# Convergence of Relaxed Inertial Subgradient Extragradient Methods for Quasimonotone Variational Inequality Problems

G. N. Ogwo[1] · C. Izuchukwu[1,2] · Y. Shehu[3] · O. T. Mewomo[1]

## Abstract

In this paper, we present two new relaxed inertial subgradient extragradient methods for solving variational inequality problems in a real Hilbert space. We establish the convergence of the sequence generated by these methods when the cost operator is quasimonotone and Lipschitz continuous, and when it is Lipschitz continuous without any form of monotonicity. The methods combine both the inertial and relaxation techniques in order to achieve high convergence speed, and the techniques used are quite different from the ones in most papers for solving variational inequality problems. Furthermore, we present some experimental results to illustrate the profits gained from the relaxed inertial steps.

✉ Y. Shehu
 yekini.shehu@zjnu.edu.cn

 G. N. Ogwo
 graceogwo@aims.ac.za; 219095374@stu.ukzn.ac.za

 C. Izuchukwu
 izuchukwuc@ukzn.ac.za; izuchukwu_c@yahoo.com

 O. T. Mewomo
 mewomoo@ukzn.ac.za

[1] School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

[2] DSI-NRF Center of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), Johannesburg, South Africa

[3] Department of Mathematics, Zhejiang Normal University, Jinhua 321004, People's Republic of China

## 1 Introduction

Let $\mathcal{C}$ be a nonempty closed and convex subset of a real Hilbert space $\mathcal{H}$ with inner product $\langle \cdot , \cdot \rangle$, and associated norm $|| \cdot ||$. Let $A : \mathcal{H} \to \mathcal{H}$ be a continuous mapping. The Variational Inequality Problem (VIP) is formulated as:

$$\text{Find } x \in \mathcal{C} \text{ such that } \langle Ax, y - x \rangle \geq 0, \ \forall y \in \mathcal{C}. \tag{1.1}$$

The VIP is a fundamental problem in optimization and known to have numerous applications in diverse fields (see, for example [16,17,19,20,24,26,27] and the references therein). Its theory provides a simple, natural and unified framework for a general treatment of unrelated problems. The VIP was first considered for the purpose of modeling problems in mechanics by Stampacchia [41] (also independently by Fichera [16]).

One of the simplest methods for solving VIP (1.1) is the following gradient-projection method:

$$x_{n+1} = P_{\mathcal{C}}(x_n - \lambda Ax_n), n \geq 1. \tag{1.2}$$

This method converges strongly to a unique solution of the VIP (1.1) if $A$ is $\eta$-strongly monotone and $L$-Lipschitz continuous, with $\lambda \in \left(0, \frac{2\eta}{L^2}\right)$. However, if $A$ is monotone and Lipschitz continuous, the gradient-projection method may fail to converge. For instance, take $C = \mathbb{R}^2$ and $A$ to be a rotation with $\frac{\pi}{2}$ angle. Then, $A$ is monotone and Lipschitz continuous with $(0, 0)$ being the unique solution of (1.1). But the sequence $\{x_n\}$ generated by (1.2) satisfies $||x_{n+1}|| > ||x_n||$, $\forall n$. Thus, the gradient-projection method does not always work for monotone and Lipschitz continuous operators.

In [28], Korpelevich proposed a method in finite dimensional spaces which converges to a solution of the VIP (1.1) when the operator is monotone and Lipchitz continuous. This method which is known as the extragradient method is given as:

$$\begin{cases} x_1 \in \mathcal{C}, \\ y_n = P_{\mathcal{C}}(x_n - \lambda_n Ax_n) \\ x_{n+1} = P_{\mathcal{C}}(x_n - \lambda_n Ay_n), n \geq 1, \end{cases} \tag{1.3}$$

where $\lambda_n \in \left(0, \frac{1}{L}\right)$. Since the introduction of the extragradient method, many authors have studied it in infinite dimensional spaces when $A$ is monotone (see [5,9,19]) and when $A$ is pseudomonotone (see [36,48]).

However, the major drawback of this method is that it requires two projections onto the feasible set $\mathcal{C}$ per iteration, which can be costly if $\mathcal{C}$ has complex structure. In this case, $P_{\mathcal{C}}$ may not have a closed form formula, and a minimization problem has to be solved twice per iteration in implementing (1.3). Hence, the efficiency of the method may be affected.

To overcome this setback, Censor et al. [12] introduced the following subgradient extragradient method: $x_1 \in \mathcal{H}$,

$$\begin{cases} y_n = P_{\mathcal{C}}(x_n - \lambda_n Ax_n) \\ T_n := \{w \in \mathcal{H} : \langle x_n - \lambda_n Ax_n - y_n, w - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \lambda_n Ay_n), n \geq 1. \end{cases} \tag{1.4}$$

They proved that $\{x_n\}$ generated by (1.4) converges weakly to a solution of problem (1.1) when $A$ is monotone and Lipschitz continuous.

Unlike (1.3), Algorithm (1.4) requires only one projection onto $\mathcal{C}$ per iteration since the second projection is onto a half space $T_n$, which has an explicit formula. Hence, subgradient extragradient methods are less computationally expensive than extragradient methods. This method has also been considered by many authors for solving the VIP (1.1) in infinite dimensional spaces, when $A$ is monotone [1,10,11,25,30] and when $A$ is pseudomonotone [45,51].

In recent years, algorithms with fast convergence rate have been of great interest. To accelerate the convergence of iterative methods for solving (1.1) and other related optimization problems, there exist different modifications of such iterative methods. In this paper, we consider the two very important modifications, namely; inertial and relaxation techniques. The inertial technique is based upon a discrete analogue of a second order dissipative dynamical system and is known for its efficiency in improving the convergence rate of iterative methods. The method was first considered by Polyak [38] for solving the smooth convex minimization problems. It was later made very popular by Nesterov's acceleration gradient method [35], and was further developed by Beck and Teboulle in the case of structured convex minimization problem [6]. Since then, many authors have employed the use of inertial techniques for improving the convergence of their iterative methods (see for example [14,15,18,21,31,32,34,37–40,46,47] and the references therein). On the other hand, the relaxation technique has proven to be an essential ingredient in the resolution of optimization problems due to the improved convergence rate that it contribute to iterative schemes. Moreover, both inertial and relaxation techniques naturally come from an explicit discretization of a dynamical system in time (see, for example [4,49]).

Some authors have considered incorporating these two techniques into known methods in order to achieve high convergence speed of the resulting methods (see, [3,4]). Also, in [23], the influence of inertial and relaxation techniques on the numerical performance of iterative schemes was studied.

Given the importance of these two techniques (inertial and relaxation) for solving optimization problems, our aim in this paper is to incorporate the inertial and relaxation techniques into Algorithm (1.4), when $A$ is not necessarily pseudomonotone. That is, we design two new relaxed inertial subgradient extragradient methods, and prove that they converge weakly to a solution of VIP (1.1) when the operator $A$ is quasimonotone and Lipschitz continuous, and when it is Lipschitz continuous without any form of monotonicity. The techniques employed in this paper are quite different from the ones used in most papers (see for example [14,15,22,31,32,34,37–40,44,46,47]). Moreover, the assumptions on the inertial and relaxation factors in this paper, are weaker than those in many papers for solving VIPs. Finally, we provide some numerical implementations of our methods and compare them with other methods, in order to show the profits gained from the inertial and relaxation techniques.

The rest of the paper is organized as follows: Sect. 2 contains basic definitions and results needed in subsequent sections. In Sect. 3, we present and discuss the proposed methods. The convergence analysis of these methods are investigated in Sect. 4. In Sect. 5, we perform some numerical analysis of our methods in comparison with other methods in the literature. We then conclude in Sect. 6.

## 2 Preliminaries

In this section, we recall some concepts and results needed in subsequent sections. Henceforth, we denote the weak convergence of $\{x_n\}$ to a point $x^*$ by $x_n \rightharpoonup x^*$.

Let $\mathcal{H}$ be a real Hilbert space and $A : \mathcal{H} \to \mathcal{H}$ be a mapping. Then, $A$ is said to be

(i) *L−Lipschitz continuous*, if there exists $L > 0$ such that

$$\|Ax - Ay\| \le L\|x - y\|, \ \forall x, y \in \mathcal{H},$$

(ii) *L-strongly monotone*, if there exists $L > 0$ such that

$$\langle Ax - Ay, x - y \rangle \ge L\|x - y\|^2, \ \forall x, y \in \mathcal{H},$$

(iii) *Monotone*, if

$$\langle Ax - Ay, x - y \rangle \ge 0, \ \forall x, y \in \mathcal{H},$$

(iv) *Pseudomonotone*, if

$$\langle Ay, x - y \rangle \ge 0 \implies \langle Ax, x - y \rangle \ge 0, \ \forall x, y \in \mathcal{H},$$

(v) Quasimonotone, if

$$\langle Ay, x - y \rangle > 0 \implies \langle Ax, x - y \rangle \ge 0, \ \forall x, y \in \mathcal{H},$$

(vi) *Sequentially weakly-strongly continuous,* if for every sequence $\{x_n\}$ that converges weakly to a point $x$, the sequence $\{Ax_n\}$ converges strongly to $Ax$,

(vii) *Sequentially weakly continuous,* if for every sequence $\{x_n\}$ that converges weakly to a point $x$, the sequence $\{Ax_n\}$ converges weakly to $Ax$.

Clearly, $(ii) \implies (iii) \implies (iv) \implies (v)$. But the converses are not always true.
Let $VI(\mathcal{C}, A)$ be the solution set of problem (1.1) and $\Gamma$ be the solution set of the following problem:

$$\text{Find } x \in \mathcal{C} \text{ such that } \langle Ay, y - x \rangle \ge 0, \ \forall y \in \mathcal{C}. \tag{2.1}$$

Then, $\Gamma$ is a closed and convex subset of $\mathcal{C}$, and since $\mathcal{C}$ is convex and $A$ is continuous, we have the following relation

$$\Gamma \subseteq VI(\mathcal{C}, A). \tag{2.2}$$

**Lemma 2.1** [50] *Let $\mathcal{C}$ be a nonempty closed and convex subset of $\mathcal{H}$. If either*

(1) *$A$ is pseudomonotone on $\mathcal{C}$ and $VI(\mathcal{C}, A) \ne \emptyset$,*
(2) *$A$ is the gradient of $G$, where $G$ is a differential quasiconvex function on an open set $\mathcal{K} \supset \mathcal{C}$ and attains its global minimum on $\mathcal{C}$,*
(3) *$A$ is quasimonotone on $\mathcal{C}$, $A \ne 0$ on $\mathcal{C}$ and $\mathcal{C}$ is bounded,*
(4) *$A$ is quasimonotone on $\mathcal{C}$, $A \ne 0$ on $\mathcal{C}$ and there exists a positive number $r$ such that, for every $x \in \mathcal{C}$ with $\|x\| \ge r$, there exists $y \in \mathcal{C}$ such that $\|y\| \le r$ and $\langle Ax, y - x \rangle \le 0$,*
(5) *$A$ is quasimonotone on $\mathcal{C}$, int $\mathcal{C}$ is nonempty and there exists $x^* \in VI(\mathcal{C}, A)$ such that $Ax^* \ne 0$.*

*Then, $\Gamma$ is nonempty.*

Recall that for a nonempty closed and convex subset $\mathcal{C}$ of $\mathcal{H}$, the metric projection denoted by $P_{\mathcal{C}}$ (see [43]), is a map defined on $\mathcal{H}$ onto $\mathcal{C}$ which assigns to each $x \in \mathcal{H}$, the unique point in $\mathcal{C}$, denoted by $P_{\mathcal{C}}x$ such that

$$\|x - P_{\mathcal{C}}x\| = \inf\{\|x - y\| : \ y \in \mathcal{C}\}.$$

It is well known that $P_C$ is characterized by the inequality

$$\langle x - P_C x, y - P_C x \rangle \leq 0, \quad \forall y \in C. \tag{2.3}$$

Following Attouch and Cabot [3, pages 5, 10], we note that if $x_{n+1} = x_n + \theta_n(x_n - x_{n-1})$, then for all $n \geq 1$, we have that

$$x_{n+1} - x_n = \left( \prod_{j=1}^{n} \theta_j \right) (x_1 - x_0),$$

which implies that

$$x_n = x_1 + \left( \sum_{j=1}^{n-1} \prod_{j=1}^{l} \theta_j \right) (x_1 - x_0).$$

Thus, $\{x_n\}$ converges if and only if $x_1 = x_0$ or if $\sum_{l=1}^{\infty} \prod_{j=1}^{l} \theta_j < \infty$.

Therefore, we assume henceforth that

$$\sum_{l=i}^{\infty} \left( \prod_{j=i}^{l} \theta_j \right) < \infty \ \forall i \geq 1. \tag{2.4}$$

Hence, we can define the sequence $\{t_i\}$ in $\mathbb{R}$ by

$$t_i := \sum_{l=i-1}^{\infty} \left( \prod_{j=i}^{l} \theta_j \right) = 1 + \sum_{l=i}^{\infty} \left( \prod_{j=i}^{l} \theta_j \right), \tag{2.5}$$

with the convention $\prod_{j=i}^{i-1} \theta_j = 1 \ \forall i \geq 1$.

**Remark 2.2** (*See also* [3]).
Assumption (2.4) ensures that the sequence $\{t_i\}$ given by (2.5) is well-defined, and

$$t_i = 1 + \theta_i t_{i+1}, \ \forall i \geq 1. \tag{2.6}$$

The following proposition provides a criterion for ensuring assumption (2.4).

**Proposition 2.3** [3, Proposition 3.1] *Let $\{\theta_n\}$ be a sequence such that $\theta_n \in [0, 1)$ for every $n \geq 1$. Assume that*

$$\lim_{n \to \infty} \left( \frac{1}{1 - \theta_{n+1}} - \frac{1}{1 - \theta_n} \right) = c,$$

*for some $c \in [0, 1)$. Then,*

(i) *Assumption (2.4) holds, and $t_{n+1} \sim \frac{1}{(1-c)(1-\theta_n)}$ as $n \to \infty$.*
(ii) *The equivalence $1 - \theta_n \sim 1 - \theta_{n+1}$ holds true as $n \to \infty$. Hence, $t_{n+1} \sim t_{n+2}$ as $n \to \infty$.*

**Remark 2.4** Using Proposition 2.3, we can see that $\theta_n = 1 - \frac{\bar{\theta}}{n}$, $\bar{\theta} > 1$, is a typical example of a sequence satisfying assumption (2.4).

