# Projection Methods with Alternating Inertial Steps for Variational Inequalities: Weak and Linear Convergence

Yekini Shehu*, Olaniyi S. Iyiola†

July 2, 2020

**Abstract**

The projection methods with vanilla inertial extrapolation step for variational inequalities have been of interest to many authors recently due to the improved convergence speed contributed by the presence of inertial extrapolation step. However, it is discovered that these projection methods with inertial steps lose the Fejér monotonicity of the iterates with respect to the solution, which is being enjoyed by their corresponding non-inertial projection methods for variational inequalities. This lack of Fejér monotonicity makes projection methods with vanilla inertial extrapolation step for variational inequalities not to converge faster than their corresponding non-inertial projection methods at times. Also, it has recently been proved that the projection methods with vanilla inertial extrapolation step may provide convergence rates that are worse than the classical projected gradient methods for strongly convex functions. In this paper, we introduce projection methods with alternated inertial extrapolation step for solving variational inequalities. We show that the sequence of iterates generated by our methods converges weakly to a solution of the variational inequality under some appropriate conditions. The Fejér monotonicity of even subsequence is recovered in these methods and linear rate of convergence is obtained. The numerical implementations of our methods compared with some other inertial projection methods show that our method is more efficient and outperforms some of these inertial projection methods.

**Keywords:** alternated inertial; projection methods; variational inequality; pseudo-monotone operator.

**2010 MSC classification:** 47H05, 47J20, 47J25, 65K15, 90C25.

*Department of Mathematics, Zhejiang Normal University, Jinhua, 321004, China; Institute of Science and Technology (IST), Am Campus 1, 3400, Klosterneuburg, Austria; e-mail: yekini.shehu@unn.edu.ng

†Department of Mathematics, Computer Science and Information Systems, California University of Pennsylvania, PA, USA; e-mail:niyi4oau@gmail.com; iyiola@calu.edu

## 1 Introduction

Let $H$ denote a real Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. Consider $C$ as a nonempty, closed and convex subset of $H$ and $A : C \to H$ a continuous mapping. The variational inequality problem (for short, VI$(A, C)$) is defined as: find $x \in C$ such that

$$\langle Ax, y - x \rangle \geq 0, \quad \forall y \in C. \tag{1}$$

We shall denote by $S$ the set of solutions of VI$(A, C)$ (1). Several applications of VI$(A, C)$ (1) are discussed in [2, 4, 18, 25, 26, 27, 32].

A point $x \in C$ is a solution of VI$(A, C)$ (1) if and only if (see [18] for the details)

$$x = P_C(x - \gamma Ax), \gamma > 0 \;\text{ and }\; r_\gamma(x) := x - P_C(x - \gamma Ax) = 0.$$

This led to the introduction of fixed point approach to solve VI$(A, C)$ (1) ((see, e.g., [16, 27, 36]).

If $A$ is $\eta$-strongly monotone and $L$-Lipschitz-continuous, then the sequence generated by gradient-projection method

$$x_{n+1} = P_C(x_n - \lambda Ax_n) \tag{2}$$

converges to a solution of VI$(A, C)$ (1) if the step-size $\lambda \in (0, \frac{2\eta}{L^2})$. The gradient-projection method (2) fails if $A$ is monotone. For example, take $C = \mathbb{R}^2$ and $A$ a rotation with $\frac{\pi}{2}$ angle. Then $A$ is monotone and $L$-Lipschitz-continuous and $(0, 0)$ is the unique solution of VI$(A, C)$ (1). However, $\{x_n\}$ generated by gradient-projection method (2) satisfies the property $\|x_{n+1}\| > \|x_n\|$ for all $n$.

In [28], Korpelevich introduced the extra-gradient method, which is:

$$\begin{cases} x_1 \in C, \\ y_n = P_C(x_n - \lambda_n Ax_n) \\ x_{n+1} = P_C(x_n - \lambda_n Ay_n), \quad n \geq 1, \end{cases} \tag{3}$$

where $\lambda_n \in (0, \frac{1}{L})$. It is shown in [28] that $\{x_n\}$ converges to a solution of VI$(A, C)$ (1) when $A$ is monotone and $L$-Lipschitz-continuous. Similar results are found in [1, 8, 11, 17, 20, 21, 35, 38, 42, 44].

A question of interest in projection method for of VI$(A, C)$ (1) is how to minimize the number of projections per iterations in extra-gradient method (3). This is because if $P_C$ does not have a closed form formula, then a minimization problem has to be solved twice per iteration in implementing extra-gradient method (3). In such situation, the efficiency of the extra-gradient method (3) is affected. This has led to introducing some projection methods with one projection per iteration in solving VI$(A, C)$ (1).

In [12], Censor et al. introduced the subgradient extragradient method: $x_1 \in H$,

$$\begin{cases} y_n = P_C(x_n - \lambda_n A x_n), \\ T_n := \{w \in H : \langle x_n - \lambda_n A x_n - y_n, w - y_n \rangle \leq 0\}, \\ x_{n+1} = P_{T_n}(x_n - \lambda_n A y_n). \end{cases} \tag{4}$$

Censor et al. [12] proved that $\{x_n\}$ generated by (4) converges weakly to a solution of VI$(A, C)$ (1) under some appropriate conditions on $\{\lambda_n\}$.

Recently, Maingé and Gobinddass [30], motivated by Malitsky [31] introduced the following iterative method for solving VI$(A, C)$ (1): choose $\delta \in (0, 1], \lambda_{-1} \in (0, \infty)$ and $x_1, x_0 \in C$

$$\begin{cases} y_n = x_n + \frac{\lambda_n}{\delta \lambda_{n-1}}(x_n - x_{n-1}), \\ x_{n+1} = P_C(x_n - \lambda_n A y_n), \end{cases} \tag{5}$$

and showed that $\{x_n\}$ generated by (5) converges weakly to a solution of VI$(A, C)$ (1) when $A$ is monotone and $L$-Lipschitz-continuous. The method (5) requires one projection $P_C$ onto $C$ and no further projections onto the half-space unlike (4).

When $A$ is pseudo-monotone and $L$-Lipschitz-continuous in VI$(A, C)$ (1), Ceng *et al.* [9] introduced the following method (here, we take $\alpha_n = 0$ and $S_n =$ the identity mapping in [9, Theorem 3.1]):

$$\begin{cases} x_1 \in C, \\ y_n = P_C(x_n - \lambda_n A x_n), \\ z_n = P_C(x_n - \lambda_n A y_n), \\ C_n = \{z \in C : \|z_n - z\| \leq \|x_n - z\|\}, \\ \text{find } x_{n+1} \in C_n \text{ such that} \\ \langle x_n - x_{n+1} + e_n - \sigma_n A x_{n+1}, x_{n+1} - x \rangle \geq -\epsilon_n, \quad \forall x \in C_n, \end{cases} \tag{6}$$

where $\{e_n\}$ is an error sequence, $\{\sigma_n\} \subset (0, \frac{1}{L})$ and $\{\epsilon_n\} \subset [0, \infty)$. The intuition of the last step of (6) comes from the approximate proximal methods. Ceng *et al.* [9] showed that $\{x_n\}$ generated by (6) converges weakly to a solution of VI$(A, C)$ (1) and under the condition (see Yao and Postolache [45] also) that $0 \leq \liminf_{n\to\infty} \langle Ax_n, z - x_n \rangle, \forall z \in C$. This method (6) of Ceng *et al.* [9] requires computations of projection twice per iteration.

The extragradient method (3) and subgradient extragradient method (4) are mentioned for a brief history lesson of our introduction. Equation (5) and (6) are mentioned as methods with weaknesses (lack of convergence proof for pseudo-monotone for (5) and two projection steps for (6)). Methods from [14, 15, 39, 41] are compared against our proposed methods because they only require one projection per step and are proven to converge for pseudo-monotone problems. Now, we give some discussion on the methods proposed in [14, 15, 41].

## 1.1 Contributions and Related Work

He [19] introduced the projection and contraction method for solving VI$(A, C)$ (1):

$$\begin{cases} y_n = P_C(x_n - \lambda A x_n), \\ d_n = x_n - y_n - \lambda(Ax_n - Ay_n), \\ x_{n+1} = x_n - \gamma \eta_n d_n, \end{cases} \tag{7}$$

where $\gamma \in (0, 2)$ and $\{\eta_n\}$ is given by

$$\eta_n = \begin{cases} \frac{\langle x_n - y_n, d_n \rangle}{\|d_n\|^2}, & d_n \neq 0 \\ 0, & d_n = 0. \end{cases}$$

In [15], Dong *et al.* proposed the following inertial projection and contraction method for VI$(A, C)$ (1):

$$\begin{cases} w_n = x_n + \alpha_n(x_n - x_{n-1}), \\ y_n = P_C(w_n - \lambda A w_n), \\ d_n = w_n - y_n - \lambda(Aw_n - Ay_n), \\ x_{n+1} = w_n - \gamma \eta_n d_n, \end{cases} \tag{8}$$

where $w_n = x_n + \alpha_n(x_n - x_{n-1})$ is the inertial step and $\{\eta_n\}$ is given by

$$\eta_n = \begin{cases} \frac{\langle w_n - y_n, d_n \rangle}{\|d_n\|^2}, & d_n \neq 0 \\ 0, & d_n = 0, \end{cases}$$

$0 \leq \alpha_n \leq \alpha_{n+1} \leq \alpha < 1$ with $\sigma, \delta > 0$ such that
(a) $\delta > \frac{\alpha^2(1+\alpha)+\alpha\sigma}{1-\alpha^2}$; and
(b) $0 < \gamma \leq \frac{2[\delta - \alpha[\alpha(1+\alpha)+\alpha\delta+\sigma]]}{\delta[1+\alpha(1+\alpha)+\alpha\delta+\sigma]}$. Under conditions (a) and (b), Dong *et al.* [15] showed that $\{x_n\}$ generated by (8) converges weakly to a solution of VI$(A, C)$ (1).

When the inertial factor $\{\alpha_n\}$ is chosen such that $0 \leq \alpha_n \leq \bar{\alpha_n}$, where

$$\bar{\alpha_n} = \begin{cases} \min\left\{\alpha, \frac{\epsilon_n}{\|x_n - x_{n-1}\|}\right\}, & x_n \neq x_{n-1} \\ \alpha, & \text{otherwise} \end{cases}$$

with $\alpha \in [0, 1)$ and $\sum_{n=1}^{\infty} \epsilon_n < \infty$, appropriate convergence results of (8) to a solution of VI$(A, C)$ (1) have been obtained in [14, 41]. The results in [14, 15, 41] all extend the result of He [19] when $\alpha_n = 0$ and they have been shown numerically to improve the speed of convergence of the projection and contraction method studied in [7, 19].

