambda_n M u_n)$. If $u_n = y_n$ or $My_n = 0$, then stop and y_n is a solution of (VIP). Otherwise, go to **Step 3**.
Step 3. Compute $z_n = P_{H_n}(u_n - \lambda_n M y_n)$, where the half space H_n is defined by

$$H_n = \{x \in \mathscr{H} : \langle u_n - \beta\lambda_n M u_n - y_n, x - y_n\rangle \leq 0\},$$

and $\lambda_n := \delta\ell^{m_n}$ and m_n is the smallest nonnegative integer m satisfying (3.29).
Step 4. Compute $x_{n+1} = z_n - \alpha_n \gamma F z_n$.
Set $n := n+1$ and go to **Step 1**.

Lemma 3.6. *Assume that Condition (C6) hold. Let $\{z_n\}$ be a sequence generated by Algorithm 3.4. Then,*

$$\|z_n - p\|^2 \leq \|u_n - p\|^2 - \beta^\ddagger\left(\|u_n - y_n\|^2 + \|z_n - y_n\|^2\right), \quad \forall p \in \mathrm{VI}(C, M),$$

where $\beta^\ddagger = 2 - \frac{1}{\beta} - \mu$ if $\beta \in (1/(2-\mu), 1]$ and $\beta^\ddagger = \frac{1}{\beta} - \mu$ if $\beta \in (1, 1/\mu)$.

Proof. The proof follows that of Lemma 3.5. So it is omitted here. □

Theorem 3.4. *Assume that Conditions (C1)–(C4) and (C6) holds. Then the sequence $\{x_n\}$ generated by Algorithm 3.4 converges strongly to the unique solution of the (BVIP).*

Proof. The proof follows almost in the same way as that of Theorem 3.1 but we apply Lemma 3.6 in place of Lemma 3.2. □

Remark 3.6. We have the following observations for the proposed algorithms.

- Notice that if $\beta = 1$, then the proposed Algorithm 3.1 (respectively, Algorithm 3.2) and Algorithm 3.3 (respectively, Algorithm 3.4) are equivalent. If $\beta = 1$ and $\xi_n = 0$ in the proposed Algorithm 3.1, then it degenerates to Algorithm 3.1 introduced in [18].
- The four algorithms obtained in this paper can solve the bilevel pseudomonotone variational inequality problem, while the algorithms suggested in [13, 15] can only solve the bilevel monotone variational inequality problem. On the other hand, we replace the Lipschitz continuity of the mapping M in the literature [13–18] with the uniform continuity of the mapping M in the proposed Algorithms 3.2 and 3.4. Therefore, our suggested schemes have a wider range of applications. In addition, we make two modifications to the subgradient extragradient method introduced by Censor, Gibali, and Reich [5–7] and apply a new non-monotonic step size criterion. These changes allow our algorithms to converge faster than some known ones (see numerical experiments in Sect. 4).

- Set $F(x) = x - f(x)$ in Algorithms 3.1–3.4, where the mapping $f : \mathscr{H} \to \mathscr{H}$ is ρ-contraction. It can be easily verified that the mapping $F : \mathscr{H} \to \mathscr{H}$ is $(1+\rho)$-Lipschitz continuous and $(1-\rho)$-strongly monotone. In this situation, by picking $\gamma = 1$, we obtain

$$x_{n+1} = z_n - \alpha_n \gamma F z_n = (1-\alpha_n) z_n + \alpha_n f(z_n).$$

Thus, we obtain four new inertial modified subgradient extragradient algorithms for solving the (VIP) in real Hilbert spaces.

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 compare them with some algorithms in the literature [17, 18]. 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.

4.1. Theoretical examples.

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 an $m \times m$ matrix with their entries being generated in $(0,1)$, D is an $m \times m$ skew-symmetric matrix with their entries being generated in $(-1,1)$, K is an $m \times m$ 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 ζ_F-strongly monotone with $L_F = \max\{\mathrm{eig}(G)\}$ and $\zeta_F = \min\{\mathrm{eig}(G)\}$, where $\mathrm{eig}(G)$ represents all eigenvalues of G. Next, we consider the following fractional programming problem:

