

# Tseng Splitting Method with Double Inertial Steps for Solving Monotone Inclusion Problems

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

· Zhang-you Chen

Received: date / Accepted: date

## Abstract

In this paper, based on double inertial extrapolation steps strategy and relaxed techniques, we introduce a new Tseng splitting method with double inertial extrapolation steps and self-adaptive step sizes for solving monotone inclusion problems in real Hilbert spaces. Under mild and standard assumptions, we establish successively the weak convergence, nonasymptotic  $O(\frac{1}{\sqrt{n}})$  convergence rate, strong convergence and linear convergence rate of the proposed algorithm. Finally, several numerical experiments are provided to illustrate the performance and theoretical outcomes of our algorithm.

**Keywords** Monotone inclusion problem; Tseng splitting method; Double inertial extrapolation steps; Strong and weak convergence; Linear convergence rate

## 1 Introduction

---

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

Zhang-you Chen

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

zhangyouchen@swjtu.edu.cn

Let  $H$  be a real Hilbert space with the inner product  $\langle \cdot, \cdot \rangle$  and the induced norm  $\|\cdot\|$ . The monotone inclusion problem (MIP) is as follows

$$\text{find } x^* \in H \text{ such that } 0 \in (A + B)x^*. \quad (1.1)$$

where  $A : H \rightarrow H$  is a single mapping and  $B : H \rightarrow 2^H$  is a multivalued mapping. The solution set is denoted by  $\Omega := (A + B)^{-1}(0)$ .

The monotone inclusion problem has drawn much attention because it provides a broad unifying frame for variational inequalities, convex minimization problems, split feasibility problems and equilibrium problems, and has been applied to solve several real-world problems from machine learning, signal processing and image restoration, see [1, 3, 4, 7, 10, 19, 24].

One of famous methods for solving MIP (1.1) is forward-backward splitting method, which was introduced by Passty [18] and Lions et al. [13]. This method generates an iterative sequence  $\{x_n\}$  in following way

$$x_{n+1} = (I + \lambda_n B)^{-1}(I - \lambda_n A)x_n. \quad (1.2)$$

where the mapping  $A$  is  $\frac{1}{L}$ -co-coercive,  $B$  is maximal monotone,  $I$  is an identity mapping on  $H$  and  $\lambda_n > 0$ . The operator  $(I - \lambda_n A)$  is called an forward operator and  $(I + \lambda_n B)^{-1}$  is said to be a backward operator.

Tseng [25] proposed a modified forward-backward splitting method (also known as Tseng splitting algorithm), whose iterative formula is as follows

$$\begin{cases} y_n = (I + \lambda_n B)^{-1}(I - \lambda_n A)x_n \\ x_{n+1} = y_n - \lambda_n(Ay_n - Ax_n), \end{cases} \quad (1.3)$$

where  $A$  is  $L$ -lipschitz continuous and  $\{\lambda_n\} \subset (0, 1/L)$ . However, the Lipschitz constant of an operator is often unknown or difficult to estimate in nonlinear problems. To overcome this drawback, Cholamjiak et al. [5] introduce a relaxed forward-backward splitting method, which uses a simple step-size rule without the prior knowledge of Lipschitz constant of the operator, for solving MIP (1.1) and prove the linear convergence rate of the proposed algorithm.

In recent year, the inertial method was introduced in [2], which can be regarded as a procedure of speeding up the convergence rate of algorithms. Many researchers utilize inertial methods to design algorithm for solving monotone inclusion problems and variational inequalities, see, for example, [1, 4, 5, 6, 8, 9, 12, 17, 21, 23, 29]. To enhance the numerical efficiency, Çopur et al. [8] introduce firstly the double inertial extrapolation steps for solving quasi-variational inequalities in real Hilbert spaces. Combining relaxation techniques with the inertial methods, Cholamjiak et al. [5] modify Tseng splitting method to solve MIP (1.1) in real Hilbert spaces. Very recently, incorporating double inertial extrapolation steps and relaxation techniques, Yao et al. [29] present a novel subgradients extragradient method to solve variational inequalities, and prove its strong convergence, weak convergence and linear convergence, respectively. However, the linear convergence of [29] is obtained under a single inertia rather than double inertias.

This paper devotes to further modifying Tseng splitting method for solving MIP (1.1) in real Hilbert spaces. We obtain successively the weak convergence, nonasymptotic  $O(\frac{1}{\sqrt{n}})$  convergence rate, strong convergence and linear convergence rate of the proposed algorithm. Our results obtained in this paper improve the corresponding results in [1, 5, 6, 26, 29] as follows:

- Combining double inertial extrapolation steps strategy and relaxed techniques, we propose a new Tseng splitting method, which include the corresponding methods considered in [1, 5, 29] as special cases. The two inertial factors in our algorithm are variable sequences, different from the constant inertial factor in [1, 5, 6, 29]. Especially, when our algorithm is applied to solving variational inequalities, its some parameters have larger choosing interval than the ones of [29]. In addition, one of our inertial factors can be equal to 1, which is not allowed in the single inertial methods [5, 6], which require that the inertial factor must be strictly less than 1.
- We prove the strong convergence, nonasymptotic  $O(\frac{1}{\sqrt{n}})$  convergence rate and linear convergence rate of the proposed algorithm. Note that that the strong convergence does not require to know the

modulus of strong monotonicity and the Lipschitz constant in advance. As far as we know, there is no convergence rate results in the literature for methods with double inertial extrapolation steps for solving MIP (1.1) in infinite-dimensional Hilbert spaces.

- Our algorithm use double inertial extrapolation steps to accelerate the speed of the algorithm. The step sizes of our algorithm are updated by a simple calculation without knowing the Lipschitz constant of the underlying operator. Some numerical experiments show that our algorithm has better efficiency than the corresponding algorithms in [1, 5, 6, 26, 29].

The structure of this article is as follows. In section 2, we recall some essential definitions and results which is relate to this paper. In section 3, we present our algorithm and analyze its weak convergence. In section 4, we establish the strong convergence and the linear convergence rate of our method. In section 5, we present some numerical experiments to demonstrate the performance of our algorithm. We give some concluding remarks in section 6.

## 2 Preliminaries

In this section, we first give some definitions and results that will be used in this paper. The weak convergence and strong convergence of sequences are denoted by  $\rightharpoonup$  and  $\rightarrow$ , respectively.

**Definition 2.1** The mapping  $A : H \rightarrow H$  is called

- (i) pseudomonotone on  $H$  if  $\langle Ax, y - x \rangle \geq 0$  implies that  $\langle Ay, y - x \rangle \geq 0$ ,  $\forall x, y \in H$ ;
- (ii) monotone on  $H$  if

$$\langle Ax - Ay, x - y \rangle \geq 0, \quad \forall x, y \in H;$$

- (iii)  $\mu$ -strongly monotone on  $H$  if there exists a positive constant  $\mu > 0$

$$\langle Ax - Ay, x - y \rangle \geq \mu \|x - y\|^2, \quad \forall x, y \in H;$$

- (iv)  $L$ -lipschitz continuous on  $H$  if there exists a scalar  $L > 0$  satisfying

$$\|Ax - Ay\| \leq L \|x - y\|, \quad \forall x, y \in H;$$

- (v)  $r$ -strongly pseudomonotone on  $H$  if there exists a positive constant  $r > 0$  such that

$$\langle Ay, x - y \rangle \geq 0 \text{ implies that } \langle Ax, x - y \rangle \geq r \|x - y\|^2, \quad \forall x, y \in H.$$

**Definition 2.2** The graph of  $A$  is the set in  $H \times H$  defined by

$$\text{Graph}(A) := \{(x, u) : x \in H, u \in Ax\}.$$

Let  $C \subset H$  be a nonempty, closed and convex set. The normal cone  $N_C(x)$  of  $C$  at  $x$  is represented by

$$N_C(x) := \begin{cases} \{z \in H : \langle z, y - x \rangle \leq 0, \forall y \in C\}, & \text{if } x \in C, \\ \emptyset, & \text{otherwise.} \end{cases}$$

The projection of  $x \in H$  onto  $C$ , denoted by  $P_C(x)$ , is defined as

$$P_C(x) := \arg \min_{y \in C} \|x - y\|.$$

and is has  $\langle x - P_C(x), y - P_C(x) \rangle \leq 0, \forall y \in C$ .

The sequence  $\{u_n\}$  is  $Q$ -linear convergence if there is  $q \in (0, 1)$  such that  $\|u^{k+1} - u\| \leq q \|u^k - u\|$  for all  $k$  large enough.

**Definition 2.3** The set-valued mapping  $A : H \rightarrow 2^H$  is called

- (i) monotone on  $H$  if for all  $x, y \in H$ ,  $u \in Ax$  and  $v \in Ay$  implies that

$$\langle u - v, x - y \rangle \geq 0;$$

- (ii) maximal monotone on  $H$  if it is monotone and for any  $(x, u) \in H \times H$ ,  $\langle u - v, x - y \rangle \geq 0$  for every  $(y, v) \in \text{Graph}(A)$  implies that  $u \in Ax$ ;
- (iii)  $\mu$ -strongly monotone on  $H$  if for all  $x, y \in H$ ,  $u \in Ax$  and  $v \in Ay$  implies that

$$\langle u - v, x - y \rangle \geq \mu \|x - y\|^2, \quad \forall x, y \in H.$$

**Lemma 2.1** ([15]) Let  $\{\varphi_n\}, \{\delta_n\}$  and  $\{\alpha_n\}$  be sequences in  $[0, +\infty)$  such that

$$\varphi_{n+1} \leq \varphi_n + \alpha_n(\varphi_n - \varphi_{n-1}) + \delta_n, \quad \forall n \geq 1, \quad \sum_{n=1}^{\infty} \delta_n < +\infty$$

and there exists a real number  $\alpha$  with  $0 \leq \alpha_n \leq \alpha < 1$  for all  $n \in N$ . Then the following hold :

- (i)  $\sum_{n=1}^{+\infty} [\varphi_n - \varphi_{n-1}]_+ < +\infty$  where  $[t]_+ := \max\{t, 0\}$ ;
- (ii) there exists  $\varphi^* \in [0, +\infty)$  such that  $\lim_{n \rightarrow \infty} \varphi_n = \varphi^*$ .

**Lemma 2.2** ([16]) Let  $C$  be a nonempty set of  $H$  and  $\{x_n\}$  be a sequence in  $H$  such that the following two conditions hold:

- (i) for every  $x \in C$ ,  $\lim_{n \rightarrow \infty} \|x_n - x\|$  exists;
- (ii) every sequential weak cluster point of  $\{x_n\}$  is in  $C$ .

Then  $\{x_n\}$  converges weakly to a point in  $C$ .

**Lemma 2.3** [22] Let  $A : H \rightarrow H$  be a maximal monotone mapping and  $B : H \rightarrow 2^H$  be a Lipschitz continuous and monotone mapping. Then the mapping  $A + B$  is a maximal monotone mapping.

**Lemma 2.4** [14] Let  $\{a_n\}$  and  $\{b_n\}$  be nonnegative real numbers sequences for which there exists  $0 \leq q < 1$ , so that

$$a_{n+1} \leq qa_n + b_n \quad \text{for every } n \in N.$$

If  $\lim_{n \rightarrow \infty} b_n = 0$ , then  $\lim_{n \rightarrow \infty} a_n = 0$ .

**Lemma 2.5** [28] Let  $\{\alpha_n\}, \{a_n\}, \{b_n\}$  and  $\{c_n\}$  be nonnegative real numbers sequences and there exists  $n_0 \in N$  such that

$$a_{n+1} \leq (1 - \alpha_n)a_n + \alpha_n b_n + c_n \quad \forall n \geq 1,$$

where  $\{\alpha_n\}, \{b_n\}$  and  $\{c_n\}$  satisfy the following conditions

- (i)  $\{\alpha_n\} \subset (0, 1)$  and  $\sum_{n=0}^{\infty} \alpha_n = \infty$ ;
- (ii)  $\limsup_{n \rightarrow \infty} b_n \leq 0$ ;
- (iii)  $c_n \geq 0$ ,  $\forall n \geq 0$ ,  $\sum_{n=1}^{\infty} c_n < \infty$ .

Then  $\lim_{n \rightarrow \infty} a_n = 0$ .

**Lemma 2.6** [11] Let  $B : H \rightarrow 2^H$  be a set-valued maximal monotone mapping and  $A : H \rightarrow H$  is a mapping. Define  $T_\lambda := (I + \lambda B)^{-1}(I - \lambda A)$ ,  $\lambda > 0$ . Then  $\text{Fix}(T_\lambda) = (A + B)^{-1}0$ .

**Lemma 2.7** ([3], Corollary 2.14) For all  $x, y \in H$  and  $\alpha \in R$ , the following equality holds:

$$\|\alpha x + (1 - \alpha)y\|^2 = \alpha\|x\|^2 + (1 - \alpha)\|y\|^2 - \alpha(1 - \alpha)\|x - y\|^2.$$

### 3 Weak convergence

In this section, we introduce the Tseng splitting method with double inertial steps to solve MIP (1.1) and discuss convergence and convergence rate of the new algorithm. We firstly give the following conditions.