In all the inertial projection methods proposed in [14, 15, 41], it is seen that the Fejér monotonicity of $\|x_n - x^*\|, x^* \in S$ is lost and this makes $\{x_n\}$ generated by the methods in [14, 15, 41] to move or swing back and forth around $S$. This furthermore makes these methods sometimes not converge faster than their counterpart non-inertial methods.

It is our aim in this paper to propose an inertial projection method in which Fejér monotonicity of $\|x_n - x^*\|, x^* \in S$ is regained to some extent. We show that $\{x_n\}$ generated by our proposed method converges to a point in $S$ under some mild assumptions. In simple terms, our contributions in this paper are:

- We propose a projection-type algorithm with alternated inertial step which requires one evaluation of projection onto the feasible set $C$ per iteration. The inertial extrapolation step proposed is different from the famous vanilla inertial extrapolation step proposed in [14, 15, 41] for solving VI$(A, C)$ (1). Our method is particularly useful in cases when computation of projection onto the feasible set is difficult.

- Our proposed method for solving VI$(A, C)$ (1) assumes the operator is pseudo-monotone and therefore more applicable than the methods in [29, 31]) where $A$ is assumed to be monotone.

- In our proposed algorithm, the inertial factor $\alpha_n \geq 1$ is possible. This is not allowed in many other proposed projection-type methods with inertial extrapolation step in the literature, where $\alpha_n < 1$ (see, for example, [14, 15, 41]). Therefore, our proposed method brings novelty and state of the art contributions to inertial projection-type method for solving VI$(A, C)$ (1).

- Linear convergence, a priori and a posteriori error estimates of generated sequences are given in the special case when $A$ is strongly-pseudo-monotone.

- We give some carefully designed computational experiments to show that our proposed method is efficient and outperforms some related inertial projection-type methods.

Organization of the paper: Some definitions and results are given in Section 2. Some discussions about our proposed methods are given in Section 3. The proof of global convergence of our algorithms are given in Section 4 and the linear convergence analysis is given in Section 5. We give numerical implementations in Section 6. We end the paper with some remarks in Section 7.

# 2 Preliminaries

**Definition 2.1.** *A mapping $A : H \to H$ is called*

(a) $\eta$-strongly monotone *on $H$ if there exists a constant $\eta > 0$ such that $\langle Ax - Ay, x - y\rangle \geq \eta\|x - y\|^2$, for all $x, y \in H$;*

(b) monotone *on $H$ if $\langle Ax - Ay, x - y\rangle \geq 0$ for all $x, y \in H$;*

(c) $\delta$-pseudo-monotone *on $H$ if there exists $\delta > 0$ such that $\langle Ay, x - y\rangle \geq 0 \Rightarrow \langle Ax, x - y\rangle \geq \delta\|x - y\|^2, \quad x, y \in H$;*

(d) pseudo-monotone *on $H$ if, for all $x, y \in H$, $\langle Ay, x-y\rangle \geq 0 \Rightarrow \langle Ax, x-y\rangle \geq 0$;*

(e) $L$-Lipschitz-continuous *on $H$ if there exists a constant $L > 0$ such that $\|Ax - Ay\| \leq L\|x - y\|$ for all $x, y \in H$.*

(f) *sequentially weakly continuous if for each sequence $\{x_n\}$ we have: $\{x_n\}$ converges weakly to $x$ implies $\{Ax_n\}$ converges weakly to $Ax$.*

**Remark 2.2.** Note that (a) implies (b), (a) implies (c), (c) implies (d) and (b) implies (d) in the above definitions. Furthermore, if (c) is satisfied, then VI($A, C$) (1) has a unique solution.

**Definition 2.3.** *Let $C$ be a nonempty, closed and convex subset of $H$. $P_C$ is called the metric projection of $H$ onto $C$ if, for any point $u \in H$, there exists a unique point $P_C u \in C$ such that*

$$\|u - P_C u\| \leq \|u - y\| \quad \forall y \in C.$$

$P_C$ satisfies (see, e.g., [5])

$$\langle x - y, P_C x - P_C y \rangle \geq \|P_C x - P_C y\|^2 \quad \forall x, y \in H. \tag{9}$$

Furthermore, $P_C x$ is characterized by the properties

$$P_C x \in C \quad \text{and} \quad \langle x - P_C x, P_C x - y \rangle \geq 0 \quad \forall y \in C. \tag{10}$$

This characterization implies that

$$\|x - y\|^2 \geq \|x - P_C x\|^2 + \|y - P_C x\|^2 \quad \forall x \in H, \forall y \in C. \tag{11}$$

**Lemma 2.4.** *The following statements hold in $H$:*

(i) $\|x + y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2$ *for all* $x, y \in H$*;*

(ii) $\|x + y\|^2 \leq \|x\|^2 + 2\langle y, x + y \rangle$ *for all* $x, y \in H$

*(iii)* $\|\alpha x + \beta y\|^2 = \alpha(\alpha + \beta)\|x\|^2 + \beta(\alpha + \beta)\|y\|^2 - \alpha\beta\|x - y\|^2, \quad \forall x, y \in H, \forall \alpha, \beta \in \mathbb{R}.$

**Lemma 2.5.** *([33, Lem. 2.2]) Suppose $A$ is pseudo-monotone in VI($A, C$)* (1)*. Then $S$ is closed, convex and $M(A, C) = S$, where $M(A, C) := \{x \in C : \langle Ay, y - x \rangle \geq 0, \ \ \forall y \in C\}$.*

We remark that the definition of $M(A, C)$ in Lemma 2.5 differs from the definition of $S$ in VI($A, C$) (1). Observe that this change is necessary when considering pseudo-monotone rather than monotone variational inequalities. Using Definition (1) would make all fixed points solutions to the variational inequalities. For example, consider VI($A, C$) where $A = d/dx[-\cos(x)]$ with $C = [-\pi, \pi]$.

**Definition 2.6.** *A sequence $\{x_n\}$ in $H$ is said to converge weakly to $p \in H$ if*

$$\forall z \in H, \lim_{n \to \infty} \langle x_n, z \rangle = \langle p, z \rangle.$$

**Lemma 2.7.** *([37]) 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 any $x \in C$, $\lim_{n \to \infty} \|x_n - x\|$ exists;*
*(ii) every sequential weak cluster point of $\{x_n\}$ is in $C$.*
*Then $\{x_n\}$ converges weakly to a point in $C$.*

**Definition 2.8.** *A sequence $\{x_n\}$ is Fejér monotone with respect to a set $S$ if each point in the sequence is not strictly farther from any point in $S$ than its predecessor. In other words,*

$$\|x_{n+1} - z\| \leq \|x_n - z\|, \forall z \in S.$$

**Definition 2.9.** *Suppose a sequence $\{x_n\}$ in $H$ converges in norm to $x^* \in H$. We say that $\{x_n\}$ converges to $x^*$ R-linearly if $\limsup_{n\to\infty} \|x_n - x^*\|^{\frac{1}{n}} < 1$. We say that $\{x_n\}$ converges to $x^*$ Q-linearly if there exists $\mu \in (0,1)$ such that $\|x_{n+1} - x^*\| \leq \mu\|x_n - x^*\|$ for all sufficiently large $n$. It is well known that Q-linear convergence implies R-linear convergence, but the reverse implication is not true.*

# 3 Proposed Methods

In this section, we introduce our projection-type method with alternated inertial extrapolation step and give some discussions.

**Assumption 3.1.** In this section and the next, let us assume that the following assumptions are satisfied:

(a) The feasible set $C$ is a nonempty, closed, and convex subset of $H$.

(b) $A : H \to H$ is pseudo-monotone, sequentially weakly continuous and $L$-Lipschitz-continuous.

(c) The solution set $S$ of VI$(A, C)$ (1) is nonempty.

**Assumption 3.2.** For the first proposed method, we assume that the iterative parameters satisfy these conditions:

(a) $\gamma \in (0, 2)$

(b) $0 \leq \alpha_n \leq \alpha < \frac{2-\gamma}{\gamma}$

(c) $\lambda \in (0, \frac{1}{L})$

Now, our first proposed method is introduced.

**Algorithm 1** Inertial Projection Method with Fixed Step-size

1: Choose the iterative parameters $\alpha_n, \gamma$ and $\lambda$ such that Assumption 3.2 hold. Let $x_0, x_1 \in H$ be given starting points. Set $n := 1$.

2: Compute

$$w_n = \begin{cases} x_n, & n = \text{even} \\ x_n + \alpha_n(x_n - x_{n-1}), & n = \text{odd}. \end{cases}$$

and

$$y_n := P_C(w_n - \lambda A w_n). \tag{12}$$

If $\|w_n - y_n\| = 0$ or $\|Ay_n\| = 0$, STOP. Otherwise

3: Compute

$$d_n = w_n - y_n - \lambda(Aw_n - Ay_n), \quad \forall n \geq 1. \tag{13}$$

4: Compute

$$x_{n+1} = w_n - \gamma \eta_n d_n, \quad n \geq 1, \tag{14}$$

where $\{\eta_n\}$ is given by

$$\eta_n = \begin{cases} \frac{\langle w_n - y_n, d_n \rangle}{\|d_n\|^2}, & d_n \neq 0 \\ 0, & d_n = 0. \end{cases}$$

5: Set $n \leftarrow n + 1$, and **go to 2**.

**Remark 3.3.**

(a) We give some intuition for Step 3 (Equation 13) of Algorithm 1. This step can be considered a weighted average of $(w_n - y_n \sim \lambda A w_n)$ and a hypothetical $(w'_n - y'_n \sim \lambda A w'_n)$ where $w'_n = w_n - \lambda A w_n$ and $y'_n = y_n - \lambda A y_n$. This looks similar to Heun's method or "improved" Euler from numerical methods for ODEs (please see [40, page 328] for more details).

(b) From Algorithm 1, we have that $x_n \in H$ and $y_n \in C$. In Step 4 of Algorithm 1, $\eta_n$ is the vector projection of $(w_n - y_n)$ onto the direction $d_n$. Thus, large steps in the direction $d_n$ are only taken if $d_n$ and the vanilla projected gradient direction agree.

(c) Algorithm 1 requires, at each iteration, only one projection onto the feasible set $C$ and it is different from other methods in [8, 9, 10, 11, 13, 29, 45] where more than one projection per iteration is needed.

(d) Our proposed Algorithm 1 allows the inertial factor $\alpha_n \geq 1$ (e.g., take $\gamma = \frac{1}{2}$) which is not allowed in many other proposed projection-type methods with inertial extrapolation step in the literature, where $\alpha_n < 1$ (see, for example, [14, 15, 41]). In fact, in our method, one can choose $\alpha_n > 1$ (when $\gamma < 1$). This brings novelty and state of the art contributions to inertial projection-type methods for solving VI$(A,C)$ (1) in terms of empirical convergence rate (as confirmed by the numerical examples in Section 6). $\Diamond$

In the case when the Lipschitz constant $L$ of $A$ is not available, we propose the following method with adaptive step-size.