Indeed, we have that

$$\lim_{n\to\infty}\left(\frac{1}{1-\theta_{n+1}}-\frac{1}{1-\theta_n}\right)=\lim_{n\to\infty}\left(\frac{1}{\bar\theta}(n+1)-\frac{1}{\bar\theta}n\right)=\frac{1}{\bar\theta}\in[0,1),$$

which satisfies the assumption of Proposition 2.3. Hence by Proposition 2.3(i), assumption (2.4) holds.

It is worthy of note that the example $\theta_n = 1 - \frac{\bar\theta}{n}$, $\bar\theta > 1$, falls within the setting of Nesterov's extrapolation methods. In fact, many practical choices for $\theta_n$ satisfy assumption (2.4) (for instance, see [3,6,13,35]).

The corresponding finite sum expression for $\{t_i\}$ is defined for $i, n \geq 1$, by

$$t_{i,n}:=\begin{cases}\sum_{l=i-1}^{n-1}\left(\prod_{j=i}^{l}\theta_j\right)=1+\sum_{l=i}^{n-1}\left(\prod_{j=i}^{l}\theta_j\right), & i\leq n,\\ 0, & \text{otherwise.}\end{cases} \tag{2.7}$$

In the same manner, we have that $\{t_{i,n}\}$ is well-defined and

$$t_{i,n}=1+\theta_i t_{i+1,n}\ \forall i\geq 1,\ n\geq i+1. \tag{2.8}$$

The sequences $\{t_i\}$ and $\{t_{i,n}\}$ are very crucial to our convergence analysis. In fact, their effect can be seen in the following lemma which also plays a crucial role in establishing our convergence results.

**Lemma 2.5** [3, page 42, Lemma B.1] *Let $\{a_n\}$, $\{\theta_n\}$ and $\{b_n\}$ be sequences of real numbers satisfying*

$$a_{n+1}\leq\theta_n a_n+b_n\ \ \text{for every}\ n\geq 1.$$

*Assume that $\theta_n\geq 0$ for every $n\geq 1$.*

(a) *For every $n\geq 1$, we have*

$$\sum_{i=1}^{n}a_i\leq t_{1,n}a_1+\sum_{i=1}^{n-1}t_{i+1,n}b_i,$$

   *where the double sequence $\{t_{i,n}\}$ is defined by (2.7).*

(b) *Under (2.4), assume that the sequence $\{t_i\}$ defined by (2.5) satisfies $\sum_{i=1}^{\infty}t_{i+1}[b_i]_+ < \infty$. Then, the series $\sum_{i\geq 1}[a_i]_+$ is convergent, and*

$$\sum_{i=1}^{\infty}[a_i]_+\leq t_1[a_1]_++\sum_{i=1}^{\infty}t_{i+1}[b_i]_+\ ,$$

   *where $[t]_+:=\max\{t,0\}$ for any $t\in\mathbb{R}$.*

**Lemma 2.6** [3, page 7, Lemma 2.1] *Let $\{x_n\}$ be a sequence in $\mathcal{H}$ and $\{\theta_n\}$ be a sequence of real numbers. Given $z\in\mathcal{H}$, define the sequence $\{\Gamma_n\}$ by $\Gamma_n:=\frac{1}{2}\|x_n-z\|^2$. Then*

$$\Gamma_{n+1}-\Gamma_n-\theta_n(\Gamma_n-\Gamma_{n-1})=\frac{1}{2}(\theta_n+\theta_n^2)\|x_n-x_{n-1}\|^2+\langle x_{n+1}-w_n,w_n-z\rangle$$

$$+\frac{1}{2}\|x_{n+1}-w_n\|^2, \tag{2.9}$$

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

The following lemmas are well-known.

**Lemma 2.7** *Let $\{x_n\}$ be any sequence in $\mathcal{H}$ such that $x_n \rightharpoonup x$. Then,*

$$\liminf_{n\to\infty} \|x_n - x\| < \liminf_{n\to\infty} \|x_n - y\|, \forall y \neq x.$$

**Lemma 2.8** *Let $x, y \in \mathcal{H}$. Then,*

$$2\langle x, y \rangle = \|x\|^2 + \|y\|^2 - \|x - y\|^2 = \|x + y\|^2 - \|x\|^2 - \|y\|^2.$$

## 3 Proposed Methods

In this section, we present our methods and discuss their features. We begin with the following assumptions under which we obtain our convergence results.

**Assumption 3.1** Conditions on the inertial and relaxation factors:
Suppose that $\theta_n \in [0, 1)$ and $\rho_n \in (0, 1]$ for all $n \geq 1$ such that $\liminf_{n\to\infty} \rho_n > 0$. Assume that there exists $\epsilon \in (0, 1)$ such that for $n$ large enough,

$$\begin{cases} 2(1 - \epsilon)\frac{1-\rho_n}{2\rho_n}(1 - \theta_{n-1}) \geq \theta_n t_{n+1}\left(1 + \theta_n + 2\left[\frac{1-\rho_n}{2\rho_n}(1 - \theta_n) - \frac{1-\rho_{n-1}}{2\rho_{n-1}}(1 - \theta_{n-1})\right]_+\right), & \text{if } \rho_n \in (0, 1), \\ (1 - \epsilon)(1 - \theta_{n-1}) \geq \theta_n t_{n+1}\left(1 + \theta_n + \left[\theta_{n-1} - \theta_n\right]_+\right), & \text{if } \rho_n = 1. \end{cases}$$
(3.1)

**Assumption 3.2** We further make the following assumptions:

(a) $\Gamma \neq \emptyset$,
(b) $A$ is Lipschitz-continuous on $\mathcal{H}$ with constant $L > 0$,
(c) $A$ is sequentially weakly continuous on $\mathcal{C}$,
(d) $A$ is quasimonotone on $\mathcal{H}$,
(e) The set $\{z \in \mathcal{C} : Az = 0\} \setminus \Gamma$ is a finite set.

We now present the proposed methods of this paper.
When the Lipschitz constant $L$ is known, we present the following method for solving the VIP (1.1).

**Algorithm 3.3** *Relaxed inertial subgradient extragradient method with fixed stepsize.*

**Step 0:** *Choose sequences $\{\theta_n\}$ and $\{\rho_n\}$ such that $\theta_n \in [0, 1)$ and $\rho_n \in (0, 1]$ for all $n \geq 1$. Let $\lambda \in (0, \frac{1}{L})$ and $x_0, x_1 \in \mathcal{H}$ be given arbitrarily. Set $n := 1$.*
**Step 1:** *Given the current iterates $x_{n-1}$ and $x_n$  $(n \geq 1)$, compute*

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

*and*

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

*If $w_n = y_n$: STOP. Otherwise, go to **Step 2.***
**Step 2:** *Construct the half-space*

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

*Then, compute*

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

*and*

$$x_{n+1} = (1 - \rho_n)w_n + \rho_n z_n.$$

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

In situations where the Lipschitz constant $L$ is not known, we present the following method with adaptive stepsize for solving the VIP (1.1).

**Algorithm 3.4** *Relaxed inertial subgradient extragradient method with adaptive stepsize strategy.*

**Step 0:** *Choose sequences $\{\theta_n\}$ and $\{\rho_n\}$ such that $\theta_n \in [0, 1)$ and $\rho_n \in (0, 1]$ for all $n \geq 1$. Let $\lambda_1 > 0$, $\mu \in (0, 1)$ and $x_0, x_1 \in \mathcal{H}$ be given arbitrarily. Choose a nonnegative real sequence $\{d_n\}$ such that $\sum_{n=1}^{\infty} d_n < \infty$. Set $n := 1$.*
**Step 1:** *Given the current iterates $x_{n-1}$ and $x_n$  $(n \geq 1)$, compute*

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

*and*

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

*If $w_n = y_n$: STOP. Otherwise, go to **Step 2.***

**Step 2:** *Construct the half-space*

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

*Then, compute*

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

*and*

$$x_{n+1} = (1 - \rho_n)w_n + \rho_n z_n,$$

*where*

$$
\lambda_{n+1} = 
\begin{cases}
\min\left\{ \dfrac{\mu\left(\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right)}{2\langle Aw_n - Ay_n, z_n - y_n\rangle}, \ \lambda_n + d_n \right\}, & \text{if } \langle Aw_n - Ay_n, z_n - y_n\rangle > 0, \\[2mm]
\lambda_n + d_n, & \text{otherwise.}
\end{cases}
\tag{3.2}
$$

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

**Remark 3.5** In a special case when $\rho_n = 1$, $\forall n \geq 1$, we have that $x_{n+1} = z_n$. Thus, Algorithms 3.3 and 3.4 reduce to the inertial version of the subgradient extragradient method of Censor et al. [10–12]. In such case, we have the following simple criteria which guarantee Assumption 3.1, as well as assumption (2.4).

We give the following remarks on the difference between our proposed Algorithms 3.3 and 3.4 and [29, Algorithm 3.3] below.

**Remark 3.6** The method considered in [29, Algorithm 3.3] is a subgradient extragradient method without an inertial extrapolation step and without relation parameter. Our proposed Algorithms 3.3 and 3.4 are combinations of subgradient extragradient method, inertial extrapolation step and relaxation parameter. It can also be easily seen that [29, Algorithm 3.3] is a special case of our proposed Algorithm 3.4 when $\theta_n = 0$ and $\rho_n = 1$ in Algorithm 3.4. As pointed out in the Introduction, the aim of adding inertial extrapolation step and relaxation parameter to [29, Algorithm 3.3] (as seen in our proposed Algorithms 3.3 and 3.4) is to increase the speed of convergence of [29, Algorithm 3.3] in terms of reduction in the number of required iterations and CPU time. Consequently, we have given the numerical comparisons in terms of number of iterations and CPU time of our proposed Algorithms 3.3, 3.4 and [29, Algorithm 3.3] (kindly see Tables 3, 4, 5 and Figs. 3, 4, 5 in Section 5). This is one of the novelties of this paper.

**Proposition 3.7** *Assume that* $\{\theta_n\}$ *is a nondecreasing sequence that satisfies* $\theta_n \in [0, 1)$ $\forall n \geq 1$ *with* $\lim_{n\to\infty} \theta_n = \theta$ *such that the following condition holds:*

$$
1 - 3\theta > 0.
\tag{3.3}
$$

*Then, Assumptions* (2.4) *and* 3.1 *hold.*

**Proof** Observe that $\theta_n \leq \theta$, $\forall n \geq 1$. Thus, we have that assumption (2.4) is satisfied and $t_n \leq \frac{1}{1-\theta}$, $\forall n \geq 1$ (see [3]). Now, observe that $1 - 3\theta > 0$ implies that $(1 - \theta) > \frac{\theta(1+\theta)}{1-\theta}$. This further implies that there exists $\epsilon \in (0, 1)$ such that

$$
(1 - \epsilon)(1 - \theta) \geq \frac{\theta(1 + \theta)}{1 - \theta}.
\tag{3.4}
$$

Since $\theta_n \leq \theta$, $\forall n \geq 1$, we obtain from (3.4) that

$$
(1 - \epsilon)(1 - \theta_{n-1}) \geq \frac{\theta(1 + \theta)}{1 - \theta} \geq \theta_n t_{n+1}(1 + \theta_n),
\tag{3.5}
$$

for some $\epsilon \in (0, 1)$. Since $\theta_{n-1} \leq \theta_n$, $\forall n \geq 1$, we obtain that

$$
\theta_n t_{n+1}(1 + \theta_n) = \theta_n t_{n+1}(1 + \theta_n + [\theta_{n-1} - \theta_n]_+).
$$

Combining this with (3.5), we get that Assumption 3.1 is satisfied. $\qquad\square$

**Proposition 3.8** *Suppose that $\theta_n \in [0, 1)$, $\forall n \geq 1$ and there exists $c \in [0, \frac{1}{2})$ such that*

$$\lim_{n \to \infty} \left( \frac{1}{1 - \theta_{n+1}} - \frac{1}{1 - \theta_n} \right) = c \qquad (3.6)$$

*and*

$$\liminf_{n \to \infty} (1 - \theta_n)^2 > \limsup_{n \to \infty} \frac{\theta_n (1 + \theta_n)}{1 - 2c}. \qquad (3.7)$$

*Then, Assumptions (2.4) and 3.1 hold.*

**Proof**  It follows from Proposition 2.3(i) that Assumption (2.4) holds.
Now, from (3.7), we obtain that

$$\liminf_{n \to \infty} (1 - \theta_{n-1})^2 > \limsup_{n \to \infty} \frac{\theta_n (1 + \theta_n)}{1 - 2c}. \qquad (3.8)$$

Thus, there exists $\epsilon \in (0, 1)$ sufficiently small enough such that

$$\liminf_{n \to \infty} (1 - \theta_{n-1})^2 > \limsup_{n \to \infty} \frac{\theta_n (1 + \theta_n)}{1 - 2c - \epsilon(1 - c)} > \limsup_{n \to \infty} \frac{\theta_n (1 + \theta_n)}{1 - 2c}. \qquad (3.9)$$

This implies that

$$
\begin{aligned}
(1 + o(1))\theta_n (1 + \theta_n) &\leq [1 - 2c - \epsilon(1 - c) + o(1)](1 - \theta_{n-1})^2 \\
&= [(1 - \epsilon)(1 - c) - (2c - c + o(1))](1 - \theta_{n-1})^2 \\
&\leq [(1 - \epsilon)(1 - c) - \theta_n(c + o(1))](1 - \theta_{n-1})^2,
\end{aligned}
$$

which implies that

$$(1 - \epsilon)(1 - c)(1 - \theta_{n-1})^2 \geq (1 + o(1))\theta_n \left( 1 + \theta_n + (1 - \theta_{n-1})^2 + o\left((1 - \theta_{n-1})^2\right) \right). \qquad (3.10)$$

Now, observe from (3.6) that

$$\theta_{n-1} - \theta_n + c(1 - \theta_{n-1})(1 - \theta_n) = o\left((1 - \theta_{n-1})(1 - \theta_n)\right),$$

which implies from Proposition 2.3(ii) that

$$
\begin{aligned}
\theta_{n-1} - \theta_n &= -c(1 - \theta_{n-1})(1 - \theta_n) + o\left((1 - \theta_{n-1})(1 - \theta_n)\right) \\
&= -c(1 - \theta_{n-1})^2 + o(1 - \theta_{n-1})^2 \quad \text{as} \quad n \to \infty.
\end{aligned}
$$

This implies that

$$
\begin{aligned}
|\theta_{n-1} - \theta_n| &= |-c(1 - \theta_{n-1})^2 + o(1 - \theta_{n-1})^2| \\
&\leq c(1 - \theta_{n-1})^2 + o(1 - \theta_{n-1})^2 \quad \text{as} \quad n \to \infty. \qquad (3.11)
\end{aligned}
$$

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

$$(1 - \epsilon)(1 - c)(1 - \theta_{n-1})^2 \geq (1 + o(1))\,\theta_n \left( 1 + \theta_n + [\theta_{n-1} - \theta_n]_+ \right). \qquad (3.12)$$

By Proposition 2.3, we have that $t_{n+1} \sim t_n \sim \frac{1}{(1-c)(1-\theta_{n-1})}$ as $n \to \infty$. Hence, (3.12) is equivalent to

$$(1 - \epsilon)(1 - c)(1 - \theta_{n-1})^2 \geq \frac{\theta_n}{(1 - c)(1 - \theta_{n-1})} t_{n+1} \left( 1 + \theta_n + [\theta_{n-1} - \theta_n]_+ \right),$$

which further implies Assumption 3.1.                                                  □

When $\rho_n \neq 1$, then we have the following proposition which provides some criteria for ensuring Assumption 3.1, as well as Assumption (2.4).