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

where

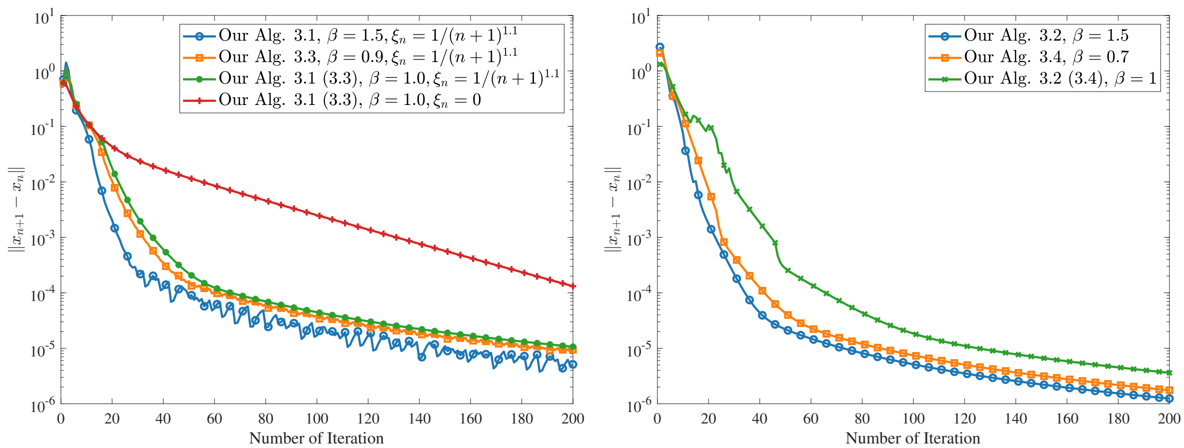
$$Q = \begin{bmatrix} 5 & -1 & 2 & 0 & 2 \\ -1 & 6 & -1 & 3 & 0 \\ 2 & -1 & 3 & 0 & 1 \\ 0 & 3 & 0 & 5 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{bmatrix}, a = \begin{bmatrix} 1 \\ 2 \\ -1 \\ -2 \\ 1 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 1 \end{bmatrix}, a_0 = -2, b_0 = 20.$$

It is easy to check that Q is symmetric and positive definite in $\mathbb{R}^5$ and hence f is pseudo-convex on C. Let $M(x) := \nabla f(x)$, where $\nabla f(x)$ denotes the gradient of the function $f(x)$. It is known that the mapping M is pseudomonotone and Lipschitz continuous. We compare the proposed Algorithms 3.1–3.4 with the Algorithm 1 introduced by Thong et al. [17] and the Algorithm 3.2 suggested by Tan, Liu, and Qin [18]. The parameters of all the algorithms are set as follows.

- In the proposed Algorithms 3.1–3.4, we set $\theta = 0.3$, $\varepsilon_n = 100/(n+1)^2$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\zeta_F/L_F^2$. Choose $\mu = 0.1$ and $\lambda_1 = 1$ for Algorithms 3.1 and 3.3. Pick $\delta = 1$, $\ell = 0.5$, and $\mu = 0.1$ for Algorithms 3.2 and 3.4.
- In the Algorithm 1 introduced by Thong et al. [17], we choose $\mu = 0.1$, $\lambda_1 = 1$, $\phi = 1.5$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\zeta_F/L_F^2$.