**Algorithm 2** Inertial Projection Method with Adaptive Step-size

1: Choose the iterative parameters $\alpha_n, \gamma$ such that Assumption 3.2 (a)-(b) hold, $\mu \in (0,1)$ and $\lambda_1 > 0$. Let $x_0, x_1 \in H$ be given starting points. Set $n := 1$.

2: Compute

$$w_n = \begin{cases} x_n, & n = \text{even} \\ x_n + \alpha_n(x_n - x_{n-1}), & n = \text{odd}. \end{cases}$$

and

$$y_n := P_C(w_n - \lambda_n A w_n), \tag{15}$$

where

$$\lambda_{n+1} = \begin{cases} \min\left\{\frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|}, \lambda_n\right\}, & Aw_n \neq Ay_n \\ \lambda_n, & \text{otherwise.} \end{cases} \tag{16}$$

If $\|w_n - y_n\| = 0$ or $\|Ay_n\| = 0$, STOP. Otherwise

3: Compute

$$d_n = w_n - y_n - \lambda_n(Aw_n - Ay_n), \quad \forall n \geq 1. \tag{17}$$

4: Compute

$$x_{n+1} = w_n - \gamma\eta_n d_n, \quad n \geq 1, \tag{18}$$

where $\{\eta_n\}$ is given by

$$\eta_n = \begin{cases} \frac{\langle w_n - y_n, d_n\rangle}{\|d_n\|^2}, & d_n \neq 0 \\ 0, & d_n = 0. \end{cases}$$

5: Set $n \leftarrow n+1$, and **go to 2**.

**Remark 3.4.** Note that by (16), $\lambda_{n+1} \leq \lambda_n, \quad \forall n \geq 1$. Also, observe in Algorithm 2 that if $Aw_n \neq Ay_n$, then

$$\frac{\mu\|w_n - y_n\|}{\|Aw_n - Ay_n\|} \geq \frac{\mu}{L}\frac{\|w_n - y_n\|}{\|w_n - y_n\|} = \frac{\mu}{L}$$

which implies that $0 < \min\left\{\lambda_1, \frac{\mu}{L}\right\} \leq \lambda_n \quad \forall n \geq 1$. This means that $\lim_{n\to\infty} \lambda_n$ exists. Thus, there exists $\lambda > 0$ such that $\lim_{n\to\infty} \lambda_n = \lambda$.

## 4 Convergence Analysis

Next, we show that the sequence $\{x_n\}$ generated by our Algorithm 1 and Algorithm 2 converges weakly to a point in $S$ under Assumptions 3.1 and 3.2. To achieve this, we first establish some lemmas below.

**Lemma 4.1.** *Suppose $\{x_n\}$ is generated by Algorithm 1. Then under under Assumptions 3.1 and 3.2, $\{x_{2n}\}$ is Fejér monotone with respect to $S$ and $\lim_{n\to\infty} \|x_{2n} - x^*\|$ exists, where $x^* \in S$, the solution set of VI$(A, C)$ (1).*

*Proof.* Choose $x^* \in S$. Then

$$x_{2n+2} = w_{2n+1} - \gamma\eta_{2n+1} d_{2n+1},$$

$$d_{2n+1} = w_{2n+1} - y_{2n+1} - \lambda(Aw_{2n+1} - Ay_{2n+1})$$

and

$$\eta_{2n+1} = \begin{cases} \frac{\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle}{\|d_{2n+1}\|^2}, & d_{2n+1} \neq 0 \\ 0, & d_{2n+1} = 0. \end{cases}$$

So,

$$\begin{aligned} \|x_{2n+2} - x^*\|^2 &= \|w_{2n+1} - \gamma\eta_{2n+1} d_{2n+1} - x^*\|^2 \\ &= \|(w_{2n+1} - x^*) - \gamma\eta_{2n+1} d_{2n+1}\|^2 \\ &= \|w_{2n+1} - x^*\|^2 - 2\gamma\eta_{2n+1}\langle w_{2n+1} - x^*, d_{2n+1}\rangle \\ &\quad +\gamma^2\eta_{2n+1}^2\|d_{2n+1}\|^2. \end{aligned} \tag{19}$$

Note that

$$\langle w_{2n+1} - x^*, d_{2n+1}\rangle = \langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle + \langle y_{2n+1} - x^*, d_{2n+1}\rangle. \tag{20}$$

By $y_{2n+1} = P_C(w_{2n+1} - \lambda Aw_{2n+1})$ and property (10), we get (since $x^* \in S$)

$$\langle y_{2n+1} - x^*, w_{2n+1} - y_{2n+1} - \lambda Aw_{2n+1}\rangle \geq 0. \tag{21}$$

By the pseudomonotonicity of $A$:

$$\langle Ay_{2n+1}, y_{2n+1} - x^*\rangle \geq 0. \tag{22}$$

$\langle Ax^*, y_{2n+1} - x^*\rangle \geq 0$ (see Equation (1)) and $\lambda > 0$, therefore,

$$\langle \lambda Ay_{2n+1}, y_{2n+1} - x^*\rangle \geq 0. \tag{23}$$

Adding (21) and (23), we obtain

$$\langle y_{2n+1} - x^*, w_{2n+1} - y_{2n+1} - \lambda Aw_{2n+1} + \lambda Ay_{2n+1}\rangle \geq 0.$$

i.e.,

$$\langle y_{2n+1} - x^*, d_{2n+1}\rangle \geq 0. \tag{24}$$

Combining (20) and (24) gives

$$\langle w_{2n+1} - x^*, d_{2n+1}\rangle \geq \langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle. \tag{25}$$

Put (25) into (19) $\left(\text{noting that } \eta_{2n+1} = \frac{\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle}{\|d_{2n+1}\|^2}\right)$:

$$\begin{aligned}
\|x_{2n+2} - x^*\|^2 &\leq \|w_{2n+1} - x^*\|^2 - 2\gamma\eta_{2n+1}\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle + \gamma^2\eta_{2n+1}^2\|d_{2n+1}\|^2 \\
&= \|w_{2n+1} - x^*\|^2 - 2\gamma\eta_{2n+1}\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle \\
&\quad +\gamma^2\eta_{2n+1}\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle \\
&= \|w_{2n+1} - x^*\|^2 - \gamma(2-\gamma)\eta_{2n+1}\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle. \qquad (26)
\end{aligned}$$

Observe that

$$\begin{aligned}
\eta_{2n+1}\langle w_{2n+1} - y_{2n+1}, d_{2n+1}\rangle &= \|\eta_{2n+1}d_{2n+1}\|^2 \\
&= \frac{1}{\gamma^2}\|x_{2n+2} - w_{2n+1}\|^2. \qquad (27)
\end{aligned}$$

From (26) and (27), we get

$$\|x_{2n+2} - x^*\|^2 \leq \|w_{2n+1} - x^*\|^2 - \frac{(2-\gamma)}{\gamma}\|x_{2n+2} - w_{2n+1}\|^2. \tag{28}$$

Now, by Lemma 2.4 (iii), we get

$$\begin{aligned}
\|w_{2n+1} - x^*\|^2 &= \|x_{2n+1} + \alpha_{2n+1}(x_{2n+1} - x_{2n}) - x^*\|^2 \\
&= \|(1+\alpha_{2n+1})(x_{2n+1} - x^*) - \alpha_{2n+1}(x_{2n} - x^*)\|^2 \\
&= (1+\alpha_{2n+1})\|x_{2n+1} - x^*\|^2 - \alpha_{2n+1}\|x_{2n} - x^*\|^2 \\
&\quad +\alpha_{2n+1}(1+\alpha_{2n+1})\|x_{2n+1} - x_{2n}\|^2. \qquad (29)
\end{aligned}$$

By following the same line of argument in obtaining (28), one can show:

$$\begin{aligned}
\|x_{2n+1} - x^*\|^2 &\leq \|w_{2n} - x^*\| - \frac{(2-\gamma)}{\gamma}\|x_{2n+1} - w_{2n}\|^2 \\
&= \|x_{2n} - x^*\|^2 - \frac{(2-\gamma)}{\gamma}\|x_{2n+1} - x_{2n}\|^2. \qquad (30)
\end{aligned}$$

Using (30) and (29):

$$\begin{aligned}
\|w_{2n+1} - x^*\|^2 &\leq (1+\alpha_{2n+1})\left[\|x_{2n} - x^*\|^2 - \frac{2-\gamma}{\gamma}\|x_{2n+1} - x_{2n}\|^2\right] \\
&\quad -\alpha_{2n+1}\|x_{2n} - x^*\|^2 + \alpha_{2n+1}(1+\alpha_{2n+1})\|x_{2n+1} - x_{2n}\|^2 \\
&= \|x_{2n} - x^*\|^2 - (1+\alpha_{2n+1})\left(\frac{2-\gamma}{\gamma} - \alpha_{2n+1}\right)\|x_{2n+1} - x_{2n}\|^2 \qquad (31)
\end{aligned}$$

Using (31) in (28), we have

$$\|x_{2n+2} - x^*\|^2 \leq \|x_{2n} - x^*\|^2 - (1+\alpha_{2n+1})\left(\frac{2-\gamma}{\gamma} - \alpha_{2n+1}\right)\|x_{2n+1} - x_{2n}\|^2$$

$$-\frac{(2-\gamma)}{\gamma}\|x_{2n+2}-w_{2n+1}\|^2. \quad (32)$$

Since $\alpha_{2n} \leq \alpha < \frac{2-\gamma}{\gamma}$, we get from (32) that

$$\|x_{2n+2}-x^*\| \leq \|x_{2n}-x^*\|.$$

This implies that $\{\|x_{2n}-x^*\|\}$ and $\{x_{2n}\}$ are bounded. Furthermore, $\lim_{n\to\infty}\|x_{2n}-x^*\|$ exists.

□

**Lemma 4.2.** *Suppose $\{x_n\}$ and $\{y_n\}$ are generated by Algorithm 1. Then under under Assumptions 3.1 and 3.2,* $\lim_{n\to\infty}\|x_{2n}-y_{2n}\| = 0$.