**Proposition 3.9** (See for example, [3, Proposition 3.3])
*Suppose that $\theta_n \in [0, 1)$ and $\rho_n \in (0, 1)$ for all $n \geq 1$. Assume that there exist $c \in [0, 1)$ and $\bar{c} \in [c, 1)$ such that*

$$\lim_{n \to \infty} \left( \frac{1}{1 - \theta_{n+1}} - \frac{1}{1 - \theta_n} \right) = c,$$

$$\lim_{n \to \infty} \frac{\gamma_{n+1} - \gamma_n}{\gamma_n (1 - \theta_n)} = \bar{c}$$

*and*

$$\liminf_{n \to \infty} \gamma_n (1 - \theta_n)^2 > \limsup_{n \to \infty} \frac{\theta_n (1 + \theta_n)}{1 - \bar{c}},$$

*where $\gamma_n = \frac{1 - \rho_n}{2\rho_n}$. Then, Assumptions (2.4) and 3.1 hold.*

**Remark 3.10** It is worthy of note that many practical choices for the inertial and relaxation factors $\theta_n$ and $\rho_n$, respectively, satisfy Assumption 3.1. In fact, similar conditions as in Propositions 3.7–3.9 have already been used in the literature to ensure the convergence of inertial and relaxation methods (see [31,46,47] and the references therein). Thus, Assumption 3.1, as well as assumption (2.4) are much more weaker than the assumptions in those papers.

Moreover, we shall give in Sect. 5, some typical examples of $\theta_n$ and $\rho_n$ which satisfy the conditions in Propositions 3.7–3.9 (therefore, satisfying assumption (2.4) and Assumption 3.1). Then, we check the sensitivity of both $\theta_n$ and $\rho_n$ in order to find numerically, the optimum choice of these parameters with respect to the convergence speed of our proposed methods.

# 4 Convergence Analysis

**Lemma 4.1** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumption 3.2(a)-(b) hold. Then,*

$$\Gamma_{n+1} - \Gamma_n - \theta_n (\Gamma_n - \Gamma_{n-1})$$
$$\leq \frac{1}{2} (\theta_n + \theta_n^2) \|x_n - x_{n-1}\|^2 - \gamma_n \|x_{n+1} - w_n\|^2 - \frac{\rho_n}{2} (1 - \lambda L) \left[ \|w_n - y_n\|^2 + \|z_n - y_n\|^2 \right],$$

*where $\gamma_n := \frac{1 - \rho_n}{2\rho_n}$ and $\Gamma_n := \frac{1}{2} \|x_n - z\|^2, \ \forall z \in \Gamma$.*

**Proof** Let $z \in \Gamma$. Since $x_{n+1} - w_n = \rho_n (z_n - w_n)$, we obtain

$$\langle x_{n+1} - w_n, \ w_n - z \rangle = \rho_n \langle z_n - w_n, \ w_n - z_n \rangle + \rho_n \langle z_n - w_n, \ z_n - z \rangle$$
$$= -\rho_n \|z_n - w_n\|^2 + \rho_n \langle z_n - w_n, \ z_n - z \rangle$$
$$= -\rho_n^{-1} \|x_{n+1} - w_n\|^2 + \rho_n \langle z_n - w_n + \lambda A y_n, \ z_n - z \rangle - \lambda \rho_n \langle A y_n, \ z_n - z \rangle$$
$$\leq -\rho_n^{-1} \|x_{n+1} - w_n\|^2 + \lambda \rho_n \langle A y_n, \ z - z_n \rangle, \qquad (4.1)$$

where the last inequality follows from $z \in \Gamma \subseteq T_n$ and the characteristic property of $P_{T_n}$ (see (2.3)).

Now, using (4.1) in Lemma 2.6, we obtain

$$
\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1})
$$

$$
\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \rho_n^{-1}\|x_{n+1} - w_n\|^2 + \lambda\rho_n\langle Ay_n,\ z - z_n\rangle + \frac{1}{2}\|x_{n+1} - w_n\|^2
$$

$$
= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 - \frac{1}{2\rho_n}\|x_{n+1} - w_n\|^2
$$

$$
+ \lambda\rho_n\langle Ay_n,\ z - z_n\rangle
$$

$$
= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 + \lambda\rho_n\langle Ay_n,\ z - z_n\rangle
$$

$$
- \frac{1}{2\rho_n}\cdot\rho_n^2\|w_n - z_n\|^2
$$

$$
= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 + \lambda\rho_n\langle Ay_n,\ z - z_n\rangle
$$

$$
- \frac{\rho_n}{2}\|w_n - y_n\|^2 - \frac{\rho_n}{2}\|y_n - z_n\|^2 - \rho_n\langle w_n - y_n,\ y_n - z_n\rangle, \tag{4.2}
$$

where the last equality follows from Lemma 2.8.

Since $y_n \subset \mathcal{C}$ and $z \in \Gamma$, we get from (2.1) that $\langle Ay_n,\ y_n - z\rangle \geq 0, \forall n \geq 1$. That is $\langle Ay_n,\ y_n - z_n + z_n - z\rangle \geq 0, \forall n \geq 1$. This implies that

$$
\lambda\langle Ay_n,\ z - z_n\rangle - \langle w_n - y_n,\ y_n - z_n\rangle \leq \lambda\langle Ay_n,\ y_n - z_n\rangle - \langle w_n - y_n,\ y_n - z_n\rangle
$$

$$
= \langle \lambda Ay_n - w_n + y_n,\ y_n - z_n\rangle. \tag{4.3}
$$

Substituting (4.3) into (4.2), we obtain that

$$
\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1})
$$

$$
\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2
$$

$$
- \frac{\rho_n}{2}\left(\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right) + \rho_n\langle\lambda Ay_n - w_n + y_n,\ y_n - z_n\rangle
$$

$$
= \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 - \frac{\rho_n}{2}\left(\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right)
$$

$$
+ \rho_n\langle w_n - \lambda Aw_n - y_n,\ z_n - y_n\rangle + \lambda\rho_n\langle Aw_n - Ay_n,\ z_n - y_n\rangle
$$

$$
\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 - \frac{\rho_n}{2}\left(\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right)
$$

$$
+ \lambda\rho_n\langle Aw_n - Ay_n,\ z_n - y_n\rangle, \tag{4.4}
$$

where the last inequality follows from the definition of $T_n$.

Now, from the Lipschitz continuity of $A$, we obtain

$$
\langle Aw_n - Ay_n, z_n - y_n\rangle \leq L\|w_n - y_n\|\|z_n - y_n\|
$$

$$
= \frac{L}{2}\left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2 - (\|w_n - y_n\| - \|z_n - y_n\|)^2\right]
$$

$$
\leq \frac{L}{2}\left(\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right). \tag{4.5}
$$

Substituting (4.5) into (4.4), we obtain the desired conclusion.   □

**Lemma 4.2** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumption (2.4) and Assumption 3.2(a)-(b) hold. Then, the following inequality holds:*

$$\sum_{i=1}^{n-1} t_{i+1,n}\Big[ \big(2\gamma_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\big)\, \|x_i - x_{i-1}\|^2 + 2\gamma_i(1-\theta_i)$$
$$\big(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\big)\Big] \leq 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0,$$

*where $t_{i,n}$ is as defined in (2.7).*

**Proof** By Lemma 2.8, we obtain

$$\begin{aligned}
\|x_{n+1} - w_n\|^2 &= \|x_{n+1} - x_n - (x_n - x_{n-1}) + (1-\theta_n)(x_n - x_{n-1})\|^2 \\
&= \|x_{n+1} - 2x_n + x_{n-1}\|^2 + (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\
&\quad + 2(1-\theta_n)\langle x_{n+1} - 2x_n + x_{n-1}, x_n - x_{n-1}\rangle \\
&= \|x_{n+1} - 2x_n + x_{n-1}\|^2 + (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\
&\quad + (1-\theta_n)\big[\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2 - \|x_{n+1} - 2x_n + x_{n-1}\|^2\big] \\
&= \theta_n\|x_{n+1} - 2x_n + x_{n-1}\|^2 + (1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\
&\quad + (1-\theta_n)\big[\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2\big] \\
&\geq (1-\theta_n)^2\|x_n - x_{n-1}\|^2 + (1-\theta_n)\big[\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2\big].
\end{aligned}$$
$$(4.6)$$

Using (4.6) in Lemma 4.1, and noting that $1 - \lambda L > 0$, we obtain

$$\begin{aligned}
\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1}) &\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \gamma_n(1-\theta_n)^2\|x_n - x_{n-1}\|^2 \\
&\quad - \gamma_n(1-\theta_n)\Big[\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2\Big] \\
&= \Big[\frac{1}{2}(\theta_n + \theta_n^2) - \gamma_n(1-\theta_n)^2\Big]\|x_n - x_{n-1}\|^2 \\
&\quad - \gamma_n(1-\theta_n)\Big[\|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2\Big].
\end{aligned}$$

This together with Lemma 2.5(a) imply that

$$\begin{aligned}
\Gamma_n - \Gamma_0 &= \sum_{i=1}^{n}(\Gamma_i - \Gamma_{i-1}) \\
&\leq t_{1,n}(\Gamma_1 - \Gamma_0) \\
&\quad + \sum_{i=1}^{n-1} t_{i+1,n}\Big[\Big(\frac{1}{2}(\theta_i + \theta_i^2) - \gamma_i(1-\theta_i)^2\Big) \\
&\quad \|x_i - x_{i-1}\|^2 - \gamma_i(1-\theta_i)\Big(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\Big)\Big].
\end{aligned}$$

Noting that $t_{1,n} \leq t_1$, we obtain

$$\sum_{i=1}^{n-1} t_{i+1,n} \Big[ \big(2\gamma_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\big) \|x_i - x_{i-1}\|^2 + 2\gamma_i(1-\theta_i)$$

$$\Big( \|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2 \Big) \Big]$$

$$\leq 2t_{1,n}(\Gamma_1 - \Gamma_0) + 2(\Gamma_0 - \Gamma_n)$$

$$\leq 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0 - 2\Gamma_n.$$

Using $\Gamma_n = \frac{1}{2}\|x_n - z\|^2 \geq 0$, $\forall n \geq 1$ in the above inequality, we obtain the desired conclusion $\qquad\square$

**Lemma 4.3** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumptions (2.4) and 3.2(a)-(b) hold. Then, the following inequality holds:*

$$\sum_{i=1}^{n-1} 2\gamma_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}\Big(1 + \theta_i + 2\big[\gamma_i(1-\theta_i) - \gamma_{i-1}(1-\theta_{i-1})\big]_+\Big)\|x_i - x_{i-1}\|^2$$

$$\leq 2\Big[t_1|\Gamma_1 - \Gamma_0| + \Gamma_0 + t_1\gamma_0(1-\theta_0)\|x_1 - x_0\|^2\Big].$$

***Proof*** Observe that

$$\sum_{i=1}^{n-1} t_{i+1,n} \cdot 2\gamma_i(1-\theta_i)\Big[\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\Big]$$

$$= 2\sum_{i=1}^{n-1} \Big(t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\gamma_i(1-\theta_i)\Big)\|x_i - x_{i-1}\|^2$$

$$+ 2t_{1,n}\gamma_{n-1}(1-\theta_{n-1})\|x_n - x_{n-1}\|^2 - 2t_{1,n}\gamma_0(1-\theta_0)\|x_1 - x_0\|^2$$

$$\geq 2\sum_{i=1}^{n-1} \Big(t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\gamma_i(1-\theta_i)\Big)\|x_i - x_{i-1}\|^2$$

$$- 2t_{1,n}\gamma_0(1-\theta_0)\|x_1 - x_0\|^2$$

$$\geq 2\sum_{i=1}^{n-1} \Big(t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\gamma_i(1-\theta_i)\Big)\|x_i - x_{i-1}\|^2$$

$$- 2t_1\gamma_0(1-\theta_0)\|x_1 - x_0\|^2, \tag{4.7}$$

where the last inequality follows from $t_{1,n} \leq t_1$.
Now, using (4.7) in Lemma 4.2, we obtain that

$$\sum_{i=1}^{n-1} t_{i+1,n}\big(2\gamma_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\big)\|x_i - x_{i-1}\|^2$$

$$+ 2\sum_{i=1}^{n-1} \big(t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) - t_{i+1,n}\gamma_i(1-\theta_i)\big)\|x_i - x_{i-1}\|^2$$

$$\leq 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0 + 2t_1\gamma_0(1-\theta_0)\|x_1 - x_0\|^2.$$

