Noname manuscript No.
(will be inserted by the editor)

An novel adaptive inertial algorithm for solving bilevel variational inequalities with pseudomonotone multivalued operators

Zhong-bao Wang · Zhen-yin Lei · Xin Long · Yi Jiang

the date of receipt and acceptance should be inserted later

Abstract This paper aims to develop an adaptive inertial algorithm for solving bilevel variational inequalities with multivalued pseudomonotone operators in real Hilbert spaces and establish its strong convergence property. The algorithm does not need to know the prior information of the Lipschitz constants and strong monotonicity coefficients of the associated mappings, incorporates inertial techniques and involves only one projection per iteration. The step sizes of the proposed algorithm are updated at each iteration by a cheap computation without any linesearch procedure. Some numerical experiments show that the proposed algorithm has competitive advantages over some existing algorithms.

Keywords Bilevel variational inequality · Adaptive inertial algorithm · Pseudomonotone multivalued operator · Strong convergence

Mathematics Subject Classification (2010) 47H05 · 47J20 · 47J25 · 90C25

This work was supported by the National Natural Science Foundation of China (11701479, 11701478), the Chinese Postdoctoral Science Foundation (2018M643434) and the Fundamental Research Funds for the Central Universities (2682021ZT-PY040).

Zhong-bao Wang, Corresponding author
1. Department of Mathematics, Southwest Jiaotong University
Chengdu, Sichuan 611756, China
2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology
Chengdu, Sichuan 611756, China
3. School of Mathematical Sciences, University of Electronic Science and Technology of China
Chengdu, Sichuan, 611731, China
zhongbaowang@hotmail.com

Zhen-yin Lei
1. Department of Mathematics, Southwest Jiaotong University
Chengdu, Sichuan 611756, China
2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology
Chengdu, Sichuan 611756, China
lzy0615edu@163.com

Xin Long
1. Department of Mathematics, Southwest Jiaotong University
Chengdu, Sichuan 611756, China
2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology
Chengdu, Sichuan 611756, China
1065123400@qq.com

Yi Jiang
1. Department of Mathematics, Southwest Jiaotong University
Chengdu, Sichuan 611756, China
2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology
Chengdu, Sichuan 611756, China
1015441963@qq.com

1 Introduction

Suppose that H is a real Hilbert space and C is a nonempty closed convex subset in H. The inner product and norm of the Hilbert space H are represented by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. Let $A: H \to 2^H$ be a multivalued operator, and $F: H \to H$ be a single-valued mapping. We pay attention to the following bilevel multivalued variational inequality problems (shortly, BMVIPs): find $x^* \in VI(C, A)$, such that

$$\langle F(x^*), y - x^* \rangle \geq 0, \ \forall y \in VI(C, A), \tag{1.1}$$

where $VI(C, A)$ denotes the solution set of the multivalued variational inequality problems (shortly, MVIPs), that is,

$$VI(C, A) = \{y^* \in C : \exists\, w^* \in A(y^*) \text{ such that} \langle w^*, z - y^* \rangle \geq 0, \ \forall\, z \in C\}. \tag{1.2}$$

It's worth noting that the BMVIPs (1.1)-(1.2) serve as computational frameworks for important applications from machine learning, image processing, transportation, economics, engineering, circuits in electronics and other applied fields, see [1–5]. Thus, there are growing interests for studying numerical algorithms to solve BMVIPs and related problems, see [6–23].

The extragradient method (shortly, EGM) was introduced by Korpelevich[24] for solving monotone variational inequality problems (shortly, VIPs). It needs to compute two projections per iteration, which may have a bad impact on the computational efficiency of the method when the projection on the feasible set is very difficult to calculate. The projection and contraction method (shortly, PCM) [11, 25, 26] and the subgradiend extragradient method (shortly, SEGM) [27] are outstanding algorithms in improving the extragradient methods and calculate only one projection in every iteration. Inspired by [28], combining advantages of PCM and SEGM, Thong and Vuong [29] introduced improved subgradient extragradient methods (shortly, ISEGM) for solving pseudomonotone variational inequalities in Hilbert spaces. The main spirit of ISEGM is adopting the step sizes rule of PCM in the second explicit projection step of SEGM.

Multivalued variational inequality problems (shortly, MVIPs) are generalizations of VIPs and numerical methods for solving MVIPs have attracted much attention among researchers, see [5–13]. Under the condition that multivalued operators are continuous, most of numerical algorithms that solve MVIPs need to use the linesearch procedure, which contains additional operations of the projection and mappings values and magnifies notably the computational burden of the corresponding algorithms, see [5–7, 9–11]. Recently, Anh, Thang and Thach [12] proposed Halpern projection methods without any linesearch procedure for solving pseudomonotone MVIPs. Under the condition that the multivalued operators are Lipschitz continuous, their algorithms are strongly convergent.

The inertial technique, which traces back to a discrete form of a second-order dissipative dynamical system [30, 31], has aroused great interests of scholars as a method to expedite the convergence rate of algorithms. It has been used widely to create algorithms of variational inequalities and related problems, and can improve distinctly computational efficiency of numerical methods, which has been confirmed in many numerical experiments, see [11, 17, 18, 21–23].

If the multivalued operator A is a single-valued mapping, then BMVIPs (1.1)-(1.2) reduces to bilevel variational inequality problems (shortly, BVIPs). Next we state some algorithms for solving BVIPs, which stimulate us to establish new efficient iterative schemes. Based on SEGM, Thong et al. [15] proposed an extragradient method for solving bilevel pseudo-monotone variational inequality problems in real Hilbert spaces. Tan, Liu and Qin [17] introduced inertial terms into extragradient algorithms and extended the corresponding results of [15]. Combining PCM and the inertial technique, Tan, Qin and Yao [18] established a new modified inertial projection and contraction algorithm with the linesearch procedure for solving BVIPs. Tan and Cho [21] investigated BVIPs involving a pseudomonotone operator and constructed the inertial algorithm for solving them by combining SEGM and PCM. Note that the methods of [15–18, 21–23] need to know in advance the Lipschitz constant and strong monotonicity coefficient of F. However, the Lipschitz constants and strong monotonicity coefficients are sometimes not available in practical applications, see [32]. Recently, without the knowledge of the Lipschitz constants and the strong monotonicity coefficients of the underlying operators, Hieu and Moudafi [19] introduced a new regularization projection method and Yang [20] established no inertial acceleration algorithms, for solving the BVIPs.

An interesting problem is whether one can further improve the methods considered by [17–23] to obtain a new algorithm with inertial steps for BMVIPs (1.1)-(1.2) without any prior information of the Lipschitz constants and the strong monotonicity coefficients of the associated mappings in infinite dimensional Hilbert spaces.

In the paper, we introduce a novel adaptive inertial algorithm for solving bilevel variational inequalities with multivalued pseudomonotone operators and prove the sequence from the proposed algorithm converges strongly to the unique solution of BMVIPs (1.1)-(1.2) in real Hilbert spaces. The novel algorithm requires only one projection per iteration, owns inertial terms and adaptive step sizes that can run without any linesearch process or the prior information of the Lipschitz constants and strong monotonicity coefficients of the associated mappings.

The arrangement of this paper is as follows. In Section 2, we recall some preliminary results and lemmas for further use. In Section 3, we present the new algorithm and analyze its convergence. In Section 4, we report numerical experiments to illustrate the feasibility and efficiency of the proposed algorithm. Some concluding remarks are given in the last section.

2 Preliminaries

Let us recall some notations, definitions and well-known results that will be used in this paper. The weak convergence of x_n to x is denoted by $x_n \rightharpoonup x$ and the strong convergence of x_n to x is written $x_n \to x$ as $n \to \infty$.

Definition 2.1 Let $A:\ H \to 2^H$ be a multivalued mapping such that $A(x)$ is a nonempty closed convex set for each $x \in H$.

(a) The mapping A is called

(i) monotone on H if

$$\langle u - v,\ x - y\rangle \geq 0,\ \forall x,\ y \in H,\ u \in A(x),\ v \in A(y);$$

(ii) pseudomonotone on H if

$$\langle v,\ x - y\rangle \geq 0 \Longrightarrow \langle u,\ x - y\rangle \geq 0,\ \forall x,\ y \in H,\ u \in A(x),\ v \in A(y);$$

(iii) L-Lipschitz continuous if there exists a constant $L > 0$ on H such that

$$\rho(A(x),\ A(y)) \leq L\|x - y\|,\ \forall x,\ y \in H,$$

where ρ denotes the Hausdorff distance, that is, for any nonempty subsets D_1 and D_2 of H the Hausdorff distance $\rho(D_1,\ D_2)$ between D_1 and D_2 is defined by

$$\rho(D_1,\ D_2) :=\ \max\{\sup_{x\in D_1,\ y\in D_2}\inf \|x - y\|,\ \inf_{x\in D_1,\ y\in D_2}\sup \|x - y\|\}.$$

(b) Let $F:\ H \to H$ be a single-valued mapping. The mapping F is called

(i) L-Lipschitz continuous if there exists a constant $L > 0$ on H such that

$$\|F(x) - F(y)\| \leq L\|x - y\|,\ \forall x,\ y \in H;$$

if $L \in (0, 1)$, then F is also called a contraction mapping;

(ii) β-strongly monotone, if there exists a constant $\beta > 0$ such that

$$\langle F(x) - F(y),\ x - y\rangle \geq \beta\|x - y\|^2,\ \forall x,\ y \in H.$$

(c) Given a vector $x \in H$, the projection of x onto C, denoted by $P_C(x)$, is defined by

$$P_C(x) :=\ \text{argmin}_{y\in C}\|x - y\|.$$

Lemma 2.1 *[33] The mapping P_C has the following characteristic properties.*

(i) For any $x \in H,\ z = P_C(x)$ if and only if $\langle z - x,\ y - z\rangle \geq 0,\ \forall y \in C$;
(ii) $\|P_C(x) - P_C(y)\|^2 \leq \langle P_C(x) - P_C(y),\ x - y\rangle,\ \forall x,\ y \in H$;
(iii) $\|P_C(x) - y\|^2 \leq \|x - y\|^2 - \|x - P_C(x)\|^2,\ \forall x \in H,\ \forall y \in C$.

Lemma 2.2 *[34] For any $x, y \in H$, $\|x + y\|^2 \leq \|x\|^2 + 2\langle y,\ x + y\rangle$.*

Lemma 2.3 *[35] Let sequences of real numbers $\{b_n\}$, $\{\psi_n\}$ and $\{\beta_n\}$ satisfy $\psi_n \geq 0$, $\beta_n \in (0,1)$ and $\sum_{n=1}^{\infty} \beta_n = \infty$. Assume that*

$$\psi_{n+1} \leq (1 - \beta_n)\psi_n + \beta_n b_n, \quad \forall\ n \geq 1.$$

If $\limsup_{k\to\infty} b_{n_k} \leq 0$, for every subsequence $\{\psi_{n_k}\}$ of $\{\psi_n\}$ satisfying $\liminf_{k\to\infty}(\psi_{n_k+1} - \psi_{n_k}) \geq 0$, then $\lim_{n\to\infty}\psi_n = 0$.

3 Main results

In this section, we introduce an adaptive inertial algorithm to solve BMVIPs (1.1)-(1.2) and establish the strong convergence of the proposed algorithm. To analyse the strong convergence of the proposed algorithm, the following assumptions are presented.

