R-linear convergence analysis of inertial extragradient algorithms for strongly pseudo-monotone variational inequalities

Duong Viet Thong * Phan Tu Vuong†

Abstract

Some extragradient-type algorithms with inertial effect for solving strongly pseudo-monotone variational inequalities have been proposed and investigated recently. While the convergence of these algorithms was established, it is unclear if the linear rate is guaranteed. In this paper, we provide R-linear convergence analysis for two extragradient-type algorithms for solving strongly pseudo-monotone, Lipschitz continuous variational inequality in Hilbert spaces. The linear convergence rate of is obtained without the prior knowledge of the Lipschitz constants of the variational inequality mapping and the stepsize is bounded from below by a positive number. Some numerical results are provided to show the computational effectiveness of the algorithms.

Keywords: Inertial subgradient extragradient method; forward-backward-forward method; strongly pseudo-monotone mapping; Lipschitz continuity; R-linear rate

AMS class: 65Y05; 65K15; 68W10; 47H05; 47J25

1 Introduction

Let C be a nonempty closed and convex subset of a Hilbert space H and $F : H \to H$ be a continuous operator. The variational inequality (VI) as follows.

$$\mathrm{VI(C,F)}: \quad \text{Find } x^* \in C \ \text{ s.t. } \langle Fx^*, y - x^*\rangle \geq 0 \quad \forall y \in C.$$

Variational inequality theory is an important tool in economics, engineering mechanics, mathematical programming, transportation and others (see, for example, [1, 2, 3, 4, 5]). Numerous numerical algorithms for solving VI(C,F) have been proposed for solving variational inequalities and related optimization problems, see [6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17]. In general, each solution method is designed for certain class of VIs so that the convergence of the algorithms can be guaranteed. The extragradient type algorithms are well designed for solving pseudo-monotone VIs. The classical extragardient method for solving monotone VIs in finite dimenstional spaces was proposed by Korpelevich [18], Antipin [19] and improved by Popov in [20]. Extension to pseudo-monotone VIs in infinite dimensional Hilbert space has been studied recently in [14]. The extragradient method with inertial effect was studied in [21], where the inertial constant depends heavily on the stepsize via a very complicated formula. This method was improved in [32], where the authors proposed a subgradient extragradient method with constant stepsize for solving monotone VIs. Note that all the extragradient type algorithms appeared before [14] can only be applied

*Faculty of Mathematical Economics, National Economics University, Hanoi City, Viet Nam. Email: thongduongviet@gmail.com.

†Mathematical Sciences, University of Southampton, SO17 1BJ, Southampton, UK. Email: t.v.phan@soton.ac.uk.

for solving monotone VIs in infinite dimensional Hilbert spaces, but not pseudo-monotone VIs.

We focus on the class of strongly pseudo-monotone VIs, which has been attracting a lot of attentions in recent years, see e.g., [22, 23, 24, 25, 26, 27, 28]. The existence and uniqueness as well as stability of this problem were studied in [27]. It was proved in [27] that if F is strongly pseudo-monotone and Lipschitz continuous, then VI(C,F) has a unique solution. [25] proved that the gradient projection method converges linearly to the unique solution provided that the step-size is sufficiently small, depending on the strong pseudo-monotonicity and Lipschitz continuity constants of the considered operator. Some modifications of the gradient projection method were recently considered in [23], where diminishing stepsizes were required but the Lipschitz continuity was relaxed to continuity. Variants of the extragradient method, where two projections per iteration are needed, were studied in [26]. When the knowledge of Lipschitz and strong pseudo-monotonicity constants is not available, one has to consider diminishing stepsizes [23, 24, 25, 28], which slows down the convergence speed. To improve the convergence speed, we need to avoid diminishing stepsize and/or combine the algorithms with inertial effect. Inertial algorithms for variational inequality and optimization problems has recently received considerable attention, see, e.g, [21, 29, 30, 31] and the references therein.

In this paper, we consider two extragradient-type algorithms: the inertial subgradient algorithm proposed in [32] and the inertial forward-backward-forward algorithm studied in [33]. In contrast to the inertial extragradient algorithm proposed in [28], these algorithms require only one projection onto the feasible set per iteration. Under strong pseudo-monotonicity and Lipschitz continuity assumptions, we prove that the iterative sequence generated by the aforementioned algorithms converges linearly to the unique solution of VI(C,F). These results are new and have not been considered in [32, 33]. Moreover, in comparison with [23, 25] the stepsizes considered in these algorithms are not diminishing but instead bounded from below by a positive constant, which avoid the slow convergence speed.