*Proof.* Rearranging (32) and using the fact that $\|x_{2n}-x^*\|$ is bounded, we get

$$\lim_{n\to\infty}\|x_{2n}-x_{2n+1}\| = 0. \quad (33)$$

Now,

$$\begin{aligned}\|d_{2n}\| &= \|w_{2n}-y_{2n}-\lambda(Aw_{2n}-Ay_{2n})\|\\ &\leq \|w_{2n}-y_{2n}\|+\lambda\|Aw_{2n}-Ay_{2n}\|\\ &\leq (1+\lambda L)\|w_{2n}-y_{2n}\|.\end{aligned}$$

This means

$$\frac{1}{\|d_{2n}\|} \geq \frac{1}{(1+\lambda L)\|w_{2n}-y_{2n}\|}. \quad (34)$$

Now

$$\begin{aligned}\langle w_{2n}-y_{2n}, d_{2n}\rangle &= \langle w_{2n}-y_{2n}, w_{2n}-y_{2n}-\lambda(Aw_{2n}-Ay_{2n})\rangle\\ &= \|w_{2n}-y_{2n}\|^2-\langle w_{2n}-y_{2n}, \lambda(Aw_{2n}-Ay_{2n})\rangle\\ &\geq \|w_{2n}-y_{2n}\|^2-\lambda L\|w_{2n}-y_{2n}\|^2\\ &= (1-\lambda L)\|w_{2n}-y_{2n}\|^2. \quad (35)\end{aligned}$$

Using (34) and (35),

$$\begin{aligned}\|x_{2n+1}-w_{2n}\| &= \gamma\eta_{2n}\|d_{2n}\| = \gamma\frac{\langle w_{2n}-y_{2n}, d_{2n}\rangle}{\|d_{2n}\|}\\ &\geq \gamma\frac{1-\lambda L}{1+\lambda L}\|w_{2n}-y_{2n}\|. \quad (36)\end{aligned}$$

Using (33) in (36), we get (noting that $w_{2n} = x_{2n}$)

$$\lim_{n\to\infty}\|x_{2n}-y_{2n}\| = 0.$$

□

We develop the technique in [43] to obtain the following result.

**Lemma 4.3.** *Assume that $\{x_n\}$ is generated by Algorithm 1. Let $p \in H$ denote the weak limit of the subsequence $\{x_{2n_k}\}$ of $\{x_{2n}\}$. Then $p \in S$.*

*Proof.* Since $\{x_{2n}\}$ is bounded by Lemma 4.1, then there exits a subsequence $\{x_{2n_k}\} \subset \{x_{2n}\}$ such that $x_{2n_k} \rightharpoonup p \in H$. Using the definition of $y_{2n_k}$ and (10) (noting that $w_{2n} = x_{2n}$), we have

$$\langle x_{2n_k} - \lambda A x_{2n_k} - y_{2n_k}, x - y_{2n_k} \rangle \leq 0, \ \forall x \in C,$$

and so

$$\frac{1}{\lambda} \langle x_{2n_k} - y_{2n_k}, x - y_{2n_k} \rangle \leq \langle A x_{2n_k}, x - y_{2n_k} \rangle, \ \forall x \in C.$$

Hence,

$$\frac{1}{\lambda} \langle x_{2n_k} - y_{2n_k}, x - y_{2n_k} \rangle + \langle A x_{2n_k}, y_{2n_k} - x_{2n_k} \rangle \leq \langle A x_{2n_k}, x - x_{2n_k} \rangle, \ \forall x \in C. \quad (37)$$

Fixing $x \in C$ and letting $k \to \infty$ in (37), we get (noting $\lim_{k\to\infty} \|x_{2n_k} - y_{2n_k}\| = 0$ by Lemma 4.2), we obtain

$$0 \leq \liminf_{k\to\infty} \langle A x_{2n_k}, x - x_{2n_k} \rangle, \quad \forall x \in C. \quad (38)$$

Let us choose a decreasing sequence $\{\epsilon_k\} \subset (0, \infty)$ such that $\lim\limits_{k\to\infty} \epsilon_k = 0$. For each $\epsilon_k$, we denote by $N_k$ the smallest positive integer such that

$$\left\langle A x_{2n_j}, x - x_{2n_j} \right\rangle + \epsilon_k \geq 0 \quad \forall j \geq N_k, \quad (39)$$

where the existence of $N_k$ follows from (38). Since $\{\epsilon_k\}$ is decreasing, then $\{N_k\}$ is increasing. Also, for each $k$, $A x_{2N_k} \neq 0$ and, setting

$$v_{2N_k} = \frac{A x_{2N_k}}{\|A x_{2N_k}\|^2},$$

one gets $\langle A x_{2N_k}, v_{2N_k} \rangle = 1$ for each $k$. Then by (39), we have for each $k$

$$\langle A x_{2N_k}, x + \epsilon_k v_{2N_k} - x_{2N_k} \rangle \geq 0.$$

By the fact that $A$ is pseudo-monotone, we get

$$\langle A(x + \epsilon_k v_{2N_k}), x + \epsilon_k v_{2N_k} - x_{2N_k} \rangle \geq 0. \quad (40)$$

Since $\{x_{2n_k}\}$ converges weakly to $p$ as $k \to \infty$ and $A$ is sequentially weakly continuous on $H$, we have that $\{A x_{2n_k}\}$ converges weakly to $Ap$. Suppose $Ap \neq 0$ (otherwise, $p \in S$). Then by the sequentially weakly lower semicontinuity of norm, we get

$$0 < \|Ap\| = \liminf_{k\to\infty} \|A x_{2n_k}\|.$$

Since $\{x_{N_k}\} \subset \{x_{n_k}\}$ and $\epsilon_k \to 0$ as $k \to \infty$, we get

$$0 \ \leq \ \limsup_{k\to\infty} \|\epsilon_k v_{2N_k}\| = \limsup_{k\to\infty} \left( \frac{\epsilon_k}{\|A x_{2n_k}\|} \right)$$

$$\leq \frac{\limsup_{k\to\infty}\epsilon_k}{\liminf_{k\to\infty}\|Ax_{2n_k}\|} \leq \frac{0}{\|Ap\|} = 0,$$

and this means $\lim_{k\to\infty}\|\epsilon_k v_{2N_k}\| = 0$. Passing the limit $k \to \infty$ in (40), we get

$$\langle Ax, x-p\rangle \geq 0.$$

In view of Lemma 2.5, this implies $p \in S$. □

**Theorem 4.4.** *Suppose $\{x_n\}$ is generated by Algorithm 1. Then under under Assumptions 3.1 and 3.2, $\{x_n\}$ converges weakly to a point in $S$.*

*Proof.* Since $\{x_{2n}\}$ is bounded by Lemma 4.1, then $\{x_{2n}\}$ has weakly convergent subsequences. Suppose $p \in H$ denotes the weak limit of such a subsequence $\{x_{2n_k}\}$ of $\{x_{2n}\}$. By Lemma 4.3, we have $p \in S$. Also, by Lemma 4.1, we get $\lim_{n\to\infty}\|x_{2n}-p\|$ exists. It now follows from Lemma 2.7 that the whole sequence $\{x_{2n}\}$ converges weakly to a point in $S$. Suppose $\{x_{2n}\}$ converges weakly to $p \in S$ and $\{x_{2n}\}$ converges weakly to $q \in S$. Then

$$\begin{aligned}
\|p-q\|^2 &= \langle p-q, p-q\rangle \\
&= \langle p, p-q\rangle - \langle q, p-q\rangle \\
&= \lim_{n\to\infty}\langle x_{2n}, p-q\rangle - \lim_{n\to\infty}\langle x_{2n}, p-q\rangle \\
&= \lim_{n\to\infty}\langle x_{2n}-x_{2n}, p-q\rangle = 0.
\end{aligned}$$

Thus, the weak limit $p$ is unique. By definition, we have that for all $z \in H$,

$$\lim_{n\to\infty}\langle x_{2n}-p, z\rangle = 0.$$

Furthermore, by (33), we have for all $z \in H$,

$$\begin{aligned}
|\langle x_{2n+1}-p, z\rangle| &= |\langle x_{2n+1}-p+x_{2n}-x_{2n}, z\rangle| \\
&\leq |\langle x_{2n}-p, z\rangle| + |\langle x_{2n+1}-x_{2n}, z\rangle| \\
&\leq |\langle x_{2n}-p, z\rangle| + \|x_{2n+1}-x_{2n}\|\|z\| \to 0, \quad n\to\infty.
\end{aligned}$$

Therefore, $\{x_{2n+1}\}$ converges weakly to $p$ in $S$. Hence, $\{x_n\}$ converges weakly to a point $p \in S$.

□

The following remarks are in order.

**Remark 4.5.**
(a) Our results in Lemma 4.1, Lemma 4.2, Lemma 4.3 and Theorem 4.4 still hold when $\lambda > 0$ in Algorithm 1 is replaced with $\lambda_n$ such that

$$0 < \inf_{n\geq 1}\lambda_n \leq \sup_{n\geq 1}\lambda_n < \frac{1}{L}.$$

(b) In the convergence analysis in this paper, we do not assume that condition

$$0 \leq \liminf_{n\to\infty}\langle Ax_n, z-x_n\rangle, \forall z \in C$$

assumed by Ceng *et al.* [9] and Yao and Postolache [45] for VI$(A, C)$ (1) with $A$ being pseudo-monotone. Also, there is one computation of projection per iteration in our proposed methods unlike Ceng *et al.* [9] and Yao and Postolache [45], where their methods involved twice computations of projection per iterations. This is an improvement over the results [9, 45].

(c) The assumptions imposed on the iterative parameters given in Assumption 3.2 appear simpler and easier than the assumption imposed on inertial projection method introduced by Dong *et al.* [15]. Also we do not need the inertial factor $\{\alpha_n\}$ to be monotone nondecreasing in our result.

(d) Lemma 4.1 proves that our proposed method produces the Fejér monotonicity of iterates with respect to the solution unlike the other inertial projection methods in [14, 15, 41] and other related papers.

In the light of above Remark 3.4, we give the following result.