Assumption 3.1 *(i) The solution set VI$(C,\ A)$ is nonempty.*
(ii) The mapping $A:\ H \to 2^H$ is pseudomonotone, and L_1-Lipschitz continuous with nonempty closed and convex values.
(iii) If $x_k \rightharpoonup x^*$ and $w_k \in A(x_k)$, then there exists a subsequence $\{w_{k_j}\}$ of $\{w_k\}$ such that $w_{k_j} \rightharpoonup w^* \in A(x^*)$ as $j \to +\infty$.
(iv) The mapping $F:\ H \to H$ is β-strongly monotone and L_2-Lipschitz continuous.

Assumption 3.2 *The positive real sequences $\{\epsilon_n\}$, $\{\gamma_n\}$, $\{q_n\}$, $\{q_n'\}$ and $\{p_n\}$ satisfy*
(i) $\gamma_n \in (0,\ 1)$, $\sum_{n=1}^{\infty} \gamma_n = \infty$, $\lim_{n\to\infty} \gamma_n = 0$ and $\lim_{n\to\infty} \frac{\epsilon_n}{\gamma_n} = 0$;
(ii) $\sum_{n=1}^{\infty} p_n < \infty$, $\{q_n\} \subset [1, +\infty)$ and $\{q_n'\} \subset [1, +\infty)$ with $\lim_{n\to\infty} q_n = \lim_{n\to\infty} q_n' = 1$.

If $A = A_1$ is a single-valued mapping in BMVIPs (1.1)-(1.2), then BMVIPs (1.1)-(1.2) reduces to the following BVIPs : find $x^* \in VI(C,\ A_1)$, such that

$$\langle F(x^*),\ y - x^*\rangle \geq 0,\ \forall y \in VI(C,\ A_1), \tag{3.1}$$

where $VI(C,\ A_1)$ is the solution set of the VIPs, that is,

$$VI(C,\ A_1) = \{\bar{P}:\ \langle A_1(\bar{P}), y - \bar{P}\rangle \geq 0,\ \forall\ y \in C\}. \tag{3.2}$$

Assumption 3.3 *(i) The mapping $A_1 : H \to H$ is pseudomonotone, Lipschitz continuous and for any $\{x_n\} \subset H$ with $x_n \rightharpoonup w^*$, one has $\|A_1 w^*\| \leq \liminf_{n\to\infty} \|A_1 x_n\|$.*
(ii) The solution set VI$(C,\ A_1)$ is nonempty.
(iii) The mapping $F:\ H \to H$ is β-strongly monotone and L_2-Lipschitz continuous.

Algorithm 3.1 *Take $x_0,\ x_1 \in H$, $\mu,\ \mu' \in (0,1)$ with $\mu < \mu'$, $\delta,\ \delta' \in (0, \frac{2}{\mu'})$ with $\delta < \delta'$, $\alpha \in (0,1)$, $\tau \in (\frac{\delta'}{2}, \frac{1}{\mu'})$, $\gamma > 0$ and $\lambda_1 > 0$, and choose the sequences $\{\epsilon_n\}$, $\{\gamma_n\}$, $\{q_n\}$, $\{q_n'\}$ and $\{p_n\}$ such that Assumption 3.2 holds.*

Step 1. Let $\{\alpha_n\}$ be a sequence such that $0 \leq \alpha_n \leq \bar{\alpha}_n$, where

$$\bar{\alpha}_n = \begin{cases} \min\{\frac{\epsilon_n}{\|x_n - x_{n-1}\|},\ \alpha\}, & \text{if } x_n \neq x_{n-1}, \\ \alpha, & \text{otherwise.} \end{cases} \tag{3.3}$$

Step 2. Compute

$$w_n = x_n + \alpha_n(x_n - x_{n-1}). \tag{3.4}$$

Step 3. *Taking $u_n \in A(w_n)$, compute*

$$y_n = P_C(w_n - \tau\lambda_n u_n). \tag{3.5}$$

Step 4. *Taking $v_n = P_{A(y_n)}(u_n)$, compute*

$$d_n = w_n - y_n - \tau\lambda_n(u_n - v_n)$$

and

$$z_n = P_{T_n}(w_n - \lambda_n\eta_n v_n), \tag{3.6}$$

where

$$T_n = \{x \in H | \langle w_n - \tau\lambda_n u_n - y_n, x - y_n\rangle \leq 0\}$$

and

$$\eta_n = \begin{cases} \delta q_n' \frac{\langle w_n - y_n,\ d_n\rangle}{\|d_n\|^2}, & \text{if } d_n \neq 0, \\ \bar{\eta} \in [\delta\frac{1-\tau\mu'}{(1+\tau\mu')^2}, \frac{\delta'}{1-\tau\mu'}] & \text{otherwise.} \end{cases}$$

Step 5. *Update $x_{n+1} = z_n - \gamma\gamma_n F(z_n)$ and*

$$\lambda_{n+1} = \begin{cases} \min\{\mu q_n \frac{\|w_n - y_n\|}{\|u_n - v_n\|},\ \lambda_n + p_n\}, & \text{if } \|u_n - v_n\| \neq 0, \\ \lambda_n + p_n, & \text{otherwise.} \end{cases} \tag{3.7}$$

Let $n = n + 1$ and return to Step 1.

Remark 3.1 We have the following comments on the Assumptions and Algorithm 3.1.

• The conditions (ii) and (iii) of Assumption 3.1 have been used and given the corresponding concrete example in [12]. In Assumption 3.3, the condition "for any $\{x_n\} \subset H$ with $x_n \rightharpoonup w^*$, one has $\|A_1 w^*\| \leq \liminf_{n\to\infty} \|A_1 x_n\|$" is strictly weaker than the sequentially weakly continuous assumption, see [36,37].

• In order to get larger step sizes, the sequences $\{q_n\}$ and $\{q_n'\}$ are used to relax the parameter μ and δ, respectively. The equality (3.4) is called an inertial term, which can speed up notably convergence rate of numerical methods, see [11,17,18,21–23]. If $\alpha_n = 0$, $A = A_1$ is a single-valued mapping, then the step size λ_n is the same as the one of Algorithms 1 and 2 of [20].

• Computing the projection on the half-space T_n has an explicit formula, see [18]. Hence Algorithm 3.1 needs to calculate only one projection in each iteration except computing v_n. Motivated by [21], we use the step size $\tau\lambda_n$ to compute the y_n, which can improve the convergence rate of the algorithm.

• We only require the parameter $\gamma > 0$, other than $\gamma \in (0, \frac{2\beta}{L_2^2})$ as in [15–18,21–23]. Thus we do not need to know in advance the Lipschitz constant and strong monotonicity coefficient of F, which is an important innovation of Algorithm 3.1.

Remark 3.2 It follows from (3.3) that $\alpha_n\|x_n - x_{n-1}\| \leq \epsilon_n = \gamma_n \frac{\epsilon_n}{\gamma_n}$, this and Assumption 3.2 mean that

$$\lim_{n\to\infty} \alpha_n\|x_n - x_{n-1}\| = 0. \tag{3.8}$$

Remark 3.3 Let $\{p_n\}$ and $\{\lambda_n\}$ be generated by Algorithm 3.1, and $P = \sum_{n=1}^{\infty} p_n$. Thus

$$\lim_{n\to\infty} \lambda_n = \lambda \text{ and } \lambda_n,\ \lambda \in [\min\{\tfrac{\mu}{L_1}, \lambda_1\}, \lambda_1 + P].$$

Indeed, since A is L_1-Lipschitz continuous, according to the definition of v_n we get

$$\|u_n - v_n\| = \inf_{y'\in A(y_n)} \|u_n - y'\| \leq \sup_{x'\in A(w_n)} \inf_{y'\in A(y_n)} \|x' - y'\| \leq \rho(A(w_n), A(y_n)) \leq L_1\|w_n - y_n\|. \tag{3.9}$$

Thus

$$\mu q_n \frac{\|w_n - y_n\|}{\|u_n - v_n\|} \geq \mu q_n \frac{\|w_n - y_n\|}{L_1\|w_n - y_n\|} \geq \frac{\mu}{L_1}.$$

The rest of the proof is the same as in Lemma 3.1 of [38], and we omit it.

In order to obtain the strong convergence of Algorithm 3.1, we need the following lemmas.

Lemma 3.1 *Let $\{\eta_n\}$ and $\{d_n\}$ be from Algorithm 3.1. Then the following statements hold*

(i) there exists a positive integer $N > 0$ such that

$$(1+\tau\mu')\|w_n - y_n\| \geq \|d_n\| \geq (1-\tau\mu')\|w_n - y_n\| \text{ and } \tfrac{\delta'}{1-\tau\mu'} \geq \eta_n \geq \delta\tfrac{1-\tau\mu'}{(1+\tau\mu')^2},\ \forall\, n \geq N;$$

(ii) For all $n \geq N$, $d_n = 0$ if and only if $w_n = y_n$.

Proof It follows that from Algorithm 3.1, for each $n \geq 1$,

$$\langle w_n - y_n, d_n\rangle = \langle w_n - y_n, w_n - y_n - \tau\lambda_n(u_n - v_n)\rangle \geq \|w_n - y_n\|^2 - \tau\lambda_n\|w_n - y_n\|\|u_n - v_n\| \tag{3.10}$$

and

$$\|d_n\| = \|w_n - y_n - \tau\lambda_n(u_n - v_n)\| \geq \|w_n - y_n\| - \tau\lambda_n\|u_n - v_n\|. \tag{3.11}$$

Now, we consider two cases:

Case 1 If $\|u_n - v_n\| = 0$, then by (3.10) and (3.11), we have

$$\langle w_n - y_n, d_n\rangle = \|w_n - y_n\|^2 \geq (1-\tau\mu')\|w_n - y_n\|^2 \text{ and } \|d_n\| = \|w_n - y_n\| \geq (1-\tau\mu')\|w_n - y_n\|. \tag{3.12}$$

Case 2 If $\|u_n - v_n\| \neq 0$, then according to (3.10), (3.11), Remark 3.3 and the definition of λ_{n+1}, we have

$$\langle w_n - y_n, d_n\rangle \geq \|w_n - y_n\|^2 - \frac{\tau\lambda_n}{\lambda_{n+1}}\lambda_{n+1}\|w_n - y_n\|\|u_n - v_n\| \geq (1 - \frac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu)\|w_n - y_n\|^2 \tag{3.13}$$

and

$$\|d_n\| \geq (1 - \frac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu)\|w_n - y_n\|. \tag{3.14}$$

In addition,

$$\|d_n\| = \|w_n - y_n - \tau\lambda_n(u_n - v_n)\| \leq \|w_n - y_n\| + \tau\lambda_n\|u_n - v_n\| \leq (1 + \frac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu)\|w_n - y_n\|. \tag{3.15}$$

In view of Assumption 3.2 and Remark 3.3, we know that

$$\lim_{n\to\infty} q'_n\delta = \delta < \delta',\ \lim_{n\to\infty} 1 - \tfrac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu = 1-\tau\mu > 1-\tau\mu' \text{ and } \lim_{n\to\infty} \tfrac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu = \tau\mu < \tau\mu'.$$

Thus there is a positive integer N such that

$$q'_n\delta < \delta',\ 1 - \frac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu > 1-\tau\mu' \text{ and } 1 + \frac{\tau\lambda_n}{\lambda_{n+1}}q_n\mu < 1+\tau\mu',\ \forall\, n \geq N. \tag{3.16}$$

It follows from (3.12), (3.14), (3.15) and (3.16) that

$$(1+\tau\mu')\|w_n - y_n\| \geq \|d_n\| \geq (1-\tau\mu')\|w_n - y_n\|,\ \forall\, n \geq N. \tag{3.17}$$

If $d_n \neq 0$, then

$$\eta_n = q'_n\delta\frac{\langle w_n - y_n,\ d_n\rangle}{\|d_n\|^2} \leq q'_n\delta\frac{\|w_n - y_n\|}{\|d_n\|} \leq \frac{\delta'}{1-\tau\mu'},\ \forall\, n \geq N.$$

By (3.12), (3.13), (3.15) and (3.16), we have

$$\eta_n = q'_n\delta\frac{\langle w_n - y_n,\ d_n\rangle}{\|d_n\|^2} \geq \delta\frac{\langle w_n - y_n,\ d_n\rangle}{\|d_n\|^2} \geq \delta\frac{1-\tau\mu'}{(1+\tau\mu')^2},\ \forall\, n \geq N. \tag{3.18}$$

If $d_n = 0$, then $\frac{\delta'}{1-\tau\mu'} \geq \eta_n \geq \delta\frac{1-\tau\mu'}{(1+\tau\mu')^2}$ hold naturally.