The rest of the paper is organized as follows: In Sect. 2, we recall some definitions and preliminary results for the convergence analysis. Sect. 3 deals with the linear convergence analysis of the proposed algorithms. Finally, in Sect. 4, we present some numerical experiments to illustrate the behaviors of the proposed algorithms.

2 Preliminaries

Let H be a real Hilbert space and let C be a nonempty closed convex subset of H. For every point $x \in H$, there exists a unique nearest point in C, denoted by $P_C x$ such that

$$\|x - P_C x\| \leq \|x - y\| \ \forall y \in C.$$

P_C is the metric projection of H onto C. It is known that P_C is nonexpansive. Moreover the following property hold:

Fact 2.1. ([34]) *Let C be a nonempty closed convex subset of a real Hilbert space H and P_C be the metric projection onto C. Given $x \in H$ and $z \in C$. Then $z = P_C x \iff \langle x - z, z - y \rangle \geq 0 \ \ \forall y \in C$.*

Fact 2.2. ([34]) *Let C be a closed and convex subset in a real Hilbert space H, $x \in H$. Then*
i) $\|P_C x - P_C y\|^2 \leq \langle P_C x - P_C y, x - y \rangle \ \forall y \in H$;
ii) $\|P_C x - y\|^2 \leq \|x - y\|^2 - \|x - P_C x\|^2 \ \forall y \in C$.

For properties of the metric projection, the interested readers are referred to [34, Section 3].

Definition 2.1. *[35] A sequence $\{x_n\}$ in H is said to converge R-linearly to x^* with rate $\rho \in [0,1)$ if there is a constant $c > 0$ such that*

$$\|x_n - x^*\| \leq c\rho^n \quad \forall n \in \mathbb{N}.$$

Definition 2.2. *Let $F : H \to H$ be a mapping. Then F is called*

1. *L-Lipschitz continuous if there exists a constant $L > 0$ such that*

$$\|Fx - Fy\| \leq L\|x - y\| \quad \forall x, y \in H.$$

2. *κ-strongly monotone if there exists a constant $\kappa > 0$ such that*

$$\langle Fx - Fy, x - y\rangle \geq \kappa\|x - y\|^2 \quad \forall x, y \in H.$$

3. *κ-strongly pseudo-monotone if there exists a constant $\kappa > 0$ such that*

$$\langle Fx, y - x\rangle \geq 0 \Longrightarrow \langle Fy, y - x\rangle \geq \kappa\|x - y\|^2 \quad \forall x, y \in H.$$

Remark 2.1. *Obviously, if the mapping F is strongly monotone, then it is strongly pseudo-monotone with the same modulus. In addition, if the mapping F is strongly pseudo-monotone and Lipschitz continuous then the problem* $\mathrm{VI(C,F)}$ *has a unique solution [27].*

3 R-linear Convergence Analysis

In this section, we consider two extragradient-type algorithms: the inertial subgradient extragradient algorithm [28, 32] and the inertial forward-backward-forward algorithm [33]. We prove that the iterative sequence generated by the these algorithms converge to the unique solution of VIs with an R-linear rate.

3.1 The inertial subgradient extragradient algorithm

We consider first the inertial subgradient extragradient algorithm studied in [28, 32]. Even though it does not require the prior knowledge of the Lipschitz constant associated with the variational inequality mapping, the stepsize is still bounded from below by a positive constant. This is in contrast with the gradient projection algorithms considered in [23, 24, 25].

Algorithm 1. (**iSEG**)

Initialization: *Given $\alpha \geq 0, \tau_1 > 0, \mu \in (0,1)$. Let $x_0, x_1 \in H$ be arbitrary*

Step 1. *Given the iterates x_{n-1} and x_n $(n \geq 1)$. Set $w_n = x_n + \alpha(x_n - x_{n-1})$ and compute*

$$y_n = P_C(w_n - \tau_n F w_n).$$

If $y_n = w_n$ or $Fy_n = 0$ then stop: y_n is a solution to the problem $\mathrm{VI(C,F)}$.
Otherwise, go to **Step 2**.
Step 2. *Compute*

$$x_{n+1} = P_{T_n}(w_n - \tau_n F y_n),$$

where

$$T_n := \{x \in H | \langle w_n - \tau_n F w_n - y_n, x - y_n\rangle \leq 0\}.$$

Update

$$\tau_{n+1} = \begin{cases} \min\left\{ \dfrac{\mu\|w_n - y_n\|}{\|Fw_n - Fy_n\|}, \tau_n \right\} & \text{if } Fw_n - Fy_n \neq 0, \\ \tau_n & \text{otherwise.} \end{cases} \tag{3.1}$$

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

The following lemmas are quite helpful to analyze the convergence of algorithm.