**Theorem 4.6.** *Let $\{x_n\}$ be generated by Algorithm 2. Then under Assumptions 3.1 and 3.2 (a)-(b), $\{x_n\}$ converges weakly to a point in $S$.*

*Proof.* Let $x^* \in S$ . Replacing $\lambda$ with $\lambda_n$ in Lemma 4.1, one can easily obtain that $\{x_{2n}\}$ is Fejér monotone with respect to $S$ and $\lim_{n\to\infty} \|x_{2n} - x^*\|$ exists. Furthermore, by Algorithm 2, we get

$$\begin{aligned} \|d_{2n}\| &= \|w_{2n} - y_{2n} - \lambda_{2n}(Aw_{2n} - Ay_{2n})\| \\ &\leq \|w_{2n} - y_{2n}\| + \lambda_{2n}\|Aw_{2n} - Ay_{2n}\| \\ &\leq \Big(1 + \frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\Big)\|w_{2n} - y_{2n}\|. \end{aligned}$$

So,

$$\frac{1}{\|d_{2n}\|} \geq \frac{1}{\Big(1 + \frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\Big)\|w_{2n} - y_{2n}\|}. \tag{41}$$

Also, using similar ideas as in (35), we get

$$\begin{aligned} \langle w_{2n} - y_{2n}, d_{2n}\rangle &= \langle w_{2n} - y_{2n}, w_{2n} - y_{2n} - \lambda_{2n}(Aw_{2n} - Ay_{2n})\rangle \\ &= \|w_{2n} - y_{2n}\|^2 - \langle w_{2n} - y_{2n}, \lambda_{2n}(Aw_{2n} - Ay_{2n})\rangle \\ &\geq \|w_{2n} - y_{2n}\|^2 - \lambda_{2n}\|Aw_{2n} - Ay_{2n}\|\|w_{2n} - y_{2n}\| \\ &\geq \|w_{2n} - y_{2n}\|^2 - \frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\|w_{2n} - y_{2n}\|^2 \\ &= \Big(1 - \frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\Big)\|w_{2n} - y_{2n}\|^2. \end{aligned} \tag{42}$$

By (41) and (42), we get

$$\begin{aligned} \|x_{2n+1} - w_{2n}\| &= \gamma\eta_{2n}\|d_{2n}\| \\ &= \gamma\frac{\langle w_{2n} - y_{2n}, d_{2n}\rangle}{\|d_{2n}\|} \end{aligned}$$

$$\geq \gamma\Big[\frac{\Big(1-\frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\Big)}{\Big(1+\frac{\lambda_{2n}\mu}{\lambda_{2n+1}}\Big)}\Big]\|w_{2n}-y_{2n}\|. \tag{43}$$

By (33), we get from (43) (noting that $\lim_{n\to\infty}\lambda_n=\lambda$ and $w_{2n}=x_{2n}$)

$$\lim_{n\to\infty}\|x_{2n}-y_{2n}\|=0.$$

The remaining part of the proof follows the same arguments as in Lemma 4.3 and Theorem 4.4. □

# 5 Linear Convergence

In a special case when the operator $A$ in VI$(A,C)$ (1) is $\delta$-strongly pseudo-monotone, then Step 3 and Step 4 in Algorithm 1 and Algorithm 2 are not needed to obtain convergence. We propose the following method for the case when $A$ is $\delta$-strongly pseudo-monotone.

**Algorithm 3** Inertial Projection Method for Strongly Pseudo-monotone

1: Define $q:=\frac{1}{\sqrt{1+\lambda(2\delta-\lambda L^2)}}$, where $0<\lambda<\frac{2\delta}{L^2}$. Choose the iterative parameter $\alpha$ such that $0\leq\alpha\leq\frac{1-q}{1+q}$. Let $x_0,x_1\in H$ be given starting points. Set $n:=1$.

2: Compute

$$w_n=\begin{cases} x_n, & n=\text{even}\\ x_n+\alpha(x_n-x_{n-1}), & n=\text{odd}.\end{cases}$$

and

$$x_{n+1}=P_C(w_n-\lambda Aw_n). \tag{44}$$

3: Set $n\leftarrow n+1$, and **go to 2**.

Our focus here is to give theoretical linear rate of convergence of our proposed Algorithm 3. Using Algorithm 3, we have the following result

**Theorem 5.1.** *Suppose $\{x_n\}$ is generated by Algorithm 3. If $A$ is $\delta$-strongly pseudo-monotone on $H$, then $\{x_n\}$ converges at least R-linearly to the unique solution $x^*$ of VI$(A,C)$ (1) and*

$$\|x_n-x^*\|\leq\begin{cases}\frac{\|x_2-x^*\|}{q}q^{\frac{n}{2}}, & n=\text{even}\\ \frac{\|x_2-x^*\|}{q}q^{\frac{n-1}{2}}, & n=\text{odd}\end{cases}$$

*Proof.* By Algorithm 3, we have

$$x_{n+1}=P_C(w_n-\lambda Aw_n),\forall n\geq 1$$

and by (10), we get

$$\langle w_n-\lambda Aw_n-x_{n+1},u-x_{n+1}\rangle\leq 0,\forall u\in C. \tag{45}$$

In particular, from (45), we obtain

$$\langle w_n - \lambda A w_n - x_{n+1}, x^* - x_{n+1}\rangle \leq 0.$$

Thus,

$$2\langle w_n - x_{n+1}, x^* - x_{n+1}\rangle \leq 2\lambda\langle Aw_n, x^* - x_{n+1}\rangle. \tag{46}$$

Since $x^* \in S$, we get $\langle Ax^*, y - x^*\rangle \geq 0, \forall y \in C$. Using the strong pseudo-monotonicity of $A$, we have $\langle Ay, y-x^*\rangle \geq \delta\|y-x^*\|^2, \forall y \in C$. Using Cauchy-schwarz inequality, the Lipschitz continuity of $A$ and the inequality $2ab \leq a^2+b^2, a, b \in H$, we get

$$\begin{aligned}
2\lambda\langle Aw_n, x^* - x_{n+1}\rangle &= -2\lambda\langle Ax_{n+1}, x_{n+1} - x^*\rangle \\
&\quad +2\lambda\langle Aw_n - Ax_{n+1}, x^* - x_{n+1}\rangle \\
&\leq -2\lambda\delta\|x_{n+1} - x^*\|^2 + 2\lambda\|Aw_n - Ax_{n+1}\|\|x_{n+1} - x^*\| \\
&\leq -2\lambda\delta\|x_{n+1} - x^*\|^2 + 2\lambda L\|w_n - x_{n+1}\|\|x_{n+1} - x^*\| \\
&\leq -2\lambda\delta\|x_{n+1} - x^*\|^2 + \|w_n - x_{n+1}\|^2 \\
&\quad +(\lambda L)^2\|x_{n+1} - x^*\|^2. 
\end{aligned} \tag{47}$$

On the other hand, observe that

$$\begin{aligned}
& 2\langle w_n - x_{n+1}, x^* - x_{n+1}\rangle = \|w_n - x_{n+1}\|^2 \\
& +\|x^* - x_{n+1}\|^2 - \|(w_n - x_{n+1}) - (x^* - x_{n+1})\|^2 \\
= & \ \|w_n - x_{n+1}\|^2 + \|x_{n+1} - x^*\|^2 - \|w_n - x^*\|^2.
\end{aligned} \tag{48}$$

Putting (47) and (48) into (46), we get

$$\begin{aligned}
&\|w_n - x_{n+1}\|^2 + \|x_{n+1} - x^*\|^2 - \|w_n - x^*\|^2 \leq \\
&-2\lambda\delta\|x_{n+1} - x^*\|^2 + \|w_n - x_{n+1}\|^2 + (\lambda L)^2\|x_{n+1} - x^*\|^2.
\end{aligned}$$

Therefore, for all $n \geq 1$,

$$\Big[1 + \lambda(2\delta - \lambda L^2)\Big]\|x_{n+1} - x^*\|^2 \leq \|w_n - x^*\|^2. \tag{49}$$

Thus,

$$\begin{aligned}
\|x_{n+1} - x^*\|^2 &\leq \frac{1}{\Big[1 + \lambda(2\delta - \lambda L^2)\Big]}\|w_n - x^*\|^2 \\
&= q^2\|w_n - x^*\|^2, \forall n \geq 1.
\end{aligned} \tag{50}$$

Now,

$$\begin{aligned}
\|w_{2n+1} - x^*\|^2 &= \|x_{2n+1} + \alpha(x_{2n+1} - x_{2n}) - x^*\|^2 \\
&= \|(1+\alpha)(x_{2n+1} - x^*) - \alpha(x_{2n} - x^*)\|^2 \\
&= (1+\alpha)\|x_{2n+1} - x^*\|^2 - \alpha\|x_{2n} - x^*\|^2 \\
&\quad +\alpha(1+\alpha)\|x_{2n+1} - x_{2n}\|^2.
\end{aligned} \tag{51}$$

Putting $n = 2n$ in (50), we obtain

$$\|x_{2n+1} - x^*\|^2 \le q^2\|w_{2n} - x^*\|^2 = q^2\|x_{2n} - x^*\|^2. \tag{52}$$

Again putting $n = 2n+1$ in (50) and using both (51) and (52), we have

$$\begin{aligned}
& \|x_{2n+2} - x^*\|^2 \le q^2\|w_{2n+1} - x^*\|^2 \\
= \;& q^2\Big[(1+\alpha)\|x_{2n+1} - x^*\|^2 - \alpha\|x_{2n} - x^*\|^2 + \alpha(1+\alpha)\|x_{2n+1} - x_{2n}\|^2\Big] \\
\le \;& q^2\Big[q^2(1+\alpha)\|x_{2n} - x^*\|^2 - \alpha\|x_{2n} - x^*\|^2 + \alpha(1+\alpha)\|x_{2n+1} - x_{2n}\|^2\Big] \\
\le \;& q^2\Big[q^2(1+\alpha)\|x_{2n} - x^*\|^2 - \alpha\|x_{2n} - x^*\|^2 + \alpha(1+\alpha)(\|x_{2n+1} - x^*\| + \|x_{2n} - x^*\|)^2\Big] \\
\le \;& q^2\Big[q^2(1+\alpha)\|x_{2n} - x^*\|^2 - \alpha\|x_{2n} - x^*\|^2 + \alpha(1+\alpha)(1+q)^2\|x_{2n} - x^*\|^2\Big] \\
= \;& q^2\Big[q^2(1+\alpha) - \alpha + \alpha(1+\alpha)(1+q)^2\Big]\|x_{2n} - x^*\|^2 \\
\le \;& q^2\|x_{2n} - x^*\|^2.
\end{aligned} \tag{53}$$

By (53), we have

$$\begin{aligned}
\|x_{2n+2} - x^*\| &\le q\|x_{2n} - x^*\| \\
&\le q^2\|x_{2n-2} - x^*\| \\
&\;\vdots \\
&\le q^n\|x_2 - x^*\|, \quad \forall n \ge 1.
\end{aligned} \tag{54}$$

This implies that

$$\|x_{2n} - x^*\| \le \frac{\|x_2 - x^*\|}{q} q^n, \quad \forall n \ge 1. \tag{55}$$

Using (55) in (52), we have

$$\begin{aligned}
\|x_{2n+1} - x^*\| &\le q\|x_{2n} - x^*\| \le \|x_{2n} - x^*\| \\
&\le \frac{\|x_2 - x^*\|}{q} q^n, \quad \forall n \ge 1.
\end{aligned} \tag{56}$$

It follows from (55) and (56) that $\{x_n\}$ converges $R$-linearly to $x^*$ and the desired conclusion is obtained. □

The following result gives priori and posteriori error estimates of the subsequences generated by Algorithm 3.

