Japan Journal of Industrial and Applied Mathematics
https://doi.org/10.1007/s13160-020-00450-y

ORIGINAL PAPER

Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems

Bing Tan[1] · **Liya Liu**[2] · **Xiaolong Qin**[3]

Received: 14 June 2020 / Revised: 1 October 2020 / Accepted: 15 November 2020

Abstract
We introduce two inertial extragradient algorithms for solving a bilevel pseudomonotone variational inequality problem in real Hilbert spaces. The advantages of the proposed algorithms are that they can work without the prior knowledge of the Lipschitz constant of the involving operator and only one projection onto the feasible set is required. Strong convergence theorems of the suggested algorithms are obtained under suitable conditions. Finally, some numerical examples are provided to show the efficiency of the proposed algorithms.

Keywords Inertial subgradient extragradient method · Inertial Tseng's extragradient method · Steepest descent method · Bilevel variational inequality problem · Pseudomonotone mapping

Mathematics Subject Classification 47H05 · 47H09 · 65K15

✉ Xiaolong Qin
qxlxajh@163.com

Bing Tan
bingtan72@gmail.com

Liya Liu
liya42@qq.com

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

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

3 Department of Mathematics, Zhejiang Normal University, Jinhua, Zhejiang, China

Published online: 04 January 2021

1 Introduction

Throughout this paper, C is assumed to be a convex and closed nonempty set in a Hilbert space H with the inner product $\langle\cdot,\cdot\rangle$ and the norm $\|\cdot\|$. One gives two single-valued mappings $A : H \to H$ and $F : H \to H$ on H. The classical variational inequality problem (VIP) is described as follows.

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

One denotes by $\mathrm{VI}(C,A)$ the set of all solutions of (VIP). In this paper, one investigates two numerical methods to find solutions of the following bilevel variational inequality problem (BVIP), which reads as follows.

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

Bilevel variational inequality problems, which have been extensively investigated by numerical methods, cover a number of nonlinear optimization problems, such as, fixed point problems, quasi-variational inequality problems, complementary problems, saddle problems and minimum norm problems, see, e.g., [1–5]. It is known that (VIP) is equivalent to the fixed point problem of finding a point x^* in C such that $x^* = P_C(x^* - \lambda Ax^*)$, where λ is any positive real number and P_C represents the metric projection from H onto C (see the definition in Sect. 2). Recently, a number of authors proposed and analyzed various methods to solve the (VIP). Two notable methods to solve (VIP) are the regularization method and the projection method. In this paper, we focus on the second approach involving projection methods. The simplest and oldest projection method is the gradient projection method:

$$x_{n+1} = P_C\big(x_n - \lambda Ax_n\big). \tag{1.1}$$

It is known that the iterative sequence defined by (1.1) converges to an element of $\mathrm{VI}(C,A)$ when A is L-Lipschitz continuous and α-strongly monotone and $\lambda \in \big(0, 2\alpha/L^2\big)$. For avoiding the use of such assumptions, the extragradient method (EGM) [6] has been proposed for a monotone and L-Lipschitz continuous mapping A. The algorithm can be presented as follows:

$$\begin{cases} y_n = P_C\big(x_n - \lambda Ax_n\big), \\ x_{n+1} = P_C\big(x_n - \lambda Ay_n\big), \end{cases} \tag{1.2}$$

where $\lambda \in (0, 1/L)$. The algorithm defined by (1.2) converges to an element of $\mathrm{VI}(C,A)$ provided that $\mathrm{VI}(C,A)$ is nonempty.

We see that the EGM needs to compute two projections onto the feasible set C and two evaluations of operator A in each iteration. Generally, this is expensive, and when operator A and feasible set C have a complicated structure, it will affect the efficiency of the method used. To overcome one of these shortcomings, there are two notable methods in the literature. The first one is the subgradient extragradient method (SEGM) [7], which can be considered as an improvement of the EGM. The algorithm reads as follows:

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ T_n = \left\{x \in H \mid \langle x_n - \lambda A x_n - y_n, x - y_n\rangle \le 0\right\}, \\ x_{n+1} = P_{T_n}\left(x_n - \lambda A y_n\right), \end{cases} \tag{1.3}$$

where $\lambda \in (0, 1/L)$. The main advantage of SEGM is that it replaces the second projection from the closed and convex subset C to the half-space T_n, and this projection can be calculated by an explicit formula.

The second one is called the Tseng's extragradient method [8], which is described as follows:

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ x_{n+1} = y_n - \lambda\left(A y_n - A x_n\right). \end{cases} \tag{1.4}$$

It is worth noting that the algorithm defined in (1.4) only needs to calculate one projection onto the feasible set C and two evaluations of A in each step. Since subgradient extragradient method and Tseng's extragradient method only need to calculate once projection onto the feasible set C in each step, they have received a lot of attention and research from scientific researchers, who have improved and extended in various ways to obtain the weak and strong convergence of these methods, see [9–11] and the references therein. However, the subgradient extragradient method and Tseng's extragradient method need to know the Lipschitz constant of the operator A, which limits the applicability of the algorithms. To handle the case where the Lipschitz constant of the operator A is unknown, Yang et al. [12, 13] introduced the following self-adaptive step size strategy for (1.3) and (1.4), respectively.

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\left(\|x_n - y_n\|^2 + \|x_{n+1} - y_n\|^2\right)}{2\left\langle A x_n - A y_n, x_{n+1} - y_n\right\rangle}, \lambda_n\right\}, & \text{if } \left\langle A x_n - A y_n, x_{n+1} - y_n\right\rangle > 0; \\ \lambda_n, & \text{otherwise}, \end{cases}$$

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\|x_n - y_n\|}{\|A x_n - A y_n\|}, \lambda_n\right\}, & \text{if } A x_n - A y_n \ne 0; \\ \lambda_n, & \text{otherwise}, \end{cases}$$

where $\mu \in (0, 1)$ and $\lambda_0 > 0$.

Note that computation of the metric projection P_C onto C is not necessarily easy. In order to reduce the complexity probably caused by the projection P_C, Yamada [14] introduced the following hybrid steepest descent method for solving the variational inequality $\mathrm{VI}(C, F)$. Recall that T is nonexpansive if $\|Tx - Ty\| \le \|x - y\|,\ \forall x, y \in H$. Let $\mathrm{Fix}(T) = \{x \in H : Tx = x\}$ denote the fixed point set of T. Assume that C is the fixed point set of a nonexpansive mapping $T : H \to H$, that is, $C = \{x \in H : Tx = x\}$. Let F be a mapping of η-strongly monotone and κ-Lipschitzian on C. Fix a constant $\mu \in (0, 2\eta/\kappa^2)$ and a sequence $\{\lambda_n\}$ of real numbers in (0, 1) satisfying the following conditions: (i)

$\lim_{n\to\infty}\lambda_n = 0$, (ii) $\sum_{n=1}^{\infty}\lambda_n = \infty$, (iii) $\lim_{n\to\infty}(\lambda_n - \lambda_{n+1})/\lambda_{n+1}^2 = 0$. For any initial data $x_0 \in H$, one can generate a sequence $\{x_n\}$ via the following algorithm:

$$x_{n+1} = Tx_n - \lambda_{n+1}\mu F(Tx_n), \quad n \geq 0\,.$$

Yamada proved that $\{x_n\}$ converges to the unique solution of the VI(C, F) in norm. In recent years, there are many papers dealing with the variational inequality problems by using the steepest descent method, see [15–17]. Note that the methods suggested in (1.2), (1.3) and (1.4) all achieve weak convergence in infinite-dimensional spaces. Examples in CT reconstruction and machine learning tell us that strong convergence is preferable to weak convergence in an infinite-dimensional space. Therefore, a natural question is how to modify the methods (1.3) and (1.4) such that they can achieve strong convergence in infinite-dimensional spaces. Recently, based on the subgradient extragradient algorithm (1.3), the Tseng's extragradient method (1.4) and the hybrid steepest descent method [14], Thong et al. [18] proposed two new modified extragradient algorithms with strong convergence to solve the (BVIP) in a real Hilbert space.