Lemma 3.1. ([17]) *The sequence $\{\tau_n\}$ generated by (3.1) is a nonincreasing sequence and*

$$\lim_{n\to\infty} \tau_n = \tau \geq \min\{\tau_1, \frac{\mu}{L}\}.$$

Lemma 3.2. *Assume that $F : H \to H$ is L-Lipschitz continuous on H and δ-strongly pseudo-monotone on C. Let p be the unique solution of $VI(C, F)$ and $\{x_n\}$ be generated by Algorithm 1. Then the following inequality holds:*

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - (1 - \mu\frac{\tau_n}{\tau_{n+1}})\|y_n - w_n\|^2 - (1 - \mu\frac{\tau_n}{\tau_{n+1}})\|x_{n+1} - y_n\|^2 - 2\tau_n\delta\|y_n - p\|^2.$$

Proof: Since $p \in Sol(C, F) \subset C \subset T_n$, we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|P_{T_n}(w_n - \tau_n F y_n) - P_{T_n} p\|^2 \\ &\leq \langle x_{n+1} - p, w_n - \tau_n F y_n - p\rangle \\ &= \frac{1}{2}\|x_{n+1} - p\|^2 + \frac{1}{2}\|w_n - \tau_n F y_n - p\|^2 - \frac{1}{2}\|x_{n+1} - w_n + \tau_n F y_n\|^2 \\ &= \frac{1}{2}\|x_{n+1} - p\|^2 + \frac{1}{2}\|w_n - p\|^2 + \frac{1}{2}\tau_n^2\|F y_n\|^2 - \langle w_n - p, \tau_n F y_n\rangle \\ &\quad - \frac{1}{2}\|x_{n+1} - w_n\|^2 - \frac{1}{2}\tau_n^2\|F y_n\|^2 - \langle x_{n+1} - w_n, \tau_n F y_n\rangle \\ &= \frac{1}{2}\|x_{n+1} - p\|^2 + \frac{1}{2}\|w_n - p\|^2 - \frac{1}{2}\|x_{n+1} - w_n\|^2 - \langle x_{n+1} - p, \tau_n F y_n\rangle. \end{aligned}$$

This implies that

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - \|x_{n+1} - w_n\|^2 - 2\langle x_{n+1} - p, \tau_n F y_n\rangle. \tag{3.2}$$

Since p is the solution of the problem (VI), we have $\langle Fp, x - p\rangle \geq 0$ for all $x \in C$. By the strong pseudo-montonicity of F on C, we have $\langle Fx, x - p\rangle \geq \delta\|x - p\|^2$ for all $x \in C$. Taking $x := y_n \in C$, we get

$$\langle F y_n, p - y_n\rangle \leq -\delta\|y_n - p\|^2.$$

Thus we have

$$\begin{aligned} \langle F y_n, p - x_{n+1}\rangle &= \langle F y_n, p - y_n\rangle + \langle F y_n, y_n - x_{n+1}\rangle \\ &\leq -\delta\|y_n - p\|^2 + \langle F y_n, y_n - x_{n+1}\rangle. \end{aligned} \tag{3.3}$$

From (3.2) and (3.3), we obtain

$$\begin{aligned}
&\|x_{n+1}-p\|^2\\
&\le \|w_n-p\|^2-\|x_{n+1}-w_n\|^2+2\tau_n\langle Fy_n, y_n-x_{n+1}\rangle-2\tau_n\delta\|y_n-p\|^2\\
&=\|w_n-p\|^2-\|x_{n+1}-y_n\|^2-\|y_n-w_n\|^2-2\langle x_{n+1}-y_n, y_n-w_n\rangle\\
&\quad+2\tau_n\langle Fy_n, y_n-x_{n+1}\rangle-2\tau_n\delta\|y_n-p\|^2\\
&=\|w_n-p\|^2-\|x_{n+1}-y_n\|^2-\|y_n-w_n\|^2\\
&\quad+2\langle w_n-\tau_nFy_n-y_n, x_{n+1}-y_n\rangle-2\tau_n\delta\|y_n-p\|^2. \qquad (3.4)
\end{aligned}$$

Since $y_n = P_{T_n}(w_n - \lambda_n Fw_n)$ and $x_{n+1} \in T_n$ we have

$$\begin{aligned}
2\langle w_n-\lambda_nFy_n-y_n, &x_{n+1}-y_n\rangle\\
&=2\langle w_n-\lambda_nFw_n-y_n, x_{n+1}-y_n\rangle+2\lambda_n\langle Fw_n-Fy_n, x_{n+1}-y_n\rangle\\
&\le 2\lambda_n\langle Fw_n-Fy_n, x_{n+1}-y_n\rangle. \qquad (3.5)
\end{aligned}$$