- (C<sub>1</sub>) The solution set of the inclusion problem (1.1) is nonempty, that is,  $\Omega \neq \emptyset$ .
- (C<sub>2</sub>) The mappings  $A : H \rightarrow H$  is  $L$ -Lipschitz continuous and monotone and the set-valued mapping  $B : H \rightarrow 2^H$  is maximal monotone.
- (C<sub>3</sub>) The real sequences  $\{\alpha_n\}$ ,  $\{\beta_n\}$ ,  $\{\beta_n\}$ ,  $\{a_n\}$ ,  $\{p_n\}$  and  $\{\mu_n\}$  satisfy the following conditions
  - (i)  $0 \leq \alpha_n \leq 1$ ;
  - (ii)  $0 \leq \beta_n \leq \beta_{n+1} \leq \beta < \frac{3+2\varepsilon-\sqrt{8\varepsilon+17}}{2\varepsilon}$ ,  $\varepsilon \in (1, +\infty)$ ;
  - (iii)  $0 < \theta < \theta_n \leq \theta_{n+1} \leq \frac{1}{1+\varepsilon}$ ,  $\varepsilon \in (1, +\infty)$ ;
  - (iv)  $a_n = (1 - \theta_n)\beta_n + \theta_n\alpha_n$  is a non-decreasing sequence;
  - (v)  $\sum_{n=1}^{\infty} p_n < \infty$  and  $\lim_{n \rightarrow \infty} \mu_n = 0$ .

*Remark 3.1* Note that if  $\{\alpha_n\}$  is a non-decreasing sequence,  $\tilde{\delta} \geq 0$  and  $\beta_n = \tilde{\delta} \leq \alpha_1$ , then the condition (C<sub>3</sub>)(iv) holds naturally. In addition, if we choose  $\alpha_n = \frac{1}{5} + \frac{1}{6+n}$ ,  $\beta_n = \frac{1}{6} - \frac{1}{6+n}$  and  $\theta_n = \frac{1}{4} - \frac{1}{6+n}$ , then the condition (C<sub>3</sub>)(iv) is also true.

**Algorithm 3.1** Choose  $x_0, x_1 \in H$ ,  $\mu \in (0, 1)$  and  $\lambda_1 > 0$

Step 1. Compute

$$\begin{aligned} w_n &= x_n + \alpha_n(x_n - x_{n-1}) \\ z_n &= x_n + \beta_n(x_n - x_{n-1}) \\ y_n &= (I + \lambda_n B)^{-1}(I - \lambda_n A)w_n \end{aligned}$$

where

$$\lambda_{n+1} = \begin{cases} \min\left\{\frac{(\mu_n+\mu)\|w_n-y_n\|}{\|Aw_n-Ay_n\|}, \lambda_n + p_n\right\}, & Aw_n \neq Ay_n \\ \lambda_n + p_n, & \text{otherwise.} \end{cases}$$

If  $w_n = y_n$ , stop and  $y_n$  is a solution of the problem (1.1). Otherwise

Step 2. Compute

$$x_{n+1} = (1 - \theta_n)z_n + \theta_n(y_n - \lambda_n(Ay_n - Aw_n))$$

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

*Remark 3.2* (i) In our Algorithm 3.1, we can take the inertial factor  $\alpha_n = 1$ . This is not allowed in the corresponding algorithms of [5, 6], where the only single inertial extrapolation step is considered and the inertia is bounded away from 1.

- (ii) In order to get the larger step sizes, being similar to the sequence  $\{\theta_n\}$  in Algorithm 3.1 of [27], the sequence  $\{\mu_n\}$  is used to relax the parameter  $\mu$ . The sequence  $\{\theta_n\}$  can be called a relaxed parameter sequence, which can often improve numerical efficiency of algorithms, see [5]. If  $\mu_n = 0$ , then the step size  $\lambda_n$  is the same as the one of Algorithm 4.1 in [6]. If  $\mu_n = 0$  and  $p_n = 0$ , then the step size  $\lambda_n$  is the same as the one of Algorithm 3.1 of [26].
- (iii) Note that if  $\beta_n = 0$ , then the condition (C<sub>3</sub>)(iii) can be relaxed as  $0 < \theta < \theta_n \leq \theta_{n+1} \leq \frac{1}{1+\varepsilon}$ ,  $\varepsilon \in [0, +\infty)$ , which indicates that  $\theta_n = \hat{\theta}$  can equal 1. Setting  $\alpha_n = \alpha$ ,  $\theta_n = \hat{\theta}$ ,  $\mu_n = 0$  and  $\beta_n = p_n = 0$ , our algorithm can reduce to Algorithm 2 of [5]. In addition, if  $\alpha_n = \theta_n = \alpha$ ,  $\mu_n = 0$  and  $\beta_n = p_n = 0$ , Algorithm 3.1 can reduce to Algorithm 1 of [1].

**Lemma 3.1** The sequences  $\{\lambda_n\}$  from Algorithm 3.1 is bounded and  $\lambda_n \in [\min\{\frac{\mu}{L}, \lambda_1\}, \lambda_1 + P]$ . Furthermore there exists  $\lambda \in [\min\{\frac{\mu}{L}, \lambda_1\}, \lambda_1 + P]$  such that  $\lim_{n \rightarrow \infty} \lambda_n = \lambda$ , where  $P = \sum_{n=1}^{\infty} p_n$ .

*Proof* By the definition of  $\lambda_n$ , if  $Aw_n \neq Ay_n$ , we get

$$\lambda_n \geq \frac{(\mu+\mu_n)\|w_n-y_n\|}{\|Aw_n-Ay_n\|} \geq \frac{\mu+\mu_n}{L} \geq \frac{\mu}{L}. \quad (3.1)$$

Since  $P = \sum_{n=1}^{\infty} p_n$ , we have

$$\lambda_{n+1} \leq \lambda_n + p_n \leq \lambda_1 + \sum_{n=1}^{\infty} p_n = \lambda_1 + P. \quad (3.2)$$

It implies that  $\{\frac{\mu}{L}, \lambda_1\} \leq \lambda_n \leq \lambda_1 + P$ .

We have

$$\lambda_{n+1} - \lambda_n = [\lambda_{n+1} - \lambda_n]_+ - [\lambda_{n+1} - \lambda_n]_-. \quad (3.3)$$

Thus

$$\lambda_{n+1} - \lambda_1 = \sum_{i=1}^n [\lambda_{i+1} - \lambda_i]_+ - \sum_{i=1}^n [\lambda_{i+1} - \lambda_i]_-. \quad (3.4)$$

Since  $\{\lambda_n\}$  is bounded and  $\sum_{n=1}^{\infty} [\lambda_{n+1} - \lambda_n]_+ \leq \sum_{n=1}^{\infty} p_n < \infty$ , we get  $\sum_{n=1}^{\infty} [\lambda_{n+1} - \lambda_n]_-$  is convergent. Therefore, there exists  $\lambda \in [\min\{\frac{\mu}{L}, \lambda_1\}, \lambda_1 + P]$  such that  $\lim_{n \rightarrow \infty} \lambda_n = \lambda$ . This proof is completed.

*Remark 3.3* If  $w_n = y_n$ , then  $y_n = (I + \lambda_n B)^{-1}(I - \lambda_n A)y_n$ . By Lemma 2.6, we know  $y_n \in \Omega$ .

**Lemma 3.2** Suppose that the sequence  $\{y_n\}$  is generated by Algorithm 3.1. Thus the following assertions hold:

(i) if the conditions (C<sub>1</sub>) and (C<sub>2</sub>) hold, then

$$\|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 \leq \|w_n - p\|^2 - (1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2, \forall p \in \Omega;$$

(ii) if the conditions (C<sub>1</sub>) and (C<sub>2</sub>) hold and A or B is r-strongly monotone, then

$$\begin{aligned} \|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 &\leq \|w_n - p\|^2 - (1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \\ &\quad - 2r\lambda_n\|y_n - p\|^2, \forall p \in \Omega. \end{aligned}$$

*Proof* (i) According to the definition of  $\lambda_n$ , we have that

$$\begin{aligned} \|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 &= \|y_n - p\|^2 + \lambda_n^2 \|Ay_n - Aw_n\|^2 - 2\lambda_n \langle Ay_n - Aw_n, y_n - p \rangle \\ &\leq \|y_n - w_n\|^2 + \|w_n - p\|^2 + 2 \langle y_n - w_n, w_n - p \rangle \\ &\quad + \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} \|y_n - w_n\|^2 - 2\lambda_n \langle Ay_n - Aw_n, y_n - p \rangle \\ &= (1 + \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2 + \|w_n - p\|^2 + 2 \langle y_n - w_n, w_n - y_n \rangle \\ &\quad + 2 \langle y_n - w_n, y_n - p \rangle - 2\lambda_n \langle Ay_n - Aw_n, y_n - p \rangle \\ &= (-1 + \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|y_n - w_n\|^2 + \|w_n - p\|^2 \\ &\quad - 2 \langle w_n - y_n - \lambda_n(Aw_n - Ay_n), y_n - p \rangle. \end{aligned} \quad (3.5)$$

Since  $y_n = (I + \lambda_n B)^{-1}(I - \lambda_n A)w_n$  and B is maximal monotone, we obtain

$$\frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} \in By_n. \quad (3.6)$$

Thus

$$\frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} + Ay_n \in (A + B)y_n.$$

Since  $A : H \rightarrow H$  is monotone and  $B : H \rightarrow 2^H$  is a maximal monotone operator, Lemma 2.3 implies that  $A + B$  is maximal monotone. Since  $p \in \Omega$ ,  $0 \in (A + B)p$  and so

$$\left\langle \frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} + Ay_n - 0, y_n - p \right\rangle \geq 0.$$

Hence

$$\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), y_n - p \rangle \geq 0. \quad (3.7)$$

Combining (3.5) with (3.7), we get

$$\|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 \leq \|w_n - p\|^2 - (1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2. \quad (3.8)$$

The proof of the part (i) is completed.

(ii) **Case I B** is  $r$ -strongly monotone.

Since  $p \in \omega$ , we have  $0 \in (A + B)p$  and thus  $-Ap \in Bp$ . Since  $B$  is  $r$ -strongly monotone, by (3.6), we have

$$\left\langle \frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} + Ap, y_n - p \right\rangle \geq r\|y_n - p\|^2. \quad (3.9)$$

The monotonicity of  $A$  implies that

$$\langle Ay_n - Ap, y_n - p \rangle \geq 0. \quad (3.10)$$

Adding together (3.9) and (3.10), we have

$$\left\langle \frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} + Ay_n, y_n - p \right\rangle \geq r\|y_n - p\|^2,$$

which implies

$$\langle w_n - y_n - \lambda_n Aw_n + \lambda_n Ay_n, y_n - p \rangle \geq r\lambda_n\|y_n - p\|^2. \quad (3.11)$$

Utilizing (3.5) and (3.11), we get

$$\begin{aligned} \|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 &\leq \|w_n - p\|^2 - (1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \\ &\quad - 2r\lambda_n\|y_n - p\|^2. \end{aligned} \quad (3.12)$$

**Case II A** is  $r$ -strongly monotone.

The strong monotonicity of  $A$  implies that

$$\langle Ay_n - Ap, y_n - p \rangle \geq r\|y_n - p\|^2. \quad (3.13)$$

Since  $p \in \omega$ , we have  $0 \in (A + B)p$  and thus  $-Ap \in Bp$ . Since  $B$  is monotone, by (3.6), we have

$$\left\langle \frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n} + Ap, y_n - p \right\rangle \geq 0. \quad (3.14)$$

The rest of proof is the same as in Case I. This completes the proof of Lemma 3.2.

**Lemma 3.3** Assume that the conditions (C<sub>1</sub>) and (C<sub>2</sub>) hold, and  $\{w_n\}$  and  $\{y_n\}$  are sequences generated by Algorithm 3.1. If  $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$  and  $\{w_n\}$  converges weakly to some  $z \in H$ , then  $z \in \Omega$ .

*Proof* Letting  $(u, v) \in \text{Graph}(A + B)$ , we get  $v - Au \in Bu$ . Since  $B$  is maximal monotone, by (3.6), we have

$$\left\langle v - Au - \frac{w_n - y_n - \lambda_n Aw_n}{\lambda_n}, u - y_n \right\rangle \geq 0. \quad (3.15)$$

This implies that

$$\begin{aligned} \langle v, u - y_n \rangle &\geq \frac{1}{\lambda_n} \langle \lambda_n Au + w_n - y_n - \lambda_n Aw_n, u - y_n \rangle \\ &= \frac{1}{\lambda_n} \langle u - y_n, w_n - y_n \rangle + \langle u - y_n, Au - Aw_n \rangle \\ &= \frac{1}{\lambda_n} \langle u - y_n, w_n - y_n \rangle + \langle u - y_n, Au - Ay_n \rangle + \langle u - y_n, Ay_n - Aw_n \rangle. \end{aligned} \quad (3.16)$$

From the Lipschitz continuity and monotonicity of  $A$ , it follows that

$$\begin{aligned} \langle u - y_n, v \rangle &\geq \frac{1}{\lambda_n} \langle u - y_n, w_n - y_n \rangle + \langle u - y_n, Ay_n - Aw_n \rangle \\ &\geq \frac{1}{\lambda_n} \langle u - y_n, w_n - y_n \rangle - \|u - y_n\| \|Ay_n - Aw_n\| \\ &\geq \frac{1}{\lambda_n} \langle u - y_n, w_n - y_n \rangle - L \|u - y_n\| \|y_n - w_n\|. \end{aligned} \quad (3.17)$$

Since  $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$ , by Lemma 3.1 and (3.17), we have

$$\langle u - z, v \rangle = \lim_{n \rightarrow \infty} \langle u - w_n, v \rangle = \lim_{n \rightarrow \infty} \langle u - y_n, v \rangle \geq 0$$

Since  $v \in (A + B)u$  and  $A + B$  is maximal monotone, we know  $0 \in (A + B)z$ , that is,  $z \in \Omega$ . This proof is completed.