That is

$$\sum_{i=1}^{n-1} \left[ t_{i+1,n}(2\gamma_i(1-\theta_i)^2 - (\theta_i+\theta_i^2) - 2\gamma_i(1-\theta_i)) + 2t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) \right] \|x_i - x_{i-1}\|^2$$

$$\leq 2\left[ t_1|\Gamma_1 - \Gamma_0| + \Gamma_0 + t_1\gamma_0(1-\theta_0)\|x_1 - x_0\|^2 \right]. \tag{4.8}$$

Now, recall from (2.8) that $t_{i,n} = 1 + \theta_i t_{i+1,n}$ for all $i \geq 1$ and $n \geq i+1$. Hence, we have

$$2t_{i,n}\gamma_{i-1}(1-\theta_{i-1}) = 2\left[ \gamma_{i-1}(1-\theta_{i-1}) + \theta_i t_{i+1,n}\gamma_{i-1}(1-\theta_{i-1}) \right].$$

This implies that

$$t_{i+1,n}\left[ 2\gamma_i(1-\theta_i)^2 - (\theta_i+\theta_i^2) - 2\gamma_i(1-\theta_i) \right] + 2t_{i,n}\gamma_{i-1}(1-\theta_{i-1})$$

$$= t_{i+1,n}\left[ 2\gamma_i(1-\theta_i)^2 - (\theta_i+\theta_i^2) - 2\gamma_i(1-\theta_i) + 2\theta_i\gamma_{i-1}(1-\theta_{i-1}) \right]$$

$$\quad + 2\gamma_{i-1}(1-\theta_{i-1})$$

$$= 2\gamma_{i-1}(1-\theta_{i-1}) + t_{i+1,n}\left( -2\gamma_i\theta_i(1-\theta_i) - (\theta_i+\theta_i^2) + 2\theta_i\gamma_{i-1}(1-\theta_{i-1}) \right)$$

$$= 2\gamma_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1,n}\left( 2\gamma_i(1-\theta_i) + 1 + \theta_i - 2\gamma_{i-1}(1-\theta_{i-1}) \right)$$

$$\geq 2\gamma_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1,n}\left( 1 + \theta_i + 2\left[ \gamma_i(1-\theta_i) - \gamma_{i-1}(1-\theta_{i-1}) \right]_+ \right)$$

$$\geq 2\gamma_{i-1}(1-\theta_{i-1}) - \theta_i t_{i+1}\left( 1 + \theta_i + 2\left[ \gamma_i(1-\theta_i) - \gamma_{i-1}(1-\theta_{i-1}) \right]_+ \right), \tag{4.9}$$

where the last inequality follows from $t_{i+1,n} \leq t_{i+1}$.
Now, using (4.9) in (4.8), we obtain that the desired conclusion.    □

**Lemma 4.4** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumptions (2.4), 3.1 and 3.2(a)-(b) hold. Then,*

$$\sum_{n=1}^{\infty} \theta_n t_{n+1}\|x_n - x_{n-1}\|^2 < \infty.$$

**Proof** Without loss of generality, we may assume that inequality (3.1) holds true for all $n \geq 1$. That is,

$$\begin{cases} 2\epsilon\gamma_{n-1}(1-\theta_{n-1}) \leq 2\gamma_{n-1}(1-\theta_{n-1}) - \theta_n t_{n+1}\left( 1 + \theta_n + 2\left[ \gamma_n(1-\theta_n) - \gamma_{n-1}(1-\theta_{n-1}) \right]_+ \right), & \text{if } \rho_n \in (0,1), \\ \epsilon(1-\theta_{n-1}) \leq (1-\theta_{n-1}) - \theta_n t_{n+1}\left( 1 + \theta_n + \left[ \theta_{n-1} - \theta_n \right]_+ \right), & \text{if } \rho_n = 1, \end{cases} \tag{4.10}$$

where $\gamma_n = \frac{1-\rho_n}{2\rho_n}$, for all $n \geq 1$.
At this point, we divide our proof into two cases.
**Case 1:** Suppose that $\rho_n \in (0,1), \ \forall n \geq 1$. Then, using (4.10) in Lemma 4.3, we obtain

$$\sum_{i=1}^{n-1} \epsilon\gamma_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 \leq t_1|\Gamma_1 - \Gamma_0| + \Gamma_0 + t_1\gamma_0(1-\theta_0)\|x_1 - x_0\|^2. \tag{4.11}$$

Now, taking limit as $n \to \infty$ in (4.11), we obtain that

$$\sum_{i=1}^{\infty} \gamma_{i-1}(1-\theta_{i-1})\|x_i - x_{i-1}\|^2 < \infty. \tag{4.12}$$

Again, from (4.10), we obtain that

$$
\frac{1}{2}\theta_n t_{n+1} \le \gamma_{n-1}(1-\theta_{n-1}) - \frac{1}{2}\theta_n t_{n+1}\left(\theta_n + 2\left[\gamma_n(1-\theta_n) - \gamma_{n-1}(1-\theta_{n-1})\right]_+\right)
$$
$$
- \epsilon\gamma_{n-1}(1-\theta_{n-1})
$$
$$
\le \gamma_{n-1}(1-\theta_{n-1}). \tag{4.13}
$$

Using (4.13) in (4.12), we obtain

$$
\sum_{i=1}^{\infty}\theta_i t_{i+1}\|x_i - x_{i-1}\|^2 < \infty.
$$

**Case 2:** Suppose that $\rho_n = 1$, $\forall n \ge 1$. Then, $x_{n+1} = z_n$ and $\gamma_n = 0$, $\forall n \ge 1$.
Setting $x = w_n - y_n$ and $y = x_{n+1} - y_n$ in Lemma 2.8, and using its result in Lemma 4.1, we obtain

$$
\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1}) \le \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2
$$
$$
- \frac{1}{2}\left[\|w_n - y_n + x_{n+1} - y_n\|^2 + \|w_n - x_{n+1}\|^2\right]
$$
$$
\le \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \frac{1}{2}\|w_n - x_{n+1}\|^2. \tag{4.14}
$$

Now, using (4.6) in (4.14), and repeating the same line of proof as in Lemma 4.2, we get

$$
\sum_{i=1}^{n-1} t_{i+1,n}\left[\left((1-\theta_i)^2 - (\theta_i+\theta_i^2)\right)\|x_i - x_{i-1}\|^2 + (1-\theta_i)\left(\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\right)\right]
$$
$$
\le 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0. \tag{4.15}
$$

Similar to (4.7), we have

$$
\sum_{i=1}^{n-1} t_{i+1,n}(1-\theta_i)\left[\|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2\right]
$$
$$
\ge \sum_{i=1}^{n-1}\left(t_{i,n}(1-\theta_{i-1}) - t_{i+1,n}(1-\theta_i)\right)\|x_i - x_{i-1}\|^2 - t_1(1-\theta_0)\|x_1 - x_0\|^2 \tag{4.16}
$$

Also, using (4.16) in (4.15), and repeating the same line of proof as in Lemma 4.3, we get

$$
\sum_{i=1}^{n-1}(1-\theta_{i-1}) - \theta_i t_{i+1}\left(1 + \theta_i + \left[\theta_{i-1} - \theta_i\right]_+\right)\|x_i - x_{i-1}\|^2
$$
$$
\le 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0 + t_1(1-\theta_0)\|x_1 - x_0\|^2. \tag{4.17}
$$

Using (4.10) in (4.17), and repeating the same line of proof as in **Case 1**, we obtain

$$
\sum_{i=1}^{\infty}\theta_i t_{i+1}\|x_i - x_{i-1}\|^2 < \infty,
$$

which yields the desired conclusion. □

**Lemma 4.5** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumption (2.4), Assumptions 3.1 and 3.2(a)-(b) hold. Then,*

(a) $\lim\limits_{n\to\infty} \|x_n - z\|$ *exists for all* $z \in \Gamma$, *and consequently,* $\{x_n\}$ *is bounded.*

(b) $\lim\limits_{n\to\infty} \left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right] = 0,$

(c) $\lim\limits_{n\to\infty} \|x_n - w_n\| = 0.$

**Proof** (a) From Lemma 4.1, we obtain

$$\Gamma_{n+1} - \Gamma_n \le \theta_n\left(\Gamma_n - \Gamma_{n-1}\right) + \frac{1}{2}\left(\theta_n + \theta_n^2\right)\|x_n - x_{n-1}\|^2$$
$$- \frac{\rho_n}{2}(1 - \lambda L)\left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right]$$
$$\le \theta_n\left(\Gamma_n - \Gamma_{n-1}\right) + \theta_n\|x_n - x_{n-1}\|^2$$
$$- \frac{\rho_n}{2}(1 - \lambda L)\left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right] \qquad (4.18)$$
$$\le \theta_n\left(\Gamma_n - \Gamma_{n-1}\right) + \theta_n\|x_n - x_{n-1}\|^2, \qquad (4.19)$$

where the second inequality follows from $\theta_n^2 \le \theta_n$ and the third inequality follows from $1 - \lambda L > 0$.

Now, applying Lemmas 2.5(b) and 4.4 in (4.19), we obtain that $\sum\limits_{n=1}^{\infty}\left[\Gamma_n - \Gamma_{n-1}\right]_+ < \infty$.

Since $\Gamma_n = \frac{1}{2}\|x_n - z\|^2$, we get that $\lim\limits_{n\to\infty}\|x_n - z\|^2$ exists. Hence, $\{x_n\}$ is bounded.

(b) By applying Lemma 2.5(a) in (4.18), we obtain that

$$\Gamma_n - \Gamma_0 = \sum_{i=1}^{n}\left(\Gamma_i - \Gamma_{i-1}\right)$$
$$\le t_{1,n}\left(\Gamma_1 - \Gamma_0\right) + \sum_{i=1}^{n-1}t_{i+1,n}\Big[\theta_i\|x_i - x_{i-1}\|^2$$
$$- \frac{\rho_i}{2}(1 - \lambda L)\left[\|w_i - y_i\|^2 + \|z_i - y_i\|^2\right]\Big]$$
$$\le t_1\left(\Gamma_1 - \Gamma_0\right) + \sum_{i=1}^{n-1}t_{i+1}\theta_i\|x_i - x_{i-1}\|^2 - \sum_{i=1}^{n-1}t_{i+1,n}$$
$$\frac{\rho_i}{2}(1 - \lambda L)\left[\|w_i - y_i\|^2 + \|z_i - y_i\|^2\right],$$

which implies that

$$\sum_{i=1}^{n-1}t_{i+1,n}\frac{\rho_i}{2}(1 - \lambda L)\left[\|w_i - y_i\|^2 + \|z_i - y_i\|^2\right]\Big]$$
$$\le \Gamma_0 - \Gamma_n + t_1(\Gamma_1 - \Gamma_0) + \sum_{i=1}^{n-1}t_{i+1}\theta_i\|x_i - x_{i-1}\|^2$$
$$\le \Gamma_0 + t_1|\Gamma_1 - \Gamma_0| + \sum_{i=1}^{n-1}t_{i+1}\theta_i\|x_i - x_{i-1}\|^2$$
$$< \infty,$$

where the last inequality follows from Lemma 4.4.

Since $t_{i+1,n} = 0$ for $i \geq n$, letting $n$ tend to $\infty$, the monotone convergence theorem then implies that

$$\sum_{i=1}^{\infty} t_{i+1} \frac{\rho_i}{2} (1 - \lambda L) \left[ ||w_i - y_i||^2 + ||z_i - y_i||^2 \right] < \infty. \tag{4.20}$$

Since $\liminf_{n \to \infty} \rho_n > 0$, there exists $M > 0$ such that $\rho_n \geq M$ for $n$ large enough. Now, replacing $i$ with $n$ in (4.20) and noting that $t_n \geq 1, \forall n \geq 1$, we obtain from (4.20) that $\sum_{n=1}^{\infty} \frac{M}{2} (1 - \lambda L) \left[ ||w_n - y_n||^2 + ||z_n - y_n||^2 \right] < \infty$, which further gives that $\lim_{n \to \infty} \left[ ||w_n - y_n||^2 + ||z_n - y_n||^2 \right] = 0$.

(c) Since $w_n = x_n + \theta_n (x_n - x_{n-1})$, we obtain from Lemma 4.4 that

$$\sum_{n=1}^{\infty} t_{n+1} ||w_n - x_n||^2 \leq \sum_{n=1}^{\infty} \theta_n t_{n+1} ||x_n - x_{n-1}||^2 < \infty.$$

Noting that $t_n \geq 1, \forall n \geq 1$, we conclude immediately that $\lim_{n \to \infty} ||w_n - x_n|| = 0$. $\qquad \square$

**Lemma 4.6** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumption (2.4), Assumptions 3.1 and 3.2(a)-(d) hold. If $x^*$ is one of the weak cluster points of $\{x_n\}$, then we have at least one of the following: $x^* \in \Gamma$ or $Ax^* = 0$.*

**Proof** By Lemma 4.5(a), we can choose a subsequence of $\{x_n\}$ denoted by $\{x_{n_k}\}$ such that $x_{n_k} \rightharpoonup x^* \in \mathcal{H}$. Also, from Lemma 4.5(b),(c), we obtain that $\lim_{n \to \infty} ||y_n - x_n|| = 0$. Hence, we can choose a subsequence $\{y_{n_k}\}$ of $\{y_n\}$ such that $y_{n_k} \rightharpoonup x^*$. Note that $x^* \in \mathcal{C}$ since $\{y_{n_k}\} \subset \mathcal{C}$. We now consider two possible cases.