We estimate $2\tau_n\langle Fw_n - Fy_n, x_{n+1} - y_n\rangle$ as follows

$$\begin{aligned}
2\tau_n\langle Fw_n-Fy_n, x_{n+1}-y_n\rangle &\le 2\tau_n\|Fy_n-Fw_n\|\|y_n-x_{n+1}\|\\
&\le 2\mu\frac{\tau_n}{\tau_{n+1}}\|w_n-y_n\|\|y_n-x_{n+1}\|\\
&\le \mu\frac{\tau_n}{\tau_{n+1}}\|w_n-y_n\|^2+\mu\frac{\tau_n}{\tau_{n+1}}\|y_n-x_{n+1}\|^2. \qquad (3.6)
\end{aligned}$$

Combining (3.5) and (3.6), we obtain

$$2\langle w_n-\lambda_nFy_n-y_n, x_{n+1}-y_n\rangle \le \mu\frac{\tau_n}{\tau_{n+1}}\|w_n-y_n\|^2+\mu\frac{\tau_n}{\tau_{n+1}}\|y_n-x_{n+1}\|^2. \qquad (3.7)$$

Substituting (3.7) into (3.4) it holds

$$\|x_{n+1}-p\|^2 \le \|w_n-p\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})\|y_n-w_n\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})\|x_{n+1}-y_n\|^2-2\tau_n\delta\|y_n-p\|^2.$$

Lemma 3.3. *Assume that $F : H \to H$ is L-Lipschitz continuous on H and δ-strongly pseudo-monotone on C. Let p be the unique solution of $VI(C, F)$ and $\{x_n\}$ be generated by Algorithm 1. Then there exist $N \in \mathbb{N}$ and $\rho, \xi \in (0, 1)$ such that*

$$\|x_{n+1}-p\|^2 \le\rho\|w_n-p\|^2-\xi\|x_{n+1}-w_n\|^2 \quad \forall n \ge N.$$

Proof: Indeed, thanks to Lemma 3.2 we get for any $\theta \in (0, 1)$ that

$$\begin{aligned}
\|x_{n+1}-p\|^2 &\le \|w_n-p\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})\|y_n-w_n\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})(1-\theta)\|x_{n+1}-y_n\|^2-2\tau_n\delta\|y_n-p\|^2\\
&=\|w_n-p\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})\theta\|y_n-w_n\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})(1-\theta)[\|y_n-w_n\|^2\\
&\quad+\|x_{n+1}-y_n\|^2]-2\tau_n\delta\|y_n-p\|^2\\
&\le \|w_n-p\|^2-(1-\mu\frac{\tau_n}{\tau_{n+1}})\theta\|y_n-w_n\|^2-\frac{1}{2}(1-\mu\frac{\tau_n}{\tau_{n+1}})(1-\theta)\|x_{n+1}-w_n\|^2-2\tau_n\delta\|y_n-p\|^2.
\end{aligned}$$

Let $\beta := \min\left\{\frac{(1-\mu)\theta}{4}, \frac{\tau\delta}{2}\right\}$, where $\tau := \lim_{n\to\infty}\tau_n$, we obtain

$$\lim_{n\to\infty}\frac{1}{2}(1-\mu\frac{\tau_n}{\tau_{n+1}})(1-\theta) = \frac{1}{2}(1-\mu)(1-\theta) > \frac{1}{2}(1-\mu)(1-\theta)\theta;$$
$$\lim_{n\to\infty}(1-\mu\frac{\tau_n}{\tau_{n+1}})\theta = (1-\mu)\theta \geq 4\beta;$$
$$\lim_{n\to\infty}\tau_n\delta = \tau\delta \geq 2\beta.$$

Thus, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ it holds

$$\frac{1}{2}(1-\mu\frac{\tau_n}{\tau_{n+1}})(1-\theta) \geq \frac{1}{2}(1-\mu)(1-\theta)\theta;$$
$$(1-\mu\frac{\tau_n}{\tau_{n+1}})\theta \geq 2\beta;$$
$$\tau_n\delta \geq \beta.$$