To accelerate the convergence rate of the algorithms, Polyak [19] considered the second-order dynamical system $\ddot{x}(t) + \gamma\dot{x}(t) + \nabla f(x(t)) = 0$, where $\gamma > 0$, ∇f represents the gradient of f, $\dot{x}(t)$ and $\ddot{x}(t)$ denote the first and second derivatives of x at t, respectively. This dynamic system is called the Heavy Ball with Friction (HBF). Next, we consider the discretization of this dynamic system (HBF), that is,

$$\frac{x_{n+1} - 2x_n + x_{n-1}}{h^2} + \gamma\frac{x_n - x_{n-1}}{h} + \nabla f\left(x_n\right) = 0, \quad \forall n \geq 0\,.$$

Through a direct calculation, we can get the following form:

$$x_{n+1} = x_n + \tau\left(x_n - x_{n-1}\right) - \varphi\nabla f\left(x_n\right), \quad \forall n \geq 0\,,$$

where $\tau = 1 - \gamma h$ and $\varphi = h^2$. This can be considered as the following two-step iteration scheme:

$$\begin{cases} y_n = x_n + \tau\left(x_n - x_{n-1}\right), \\ x_{n+1} = y_n - \varphi\nabla f\left(x_n\right), \quad \forall n \geq 0\,. \end{cases}$$

This iteration is now called the inertial extrapolation algorithm, the term $\tau\left(x_n - x_{n-1}\right)$ is referred to as the extrapolation point. It is known that the Nesterov accelerated gradient method [20] improves the convergence rate of the gradient method from standard $O\left(k^{-1}\right)$ down to $O\left(k^{-2}\right)$. However, it should be highlighted that inertial algorithms do not guarantee that the objective function is monotone. Recently, many authors constructed a large number of inertial algorithms for solving variational inequalities and optimization problems; see, e.g., [21–24] and the references therein.

Motivated and inspired by the above works, we here introduce two new inertial extragradient methods for solving (BVIP) in real Hilbert spaces. The algorithms are inspired by the inertial method, the subgradient extragradient method, the Tseng's extragradient method and the steepest descent method. We provide a choice of inertial parameter and two new stepsize rules which allow the algorithms to work without previously knowing the Lipschitz constant of the mapping. Under some suitable conditions, we prove that the iterative sequence generated by the algorithms converges strongly to a solution of (BVIP). Some numerical experiments are carried out to support the theoretical results. Our numerical results show that the new algorithms have a better convergence speed than the existing ones presented in [18].

An outline of this paper is as follows. In Sect. 2, we recall some preliminary results and lemmas for further use. Section 3 analyzes the convergence of the proposed algorithms. In Sect. 4, some numerical examples are presented to illustrate the numerical behavior of the proposed algorithms and compare them with other ones. Finally, a brief summary is given in Sect. 5.

2 Preliminaries

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

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

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

- $\langle x - P_C x, y - P_C x\rangle \leq 0, \forall y \in C$;
- $\|P_C x - P_C y\|^2 \leq \langle P_C x - P_C y, x - y\rangle, \forall y \in H$.

We give some special cases with simple analytical solutions. These give us some explicit formulas to find the projection of any point onto the half-space and the ball.

- The Euclidean projection of x_0 onto a halfspace $H_{a,b}^{-} = \{x : \langle a, x\rangle \leq b\}$ is given by

$$P_{H_{a,b}^{-}}(x) = x - \max\left\{\frac{[\langle a, x\rangle - b]}{\|a\|^2}, 0\right\} a .$$

- The Euclidean projection of x_0 onto an Euclidean ball $B[c, r] = \{x : \|x - c\| \leq r\}$ is given by

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

It is well known that if $F : H \to H$ is L-Lipschitz continuous and β-strongly monotone on H and if $\mathrm{VI}(C, A)$ is a nonempty, convex and closed subset of H, then the (BVIP) has a unique solution (see, e.g., [25]).

The following lemmas play an important role in our proofs.

Lemma 2.1 ([26]) *Let $A : C \to H$ be a continuous and pseudomonotone operator. Then, x^* is a solution of $\mathrm{VI}(C, A)$ if and only if $\langle Ax, x - x^*\rangle \geq 0$, $\forall x \in C$.*

Lemma 2.2 ([14]) *Let $\gamma > 0$ and $\alpha \in (0, 1]$. Let $F : H \to H$ be a β-strongly monotone and L-Lipschitz continuous mapping with $0 < \beta \leq L$. Associating with a nonexpansive mapping $T : H \to H$, define a mapping $T^\gamma : H \to H$ by $T^\gamma x = (I - \alpha\gamma F)(Tx), \forall x \in H$. Then, T^γ is a contraction provided $\gamma < \frac{2\beta}{L^2}$, that is,*

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

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

Lemma 2.3 ([27]) *Let $\left\{a_n\right\}$ be a sequence of nonnegative real numbers, $\left\{\alpha_n\right\}$ be a sequence of real numbers in $(0, 1)$ with $\sum_{n=1}^{\infty} \alpha_n = \infty$, and $\left\{b_n\right\}$ be a sequence of real numbers. Assume that $a_{n+1} \leq \left(1 - \alpha_n\right)a_n + \alpha_n b_n, \forall n \geq 1$. If $\limsup_{k\to\infty} b_{n_k} \leq 0$ for every subsequence $\left\{a_{n_k}\right\}$ of $\left\{a_n\right\}$ satisfying $\liminf_{k\to\infty}\left(a_{n_k+1} - a_{n_k}\right) \geq 0$, then $\lim_{n\to\infty} a_n = 0$.*

3 Main results

In this section, we introduce two new inertial extragradient methods for solving (BVIP) and analyze their convergence. First, we assume that our proposed methods satisfy the following conditions.

- C1 The feasible set C is a nonempty, convex and closed set.
- C2 The solution set of the (VIP) is nonempty, that is $\mathrm{VI}(C, A) \neq \emptyset$.
- C3 The mapping $A : H \to H$ is L_1-Lipschitz continuous and pseudomonotone on H, and sequentially weakly continuous on C.
- C4 The mapping $F : H \to H$ is L_2-Lipschitz continuous and β-strongly monotone on H such that $L_2 \geq \beta$. In addition, we denote by p the unique solution of the (BVIP).
- 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 the following conditions: $\sum_{n=1}^{\infty} \alpha_n = \infty$ and $\lim_{n\to\infty} \alpha_n = 0$.

Remark 3.1 Our methods are embedded with the inertial terms to ensure the strong convergence of the algorithms. Condition (C5), which is routine restriction, is easily satisfied. For example, one can take $\alpha_n = 1/n$ and $\epsilon_n = 1/n^2$.

3.1 The modified inertial subgradient extragradient algorithm

Now, we introduce the new modified inertial subgradient extragradient algorithm for solving (BVIP). Algorithm 3.1 reads as follows:

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

Initialization: Set $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0,1)$, $0 < \gamma < \frac{2\beta}{L_2^2}$. Arbitrarily given $x_0, x_1 \in H$.

Iterative Steps: Calculate x_{n+1} as follows:

Step 1. Given the iterates x_{n-1} and x_n $(n \geq 1)$. Set

$$w_n = x_n + \theta_n (x_n - x_{n-1}),$$

where

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

Step 2. Compute

$$y_n = P_C(w_n - \lambda_n A w_n).$$

Step 3. Compute

$$z_n = P_{T_n}(w_n - \lambda_n A y_n),$$

where $T_n := \{x \in H \mid \langle w_n - \lambda_n A w_n - y_n, x - y_n\rangle \leq 0\}$.