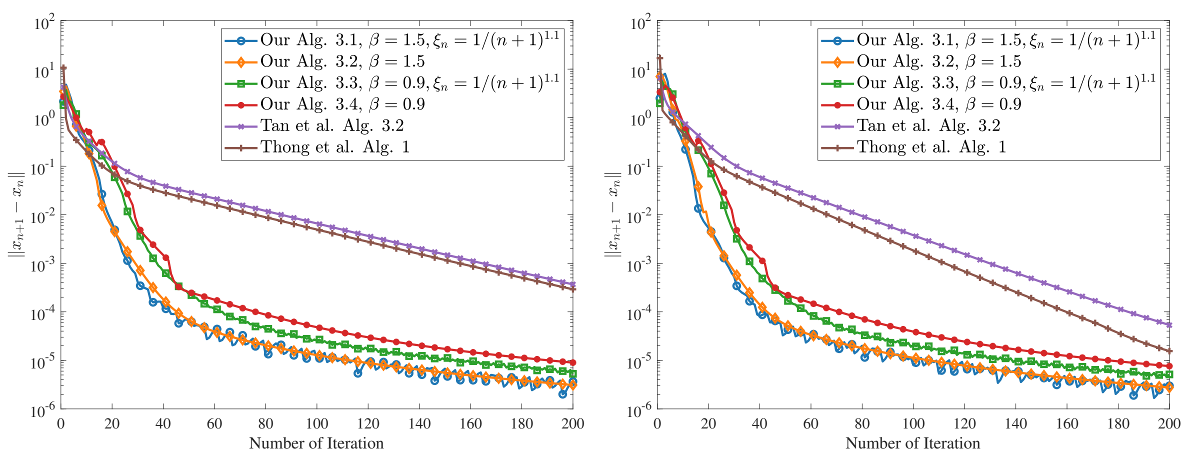
- In the Algorithm 3.2 suggested by Tan et al. [18], we take $\theta = 0.3$, $\varepsilon_n = 100/(n+1)^2$, $\mu = 0.1$, $\lambda_1 = 1$, $\alpha_n = 1/(n+1)$, and $\gamma = 1.7\zeta_F/L_F^2$.

We use $D_n = \|x_n - x_{n-1}\|$ to measure the error of the nth iteration since we do not know the exact solution to the problem (BVIP) with M and F given above. The maximum number of iterations 200 is used as a common stopping criterion for all algorithms. The numerical behavior of the proposed algorithms with different parameters β and ξ_n is shown in Fig. 1. Numerical results of all algorithms with two initial values are reported in Fig. 2.



(a) Comparison of the proposed Algorithms 3.1 and 3.3 (b) Comparison of the proposed Algorithms 3.2 and 3.4

FIGURE 1. Numerical behaviors of our algorithms with different β and ξ_n for Example 4.1



(a) $x_0 = x_1 =$ *10rand(5,1)* (b) $x_0 = x_1 =$ *20rand(5,1)*

FIGURE 2. Numerical results of all algorithms for Example 4.1

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

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

and norm

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

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 follows

$$C = \{x \in \mathscr{H} : \|x\| \le r\}.$$

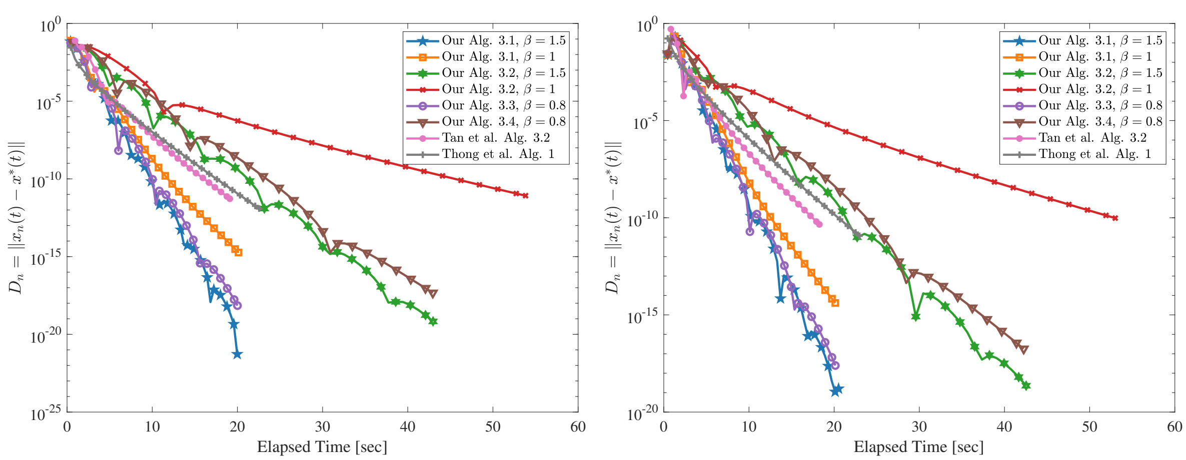
The operator $M : \mathscr{H} \to \mathscr{H}$ is given by

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

Note that the operator M is pseudomonotone rather than monotone (see [18, Example 2]). Let $F : \mathscr{H} \to \mathscr{H}$ be an operator defined by

$$(Fx)(t) = 0.5x(t), \quad t \in [0,1].$$

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 (BVIP) with M and F given above is $x^*(t) = 0$. The parameters of all algorithms remain the same as in Example 4.1, but we only adjust $\mu = 0.2$ for all the algorithms. The maximum number of iterations 50 is used as a common stopping criterion for all algorithms. Figure 3 shows the behaviors of $D_n = \|x_n(t) - x^*(t)\|$ generated by all algorithms under two different initial values.