**Theorem 5.2.** *Suppose $\{x_n\}$ is generated by Algorithm 3. If $A$ is $\delta$-strongly pseudo-monotone, then*
*(i)*

$$\|x_{2n+2} - x^*\| \le \frac{q^n}{1-q}\|x_2 - x_4\|, \quad \forall n \ge 1$$

*and*

$$\|x_{2n+2} - x^*\| \le \frac{q}{1-q}\|x_{2n} - x_{2n+2}\|, \quad \forall n \ge 1;$$

*(ii)*

$$\|x_{2n+1} - x^*\| \leq \frac{q^{n-1}}{1-q}\|x_2 - x_4\|, \quad \forall n \geq 1.$$

*and*

$$\|x_{2n+1} - x^*\| \leq \frac{q}{1-q}\|x_{2n} - x_{2n+2}\|, \quad \forall n \geq 1.$$

*Proof.* Observe that

$$\begin{aligned} \|x_{2n} - x^*\| &\leq \|x_{2n} - x_{2n+2}\| + \|x_{2n+2} - x^*\| \\ &\leq \|x_{2n} - x_{2n+2}\| + q\|x_{2n} - x^*\|, \quad \forall n \geq 1. \end{aligned} \tag{57}$$

Therefore,

$$\|x_{2n} - x^*\| \leq \frac{1}{1-q}\|x_{2n} - x_{2n+2}\|, \quad \forall n \geq 1. \tag{58}$$

From (54) and (58), we get

$$\begin{aligned} \|x_{2n+2} - x^*\| &\leq q^n\|x_2 - x^*\| \\ &\leq \frac{q^n}{1-q}\|x_2 - x_4\|, \quad \forall n \geq 1. \end{aligned} \tag{59}$$

By (54) and (58), we get

$$\begin{aligned} \|x_{2n+2} - x^*\| &\leq q\|x_{2n} - x^*\| \\ &\leq \frac{q}{1-q}\|x_{2n} - x_{2n+2}\|, \quad \forall n \geq 1. \end{aligned} \tag{60}$$

Hence, (i) is established.

By (54), we get

$$q\|x_{2n} - x^*\| \leq q^n\|x_2 - x^*\|, \quad \forall n \geq 1. \tag{61}$$

Thus

$$\|x_{2n} - x^*\| \leq q^{n-1}\|x_2 - x^*\|, \quad \forall n \geq 1. \tag{62}$$

Using (56) and (58), we get

$$\|x_{2n+1} - x^*\| \leq \frac{q^{n-1}}{1-q}\|x_2 - x_4\|, \quad \forall n \geq 1. \tag{63}$$

Also, (52) and (58) imply that

$$\begin{aligned} \|x_{2n+1} - x^*\| &\leq q\|x_{2n} - x^*\| \\ &\leq \frac{q}{1-q}\|x_{2n} - x_{2n+2}\|. \end{aligned} \tag{64}$$

This establishes (ii). □

**Corollary 5.3.** *Suppose $\{x_n\}$ is generated by Algorithm 3. If $A$ is $\delta$-strongly monotone, then $\{x_n\}$ converges at least R-linearly to the unique solution $x^*$ of VI($A, C$)* (1) *and*

$$\|x_n - x^*\| \leq \begin{cases} \frac{\|x_2-x^*\|}{q} q^{\frac{n}{2}}, & n = \text{even} \\ \frac{\|x_2-x^*\|}{q} q^{\frac{n-1}{2}}, & n = \text{odd} \end{cases}$$

*Consequently, the error estimates given in Theorem 5.2 are fulfilled.*

**Remark 5.4.** (a) It has been shown quite recently in [3] that if $w_n = x_n + \alpha(x_n - x_{n-1}), \forall n \geq 1$ in Algorithm 3, then linear convergence cannot be guaranteed. This implies that the convergence rate for projection method with vanilla inertial extrapolation step is worse than the projection method without inertial step when the cost function is strongly-pseudomonotone. In our proposed method in Algorithm 3, we modify the inertial extrapolation step $w_n$ so that linear convergence is obtained.

(b) The results in this section reduce to the results of [24] when $\alpha = 0$ in Algorithm 3.

(c) Algorithm 3 performs well on a strongly monotone variational inequality by Theorem 5.1 since every strongly monotone variational inequality is strongly pseudomonotone varaiational inequality. $\Diamond$

# 6 Numerical Examples

In this section, we provide many computational experiments and compare our proposed methods considered in Section 3 with some existing methods in the literature. All codes were written in MATLAB R2019a and performed on a PC Desktop Intel(R) Core(TM) i7-6600U CPU @ 2.60GHz 2.81 GHz, RAM 16.00 GB.

In all these examples below, we give numerical comparisons of our proposed Algorithm 1 with the methods of Cholamjiak *et.al* (Prasit Alg.) in [14], Dong *et.al* (Dong Alg.) in [15], Shehu *et.al* (Shehu Alg.) in [39] and Thong *et.al* (Thong Alg.) in [41]. In all the numerical implementations, we consider different values of $\gamma \in (0, 2)$.

**Example 6.1.** Define $A : \mathbb{R}^m \to \mathbb{R}^m$ by

$$Ax = \left(e^{-x^TQx} + \beta\right)(Px + q),$$

where $Q$ is a positive definite matrix (i.e, $x^TQx \geq \theta\|x\|^2 \ \forall x \in \mathbb{R}^m$), $P$ is a positive semi-definite matrix, $q \in \mathbb{R}^m$ and $\beta > 0$. Observe that $A$ is differentiable and there exists $M > 0$ such that $\|\nabla Ax\| \leq M, x \in \mathbb{R}^m$. Therefore, by the Mean Value Theorem $A$ is Lipschitz continuous. Also, $A$ is pseudo-monotone but not monotone (see, e.g., [6, Example 2.1]). This is a popular numerical example for variational inequalities with pseudo-monotone cost function. This example shows that the class of pseudo-monotone variational inequalities properly contains the class of monotone variational inequalities and has been considered by many authors (see, e.g., [6]).

Take $C := \{x \in \mathbb{R}^m | Bx \leq b\}$, where $B$ is a matrix of size $l^* \times m$ and $b \in \mathbb{R}^{l^*}_+$ with $l^* = 10$. Let us take $x_0 = (1, 1, ..., 1)^T$ and $x_1$ is generated randomly in $\mathbb{R}^m$. We choose $\alpha_n = \frac{2-\gamma}{1.01\gamma}, \mu = 0.05, \beta = 0.01$ and $\lambda_1 = 0.5$ in Algorithm 2.

In this example, we use the stopping criterion $\|y_n - w_n\| < 10^{-5}$ and the step-size $\{\lambda_n\}$ is generated by (16).

Table 1: Example 6.1 Proposed Algorithm 2 with different values of $\gamma$

| | $m = 100$ | | $m = 200$ | | $m = 300$ | | $m = 400$ | |
|---|---|---|---|---|---|---|---|---|
| $\gamma$ | CPU | Iter. | CPU | Iter. | CPU | Iter. | CPU | Iter. |
| 0.25 | 0.0408 | 6 | 0.0811 | 7 | 0.1260 | 7 | 0.2571 | 7 |
| 0.5 | 0.0419 | 6 | 0.0689 | 6 | 0.1055 | 6 | 0.2643 | 7 |
| 1 | 0.1130 | 6 | 0.2987 | 8 | 0.3234 | 6 | 0.4348 | 6 |
| 1.25 | 0.1075 | 6 | 0.2429 | 7 | 0.3931 | 7 | 0.9591 | 8 |
| 1.5 | 0.0881 | 6 | 0.2094 | 7 | 0.3957 | 7 | 0.4079 | 7 |

Table 2: Example 6.1 Comparison: Proposed Alg. 2, Prasit *et.al Alg., Dong et.al Alg., and Shehu et.al Alg. with* $\gamma = 0.6$

| | $m = 80$ | | $m = 120$ | |
|---|---|---|---|---|
| | CPU ($10^{-2}$) | Iter. | CPU ($10^{-2}$) | Iter. |
| Proposed Alg. 1 | 5.8005 | 6 | 9.2460 | 6 |
| Prasit *et.al Alg.* | 66.8240 | 14 | 103.6800 | 13 |
| Dong *et.al Alg.* | 20.1870 | 8 | 20.8320 | 8 |
| Shehu *et.al Alg.* | 143.4200 | 26 | 258.3300 | 28 |