Step 4. Compute

$$x_{n+1} = z_n - \alpha_n \gamma F z_n,$$

and update

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

Remark 3.2 It follows from (3.1) that

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

Indeed, we have $\theta_n\|x_n - x_{n-1}\| \leq \epsilon_n$ for all n, which together with $\lim_{n\to\infty} \frac{\epsilon_n}{\alpha_n} = 0$ implies that

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

Lemma 3.1 *Assume that Conditions (C1)–(C3) hold. Then, the sequence $\{\lambda_n\}$ generated by (3.2) is a nonincreasing sequence and*

$$\lim_{n\to\infty} \lambda_n = \lambda \geq \min\left\{\lambda_1, \frac{\mu}{L_1}\right\}.$$

Proof We can easily get that $\lambda_{n+1} \leq \lambda_n$ for all $n \in \mathbb{N}$ from (3.2). Hence, $\{\lambda_n\}$ is nonincreasing. On the other hand, we have $L_1\|w_n - y_n\| \geq \|Aw_n - Ay_n\|$, since A is L_1-Lipschitz continuous. Therefore, if $\langle Aw_n - Ay_n, z_n - y_n\rangle > 0$, then

$$\mu\frac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{2\langle Aw_n - Ay_n, z_n - y_n\rangle} \geq \mu\frac{\|w_n - y_n\|\|z_n - y_n\|}{\|Aw_n - Ay_n\|\|z_n - y_n\|} = \mu\frac{\|w_n - y_n\|}{\|Aw_n - Ay_n\|} \geq \frac{\mu}{L_1},$$

which together with (3.2) implies that

$$\lambda_n \geq \min\left\{\lambda_1, \frac{\mu}{L_1}\right\}.$$

Thus, we conclude that $\lim_{n\to\infty} \lambda_n$ exists since the sequence $\{\lambda_n\}$ is nonincreasing and lower bounded. □

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

Lemma 3.2 *Assume that Conditions (C1)–(C3) hold. Let $\{z_n\}$ be a sequence generated by Algorithm 3.1. Then, for all $p \in \mathrm{VI}(C, A)$,*

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

Proof First, from the definition of $\{\lambda_n\}$, one obtains

$$2\langle Aw_n - Ay_n, z_n - y_n\rangle \leq \frac{\mu}{\lambda_{n+1}}\|w_n - y_n\|^2 + \frac{\mu}{\lambda_{n+1}}\|z_n - y_n\|^2, \ \forall n \in \mathbb{N}. \tag{3.3}$$

Indeed, if $\langle Aw_n - Ay_n, z_n - y_n\rangle \leq 0$, then (3.3) holds. Otherwise, by (3.2) we have

$$\lambda_{n+1} = \min\left\{\mu\frac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{2\langle Aw_n - Ay_n, z_n - y_n\rangle}, \lambda_n\right\} \leq \mu\frac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{2\langle Aw_n - Ay_n, z_n - y_n\rangle},$$

which implies that

$$2\langle Aw_n - Ay_n, z_n - y_n\rangle \leq \frac{\mu}{\lambda_{n+1}}\|w_n - y_n\|^2 + \frac{\mu}{\lambda_{n+1}}\|z_n - y_n\|^2.$$

Thus, the inequality (3.3) holds. Using (3.3) and $p \in \mathrm{VI}(C, A) \subset C \subset T_n$, one has

$$\begin{aligned}2\|z_n-p\|^2 =&2\left\|P_{T_n}\left(w_n-\lambda_nAy_n\right)-P_{T_n}p\right\|^2 \le 2\langle z_n-p, w_n-\lambda_nAy_n-p\rangle\\ =&\|z_n-p\|^2+\|w_n-\lambda_nAy_n-p\|^2-\|z_n-w_n+\lambda_nAy_n\|^2\\ =&\|z_n-p\|^2+\|w_n-p\|^2+\lambda_n^2\|Ay_n\|^2-2\langle w_n-p,\lambda_nAy_n\rangle\\ &-\|z_n-w_n\|^2-\lambda_n^2\|Ay_n\|^2-2\langle z_n-w_n,\lambda_nAy_n\rangle\\ =&\|z_n-p\|^2+\|w_n-p\|^2-\|z_n-w_n\|^2-2\langle z_n-p,\lambda_nAy_n\rangle\,.\end{aligned}$$

This implies that

$$\|z_n-p\|^2 \le \|w_n-p\|^2-\|z_n-w_n\|^2-2\langle z_n-p,\lambda_nAy_n\rangle\,. \tag{3.4}$$

Since p is the solution of (VIP), we have $\langle Ap, x-p\rangle \ge 0$ for all $x \in C$. By the pseudomontonicity of A on H, we get $\langle Ax, x-p\rangle \ge 0$ for all $x \in C$. Taking $x=y_n \in C$, one infers that

$$\langle Ay_n, p-y_n\rangle \le 0\,.$$

Consequently,

$$\langle Ay_n, p-z_n\rangle = \langle Ay_n, p-y_n\rangle + \langle Ay_n, y_n-z_n\rangle \le \langle Ay_n, y_n-z_n\rangle\,. \tag{3.5}$$

Combining (3.4) and (3.5), one obtains

$$\begin{aligned}\|z_n-p\|^2 \le&\|w_n-p\|^2-\|z_n-w_n\|^2+2\lambda_n\langle Ay_n, y_n-z_n\rangle\\ =&\|w_n-p\|^2-\|z_n-y_n\|^2-\|y_n-w_n\|^2-2\langle z_n-y_n, y_n-w_n\rangle\\ &+2\lambda_n\langle Ay_n, y_n-z_n\rangle\\ =&\|w_n-p\|^2-\|z_n-y_n\|^2-\|y_n-w_n\|^2+2\langle z_n-y_n, w_n-\lambda_nAy_n-y_n\rangle\,.\end{aligned} \tag{3.6}$$

Since $z_n=P_{T_n}\left(w_n-\lambda_nAy_n\right)$ and $z_n \in T_n$, one has

$$\begin{aligned}&\langle w_n-\lambda_nAy_n-y_n, z_n-y_n\rangle\\ &= \langle w_n-\lambda_nAw_n-y_n, z_n-y_n\rangle + \lambda_n\langle Aw_n-Ay_n, z_n-y_n\rangle\\ &\le \lambda_n\langle Aw_n-Ay_n, z_n-y_n\rangle\,,\end{aligned} \tag{3.7}$$

which together with (3.3), we deduce that

$$2\langle w_n-\lambda_nAy_n-y_n, z_n-y_n\rangle \le \mu\frac{\lambda_n}{\lambda_{n+1}}\|w_n-y_n\|^2+\mu\frac{\lambda_n}{\lambda_{n+1}}\|z_n-y_n\|^2\,.$$

From (3.6) and (3.7), we obtain

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

This completes the proof of the lemma. □

Lemma 3.3 [28, *Lemma* 3.3] *Assume that Conditions (C1)–(C3) hold. Let* $\{w_n\}$ *be a sequence generated by Algorithm 3.1. If there exists a subsequence* $\{w_{n_k}\}$ *converges weakly to* $z \in H$ *and* $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$, *then* $z \in \mathrm{VI}(C, A)$.

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

Proof We divide the proof into four steps.

Setp 1. The sequence $\{x_n\}$ is bounded. Indeed, from $\lim_{n\to\infty}\left(1-\mu\frac{\lambda_n}{\lambda_{n+1}}\right) = 1-\mu > 0$, we know that there exists $n_0 \in \mathbb{N}$ such that

$$1-\mu\frac{\lambda_n}{\lambda_{n+1}} > 0, \quad \forall n \ge n_0 . \tag{3.8}$$

Combining Lemma 3.2 and (3.8), one sees that

$$\|z_n - p\| \le \|w_n - p\|, \quad \forall n \ge n_0 . \tag{3.9}$$

On the other hand, by the definition of w_n, we can write

$$\|w_n - p\| \le \alpha_n \cdot \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| + \|x_n - p\|. \tag{3.10}$$

According to Remark 3.2 we have $\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \to 0$. Therefore, there exists a constant $M_1 > 0$ such that

$$\frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| \le M_1, \quad \forall n \ge 1 . \tag{3.11}$$

Combining (3.9), (3.10) and (3.11), we obtain

$$\|z_n - p\| \le \|w_n - p\| \le \|x_n - p\| + \alpha_n M_1, \quad \forall n \ge n_0 . \tag{3.12}$$

Using Lemma 2.2 and (3.9), it follows that

$$\begin{aligned}
\|x_{n+1} - p\| &= \|\left(I - \alpha_n\gamma F\right)z_n - \left(I - \alpha_n\gamma F\right)p - \alpha_n\gamma Fp\| \\
&\le \left(1 - \alpha_n\eta\right)\|z_n - p\| + \alpha_n\gamma\|Fp\| \\
&\le \left(1 - \alpha_n\eta\right)\|x_n - p\| + \alpha_n\eta \cdot \frac{M_1}{\eta} + \alpha_n\eta \cdot \frac{\gamma}{\eta}\|Fp\| \\
&\le \max\left\{\frac{M_1 + \gamma\|Fp\|}{\eta}, \|x_n - p\|\right\} \\
&\le \cdots \le \max\left\{\frac{M_1 + \gamma\|Fp\|}{\eta}, \|x_{n_0} - p\|\right\}, \ \forall n \ge n_0 ,
\end{aligned} \tag{3.13}$$

where $\eta = 1 - \sqrt{1 - \gamma\left(2\beta - \gamma L_2^2\right)} \in (0, 1)$. That is, this implies that $\{x_n\}$ is bounded. We get that the sequences $\{z_n\}$ and $\{w_n\}$ are also bounded.