(a) $x_0 = x_1 = \cos(2t)$ (b) $x_0 = x_1 = 5^t$

FIGURE 3. Numerical results of all algorithms for Example 4.2

Remark 4.1. We have the following observations for Examples 4.1 and 4.2.

(1) From Fig. 1 and Fig. 3, it can be seen that our algorithms can achieve a faster convergence speed if the appropriate β is chosen. In addition, the proposed Algorithm 3.1 and Algorithm 3.3 apply a non-monotonic step size (i.e, $\xi_n \neq 0$ in (3.2)), which allows them to converge faster than when non-increasing step sizes (i.e., $\xi_n = 0$ in (3.2)) are applied.
(2) As shown in Figs. 2 and 3, the proposed algorithms have a higher accuracy than the previously known ones in [17, 18] under the same stopping conditions. These results are independent of the choice of initial values.
(3) It is worth noting that our Algorithms 3.2 and 3.4 use an Armijo-type step size criterion, which makes them take more execution time than the proposed Algorithms 3.1 and 3.3 and the algorithms in the literature [17, 18], due to the fact that these algorithms use a sequence of steps that can be automatically updated in each iteration using previously known information.
(4) Note that the operator M in Example 4.2 is pseudomonotone rather than monotone. The algorithms proposed in [13, 15] for solving the bilevel monotone variational inequality problem will not be applicable in this case.

In summary, the methods proposed in this paper are useful, efficient, and robust.

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

$$\begin{cases} p^*(t) \in \operatorname{Argmin}\{g(p) : p \in V\}, \\ g(p) = \Phi(x(T)), \\ V = \left\{p(t) \in L_2([0,T],\mathbb{R}^m) : p_i(t) \in \left[p_i^-, p_i^+\right], i = 1,2,\ldots,m\right\}, \\ \text{s.t. } \dot{x}(t) = Q(t)x(t) + W(t)p(t),\ \ 0 \le t \le T,\ \ x(0) = x_0, \end{cases} \tag{4.1}$$

where $g(p)$ means the terminal objective function, Φ is a convex and differentiable defined on the attainability set, $p(t)$ denotes the control function, V represents a set of feasible controls composed of m piecewise continuous functions, $x(t)$ stands for the trajectory, and $Q(t) \in \mathbb{R}^{n\times n}$ and $W(t) \in \mathbb{R}^{n\times m}$ are given continuous matrices for every $t \in [0,T]$. By the solution of the problem (4.1), we mean a control $p^*(t)$ and a corresponding (optimal) trajectory $x^*(t)$ such that its terminal value $x^*(T)$ minimizes the objective function $g(p)$. It is known that the optimal control problem (4.1) can be transformed into a variational inequality problem (see [1, 28]). We first use the classical Euler discretization method to decompose the optimal control problem (4.1) and then apply the proposed algorithms to solve the variational inequality problem corresponding to the discretized version of the problem (see [1, 28] for more details).

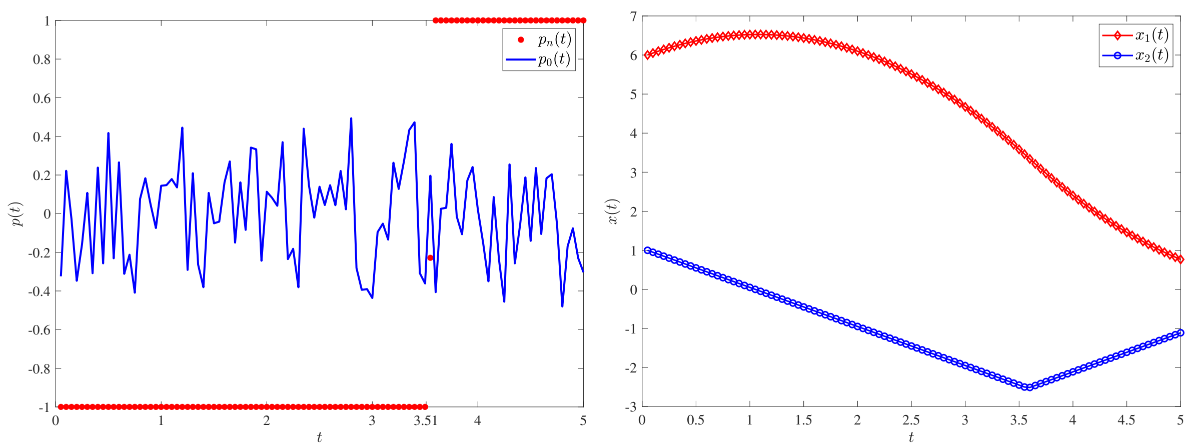
Example 4.3 (Rocket car [28]).