On the other hand, note that (3.17) implies that $d_n = 0$ if and only if $w_n = y_n$ for all $n \geq N$. This completes the proof. □

Lemma 3.2 *Let the Assumptions 3.1-3.2 hold and* $\lim_{n\to+\infty}\|y_n - w_n\| = 0$*. If the sequence* $\{x_n\}$ *is bounded and there exists a subsequence* $\{x_{n_k}\}$ *of* $\{x_n\}$ *such that* $\{x_{n_k}\}$ *converges weakly to* $\bar{P} \in H$*, then* $\bar{P} \in VI(C,\ A)$*.*

Proof By the definition of y_{n_k} and Lemma 2.1 (i), we obtain

$$\langle w_{n_k} - \tau\lambda_{n_k} u_{n_k} - y_{n_k},\ x - y_{n_k}\rangle \le 0,\ \forall\ x \in C$$

and

$$\langle w_{n_k} - y_{n_k},\ x - y_{n_k}\rangle \le \langle \tau\lambda_{n_k} u_{n_k},\ x - y_{n_k}\rangle,\ \forall\ x \in C. \tag{3.19}$$

From (3.4), Remark 3.2 and Assumption 3.2, it follows that

$$\lim_{n\to\infty}\|w_n - x_n\| = \lim_{n\to\infty}\alpha_n\|x_n - x_{n-1}\| \le \lim_{n\to\infty}\gamma_n\frac{\epsilon_n}{\gamma_n} = 0.$$

Since the sequence $\{x_n\}$ is bounded and $\lim_{n\to+\infty}\|y_n - w_n\| = 0$, we know that $\{w_n\}$ and $\{y_n\}$ are bounded. Furthermore, in view of $x_{n_k} \rightharpoonup \bar{P}$, we get $w_{n_k} \rightharpoonup \bar{P}$ and $y_{n_k} \rightharpoonup \bar{P}$ as $k \to +\infty$. Owing to Assumption 3.1 (iii), without loss of generality, we can assume that $u_{n_k} \rightharpoonup u_{\bar{P}} \in A(\bar{P})$ and hence $\{u_{n_k}\}$ is bounded.

According to $\lim_{n\to\infty}\|y_n - w_n\| = 0$, Remark 3.3 and (3.19), we can establish

$$\begin{aligned}\liminf_{k\to\infty}\langle u_{n_k},\ x - w_{n_k}\rangle &= \liminf_{k\to\infty}\langle u_{n_k},\ x - (w_{n_k} - y_{n_k}) - y_{n_k}\rangle\\ &= \liminf_{k\to\infty}\langle u_{n_k},\ x - y_{n_k}\rangle - \lim_{k\to\infty}\langle u_{n_k},\ w_{n_k} - y_{n_k}\rangle\\ &= \liminf_{k\to\infty}\langle u_{n_k},\ x - y_{n_k}\rangle \ge 0,\ \forall\ x \in C.\end{aligned} \tag{3.20}$$

Due to (3.20), for each $\frac{1}{n_k}$, there exists $k_0 > 0$ such that for all $x \in C$

$$\langle u_{n_k},\ x - w_{n_k}\rangle + \frac{1}{n_k} \ge 0,\ \forall\ k > k_0. \tag{3.21}$$

Now we consider two cases:

Case 1 If there exists a subsequence $\{u_{n_{k_j}}\}$ of $\{u_{n_k}\}$ such that $u_{n_{k_j}} = 0$ for all $j \ge 0$. Due to $u_{n_{k_j}} \in A(w_{n_{k_j}})$ and $w_{n_{k_j}} \rightharpoonup \bar{P}$, Assumption 3.1 (iii) implies that $u_{\bar{P}} = 0 \in A(\bar{P})$. Thus, $\bar{P} \in VI(C,\ A)$.

Case 2 There is not any subsequence $\{u_{n_{k_j}}\}$ of $\{u_{n_k}\}$ such that $u_{n_{k_j}} = 0$ for all $j \ge 0$. We can choose the sequence $\{u_{N_j}\} \subset \{u_{n_k}\}$ with $u_{N_j} \ne 0$. Setting $\bar{v}_{N_j} = \frac{u_{N_j}}{\|u_{N_j}\|^2}$, we have $\langle u_{N_j},\ \bar{v}_{N_j}\rangle = 1$. This, along with (3.21) implies that

$$\langle u_{N_j},\ x + \frac{1}{N_j}\bar{v}_{N_j} - w_{N_j}\rangle \ge 0,\ \forall x \in C\ , j > k_0.$$

Combining this and the pseudomonotonicity of A, we get

$$\langle u,\ x + \frac{1}{N_j}\bar{v}_{N_j} - w_{N_j}\rangle \ge 0,\ \forall x \in C,\ j > k_0,\ u \in A(x + \frac{1}{N_j}\bar{v}_{N_j}). \tag{3.22}$$

Since u_{N_j} converges weakly to $u_{\bar{P}} \in A(\bar{P})$, the weakly lower semicontinuity of the norm implies that

$$0 \le \|u_{\bar{P}}\| \le \liminf_{j\to\infty}\|u_{N_j}\|.$$

Then, if $u_{\bar{P}} = 0$, then, obviously,

$$\langle u_{\bar{P}},\ x - \bar{P}\rangle \ge 0,\ \forall x \in C$$

and thus we have $\bar{P} \in VI(C,\ A)$.

Now, if $\liminf_{j\to\infty}\|u_{N_j}\| > 0$, then

$$0 \le \limsup_{j\to\infty}\frac{1}{N_j}\|\bar{v}_{N_j}\| = \limsup_{j\to\infty}\frac{\frac{1}{N_j}}{\|u_{N_j}\|} \le \frac{\limsup_{j\to\infty}\frac{1}{N_j}}{\liminf_{j\to\infty}\|u_{N_j}\|} = 0. \tag{3.23}$$

This means $\lim_{j\to\infty}\frac{1}{N_j}\|\bar{v}_{N_j}\|=0$. Given any $x\in C$, for each $u_x\in A(x)$, set $u_x^{N_j}=P_{A(x+\frac{1}{N_j}\bar{v}_{N_j})}(u_x)$. According to the definition of the projection and Lipschitz continuity of the mapping A, we have

$$\|u_x^{N_j}-u_x\|=\inf_{y\in A(x+\frac{1}{N_j}\bar{v}_{N_j})}\|y-u_x\|\le\sup_{u'_x\in A(x)}\inf_{y\in A(x+\frac{1}{N_j}\bar{v}_{N_j})}\|y-u'_x\|\le\rho(A(x),A(x+\frac{1}{N_j}\bar{v}_{N_j}))\le\frac{L_1}{N_j}\|\bar{v}_{N_j}\|$$

and hence (3.23) implies that

$$\limsup_{j\to\infty}\|u_x^{N_j}-u_x\|\le\limsup_{j\to\infty}\frac{L_1}{N_j}\|\bar{v}_{N_j}\|=0.$$

Substituting $u:=u_x^{N_j}\in A(x+\frac{1}{N_j}\bar{v}_{N_j})$ into (3.22), we get

$$\langle u_x^{N_j},\ x+\frac{1}{N_j}\bar{v}_{N_j}-w_{N_j}\rangle\ge0,\ \forall\ x\in C. \tag{3.24}$$

Using $\lim_{j\to\infty}u_x^{N_j}=u_x$, $w_{N_j}\rightharpoonup\bar{P}$ and $\lim_{j\to\infty}\frac{1}{N_j}\|\bar{v}_{N_j}\|=0$ in (3.24), we obtain

$$\langle u_x,\ x-\bar{P}\rangle\ge0. \tag{3.25}$$

For every $\frac{1}{k}\in(0,\ 1]$, define $x_k:\ =\frac{1}{k}x+(1-\frac{1}{k})\bar{P}\in C$. According to (3.25) we have

$$\begin{aligned}0\le\langle u_k,\ x_k-\bar{P}\rangle&=\langle u_k,\ \frac{1}{k}x+(1-\frac{1}{k})\bar{P}-\bar{P}\rangle\\&=\frac{1}{k}[\langle u_k,\ x-\bar{P}\rangle],\ \forall\ u_k\in A(x_k)\end{aligned}$$

and hence

$$\langle u_k,\ x-\bar{P}\rangle\ge0. \tag{3.26}$$

Since $u_k\in A(x_k)$ and $x_k\to\bar{P}$ as $k\to\infty$, Assumption 3.1 (iii) implies that there exists a subsequence $\{u_{k_j}\}$ of $\{u_k\}$ converges weakly to $\bar{u}\in A(\bar{P})$. Taking $u_k=u_{k_j}$ and $k\to\infty$ in (3.26), we obtain

$$\langle\bar{u},\ x-\bar{P}\rangle\ge0,\ \forall x\in C,$$

that is, $\bar{P}\in VI(C,\ A)$. This completes the proof. □

Remark 3.4 In the case that the mapping A is monotone, it is not necessary to impose (iii) of Assumption 3.1 on A. Indeed, it follows from (3.21) and the monotonicity of A that for all $x\in C$ and $u_x\in A(x)$

$$\langle u_x,\ x-w_{n_k}\rangle+\frac{1}{n_k}\ge0,\ \forall\ k>k_0. \tag{3.27}$$

Letting $k\to+\infty$ in (3.27), we have

$$\langle u_x,x-\bar{P}\rangle\ge0,\ \forall\ x\in C,$$

and the rest of the proof is the same as Lemma 3.2.