Step 2.

$$\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|y_n - w_n\|^2 + \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|z_n - y_n\|^2 \\ \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n M_4,\ \forall n \geq n_0$$

for some $M_4 > 0$. Indeed, using (2.1), one has

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|\left(I - \alpha_n\gamma F\right)z_n - \left(I - \alpha_n\gamma F\right)p - \alpha_n\gamma Fp\|^2 \\ &\leq \left(1 - \alpha_n\eta\right)^2\|z_n - p\|^2 + 2\alpha_n\gamma\left\langle Fp, p - x_{n+1}\right\rangle \\ &\leq \|z_n - p\|^2 + \alpha_n M_2 \end{aligned} \tag{3.14}$$

for some $M_2 > 0$. In the light of Lemma 3.2, we obtain

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|y_n - w_n\|^2 - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|z_n - y_n\|^2 + \alpha_n M_2\,. \tag{3.15}$$

It follows from (3.12) that

$$\begin{aligned} \|w_n - p\|^2 &\leq \left(\|x_n - p\| + \alpha_n M_1\right)^2 \\ &= \|x_n - p\|^2 + \alpha_n\left(2M_1\|x_n - p\| + \alpha_n M_1^2\right) \\ &\leq \|x_n - p\|^2 + \alpha_n M_3 \end{aligned} \tag{3.16}$$

for some $M_3 > 0$. Combining (3.15) and (3.16), we get

$$\|x_{n+1} - p\|^2 \leq \|x_n - p\|^2 + \alpha_n M_3 - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|w_n - y_n\|^2 \\ - \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|z_n - y_n\|^2 + \alpha_n M_2\,,$$

which yields

$$\left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|y_n - w_n\|^2 + \left(1 - \mu\frac{\lambda_n}{\lambda_{n+1}}\right)\|z_n - y_n\|^2 \\ \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n M_4\,,$$

where $M_4 := M_2 + M_3$.

Step 3.

$$\|x_{n+1}-p\|^2 \le \left(1-\alpha_n\eta\right)\|x_n-p\|^2 \\ +\alpha_n\eta\left[\frac{2\gamma}{\eta}\left\langle Fp, p-x_{n+1}\right\rangle+\frac{3M\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|\right],\ \forall n\ge n_0$$

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

$$\|w_n-p\|^2 \le \|x_n-p\|^2+2\theta_n\|x_n-p\|\|x_n-x_{n-1}\|+\theta_n^2\|x_n-x_{n-1}\|^2. \quad (3.17)$$

Combining (3.9) and (3.14), we get

$$\|x_{n+1}-p\|^2 \le \left(1-\alpha_n\eta\right)\|w_n-p\|^2+2\alpha_n\gamma\left\langle Fp, p-x_{n+1}\right\rangle, \quad \forall n\ge n_0. \quad (3.18)$$

Substituting (3.17) into (3.18), we obtain

$$\begin{aligned}\|x_{n+1}-p\|^2 &\le \left(1-\alpha_n\eta\right)\|x_n-p\|^2+2\alpha_n\gamma\left\langle Fp, p-x_{n+1}\right\rangle \\ &\quad+\theta_n\|x_n-x_{n-1}\|\left(2\|x_n-p\|+\theta\|x_n-x_{n-1}\|\right) \\ &\le \left(1-\alpha_n\eta\right)\|x_n-p\|^2 \\ &\quad+\alpha_n\eta\left[\frac{2\gamma}{\eta}\left\langle Fp, p-x_{n+1}\right\rangle+\frac{3M\theta_n}{\alpha_n\eta}\|x_n-x_{n-1}\|\right], \forall n\ge n_0,\end{aligned}$$

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

Step 4. $\left\{\|x_n-p\|^2\right\}$ converges to zero. Indeed, by Lemma 2.3, it suffices to show that $\limsup_{k\to\infty}\left\langle Fp, p-x_{n_k+1}\right\rangle\le 0$ for every subsequence $\left\{\|x_{n_k}-p\|\right\}$ of $\left\{\|x_n-p\|\right\}$ satisfying

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

For this purpose, one assumes that $\left\{\|x_{n_k}-p\|\right\}$ is a subsequence of $\left\{\|x_n-p\|\right\}$ such that $\liminf_{k\to\infty}\left(\|x_{n_k+1}-p\|-\|x_{n_k}-p\|\right)\ge 0$. By Step 2, one has

$$\begin{aligned}&\limsup_{k\to\infty}\left[\left(1-\mu\frac{\lambda_{n_k}}{\lambda_{n_k+1}}\right)\|y_{n_k}-w_{n_k}\|^2+\left(1-\mu\frac{\lambda_{n_k}}{\lambda_{n_k+1}}\right)\|z_{n_k}-y_{n_k}\|^2\right] \\ &\le \limsup_{k\to\infty}\left[\alpha_{n_k}M_4+\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2\right] \\ &\le \limsup_{k\to\infty}\alpha_{n_k}M_4+\limsup_{k\to\infty}\left[\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2\right] \\ &= -\liminf_{k\to\infty}\left[\|x_{n_k+1}-p\|^2-\|x_{n_k}-p\|^2\right] \\ &\le 0,\end{aligned}$$

which implies that

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

Therefore

$$\lim_{k\to\infty} \|z_{n_k} - w_{n_k}\| = 0\,. \tag{3.19}$$

Moreover, we can show that

$$\|x_{n_k+1} - z_{n_k}\| = \alpha_{n_k}\gamma\|Fz_{n_k}\| \to 0 \text{ as } n \to \infty\,, \tag{3.20}$$

and

$$\|x_{n_k} - w_{n_k}\| = \theta_{n_k}\|x_{n_k} - x_{n_k-1}\| = \alpha_{n_k} \cdot \frac{\theta_{n_k}}{\alpha_{n_k}}\|x_{n_k} - x_{n_k-1}\| \to 0 \text{ as } n \to \infty\,. \tag{3.21}$$

Combining (3.19), (3.20) and (3.21), we obtain

$$\|x_{n_k+1} - x_{n_k}\| \le \|x_{n_k+1} - z_{n_k}\| + \|z_{n_k} - w_{n_k}\| + \|w_{n_k} - x_{n_k}\| \to 0 \text{ as } n \to \infty\,. \tag{3.22}$$

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 $z \in H$, such that

$$\limsup_{k\to\infty} \left\langle Fp, p - x_{n_k}\right\rangle = \lim_{j\to\infty} \left\langle Fp, p - x_{n_{k_j}}\right\rangle = \langle Fp, p - z\rangle\,. \tag{3.23}$$

By (3.21), we get $w_{n_k} \rightharpoonup z$ as $k \to \infty$. This together with $\lim_{k\to\infty} \|w_{n_k} - y_{n_k}\| = 0$ and Lemma 3.3 yields $z \in \mathrm{VI}(C, A)$. From (3.23) and the assumption that p is the unique solution of the (BVIP), we get

$$\limsup_{k\to\infty} \left\langle Fp, p - x_{n_k}\right\rangle = \langle Fp, p - z\rangle \le 0\,. \tag{3.24}$$

Combining (3.22) and (3.24), we obtain

$$\begin{aligned}\limsup_{k\to\infty} \left\langle Fp, p - x_{n_k+1}\right\rangle = \limsup_{k\to\infty} \left\langle Fp, p - x_{n_k}\right\rangle &= \langle Fp, p - z\rangle \\ &\le 0\,.\end{aligned} \tag{3.25}$$

From $\lim_{n\to\infty} \frac{\theta_n}{\alpha_n}\|x_n - x_{n-1}\| = 0$ and (3.25), we get

$$\limsup_{k\to\infty} \left[\frac{2\gamma}{\eta}\left\langle Fp, p - x_{n_k+1}\right\rangle + \frac{3M\theta_{n_k}}{\alpha_{n_k}\eta}\|x_{n_k} - x_{n_k-1}\|\right] \le 0\,. \tag{3.26}$$

Hence, combining Step 3, Condition (C5) and (3.26), in the light of Lemma 2.3, one concludes that $\lim_{n\to\infty} \|x_n - p\| = 0$. This completes the proof. □

Now applying Theorem 3.1 with $F(x) = x - x_0$, where $x_0 \in H$. It is easy to check that the mapping $F : H \to H$ is 1-Lipschitz continuous and 1-strongly monotone on H. In this case, by choosing $\gamma = 1$, the calculation of x_{n+1} in Algorithm 3.1 becomes as follows:

$$x_{n+1} = z_n - \alpha_n \gamma F z_n = z_n - \alpha_n (z_n - x_0) = \alpha_n x_0 + \left(1 - \alpha_n\right) z_n .$$

Therefore, in this special case, the calculation of x_{n+1} does not include the mapping F. The algorithm in Corollary 3.1 only contains the variational inequality mapping A, so this algorithm can solve the variational inequality problem. A similar statement applies to Corollary 3.2.