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

The exact optimal control of Example 4.3 is $p^*(t) = 1$ if $t \in (3.517, 5]$ and $p^*(t) = -1$ if $t \in (0, 3.517]$. We compare the proposed methods with the Algorithm (31) (in short, TLDCR Alg. (31)) introduced by Thong et al. [17] and the Algorithm (3.39) (in short, TLQ Alg. (3.39)) proposed by Tan, Liu, and Qin [18]. The parameters of all algorithms are set as follows.

- In the proposed Algorithms 3.1–3.4, we set $N = 100$, $\theta = 0.01$, $\varepsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\alpha_n = \frac{10^{-4}}{n+1}$, $F(x) = x - f(x)$, $f(x) = 0.1x$, and $\gamma = 1$. Choose $\mu = 0.1$, $\lambda_1 = 0.4$, and $\xi_n = 1/(n+1)^{1.1}$ for Algorithms 3.1 and 3.3. Pick $\delta = 2$, $\ell = 0.5$, and $\mu = 0.1$ for Algorithms 3.2 and 3.4.
- In the TLDCR Alg. (31), we choose $N = 100$, $\mu = 0.1$, $\lambda_1 = 0.4$, $\phi = 1.5$, and $\alpha_n = \frac{10^{-4}}{n+1}$.
- In the TLQ Alg. (3.39), we take $N = 100$, $\theta = 0.01$, $\varepsilon_n = \frac{10^{-4}}{(n+1)^2}$, $\mu = 0.1$, $\lambda_1 = 0.4$, $\alpha_n = \frac{10^{-4}}{n+1}$, and $f(x) = 0.1x$.

The initial controls $p_0(t) = p_1(t)$ are randomly generated in $[-0.5, 0.5]$. The stopping criterion is either $D_n = \|u_n - y_n\|_1 \leq 10^{-4}$, or the maximum number of iterations which is set to 300. Figure 4 shows the approximate optimal control and the corresponding trajectories of the suggested Algorithm 3.1 (with $\beta = 1.5$).



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

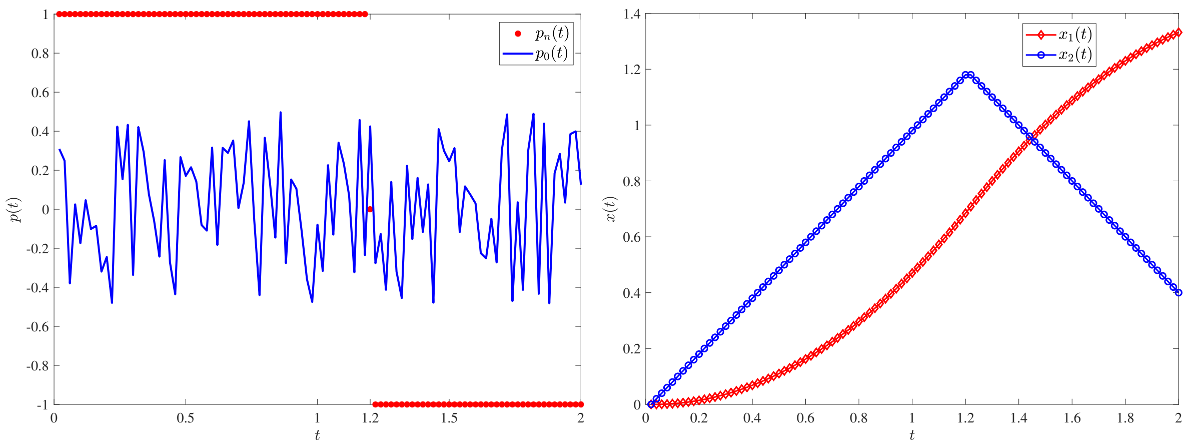
FIGURE 4. Numerical results of the proposed Algorithm 3.1 for Example 4.3

Example 4.4 (see [29]).