Lemma 3.3 *Let $\{z_n\}$ be a sequence generated by Algorithm 3.1 and $p\in VI(C,\ A)$. Then*

$$\|z_n-p\|^2\le\|w_n-p\|^2-\|w_n-z_n-\frac{1}{\tau}\eta_nd_n\|^2+(\frac{1}{\tau^2}-\frac{2}{\tau\delta'})\delta^2\frac{(1-\tau\mu')^4}{(1+\tau\mu')^4}\|w_n-y_n\|^2,\ \forall\ n\ge N.$$

Proof Since $y_n=P_C(w_n-\tau\lambda_nu_n)$, via Lemma 2.1 (i), we get

$$\langle w_n-\tau\lambda_nu_n-y_n,\ x-y_n\rangle\le0,\ \forall\ x\in C$$

and thus $C\subset T_n$. Since $p\in C\subset T_n$, Lemma 2.1 implies that

$$\begin{aligned}\|z_n-p\|^2&=\|P_{T_n}(w_n-\lambda_n\eta_nv_n)-P_{T_n}(p)\|^2\le\langle z_n-p,\ w_n-\lambda_n\eta_nv_n-p\rangle\\&=\frac{1}{2}\|z_n-p\|^2+\frac{1}{2}\|w_n-\lambda_n\eta_nv_n-p\|^2-\frac{1}{2}\|z_n-w_n+\lambda_n\eta_nv_n\|^2\\&=\frac{1}{2}\|z_n-p\|^2+\frac{1}{2}\|w_n-p\|^2-\langle w_n-p,\ \lambda_n\eta_nv_n\rangle-\frac{1}{2}\|z_n-w_n\|^2-\langle z_n-w_n,\ \lambda_n\eta_nv_n\rangle\\&=\frac{1}{2}\|z_n-p\|^2+\frac{1}{2}\|w_n-p\|^2-\frac{1}{2}\|z_n-w_n\|^2-\langle z_n-p,\ \lambda_n\eta_nv_n\rangle.\end{aligned}$$

Therefore

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \|z_n - w_n\|^2 - 2\langle \lambda_n \eta_n v_n,\ z_n - p\rangle. \tag{3.28}$$

Since $y_n \in C$ and $p \in VI(C,\ A)$, there exists $u_p \in A(p)$ such that $\langle u_p,\ y_n - p\rangle \geq 0$. The pseudomonotonicity of A implies that

$$\langle v_n,\ p - y_n\rangle \leq 0.$$

According to Lemma 3.1 and Remark 3.3, we know that

$$\begin{aligned} -2\lambda_n \tau \eta_n \langle v_n,\ z_n - y_n + y_n - p\rangle &= -2\lambda_n \tau \eta_n \langle v_n,\ y_n - p\rangle - 2\lambda_n \tau \eta_n \langle v_n,\ z_n - y_n\rangle \\ &\leq -2\lambda_n \tau \eta_n \langle v_n, z_n - y_n\rangle,\ \forall\ n \geq N. \end{aligned} \tag{3.29}$$

From $z_n = P_{T_n}(w_n - \lambda_n \eta_n v_n) \in T_n$, it follows that

$$\langle w_n - \tau\lambda_n u_n - y_n,\ z_n - y_n\rangle \leq 0$$

and thus

$$\langle w_n - y_n - \tau\lambda_n(u_n - v_n),\ z_n - y_n\rangle = \langle d_n,\ z_n - y_n\rangle \leq \tau\lambda_n\langle v_n,\ z_n - y_n\rangle.$$

This implies that

$$-2\lambda_n \tau \eta_n \langle v_n,\ z_n - y_n\rangle \leq -2\eta_n\langle d_n,\ z_n - y_n\rangle. \tag{3.30}$$

Combining (3.29) and (3.30), we get

$$-2\lambda_n \tau \eta_n \langle v_n,\ z_n - p\rangle \leq -2\eta_n\langle d_n,\ z_n - y_n\rangle = -2\eta_n\langle d_n,\ w_n - y_n\rangle + 2\eta_n\langle d_n,\ w_n - z_n\rangle,\ \forall n \geq N. \tag{3.31}$$

Now, we estimate $-2\frac{1}{\tau}\eta_n\langle d_n,\ w_n - y_n\rangle$ and $2\frac{1}{\tau}\eta_n\langle d_n,\ w_n - z_n\rangle$.

If $d_n \neq 0$, then by the definition of η_n, we have

$$\eta_n\|d_n\|^2 = \delta q'_n\langle d_n,\ w_n - y_n\rangle.$$

If $d_n = 0$, then $\eta_n\|d_n\|^2 = \delta q'_n\langle d_n,\ w_n - y_n\rangle$ holds obviously.
Thus

$$\frac{-2{\eta_n}^2\|d_n\|^2}{\tau\delta q'_n} = -2\frac{1}{\tau}\eta_n\langle d_n,\ w_n - y_n\rangle. \tag{3.32}$$

In addition, we note that

$$\begin{aligned} 2\frac{1}{\tau}\eta_n\langle d_n,\ w_n - z_n\rangle &= 2\langle \frac{1}{\tau}\eta_n d_n,\ w_n - z_n\rangle \\ &= -\|w_n - z_n - \frac{1}{\tau}\eta_n d_n\|^2 + \|w_n - z_n\|^2 + \frac{1}{\tau^2}\eta_n^2\|d_n\|^2. \end{aligned} \tag{3.33}$$

It follows from (3.31), (3.32) and (3.33) that

$$-2\lambda_n\eta_n\langle v_n,\ z_n - p\rangle \leq -\|w_n - z_n - \frac{1}{\tau}\eta_n d_n\|^2 + \|w_n - z_n\|^2 + (\frac{1}{\tau^2} - \frac{2}{\tau\delta q'_n})\eta_n^2\|d_n\|^2,\ \forall n \geq N. \tag{3.34}$$

It follows from (3.16) and $\tau \in (\frac{\delta'}{2}, \frac{1}{\mu'})$ that $\frac{1}{\tau^2} - \frac{2}{\tau\delta q'_n} < \frac{1}{\tau^2} - \frac{2}{\tau\delta'} < 0,\ \forall n \geq N$. In view of Lemma 3.1, (3.28) and (3.34), we obtain

$$\|z_n - p\|^2 \leq \|w_n - p\|^2 - \|w_n - z_n - \frac{1}{\tau}\eta_n d_n\|^2 + (\frac{1}{\tau^2} - \frac{2}{\tau\delta'})\delta^2\frac{(1-\tau\mu')^4}{(1+\tau\mu')^4}\|w_n - y_n\|^2.$$

The proof is completed. □

Next, we give the following main theorem.

Theorem 3.1 *If the Assumptions 3.1-3.2 hold, then the sequence $\{x_n\}$ generated by Algorithm 3.1 converges strongly to the unique solution of BMVIPs (1.1)-(1.2).*

Proof Due to Assumption 3.1 and [1,2], we know that BMVIPs (1.1)-(1.2) have a unique solution, which is denoted by p.

Since $F:\ H\to H$ is β-strongly monotone and L_2-Lipschitz continuous, we have

$$\beta\|x-y\|^2\le\langle F(x)-F(y),\ x-y\rangle\le\|F(x)-F(y)\|\|x-y\|\le L_2\|x-y\|^2,\ \forall\ x,\ y\in H.$$

Hence

$$L_2\ge\beta \text{ and } 1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2)\ge(1-(\gamma\gamma_n)^{\frac{1}{2}}\beta)^2.$$

Furthermore, we know that

$$\begin{aligned}&\|(I-(\gamma\gamma_n)^{\frac{1}{2}}F)(z_n)-(I-(\gamma\gamma_n)^{\frac{1}{2}}F)(p)\|^2\\
&=\|z_n-p\|^2+\gamma\gamma_n\|F(z_n)-F(p)\|^2-2(\gamma\gamma_n)^{\frac{1}{2}}\langle F(z_n)-F(p),\ z_n-p\rangle\\
&\le\|z_n-p\|^2+\gamma\gamma_n{L_2}^2\|z_n-p\|^2-2\beta(\gamma\gamma_n)^{\frac{1}{2}}\|z_n-p\|^2\\
&=(1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2))\|z_n-p\|^2.\end{aligned}$$

According to $\lim_{n\to\infty}(\gamma\gamma_n)^{\frac{1}{2}}=0$, we deduce that there exists a positive integer N_1 such that

$$(\gamma\gamma_n)^{\frac{1}{2}}<\min\{1,\frac{2\beta}{L_2^2}\},\ \forall\ n\ge N_1.$$

Let's assume for the sake of convenience $N\ge N_1$. Thus, for all $n\ge N$

$$\begin{aligned}&\|(I-\gamma\gamma_nF)(z_n)-(I-\gamma\gamma_nF)(p)\|\\
&\le(\gamma\gamma_n)^{\frac{1}{2}}\|(I-(\gamma\gamma_n)^{\frac{1}{2}}F)(z_n)-(I-(\gamma\gamma_n)^{\frac{1}{2}}F)(p)\|+(1-(\gamma\gamma_n)^{\frac{1}{2}})\|z_n-p\|\\
&\le[1-(\gamma\gamma_n)^{\frac{1}{2}}(1-\sqrt{1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2)}]\|z_n-p\|=(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|z_n-p\|,\end{aligned}\tag{3.35}$$

where $\Gamma_n=1-\sqrt{1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2)}$.

We know that for all $n\ge N$, $(\gamma\gamma_n)^{\frac{1}{2}}<\min\{1,\frac{2\beta}{L_2^2}\}$ implies that $\Gamma_n\in(0,1],\ \forall\ n\ge N$. Thanks to Lemma 3.3, we have

$$\|z_n-p\|\le\|w_n-p\|,\ \forall\ n\ge N.\tag{3.36}$$

In view of Assumption 3.2, we know $\lim_{n\to\infty}(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\frac{\alpha_n}{\gamma\gamma_n}\|x_n-x_{n-1}\|=0$ and thus there exists $M>0$ such that

$$(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\frac{\alpha_n}{\gamma\gamma_n}\|x_n-x_{n-1}\|\le M,\ \forall\ n\ge N.$$

It follows from (3.35) and (3.36) that

$$\begin{aligned}\|x_{n+1}-p\|&=\|z_n-\gamma\gamma_nF(z_n)-p\|\\
&=\|(I-\gamma\gamma_nF)(z_n)-(I-\gamma\gamma_nF)(p)-\gamma\gamma_nF(p)\|\\
&\le\|(I-\gamma\gamma_nF)(z_n)-(I-\gamma\gamma_nF)(p)\|+\|\gamma\gamma_nF(p)\|\\
&\le(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|z_n-p\|+\gamma\gamma_n\|F(p)\|\\
&\le(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|w_n-p\|+\gamma\gamma_n\|F(p)\|\\
&\le(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|x_n-p+\alpha_n(x_n-x_{n-1})\|+\gamma\gamma_n\|F(p)\|\\
&\le(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|x_n-p\|+[(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\frac{\alpha_n}{\gamma\gamma_n}\|x_n-x_{n-1}\|]\gamma\gamma_n+\gamma\gamma_n\|F(p)\|\\
&\le(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|x_n-p\|+\gamma\gamma_nM+\gamma\gamma_n\|F(p)\|\\
&=(1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|x_n-p\|+(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}(M+\|F(p)\|)\\
&\le\max\{\|x_n-p\|,\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}(M+\|F(p)\|)\},\ \forall\ n\ge N.\end{aligned}\tag{3.37}$$

By the definition of Γ_n and Assumption 3.2, we get

$$\begin{aligned}\lim_{n\to\infty}\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n} &= \lim_{n\to\infty}\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{1-\sqrt{1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2)}} \\ &= \lim_{n\to\infty}\frac{[1+\sqrt{1-(\gamma\gamma_n)^{\frac{1}{2}}(2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2)}]}{2\beta-(\gamma\gamma_n)^{\frac{1}{2}}{L_2}^2} = \frac{1}{\beta},\end{aligned} \tag{3.38}$$

which implies that there exists $M_2>0$ such that $0<\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}<M_2$. This, together with (3.37), yields

$$\|x_{n+1}-p\|\le\max\{\|x_n-p\|,M_2(M+\|F(p)\|)\}\le\cdots\le\max\{\|x_N-p\|,M_2(M+\|F(p)\|)\}. \tag{3.39}$$

This implies that the sequence $\{x_n\}$ is bounded.