**Case 1:** Suppose that $\limsup_{k \to \infty} ||Ay_{n_k}|| = 0$. Then, $\lim_{k \to \infty} ||Ay_{n_k}|| = \liminf_{k \to \infty} ||Ay_{n_k}|| = 0$. Also, by the sequentially weakly continuity of $A$ on $\mathcal{C}$, we obtain that $Ay_{n_k} \rightharpoonup Ax^*$. Thus, by the weakly lower semicontinuity of $|| \cdot ||$, we have that

$$0 < ||Ax^*|| \leq \liminf_{k \to \infty} ||Ay_{n_k}|| = 0, \tag{4.21}$$

which implies that $Ax^* = 0$.

**Case 2:** Suppose that $\limsup_{k \to \infty} ||Ay_{n_k}|| > 0$. Then, without loss of generality, we can choose a subsequence of $\{Ay_{n_k}\}$ still denoted by $\{Ay_{n_k}\}$ such that $\lim_{k \to \infty} ||Ay_{n_k}|| = M > 0$. Also by the characteristics property of $P_{\mathcal{C}}$, we obtain for all $x \in \mathcal{C}$ that

$$\langle w_{n_k} - \lambda A w_{n_k} - y_{n_k}, \ x - y_{n_k} \rangle \leq 0.$$

This implies that

$$\frac{1}{\lambda} \langle w_{n_k} - y_{n_k}, x - y_{n_k} \rangle + \langle A w_{n_k}, \ y_{n_k} - w_{n_k} \rangle \leq \langle A w_{n_k}, x - w_{n_k} \rangle. \tag{4.22}$$

Thus, we obtain from Lemma 4.5(b) that

$$0 \leq \liminf_{k \to \infty} \langle Aw_{n_k}, x - w_{n_k} \rangle \leq \limsup_{k \to \infty} \langle Aw_{n_k}, x - w_{n_k} \rangle < \infty, \ \forall x \in \mathcal{C}. \tag{4.23}$$

Now, note that

$$\langle Ay_{n_k}, x - y_{n_k} \rangle = \langle Ay_{n_k} - Aw_{n_k}, x - w_{n_k} \rangle + \langle Aw_{n_k}, x - w_{n_k} \rangle + \langle Ay_{n_k}, w_{n_k} - y_{n_k} \rangle. \tag{4.24}$$

Moreover, since $A$ is Lipschitz continuous on $\mathcal{H}$, we obtain from Lemma 4.5(b) that $\lim_{k \to \infty} \|Aw_{n_k} - Ay_{n_k}\| = 0$. Hence, we obtain from Lemmas 4.5(b), (4.23) and (4.24) that

$$0 \leq \liminf_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle \leq \limsup_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle < \infty, \ \forall x \in \mathcal{C}. \tag{4.25}$$

Based on (4.25), we consider two cases under **Case 2**, as follows:
**Case A:** Suppose that $\limsup_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle > 0, \forall x \in \mathcal{C}$. Then, we can choose a subsequence of $\{y_{n_k}\}$ denoted by $\{y_{n_{k_j}}\}$ such that $\lim_{j \to \infty} \langle Ay_{n_{k_j}}, x - y_{n_{k_j}} \rangle > 0$. Thus, there exists $j_0 \geq 1$ such that $\langle Ay_{n_{k_j}}, x - y_{n_{k_j}} \rangle > 0, \forall j \geq j_0$, which by the quasimonotonicity of $A$ on $\mathcal{H}$, implies $\langle Ax, x - y_{n_{k_j}} \rangle \geq 0, \forall x \in \mathcal{C}, j \geq j_0$. Thus, letting $j \to \infty$, we obtain that $\langle Ax, x - x^* \rangle \geq 0, \forall x \in \mathcal{C}$. Hence, $x^* \in \Gamma$.
**Case B:** Suppose that $\limsup_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle = 0, \forall x \in \mathcal{C}$. Then, by (4.25), we obtain that

$$\lim_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle = 0, \forall x \in \mathcal{C}, \tag{4.26}$$

which implies that

$$\langle Ay_{n_k}, x - y_{n_k} \rangle + |\langle Ay_{n_k}, x - y_{n_k} \rangle| + \frac{1}{k+1} > 0, \forall x \in \mathcal{C}. \tag{4.27}$$

Also, since $\lim_{k \to \infty} \|Ay_{n_k}\| = M > 0$, there exists $k_0 \geq 1$ such that $\|Ay_{n_k}\| > \frac{M}{2}, \forall k \geq k_0$. Therefore, we can set $q_{n_k} = \frac{Ay_{n_k}}{\|Ay_{n_k}\|^2}, \forall k \geq k_0$. Thus, $\langle Ay_{n_k}, q_{n_k} \rangle = 1, \forall k \geq k_0$. Hence, by (4.27), we obtain

$$\left\langle Ay_{n_k}, x + q_{n_k} \left[ |\langle Ay_{n_k}, x - y_{n_k} \rangle| + \frac{1}{k+1} - y_{n_k} \right] \right\rangle > 0,$$

which by the quasimonotonicity of $A$ on $\mathcal{H}$, implies

$$\left\langle A\left( x + q_{n_k} \left[ |\langle Ay_{n_k}, x - y_{n_k} \rangle| + \frac{1}{k+1} \right] \right), x + q_{n_k} \left[ |\langle Ay_{n_k}, x - y_{n_k} \rangle| + \frac{1}{k+1} \right] - y_{n_k} \right\rangle \geq 0.$$

This further implies that

$$\langle Ax, x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right] - y_{n_k}\rangle$$

$$\geq \langle Ax, A(x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right]), x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right] - y_{n_k}\rangle$$

$$\geq -\|Ax - A(x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right])\| \cdot \|x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right]$$

$$\quad - y_{n_k}\|$$

$$\geq -L\|q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right]\| \cdot \|x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right] - y_{n_k}\|$$

$$= \frac{-L}{\|Ay_{n_k}\|}\left(|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right) \cdot \|x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right] - y_{n_k}\|$$

$$\geq \frac{-2L}{M}\left(|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right)M_1, \tag{4.28}$$

for some $M_1 > 0$, where the existence of $M_1$ follows from the boundedness of $\{x + q_{n_k}\left[|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right] - y_{n_k}\}$. Note from (4.26) that $\lim_{k\to\infty}\left(|\langle Ay_{n_k}, x - y_{n_k}\rangle| + \frac{1}{k+1}\right) = 0$. Hence, letting $k \to \infty$ in (4.28), we get that $\langle Ax, x - x^*\rangle \geq 0, \forall x \in \mathcal{C}$. Therefore, $x^* \in \Gamma$. $\square$

**Lemma 4.7** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that Assumption (2.4), Assumption 3.1 and 3.2(a)-(d) hold. Then $\{x_n\}$ has at most one weak cluster point in $\Gamma$.*

**Proof** Suppose on the contrary that $\{x_n\}$ has at least two weak cluster points in $\Gamma$. Let $x^* \in \Gamma$ and $\bar{x} \in \Gamma$ be any two weak cluster points of $\{x_n\}$ such that $x^* \neq \bar{x}$. Also, let $\{x_{n_j}\}$ be a subsequence of $\{x_n\}$ such that $x_{n_j} \rightharpoonup \bar{x}$, as $j \to \infty$. Then, by Lemmas 4.5(a) and 2.7, we have

$$\lim_{n\to\infty} \|x_n - \bar{x}\| = \lim_{j\to\infty} \|x_{n_j} - \bar{x}\|$$

$$= \liminf_{j\to\infty} \|x_{n_j} - \bar{x}\|$$

$$< \liminf_{j\to\infty} \|x_{n_j} - x^*\|$$

$$= \lim_{n\to\infty} \|x_n - x^*\|$$

$$= \liminf_{k\to\infty} \|x_{n_k} - x^*\|$$

$$< \liminf_{k\to\infty} \|x_{n_k} - \bar{x}\| = \lim_{n\to\infty} \|x_n - \bar{x}\|,$$

which is a contradiction. Therefore, $\{x_n\}$ has at most one weak cluster point in $\Gamma$. $\square$

Using similar line of arguments as in [29, Lemma 3.5], we obtain the following weak convergence result.

**Theorem 4.8** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 such that assumption (2.4), Assumptions 3.1 and 3.2(a)-(e) hold. Then $\{x_n\}$ converges weakly to an element of $VI(\mathcal{C}, A)$.*

**Proof** By Assumption 3.2(e), $\{x \in \mathcal{C} : Ax = 0\} \setminus \Gamma$ is a finite set. Hence, by Lemma 4.6 and Lemma 4.7, we have that $\{x_n\}$ has finite weak cluster points in $VI(\mathcal{C}, A)$. Let $x^1, x^2, \cdots, x^m$

be the weak cluster points of $\{x_n\}$, and let $\{x_{n_k}^i\}$ be a subsequence of $\{x_n\}$ such that $x_{n_k}^i \rightharpoonup x^i$, as $k \to \infty$. Then, we obtain

$$\lim_{k \to \infty} \left\langle x_{n_k}^i, x^i - x^j \right\rangle = \left\langle x^i, x^i - x^j \right\rangle, \forall j \neq i. \tag{4.29}$$

Now, for $j \neq i$, set $q = \beta^{-1}(x^i - x^j)$, where $\beta = \|x^i - x^j\|$. Then,

$$\begin{aligned}
\langle x^i, q \rangle &= \beta^{-1} \langle x^i, x^i - x^j \rangle \\
&= \beta^{-1} \left( \|x^i\|^2 - \langle x^i, x^j \rangle \right) \\
&= \beta^{-1} \left[ \|x^i\|^2 - \frac{1}{2}\|x^i\|^2 - \frac{1}{2}\|x^j\|^2 + \frac{1}{2}\|x^i - x^j\|^2 \right] \\
&= \frac{1}{2\beta} \left( \|x^i\|^2 - \|x^j\|^2 \right) + \frac{1}{2}\beta \\
&> \frac{1}{2\beta} \left( \|x^i\|^2 - \|x^j\|^2 \right) + \frac{1}{4}\beta.
\end{aligned} \tag{4.30}$$

For sufficiently large $k$, we obtain from (4.29) and (4.30) that

$$x_{n_k}^i \in \left\{ x : \langle x, q \rangle > \frac{1}{2\beta} \left( \|x^i\|^2 - \|x^j\|^2 \right) + \frac{1}{4}\beta \right\}. \tag{4.31}$$

Hence, there exists $N_1 > N$ ($N \in \mathbb{N}$) such that

$$x_{n_k}^i \in B^i := \bigcap_{j=1, j \neq i}^{m} \left\{ x : \left\langle x, \frac{x^i - x^j}{\|x^i - x^j\|} \right\rangle > \frac{1}{2\|x^i - x^j\|} \left( \|x^i\|^2 - \|x^j\|^2 \right) + \epsilon_0 \right\}, \forall n \geq N_1, \tag{4.32}$$

where

$$\epsilon_0 = \min \left\{ \frac{1}{4} \|x^i - x^j\| : i, j = 1, 2, \cdots, m, i \neq j \right\}.$$

Now, from Lemma 4.5(b), we obtain that $\lim_{n \to \infty} \|z_n - w_n\| = 0$. Since $x_{n+1} - w_n = \rho_n(z_n - w_n)$ and $\{\rho_n\}$ is bounded, we obtain that $\lim_{n \to \infty} \|x_{n+1} - w_n\| = 0$. This together with Lemma 4.5(c), imply that $\lim_{n \to \infty} \|x_{n+1} - x_n\| = 0$. Hence, there exists $N_2 > N_1 > N$ such that $\|x_{n+1} - x_n\| < \epsilon_0, \forall n \geq N_2$.

Claim: $\{x_n\}$ has only one weak cluster point in $VI(\mathcal{C}, A)$.

Suppose on the contrary that $\{x_n\}$ has more than one weak cluster points in $VI(\mathcal{C}, A)$. Then, there exists $N_3 \geq N_2 > N_1 > N$ such that $x_{N_3} \in B^i$ and $x_{N_3+1} \in B^j$, where

$$B^j := \bigcap_{i=1, i \neq j}^{m} \left\{ x : \left\langle -x, \frac{x^j - x^i}{\|x^j - x^i\|} \right\rangle < \frac{1}{2\|x^j - x^i\|} \left( \|x^i\|^2 - \|x^j\|^2 \right) - \epsilon_0 \right\},$$

$$i, j \in \{1, 2, \cdots, m\} \text{ and } m \geq 2. \tag{4.33}$$

In particular, we have

$$\|x_{N_3+1} - x_{N_3}\| < \epsilon_0. \tag{4.34}$$

Since $x_{N_3} \in B^i$ and $x_{N_3+1} \in B^j$, we obtain that

$$\left\langle x_{N_3}, \frac{x^i - x^j}{\|x^i - x^j\|} \right\rangle > \frac{1}{2\|x^i - x^j\|} \left( \|x^i\|^2 - \|x^j\|^2 \right) + \epsilon_0 \tag{4.35}$$

and

$$\left\langle -x_{N_3+1}, \frac{x^i - x^j}{\|x^j - x^i\|} \right\rangle > \frac{1}{2\|x^i - x^j\|}\left(\|x^j\|^2 - \|x^i\|^2\right) + \epsilon_0. \tag{4.36}$$

Adding (4.35) and (4.36), and using (4.34), we obtain

$$2\epsilon_0 < \left\langle x_{N_3} - x_{N_3+1}, \frac{x_i - x^j}{\|x^i - x^j\|} \right\rangle \leq \|x_{N_3+1} - x_{N_3}\| < \epsilon_0, \tag{4.37}$$

which is not possible. Hence, our claim holds. That is, $\{x_n\}$ has only one weak cluster point in $VI(\mathcal{C}, A)$. Therefore, we conclude that $\{x_n\}$ converges weakly to an element of $VI(\mathcal{C}, A)$. $\square$

**Remark 4.9** (i) The conclusion of Theorem 4.8 still hold even if $\lambda \in (0, \frac{1}{L})$ in Algorithm 3.3 is replaced with a variable stepsize $\lambda_n$ such that $0 < \inf_{n \geq 1} \lambda_n \leq \sup_{n \geq 1} \lambda_n < \frac{1}{L}$. However, choosing this variable stepsize still requires the knowledge of the Lipschitz constant $L$.