Hence we have for all $n \geq N$ that

$$\begin{aligned}
\|x_{n+1}-p\|^2 \leq& \|w_n-p\|^2 - 2\beta\|y_n-w_n\|^2 - \frac{1}{2}(1-\mu)(1-\theta)\theta\|x_{n+1}-w_n\|^2 - 2\beta\|y_n-p\|^2 \\
=& \|w_n-p\|^2 - \frac{1}{2}(1-\mu)(1-\theta)\theta\|x_{n+1}-w_n\|^2 - 2\beta(\|y_n-w_n\|^2 + \|y_n-p\|^2) \\
\leq& \|w_n-p\|^2 - \frac{1}{2}(1-\mu)(1-\theta)\theta\|x_{n+1}-w_n\|^2 - \beta\|w_n-p\|^2 \\
=& (1-\beta)\|w_n-p\|^2 - \frac{1}{2}(1-\mu)(1-\theta)\theta\|x_{n+1}-w_n\|^2 \\
=& \rho\|w_n-p\|^2 - \xi\|x_{n+1}-w_n\|^2,
\end{aligned}$$

where $\rho := 1-\beta \in (0,1), \xi := \frac{1}{2}(1-\mu)(1-\theta)\theta \in (0,1)$.

Now, we prove that the iterative sequence generated by Algorithm 1 converges R-linearly to the unique solution of the problem VI(C,F).

Theorem 3.1. *Assume that $F : H \to H$ is L-Lipschitz continuous on H and δ-strongly pseudo-monotone on C. Let $\theta, \gamma \in (0,1)$ and α be such that*

$$0 \leq \alpha \leq \min\left\{1 + \frac{1-\sqrt{1+4\xi}}{2\xi}, \frac{\sqrt{(1+\gamma\xi)^2+4\gamma\xi}-(1+\gamma\xi)}{2}, (1-\gamma)\left(1-\frac{(1-\mu)\theta}{2}\right)\right\} \quad (3.8)$$

where $\xi := \frac{1}{2}(1-\mu)(1-\theta)\theta$. Then the sequence $\{x_n\}$ is generated by Algorithm 1 converges in norm to the unique solution p of the problem VI(C,F) *with an R-linear rate not worse than*

$$\rho = 1 - \min\left\{\frac{(1-\mu)\theta}{2}, \tau\delta\right\}. \quad (3.9)$$

Proof: Thanks to Lemma 3.3, we get

$$\|x_{n+1}-p\|^2 \leq \rho\|w_n-p\|^2 - \xi\|x_{n+1}-w_n\|^2 \quad \forall n \geq N. \quad (3.10)$$

On the other hand, we also have

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

and

$$\begin{aligned}\|x_{n+1} - w_n\|^2 &= \|x_{n+1} - x_n - \alpha(x_n - x_{n-1})\|^2 \\ &= \|x_{n+1} - x_n\|^2 + \alpha^2\|x_n - x_{n-1}\|^2 - 2\alpha\left\langle x_{n+1} - x_n, x_n - x_{n-1}\right\rangle \\ &\geq \|x_{n+1} - x_n\|^2 + \alpha^2\|x_n - x_{n-1}\|^2 - 2\alpha\|x_{n+1} - x_n\|\|x_n - x_{n-1}\| \\ &\geq \|x_{n+1} - x_n\|^2 + \alpha^2\|x_n - x_{n-1}\|^2 - \alpha\|x_{n+1} - x_n\|^2 - \alpha\|x_n - x_{n-1}\|^2 \\ &\geq (1-\alpha)\|x_{n+1} - x_n\|^2 - \alpha(1-\alpha)\|x_n - x_{n-1}\|^2.\end{aligned}$$

Combining these inequalities with (3.10) we obtain

$$\begin{aligned}\|x_{n+1} - p\|^2 \leq{}& \rho(1+\alpha)\|x_n - p\|^2 - \rho\alpha\|x_{n-1} - p\|^2 + \rho\alpha(1+\alpha)\|x_n - x_{n-1}\|^2 \\ &- \xi(1-\alpha)\|x_{n+1} - x_n\|^2 + \xi\alpha(1-\alpha)\|x_n - x_{n-1}\|^2 \quad \forall n \geq N,\end{aligned}$$

or equivalently

$$\begin{aligned}\|x_{n+1} - p\|^2 - \rho\alpha\|x_n - p\|^2 + &\xi(1-\alpha)\|x_{n+1} - x_n\|^2 \\ &\leq \rho\left[\|x_n - p\|^2 - \alpha\|x_{n-1} - p\|^2 + \xi(1-\alpha)\|x_n - x_{n-1}\|^2\right] \\ &\quad - \left(\rho\xi(1-\alpha) - \rho\alpha(1+\alpha) - \xi\alpha(1-\alpha)\right)\|x_n - x_{n-1}\|^2 \quad \forall n \geq N.\end{aligned}$$