By the definition of w_n, we get

$$\begin{aligned}\|w_n-p\|^2 &= \|x_n+\alpha_n(x_n-x_{n-1})-p\|^2 \\ &\le \|x_n-p\|^2+2\alpha_n\|x_n-x_{n-1}\|\|x_n-p\|+\alpha_n^2\|x_n-x_{n-1}\|^2.\end{aligned} \tag{3.40}$$

By using Lemma 2.2, (3.35), (3.36) and (3.40), we have

$$\begin{aligned}\|x_{n+1}-p\|^2 &= \|z_n-\gamma\gamma_nF(z_n)-p\|^2 \\ &= \|(I-\gamma\gamma_nF)(z_n)-(I-\gamma\gamma_nF)(p)-\gamma\gamma_nF(p)\|^2 \\ &\le \|(I-\gamma\gamma_nF)(z_n)-(I-\gamma\gamma_nF)(p)\|^2+2\gamma\gamma_n\langle F(p),\ p-x_{n+1}\rangle \\ &\le (1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)^2\|z_n-p\|^2+2\gamma\gamma_n\langle F(p),\ p-x_{n+1}\rangle \\ &\le (1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)^2\|w_n-p\|^2+2\gamma\gamma_n\langle F(p),\ p-x_{n+1}\rangle \\ &\le (1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)^2(\|x_n-p\|^2+\alpha_n\|x_n-x_{n-1}\|(2\|x_n-p\|+\alpha_n\|x_n-x_{n-1}\|)) \\ &\quad +2\gamma\gamma_n\langle F(p),\ p-x_{n+1}\rangle \\ &\le (1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)\|x_n-p\|^2+(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n[Q\frac{\alpha_n}{(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n}\|x_n-x_{n-1}\| \\ &\quad +2\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}\langle F(p),\ p-x_{n+1}\rangle],\ \forall\ n\ge N,\end{aligned} \tag{3.41}$$

where $Q=\sup_n\{2\|x_n-p\|+\alpha_n\|x_n-x_{n-1}\|\}>0$.

Set $\psi_n=\|x_n-p\|^2$, $\beta_n=(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n$ and $b_n=Q\frac{\alpha_n}{(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n}\|x_n-x_{n-1}\|+2\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}\langle F(p),\ p-x_{n+1}\rangle$. From (3.41) it follows that

$$\psi_{n+1}\le(1-\beta_n)\psi_n+\beta_nb_n,\quad\forall\ n\ge N.$$

According to (3.38), we have

$$\lim_{n\to\infty}\frac{\gamma_n}{\beta_n}=\lim_{n\to\infty}\frac{\gamma_n}{(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n}=\lim_{n\to\infty}\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\gamma\Gamma_n}=\frac{1}{\gamma\beta}.$$

Thus $\sum_{n=1}^{\infty}\gamma_n=\infty$ implies that $\sum_{n=1}^{\infty}\beta_n=\infty$.

In addition, since $(\gamma\gamma_n)^{\frac{1}{2}}<\min\{1,\frac{2\beta}{L_2^2}\}$ and $\Gamma_n\in(0,1]$, $\forall\ n\ge N$, we have for all $n\ge N$, $\beta_n\in(0,1)$. Let $\{\psi_{n_k}\}$ be a subsequence of $\{\psi_n\}$ such that $\liminf_{k\to\infty}(\psi_{n_k+1}-\psi_{n_k})\ge0$.

Utilizing Lemmas 2.2 and 3.3, (3.35), (3.40) and (3.41), we have

$$\begin{aligned}\|x_{n+1}-p\|^2 &= \|z_n-\gamma\gamma_n F(z_n)-p\|^2\\ &= \|(I-\gamma\gamma_n F)(z_n)-(I-\gamma\gamma_n F)(p)-\gamma\gamma_n F(p)\|^2\\ &\le \|(I-\gamma\gamma_n F)(z_n)-(I-\gamma\gamma_n F)(p)\|^2+2\gamma\gamma_n\langle F(p),\ p-x_{n+1}\rangle\\ &\le (1-(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n)^2\|z_n-p\|^2+2\gamma\gamma_n\|F(p)\|\|p-x_{n+1}\|\\ &\le \|z_n-p\|^2+\gamma_n Q_1\\ &\le \|w_n-p\|^2-\|w_n-z_n-\frac{1}{\tau}\eta_n d_n\|^2+(\frac{1}{\tau^2}-\frac{2}{\tau\delta'})\delta^2\frac{(1-\tau\mu')^4}{(1+\tau\mu')^4}\|w_n-y_n\|^2+\gamma_n Q_1\\ &\le \|x_n-p\|^2+\alpha_n\|x_n-x_{n-1}\|Q-\|w_n-z_n-\frac{1}{\tau}\eta_n d_n\|^2\\ &\quad+(\frac{1}{\tau^2}-\frac{2}{\tau\delta'})\delta^2\frac{(1-\tau\mu')^4}{(1+\tau\mu')^4}\|w_n-y_n\|^2+\gamma_n Q_1,\ \forall\ n\ge N,\end{aligned} \tag{3.42}$$

where $Q_1=\sup_n\{2\gamma\|F(p)\|\|p-x_{n+1}\rangle\|\}>0$. Combining (3.42) and the assumptions $\delta'\in(0,\frac{2}{\mu'})$ and $\tau\in(\frac{\delta'}{2},\frac{1}{\mu'})$, we infer that

$$\begin{aligned}&\limsup_{k\to\infty}[\|w_{n_k}-z_{n_k}-\frac{1}{\tau}\eta_{n_k}d_{n_k}\|^2-(\frac{1}{\tau^2}-\frac{2}{\tau\delta'})\delta^2\frac{(1-\tau\mu')^4}{(1+\tau\mu')^4}\|w_{n_k}-y_{n_k}\|^2]\\ &\le \limsup_{k\to\infty}(\|x_{n_k}-p\|^2-\|x_{n_k+1}-p\|^2)+\limsup_{k\to\infty}(\alpha_{n_k}\|x_{n_k}-x_{n_k-1}\|Q+\gamma_{n_k}Q_1)\\ &\le -\liminf_{k\to\infty}(\psi_{n_k+1}-\psi_{n_k})+\limsup_{k\to\infty}(\frac{\gamma_{n_k}\epsilon_{n_k}}{\gamma_{n_k}}Q+\gamma_{n_k}Q_1)\le -\liminf_{k\to\infty}(\psi_{n_k+1}-\psi_{n_k})\le 0,\end{aligned}$$

which implies

$$\lim_{k\to\infty}\|w_{n_k}-z_{n_k}-\tfrac{1}{\tau}\eta_{n_k}d_{n_k}\|=0 \text{ and } \lim_{k\to\infty}\|w_{n_k}-y_{n_k}\|=0.$$

Since the sequence $\{x_{n_k+1}\}$ is bounded, there exists a subsequence $\{x_{n_{k_j}+1}\}$ of $\{x_{n_k+1}\}$ such that $x_{n_{k_j}+1}\rightharpoonup x^*$ as $j\to\infty$ and

$$\limsup_{k\to\infty}\langle F(p),\ p-x_{n_k+1}\rangle=\lim_{j\to\infty}\langle F(p),\ p-x_{n_{k_j}+1}\rangle=\langle F(p),\ p-x^*\rangle. \tag{3.43}$$

Lemma 3.2 and $\lim_{k\to\infty}\|w_{n_k}-y_{n_k}\|=0$ imply that $x^*\in VI(C,\ A)$. From (3.43) and the assumption p is the unique solution of BMVIPs (1.1)-(1.2), it follows that

$$\limsup_{k\to\infty}\langle F(p),\ p-x_{n_k+1}\rangle=\lim_{j\to\infty}\langle F(p),\ p-x_{n_{k_j}+1}\rangle=\langle F(p),\ p-x^*\rangle\le 0. \tag{3.44}$$

Using (3.38), the assumption $\lim_{n\to\infty}\frac{\epsilon_n}{\gamma_n}=0$ and the definition of α_n, we have

$$\lim_{n\to\infty}Q\frac{\alpha_n}{(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n}\|x_n-x_{n-1}\|\le\lim_{n\to\infty}Q\frac{\epsilon_n}{(\gamma\gamma_n)^{\frac{1}{2}}\Gamma_n}=\lim_{n\to\infty}Q\frac{(\gamma\gamma_n)^{\frac{1}{2}}}{\Gamma_n}\frac{\epsilon_n}{\gamma\gamma_n}=0.$$

This, together with (3.38) and (3.44) implies that

$$\limsup_{k\to\infty}b_{n_k}\le\lim_{k\to\infty}Q\frac{\epsilon_{n_k}}{(\gamma\gamma_{n_k})^{\frac{1}{2}}\Gamma_{n_k}}+\limsup_{k\to\infty}2\frac{(\gamma\gamma_{n_k})^{\frac{1}{2}}}{\Gamma_{n_k}}\langle F(p),\ p-x_{n_k+1}\rangle\le 0.$$

Thus Lemma 2.3 ensures that the sequence $\{x_n\}$ converges strongly to p. The proof is completed. □

Proposition 3.1 *Suppose that $A=A_1$ is a single-valued mapping, the Assumptions 3.2-3.3 hold, and $\{w_n\}$ and $\{y_n\}$ are sequences generated by Algorithm 3.1. If $\lim_{n\to\infty}\|w_n-y_n\|=0$ and $\{w_n\}$ converges weakly to some $\bar{P}\in H$, then $\bar{P}\in VI(C,A_1)$.*

Proof This proof is the same as in Lemma 3.7 of [37], and we omit it.

Corollary 3.1 *Let H, C, F and A_1 be the same as the above statement. Choose $x_0,\ x_1 \in H$, $\mu,\ \mu' \in (0,1)$ with $\mu < \mu'$, $\delta,\ \delta' \in (0, \frac{2}{\mu'})$ with $\delta < \delta'$, $\alpha \in (0,1)$, $\tau \in (\frac{\delta'}{2}, \frac{1}{\mu'})$, $\gamma > 0$ and $\lambda_1 > 0$. Suppose that Assumptions 3.2 and 3.3 hold. Let $\{x_n\}$ be a sequence generated by*

$$\begin{cases} w_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \tau\lambda_n A_1(w_n)), \\ T_n = \{x \in H | \langle w_n - \tau\lambda_n A_1(w_n) - y_n, x - y_n\rangle \le 0\}, \\ z_n = P_{T_n}(w_n - \lambda_n\eta_n A_1(y_n)), \\ d_n = w_n - y_n - \tau\lambda_n(A_1(w_n) - A_1(y_n)), \\ x_{n+1} = z_n - \gamma\gamma_n F(z_n), \end{cases} \tag{3.45}$$

where $\{\alpha_n\}$ and $\{\eta_n\}$ are defined in Algorithm 3.1, and

$$\lambda_{n+1} = \begin{cases} \min\{\mu q_n \frac{\|w_n - y_n\|}{\|A_1(w_n) - A_1(y_n)\|},\ \lambda_n + p_n\}, & \text{if } \|A_1(w_n) - A_1(y_n)\| \neq 0, \\ \lambda_n + p_n, & \text{otherwise.} \end{cases}$$

Then the sequence $\{x_n\}$ converges strongly to the unique solution of BVIPs (3.1)-(3.2).