(ii) When the Lipschitz constant is not known, we refer to Algorithm 3.4, where the choice of $\lambda_n$ does not depend on its knowledge. Note from the stepsize $\lambda_n$ in (3.2), that $\lim_{n \to \infty} \lambda_n = \lambda$ and $\lambda \in [\min\{\frac{\mu}{L}, \lambda_1\}, \lambda_1 + d]$, where $d = \sum_{n=1}^{\infty} d_n$ (see also [29, Lemma 3.1]).

(iii) When $d_n = 0$, then the stepsize $\lambda_n$ generated by Algorithm 3.4 is similar to that in [52]. We recall that the stepsize in [52] is monotone non-increasing, thus, their methods may depend on the choice of the initial stepsize $\lambda_1$. However, the stepsize given in (3.2) is non-monotonic and hence, the dependence on the initial stepsize $\lambda_1$ is reduced.

In the light of the above remark, we analyze the convergence of Algorithm 3.4 in what follows.

**Lemma 4.10** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.4 such that Assumption 3.2(a)-(b) hold. Then,*

$$\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1})$$
$$\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \gamma_n\|x_{n+1} - w_n\|^2 - \frac{\rho_n}{2}\left(1 - \lambda_n\frac{\mu}{\lambda_{n+1}}\right)$$
$$\left[\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right],$$

*where $\gamma_n := \frac{1-\rho_n}{2\rho_n}$ and $\Gamma_n := \frac{1}{2}\|x_n - z\|^2, \ \forall z \in \Gamma$.*

**Proof** By following the same line of proof used in obtaining (4.4), we have

$$\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1})$$
$$\leq \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 + \frac{\rho_n - 1}{2\rho_n}\|x_{n+1} - w_n\|^2 - \frac{\rho_n}{2}\left(\|w_n - y_n\|^2 + \|y_n - z_n\|^2\right)$$
$$+ \lambda_n\rho_n\langle Aw_n - Ay_n, \ z_n - y_n\rangle. \tag{4.38}$$

If $\langle Aw_n - Ay_n, z_n - y_n\rangle \leq 0$, then we obtain the conclusion of Lemma 4.10 immediately from (4.38).

In the case that $\langle Aw_n - Ay_n, z_n - y_n\rangle > 0$, we obtain from (3.2) that

$$\langle Aw_n - Ay_n, z_n - y_n\rangle \leq \frac{\mu}{2\lambda_{n+1}}\left(\|w_n - y_n\|^2 + \|z_n - y_n\|^2\right). \tag{4.39}$$

Substituting (4.39) into (4.38), we obtain the desired conclusion. $\square$

**Lemma 4.11** *Let* $\{x_n\}$ *be a sequence generated by Algorithm* 3.3 *such that assumption* (2.4) *and Assumption* 3.2(a)-(b) *hold. Then, the following inequality holds:*

$$\sum_{i=1}^{n-1} t_{i+1,n} \Big[ \big(2\gamma_i(1-\theta_i)^2 - (\theta_i + \theta_i^2)\big) \|x_i - x_{i-1}\|^2 + 2\gamma_i(1-\theta_i)$$
$$\Big( \|x_{i+1} - x_i\|^2 - \|x_i - x_{i-1}\|^2 \Big) \Big] \le 2t_1|\Gamma_1 - \Gamma_0| + 2\Gamma_0,$$

*where* $t_{i,n}$ *is as defined in* (2.7).

**Proof** By following the same line of proof used in obtaining (4.6), we have

$$\|x_{n+1} - w_n\|^2 \ge (1-\theta_n)^2\|x_n - x_{n-1}\|^2 + (1-\theta_n)$$
$$\big[ \|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2 \big]. \tag{4.40}$$

Also, by Remark 4.9(ii), we obtain that $\lim_{n\to\infty} \lambda_n \frac{\mu}{\lambda_{n+1}} = \mu \in (0,1)$. Thus, there exists $n_0 \ge 1$ such that $\forall n \ge n_0, 0 < \lambda_n \frac{\mu}{\lambda_{n+1}} < 1$. Hence, we get from Lemmas 4.10 and (4.40) that

$$\Gamma_{n+1} - \Gamma_n - \theta_n(\Gamma_n - \Gamma_{n-1}) \le \frac{1}{2}(\theta_n + \theta_n^2)\|x_n - x_{n-1}\|^2 - \gamma_n\|x_{n+1} - w_n\|^2$$
$$\le \Big[ \frac{1}{2}(\theta_n + \theta_n^2) - \gamma_n(1-\theta_n)^2 \Big]\|x_n - x_{n-1}\|^2$$
$$- \gamma_n(1-\theta_n)\Big[ \|x_{n+1} - x_n\|^2 - \|x_n - x_{n-1}\|^2 \Big], \ \forall n \ge n_0.$$

The remaining part of the proof is the same as the proof of Lemma 4.2.                □

In similar manner, we have that Lemmas 4.3 to 4.7 hold for Algorithm 3.4. Thus, we have the following theorem whose proof is the same as the proof of Theorem 4.8.

**Theorem 4.12** *Let* $\{x_n\}$ *be a sequence generated by Algorithm* 3.4 *such that Assumption* (2.4), *Assumptions* 3.1 *and* 3.2(a)-(e) *hold. Then* $\{x_n\}$ *converges weakly to an element of* $VI(\mathcal{C}, A)$.

**Remark 4.13** From our analyses, one can see that Assumption 3.1 is mainly used to guarantee the summation:

$$\sum_{n=1}^{\infty} \theta_n t_{n+1} \|x_n - x_{n-1}\|^2 < \infty \tag{4.41}$$

obtained in Lemma 4.4. Thus, if we assume that (4.41) directly, then we do not need Assumption 3.1 for the convergence of our methods.

In the case where $\rho_n = 1, \ \forall n \ge 1$, our methods correspond to the inertial subgradient extragradient methods. Note that if $\theta_n \in [0, \theta]$ for every $n \ge 1$, where $\theta \in [0, 1)$, then $t_n \le \frac{1}{(1-\theta)} \ \forall n \ge 1$. Under these settings, (4.41) is guaranteed by the condition

$$\sum_{n=1}^{\infty} \theta_n \|x_n - x_{n-1}\|^2 < \infty. \tag{4.42}$$

Thus, instead of Assumption 3.1, we may assume directly that $\theta_n \in [0, \theta], \ \forall n \ge 1$ and that condition (4.42) holds. Recall that this condition has been used by numerous authors to ensure convergence of inertial methods (see, for example, [2,15,31,32,34] and the references therein).

**Remark 4.14** Note that we did not make use of Assumption 3.2(c)-(e) in the proof of Lemmas 4.1–4.5. Suppose that $\mathcal{H}$ is a finite dimensional Hilbert space, then under Assumption 3.2(a)-(b), we get from Lemma 4.5(a) that there exists a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $\{x_{n_k}\}$ converges to some point $x^*$. By Lemma 4.5(b)(c), we get

$$\lim_{k \to \infty} \|w_{n_k} - y_{n_k}\| = 0$$

and

$$\lim_{k \to \infty} \|w_{n_k} - x_{n_k}\| = 0.$$

Thus, by the definition of $y_n$ and the continuity of $A$, we have

$$x^* = \lim_{k \to \infty} x_{n_k} = \lim_{k \to \infty} y_{n_k} = \lim_{k \to \infty} P_C(w_{n_k} - \lambda A w_{n_k}) = P_C(x^* - \lambda A x^*),$$

which implies that $x^* \in VI(C, A)$.

Now, replacing $z$ by $x^*$ in Lemma 4.5(a), we obtain that $\lim_{n \to \infty} \|x_n - x^*\|^2$ exists. Since $x^*$ is a cluster point of $\{x_n\}$, we obtain that $\{x_n\}$ converges to $x^*$.

*In summary, in a finite dimensional Hilbert space, our methods require that $\Gamma \neq \emptyset$ and the operator $A$ only needs to be Lipschitz continuous without any form of monotonicity.*

To achieve this (convergence without any form of monotonicity) in infinite dimensional Hilbert space, we replace Assumption 3.2(d)-(e) with the following:

(d)*  If $x_n \rightharpoonup x^*$ and $\limsup\limits_{n \to \infty} \langle Ax_n, x_n \rangle \leq \langle Ax^*, \bar{x} \rangle$, then $\lim\limits_{n \to \infty} \langle Ax_n, x_n \rangle = \langle Ax^*, x^* \rangle$.

(e)*  The set $VI(C, A) \setminus \Gamma$ is a finite set.

In fact, we have the following theorem.

**Theorem 4.15** *Let $\{x_n\}$ be a sequence generated by Algorithm 3.3 (or Algorithm 3.4) such that Assumptions (2.4), 3.1 and 3.2(a)-(c) and conditions (d)\*-(e)\* hold. Then $\{x_n\}$ converges weakly to an element of $VI(C, A)$.*

**Proof** First notice that Assumption 3.2(d) was used only after (4.25) in order to establish the conclusion of Lemma 4.6.

But from (4.25), we have that

$$0 \leq \liminf_{k \to \infty} \langle Ay_{n_k}, x - y_{n_k} \rangle, \ \forall x \in C.$$

Now, let $\{c_k\}$ be a sequence of positive numbers such that $\lim_{k \to \infty} c_k = 0$ and $\langle Ay_{n_k}, x - y_{n_k} \rangle + c_k > 0, \ \forall k \geq 0, x \in C$. Then,

$$\langle Ay_{n_k}, x \rangle + c_k > \langle Ay_{n_k}, y_{n_k} \rangle, \ \forall k \geq 0, x \in C. \tag{4.43}$$

Note that $y_{n_k} \rightharpoonup x^*$ and $x^* \in C$. Thus, we have in particular,

$$\langle Ay_{n_k}, x^* \rangle + c_k > \langle Ay_{n_k}, y_{n_k} \rangle, \ \forall k \geq 0. \tag{4.44}$$

Taking limit as $k \to \infty$ in (4.44), and using the sequentially weakly continuity of $A$, we obtain that

$$\langle Ax^*, x^* \rangle \geq \limsup_{k \to \infty} \langle Ay_{n_k}, y_{n_k} \rangle,$$

which by condition (d)* and (4.43) implies that

$$
\begin{aligned}
\langle Ax^*, x^* \rangle &= \lim_{k\to\infty} \langle Ay_{n_k}, y_{n_k} \rangle \\
&= \liminf_{k\to\infty} \langle Ay_{n_k}, y_{n_k} \rangle \\
&\leq \lim_{k\to\infty} \left( \langle Ay_{n_k}, x \rangle + c_k \right) = \langle Ax^*, x \rangle.
\end{aligned}
$$

This further implies that $x^* \in VI(C, A)$.

Now, using condition (e)* and following similar line of proof as in Lemma 4.7 and Theorem 4.15, we get that $\{x_n\}$ converges weakly to $x^*$. □

**Remark 4.16** (i) If $x_n \rightharpoonup x^*$ and $A$ is sequentially weakly-strongly continuous, then $A$ satisfies condition (d)*.

(ii) In the numerical experiments, we do not need to consider condition (e) (or (e)*). First note that whenever $\|y_n - w_n\| < \epsilon$, Algorithms 3.3 and 3.4 terminate in a finite step of the iterations (and $y_n$ is the solution of the VIP (1.1)). But from Lemma 4.5(b), $\lim_{n\to\infty} \|y_n - w_n\| = 0$ and condition (e) (or (e)*) was not used in establishing it.

We now give some remarks regarding the contributions and novelties of this paper.

**Remark 4.17** (1) If we set the inertial factor $\theta_n = 0$ and relaxation factor $\rho_n = 1$, then our Algorithm 3.4 reduces to [29, Algorithm 3.3]. Note that these parameters (factors) play vital role in improving the convergence rate of iterative methods. In fact, their influence with regards to the numerical performance of iterative schemes was discussed in [23]. Moreover, the benefits gained from incorporating the steps of these two parameters in our algorithms, are further verified in Sect. 5. Thus, bearing in mind the importance of these two parameters in iterative algorithms, we can see that our methods significantly improve the methods in [29].

(2) In the study of inertial methods for solving VIPs (even for monotone mappings), the inertia parameter is usually restricted in $[0, \frac{1}{3})$ and/or required to be nondecreasing with an upper bound (see, for example, [2,31,46,47]). In many cases, to ensure convergence, authors usually require the inertial parameter to depend on the knowledge of the Lipschitz constant of the cost operator which sometimes is very difficult to estimate in practice (see, for instance, [8]). Another condition usually imposed on the inertial parameter in the literature, is condition (4.42), which rely on the sequence $\{x_n\}$. One of the novelties of this paper is that we derived a general condition (Assumption 3.1) which is weaker than the above conditions used in the literature for ensuring the convergence of inertial methods for VIPs. As a result, we developed a different technique to ensure the convergence of our proposed Algorithms 3.3 and 3.4.

(3) In addition to (2) above, bearing in mind the Nestrov's accelerated scheme ( [35]), another novelty of this paper, is that the assumptions on the inertial and relaxation parameters in Algorithms 3.3 and 3.4 allow the case where inertial factor $\theta_n$ converges to the limit very close to 1 (see Remark 2.4 and the choices in **Experiment 1** of Sect. 5), which is very crucial in the study of inertial methods. This is actually where the relaxation effect and the parameter $\rho_n$ come into play, as crucial ingredients of our methods. Thus, the novelty of this paper is not only in improving the convergence speed of the methods in [29] but to also provide weaker conditions on the inertial parameter in methods for solving VIPs and offer a different but unified technique in proving their convergence. Furthermore, we employed condition (Assumption 3.1) where joint adjustments of the inertial and relaxation parameters play crucial role (for instance, see **Experiment 1**). Indeed, this is

a new perspective in the study of inertial methods for VIPs. Moreover, our study offers many possibilities for future research in this direction; like how to modify Assumption 3.1 so that inertial factor $\theta_n$ is allowed to converge to 1.

(4) The method of proof of our proposed Algorithms 3.3 and 3.4 is different from the method of proof of [29, Algorithm 3.3]. For example, we use different arguments to show that the sequence $\{x_n\}$ is bounded in Lemma 4.5 (different from the arguments used in [29, Lemma 3.2]). We also use different arguments in Lemmas 4.1, 4.2, 4.3, and 4.4.