**Theorem 3.1** *Assume that the conditions (C<sub>1</sub>)-(C<sub>3</sub>) hold. Then the sequence  $\{x_n\}$  from Algorithm 3.1 converges weakly to some element  $p \in \Omega$ .*

*Proof* By the definition of  $x_{n+1}$  and Lemma 2.7, we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|(1 - \theta_n)z_n + \theta_n(y_n - \lambda_n(Ay_n - Aw_n)) - p\|^2 \\ &= (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 \\ &\quad - (1 - \theta_n)\theta_n\|z_n - y_n + \lambda_n(Ay_n - Aw_n)\|^2. \end{aligned} \quad (3.18)$$

The definition of  $x_{n+1}$  implies that  $y_n - \lambda_n(Ay_n - Aw_n) = \frac{x_{n+1} - (1 - \theta_n)z_n}{\theta_n}$ . Thus

$$\begin{aligned} \|z_n - y_n + \lambda_n(Ay_n - Aw_n)\|^2 &= \left\|z_n - \frac{x_{n+1} - (1 - \theta_n)z_n}{\theta_n}\right\|^2 \\ &= \frac{1}{\theta_n^2}\|x_{n+1} - z_n\|^2. \end{aligned} \quad (3.19)$$

By Lemma 3.2, we have

$$\|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 \leq \|w_n - p\|^2 - (1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2, \forall p \in \Omega. \quad (3.20)$$

Substituting (3.19) and (3.20) into (3.18), we get

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \theta_n(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \\ &\quad - \frac{1 - \theta_n}{\theta_n}\|x_{n+1} - z_n\|^2. \end{aligned} \quad (3.21)$$

Lemma 3.1 and the fact  $\lim_{n \rightarrow \infty} \mu_n = 0$  imply that  $\lim_{n \rightarrow \infty} 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} = 1 - \mu^2 > 0$ . Thus there exists a positive integer  $N \geq 1$  such that

$$\theta_n(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \geq 0, \forall n \geq N.$$

Thanks to (3.21), we have

$$\|x_{n+1} - p\|^2 \leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \frac{1 - \theta_n}{\theta_n}\|x_{n+1} - z_n\|^2, \forall n \geq N. \quad (3.22)$$

By the definition of  $z_n$  and Lemma 2.7, we obtain

$$\begin{aligned} \|z_n - p\|^2 &= \|x_n + \beta_n(x_n - x_{n-1}) - p\|^2 \\ &= \|(1 + \beta_n)(x_n - p) - \beta_n(x_{n-1} - p)\|^2 \\ &= (1 + \beta_n)\|x_n - p\|^2 - \beta_n\|x_{n-1} - p\|^2 + \beta_n(1 + \beta_n)\|x_n - x_{n-1}\|^2. \end{aligned} \quad (3.23)$$

From the definition of  $w_n$  and Lemma 2.7, it follows that

$$\|w_n - p\|^2 = (1 + \alpha_n)\|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2 + \alpha_n(1 + \alpha_n)\|x_n - x_{n-1}\|^2. \quad (3.24)$$

The definition of  $z_n$  means that

$$\begin{aligned} \|x_{n+1} - z_n\|^2 &= \|x_{n+1} - x_n - \beta_n(x_n - x_{n-1})\|^2 \\ &\geq \|x_{n+1} - x_n\|^2 + \beta_n^2\|x_n - x_{n-1}\|^2 - 2\beta_n\|x_{n+1} - x_n\|\|x_n - x_{n-1}\| \\ &\geq (1 - \beta_n)\|x_{n+1} - x_n\|^2 + (\beta_n^2 - \beta_n)\|x_n - x_{n-1}\|^2. \end{aligned} \quad (3.25)$$

Owing to (3.23), (3.24), (3.25) and (3.22), we know

$$\begin{aligned}
\|x_{n+1} - p\|^2 &\leq (1 - \theta_n)[(1 + \beta_n)\|x_n - p\|^2 - \beta_n\|x_{n-1} - p\|^2 + \beta_n(1 + \beta_n)\|x_n - x_{n-1}\|^2] \\
&\quad + \theta_n[(1 + \alpha_n)\|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2 + \alpha_n(1 + \alpha_n)\|x_n - x_{n-1}\|^2] \\
&\quad - \frac{1 - \theta_n}{\theta_n}[(1 - \beta_n)\|x_{n+1} - x_n\|^2 + (\beta_n^2 - \beta_n)\|x_n - x_{n-1}\|^2] \\
&\leq [1 + (1 - \theta_n)\beta_n + \theta_n\alpha_n]\|x_n - p\|^2 - [(1 - \theta_n)\beta_n + \theta_n\alpha_n]\|x_{n-1} - p\|^2 \\
&\quad + [(1 - \theta_n)(1 + \beta_n)\beta_n + \theta_n(1 + \alpha_n)\alpha_n - \frac{(1 - \theta_n)}{\theta_n}(\beta_n^2 - \beta_n)]\|x_n - x_{n-1}\|^2 \\
&\quad - \frac{(1 - \theta_n)}{\theta_n}(1 - \beta_n)\|x_{n+1} - x_n\|^2 \\
&\leq (1 + a_n)\|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2 + b_n\|x_n - x_{n-1}\|^2 - c_n\|x_{n+1} - x_n\|^2.
\end{aligned} \tag{3.26}$$

where  $a_n = (1 - \theta_n)\beta_n + \theta_n\alpha_n$ ,  $b_n := (1 - \theta_n)(1 + \beta_n)\beta_n + \theta_n(1 + \alpha_n)\alpha_n - \frac{1 - \theta_n}{\theta_n}(\beta_n^2 - \beta_n)$  and  $c_n := \frac{1 - \theta_n}{\theta_n}(1 - \beta_n)$ .

Define

$$\Gamma_n := \|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2 + b_n\|x_n - x_{n-1}\|^2$$

Since  $\{a_n\}$  is non-decreasing, by (3.26), we deduce that

$$\begin{aligned}
\Gamma_{n+1} - \Gamma_n &= \|x_{n+1} - p\|^2 - (1 + a_{n+1})\|x_n - p\|^2 + a_n\|x_{n-1} - p\|^2 \\
&\quad - b_n\|x_n - x_{n-1}\|^2 + b_{n+1}\|x_{n+1} - x_n\|^2 \\
&\leq \|x_{n+1} - p\|^2 - (1 + a_n)\|x_n - p\|^2 + a_n\|x_{n-1} - p\|^2 \\
&\quad - b_n\|x_n - x_{n-1}\|^2 + b_{n+1}\|x_{n+1} - x_n\|^2 \\
&\leq -(c_n - b_{n+1})\|x_{n+1} - x_n\|^2.
\end{aligned} \tag{3.27}$$

The condition  $(C_3)$  means that

$$\begin{aligned}
c_n - b_{n+1} &= \frac{1 - \theta_n}{\theta_n}(1 - \beta_n) - (1 - \theta_{n+1})(1 + \beta_{n+1})\beta_{n+1} - \theta_{n+1}(1 + \alpha_{n+1})\alpha_{n+1} \\
&\quad + \frac{1 - \theta_{n+1}}{\theta_{n+1}}(\beta_{n+1}^2 - \beta_{n+1}) \\
&\geq \frac{1 - \theta_{n+1}}{\theta_{n+1}}(1 - 2\beta_{n+1} + \beta_{n+1}^2) - (1 - \theta_{n+1})(\beta_{n+1} + \beta_{n+1}^2) - 2\theta_{n+1} \\
&\geq \varepsilon(1 - 2\beta + \beta^2) - \frac{\varepsilon}{1 + \varepsilon}(\beta + \beta^2) - \frac{2}{1 + \varepsilon} \\
&= \frac{1}{1 + \varepsilon}[\varepsilon^2\beta^2 - (3\varepsilon + 2\varepsilon^2)\beta + (\varepsilon^2 + \varepsilon - 2)].
\end{aligned} \tag{3.28}$$

Combining (3.27) and (3.28), we infer that

$$\Gamma_{n+1} - \Gamma_n \leq -\delta\|x_{n+1} - x_n\|^2. \tag{3.29}$$

where  $\delta := \frac{\varepsilon^2\beta^2 - (3\varepsilon + 2\varepsilon^2)\beta + (\varepsilon^2 + \varepsilon - 2)}{1 + \varepsilon}$ . Since  $\beta < \frac{3+2\varepsilon-\sqrt{8\varepsilon+17}}{2\varepsilon}$ ,  $\varepsilon \in (1, +\infty)$ , we conclude that  $\delta > 0$ . From (3.29), it follows that

$$\Gamma_{n+1} - \Gamma_n \leq 0.$$

Thus the sequence  $\{\Gamma_n\}$  is nonincreasing. The condition  $(C_3)$  implies that  $b_n > 0$ . Thus

$$\begin{aligned}
\Gamma_n &= \|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2 + b_n\|x_n - x_{n-1}\|^2 \\
&\geq \|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2.
\end{aligned} \tag{3.30}$$

This implies that

$$\begin{aligned}
\|x_n - p\|^2 &\leq a_n\|x_{n-1} - p\|^2 + \Gamma_n \\
&\leq a\|x_{n-1} - p\|^2 + \Gamma_1 \\
&\vdots \\
&\leq a^n\|x_0 - p\|^2 + \Gamma_1(1 + a + \dots + a^{n-1}) \\
&\leq a^n\|x_0 - p\|^2 + \frac{\Gamma_1}{1 - a}.
\end{aligned} \tag{3.31}$$

where  $a := \frac{5+2\varepsilon-\sqrt{8\varepsilon+17}}{2+2\varepsilon} < 1$ .

By definition of  $\{\Gamma_n\}$ , we have

$$\begin{aligned}\Gamma_{n+1} &= \|x_{n+1} - p\|^2 - a_{n+1}\|x_n - p\|^2 + b_{n+1}\|x_{n+1} - x_n\|^2 \\ &\geq -a_{n+1}\|x_n - p\|^2.\end{aligned}\quad (3.32)$$

According to (3.31) and (3.32), we infer that

$$-\Gamma_{n+1} \leq a_{n+1}\|x_n - p\|^2 \leq a\|x_n - p\|^2 \leq a^{n+1}\|x_0 - p\|^2 + \frac{a\Gamma_1}{1-a}. \quad (3.33)$$

By (3.33) and (3.29), we get

$$\begin{aligned}\delta \sum_{k=1}^n \|x_{k+1} - x_k\|^2 &\leq \Gamma_1 - \Gamma_{n+1} \leq a^{n+1}\|x_0 - p\|^2 + \frac{\Gamma_1}{1-a} \\ &\leq \|x_0 - p\|^2 + \frac{\Gamma_1}{1-a}.\end{aligned}\quad (3.34)$$

This implies

$$\sum_{k=1}^{\infty} \|x_{k+1} - x_k\|^2 \leq \infty. \quad (3.35)$$

Thus

$$\|x_{n+1} - x_n\| \rightarrow 0, n \rightarrow \infty. \quad (3.36)$$

The definition of  $\{\Gamma_n\}$  implies that

$$\|x_{n+1} - w_n\|^2 = \|x_{n+1} - x_n\|^2 + \alpha_n\|x_n - x_{n-1}\|^2 - 2\alpha_n \langle x_{n+1} - x_n, x_n - x_{n-1} \rangle. \quad (3.37)$$

Since  $\{\alpha_n\}$  is bounded, by (3.37), we obtain

$$\|x_{n+1} - w_n\|^2 \rightarrow 0, n \rightarrow \infty. \quad (3.38)$$

On the other hand

$$\|x_n - w_n\| \leq \|x_n - x_{n+1}\| + \|x_{n+1} - w_n\|. \quad (3.39)$$

From (3.36) and (3.38), it follows that

$$\|x_n - w_n\|^2 \rightarrow 0, n \rightarrow \infty. \quad (3.40)$$

By (3.26), we have

$$\begin{aligned}\|x_{n+1} - p\|^2 &\leq (1 + a_n)\|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2 + b_n\|x_n - x_{n-1}\|^2 - c_n\|x_{n+1} - x_n\|^2 \\ &\leq (1 + a_n)\|x_n - p\|^2 - a_n\|x_{n-1} - p\|^2 + b_n\|x_n - x_{n-1}\|^2 \\ &= \|x_n - p\|^2 + a_n(\|x_n - p\|^2 - \|x_{n-1} - p\|^2) + b_n\|x_n - x_{n-1}\|^2.\end{aligned}\quad (3.41)$$

Since  $0 \leq a_n < a < 1$  and  $\{b_n\}$  is bounded, by Lemma 2.1 and (3.35), we know there exists  $l \in [0, +\infty)$  such that

$$\lim_{n \rightarrow \infty} \|x_n - p\|^2 = l. \quad (3.42)$$

Then from (3.24), we have

$$\begin{aligned}\|w_n - p\|^2 &= (1 + \alpha_n)\|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2 + \alpha_n(1 + \alpha_n)\|x_n - x_{n-1}\|^2 \\ &= \|x_n - p\|^2 + \alpha_n(\|x_n - p\|^2 - \|x_{n-1} - p\|^2) + \alpha_n(1 + \alpha_n)\|x_n - x_{n-1}\|^2.\end{aligned}\quad (3.43)$$

Since  $\{\alpha_n\}$  is bounded, by (3.36) and (3.42), we obtain

$$\lim_{n \rightarrow \infty} \|w_n - p\|^2 = l. \quad (3.44)$$

Utilizing the similar discussion as in obtaining (3.44), we can get

$$\lim_{n \rightarrow \infty} \|z_n - p\|^2 = l. \quad (3.45)$$

Owing to (3.21), we have

$$\begin{aligned} \theta_n \left(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2}\right) \|w_n - y_n\|^2 &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 \\ &\quad - \frac{(1 - \theta_n)}{\theta_n} \|x_{n+1} - z_n\|^2 - \|x_{n+1} - p\|^2 \\ &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - \|x_{n+1} - p\|^2. \end{aligned} \quad (3.46)$$

Since  $0 < \theta_n < \frac{1}{1 + \epsilon}$ ,  $\epsilon \in (1, +\infty)$  and  $\lim_{n \rightarrow \infty} 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} = 1 - \mu^2 > 0$ , by (3.42) (3.44) and (3.45), we get  $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$ .