Setting

$$\Gamma_n := \|x_n - p\|^2 - \alpha\|x_{n-1} - p\|^2 + \xi(1-\alpha)\|x_n - x_{n-1}\|^2,$$

since $\rho \in (0,1)$ we can write

$$\begin{aligned}\Gamma_{n+1} \leq{}& \|x_{n+1} - p\|^2 - \rho\alpha\|x_n - p\|^2 + \xi(1-\alpha)\|x_{n+1} - x_n\|^2 \\ \leq{}& \rho\Gamma_n - \left(\rho\xi(1-\alpha) - \rho\alpha(1+\alpha) - \xi\alpha(1-\alpha)\right)\|x_n - x_{n-1}\|^2 \quad \forall n \geq N.\end{aligned}$$

Note that from (3.8) and Lemma 3.3 we have

$$\begin{aligned}\alpha \leq{}& (1-\gamma)\left(1 - \frac{(1-\mu)\theta}{2}\right) \\ \leq{}& (1-\gamma)(1-\beta) = (1-\delta)\rho,\end{aligned}$$

which implies

$$\xi\alpha(1-\alpha) \leq (1-\gamma)\rho\xi(1-\alpha) = \rho\xi(1-\alpha) - \gamma\rho\xi(1-\alpha). \tag{3.11}$$

Since

$$\alpha \leq \frac{\sqrt{(1+\gamma\xi)^2 + 4\gamma\xi} - (1+\gamma\xi)}{2}$$

it holds

$$\alpha^2 + (1+\gamma\xi)\alpha - \gamma\xi \leq 0,$$

or equivalently

$$\alpha(1+\alpha) \leq \gamma\xi(1-\alpha).$$

Hence

$$\rho\alpha(1+\alpha) \leq \rho\gamma\xi(1-\alpha). \tag{3.12}$$

From (3.11) and (3.12) we deduce

$$\rho\xi(1-\alpha) - \rho\alpha(1+\alpha) - \xi\alpha(1-\alpha) \geq 0.$$

which implies that

$$\Gamma_{n+1} \leq \rho\Gamma_n.$$

Now, we show that $\Gamma_n \geq 0$ for all n. Indeed, we have

$$\Gamma_n = \|x_n - p\|^2 - \alpha\|x_{n-1} - p\|^2 + \xi(1-\alpha)\|x_n - x_{n-1}\|^2 \tag{3.13}$$

On the other hand, we have

$$\begin{aligned}
\|x_{n-1} - p\|^2 = \|x_{n-1} - x_n + x_n - p\|^2 &= \|x_{n-1} - x_n\|^2 + \|x_n - p\|^2 + 2\langle x_{n-1} - x_n, x_n - p\rangle \\
&\leq \|x_{n-1} - x_n\|^2 + \|x_n - p\|^2 + 2\|x_{n-1} - x_n\|\|x_n - p\| \\
&\leq \|x_{n-1} - x_n\|^2 + \|x_n - p\|^2 + k\|x_{n-1} - x_n\|^2 + \frac{1}{k}\|x_n - p\|^2 \\
&= (1+k)\|x_{n-1} - x_n\|^2 + \left(1 + \frac{1}{k}\right)\|x_n - p\|^2,
\end{aligned} \tag{3.14}$$

for all $k > 0$. Combining (3.13) and (3.14), we get

$$\Gamma_n \geq \left[1 - \left(1 + \frac{1}{k}\right)\alpha\right]\|x_n - p\|^2 + \Big[\xi(1-\alpha) - (1+k)\alpha\Big]\|x_n - x_{n-1}\|^2. \tag{3.15}$$

We show that there exists $k > 0$ such that

$$\begin{cases} 1 - \left(1 + \frac{1}{k}\right)\alpha & > 0, \\ \xi(1-\alpha) - (1+k)\alpha & > 0. \end{cases} \tag{3.16}$$

Indeed, if $\alpha = 0$ then the inequality (3.16) is obvious. Now, we consider $\alpha > 0$. In this case, the inequality (3.16) is equivalently to

$$\begin{cases} k > \dfrac{\alpha}{1-\alpha}, \\ k < \dfrac{\xi(1-\alpha)}{\alpha} - 1. \end{cases}$$

Moreover, from (3.8) we also have

$$0 \leq \alpha \leq 1 + \frac{1 - \sqrt{1 + 4\xi}}{2\xi}$$

which follows that

$$\frac{\alpha}{1-\alpha} < \frac{\xi(1-\alpha)}{\alpha} - 1,$$

that is, there exists $k > 0$ satisfying the inequality (3.16). From (3.15), it implies that $\Gamma_n \geq 0$ for all n. Hence

$$\Gamma_{n+1} \leq \rho\Gamma_n \leq ... \leq \rho^{n-N+1}\Gamma_N.$$

$$\|x_n - p\|^2 \leq \frac{\Gamma_N}{\rho^N(1 - \xi(1-\alpha))}\rho^n,$$

which implies that $\{x_n\}$ converges R-linearly to p, the unique solution of VI(C,F).