Proof Replacing Lemma 3.2 with Proposition 3.1 and taking $A = A_1$, $u_n = A_1(w_n)$ and $v_n = A_1(y_n)$ in Lemmas 3.1 and 3.3, and Theorem 3.1, the desired conclusion holds.

Remark 3.5 We compare Corollary 3.1 with corresponding results in recent literature.

• Corollary 3.1 improves Theorem 3.1 of [20] in the following aspects: (i) we require that the mapping A_1 satisfies the condition "for any $\{x_n\} \subset H$ with $x_n \rightharpoonup w^*$, one has $\|A_1 w^*\| \le \liminf_{n\to\infty} \|A_1 x_n\|$", which is strictly weaker than the sequentially weakly continuous assumption, see [36,37]; (ii) the step size $\tau\lambda_n$ is used to compute the y_n, which can improve the convergence efficiency of the algorithm, see [21]; (iii) Instead of just relaxing μ, we relax the parameters μ and δ by using the sequences $\{q_n\}$ and $\{q'_n\}$, respectively, to get the larger step sizes; (iv) the iterative scheme (3.45) owns the inertial term $w_n = x_n + \alpha_n(x_n - x_{n-1})$, which often accelerates prominently convergence speed of numerical methods, see [11,17,18,21–23].

• The mapping A_1 may be pseudomonotone and the iterative scheme (3.45) has inertial acceleration. Thus Corollary 3.1 improves Theorem 1 of [19].

• Compared with Theorem 3.1 of [18], Theorem 2.2 of [21] and Theorems 3.1 and 3.2 of [22], though A_1 is Lipschitz continuous, the step sizes of the iterative scheme (3.45) are updated at each iteration by a cheap computation without any linesearch procedure. Since we relax the parameters μ and δ by using the sequences $\{q_n\}$ and $\{q'_n\}$, respectively, the step sizes of the iterative scheme (3.45) are larger than ones in Algorithm 2.1 of [21] and Algorithm 3.1 of [23]. Especially, we only assume that the parameter $\gamma > 0$, other than $\gamma \in (0, \frac{2\beta}{L_2^2})$ as in [15–18,21–23]. Thus we do not need the prior information of the Lipschitz constant and strong monotonicity coefficient of F.

Let ρ' be a constant such that $\rho' \in (0,1)$ and $f:\ H \to H$ be a ρ'-contraction mapping. It is easy to know that $F(x) = x - f(x),\ \forall\, x \in H$ is $(1+\rho')$-Lipschitz continuous and $(1-\rho')$-strongly monotone. By Corollary 3.1, we can get the following corollary.

Corollary 3.2 *Let H, C and A_1 be the same as Corollary 3.1, and $f:\ H \to H$ be a ρ'-contraction mapping. Take $x_0,\ x_1 \in H$, $\mu,\ \mu' \in (0,1)$ with $\mu < \mu'$, $\delta,\ \delta' \in (0, \frac{2}{\mu'})$ with $\delta < \delta'$, $\alpha \in (0,1)$, $\tau \in (\frac{\delta'}{2}, \frac{1}{\mu'})$, $\gamma > 0$, $\rho' \in (0,1)$ and $\lambda_1 > 0$. Suppose that (i) and (ii) of Assumption 3.3 and Assumption 3.2 hold. Let $\{x_n\}$ be a sequence generated by*

$$\begin{cases} w_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \tau\lambda_n A_1(w_n)), \\ T_n = \{x \in H | \langle w_n - \tau\lambda_n A_1(w_n) - y_n, x - y_n\rangle \le 0\}, \\ z_n = P_{T_n}(w_n - \lambda_n\eta_n A_1(y_n)), \\ d_n = w_n - y_n - \tau\lambda_n(A_1(w_n) - A_1(y_n)), \\ x_{n+1} = (1 - \gamma\gamma_n)z_n + \gamma\gamma_n f(z_n), \end{cases} \tag{3.46}$$

where $\{\alpha_n\}$ *and* $\{\eta_n\}$ *are defined in Algorithm 3.1, and*

$$\lambda_{n+1} = \begin{cases} \min\{\mu q_n \frac{\|w_n - y_n\|}{\|A_1(w_n) - A_1(y_n)\|},\ \lambda_n + p_n\}, & \text{if } \|A_1(w_n) - A_1(y_n)\| \neq 0, \\ \lambda_n + p_n, & \text{otherwise.} \end{cases}$$

Then the sequence $\{x_n\}$ *converges strongly to* $\bar{P} \in VI(C, A_1)$ *and* $\bar{P} = P_{VI(C,A_1)}(f(\bar{P}))$.

Remark 3.6 We only require $\rho' \in (0, 1)$, other than $\rho' \in (0, \sqrt{5} - 2)$. In addition, the parameter γ can equal 1. Thus Corollary 3.1 improves Corollary 2.1 of [21].

Taking $F(x) = x - f(x)$ for all $x \in H$ in Theorem 3.1, we can get the following corollary.

Corollary 3.3 *Let* H, C *and* A *be the same as Theorem 3.1, and* $f : H \to H$ *be a* ρ'*-contraction mapping. Take* $x_0,\ x_1 \in H$, $\mu,\ \mu' \in (0, 1)$ *with* $\mu < \mu'$, $\delta,\ \delta' \in (0, \frac{2}{\mu'})$ *with* $\delta < \delta'$, $\alpha \in (0, 1)$, $\tau \in (\frac{\delta'}{2}, \frac{1}{\mu'})$, $\gamma > 0$, $\rho' \in (0, 1)$ *and* $\lambda_1 > 0$. *Suppose that (i)-(iii) of Assumption 3.1 and Assumption 3.2 hold. Let* $\{x_n\}$ *be a sequence generated by*

$$\begin{cases} w_n = x_n + \alpha_n(x_n - x_{n-1}),\ u_n \in A(w_n) \\ y_n = P_C(w_n - \tau\lambda_n u_n),\ v_n = P_{A(y_n)}(u_n) \\ T_n = \{x \in H | \langle w_n - \tau\lambda_n u_n - y_n, x - y_n\rangle \leq 0\}, \\ z_n = P_{T_n}(w_n - \lambda_n \eta_n v_n), \\ d_n = w_n - y_n - \tau\lambda_n(u_n - v_n), \\ x_{n+1} = (1 - \gamma\gamma_n)z_n + \gamma\gamma_n f(z_n), \end{cases} \tag{3.47}$$

where $\{\alpha_n\}$, λ_n *and* $\{\eta_n\}$ *are defined in Algorithm 3.1. Then the sequence* $\{x_n\}$ *converges strongly to* $\bar{P} \in VI(C, A)$ *and* $\bar{P} = P_{VI(C,A)}(f(\bar{P}))$.

Remark 3.7 Corollary 3.3 extends Theorem 3.9 of [12] in the following aspects.

(i) The iterative scheme (3.47) has inertial acceleration ($w_n = x_n + \alpha_n(x_n - x_{n-1})$), which has been widely uesd to increase the convergence rate of algorithms, see [11, 17, 18, 21–23].

(ii) The step sizes of the iterative scheme (3.47) are updated at each iteration without any prior information of Lipschitz constants, while the step sizes of the Algorithm 3.1 of [12] must satisfy $\lambda_n \in (0, \frac{1}{L})$, where L is a Lipschitz constant of the mapping A. In some cases, the value of the Lipschitz constant may be very large, which leads to a very small step size of the algorithm and reduces the convergence speed of Algorithm 3.1 of [12]. Furthermore, the Lipschitz constant is sometimes not obtainable in practical applications, see [32].

4 Numerical experiments

In this section, we report some numerical experiments to show the numerical behaviors of the proposed algorithm, namely Algorithm 3.1 (shortly, Alg3.1), and also compare them with several other well known algorithms including the Halpern projection algorithm 3.1 (HPM3.1) in [12], the Subgradient extragradient algorithm (SEA1) in [15], the projection and contraction methods 1 and 2 (PCM1, PCM2) in [20] and the modified subgradient extragradient algorithms 2.1 and 2.2 (MSE2.1, MSE2.2) in [21], the Relaxed forward-backward splitting method (RFBSM) in [40], the inertial Mann-type Tseng algorithm 3.3 (IMTT3.3) [41], and the new iterative proximal algorithm (NIPA) [42]. If the solution of the bilevel pseudomonotone multivalued variational inequality problem is unknown, then we use the function

$$D_n(x) = \|x_n - x_{n-1}\|^2$$

to measure the error of the n-th iteration. Otherwise, when the solution x^* of the bilevel pseudomonotone multivalued variational inequality problem is known, we use the function

$$E(x) = \|x - x^*\|^2$$

to show the efficiencies of aforementioned algorithms.

All the programs are written in MATLAB 2021a on an AMD Ryzen 5 3600 6-Core Processor (12 CPUs), 3.6Hz computer with RAM 16.00GB. We denote by "Iter." the number of iterations.

Example 4.1 Assume that the operator

$$F(x) := \ Mx + q,$$

where M is a symmetric and positive-definite matrix of size 5×5 and $q \in \Re^5$ with their entries generated randomly in $(-2, 2)$. It is clear that F is β- strongly monotone and L_2-Lipschitz continuous with $\beta = min\{eig(M)\}$, $L_2 = max\{eig(M)\}$. The feasible set is $C = \{x \in \Re^5 : 1 \le x_i \le 3, i = 1, 2, \cdots, 5\}$. Consider the following fractional programming problem

$$\min_{x \in C} f(x) = \frac{x^\top Qx + a^\top x + a_0}{b^\top x + b_0},$$

where $a = (1, 2, -1, -2, 1)^\top$, $b = (1, 0, -1, 0, 1)^\top$, $a_0 = -2$, $b_0 = 20$, and

$$Q = \begin{pmatrix} 5 & -1 & 2 & 0 & 2 \\ -1 & 6 & -1 & 3 & 0 \\ 2 & -1 & 3 & 0 & 1 \\ 0 & 3 & 0 & 5 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{pmatrix}$$

It is easy to verify that Q is symmetric and positive definite and f is pseudo-convex, see [19]. Setting

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

thus A is pseudo-monotone. The exact solution of our problem is $x^* = (1, 1, 1, 1, 1)^\top$. The starting point is $x_0 = x_1 = (2, 2, 2, 2, 2)^\top$. In Example 4.1, the parameters are taken as

- Alg3.1: $\alpha = 0.4$, $\epsilon_n = \frac{1}{(n+1)^2}$, $\tau = 0.8$, $\lambda_1 = 1$, $\delta = 1.2$, $\mu = 0.9$, $\mu' = 0.95$, $\delta' = 1.5$, $\gamma = 10^{-7}$, $\gamma_n = \frac{1}{n+1}$, $p_n = \frac{1}{(n+1)^{1.1}}$, $q_n = 1 + \frac{1}{n+1}$, $q'_n = 1 + \frac{1}{n+1}$;
- MSE2.1: $\theta = 0.4$, $\tau = 0.8$, $\delta = 1.5$, $\gamma = \frac{1.7\beta}{L_2^2}$, $\lambda_1 = 0.5$, $\mu = 0.1$, $\theta_n = \frac{1}{(n+1)^2}$, $\alpha_n = \frac{1}{n+1}$, $\xi_n = \frac{1}{(n+1)^{1.1}}$;
- MSE2.2: $\theta = 0.4$, $\tau = 0.8$, $\delta = 1.5$, $\gamma = \frac{1.7\beta}{L_2^2}$, $\mu = 0.1$, $\theta_n = \frac{1}{(n+1)^2}$, $\alpha_n = \frac{1}{n+1}$, $\xi_n = \frac{1}{(n+1)^{1.1}}$, $\sigma = 2$, $\ell = 0.5$;
- PCM1: $\lambda_1 = 1$, $\delta = 1.5$, $\mu = 0.9$, $\mu' = 0.95$, $\alpha_n = \frac{10}{n+10}$, $p_n = \frac{1}{(n+1)^{1.1}}$, $q_n = 1 + \frac{1}{n+1}$;
- PCM2: $\lambda_1 = 1$, $\delta = 1.5$, $\mu = 0.9$, $\mu' = 0.95$, $\alpha_n = \frac{10}{n+10}$, $p_n = \frac{1}{(n+1)^{1.1}}$, $q_n = 1 + \frac{1}{n+1}$;
- SEA1: $\gamma = \frac{1.7\beta}{L_2^2}$, $\mu = 0.1$, $\lambda_1 = 0.5$, $\alpha_n = \frac{1}{n+1}$.