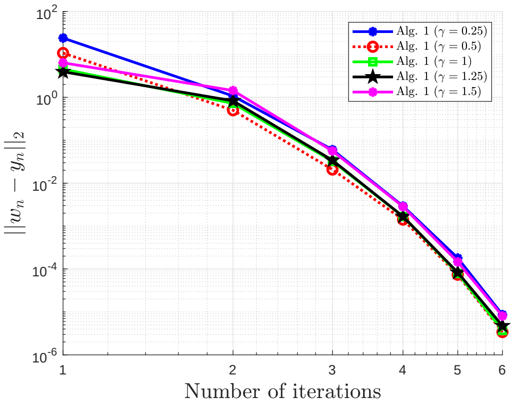


Figure 1: Example 6.1: $\gamma = 0.25$, $m = 100$

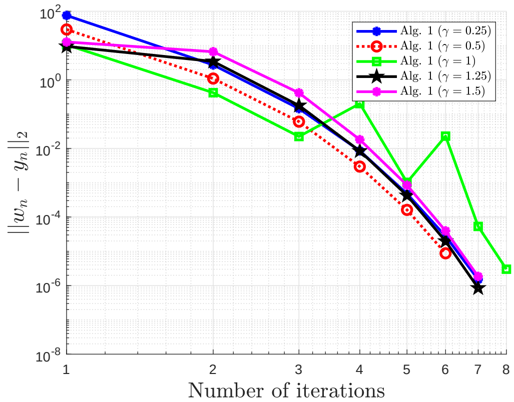


Figure 2: Example 6.1: $\gamma = 0.5$, $m = 200$

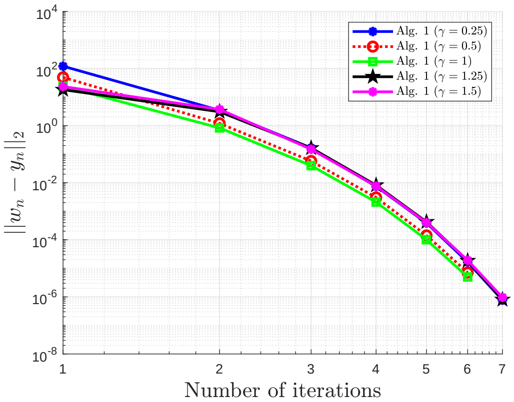


Figure 3: Example 6.1: $\gamma = 1.25$, $m = 300$

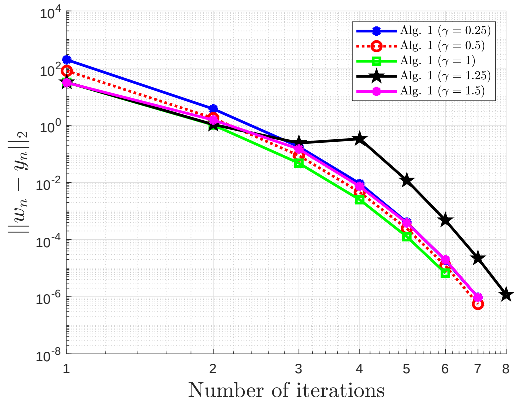


Figure 4: Example 6.1: $\gamma = 1.5$, $m = 400$

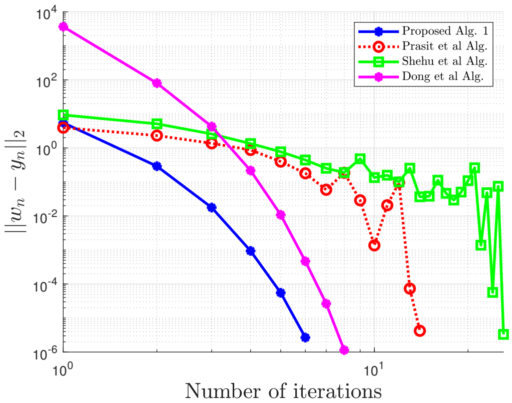


Figure 5: Example 6.1 Comparison: $\gamma = 0.6$, $m = 80$

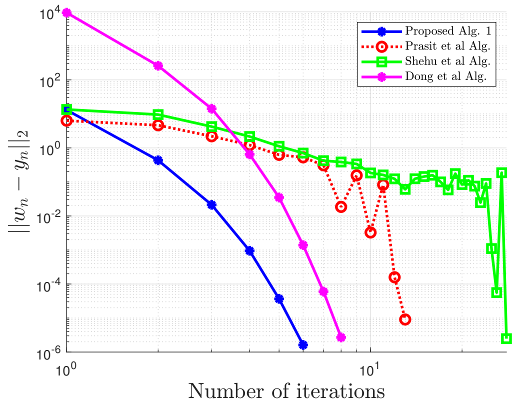


Figure 6: Example 6.1 Comparison: $\gamma = 0.6$, $m = 120$

**Example 6.2.** Define $Ax = Mx + q$, with $M = B^T B + S + D$, where $S, D \in \mathbb{R}^{m\times m}$ are randomly generated matrices such that $S$ is skew-symmetric (hence it does not arise from an optimization problem), $D$ is a positive definite diagonal matrix (hence the variational inequalities has a unique solution) and $q = 0$. Suppose the feasible set $C := \{x \in \mathbb{R}^m | Bx \leq b\}$, for some random matrix $B \in \mathbb{R}^{m\times k}$ and random vector $b \in \mathbb{R}^k$ with non-negative entries. The unique solution of VI$(A, C)$ (1) here is $x^* = \{0\}$. Here, the Lipschitz constant $L = \|M\|$, $\alpha_n = \frac{2-\gamma}{1.01\gamma}$ and $\lambda_n = \lambda = \frac{1}{1.05L}$ in Algorithm 1. We generate $x_0, x_1$ randomly in $\mathbb{R}^m$. We use the stopping criterion $\|x_n - x^*\| < 10^{-3}$.

Table 3: Example 6.2 Proposed Algorithm 1 with different values of $\gamma$

| | $m = 30$ | | $m = 50$ | |
|---|---|---|---|---|
| $\gamma$ | CPU | Iter. | CPU | Iter. |
| 0.25 | 10.8625 | 1026 | 28.9645 | 2250 |
| 0.5 | 9.7527 | 963 | 25.0423 | 1944 |
| 1 | 9.7189 | 858 | 19.2209 | 1550 |
| 1.25 | 7.7424 | 670 | 15.0703 | 1242 |
| 1.5 | 7.7202 | 681 | 13.9887 | 1158 |

Table 4: Example 6.2 Comparison: Alg. 1, Prasit *et.al Alg.*, Shehu *et.al Alg. and Thong et.al Alg. for* $k = 20$

| | $m = 30$ | | $m = 50$ | | $m = 70$ | |
|---|---|---|---|---|---|---|
| | CPU | Iter. | CPU | Iter. | CPU | Iter. |
| Proposed Alg. 1 | 5.1326 | 549 | 23.1849 | 1175 | 26.2597 | 1690 |
| Prasit *et.al Alg.* | 6.4600 | 726 | 38.6367 | 1412 | 29.7394 | 1988 |
| Shehu *et.al Alg.* | 203.2958 | 29487 | 1404.8401 | 81893 | 2519.2316 | 172916 |
| Thong *et.al Alg.* | 66.6438 | 7696 | 153.6271 | 10960 | 216.6621 | 13858 |

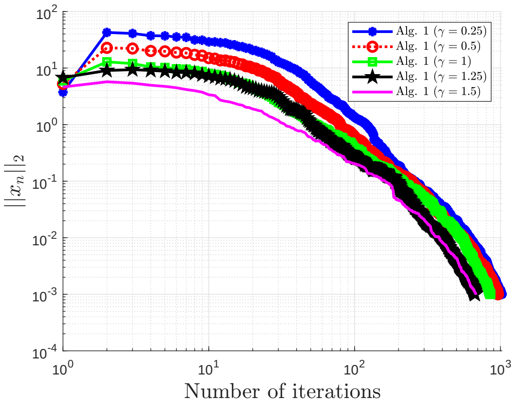


Figure 7: Example 6.2 with different $\gamma$: $m = 30$

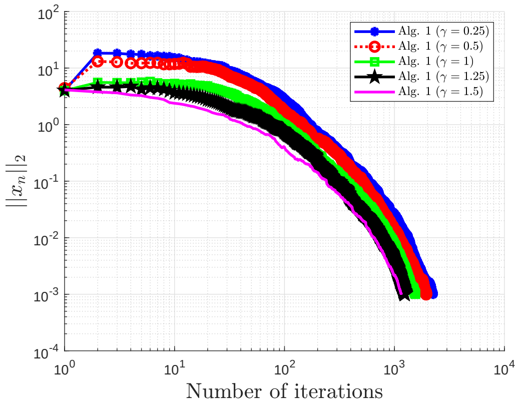


Figure 8: Example 6.2 with different $\gamma$: $m = 50$

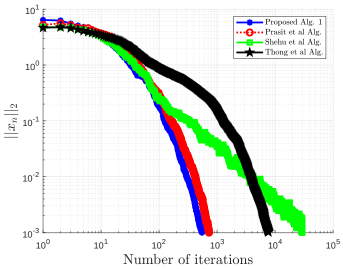


Figure 9: Example 6.2 Comparison: $m = 30$

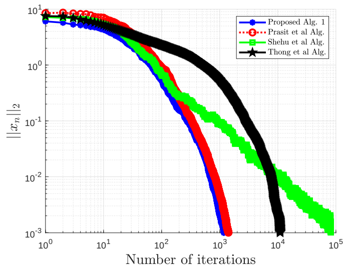


Figure 10: Example 6.2 Comparison: $m = 50$

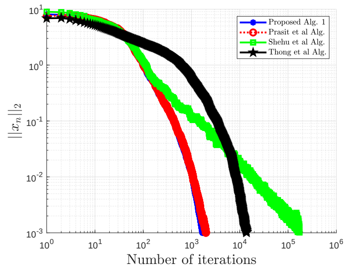


Figure 11: Example 6.2 Comparison: $m = 70$

**Example 6.3.** Define $A \in \mathbb{R}^{2\times 2}$ by

$$Ax = \begin{bmatrix} 0.5x_1x_2 - 2x_2 - 10^7 \\ -4x_1 + 0.1x_2^2 - 10^7 \end{bmatrix}$$

and $C = \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 - 2)^2 + (x_2 - 2)^2 \leq 1\}$ Then $A$ is pseudomonotone but not monotone with Lipschitz constant $L = 5$ The unique solution of VI$(A, C)$ (1) is $x^* = (2.707, 2.707)^T$. We take $\alpha_n = \frac{2-\gamma}{4\gamma}$ and $x_0, x_1$ are randomly generated in Algorithm 1.

We use the stopping criterion $\|x_n - x^*\| < 10^{-3}$ for proposed Algorithm 1 with the different choices of step-size $\lambda \in (0, \frac{1}{5})$ and $\gamma$.

Table 5: Example 6.3 Proposed Algorithm 1 with different values of $\lambda$

| | $\lambda = 0.01$ | | $\lambda = 0.1$ | | $\lambda = 0.15$ | |
|---|---|---|---|---|---|---|
| $\gamma$ | CPU ($10^{-5}$) | Iter. | CPU ($10^{-5}$) | Iter. | CPU ($10^{-5}$) | Iter. |
| 0.25 | 27.36 | 14 | 13.92 | 12 | 28.38 | 14 |
| 0.5 | 16.04 | 8 | 10.16 | 8 | 13.60 | 10 |
| 1 | 7.60 | 3 | 9.82 | 6 | 9.52 | 6 |
| 1.25 | 17.23 | 9 | 17.08 | 9 | 18.43 | 9 |
| 1.5 | 15.08 | 15 | 26.08 | 15 | 19.72 | 15 |

Table 6: Example 6.3 Comparison: Alg. 1, Prasit *et.al Alg.*, *Dong et.al Alg.*, *Shehu et.al Alg. and Thong et.al Alg. for* $\gamma = 0.1$

| | $\lambda = 0.01$ | | $\lambda = 0.1$ | | $\lambda = 0.19$ | |
|---|---|---|---|---|---|---|
| | CPU ($10^{-4}$) | Iter. | CPU ($10^{-4}$) | Iter. | CPU ($10^{-4}$) | Iter. |
| Proposed Alg. 1 | 2.836 | 22 | 3.225 | 23 | 3.499 | 24 |
| Dong *et.al Alg.* | 4.488 | 50 | 6.069 | 62 | 6.186 | 66 |
| Prasit *et.al Alg.* | 91.046 | 879 | 94.349 | 879 | 139.660 | 1273 |
| Shehu *et.al Alg.* | 324.940 | 3378 | 209.64 | 2211 | 340.220 | 3665 |
| Thong *et.al Alg.* | 43.380 | 639 | 92.075 | 1356 | 2399.400 | 13131 |