Since  $\{x_n\}$  is bounded, we assume that there exists a subsequence  $\{x_{n_k}\}$  of  $\{x_n\}$  such that  $x_{n_k} \rightharpoonup z \in H$ . The fact  $\lim_{n \rightarrow \infty} \|w_n - x_n\| = 0$  implies that  $w_{n_k} \rightharpoonup z \in H$ . By Lemma 3.3, we know  $z \in \Omega$ . The two assumptions of Lemma 2.2 are verified. Lemma 2.2 ensure that the sequence  $\{x_n\}$  converges weakly to  $\mu^* \in \Omega$ . The proof is completed.

Let  $C$  is a nonempty, closed and convex subset of  $H$ . If  $B = N_C$ , then MIP (1.1) reduce to the following variational inequality, denoted by VI( $A, C$ ): find a point  $x^* \in C$  such that

$$\langle A(x^*), y - x^* \rangle \geq 0, \quad \forall y \in C.$$

Denote the solution set of VI( $A, C$ ) by  $S$ .

- Assumption 3.1** (i) *The solution set of the problem (VI) is nonempty, that is,  $S \neq \emptyset$ .*  
(ii) *The mapping  $A : H \rightarrow H$  is pseudomonotone, Lipschitz continuous and  $A$  satisfies the condition: for any  $\{x_n\} \subset H$  with  $x_n \rightharpoonup w^*$ , one has  $\|Aw^*\| \leq \liminf_{n \rightarrow \infty} \|Ax_n\|$ .*

**Proposition 3.1** *Suppose that  $B = N_C$  in Algorithm 3.1. Thus the following statements hold:*

- (i) *if the Assumption 3.1 hold, then*

$$\|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 \leq \|w_n - p\|^2 - \left(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2}\right) \|w_n - y_n\|^2, \quad \forall p \in S;$$

- (ii) *if the Assumption 3.1 hold and  $A$  is  $\mu$ -strongly pseudomonotone, then*

$$\begin{aligned} \|y_n - \lambda_n(Ay_n - Aw_n) - p\|^2 &\leq \|w_n - p\|^2 - \left(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2}\right) \|w_n - y_n\|^2 \\ &\quad - 2\mu\lambda_n \|y_n - p\|^2, \quad \forall p \in S. \end{aligned}$$

*Proof* Since  $B = N_C$ , we know  $y_n = (I + \lambda_n B)^{-1}(I - \lambda_n A)w_n = P_C(w_n - \lambda_n Aw_n)$ . Thus we get

$$\langle y_n - w_n + \lambda_n Aw_n, y_n - p \rangle \leq 0. \quad (3.47)$$

For any given  $p \in S$ ,  $\langle Ap, y_n - p \rangle \geq 0$ . Since  $A$  is pseudomonotone, we have

$$\langle Ay_n, y_n - p \rangle \geq 0. \quad (3.48)$$

From (3.47) and (3.48), it follows that

$$\langle w_n - y_n - \lambda_n Aw_n + \lambda_n Ay_n, y_n - p \rangle \geq 0. \quad (3.49)$$

which is just (3.7). The rest of the proof follows the same arguments as in Lemma 3.2. The proof of the part (i) is completed.

- (ii) For any given  $p \in S$ ,  $\langle Ap, y_n - p \rangle \geq 0$ . The strong pseudomonotonicity of  $A$  implies that

$$\langle Ay_n, y_n - p \rangle \geq \mu \|y_n - p\|^2. \quad (3.50)$$

According to (3.47) and (3.50), we have

$$\langle w_n - y_n - \lambda_n(Aw_n - Ay_n), y_n - p \rangle \geq \mu\lambda_n \|y_n - p\|^2. \quad (3.51)$$

The rest of proof is the same as in Lemma 3.2. This completes the proof of Proposition 3.1.

**Proposition 3.2** Assume that  $B = N_C$ , the Assumption 3.1 hold, and  $\{w_n\}$  and  $\{y_n\}$  are sequences generated by Algorithm 3.1. If  $\lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$  and  $\{w_n\}$  converges weakly to some  $z \in H$ , then  $z \in S$ .

*Proof* This proof is the same as in Lemma 3.7 of [24], and we omit it.

**Corollary 3.1** If  $B = N_C$ , the conditions  $(C_3)$  and Assumption 3.1 hold, then the sequence  $\{x_n\}$  from Algorithm 3.1 converges weakly to a point  $p \in S$ .

*Proof* Replacing Lemmas 3.2 and 3.3 by Propositions 3.1 and 3.2, respectively and using the same proof as in Theorem 3.1, we obtain the desired conclusion.

**Remark 3.4** Compared with Theorem 4.2 of [29], the advantages of Corollary 3.1 have (i) the sequence  $\{\alpha_n\}$  may not be non-decreasing; (ii) the sequence  $\{\beta_n\}$  may not be a constant; (iii) we require  $0 < \theta < \theta_n \leq \theta_{n+1} < \frac{1}{1+\epsilon}, \epsilon \in (1, +\infty)$  other than  $\epsilon \in (2, +\infty)$ , which extend the taking value interval of  $\theta_n$ . From the numerical experiment in Section 5, it can be seen that the larger the values of  $\theta_n$ , the better the algorithm performs.

Motivated by the Theorem 5.1 of [20], which may be the first nonasymptotic convergence rate results of inertial projection-type algorithm for solving variational inequalities with monotone mappings, we give the nonasymptotic  $O(\frac{1}{\sqrt{n}})$  convergence rate with “ $\min_{N \leq i \leq n}$ ” for Algorithm 3.1.

**Theorem 3.2** Assume that the conditions  $(C_1)$ - $(C_3)$  hold, the sequence  $\{x_n\}$  be generated by Algorithm 3.1 and  $[t]_+ := \max\{t, 0\}$ . Then for any  $p \in \Omega$ , there exist constants  $M$  such that the following estimate holds

$$\min_{N \leq i \leq n} \|w_i - y_i\| \leq \left( \frac{(\|x_N - p\|^2 + \frac{a}{1-a}[\|x_N - p\|^2 - \|x_{N-1} - p\|^2])_+ + \frac{(2+\frac{1-\theta}{4a})M}{1-a}}{n - N + 1} \right)^{\frac{1}{2}}.$$

*Proof* Lemma 3.1 and the fact  $\lim_{n \rightarrow \infty} \mu_n = 0$  imply that  $\lim_{n \rightarrow \infty} 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} = 1 - \mu^2 > 0$ . Thus there exists a positive integer  $N \geq 1$  such that

$$\theta_n \left( 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} \right) \|w_n - y_n\|^2 \geq \theta_n (1 - \mu) \|w_n - y_n\|^2, \quad \forall n \geq N.$$

Thanks to (3.21), we have, for all  $n \geq N$

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - \theta_n (1 - \mu) \|w_n - y_n\|^2 - \frac{1 - \theta_n}{\theta_n} \|x_{n+1} - z_n\|^2 \\ &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - \theta (1 - \mu) \|w_n - y_n\|^2 - \frac{1 - \theta_n}{\theta_n} \|x_{n+1} - z_n\|^2. \end{aligned} \quad (3.52)$$

Owing to (3.23), (3.24), (3.25) and (3.52), we know, for all  $n \geq N$

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \theta_n) [(1 + \beta_n) \|x_n - p\|^2 - \beta_n \|x_{n-1} - p\|^2 + \beta_n (1 + \beta_n) \|x_n - x_{n-1}\|^2] \\ &\quad + \theta_n [(1 + \alpha_n) \|x_n - p\|^2 - \alpha_n \|x_{n-1} - p\|^2 + \alpha_n (1 + \alpha_n) \|x_n - x_{n-1}\|^2] \\ &\quad - \frac{1 - \theta_n}{\theta_n} [(1 - \beta_n) \|x_{n+1} - x_n\|^2 + (\beta_n^2 - \beta_n) \|x_n - x_{n-1}\|^2] \\ &\quad - \theta (1 - \mu) \|w_n - y_n\|^2 \\ &\leq [1 + (1 - \theta_n) \beta_n + \theta_n \alpha_n] \|x_n - p\|^2 - [(1 - \theta_n) \beta_n + \theta_n \alpha_n] \|x_{n-1} - p\|^2 \\ &\quad + [(1 - \theta_n) (1 + \beta_n) \beta_n + \theta_n (1 + \alpha_n) \alpha_n - \frac{(1 - \theta_n)}{\theta_n} (\beta_n^2 - \beta_n)] \|x_n - x_{n-1}\|^2 \\ &\quad - \frac{(1 - \theta_n)}{\theta_n} (1 - \beta_n) \|x_{n+1} - x_n\|^2 - \theta (1 - \mu) \|w_n - y_n\|^2 \\ &\leq (1 + a_n) \|x_n - p\|^2 - a_n \|x_{n-1} - p\|^2 + b_n \|x_n - x_{n-1}\|^2 - c_n \|x_{n+1} - x_n\|^2 \\ &\quad - \theta (1 - \mu) \|w_n - y_n\|^2. \\ &\leq (1 + a_n) \|x_n - p\|^2 - a_n \|x_{n-1} - p\|^2 + b_n \|x_n - x_{n-1}\|^2 - \theta (1 - \mu) \|w_n - y_n\|^2. \end{aligned} \quad (3.53)$$

where  $\{a_n\}$ ,  $\{b_n\}$  and  $\{c_n\}$  have the same definitions as in (3.26).

This implies that, for all  $n \geq N$

$$\theta(1-\mu) \|w_n - y_n\|^2 \leq \|x_n - p\|^2 - \|x_{n+1} - p\|^2 + a_n(\|x_n - p\|^2 - \|x_{n-1} - p\|^2) + b_n\|x_n - x_{n-1}\|^2. \quad (3.54)$$

Let  $\sigma_n = \|x_n - p\|^2$ ,  $K_n = \sigma_n - \sigma_{n-1}$  and  $\tau_n = b_n\|x_n - x_{n-1}\|^2$ . We get, for all  $n \geq N$

$$\begin{aligned} \theta(1-\mu) \|w_n - y_n\|^2 &\leq \sigma_n - \sigma_{n+1} + a_n K_n + \tau_n \\ &\leq \sigma_n - \sigma_{n+1} + a_n [K_n]_+ + \tau_n \\ &\leq \sigma_n - \sigma_{n+1} + a [K_n]_+ + \tau_n \end{aligned} \quad (3.55)$$

where  $a$  has been defined in (3.31).