Corollary 3.1 *Let $A : H \to H$ be L_1-Lipschitz continuous and pseudomonotone on H. Take $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0, 1)$. Assume that $\left\{\alpha_n\right\}$ is a real sequence in $(0, 1)$ such that $\lim_{n\to\infty} \alpha_n = 0$ and $\sum_{n=0}^{\infty} \alpha_n = \infty$. Let $x_0, x_1 \in H$ and $\left\{x_n\right\}$ be defined by*

$$\left\{\begin{array}{l}
w_n = x_n + \theta_n\left(x_n - x_{n-1}\right) , \\
\text{where } \theta_n = \left\{\begin{array}{ll} \min\left\{\frac{\epsilon_n}{\|x_n - x_{n-1}\|}, \theta\right\}, & \text{if } x_n \neq x_{n-1}, \\ \theta, & \text{otherwise .} \end{array}\right. \\
y_n = P_C\left(w_n - \lambda_n A w_n\right) , \\
z_n = P_{T_n}\left(w_n - \lambda_n A y_n\right) , \\
\text{where } T_n := \left\{x \in H \mid \langle w_n - \lambda_n A w_n - y_n, x - y_n\rangle \leq 0\right\} , \\
x_{n+1} = \alpha_n x_0 + \left(1 - \alpha_n\right) z_n , \\
\lambda_{n+1} = \left\{\begin{array}{ll} \min\left\{\mu \frac{\|w_n - y_n\|^2 + \|z_n - y_n\|^2}{2\langle A w_n - A y_n, z_n - y_n\rangle}, \lambda_n\right\}, & \text{if } \langle A w_n - A y_n, z_n - y_n\rangle > 0, \\ \lambda_n, & \text{otherwise.} \end{array}\right.
\end{array}\right. \tag{3.27}$$

Then the iterative sequence $\left\{x_n\right\}$ created by (3.27) converges to $p \in \mathrm{VI}(C, A)$ in norm, where $p = P_{\mathrm{VI}(C,A)} x_0$.

3.2 The modified inertial Tseng's extragradient algorithm

In this subsection, we introduce a new modified inertial Tseng's extragradient algorithm for solving (BVIP). Our algorithm is described in Algorithm 3.2.

Algorithm 3.2 Modified inertial Tseng's extragradient algorithm for solving (BVIP)

Initialization: Set $\theta > 0$, $\lambda_1 > 0$, $\mu \in (0,1)$, $0 < \gamma < \frac{2\beta}{L_2^2}$. Arbitrarily given $x_0, x_1 \in H$.
Iterative Steps: Calculate x_{n+1} as follows:
Step 1. Given the iterates x_{n-1} and x_n $(n \geq 1)$. Set

$$w_n = x_n + \theta_n (x_n - x_{n-1}) .$$

where

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

Step 2. Compute

$$y_n = P_C (w_n - \lambda_n A w_n) .$$

Step 3. Compute

$$z_n = y_n - \lambda_n (A y_n - A w_n) .$$

Step 4. Compute

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

and update

$$\lambda_{n+1} = \begin{cases} \min\left\{\dfrac{\mu \|w_n - y_n\|}{\|A w_n - A y_n\|}, \lambda_n\right\}, & \text{if } A w_n - A y_n \neq 0, \\ \lambda_n, & \text{otherwise.} \end{cases} \tag{3.28}$$

Lemma 3.4 *The sequence* $\{\lambda_n\}$ *generated by* (3.28) *is a nonincreasing sequence and*

$$\lim_{n\to\infty} \lambda_n = \lambda \geq \min\left\{\lambda_1, \frac{\mu}{L_1}\right\}.$$

Proof It follows from (3.28) that $\lambda_{n+1} \leq \lambda_n$ for all $n \in \mathbb{N}$. Hence, $\{\lambda_n\}$ is nonincreasing. On the other hand, we get $L_1 \|w_n - y_n\| \geq \|A w_n - A y_n\|$ since A is L_1-Lipschitz continuous, consequently

$$\mu \frac{\|w_n - y_n\|}{\|A w_n - A y_n\|} \geq \frac{\mu}{L_1} \quad \text{if} \quad A w_n \neq A y_n ,$$

which together with (3.28) implies that

$$\lambda_n \geq \min\left\{\lambda_1, \frac{\mu}{L_1}\right\}.$$

Therefore, $\lim_{n\to\infty} \lambda_n = \lambda \geq \min\left\{\lambda_1, \frac{\mu}{L_1}\right\}$ since the sequence $\{\lambda_n\}$ is nonincreasing and lower bounded. □

The following lemma is very helpful for analyzing the convergence of Algorithm 3.2.

Lemma 3.5 *Assume that Conditions (C1)–(C3) hold. Let* $\{z_n\}$ *be a sequence generated by Algorithm 3.2. Then,*

$$\|z_n - p\|^2 \le \|w_n - p\|^2 - \left(1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2, \quad \forall p \in \mathrm{VI}(C, A)\,.$$

Proof First, using the definition of $\{\lambda_n\}$, one obtains

$$\|Aw_n - Ay_n\| \le \frac{\mu}{\lambda_{n+1}}\|w_n - y_n\|, \quad \forall n\,. \tag{3.29}$$

Indeed, if $Aw_n = Ay_n$ then (3.29) holds. Otherwise, it implies from (3.28) that

$$\lambda_{n+1} = \min\left\{\frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n\right\} \le \frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}\,.$$

Consequently,

$$\|Aw_n - Ay_n\| \le \frac{\mu}{\lambda_{n+1}}\|w_n - y_n\|\,.$$

Therefore, the inequality (3.29) holds when $Aw_n = Ay_n$ and $Aw_n \ne Ay_n$. By the definition of z_n, one sees that

$$\begin{aligned}\|z_n - p\|^2 =& \|y_n - p\|^2 + \lambda_n^2\|Ay_n - Aw_n\|^2 - 2\lambda_n\langle y_n - p, Ay_n - Aw_n\rangle\\ =& \|w_n - p\|^2 + \|y_n - w_n\|^2 - 2\langle y_n - w_n, y_n - w_n\rangle + 2\langle y_n - w_n, y_n - p\rangle\\ &+ \lambda_n^2\|Ay_n - Aw_n\|^2 - 2\lambda_n\langle y_n - p, Ay_n - Aw_n\rangle\\ =& \|w_n - p\|^2 - \|y_n - w_n\|^2 + 2\langle y_n - w_n, y_n - p\rangle\\ &+ \lambda_n^2\|Ay_n - Aw_n\|^2 - 2\lambda_n\langle y_n - p, Ay_n - Aw_n\rangle\,.\end{aligned} \tag{3.30}$$

Since $y_n = P_C(w_n - \lambda_n Aw_n)$, using the property of projection, we obtain

$$\langle y_n - w_n + \lambda_n Aw_n, y_n - p\rangle \le 0\,,$$

or equivalently

$$\langle y_n - w_n, y_n - p\rangle \le -\lambda_n\langle Aw_n, y_n - p\rangle\,. \tag{3.31}$$

From (3.29), (3.30) and (3.31), we have

$$\begin{aligned}\|z_n - p\|^2 \le& \|w_n - p\|^2 - \|y_n - w_n\|^2 - 2\lambda_n\langle Aw_n, y_n - p\rangle + \mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2}\|w_n - y_n\|^2\\ &- 2\lambda_n\langle y_n - p, Ay_n - Aw_n\rangle\\ \le& \|w_n - p\|^2 - \left(1 - \mu^2\frac{\lambda_n^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2 - 2\lambda_n\langle y_n - p, Ay_n\rangle\,.\end{aligned} \tag{3.32}$$

Since $p \in \mathrm{VI}(C, A)$, we have $\langle Ap, y_n - p\rangle \geq 0$. From the pseudomonotonicity of A, we get

$$\langle Ay_n, y_n - p\rangle \geq 0\,. \tag{3.33}$$

Combining (3.32) and (3.33), we obtain

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

The proof of the lemma is now complete. □

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 Step 1. The sequence $\{x_n\}$ is bounded. Indeed, according to Lemma 3.4, we have $\lim_{n\to\infty}\left(1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2}\right) = 1 - \mu^2 > 0$. Thus, there exists $n_0 \in \mathbb{N}$ such that

$$1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2} > 0, \quad \forall n \geq n_0\,. \tag{3.34}$$

Combining Lemma 3.5 and (3.34), we obtain

$$\|z_n - p\| \leq \|w_n - p\|, \quad \forall n \geq n_0\,.$$

As the same as (3.10)–(3.13), we have $\{x_n\}$ is bounded. We also get $\{z_n\}$ and $\{w_n\}$ are bounded.