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

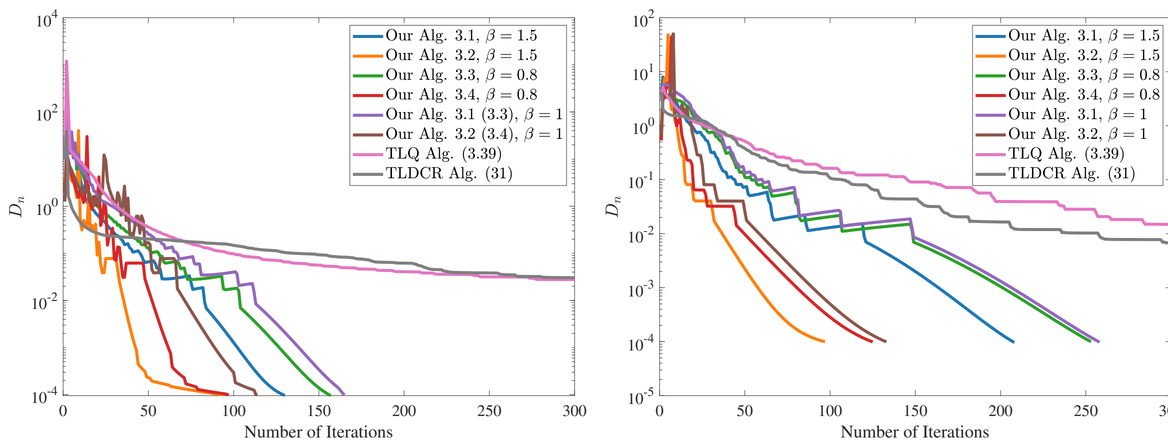
The exact optimal control of Example 4.4 is $p^*(t) = 1$ if $t \in [0, 1.2)$ and $p^*(t) = -1$ if $t \in (1.2, 2]$. The approximate optimal control and the corresponding trajectories of the suggested Algorithm 3.2 (with $\beta = 1.5$) are plotted in Fig. 5.

The numerical results of all algorithms for Examples 4.3 and 4.4 are shown in Fig. 6 and Table 1.



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

FIGURE 5. Numerical results of the proposed Algorithm 3.2 for Example 4.4



(a) Example 4.3 (b) Example 4.4

FIGURE 6. Error estimates of all algorithms for Examples 4.3 and 4.4

Remark 4.2. From Fig. 4, Fig. 5, Fig. 6, and Table 1, it can be seen that the suggested algorithms can be applied to solve optimal control problems. Notice that the proposed Algorithm 3.1 (resp., Algorithm 3.2) with $\beta = 1.5$ and the proposed Algorithm 3.3 (resp., Algorithm 3.4) with $\beta = 0.8$ converges faster than the proposed Algorithm 3.1 (resp., Algorithm 3.2) with $\beta = 1$. That is, the proposed algorithms can obtain a faster convergence speed if the appropriate β is chosen. Moreover, the proposed schemes outperform the existing methods in the literature [17, 18].

5. CONCLUSIONS

In this paper, we proposed four new modified subgradient extragradient methods for solving bilevel pseudomonotone variational inequality problems in real Hilbert spaces. Our algorithms work adaptively and do not require the prior knowledge of the Lipschitz constant of the