The numerical results are shown in Fig. 1.

Example 4.2 Suppose $H = L^2([0, 1])$ is an infinite-dimensional Hilbert space with inner product

$$\langle x, y \rangle := \int_0^1 x(t)y(t)dt, \ \forall \ x, y \in H,$$

and the induced norm

$$\|x\| := (\int_0^1 |x(t)|^2 dt)^{\frac{1}{2}}, \ \forall \ x \in H.$$

Assume r and R are two positive real numbers such that $R/(k + 1) < r/k < r < R$ for some $k > 1$. Let $C = \{x \in H : \|x\| \le r\}$. The mapping $A : H \to H$ is defined by

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

Taking $R = 1.5, r = 1, k = 1.1$, we can verify that the operator A is pseudo-monotone rather than monotone (see [39], Section 4). Let $F : H \to H$ be defined by $(Fx)(t) = \frac{1}{2}x(t), t \in [0, 1]$. Thus the mapping F is $\frac{1}{2}$-strongly monotone and $\frac{1}{2}$-Lipschitz continuous. The parameters of all algorithms are the same as in Example 4.1 except $\gamma = 27$ of Alg3.1.

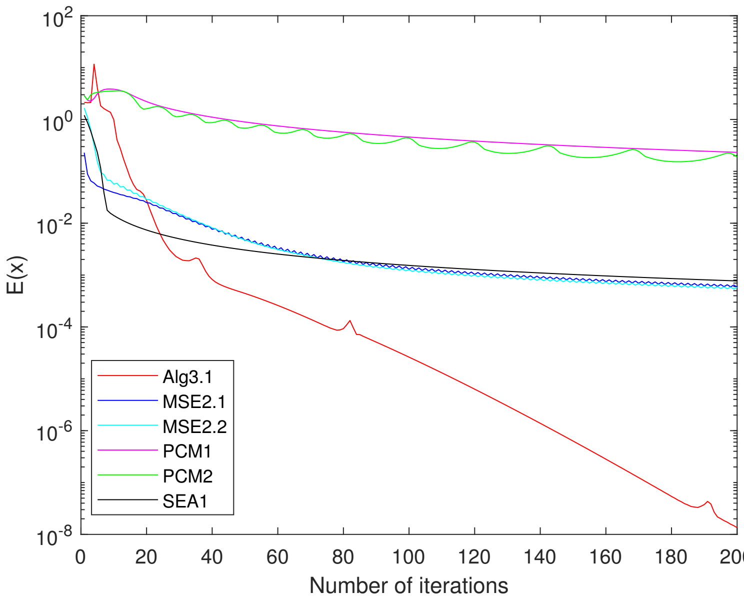
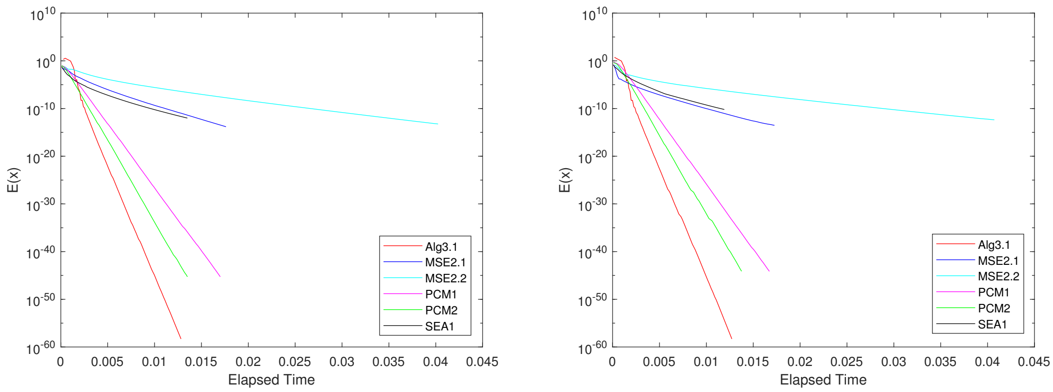


Fig. 1: Numerical results for Example 4.1



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

Fig. 2: Numerical results for Example 4.2

The solution of this problem is $x^*(t) = 0$. The maximum number of iterations 100 is used as a common stopping criterion. Fig. 2 shows the numerical behaviors of $E(x) = \|x - x^*\|^2$ of all algorithms with two starting points $x_0(t) = x_1(t)$. In Fig. 3, we give performances of Algorithm 3.1 for different values of γ in Example 4.2. This means that $\gamma > \frac{2\beta}{L_2^2}$ may be better than the case with $\gamma \in (0, \frac{2\beta}{L_2^2})$, where $\beta = L_2 = \frac{1}{2}$.

Example 4.3 Let $[0, 255]$ be the value range of each pixel, M and N be the width and height of image pixels, respectively, $D = M \times N$, $C = [0, 255]^D$, and $H = \Re^D$ be a Hilbert space with the standard Euclidean norm $\|\cdot\|_2$. If we do not consider the effects of noise, then the image deblurring problem is stated as follows:

$$find\ x \in C\ such\ that\ y = Kx, \tag{4.1}$$

where y is the observed image, K denotes the blurring matrix and x is the original image. The problem (4.1) can be seen as the convex minimization problem:

$$\min_{x \in C} h(x) = \frac{1}{2}\|Kx - y\|_2^2. \tag{4.2}$$

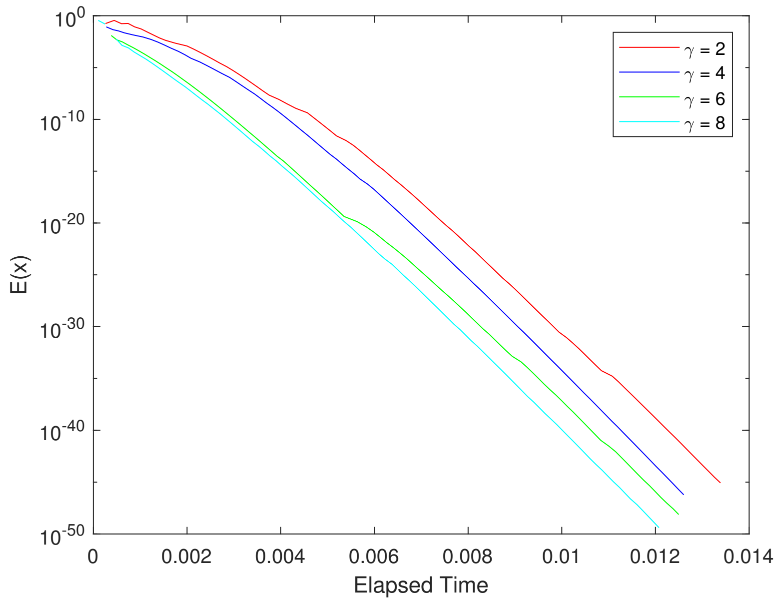


Fig. 3: The performances of Algorithm 3.1 for different values of γ in Example 4.2

Therefore, we can use Corollary 3.2 to solve above problem. Note that $A_1 = \nabla h(x)$. The signal to noise ration (SNR) in decibel(dB) is defined by

$$SNR = 10 \log_{10} \frac{\|\bar{x}\|_2^2}{\|x - \bar{x}\|_2^2},$$

where $\bar{x}$ denotes the original image and x denotes the recovered image. SNR can measure the efficiency of different algorithms to restore the image. The larger the SNR value, the better is the image restoration effectiveness. Let $x_0 = 1 \in \Re^D$ and $x_1 = 0 \in \Re^D$. We consider the Lena (512×512) as test image and use the blurring function of motion blur (”fspecial(’motion’, 45,180)”) from Matlab.

Then Table 1 reports numerical results. Fig. 4 gives the original image, blurred image and recovered images by using the methods Alg3.1, IMTT3.3, RFBSM, SEA1 and PCM1. Fig. 5 shows the SNR values of images recovered by the methods Alg3.1, IMTT3.3, RFBSM, SEA1 and PCM1, respectively. The parameters of algorithms are the same as in Example 4.1 except

- Alg3.1: $\tau = 0.9$, $\gamma = 0.3$, $f(x) = 0.5x$;
- IMTT3.3: $\theta = 0.5$, $\mu = 0.9$, $\epsilon_n = \frac{100}{(n+1)^2}$, $\lambda_1 = 1$, $\alpha_n = \frac{1}{n+1}$, $\beta_n = 0.5(1 - \alpha_n)$, $f(x) = 0.5x$;
- RFBSM: $\alpha = 0.1$, $\theta = 0.5$, $\lambda_1 = 1$ and $\mu = 0.7$.