Step 2.

$$\left(1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n M_4 \tag{3.35}$$

for some $M_4 > 0$. Indeed, from (3.14) and Lemma 3.5, we have

$$\begin{aligned}\|x_{n+1} - p\|^2 &= \|z_n - \alpha_n \gamma F z_n - p\|^2 \\ &\leq \|w_n - p\|^2 - \left(1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2 + \alpha_n M_2\end{aligned} \tag{3.36}$$

for some $M_2 > 0$. By (3.16), we obtain

$$\left(1 - \mu^2 \frac{\lambda_n^2}{\lambda_{n+1}^2}\right)\|w_n - y_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + \alpha_n M_4\,, \tag{3.37}$$

where $M_4 := M_2 + M_3$, both M_2 and M_3 are defined in Step 2 of Theorem 3.1.

Step 3.

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

for some $M>0$. Using (3.17) and (3.18), we can get the desired result immediately.

Step 4. $\{\|x_n-p\|^2\}$ converges to zero. According to Step 4 in Theorem 3.1, we suppose that $\{\|x_{n_k}-p\|\}$ is a subsequence of $\{\|x_n-p\|\}$ satisfying $\liminf_{k\to\infty}(\|x_{n_k+1}-p\|-\|x_{n_k}-p\|)\ge 0$. From Step 2 and Condition (C5), one obtains

$$\begin{aligned}\limsup_{k\to\infty}\Big(1-\mu^2\frac{\lambda_{n_k}^2}{\lambda_{n_k+1}^2}\Big)\|w_{n_k}-y_{n_k}\|^2 &\le \limsup_{k\to\infty}\{\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2+\alpha_{n_k}M_4\}\\ &\le 0\,.\end{aligned}$$

By (3.34), it follows that $\lim_{k\to\infty}\|y_{n_k}-w_{n_k}\|=0$. According to the definition of z_n in Algorithm 3.2 and (3.29), we have

$$\lim_{k\to\infty}\|z_{n_k}-y_{n_k}\|\le\lim_{k\to\infty}\lambda_{n_k}\|Ay_{n_k}-Aw_{n_k}\|\le\lim_{k\to\infty}\mu\frac{\lambda_{n_k}}{\lambda_{n_k+1}}\|w_{n_k}-y_{n_k}\|=0\,,$$

which implies that $\lim_{k\to\infty}\|y_{n_k}-z_{n_k}\|=0$. Using the same facts as (3.19)–(3.25), we obtain

$$\limsup_{k\to\infty}\Big[\frac{2\gamma}{\eta}\langle Fp,p-x_{n_k+1}\rangle+\frac{3M\theta_{n_k}}{\alpha_{n_k}\eta}\|x_{n_k}-x_{n_k-1}\|\Big]\le 0\,. \tag{3.38}$$

Therefore, using Step 3, Condition (C5) and (3.38), by means of Lemma 2.3, one concludes that $\lim_{n\to\infty}\|x_n-p\|=0$. The proof is completed. □

Now applying Theorem 3.2 with $F(x)=x-f(x)$, where $f:H\to H$ is a contraction mapping with constant $\rho\in[0,1)$. It is easy to verify that the mapping $F:H\to H$ is $(1+\rho)$-Lipschitz continuous and $(1-\rho)$-strongly monotone on H. In this case, by choosing $\gamma=1$, we obtain the following result.

Corollary 3.2 *Let $A:H\to H$ be L_1-Lipschitz continuous and pseudomonotone on H and $f: H\to H$ be a ρ-contraction mapping with $\rho\in[0,\sqrt{5}-2)$. Set $\theta>0$, $\lambda_1>0$, $\mu\in(0,1)$. Assume that $\{\alpha_n\}$ is a real sequence in $(0,1)$ such that $\lim_{n\to\infty}\alpha_n=0$ and $\sum_{n=0}^{\infty}\alpha_n=\infty$. Let $x_0,x_1\in H$ and $\{x_n\}$ be defined by*

$$\begin{cases} w_n = x_n + \theta_n\left(x_n - x_{n-1}\right), \\ \text{where } \theta_n = \begin{cases} \min\left\{\frac{\epsilon_n}{\|x_n - x_{n-1}\|}, \theta\right\}, & \text{if } x_n \neq x_{n-1}, \\ \theta, & \text{otherwise}. \end{cases} \\ y_n = P_C\left(w_n - \lambda_n A w_n\right), \\ z_n = y_n - \lambda_n\left(A y_n - A w_n\right), \\ x_{n+1} = \left(1 - \alpha_n\right) z_n + \alpha_n f\left(z_n\right), \\ \lambda_{n+1} = \begin{cases} \min\left\{\frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n\right\}, & \text{if } Aw_n - Ay_n \neq 0, \\ \lambda_n, & \text{otherwise}. \end{cases} \end{cases} \tag{3.39}$$

Then the iterative sequence $\{x_n\}$ *generated by* (3.39) *converges to* $p \in \mathrm{VI}(C, A)$ *in norm, where* $p = P_{\mathrm{VI}(C,A)} \circ f(p)$.

4 Numerical examples

In this section, we provide some numerical examples to show the numerical behavior of our proposed algorithms and compare them with Algorithm 1 and Algorithm 2 in [18]. We use the FOM Solver [29] to effectively calculate the projections onto C and T_n. 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.

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

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

and B is 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 M is positive semidefinite), $q \in \mathbb{R}^m$ is a vector with entries being generated in (0, 1) . It is clear that F is L_2-Lipschitz continuous and β-strongly monotone with $L_2 = \max\{\mathrm{eig}(M)\}$, $\beta = \min\{\mathrm{eig}(M)\}$, where $\mathrm{eig}(M)$ represents all eigenvalues of M.

Next, we consider the following fractional programming problem:

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

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 $C=\left\{x\in\mathbb{R}^5 : b^{\top}x+b_0>0\right\}$. Let

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

It is known that A is pseudomonotone and Lipschitz continuous.

We compare our proposed Algorithm 3.1 and Algorithm 3.2 with Algorithm 1 and Algorithm 2 proposed by Thong et al. [18]. Our parameters are set as follows. In all algorithms, set $\mu=0.1$, $\alpha_n=1/(n+1)$, $\gamma=1.7\beta/(L_2^2)$, $\lambda_1=0.6$. Take $\theta=0.4$, $\epsilon_n=100/(n+1)^2$ in Algorithm 3.1 and Algorithm 3.2. We use $D_n=\|x_n-x_{n-1}\|$ to measure the error of the n-th iteration since we do not know the exact solution to the problem, and the maximum iteration of 200 as the stopping criterion. Numerical results are reported in Figs. 1, 2.