In view of (3.35), we have  $\sum_{k=1}^{\infty} \|x_{k+1} - x_k\|^2 \leq \infty$ . Thus, there exists a positive constant  $M$  such that

$$\sum_{k=1}^{\infty} \|x_{k+1} - x_k\|^2 \leq M$$

Therefore,

$$\begin{aligned} \sum_{n=1}^{\infty} \tau_n &= \sum_{n=1}^{\infty} b_n \|x_n - x_{n-1}\|^2 \\ &= \sum_{n=1}^{\infty} [(1-\theta_n)(1+\beta_n)\beta_n + \theta_n(1+\alpha_n)\alpha_n + \frac{1-\theta_n}{\theta_n}(\beta_n - \beta_n^2)] \|x_n - x_{n-1}\|^2 \\ &\leq \sum_{n=1}^{\infty} (2 + \frac{1-\theta}{4\theta}) \|x_n - x_{n-1}\|^2 \\ &= (2 + \frac{1-\theta}{4\theta}) \sum_{n=1}^{\infty} \|x_n - x_{n-1}\|^2 \\ &\leq (2 + \frac{1-\theta}{4\theta}) M = C_1 \end{aligned} \quad (3.56)$$

From (3.53), it follows that, for all  $n \geq N$

$$K_{n+1} \leq a_n K_n + \tau_n \leq a_n [K_n]_+ + \tau_n.$$

Thus,

$$\begin{aligned} [K_{n+1}]_+ &\leq a [K_n]_+ + \tau_n \\ &\leq a^{n-N+1} [K_n]_+ + \sum_{j=1}^{n-N+1} a^{j-1} \tau_{n+1-j}. \end{aligned} \quad (3.57)$$

Combining (3.56) and (3.57), we obtain

$$\begin{aligned} \sum_{n=N}^{\infty} [K_{n+1}]_+ &\leq \frac{a}{1-a} [K_N]_+ + \frac{1}{1-a} \sum_{n=N}^{\infty} \tau_n \\ &\leq \frac{a}{1-a} [K_N]_+ + \frac{C_1}{1-a}. \end{aligned} \quad (3.58)$$

From (3.55), it follows that

$$\begin{aligned}
\theta(1-\mu) \sum_{n=N}^n \|w_n - y_n\|^2 &\leq \sigma_N - \sigma_{n+1} + a \sum_{n=N}^n [K_n]_+ + \sum_{n=N}^n \tau_n \\
&\leq \sigma_N + a([K_N]_+ + \sum_{n=N}^n [K_{n+1}]_+) + \sum_{n=N}^n \tau_n \\
&\leq \sigma_N + a[K_N]_+ + \frac{a^2}{1-a}[K_N]_+ + \frac{aC_1}{1-a} + C_1 \\
&= \sigma_N + \frac{a}{1-a}[K_N]_+ + \frac{C_1}{1-a} \\
&= \sigma_N + \frac{a}{1-a}[K_N]_+ + \frac{(2+\frac{1-\theta}{4\theta})M}{1-a}.
\end{aligned} \tag{3.59}$$

This implies that

$$\sum_{i=N}^n \|w_i - y_i\|^2 \leq (\|x_N - p\|^2 + \frac{a}{1-a}[\|x_N - p\|^2 - \|x_{N-1} - p\|^2]_+ + \frac{(2+\frac{1-\theta}{4\theta})M}{1-a}) \frac{1}{\theta(1-\mu)}. \tag{3.60}$$

and thus

$$\min_{N \leq i \leq n} \|w_i - y_i\|^2 \leq \frac{(\|x_N - p\|^2 + \frac{a}{1-a}[\|x_N - p\|^2 - \|x_{N-1} - p\|^2]_+ + \frac{(2+\frac{1-\theta}{4\theta})M}{1-a}) \frac{1}{\theta(1-\mu)}}{n-N+1}. \tag{3.61}$$

Since  $[\min_{N \leq i \leq n} \|w_i - y_i\|^2]^{\frac{1}{2}} = \min_{N \leq i \leq n} \|w_i - y_i\|$ , we get

$$\min_{N \leq i \leq n} \|w_i - y_i\| \leq \left( \frac{(\|x_N - p\|^2 + \frac{a}{1-a}[\|x_N - p\|^2 - \|x_{N-1} - p\|^2]_+ + \frac{(2+\frac{1-\theta}{4\theta})M}{1-a}) \frac{1}{\theta(1-\mu)}}{n-N+1} \right)^{\frac{1}{2}}. \tag{3.62}$$

This completes the proof.

*Remark 3.5* By Lemma 2.6 we know  $y_n = w_n$  implies that  $y_n$  is a solution of MIP. This means that the error estimate given in Theorem 3.2 can be regarded as the convergence rate of Algorithm 3.1.

#### 4 Strong convergence

In this section, we analyse strong convergence and linear convergence of Algorithm 3.1. Firstly, we give the following assumption.

**Assumption 4.1** *The mapping  $A : H \rightarrow H$  is  $L$ -Lipschitz continuous,  $r$ -strongly monotone and the set-valued mapping  $B : H \rightarrow 2^H$  is maximal monotone or the mapping  $A : H \rightarrow H$  is  $L$ -Lipschitz continuous, monotone and the set-valued mapping  $B : H \rightarrow 2^H$  is  $r$ -strongly monotone.*

**Theorem 4.1** *Assume that the conditions  $(C_1), (C_3)$  and Assumption 4.1 hold. Let  $\{x_n\}$  be a sequence generated by Algorithm 3.1. Then  $\{x_n\}$  converges strongly to some solution  $p \in \Omega$ .*

*Proof* By (3.18) (3.19) and Lemma 3.2 (ii), we have

$$\begin{aligned}
\|x_{n+1} - p\|^2 &\leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \theta_n(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \\
&\quad - 2\theta_n r \lambda_n \|y_n - p\|^2 - \frac{(1 - \theta_n)}{\theta_n} \|x_{n+1} - z_n\|^2 \\
&\leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \theta_n(1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2})\|w_n - y_n\|^2 \\
&\quad - 2\theta_n r \lambda_n \|y_n - p\|^2.
\end{aligned} \tag{4.1}$$

Lemma 3.1 and the fact  $\lim_{n \rightarrow \infty} \mu_n = 0$  imply that  $\lim_{n \rightarrow \infty} 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} = 1 - \mu^2 > 0$ . Thus there exists a positive integer  $N \geq 1$  such that

$$\theta_n \left( 1 - \frac{(\mu + \mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} \right) \|w_n - y_n\|^2 \geq 0, \quad \forall n \geq N.$$

It follows that from (4.1)

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - 2\theta_n r \lambda_n \|y_n - p\|^2 \\ &\leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - 2\theta r \lambda_n \|y_n - p\|^2, \quad \forall n \geq N. \end{aligned} \quad (4.2)$$

From Lemma 3.1, it follows that

$$\|x_{n+1} - p\|^2 \leq (1 - \theta_n) \|z_n - p\|^2 + \theta_n \|w_n - p\|^2 - 2\theta r \lambda^* \|y_n - p\|^2, \quad \forall n \geq N, \quad (4.3)$$

where  $\lambda^* = \min\{\frac{\mu}{L}, \lambda_1\}$ .

The definition of  $w_n$  implies that

$$\begin{aligned} \|w_n - p\|^2 &= \|x_n + \alpha_n(x_n - x_{n-1}) - p\|^2 \\ &\leq \|x_n - p\|^2 + \alpha_n^2 \|x_n - x_{n-1}\|^2 + 2\alpha_n \|x_n - p\| \|x_n - x_{n-1}\|. \end{aligned} \quad (4.4)$$

By utilizing the definition of  $z_n$ , we have

$$\|z_n - p\|^2 \leq \|x_n - p\|^2 + \beta_n^2 \|x_n - x_{n-1}\|^2 + 2\beta_n \|x_n - p\| \|x_n - x_{n-1}\|. \quad (4.5)$$

Combining (4.3) (4.4) and (4.5), we obtain

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq (1 - \theta_n)(\|x_n - p\|^2 + \beta_n^2 \|x_n - x_{n-1}\|^2 + 2\beta_n \|x_n - p\| \|x_n - x_{n-1}\|) \\ &\quad + \theta_n(\|x_n - p\|^2 + \alpha_n^2 \|x_n - x_{n-1}\|^2 + 2\alpha_n \|x_n - p\| \|x_n - x_{n-1}\|) \\ &\quad - 2\theta r \lambda^* \|y_n - p\|^2 \\ &= \|x_n - p\|^2 + ((1 - \theta_n)\beta_n^2 + \theta_n \alpha_n^2) \|x_n - x_{n-1}\|^2 \\ &\quad + 2((1 - \theta_n)\beta_n + \theta_n \alpha_n) \|x_n - p\| \|x_n - x_{n-1}\| - 2\theta r \lambda^* \|y_n - p\|^2, \quad \forall n \geq N. \end{aligned} \quad (4.6)$$

In addition,

$$\begin{aligned} \|x_n - p\|^2 &\leq 2(\|x_n - y_n\|^2 + \|y_n - p\|^2) \\ &\leq 4(\|x_n - w_n\|^2 + \|y_n - w_n\|^2) + 2\|y_n - p\|^2, \end{aligned}$$

which implies that

$$\|y_n - p\|^2 \geq \frac{1}{2} \|x_n - p\|^2 - 2\|y_n - w_n\|^2 - 2\|w_n - x_n\|^2. \quad (4.7)$$

In view of (4.6) and (4.7), we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &\leq \|x_n - p\|^2 + ((1 - \theta_n)\beta_n^2 + \theta_n \alpha_n^2) \|x_n - x_{n-1}\|^2 \\ &\quad + 2((1 - \theta_n)\beta_n + \theta_n \alpha_n) \|x_n - p\| \|x_n - x_{n-1}\| \\ &\quad - \theta r \lambda^* \|x_n - p\|^2 + 4\theta r \lambda^* \|y_n - w_n\|^2 + 4\theta r \lambda^* \|w_n - x_n\|^2 \\ &= (1 - \theta r \lambda^*) \|x_n - p\|^2 + ((1 - \theta_n)\beta_n^2 + \theta_n \alpha_n^2) \|x_n - x_{n-1}\|^2 \\ &\quad + 2((1 - \theta_n)\beta_n + \theta_n \alpha_n) \|x_n - p\| \|x_n - x_{n-1}\| \\ &\quad + 4\theta r \lambda^* \|y_n - w_n\|^2 + 4\theta r \lambda^* \|w_n - x_n\|^2, \quad \forall n \geq N. \end{aligned} \quad (4.8)$$

Since  $\lambda^* \leq \frac{\mu}{L}$  and  $r \leq L$ , we obtain  $1 - \theta r \lambda^* \in (0, 1)$ .

Since  $x_n$  is bounded, and  $\lim_{n \rightarrow \infty} \|x_{n+1} - x_n\| = \lim_{n \rightarrow \infty} \|x_n - w_n\| = \lim_{n \rightarrow \infty} \|w_n - y_n\| = 0$ , owing to Lemma 2.4, we have

$$\lim_{n \rightarrow \infty} \|x_n - p\|^2 = 0. \quad (4.9)$$

Hence

$$\lim_{n \rightarrow \infty} \|x_n - p\| = 0. \quad (4.10)$$

This proof is completed.

*Remark 4.1* To the best of our knowledge, Theorem 4.1 is one of the few available strong convergence results for algorithms with the double inertial extrapolation steps to solve MIP (1.1). In addition, we emphasize that Theorem 4.1 does not need to know the modulus of strong monotonicity and the Lipschitz constant in advance.

Owing to Proposition 3.1 and Theorem 4.1, it is easy to get the following results.

**Corollary 4.1** *Let  $B = N_C$  and  $A$  is  $\mu$ -strong pseudomonotone and  $L$ -Lipschitz continuous. If the conditions  $(C_3)$  hold, then the sequence  $\{x_n\}$  from Algorithm 3.1 converges strongly to a point  $p \in S$ .*

*Remark 4.2* Corollary 4.1 improves Theorem 5.1 of [29] in the following aspects: (i) the sequence  $\{\alpha_n\}$  may not be non-decreasing; (ii) the sequence  $\{\beta_n\}$  may not be a constant; (iii) we require  $0 < \theta < \theta_n \leq \frac{1}{\theta_{n+1}} < \frac{1}{1+\epsilon}$ ,  $\epsilon \in (1, +\infty)$  other than  $\epsilon \in (2, +\infty)$ , which extend the taking value interval of  $\theta_n$ .

In order to discuss the linear convergence rate of our algorithm, we need the following assumption.

- Assumption 4.2** (i) *The solution set of the inclusion problem (1.1) is nonempty, that is,  $\Omega \neq \emptyset$ .*  
(ii) *The mapping  $A : H \rightarrow H$  is  $L$ -Lipschitz continuous,  $r$ -strongly monotone and the set-valued mapping  $B : H \rightarrow 2^H$  is maximal monotone or the mapping  $A : H \rightarrow H$  is  $L$ -Lipschitz continuous, monotone and the set-valued mapping  $B : H \rightarrow 2^H$  is  $r$ -strongly monotone.*  
(iii) *Let  $\hat{\lambda} := \min\{\frac{\mu}{L}, \lambda_1\}$ ,  $\tau := 1 - \frac{1}{2}\min\{1 - \mu, 2\hat{\lambda}r\} \in (\frac{1}{2}, 1)$  and the following conditions hold:*

$$\begin{aligned} (c_1) \quad & 0 \leq \beta_n \leq \beta < \frac{1}{2}(\frac{1}{\tau} - 1) \\ (c_2) \quad & 0 \leq \alpha_n \leq \alpha < \frac{1-\tau}{\tau} \\ (c_3) \quad & \max\{\frac{1-\beta}{1+\alpha-\beta}, \frac{\beta}{1+\beta-\tau(1+\alpha)}\} < \theta \leq \theta_{n-1} \leq \theta_n \leq \frac{-1-\beta+\sqrt{(1+\beta)^2-4(\frac{1}{\tau}-1-2\beta)(\beta-1)}}{2(\frac{1}{\tau}-1-2\beta)} \end{aligned}$$

*Remark 4.3* The parameters set satisfying the condition is non-empty. For example, we can choose  $L = 1.5$ ,  $r = 1$ ,  $\mu = 0.45$ ,  $\beta = \beta_n = 0.1$ ,  $\alpha = \alpha_n = 0.37$  and  $\theta_n = 0.72$ .

Next, we establish the linear convergence rate of our algorithm under Assumption 4.2.