Remark 3.1. *(i) We note that the linear convergence holds for any arbitrary $\theta, \gamma \in (0,1)$. In general, estimating the maximum value of inertial parameter α satisfying (3.8) is cumbersome. We explain roughly how to choose the parameters θ, γ such that α is somehow maximal. Let*

$$a := 1 + \frac{1-\sqrt{1+4\xi}}{2\xi}, \quad b := \frac{\sqrt{(1+\gamma\xi)^2 + 4\gamma\xi} - (1+\gamma\xi)}{2}, \quad c =: (1-\gamma)\left(1 - \frac{(1-\mu)\theta}{2}\right),$$

where $\xi := \frac{1}{2}(1-\mu)(1-\theta)\theta$. Since $\theta \in (0,1)$ we have $\xi \in \left(0, \frac{1-\mu}{8}\right]$ with $\xi_{\max} = \frac{1-\mu}{8}$ when $\theta = 0.5$. It is clear that $a = a(\xi)$ is an increasing function on $\left(0, \frac{1-\mu}{8}\right]$. Hence $a_{\max}$ is attained at $\xi_{\max} = \frac{1-\mu}{8}$, i.e., when $\theta = 0.5$. Similarly, for any fixed γ, the function $b = b(\xi)$ is increasing on $\left(0, \frac{1-\mu}{8}\right]$. This implies again that $b_{\max}$ is attained at $\theta = 0.5$. Now let us fix $\theta = 0.5$, then $c = (1-\gamma)\left(1 - \frac{1-\mu}{4}\right)$. For a given $\mu \in (0,1)$, $a = a(\gamma)$ is a constant, $b = b(\gamma)$ is increasing while $c = c(\gamma)$ is decreasing on $(0,1)$. (see Figure 1 for an illustration when $\mu = 0.5$, in this case $\alpha_{\max} = 0.052$ when $\gamma = 0.94$.). Hence $\alpha_{\max}$ is attained when γ is one of the solution of

$$\{a(\gamma) = b(\gamma), a(\gamma) = c(\gamma), b(\gamma) = c(\gamma)\},$$

which are eventually quadratic equations with respect to γ.

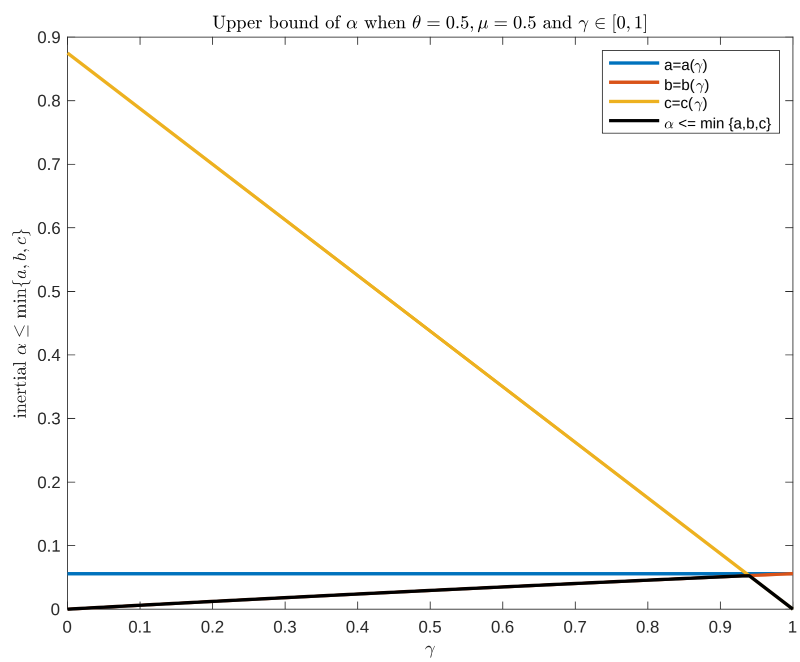

Figure 1: An illustration of upper bound for the inertial parameter.

.