Example 4.2 We consider an example that appears in the infinite dimensional Hilbert space $H=L^2[0,1]$ with the inner product $\langle x,y\rangle=\int_0^1 x(t)y(t)\mathrm{d}t$ and the induced norm $\|x\|=(\int_0^1 x(t)^2\mathrm{d}t)^{1/2}$. Let r, R be two positive real numbers such that $R/(k+1)<r/k<r<R$ for some $k>1$. Take the feasible set $C=\{x\in H : \|x\|\le r\}$ and the operator $A : H\to H$ given by

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

Note that A is not monotone. Taking a particular pair $(\tilde{x},\tilde{y})=(\tilde{x},k\tilde{x})$ and choosing $\tilde{x}\in C$ such that $R/(k+1)<\|\tilde{x}\|<r/k$, one can see that $k\|\tilde{x}\|\in C$. By a straightforward computation, we have

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

Hence, the operator A is not monotone on C. Next we show that A is pseudomonotone. Indeed, assume $\langle A(x),y-x\rangle\ge 0$ for all $x,y\in C$, that is, $\langle (R-\|x\|)x,y-x\rangle\ge 0$. Since $\|x\|<R$, we have $\langle x,y-x\rangle\ge 0$. Therefore,

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

Let $F : H\to H$ 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,

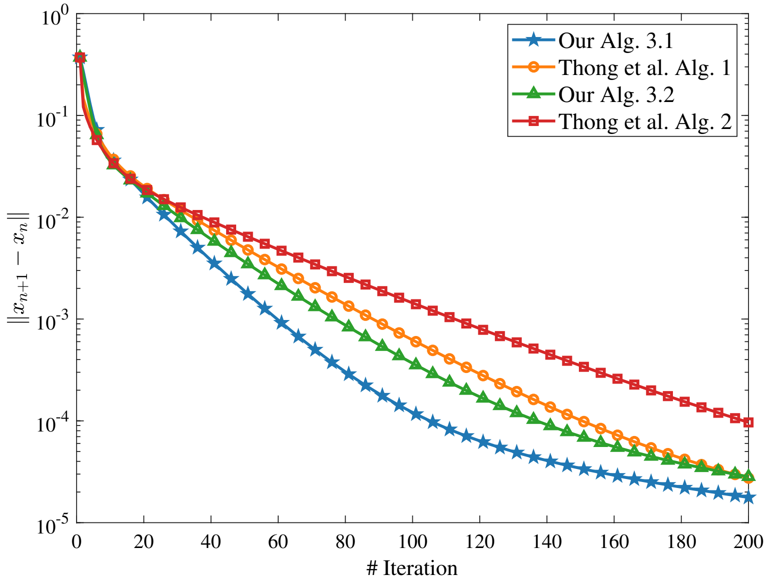

Fig. 1 Comparison of the number of iterations of all algorithms for Example 4.1

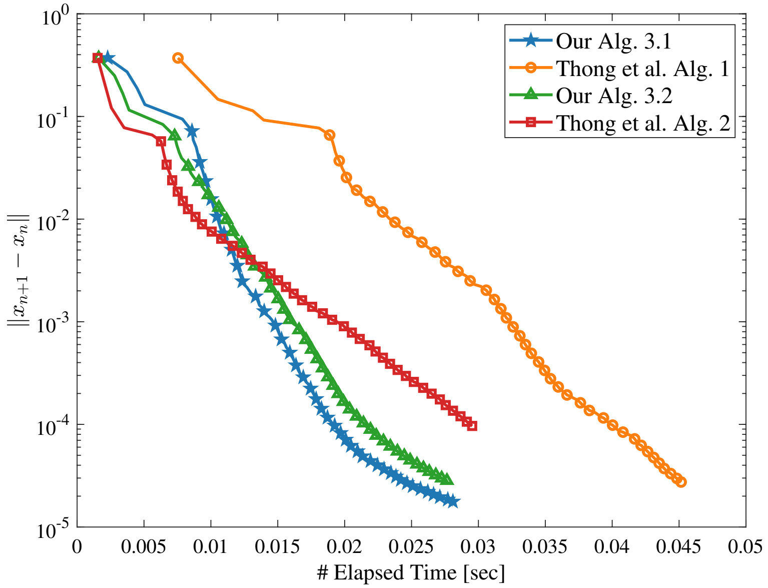

Fig. 2 Comparison of the elapsed time of all algorithms for Example 4.1

we choose $R = 1.5$, $r = 1$, $k = 1.1$. The solution of this problem is $x^*(t) = 0$. Our parameters are the same as in Example 4.1. The maximum iteration of 50 as the stopping criterion. Figs. 3, 4 show the behaviors of $D_n = \|x_n(t) - x^*(t)\|$ generated by all the algorithms with the starting points $x_0(t) = x_1(t) = t^2$. Moreover, we adjust the inertial parameters of the proposed algorithms to $\theta = 0.2$ and keep the other parameters the same as in Example 4.1. Figs. 5, 6 show the numerical behavior of all algorithms in this case.

Remark 4.1 We have the following comments on Examples 4.1 and 4.2.

- As shown in Figs. 1,2, 3, 4, 5, 6, in terms of the number of iterations and execution time, we can intuitively see that our proposed Algorithm 3.1 and Algorithm 3.2 are superior to the Algorithm 1 and the Algorithm 2 proposed by Thong et al. [18], respectively. It is worth noting that, due to the large inertial parameters we choose, our algorithms have higher accuracy and there are also oscillations. How to reduce oscillation is the next issue we need to consider in the future.
- The two algorithms proposed in this paper are semi-adaptive. That is, they can work without knowing the prior information of the Lipschitz constant of mapping A. However, in order to guarantee the strong convergence of the algorithms, we need to calculate $x_{n+1} = z_n - \alpha_n \gamma F z_n$, where $0 < \gamma < \frac{2\beta}{L_2^2}$, which requires the restriction that parameters β and L_2 must be known.

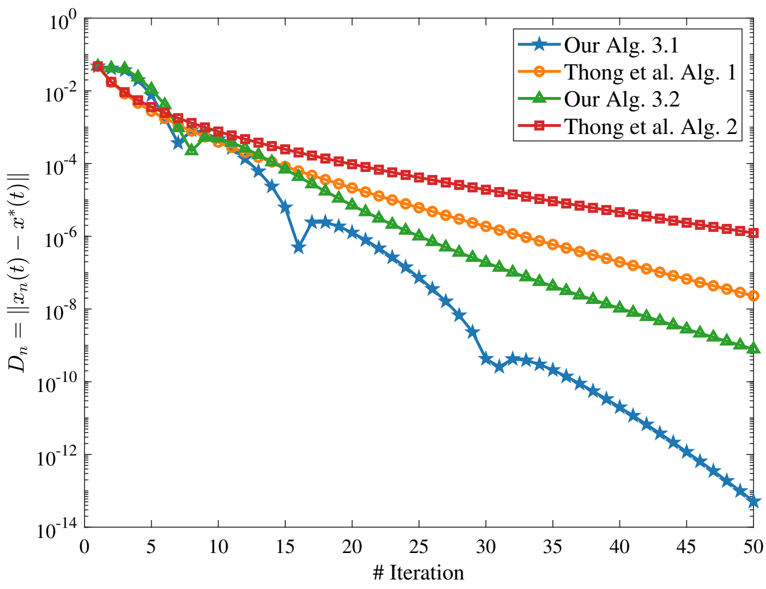

Fig. 3 Comparison of the number of iterations of all algorithms for Example 4.2 (inertial $\theta = 0.4$)

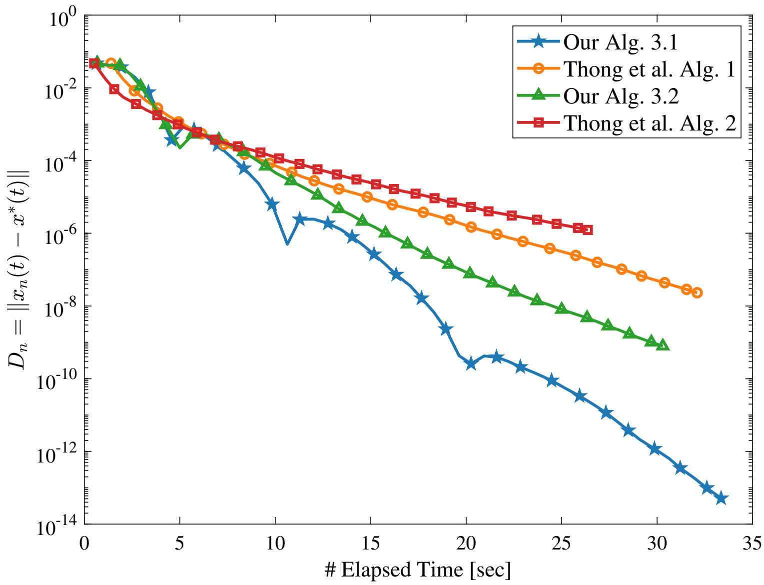

Fig. 4 Comparison of the elapsed time of all algorithms for Example 4.2 (inertial $\theta = 0.4$)