**Theorem 4.2** *Suppose that Assumption 4.2 hold. Then  $\{x_n\}$  generated by Algorithm 3.1 converges linearly to some point in  $\Omega$ .*

*Proof* By Lemma 3.1 we know  $\lim_{n \rightarrow \infty} \lambda_n = \lambda \geq \hat{\lambda} = \min\{\frac{\mu}{L}, \lambda_1\}$ . Since  $\lim_{n \rightarrow \infty} \mu_n = 0$ ,  $\lim_{n \rightarrow \infty} 1 - \frac{(\mu+\mu_n)^2 \lambda_n^2}{\lambda_{n+1}^2} = 1 - \mu^2 > 1 - \mu > 0$ . Thus it follows that from (4.1) there exists a positive integer  $N \geq 1$  such that

$$\begin{aligned} \|x_{n+1} - p\|^2 & \leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \frac{(1 - \theta_n)}{\theta_n}\|x_{n+1} - z_n\|^2 \\ & \quad - \theta_n(1 - \mu)\|w_n - y_n\|^2 - 2\theta_n r \hat{\lambda} \|y_n - p\|^2 \\ & \leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n\|w_n - p\|^2 - \frac{(1 - \theta_n)}{\theta_n}\|x_{n+1} - z_n\|^2 \\ & \quad - \theta_n \min\{1 - \mu, 2\hat{\lambda}r\}(\|w_n - y_n\|^2 + \|y_n - p\|^2) \\ & \leq (1 - \theta_n)\|z_n - p\|^2 + \theta_n \tau \|w_n - p\|^2 - \frac{(1 - \theta_n)}{\theta_n}\|x_{n+1} - z_n\|^2, \quad \forall n \geq N. \end{aligned} \tag{4.11}$$

Substituting (3.23), (3.24) and (3.25) into (4.11), we have

$$\begin{aligned}
\|x_{n+1} - p\|^2 &\leq (1 - \theta_n)[(1 + \beta_n)\|x_n - p\|^2 - \beta_n\|x_{n-1} - p\|^2 + (1 + \beta_n)\beta_n\|x_n - x_{n-1}\|^2] \\
&\quad + \theta_n\tau[(1 + \alpha_n)\|x_n - p\|^2 - \alpha_n\|x_{n-1} - p\|^2 + (1 + \alpha_n)\alpha_n\|x_n - x_{n-1}\|^2] \\
&\quad - \frac{(1 - \theta_n)}{\theta_n}[(1 - \beta_n)\|x_{n+1} - x_n\|^2 + (\beta_n^2 - \beta_n)\|x_n - x_{n-1}\|^2] \\
&\leq [(1 - \theta_n)(1 + \beta_n) + \theta_n\tau(1 + \alpha_n)]\|x_n - p\|^2 - [(1 - \theta_n)\beta_n + \theta_n\tau\alpha_n]\|x_{n-1} - p\|^2 \\
&\quad + [(1 - \theta_n)(1 + \beta_n)\beta_n + \theta_n\tau(1 + \alpha_n)\alpha_n + \frac{(1 - \theta_n)}{\theta_n}(\beta_n - \beta_n^2)]\|x_n - x_{n-1}\|^2 \\
&\quad - \frac{(1 - \theta_n)}{\theta_n}(1 - \beta_n)\|x_{n+1} - x_n\|^2 \\
&\leq [(1 - \theta_n)(1 + \beta_n) + \theta_n\tau(1 + \alpha_n)]\|x_n - p\|^2 \\
&\quad + [(1 - \theta_n)(1 + \beta_n)\beta_n + \theta_n\tau(1 + \alpha_n)\alpha_n + \frac{(1 - \theta_n)}{\theta_n}(\beta_n - \beta_n^2)]\|x_n - x_{n-1}\|^2 \\
&\quad - \frac{(1 - \theta_n)}{\theta_n}(1 - \beta_n)\|x_{n+1} - x_n\|^2 \\
&\leq [(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)]\|x_n - p\|^2 \\
&\quad + [(1 - \theta_n)(1 + \beta)\beta + \theta_n\tau(1 + \alpha)\alpha + \frac{(1 - \theta_n)}{\theta_n}(\beta - \beta^2)]\|x_n - x_{n-1}\|^2 \\
&\quad - \frac{(1 - \theta_n)}{\theta_n}(1 - \beta)\|x_{n+1} - x_n\|^2, \quad \forall n \geq N.
\end{aligned} \tag{4.12}$$

This implies that

$$\|x_{n+1} - p\|^2 + \sigma_n\|x_{n+1} - x_n\|^2 \leq [(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)](\|x_n - p\|^2 + \delta_n\|x_n - x_{n-1}\|^2). \tag{4.13}$$

where  $\delta_n = \frac{(1 - \theta_n)(1 + \beta)\beta + \theta_n\tau(1 + \alpha)\alpha + \frac{(1 - \theta_n)}{\theta_n}(\beta - \beta^2)}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)}$  and  $\sigma_n = \frac{(1 - \theta_n)}{\theta_n}(1 - \beta)$ .

Now, we will show that  $\delta_n < \sigma_n$ :

$$\begin{aligned}
\delta_n - \sigma_n &= \frac{(1 - \theta_n)(1 + \beta)\beta + \theta_n\tau(1 + \alpha)\alpha + \frac{1 - \theta_n}{\theta_n}(\beta - \beta^2)}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} - \frac{1 - \theta_n}{\theta_n}(1 - \beta) \\
&= \frac{(1 - \theta_n)(1 + \beta)\beta + \theta_n\tau(1 + \alpha)\alpha + \frac{1 - \theta_n}{\theta_n}(\beta - \beta^2)}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} \\
&\quad - \frac{[(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)]\frac{1 - \theta_n}{\theta_n}(1 - \beta)}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} \\
&= \frac{\frac{1 - \theta_n}{\theta_n}[\theta_n(1 + \beta)\beta + \beta - \beta^2 - (1 - \theta_n)(1 - \beta^2)]}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} \\
&\quad + \frac{\tau(1 + \alpha)[\theta_n\alpha - (1 - \theta_n)(1 - \beta)]}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} \\
&= \frac{\frac{1 - \theta_n}{\theta_n}[\theta_n\beta + \beta - 1 + \theta_n]}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} + \frac{\tau(1 + \alpha)[\theta_n\alpha - 1 + \theta_n + \beta - \theta_n\beta]}{(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)} \\
&= \frac{[-(1 + \beta)\theta_n^2 + 2\theta_n + \beta - 1] + \theta_n\tau(1 + \alpha)[\theta_n(1 + \alpha) - 1 + \beta - \theta_n\beta]}{\theta_n((1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha))} \\
&\leq \frac{[-(1 + \beta)\theta_n^2 + 2\theta_n + \beta - 1] + [\frac{1}{\tau}\theta_n^2 - \theta_n + \beta\theta_n - \beta\theta_n^2]}{\theta_n((1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha))} \\
&= \frac{(\frac{1}{\tau} - 1 - 2\beta)\theta_n^2 + (1 + \beta)\theta_n + \beta - 1}{\theta_n((1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha))}
\end{aligned} \tag{4.14}$$

Since  $\theta_n \leq \frac{-1 - \beta + \sqrt{(1 + \beta)^2 - 4(\frac{1}{\tau} - 1 - 2\beta)(\beta - 1)}}{2(\frac{1}{\tau} - 1 - 2\beta)}$ , we get  $\delta_n \leq \sigma_n$ . From  $\theta_{n-1} \leq \theta_n$ , it follows that  $\sigma_n \leq \sigma_{n-1}$ . By (4.13), we have

$$\begin{aligned}
\|x_{n+1} - p\|^2 + \sigma_n\|x_{n+1} - x_n\|^2 &\leq [(1 - \theta_n)(1 + \beta) + \theta_n\tau(1 + \alpha)](\|x_n - p\|^2 + \sigma_n\|x_n - x_{n-1}\|^2) \\
&= [(1 + \beta) + \theta_n[\tau(1 + \alpha) - (1 + \beta)](\|x_n - p\|^2 + \sigma_n\|x_n - x_{n-1}\|^2) \\
&< [(1 + \beta) + \theta[\tau(1 + \alpha) - (1 + \beta)](\|x_n - p\|^2 + \sigma_{n-1}\|x_n - x_{n-1}\|^2),
\end{aligned} \tag{4.15}$$

where the last inequality follows from  $\sigma_n \leq \sigma_{n-1}$  and  $\alpha < \frac{1-\tau}{\tau} - \frac{\beta}{\tau}$ .

Since  $\frac{\beta}{(1+\beta)-\tau(1+\alpha)} < \theta$ , we have  $(1+\beta) + \theta[\tau(1+\alpha) - (1+\beta)] \in (0, 1)$ . Therefore we get

$$\|x_{n+1} - p\|^2 + \sigma_n \|x_{n+1} - x_n\|^2 \leq [(1+\beta) + \theta[\tau(1+\alpha) - (1+\beta)]]^{n-1} (\|x_2 - p\|^2 + \sigma_1 \|x_2 - x_1\|^2). \quad (4.16)$$

This implies that

$$\|x_{n+1} - p\|^2 \leq [(1+\beta) + \theta[\tau(1+\alpha) - (1+\beta)]]^{n-1} (\|x_2 - p\|^2 + \sigma_1 \|x_2 - x_1\|^2). \quad (4.17)$$

Hence, the proof is completed.

*Remark 4.4* Theorem 6.2 of [29] gives a linear convergence rate result of the subgradient extragradient method with double inertial steps for solving variational inequalities. However, it only discuss the single inertial case in [29]. As far as we know, algorithms with the double inertial extrapolation steps for solving variational inequalities and monotone inclusions have no linear convergent result. Hence Theorem 4.2 is a new result.

## 5 Numerical experiments

In this section, we provide some numerical examples to show the performance of our Algorithm 3.1 (shortly Alg1), and compare it with others, including Abubakar et al's Algorithm 3.1 (shortly AKHAlg1) [1], Cholamjiak et al's Algorithm 3.2(CHCAlg2) [5], Cholamjiak et al's Algorithm 3.1 (CHMAlg1) [6], Hieu et al's Algorithm 3.1 (HAMAlg1) [26] and Yao et al's Algorithm 1 (YISAlg1) [29].

All the programs were implemented in MATLAB R2021b on Intel(R) Core(TM) i7-7700HQ CPU@ 2.80GHZ computer with RAM 8.00GB. We denote the number of iterations by "Iter." and the CPU time seconds by "CPU(s)".

*Example 5.1* We consider the signal recovery in compress sensing. This problem can be modeled as

$$y = Ax + \varepsilon. \quad (5.1)$$

where  $y \in \mathbb{R}^M$  is observed or measured data,  $A : \mathbb{R}^N \rightarrow \mathbb{R}^M$  is bounded linear operator,  $x \in \mathbb{R}^N$  is a vector with  $K$  ( $K \ll N$ ) nonzero components and  $\varepsilon$  is the noise. It is known that the problem (5.1) can be viewed as the LASSO problem [5]

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 \quad (\lambda > 0). \quad (5.2)$$

The minimization problem 5.2 is equivalent to the following monotone inclusion problem

$$\text{find } x \in \mathbb{R}^N \text{ such that } 0 \in (B + C)x. \quad (5.3)$$

where  $B = A^T(Ax - y)$  and  $C = \partial(\lambda\|x\|_1)$ . In this experiment, the vector  $x \in \mathbb{R}^N$  is from uniform distribution in the interval  $[-1, 1]$ .

The matrix  $A \in \mathbb{R}^{M \times N}$  is produced by a normal distribution with mean zero and one variance. The vector  $y$  is generated by Gaussian noise  $\varepsilon$  with variance  $10^{-4}$ . The initial points  $x_0$  and  $x_1$  are both zero. We use  $E_n = \|x_n - x_{n-1}\|$  to measure the restoration accuracy. And the stopping criterion is  $E_n \leq 10^{-5}$ .

In the first experiment we consider the influence of different  $\alpha_n$  and  $\beta_n$  on the performance of our algorithm. We take  $\mu = 0.9$ ,  $\theta_n = 0.45$ ,  $\lambda_1 = 0.1$ ,  $\mu_n = 0$  and  $p_n = \frac{1}{n^2}$ .

It can be seen from Table 1 that Algorithm 3.1 with double inertial extrapolation steps, that is,  $\beta_n \neq 0$  outperform the one with single inertia, moreover, the increase of  $\alpha_n$  and  $\beta_n$  significantly improves the convergence speed of the algorithm. This implies that it is important to investigate double inertial methods from both theoretical and numerical viewpoint.