Table 1: Numerical comparison for the methods Alg3.1, IMTT3.3, RFBSM, SEA1 and PCM1

Image	SNR(dB)					
	Iter.	Alg3.1	IMTT3.3	RFBSM	SEA1	PCM1
Lena	2500	28.1640	23.0983	23.3417	18.7880	25.5059
	3000	28.4122	23.5398	23.7863	19.1445	25.8858
Size = 512 × 512	4000	28.7557	24.2262	24.4735	19.7243	26.4382
	5000	28.9807	24.7401	24.9829	20.1905	26.8475

Example 4.4 Let N be a matrix of order m, B be an $m \times m$ skew-symmetric matrix, Q be an $m \times m$ diagonal matrix and

$$A(x) = \{f(t)Mx : \; t \in [0, 1]\}, F(x) = 0.1x, \; \forall x \in \Re^m,$$

(a) original image

(b) image motion blur

(c) the method (Alg3.1)

(d) the method (IMTT3.3)

(f)the method (RFBSM)

(f)the method (SEA1)

(f)the method (PCM1)

Fig. 4: Comparison of Lena images recovered by using different algorithms

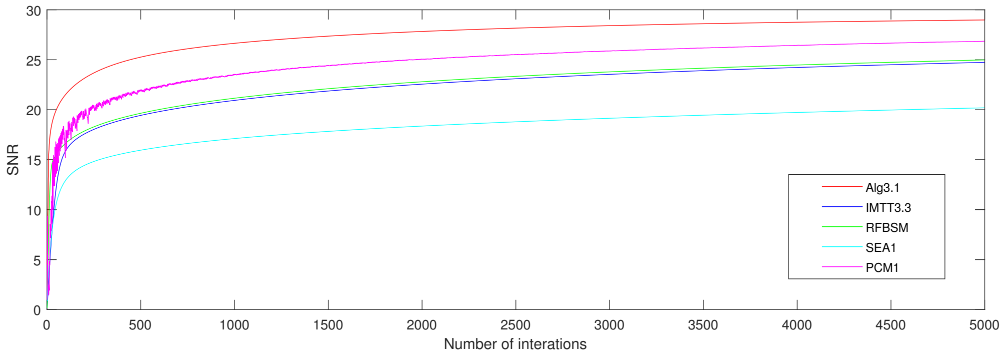


Fig. 5: Graphs of SNR for the methods Alg3.1, IMTT3.3, RFBSM, SEA1 and PCM1 of Lena image

where $M = N * N^\top + B + Q$ and

$$f(t) = 3t^2 - 2t + 1, \; t \in [0, 1].$$

Example 4.4 has been considered in [12]. In this test, the matrices N, B, Q are randomly generated by using commands $N = rand(m, m) * 2 * m - m$, $B = skewdec(m, 1), Q = diag(1 : m)$. Take $C = \{x \in \Re^m : \|x\| \leq 2\}$ and $m = 100$. The numerical behaviors of $D_n(x) = \|x_n - x_{n-1}\|^2$ of all algorithms with initial point $x_0 = x_1 = rand(m, 1)$ are shown in Fig. 6. In Example 4.4, the parameters are the same as in Example 4.1 except

- Alg3.1: $\gamma = 27$;
- HPM3.1: $\gamma = 1.5$, $\lambda_1 = \frac{1}{8\|M\|+5}$, $\alpha_n = \frac{1}{3n+2000}$;
- NIPA: $\alpha = 0.5$, $\epsilon_n = \frac{1}{(n+1)(n+2)}$, $\lambda_1 = \frac{1}{8\|M\|+5}$.

5 Conclusions

In the paper, we present an adaptive inertial algorithm to approximate the solution of bilevel variational inequalities with multivalued pseudomonotone operators in real Hilbert spaces. The features of the proposed algorithm are that (1) does not need to know any prior information of the Lipschitz constants and

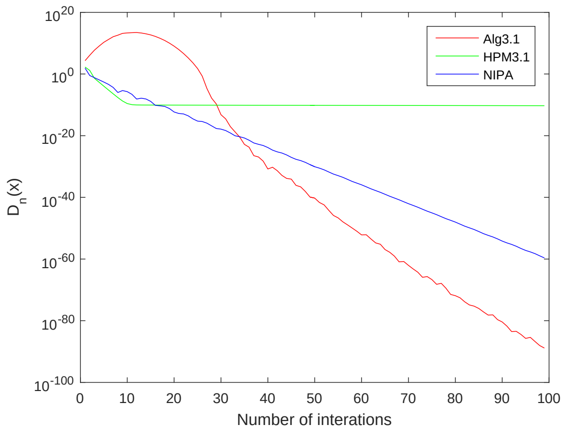


Fig. 6: Numerical results for Example 4.4

strong monotonicity coefficients of the associated mappings; (2) requires only one projection per iteration except computing v_n; (3) the operator involved is pseudomonotone and Lipschitz continuous; (4) its step sizes are updated at each iteration by a cheap computation without any linesearch procedure; (5) the embedding of the inertial terms speeds up the convergence rate of the algorithm. We prove the strong convergence of the proposed algorithm under mild conditions, and give several numerical examples to show the proposed algorithms have competitive advantages in comparison with the known methods in the literature. The results obtained in this paper improve and extend some corresponding ones in [15–23].

Declaration of Competing Interest

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

References

1. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, Berlin (2003)
2. Kinderlehrer, D., Stampacchia, G.: An Introduction to Variational Inequalities and Their Applications. Academic Press, New York (1980)
3. Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Springer, Berlin (2000)
4. Konnov, I.V.: Equilibrium Models and Variational Inequalities. Elsevier, Amsterdam (2007)
5. Wang, Z.B., Chen, Z.Y., Xiao, X.B., Zhang, C.: A new projection-type method for solving multi-valued mixed variational inequalities without monotonicity. Appl. Anal. **99**(9), 1453-1466 (2020)
6. Li, F.L., He, Y.R.: An algorithm for generalized variational inequality with pseudomonotone mapping. J. Comput. Appl. Math. **228,** 212-218 (2009)
7. Fang, C.J., Chen, S.L.: A subgradient extragradient algorithm for solving multi-valued variational inequality. Appl. Math. Comput. **229,** 123-130 (2014)
8. Ye, M.L.: A cutting hyperplane projection method for solving generalized quasivariational inequalities. J. Oper. Res. Soc. China **4,** 483-501 (2016)
9. Dong, Q.L., Lu, Y.Y., Yang, J., et al.: Approximately solving multi-valued variational inequalities by using a projection and contraction algorithm. Numer. Algor. **76,** 799-812 (2017)
10. Ye, M.L.: An improved projection method for solving generalized variational inequality problems. Optimization **67,** 1523-1533 (2018)
11. He, X., Huang, N.J., Li, X.S.: Modified projection methods for solving multi-valued variational inequality without monotonicity. Netw. Spat. Econ. **22,** 361-377 (2022)
12. Anh, P.N., Thang, T.V., Thach, H.T.C.: Halpern projection methods for solving pseudomonotone multivalued variational inequalities in Hilbert spaces. Numer. Algor. **87,** 335-363 (2021)
13. Anh, P.N., Kim, J.K., Muu, L.D.: An extragradient algorithm for solving bilevel pseudomonotone variational inequalities. J. Glob. Optim. **52,** 627-639 (2012)
14. Anh, T.V., Muu, L.D., Anh, T.V.: A projection-fixed point method for a class of bilevel variational inequalities with split fixed point constraints. Optimization **65**(6), 1229-1243 (2016)
15. Thong, V.T., Triet, N.A., Li, X.H., Dong, Q.L.: Strong convergence of extragradient methods for solving bilevel pseudomonotone variational inequality problems. Numer. Algor. **83**(3), 1123-1143 (2020)

16. Anh, T.V.: A strongly convergent subgradient extragradient-halpern method for solving a class of bilevel pseudomonotone variational inequalities. Vietnam J. Math. **45**(3), 317-332 (2017)
17. Tan, B., Liu, L.Y., Qin, X.Y.: Self adaptive inertial extragradient algorithms for solving bilevel pseudomonotone variational inequality problems. Jpn. J. Ind. Appl. Math. **38,** 519-543 (2021)
18. Tan, B., Qin, X.Y., Yao,J.C.: Two modified inertial projection algorithms for bilevel pseudomonotone variational inequalities with applications to optimal control problems. Numer. Algor. **88,** 1757-1786 (2021)
19. Hieu, D.V., Moudafi, A.: Regularization projection method for solving bilevel variational inequality problem. Optim. Lett. **15**(1), 205-229 (2021)
20. Yang, J.: Projection and contraction methods for solving bilevel pseudomonotone variational inequalities. Acta Appl. Math. **177,** 7 (2022)
21. Tan, B., Cho, S.Y.: Two adaptive modified subgradient extragradient methods for bilevel pseudomonotone variational inequalities with applications. Commun. Nonlinear Sci. Numer. Simul. **107,** 106160 (2022)
22. Tan, B., Cho, S.Y.: Two projection-based methods for bilevel pseudomonotone variational inequalities involving non-Lipschitz operators. Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **116,** 64 (2022)
23. Tan, B., Li, S., Qin, X.: An accelerated extragradient algorithm for bilevel pseudomonotone variational inequality problems with application to optimal control problems. Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat. **115,** 174 (2021)
24. Korpelevich, G.M.: The extragradient method for finding saddle points and other problem. Ekon. Mat. Metody 12, 747-756 (1976)
25. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35,** 69-76 (1997)
26. He, B.S., Liao, L.Z.: Improvements of some projection methods for monotone nonlinear variational inequalities. J. Optim. Theory Appl. **112,** 111-128 (2002)
27. Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalbities in Hilbert space. J. Optim. Theory Appl. **148,** 318-335 (2011)
28. Dong, Q.L., Jiang, D., Gibali, A.: A modified subgradient extragradient method for solving the variational inequality problem. Numer. Algor. **49,** 927-940 (2018)
29. Thong, D.V., Vuong, P.T.: Improved subgradient extragradient methods for solving pseudomonotone variational inequalities in Hilbert spaces. Appl. Nume. Math. **163,** 221-238 (2021)
30. Attouch, H., Goudon, X., Redont, P.: The heavy ball with friction. I. The continuous dynamical system. Commun. Contemp. Math. **2,** 1-34 (2000)
31. Attouch, H., Czarnecki, M.O.: Asymptotic control and stabilization of nonlinear oscillators with nonisolated equilibria. J. Differ. Equ. **179,** 278-310 (2002)
32. Hieu, D.V., Anh, P.K., Muu, L.D.: Modified extragradient-like algorithms with new stepsizes for variational inequalities. Comput. Optim. Appl. **73,** 913-932 (2019)
33. Goebel, K., Reich, S.: Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings. Marcel Dekker, New York (1984)
34. Zarantonello, E.H.: Projections on convex sets in Hilbert space and spectral theory. In: Zarantonello, E.H. (ed.) Contributions to Nonlinear Functional Analysis. Academic Press, New York (1971)
35. Saejung, S., Yotkaew, P.: Appoximation of zeros of inverse strongly monotone operators in Banach spaces. Nonlinear Anal. **75,** 742-750 (2012)
36. Yao, Y., Iyiola, O.S. Shehu, Y.: Subgradient extragradient method with double inertial steps for variational inequalities. J. Sci. Comput. **90,** 71 (2022)
37. Thong, D.V., Yang, J., Cho, Y.J., Rassias, T.M.: Explicit extragradient-like method with adaptive stepsizes for pseudomonotone variational inequalities. Optim. Lett. **15,** 2181-2199 (2021)
38. Liu, H.W., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput. Optim. Appl. **77,** 491-508 (2020)
39. Hieu, D.V., Cho, Y.J., Xiao, Y.B., Kumam, P.: Modified extragradient method for pseudomonotone variational inequalities in infinite dimensional Hilbert spaces. Vietnam J. Math. **49**, 1165-1183 (2021)
40. Cholamjiak, P., Van Hieu, D., Cho, Y.J.: Relaxed forward-backward splitting methods for solving variational inclusions and applications. J. Sci. Comput. **88**, 85 (2021)
41. Bing, T., Cho, S.Y.: Strong convergence of inertial forward-backward methods for solving monotone inclusions. Applicable Analysis **101,** 5386 - 5414 (2021)
42. Thang, T.V., Hien, N.D., Thach, H.T., Anh, P.N.: Weak convergence of inertial proximal algorithms with self adaptive stepsize for solving multivalued variational inequalities. Optimization, https://doi.org/10.1080/02331934.2022.2135966 (2022)