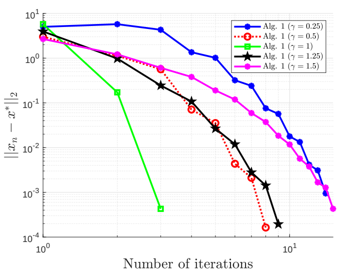


Figure 12: Example 6.3 with different $\gamma$: $\lambda = 0.01$

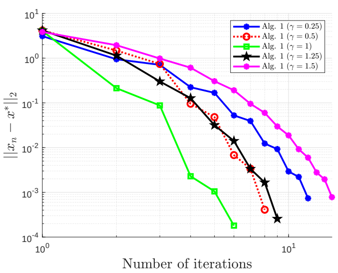


Figure 13: Example 6.3 with different $\gamma$: $\lambda = 0.1$

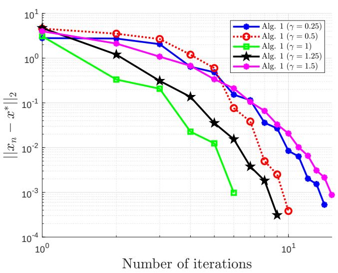


Figure 14: Example 6.3 with different $\gamma$: $\lambda = 0.15$

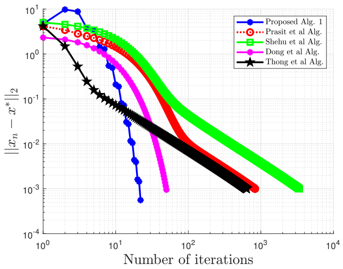


Figure 15: Example 6.3 Comparison: $\lambda = 0.01$

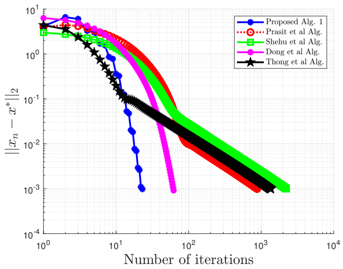


Figure 16: Example 6.3 Comparison: $\lambda = 0.1$

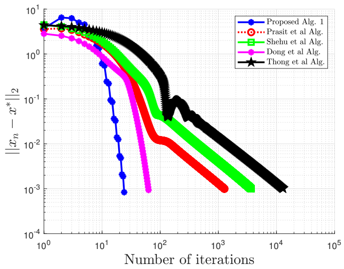


Figure 17: Example 6.3 Comparison: $\lambda = 0.19$

We check if the choices of $\gamma$ affect the efficiency of our methods or whether any choice of $\gamma$ suffices in our methods. This is why we give Table 1, Table 3 and Table 5. We also judge the sensitivity to $\gamma$ by showing the sensitivity to the other hyperparameter $\lambda$ in Table 5 and Table 6. In Example 6.1, the optimum choice of $\gamma$ is $\gamma = 0.6$ using Algorithm 2. In Example 6.2, the optimum choice of $\gamma$ is $\gamma = 1.5$ using Algorithm 1 with $\lambda$ chosen close to the Lipschitz constant of the cost function. The optimum choice of $\gamma$ is $\gamma = 1$ in Example 6.3 for $\lambda \in (0, \frac{1}{L})$ using Algorithm 1. The performance of Thong et al. is worse in Example 6.1. So, we decided not to add the numerical results for Thong et al. in Example 6.1.

# 7 Final Remarks

A projection method with alternated inertial extrapolation step is introduced and studied in this paper for solving variational inequality problems. The Fejér monotonicity, which is lost in many other projection methods with inertial extrapolation step is regained to some extent (as it appears to be lost in Figures 7, 8, 12, 15, 16, and 17, but retained in the others) and convergence analysis of the proposed method is given under some simpler conditions than other inertial projection methods for solving variational inequality available in the literature. Another contribution in this paper is that the inertial factor can be chosen bigger than 1 and this might make our proposed method converges faster than other inertial projection methods as shown in our numerical examples. Modifications of our result in solving Nash equilibrium problems would be studied in the future.

# Acknowledgments

The authors are grateful to the two anonymous referees for their insightful comments and suggestions which have improved the earlier version of the manuscript greatly. The first author has received funding from the European Research Council (ERC) under the European Union's Seventh Framework Program (FP7 - 2007-2013) (Grant agreement No. 616160).

# References


[1] R.Y. Apostol, A.A. Grynenko, and V.V. Semenov; Iterative algorithms for monotone bilevel variational inequalities, *J. Comp. Appl. Math.* **107** (2012), 3-14.

[2] J.-P. Aubin and I. Ekeland; *Applied Nonlinear Analysis*, Wiley, New York, 1984.

[3] J.-F. Aujol, C. Dossal, and A. Rondepierre; Optimal Convergence Rates for Nesterov Acceleration, *SIAM J. Optim.* doi:10.1137/18M1186757.

[4] C. Baiocchi and A. Capelo; *Variational and Quasivariational Inequalities; Applications to Free Boundary Problems*, Wiley, New York (1984).

[5] H.H. Bauschke and P.L. Combettes; *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, CMS Books in Mathematics, Springer, New York (2011).

[6] R.I. Boţ, E. R. Csetnek and P.T. Vuong; The forward-backward-forward method from discrete and continuous perspective for pseudomonotone variational inequalities in Hilbert spaces, 2018, Available from: https://arxiv.org/abs/1808.08084.

[7] X. Cai, G. Gu and B. He; On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators, *Comput. Optim. Appl.* **57** (2014), 339-363.

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

[9] L.C. Ceng, M. Teboulle, and J.C. Yao; Weak convergence of an iterative method for pseudo-monotone variational inequalities and fixed-point problems, *J. Optim. Theory Appl.* **146** (2010), 19-31.

[10] L. C. Ceng and J. C. Yao; An extragradient-like approximation method for variational inequality problems and fixed point problems, *Appl. Math. Comput.* **190** (2007), 205-215.

[11] Y. Censor, A. Gibali, and S. Reich; Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space, *Optim. Methods Softw.* **26** (2011), 827-845.

[12] Y. Censor, A. Gibali, S. Reich; The subgradient extragradient method for solving variational inequalities in Hilbert space, *J. Optim. Theory Appl.*, **148** (2011), 318-335.

[13] Y. Censor, A. Gibali, and S. Reich; The subgradient extragradient method for solving variational inequalities in Hilbert space, *J. Optim. Theory Appl.* **148** (2011), 318-335.

[14] P. Cholamjiak, D. V. Thong, and Y. J. Cho; A Novel Inertial Projection and Contraction Method for Solving Pseudomonotone Variational Inequality Problems, *Acta Appl. Math.* (2019). https://doi.org/10.1007/s10440-019-00297-7.

[15] Q. L. Dong, Y. J. Cho, L. L. Zhong and Th. M. Rassias; Inertial projection and contraction algorithms for variational inequalities, *J. Glob. Optim.* **70** (2018) 687-704.

[16] F. Facchinei and J.-S. Pang; *Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume II.* Springer Series in Operations Research, Springer, New York, 2003.

[17] A. Gibali; Two simple relaxed perturbed extragradient methods for solving variational inequalities in Euclidean spaces, *J. Nonlinear Var. Anal.* **2** (2018), 49-61.

[18] R. Glowinski, J.-L. Lions and R. Trémolières; *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam (1981).

[19] B. S. He; A class of projection and contraction methods for monotone variational inequalities, *Appl. Math. Optim.* **35** (1997), 69-76.

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

[21] A.N. Iusem and B.F. Svaiter, A variant of Korpelevich's method for variational inequalities with a new search strategy, *Optimization* **42** (1997), 309-321.

[22] F. Iutzeler and J. M. Hendrickx; A generic online acceleration scheme for optimization algorithms via relaxation and inertia, *Optim. Methods Softw.* **34** (2019), no. 2, 383–405.

[23] F. Iutzeler and J. Malick; On the proximal gradient algorithm with alternated inertia, *J. Optim. Theory Appl.* **176** (2018), no. 3, 688–710.

[24] P. D. Khanh and P. T. Vuong; Modified projection method for strongly pseudomonotone variational inequalities, *J. Global Optim.* **58** (2014), no. 2, 341-350.

[25] E.N. Khobotov; Modification of the extragradient method for solving variational inequalities and certain optimization problems, *USSR Comput. Math. Math. Phys.* **27** (1989), 120-127.

[26] D. Kinderlehrer and G. Stampacchia; *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York (1980).

[27] I.V. Konnov; *Combined Relaxation Methods for Variational Inequalities*, Springer-Verlag, Berlin (2001).

[28] G.M. Korpelevich; The extragradient method for finding saddle points and other problems, *Ekonomika i Mat. Metody* **12** (1976), 747-756.

[29] R. Kraikaew and S. Saejung; Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces, *J. Optim. Theory Appl.* **163** (2014), 399-412.

[30] P.-E. Maingé and M.L. Gobinddass; Convergence of one-step projected gradient methods for variational inequalities, *J. Optim. Theory Appl.* **171** (2016), 146-168.

[31] Yu.V. Malitsky; Projected reflected gradient methods for monotone variational inequalities, *SIAM J. Optim.* **25** (2015), 502-520.

[32] P. Marcotte; Applications of Khobotov's algorithm to variational and network equlibrium problems, *Inf. Syst. Oper. Res.* **29** (1991), 258-270.

[33] J. Mashreghi and M. Nasri; Forcing strong convergence of Korpelevich's method in Banach spaces with its applications in game theory, *Nonlinear Analysis* **72** (2010), 2086-2099.

[34] Z. Mu and Y. Peng; A note on the inertial proximal point method, *Stat. Optim. Inf. Comput.* **3** (2015), no. 3, 241–248.

[35] N. Nadezhkina and W. Takahashi; Weak Convergence Theorem by an Extragradient Method for Nonexpansive Mappings and Monotone Mappings, *J. Optim. Theory Appl.* **128**, (2006), 191-201.

[36] A. Nagurney; *Network Economics: A Variational Inequality Approach*, Kluwer Academic Publishers, Dordrecht (1999).

[37] Z. Opial; Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Am. Math. Soc.* **73** (1967), 591-597.

[38] M.V. Solodov and B.F. Svaiter; A new projection method for variational inequality problems, *SIAM J. Control Optim.* **37** (1999), 765-776.

[39] Y. Shehu, Q.-L. Dong and D. Jiang; Single projection method for pseudo-monotone Variational Inequality in Hilbert Spaces, *Optimization* **68** (2019), 385-409.

[40] E. Süli and D. F. Mayers; An introduction to numerical analysis, Cambridge University Press, Cambridge, 2003. x+433 pp.

[41] D.V. Thong, N.T. Vinh and Y.J. Cho; New strong convergence theorem of the inertial projection and contraction method for variational inequality problems, *Numer. Algorithms* (2019). https://doi.org/10.1007/ s11075-019-00755-1.

[42] P. Tseng; A modified forward-backward splitting method for maximal monotone mappings, *SIAM J. Control Optim.* **38** (2000), 431-446.

[43] P. T. Vuong; On the weak convergence of the extragradient method for solving pseudomonotone variational inequalities, *J. Optim. Theory Appl.* **176** (2018), 399-409.

[44] Y. Yao, G. Marino and L. Muglia; A modified Korpelevich's method convergent to the minimum-norm solution of a variational inequality, *Optimization* **63** (2014), 559-569.

[45] Y. Yao and M. Postolache; Iterative methods for pseudo-monotone variational inequalities and fixed-point problems, *J. Optim. Theory Appl.* **155** (2012), 273-287.