In second experiment we compared the performance of our algorithm with other algorithms. The following two cases are considered

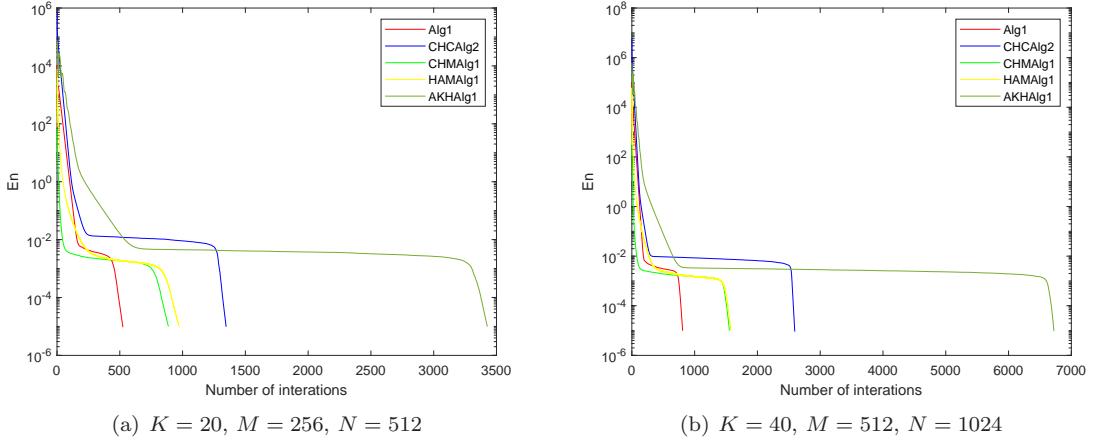
**Case 1:**  $K = 20$ ,  $M = 256$ ,  $N = 512$ ;

Table 1: The performances of Algorithm 3.1 for different values of  $\alpha_n$  and  $\beta_n$  in Example 5.1

Iter. \ \begin{array}{c} \alpha_n \\ \diagdown \\ \beta_n \end{array}	0.2	0.4	0.6	0.8	0.9	1
0	966	872	777	681	632	584
0.02	954	859	764	668	620	572
0.04	942	848	752	656	608	559
0.06	930	836	740	644	596	547
0.08	918	823	728	632	583	535
0.1	906	811	716	620	571	522

Table 2: Numerical results for Example 5.1

Algorithms	Case 1		Case 2	
	Iter.	CPU(s)	Iter.	CPU(s)
Alg1	525	0.2447	809	1.2297
CHCAlg2	1347	0.5340	2595	3.0049
CHMAlg1	887	0.3496	1554	1.9651
HAMAlg1	971	0.3904	1577	1.5138
AKHAlg1	3427	1.3142	6723	6.4723

Fig. 1: Numerical behavior of  $E_n$  for Example 5.1

**Case 2:**  $K = 40, M = 512, N = 1024$ .

The parameters for algorithms are chosen as

Alg1:  $\mu = 0.9, \alpha_n = 1 - \frac{1}{10^n}, \beta_n = 0.1 - \frac{1}{1000+n}, \theta_n = 0.45 - \frac{1}{1000+n}, \lambda_1 = 0.1, \mu_n = \frac{1}{n^2}$  and  $p_n = \frac{1}{n^2}$ ;

CHCAlg2:  $\mu = 0.9, \alpha = 0.1, \theta = 1$  and  $\lambda_1 = 1$ ;

CHMAlg1:  $\mu = 0.4, \alpha_1 = 0.01, \alpha_2 = 0.02$  and  $\lambda_0 = 0.01$ ;

HAMAlg1:  $\mu = 0.4, \lambda_{-1} = \lambda_0 = 0.1$ ;

AKHAlg1:  $\mu = 0.3, \rho = 0.1, \varrho = 0.9$  and  $\lambda_0 = 1$ .

Fig.1 and Table.2 show the numerical results of our algorithm and other algorithms in two cases respectively. We give the graphs of original signal and recovered signal in Fig. 2.

*Remark 5.1* It can be seen from Fig.1 and Table 2 that the number of iterations and CPU time of our algorithm are better than other algorithms, indicating that our algorithm has better performance.

*Example 5.2* We consider our algorithm to solve the variational inequality problem . The operator  $A : R^m \rightarrow R^m$  is defined by  $A(x) := Mx + q$  with  $M = NN^T + S + D$  and  $q \in R^m$ , where  $N$  is an  $m \times m$  matrix,  $S$  is an  $m \times m$  skew-symmetric matrix and  $D$  is an  $m \times m$  diagonal matrix. All entries of  $N$  and  $S$  are uniformly generated from  $(-5, 5)$  and all diagonal entries of  $D$  are uniformly generated from

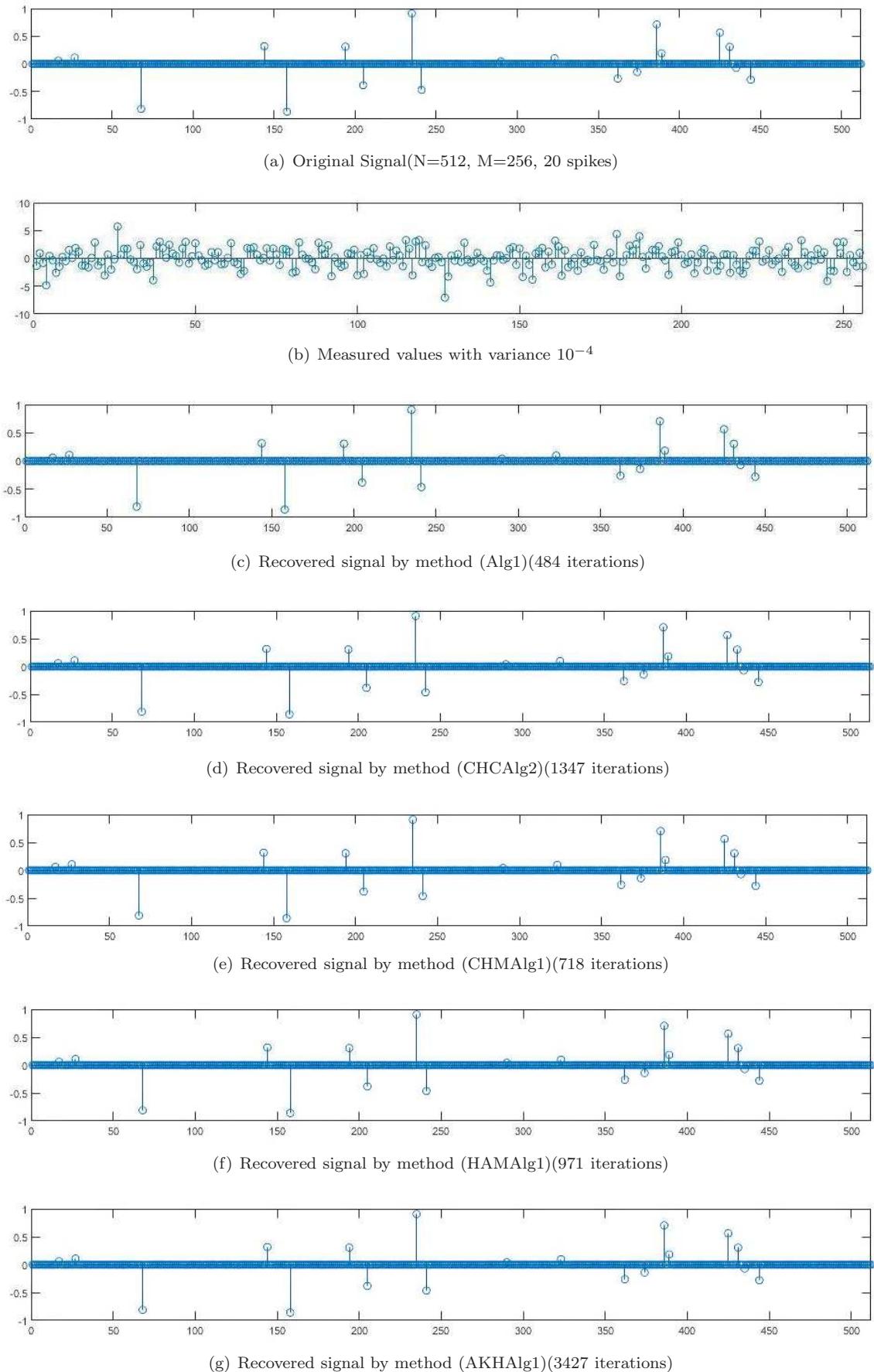


Fig. 2: From top to bottom: original signal, observation data, recovered signal by the methods (Alg1), (CHCAlg2), (CHMAlg1), (HAMAlg1) and (AKHAlg1) in Case 1, respectively

Table 3: The performances of Algorithm 3.1 for different values of  $\theta_n$  in Example 5.2

$\theta_n$	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45
Iter.	16988	8521	5562	4035	3095	2454	1987	1360	1346
CPU(s)	1.8752	0.9888	0.6785	0.4804	0.3879	0.2910	0.2436	0.2246	0.1674

Table 4: Numerical results for Example 5.2

Algorithms	$m = 50$		$m = 100$		$m = 150$		$m = 200$	
	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)
Alg1	448	0.0191	642	0.0380	759	0.0968	1012	0.1681
CHCAlg2	723	0.0224	1048	0.0499	1234	0.1105	1644	0.2128
YISAlg1	600	0.0917	963	0.0635	1214	0.1374	1595	0.2125
HAMAlg1	779	0.0203	1142	0.0475	1355	0.1410	1751	0.1940
AKHAlg1	1049	0.0267	1425	0.0659	1691	0.1625	2371	0.2839
CHMAlg1	871	0.0331	1182	0.0637	1408	0.1366	1879	0.2158

(0, 0.3). It is easy to see that  $M$  is positive definite. Define the feasible set  $C := R_m^+$  and use  $E_n = \|x_n\|$  to measure the accuracy, and the stoping criterion is  $E_n \leq 10^{-3}$ .

In the first experiment we consider the effect of relaxation coefficients  $\theta_n$  on the performance of Algorithm 3.1. We choose  $\mu = 0.9$ ,  $\alpha_n = 1$ ,  $\beta_n = 0.1$ ,  $\lambda_1 = 0.1$ ,  $\mu_n = 0$  and  $p_n = \frac{1}{n^2}$ .

Table 3 shows that the performance comparison of Algorithm 3.1 for different values of  $\theta_n$ . It can be seen that the performance of the algorithm becomes better with the increase of the relaxed parameter  $\theta_n$ . This indicates that the increase of the relaxation coefficient value range is of great significance to improve the performance of the algorithm.

In second experiment we compared the performance of our algorithm with other algorithms. The following cases are considered

**Case 1:**  $m = 50$ ;    **Case 2:**  $m = 100$ ;    **Case 3:**  $m = 150$ ;    **Case 4:**  $m = 200$ .

The parameters for algorithms are chosen as

Alg1:  $\mu = 0.9$ ,  $\alpha_n = 1 - \frac{1}{10^n}$ ,  $\beta_n = 0.1 - \frac{1}{1000+n}$ ,  $\theta_n = 0.45 - \frac{1}{1000+n}$ ,  $\mu_n = 0$  and  $p_n = \frac{1}{n^2}$ ;

HAMAlg1:  $\mu = 0.4$ ,  $\lambda_{-1} = \lambda_0 = 0.3$ ;

CHCAlg2:  $\mu = 0.9$ ,  $\alpha = 0.3$ ,  $\theta = 0.4$  and  $\lambda_1 = 1$ ;

YISAlg1:  $\mu = 0.9$ ,  $\alpha_n = 0.2903$ ,  $\delta = 0.0241$ ,  $\theta_n = 1$  and  $\lambda_1 = 0.1$ ;

AKHAlg1:  $\mu = 0.3$ ,  $\rho = 0.1$ ,  $\varrho = 0.9$  and  $\lambda_0 = 1$ ;

CHMAlg1:  $\mu = 0.4$ ,  $\alpha_1 = 0.01$ ,  $\alpha_2 = 0.02$  and  $\lambda_0 = 1$ .

Fig.3 and Table.4 show the performance comparison of our algorithm with other algorithms in four cases respectively.

*Remark 5.2* It can be seen that our algorithm performs better than other algorithms for such problems as Example 5.2 from Fig.3 and Table 4.

*Example 5.3* Let  $H := L_2([1, 2])$  with the norm

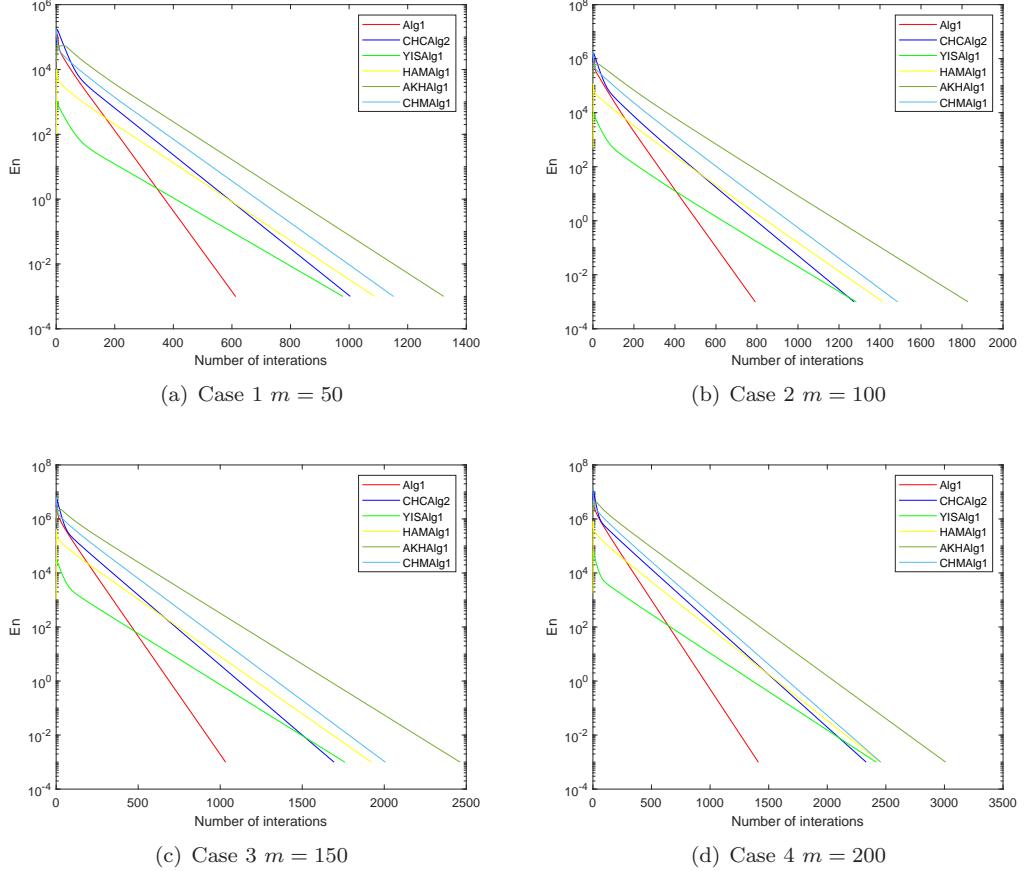
$$\|x\| := \left( \int_0^1 x(t)^2 dt \right)^{\frac{1}{2}}.$$

and the inner product

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

Let  $C := \{x \in L^2([0, 1]) : \int_0^1 tx(t)dt = 2\}$  and define  $A : L^2([0, 1]) \rightarrow L^2([0, 1])$  by

$$Ax(t) := \max\{x(t), 0\}, x \in L^2([0, 1]), t \in [0, 1].$$

Fig. 3: Numerical behavior of  $E_n$  for Example 5.2

It is clear that  $A$  is monotone and Lipschitz with  $L = 1$ . The orthogonal projection onto  $C$  have the following explicit formula

$$P_C(x)(t) := x(t) - \frac{\int_0^1 tx(t)dt - 2}{\int_0^1 t^2 dt} t.$$

The example is taken from [29]. Use  $E_n = \|x_{n+1} - x_n\|$  to measure the accuracy. The stoping criterion is  $E_n \leq 10^{-4}$ .