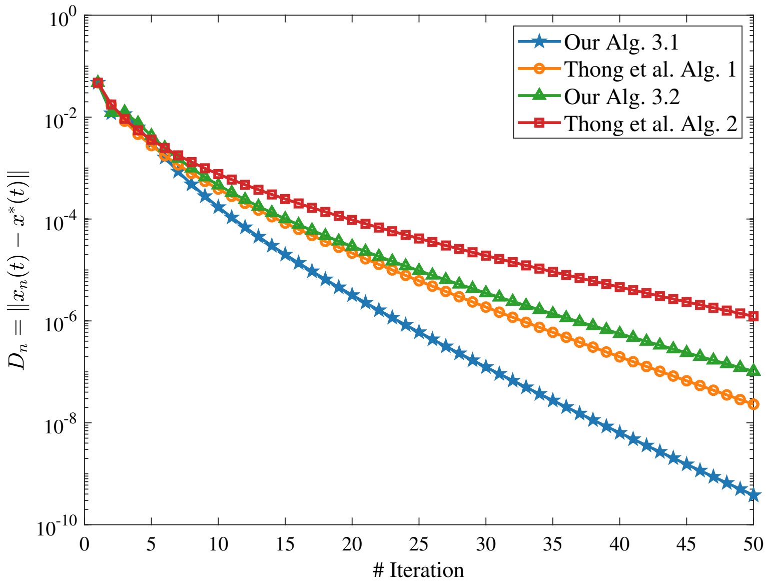

Fig. 5 Comparison of the number of iterations of all algorithms for Example 4.2 (inertial $\theta = 0.2$)

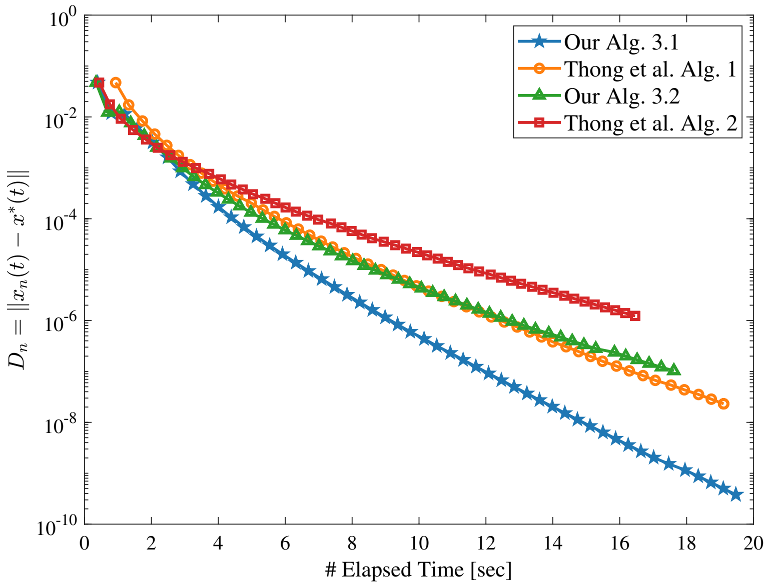

Fig. 6 Comparison of the elapsed time of all algorithms for Example 4.2 (inertial $\theta = 0.2$)

5 Conclusions

In this paper, we introduced two new extragradient algorithms to solve bilevel variational inequality problems in a Hilbert space. The algorithms were constructed with the aid of the inertial technique, the subgradient extragradient method, the Tseng's extragradient method and the steepest descent method. Only one projection onto the feasible set is needed at each iteration. Two new stepsize rules are used in our algorithms, which makes them easier to work without knowing the knowledge of the Lipschitz constant of the involved mapping. Two strong convergence theorems of the iterative sequences generated by the algorithms were proved. The theoretical results are also confirmed by some numerical examples.

Acknowledgements The authors are very grateful to the editor and anonymous referee for their useful comments, which significantly improved the original manuscript. This paper was supported by the National Natural Science Foundation of China under Grant No.11401152.

References

1. Xu, M.H., Li, M., Yang, C.C.: Neural networks for a class of bi-level variational inequalities. J. Global Optim. **44**, 535–552 (2009)
2. Anh, T.V.: Linesearch methods for bilevel split pseudomonotone variational inequality problems. Numer. Algorithm **81**, 1067–1087 (2019)
3. Cho, S.Y., Li, W., Kang, S.M.: Convergence analysis of an iterative algorithm for monotone operators. J. Inequal. Appl. **2013**, 199 (2013)
4. Hieu, D.V., Moudafi, A.: Regularization projection method for solving bilevel variational inequality problem. Optim. Lett. (2020). https://doi.org/10.1007/s11590-020-01580-5

5. Shehu, Y., Vuong, P.T., Zemkoho, A.: An inertial extrapolation method for convex simple bilevel optimization. Optim. Methods Softw. (2019). https://doi.org/10.1080/10556788.2019.1619729
6. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Ekonomikai Matematicheskie Metody. **12**, 747–756 (1976)
7. Censor, Y., Gibali, A., Reich, S.: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. Optim. Methods Softw. **26**, 827–845 (2011)
8. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control Optim. **38**, 431–446 (2000)
9. Liu, L., Tan, B., Cho, S.Y.: On the resolution of variational inequality problems with a double-hierarchical structure. J. Nonlinear Convex Anal. **21**, 377–386 (2020)
10. Sahu, D.R., Yao, J.C., Verma, M., Shukla, K.K.: Convergence rate analysis of proximal gradient methods with applications to composite minimization problems. Optimization (2020). https://doi.org/10.1080/02331934.2019.1702040
11. Cao, Y., Guo, K.: On the convergence of inertial two-subgradient extragradient method for variational inequality problems. Optimization **69**, 1237–1253 (2020)
12. Yang, J., Liu, H., Liu, Z.: Modified subgradient extragradient algorithms for solving monotone variational inequalities. Optimization **67**, 2247–2258 (2018)
13. Yang, J., Liu, H.: Strong convergence result for solving monotone variational inequalities in Hilbert space. Numer. Algorithms **80**, 741–752 (2019)
14. Yamada, I.: The hybrid steepest descent method for the variational inequality problem over the intersection of fixed point sets of nonexpansive mappings. Inherently Parallel Algorithms Feasibility Optimiz. Appl. **8**, 473–504 (2001)
15. Liu, L.: A hybrid steepest descent method for solving split feasibility problems involving nonexpansive mappings. J. Nonlinear Convex Anal. **20**, 471–488 (2019)
16. Jung, J.S.: Iterative algorithms based on the hybrid steepest descent method for the split feasibility problem. J. Nonlinear Sci. Appl. **9**, 4214–4225 (2016)
17. Takahashi, W., Wen, C.F., Yao, J.C.: The shrinking projection method for a finite family of demimetric mappings with variational inequality problems in a Hilbert space. Fixed Point Theory **19**, 407–419 (2018)
18. Thong, D.V., Triet, N.A., Li, X.-H., Dong, Q.-L.: Strong convergence of extragradient methods for solving bilevel pseudo-monotone variational inequality problems. Numer. Algorithms **83**, 1123–1143 (2020)
19. Polyak, B.T.: Some methods of speeding up the convergence of iteration methods. USSR. Comput. Math. Math. Phys. **4**, 1–17 (1964)
20. Nesterov, Y.: Gradient methods for minimizing composite functions. Math. Program. **140**, 125–161 (2013)
21. Dong, Q.-L., Lu, Y.Y., Yang, J.F.: The extragradient algorithm with inertial effects for solving the variational inequality. Optimization **65**, 2217–2226 (2016)
22. Tan, B., Xu, S., Li, S.: Inertial shrinking projection algorithms for solving hierarchical variational inequality problems. J. Nonlinear Convex Anal. **21**, 871–884 (2020)
23. Tan, B., Xu, S.: Strong convergence of two inertial projection algorithms in Hilbert spaces. J. Appl. Numer. Optim. **2**, 171–186 (2020)
24. Shehu, Y., Gibali, A.: New inertial relaxed method for solving split feasibilities. Optim. Lett. (2020). https://doi.org/10.1007/s11590-020-01603-1
25. Muu, L.D., Quy, N.V.: On existence and solution methods for strongly pseudomonotone equilibrium problems. Vietnam J. Math. **43**, 229–238 (2015)
26. Cottle, R.W., Yao, J.C.: Pseudo-monotone complementarity problems in Hilbert space. J. Optim. Theory Appl. **75**, 281–295 (1992)
27. Saejung, S., Yotkaew, P.: Approximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. **75**, 742–750 (2012)
28. Thong, D.V., Hieu, D.V., Rassias, T.M.: Self adaptive inertial subgradient extragradient algorithms for solving pseudomonotone variational inequality problems. Optim. Lett. **14**, 115–144 (2020)
29. Beck, A., Guttmann-Beck, N.: FOM–a MATLAB toolbox of first-order methods for solving convex optimization problems. Optim. Methods Softw. **34**, 172–193 (2019)