## 5 Numerical Experiments

In this section, we give some numerical examples to show the implementation of our proposed methods (Algorithms 3.3 and 3.4). We also compare our new methods with Algorithms 3.1 and 3.3 in [29], and Algorithm 2.1 in [50].

The codes are written in Matlab 2016 (b) and performed on a personal computer with an Intel(R) Core(TM) i5-2600 CPU at 2.30 GHz and 8.00 Gb-RAM. In Tables 1, 2, 3, 4, 5, "Iter." means the number of iterations while "CPU" means the CPU time in seconds.

In our computations, we define $\mathrm{TOL}_n := ||y_n - w_n||$ for Algorithms 3.3 and 3.3. While for Algorithms 3.1 and 3.3 in [29], we define $\mathrm{TOL}_n := \|y_n - x_n\|/\min\{\lambda_n, 1\}$ and for Algorithm 2.1 in [50], we define $\mathrm{TOL}_n := \|x_n - z_n\| = \|x_n - P_{\mathcal{C}}(x_n - Ax_n)\|$ (as done in [29] and [50], respectively). Then, we use the stopping criterion $\mathrm{TOL}_n < \varepsilon$ for the iterative processes, where $\varepsilon$ is the predetermined error. These choices of stopping criterion for these methods are the best to be able to terminate the algorithms based on the examples we consider. As done in [29], we take $\varepsilon = 10^{-6}$ for all Algorithms.

We choose $\mu = 0.5$, $d_n = \frac{100}{(n+1)^{1.1}}$ and $\lambda_1 = 1$ for Algorithm 3.4 and Algorithms 3.1 and 3.3 in [29]; $\lambda = \frac{1}{2L}$ for Algorithm 3.3; $\gamma = 0.4$ and $\sigma = 0.99$ for Algorithm 2.1 in [50]. These choices are the same as in [29,50] and are optimal values for these parameters.

**Example 5.1** Let $\mathcal{C} = [-1, 1]$ and

$$Ax = \begin{cases} 2x - 1, & x > 1, \\ x^2, & x \in [-1, 1], \\ -2x - 1, & x < -1. \end{cases}$$

Then, $A$ is quasimonotone and 2-Lipschitz continuous. Also, $\Gamma = \{-1\}$ and $VI(\mathcal{C}, A) = \{-1, 0\}$.

**Example 5.2** Let $\mathcal{C} = \{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1, \ 0 \leq x_1\}$ and $A(x_1, x_2) = (-x_1 e^{x_2}, x_2)$. Then, $A$ may not be quasimonotone. We can also check that $(1, 0) \in \Gamma$ and $VI(\mathcal{C}, A) = \{(1, 0), (0, 0)\}$ (see [29, Section 4]), which by Remark 4.14 satisfy our assumptions. This example was also tested in [29].

**Example 5.3** We next consider the following problem which was also considered in [29,33, 42].
Let $\mathcal{C} = [0, 1]^m$ and $Ax = (f_1 x, f_2 x, \cdots, f_m x)$,
where $f_i x = x_{i-1}^2 + x_i^2 + x_{i-1} x_i + x_i x_{i+1} - 2x_{i-1} + 4x_i + x_{i+1} - 1, \ i = 1, 2, \cdots, m$, $x_0 = x_{m+1} = 0$.

We test these examples under the following experiments.

**Table 1** Numerical results for Example 1 (**Experiment 1**)

| $\rho_n$ | | Algorithm 3.3 | | Algorithm 3.4 | |
|---|---|---|---|---|---|
| 1 | CPU Iter. | 0.0016 | 7 | 0.0015 | 3 |
| $\frac{1}{20} + \frac{1}{(n+1)^2}$ | CPU Iter. | 0.0022 | 13 | 0.0027 | 6 |
| $\frac{1}{20} + \frac{2}{(n+1)^3}$ | CPU Iter. | 0.0026 | 8 | 0.0020 | 9 |
| $\frac{1}{20} + \frac{3}{(n+1)^4}$ | CPU Iter. | 0.0028 | 11 | 0.0028 | 12 |
| $\frac{1}{20} + \frac{5}{(n+1)^4}$ | CPU Iter. | 0.0597 | 21 | 0.0146 | 20 |

**Experiment 1**

In this first experiment, we check the behavior of our methods by fixing the inertial parameter and varying the relaxation parameter. We do this in order to check the effects of the relaxation parameter on our methods.

**For Example** 1: We take $\theta_n = \frac{3n+1}{10n+5}$ with $\rho_n = 1$ (which by Proposition 3.7, satisfies Assumptions (2.4) and 3.1), and $\theta_n = \frac{19}{20} - \frac{1}{(n+1)^{\frac{1}{2}}}$ with $\rho_n \in \{\frac{1}{20} + \frac{1}{(n+1)^2}, \frac{1}{20} + \frac{2}{(n+1)^3}, \frac{1}{20} + \frac{3}{(n+1)^4}, \frac{1}{20} + \frac{5}{(n+1)^4}\}$ (which also satisfies Assumptions (2.4) and 3.1 by Proposition 3.9).

Also, we take $x_1 = 1$ and $x_0 = 0.5$ for this example. Since the Lipshitz constant is known for this example, we use Algorithms 3.3 and 3.4 for the experiment and obtain the numerical results listed in Table 1 and Fig. 1. From the table and graph, we can see that $\rho_n = 1$ performs better than other choices made for Algorithms 3.3 and 3.4.

**For Example** 2 and Example 3: We take $\theta_n = \frac{9.9}{10} - \frac{1}{n+1}$ with $\rho_n = 1$ (which by Proposition 3.8 (see also Remark 2.4), satisfies Assumption (2.4) and 3.1), and $\theta_n = \frac{9}{10} - \frac{1}{(n+1)^{\frac{1}{3}}}$ with $\rho_n \in \{\frac{1}{10} + \frac{1}{n+1}, \frac{1}{11} + \frac{1}{n+1}, \frac{1}{12} + \frac{1}{n+1}, \frac{1}{13} + \frac{1}{n+1}\}$. Also, we take $x_1 = (0.1, 0.5)$ and $x_0 = (0.2, 0.1)$ for Example 2 while for Example 3, we choose $x_1$ and $x_0$ randomly with $m = 50$. For these examples, we use Algorithm 3.4 for the experiment and obtain the numerical results listed in Table 2 and Fig. 2. From the table and graph, we can see that the $\rho_n = \frac{1}{11} + \frac{1}{n+1}$ performs better than other choices made for Example 2 while $\rho_n = \frac{1}{10} + \frac{1}{n+1}$ performs better for Example 3, which validates the benefits sometimes brought by the relaxation parameter.

**Experiment 2**

In this experiment, we compare our methods with Algorithms 3.1 and 3.3 in [29], and Algorithm 2.1 in [50]. Here, we randomly choose the $\theta_n$ and $\rho_n$ from Experiment 1, and then, consider the following cases for the starting points in each example.

**For Example** 1: Case I: $x_1 = 0.2$, $x_0 = 0.1$ and Case II: $x_1 = 0.6$, $x_0 = 0.2$.

**For Example** 2: Case I: $x_1 = (1, 0.5)$, $x_0 = (0.2, 0.1)$ and Case II: $x_1 = (0.1, 0.2)$, $x_0 = (0.3, 0.5)$.

**For Example** 3: Case I: $m = 100$ and Case II: $m = 150$ ($x_1$ and $x_0$ are randomly taken). We use Algorithms 3.1 and 3.4 for the comparison in Example 1 while we use only Algorithm 3.4 for the comparison in Example 2 and Example 3. The numerical results are given in Tables 3, 4, 5 and Figs. 3, 4, 5. The results show that our methods perform better than Algorithms 3.1 and 3.3 in [29], and Algorithm 2.1 in [50].
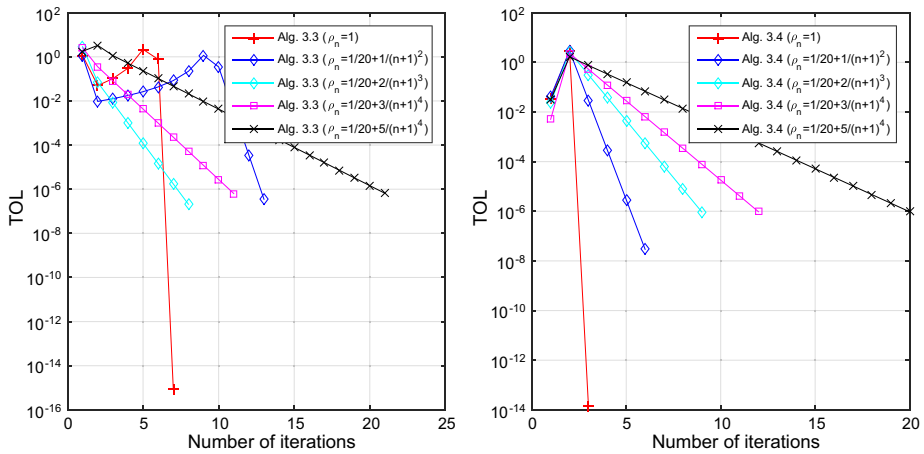
**Fig. 1** The behavior of Algorithms 3.3 and 3.4 for Example 1 (**Experiment 1**)

| $\rho_n$ | | Example 2 | | Example 3 | |
|---|---|---|---|---|---|
| | | | | | |
| $1$ | CPU Iter. | 0.0123 | 20 | 0.0232 | 68 |
| $\frac{1}{10} + \frac{1}{n+1}$ | CPU Iter. | 0.0162 | 23 | 0.0107 | 39 |
| $\frac{1}{11} + \frac{1}{n+1}$ | CPU Iter. | 0.0050 | 16 | 0.0304 | 45 |
| $\frac{1}{12} + \frac{1}{n+1}$ | CPU Iter. | 0.0149 | 30 | 0.0346 | 51 |
| $\frac{1}{13} + \frac{1}{n+1}$ | CPU Iter. | 0.0163 | 45 | 0.0257 | 71 |

**Table 2** Numerical results for Algorithm 3.4 (**Experiment 1**)
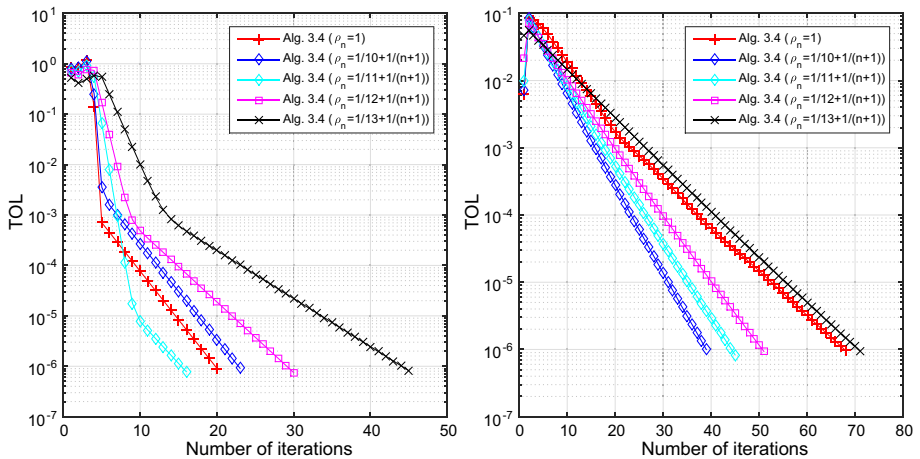


**Fig. 2** The behavior of Algorithm 3.4 for Examples 2 and 3 (**Experiment 1**)

**Table 3** Numerical results for Example 1 (**Experiment 2**)

| Cases | | Algorithm 3.3 | | Algorithm 3.4 | | Algorithm 3.1 in [29] | | Algorithm 3.3 in [29] | | Algorithm 2.1 in [50] | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| I | CPU Iter. | 0.0223 | 795 | 0.0289 | 922 | 0.0612 | 1002 | 0.0621 | 998 | 0.0923 | 1848 |
| II | CPU Iter. | 0.0244 | 788 | 0.0272 | 936 | 0.0672 | 996 | 0.0662 | 992 | 0.0900 | 1828 |

**Table 4** Numerical results for Example 2 (**Experiment 2**)

| Cases | | Algorithm 3.4 | | Algorithm 3.1 in [29] | | Algorithm 3.3 in [29] | | Algorithm 2.1 in [50] | |
|---|---|---|---|---|---|---|---|---|---|
| I | CPU Iter. | 0.0036 | 16 | 0.0200 | 55 | 0.0149 | 52 | 0.1780 | 927 |
| II | CPU Iter. | 0.0062 | 44 | 0.0185 | 80 | 0.0181 | 77 | 0.1289 | 1427 |