The following cases are considered

$$\text{Case 1: } x_0 = \frac{97t^2 + 4t}{13}, x_1 = \frac{t^2 - e^{-7t}}{250};$$

$$\text{Case 2: } x_0 = \frac{97t^2 + 4t}{13}, x_1 = \frac{\sin(3t) + \cos(10t)}{100};$$

$$\text{Case 3: } x_0 = \frac{t^2 - e^{-7t}}{250}, x_1 = \frac{\sin(3t) + \cos(10t)}{100};$$

$$\text{Case 4: } x_0 = \frac{\sin(3t) + \cos(10t)}{100}, x_1 = \frac{97t^2 + 4t}{13}.$$

The parameters for algorithms are chosen as

$$\text{Alg1: } \mu = 0.4, \alpha_n = 1 - \frac{1}{10^n}, \beta_n = 0.1 - \frac{1}{1000+n}, \theta_n = 0.45 - \frac{1}{1000+n} \text{ and } p_n = \frac{1}{n^2};$$

$$\text{CHCAlg2: } \mu = 0.4, \alpha = 0.3, \theta = 0.4 \text{ and } \lambda_1 = 1;$$

$$\text{HAMAlg1: } \mu = 0.4, \lambda_{-1} = \lambda_0 = 0.1;$$

$$\text{YISAlg1: } \mu = 0.4, \alpha_n = 0.2250, \delta = 0.4950, \theta_n = 1 \text{ and } \lambda_1 = 1.1;$$

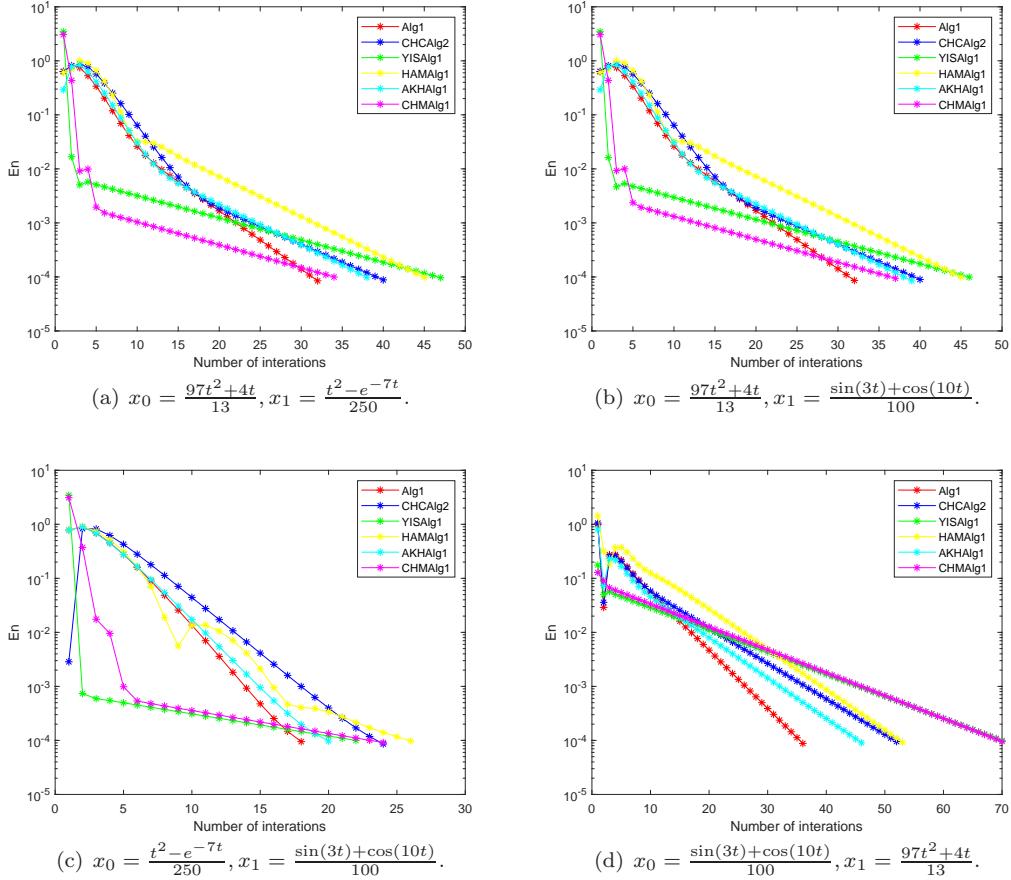
$$\text{AKHAlg1: } \mu = 0.4, \rho = 0.45, \varrho = 0.3 \text{ and } \lambda_0 = 0.5;$$

$$\text{CHMAlg1: } \mu = 0.4, \alpha_1 = 0.01, \alpha_2 = 0.02 \text{ and } \lambda_0 = 1.$$

*Remark 5.3* From Table 5 and Fig.4, we observe that our Algorithm 3.1 performs better and converges faster.

Table 5: Numerical results for Example 5.3

Algorithms	Case1		Case2		Case3		Case4	
	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)
Alg1	32	0.0554	32	0.0499	18	0.0296	36	0.0416
CHCAlg2	40	0.0601	40	0.0616	24	0.0361	52	0.0664
VAMAlg1	47	0.0616	46	0.0558	22	0.0311	70	0.0974
YISAlg1	45	0.0894	45	0.0629	26	0.0457	53	0.0912
AKHAlg1	38	0.0746	39	0.0506	20	0.0277	46	0.0697
HAMAlg1	34	0.0651	37	0.0498	24	0.0343	70	0.0802

Fig. 4: Numerical behavior of  $E_n$  for Example 5.3

## 6 Conclusions

In this paper, we propose a new Tseng splitting method with double inertial extrapolation steps for solving monotone inclusion problems in real Hilbert spaces and establish the weak convergence, nonasymptotic  $O(\frac{1}{\sqrt{n}})$  convergence rate, strong convergence and linear convergence rate of the proposed algorithm, respectively. Our method has the following advantages:

- (i) Our method uses adaptive step sizes, which can be updated by a simple calculation without knowing the Lipschitz constant of the underlying operator.
- (ii) Our method own double inertial extrapolation steps, in which inertial factor  $\alpha_n$  can equal 1. This is not allowed in the corresponding algorithms of [5, 6], where the only single inertial extrapolation step is considered and the inertial factor is bounded away from 1. From Table 1 in section 5, it can be seen Algorithm 3.1 with double inertial extrapolation steps outperforms the one with the single inertia.
- (iii) Our method includes the corresponding methods considered in [1, 5, 29] as special cases. Especially, when our algorithm is used to solve variational inequalities, the relaxed parameter sequence  $\{\theta_n\}$  have

larger choosing interval than the ones of [29]. Via Table 3 in section 5, we observe that the performance of the algorithm becomes better with the increase of the relaxed parameter  $\theta_n$ .

(iv) To the best of our knowledge, there are few available convergence rate results for algorithms with the double inertial extrapolation steps for solving variational inequalities and monotone inclusions. From numerical experiments in section 5, we can see that our algorithm has better efficiency than the corresponding algorithms in [1, 5, 6, 26, 29].

**Acknowledgements** 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 (2682021ZTPY040).

## References

1. Abubakar, J., Kumam, P., Hassan Ibrahim, A., Padcharoen, A.: Relaxed inertial Tsengs type method for solving the inclusion problem with application to image restoration. *Mathematics*. **8**, 818 (2020)
2. Alvarez, F., Attouch, H.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. *Set-Valued Analysis*. **9**, 3-11 (2001)
3. Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces. Springer, New York (2011)
4. Cai, G., Dong, Q.L., Peng, Y.: Strong convergence theorems for inertial Tseng's extragradient method for solving variational inequality problems and fixed point problems. *Optim. Lett.* **15**, 1457-1474 (2021)
5. Cholamjiak, P., Hieu, D.V., Cho, Y.J.: Relaxed forward-backward splitting methods for solving variational inclusions and applications. *J. Sci. Comput.* **88**, 85 (2021)
6. Cholamjiak, P., Hieu, D.V., Muu, L.D.: Inertial splitting methods without prior constants for solving variational inclusions of two operators. *Bull. Iran. Math. Soc.* (2022).<https://doi.org/10.1007/s41980-022-00682-3>
7. Combettes, P.L., Wajs, V.: Signal recovery by proximal forward-backward splitting. *SIAM Multiscale Model Simul.* **4**, 1168-1200 (2005)
8. Çopur AK, Hacıoglu E, Gürsoy F, et al.: An efficient inertial type iterative algorithm to approximate the solutions of quasi variational inequalities in real Hilbert spaces. *J. Sci. Comput.* **89**:50(2022)
9. Dong, Q., Jiang, D., Cholamjiak, P.: A strong convergence result involving an inertial forward-backward algorithm for monotone inclusions. *J. Fixed Point Theory Appl.* **19**, 3097-3118 (2017)
10. Duchi, J., Singer, Y.: Efficient online and batch learning using forward-backward splitting. *J. Mach. Learn Res.* **10**, 2899-2934 (2009)
11. Gibali, A., Thong, D.V.: Tseng type methods for solving inclusion problems and its applications. *Calcolo* **55**, 55:49 (2018)
12. Linh, N.X., Thong, D.V., Cholamjiak, P.: Strong convergence of an inertial extragradient method with an adaptive nondecreasing step size for solving variational inequalities. *Acta Math. Sci.* **42**, 795-812 (2022)
13. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964-979 (1979)
14. Liu, Q.: A convergence theorem of the sequence of Ishikawa iterates for quasi-contractive mappings. *J. Math. Anal. Appl.* **146**, 301-305 (1990)
15. Maingé, P. E.: Convergence theorems for inertial KM-type algorithms. *J. Comput. Appl. Math.* **219**, 223-236 (2008)
16. Opial, Z.: Weak convergence of the sequence of successive approximations for nonexpansive mappings. *Bull. Am. Math. Soc.* **73**, 591-597 (1967)
17. Padcharoen, A., Kitkuan, D., Kumam, W., Kumam, P.: Tseng methods with inertial for solving inclusion problems and application to image deblurring and image recovery problems. *Comput. Math. Methods* (2020). <https://doi.org/10.1002/cmm4.1088>
18. Passty, G.B.: Ergodic convergence to a zero of the sum of monotone operators in Hilbert space. *J. Math. Anal. Appl.* **72**, 383-390 (1979)
19. Raguet, H., Fadili, J., Peyré, G.: A generalized forward-backward splitting. *SIAM J. Imaging Sci.* **6**, 1199-1226 (2013)
20. Shehu, Y., Iyiola, O.S., Li, X.H., et al.: Convergence analysis of projection method for variational inequalities. *Comput. Appl. Math.* **38**, 161 (2019)
21. Shehu, Y., Iyiola, O.S., Reich, S.: A modified inertial subgradient extragradient method for solving variational inequalities. *Optimization and Engineering*, **23**, 421-449.(2021)
22. Tan, K.K., Xu, H.K.: Approximating fixed points of nonexpansive mappings by the Ishikawa iteration process. *J. Math. Anal. Appl.* **178**, 301-308 (1993)
23. Thong, D.V., Vinh, N.T. Cho, Y.J.: A strong convergence theorem for Tseng's extragradient method for solving variational inequality problems. *Optim. Lett.* **14**, 1157-1175 (2020)
24. Thong, D.V., Yang, J., Cho, Y.J., Rassias, T.M.: Explicit extragradient-like method with adaptive stepsizes for pseudo-monotone variational inequalities. *Optim. Lett.* **15**, 2181-2199 (2021)
25. Tseng, P.: A modified forward-backward splitting method for maximal monotone mapping. *SIAM J. Control Optim.* **38**, 431-446 (2000)

26. Van Hieu, D., Anh, P.K. Muu, L.D.: Modified forward-backward splitting method for variational inclusions. *4OR-Q J. Oper. Res.* **19**, 127-151 (2021)
27. Wang, Z.B., Chen, X., Jiang, Y., et al.: Inertial projection and contraction algorithms with larger step sizes for solving quasimonotone variational inequalities. *J. Glob. Optim.* **82**:499-522 (2022)
28. Xu, H.K.: Iterative algorithms for nonlinear operators. *J. Lond. Math. Soc.* **66**, 240-256 (2002)
29. Yao, Y., Iyiola, O.S. Shehu, Y.: Subgradient extragradient method with double inertial steps for variational inequalities. *J. Sci. Comput.* **90**, 71 (2022)