TABLE 1. Numerical results of all algorithms for Examples 4.3 and 4.4

Algorithms	Example 4.3			Example 4.4		
	Iter.	Time (s)	D_n	Iter.	Time (s)	D_n
Our Alg. 3.1, $\beta = 1.5$	129	0.0593	9.67E-05	207	0.0710	9.65E-05
Our Alg. 3.2, $\beta = 1.5$	93	0.0850	9.95E-05	96	0.0389	9.96E-05
Our Alg. 3.3, $\beta = 0.8$	156	0.0646	9.56E-05	252	0.1089	9.99E-05
Our Alg. 3.4, $\beta = 0.8$	96	0.0872	9.73E-05	124	0.0812	9.87E-05
Our Alg. 3.1 (3.3), $\beta = 1$	164	0.0555	9.84E-05	257	0.0894	9.79E-05
Our Alg. 3.2 (3.4), $\beta = 1$	113	0.0835	9.93E-05	132	0.0563	9.92E-05
TLQ Alg. (3.39)	300	0.1181	2.77E-02	300	0.1291	1.48E-02
TLDCR Alg. (31)	300	0.1023	3.04E-02	300	0.1075	6.72E-03

pseudomonotone mapping. The strong convergence theorems of the proposed algorithms are established under some mild conditions imposed by the mappings and parameters. Finally, the theoretical results are verified by some numerical experiments and applications. It would be interesting if the algorithms proposed in this paper could be applied to some practical bilevel optimization problems.

Acknowledgments
Bing Tan thanks the China Scholarship Council for the financial support (CSC No. 202106070094).

REFERENCES

[1] 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.
[2] P. Cubiotti, J.C. Yao, On the Cauchy problem for a class of differential inclusions with applications, Appl. Anal. 99 (2020), 2543-2554.
[3] Q.H. Ansari, M. Islam, J.C. Yao, Nonsmooth variational inequalities on Hadamard manifolds, Appl. Anal. 99 (2020), 340-358.
[4] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), 431-446.
[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] Y. Censor, A. Gibali, S. Reich, Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space, Optimization 61 (2012), 1119-1132.
[7] Y. Censor, A. Gibali, S. Reich, Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space, Optim. Methods Softw. 26 (2011), 827-845.
[8] Y. Malitsky, Projected reflected gradient methods for monotone variational inequalities, SIAM J. Optim. 25 (2015), 502-520.
[9] Q.L. Dong, D. Jiang, A. Gibali, A modified subgradient extragradient method for solving the variational inequality problem, Numer. Algorithms 79 (2018), 927-940.
[10] Y. Shehu, Q.L. Dong, L.L. Liu, Fast alternated inertial projection algorithms for pseudo-monotone variational inequalities, J. Comput. Appl. Math. 415 (2022), 114517.

[11] Y. Shehu, O.S. Iyiola, Projection methods with alternating inertial steps for variational inequalities: Weak and linear convergence. Appl. Numer. Math. 157 (2020), 315-337.

[12] L.C. Ceng, M. Shang, Hybrid inertial subgradient extragradient methods for variational inequalities and fixed point problems involving asymptotically nonexpansive mappings. Optimization 70 (2021), 715-740.

[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] P.N. Anh, J.K. Kim, L.D. Muu, An extragradient algorithm for solving bilevel pseudomonotone variational inequalities, J. Global Optim. 52 (2012), 627-639.

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

[16] D.V. Thong, N.A. Triet, X.H. Li, Q.L. Dong, Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems, Numer. Algorithms 83 (2020), 1123-1143.

[17] 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.

[18] 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.

[19] H. Liu, J. Yang, Weak convergence of iterative methods for solving quasimonotone variational inequalities, Comput. Optim. Appl. 77 (2020), 491-508.

[20] 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.

[21] Y. Shehu, A. Gibali, New inertial relaxed method for solving split feasibilities, Optim. Lett. 15 (2021), 2109-2126.

[22] Y. Shehu, O.S. Iyiola, S. Reich, A modified inertial subgradient extragradient method for solving variational inequalities, Optim. Eng. 23 (2022), 421-449.

[23] 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, Elsevier, Amsterdam, 2001.

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

[25] L.D. Muu, N.V. Quy, On existence and solution methods for strongly pseudomonotone equilibrium problems, Vietnam J. Math. 43 (2015), 229-238.

[26] B. Tan, X. Qin, J.C. Yao, Strong convergence of inertial projection and contraction methods for pseudomonotone variational inequalities with applications to optimal control problems, J. Global Optim. 82 (2022), 523-557.

[27] B. Tan, S.Y. Cho, Inertial extragradient methods for solving pseudomonotone variational inequalities with non-Lipschitz mappings and their optimization applications, Appl. Set-Valued Anal. Optim. 3 (2021), 165-192.

[28] 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.

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