**Table 5** Numerical results for Example 3 (**Experiment 2**)

| Cases | | Algorithm 3.4 | | Algorithm 3.1 in [29] | | Algorithm 3.3 in [29] | | Algorithm 2.1 in [50] | |
|---|---|---|---|---|---|---|---|---|---|
| I | CPU Iter. | 0.0049 | 37 | 0.0931 | 65 | 0.0911 | 62 | 0.2021 | 952 |
| II | CPU Iter. | 0.0041 | 42 | 0.0916 | 56 | 0.0900 | 54 | 0.2102 | 2427 |
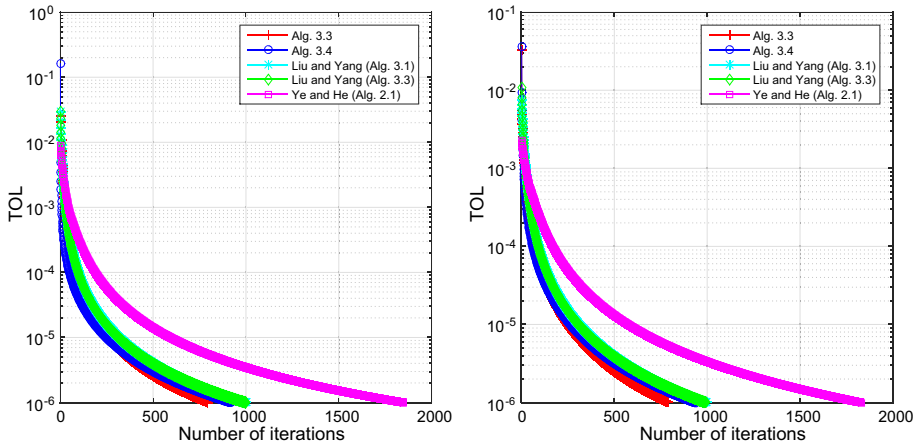
**Fig. 3** The behavior of $\text{TOL}_n$ for Example 1 (**Experiment 2**): Left: **Case I**; Right: **Case II**
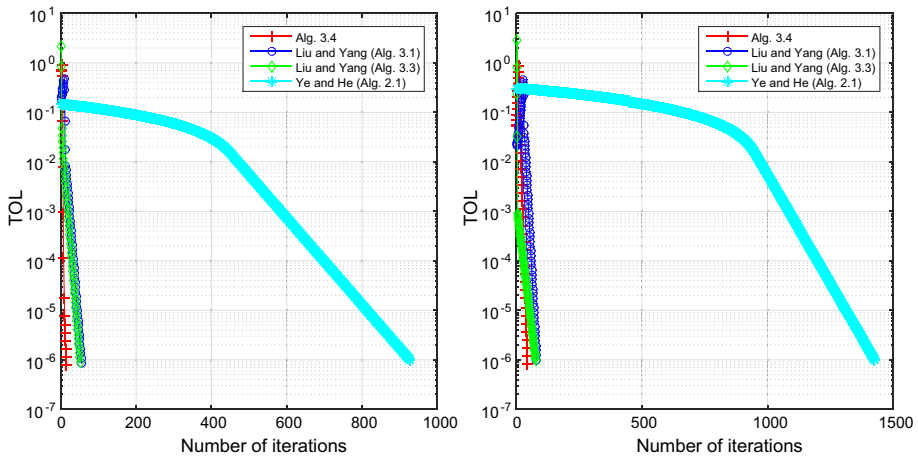


**Fig. 4** The behavior of $\text{TOL}_n$ for Example 2 (**Experiment 2**): Left: **Case I**; Right: **Case II**

**Remark 5.4** From Tables 3, 4, 5 and Figs. 3, 4, 5, one can see that our proposed Algorithms 3.3 and 3.4 are more efficient than [29, Algorithm 3.3] and outperform [29, Algorithm 3.3] in terms of number of iterations and CPU time based on our test problems due to the presence of inertial extrapolation step in Algorithms 3.3 and 3.4.

## 6 Conclusion

We propose and study two new subgradient extragradient methods with relaxed inertial steps for solving the variational inequality problems in real Hilbert spaces when the cost operator is not necessarily pseudomonotone. The first method is proposed when the Lipschitz constant of the cost operator is known while the second method which involves adaptive stepsize strategy, is proposed when the Lipschitz constant of the operator is not available. The techniques employed in this paper are quite different from the ones used in most papers,
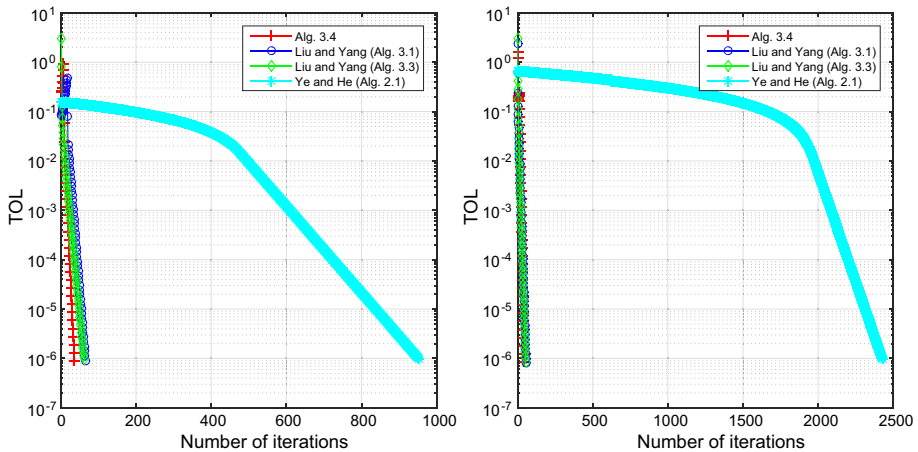
**Fig. 5** The behavior of $\text{TOL}_n$ for Example 3 (**Experiment 2**): Left: **Case I**; Right: **Case II**

and the assumptions on the inertial and relaxation factors, are weaker than those in many papers for solving variational inequality problems. The benefits brought from the relaxation are further verified in **Experiment 1**, and as seen in **Experiment 2**, our methods performs better when compared with other methods for solving the non-pseudomonotone variational inequality problems.

## Declarations

**Declaration** The authors declare that they have no competing interests.

## References

1. Alakoya, T.O., Jolaoso, L.O., Mewomo, O.T.: Modified inertia subgradient extragradient method with self adaptive stepsize for solving monotone variational inequality and fixed point problems. Optimization (2020). https://doi.org/10.1080/02331934.2020.1723586
2. Alvarez, F., Attouch, H.: An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping. Set-Valued Anal. **9**, 3–11 (2001)
3. Attouch, H., Cabot, A.: Convergence of a relaxed inertial proximal algorithm for maximally monotone operators. Math. Program. (2019). https://doi.org/10.1007/s10107-019-01412-0
4. Attouch, H., Cabot, A.: Convergence of a relaxed inertial forward-backward algorithm for structured monotone inclusions. Appl. Math. Optim. **80**(3), 547–598 (2019)

5. Apostol, R.Y., Grynenko, A.A., Semenov, V.V.: Iterative algorithms for monotone bilevel variational inequalities. J. Comput. Appl Math. **107**, 3–14 (2012)

6. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)

7. Bot, R.I., Csetnek, E.R., Hendrich, C.: Inertial Douglas–Rachford splitting for monotone inclusion. Appl. Math. Comput. **256**, 472–487 (2015)

8. Dong, Q.L., Lu, Y.Y., Yang, J.: The extragradient algorithm with inertial effects for solving the variational inequality. Optimization **65**, 2217–2226 (2016)

9. Ceng, L.C., Hadjisavvas, N., Wong, N.-C.: Strong convergence theorem by a hybrid extragradient-like approximation method for variational inequalities and fixed point problems. J. Glob. Optim. **46**, 635–646 (2010)

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

11. Censor, Y., Gibali, A., Reich, S.: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. Optim. Meth Softw. **26**, 827–845 (2011)

12. Censor, Y., Gibali, A., Reich, S.: Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. Optimization **61**, 1119–1132 (2011)

13. Chambolle, A., Dossal, Ch.: On the convergence of the iterates of the fast iterative shrinkage/thresholding algorithm. J. Optim. Theory Appl. **166**, 968–982 (2015)

14. Cholamjiak, W., Cholamjiak, P., Suantai, S.: An inertial forward-backward splitting method for solving inclusion problems in Hilbert spaces. J. Fixed Point Theory Appl. (2018). https://doi.org/10.1007/s11784-018-0526-5

15. Chuang, C.S.: Hybrid inertial proximal algorithm for the split variational inclusion problem in Hilbert spaces with applications. Optimization **66**(5), 777–792 (2017)

16. Fichera, G.: Sul pproblem elastostatico di signorini con ambigue condizioni al contorno. Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. **34** (1963), 138-142

17. Gibali, A., Jolaoso, L.O., Mewomo, O.T., Taiwo, A.: Fast and simple Bregman projection methods for solving variational inequalities and related problems in Banach spaces. Results Math., **75** (2020), Art. No. 179, 36 pp

18. Godwin, E.C., Izuchukwu, C., Mewomo, O.T.: An inertial extrapolation method for solving generalized split feasibility problems in real Hilbert spaces. Boll. Unione Mat. Ital. (2020). https://doi.org/10.1007/s40574-020-00

19. He, B.-S., Yang, Z.-H., Yuan, X.-M.: An approximate proximal-extragradient type method for monotone variational inequalities. J. Math. Anal. Appl. **300**, 362–374 (2004)

20. Izuchukwu, C., Mebawondu, A.A., Mewomo, O.T.: A new method for solving split variational inequality problems without co-coerciveness. J. Fixed Point Theory Appl.,**22** (4), (2020), Art. No. 98, 23 pp

21. Izuchukwu, C., Ogwo, G.N., Mewomo, O.T.: An inertial method for solving generalized split feasibility problems over the solution set of monotone variational inclusions. Optimization (2020). https://doi.org/10.1080/02331934.2020.1808648

22. Izuchukwu, C., Okeke, C.C., Mewomo, O.T.: Systems of Variational Inequalities and multiple-set split equality fixed point problems for countable families of multivalued type-one demicontractive-type mappings. Ukraïn. Mat. Zh. **71**(11), 1480–1501 (2019)

23. Iutzeler, F., Hendrickx, J.M.: A generic online acceleration scheme for optimization algorithms via relaxation and inertia. Optim. Methods Softw. **34**(2), 383–405 (2019)

24. Jolaoso, L.O., Taiwo, A., Alakoya, T.O., Mewomo, O.T.: Strong convergence theorem for solving pseudo-monotone variational inequality problem using projection method in a reflexive Banach space. J. Optim. Theory Appl. **185**(3), 744–766 (2020)

25. Khan, S.H., Alakoya, T.O., Mewomo, O.T.: Relaxed projection methods with self-adaptive step size for solving variational inequality and fixed point problems for an infinite family of multivalued relatively nonexpansive mappings in Banach Spaces. Math. Comput. Appl.,**25** (2020), Art. 54, 25 pp

26. Khobotov, E.N.: Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR Comput. Math. Math. Phys. **27**, 120–127 (1989)

27. Kinderlehrer, D., Stampacchia, G.: An introduction to variational inequalities and their applications. Academic Press, New York (1980)

28. Korpelevich, G.M.: An extragradient method for finding saddle points and for other problems. Ekon. Mat. Metody **12**, 747–756 (1976)

29. Liu anad, H., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput. Optim. Appl., **77** (2), (2020) 491-508

30. Kraikaew, R., Saejung, S.: Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces. J. Optim. Theory Appl. **163**(2), 399–412 (2014)

31. Lorenz, D.A., Pock, T.: An inertial forward-backward algorithm for monotone inclusions. J. Math. Imaging Vis. **51**, 311–325 (2015)
32. Mainge, P.E.: Convergence theorems for inertial KM-type algorithms. J. Comput. Appl. Math. **219**(1), 223–236 (2008)
33. Malitsky, Y.: Projected reflected gradient methods for monotone variational inequalities. SIAM J. Optim. **25**(1), 502–520 (2015)
34. Moudafi, A., Oliny, M.: Convergence of a splitting inertial proximal method for monotone operators. J. Comput. Appl. Math. **155**, 447–454 (2003)
35. Nesterov, Y.: A method of solving a convex programming problem with convergence rate O($1/k^2$). Soviet Math. Doklady **27**, 372–376 (1983)
36. Noor, M.: Extragradient Methods for pseudomonotone variational inequalities. J. Optim. Theory Appl. **117**, 475–488 (2003)
37. Ochs, P., Brox, T., Pock, T.: iPiasco: inertial proximal algorithm for strongly convex optimization. J. Math. Imaging Vis **53**, 171–181 (2015)
38. Polyak, B.T.: Some methods of speeding up the convergence of iterates methods. U.S.S.R Comput. Math. Phys.,**4** (5) (1964), 1-17
39. Shehu, Y., Li, X.H., Dong, Q.L.: An efficient projection-type method for monotone variational inequalities in Hilbert spaces. Numer. Algorithms **84**, 365–388 (2020)
40. Shehu, Y., Vuong, P.T., Zemkoho, A.: An inertial extrapolation method for convex simple bilevel optimization. Optim Methods. Softw. (2019) https://doi.org/10.1080/10556788.2019.1619729
41. Stampacchia, G.: Variational inequalities. In: Theory and Applications of Monotone Operators,Proceedings of the NATO Advanced Study Institute, Venice, Italy (Edizioni Odersi, Gubbio, Italy, 1968), 102–192
42. Sun, D.: A new step-size skill for solving a class of nonlinear projection equations. J. Comput. Math. **13**, 357–368 (1995)
43. Taiwo, A., Alakoya, T.O., Mewomo, O.T.: Halpern-type iterative process for solving split common fixed point and monotone variational inclusion problem between Banach spaces. Numer. Algorithms (2020). https://doi.org/10.1007/s11075-020-00937-2
44. Taiwo, A., Owolabi, A.O.-E., Jolaoso, L.O., Mewomo, O.T., Gibali, A.: A new approximation scheme for solving various split inverse problems. Afr. Mat. (2020). https://doi.org/10.1007/s13370-020-00832-y
45. Thong, D.V., Hieu, D.V.: A strong convergence of modified subgradient extragradient method for solving bilevel pseudomonotone variational inequality problems. Optimization **69**, 1313–1334 (2019)
46. Thong, D.V., Hieu, D.V.: Inertial subgradient extragradient algorithms with line-search process for solving variational inequality problems and fixed point problems. Numer. Algorithms **80**, 1283–1307 (2019)
47. Thong, D.V., Hieu, D.V.: An inertial method for solving split common fixed point problems. J. Fixed Point Theory Appl. **19**(4), 3029–3051 (2017)
48. Vuong, P.T.: On the weak convergence of the extragradient method for solving pseudomonotone variational inequalities. J. Optim. Theory Appl. **176**, 399–409 (2018)
49. Xia, Y., Wang, J.: A general methodology for designing globally convergent optimization neural networks. IEEE Trans. Neural Netw. **9**(6), 1331–1343 (1998)
50. Ye, M., He, Y.: A double projection method for solving variational inequalities without monotonicity. Comput. Optim. Appl. **60**(1), 141–150 (2015)
51. Yang, J.: Self-adaptive inertial subgradient extragradient algorithm for solving pseudomonotone variational inequalities. Appl. Anal. (2019). https://doi.org/10.1080/00036811.2019.1634257
52. Yang, J., Liu, H., Zexian, L.: Modified subgradient extragradient algorithms for solving monotone variational inequalities. Optimization **67**, 2247–2258 (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.