(ii) The value of ρ estimated in (3.9) only provides an upper bound of the actual linear rate. Theoretically, it is unclear how the inertial parameter α effects the linear rate. In practice, it is observed that the bigger inertia implies better rate. From (i) we know that the optimal choice for θ is 0.5. Then

$$\rho = 1 - \min\left\{\frac{1-\mu}{4}, \tau\delta\right\} = \max\left\{\frac{3+\mu}{4}, 1-\tau\delta\right\}.$$

This estimation, however, is not optimal since $\lim_{\mu\to 1}\rho = 1$.

3.2 The inertial forward-backward-forward algorithm

The inertial forward-backward-forward algorithm has been recently studied in [33, Remark 3]. In this section, we present the convergence rate analysis of this algorithm.

Algorithm 2. (**iFBF**)

Initialization: *Given* $\alpha \geq 0, \tau_1 > 0, \mu \in (0,1)$. *Let* $x_0, x_1 \in H$ *be arbitrary*

Step 1. *Given the iterates* x_{n-1} *and* x_n $(n \geq 1)$. *Set* $w_n = x_n + \alpha(x_n - x_{n-1})$ *and compute*

$$y_n = P_C(w_n - \tau_n F w_n).$$

If $y_n = w_n$ *or* $Fy_n = 0$ *then stop:* y_n *is a solution to the problem* $\mathrm{VI(C,F)}$.
Otherwise, go to **Step 2**.
Step 2. *Compute*

$$x_{n+1} = y_n - \tau_n(Fy_n - Fw_n).$$

Update

$$\tau_{n+1} = \begin{cases} \min\left\{\dfrac{\mu\|w_n - y_n\|}{\|Fw_n - Fy_n\|}, \tau_n\right\} & \text{if } Fw_n - Fy_n \neq 0, \\ \tau_n & \text{otherwise.} \end{cases}$$

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

Lemma 3.4. *Assume that* $F : H \to H$ *is L-Lipschitz continuous on* H *and* δ*-strongly pseudo-monotone on* C. *Let* p *be the unique solution of* $VI(C,F)$ *and* $\{x_n\}$ *be generated by Algorithm 2. Then the following inequality holds:*

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - (1 - \mu^2\frac{\tau_n^2}{\tau_{n+1}^2})\|y_n - w_n\|^2 - 2\tau_n\delta\|y_n - p\|^2.$$

Proof: Indeed, we have

$$\begin{aligned} \|x_{n+1} - p\|^2 &= \|y_n - \tau_n(Fy_n - Fw_n) - p\|^2 \\ &= \|y_n - p\|^2 + \tau_n^2\|Fy_n - Fw_n\|^2 - 2 - \tau_n\langle y_n - p, Fy_n - Fw_n\rangle \\ &= \|w_n - p\|^2 + \|w_n - y_n\|^2 + 2\langle y_n - w_n, w_n - p\rangle + \tau_n^2\|Fy_n - Fw_n\|^2 \\ &\quad - 2 - \tau_n\langle y_n - p, Fy_n - Fw_n\rangle \\ &= \|w_n - p\|^2 + \|w_n - y_n\|^2 - 2\langle y_n - w_n, y_n - w_n\rangle + 2\langle y_n - w_n, y_n - p\rangle \\ &\quad + \tau_n^2\|Fy_n - Fw_n\|^2 - 2\tau_n\langle y_n - p, Fy_n - Fw_n\rangle \\ &= \|w_n - p\|^2 - \|w_n - y_n\|^2 + 2\langle y_n - w_n, y_n - p\rangle + \tau_n^2\|Fy_n - Fw_n\|^2 \\ &\quad - 2\tau_n\langle y_n - p, Fy_n - Fw_n\rangle. \end{aligned} \tag{3.17}$$

Since $y_n = P_C(w_n - \tau_n Fw_n)$, it holds

$$\langle y_n - w_n + \tau_n Fw_n, y_n - p\rangle \leq 0$$

or equivalently,

$$\langle y_n - w_n, y_n - p\rangle \leq -\tau_n\langle Fw_n, y_n - p\rangle. \tag{3.18}$$

From (3.17) and (3.18), it follows that

$$\begin{aligned}\|x_{n+1} - p\|^2 \leq &\|w_n - p\|^2 - \|w_n - y_n\|^2 - 2\tau_n\langle Fw_n, y_n - p\rangle + \tau_n^2\|Fy_n - Fw_n\|^2 \\ &- 2\tau_n\langle y_n - p, Fy_n - Fw_n\rangle \\ =&\|w_n - p\|^2 - \|w_n - y_n\|^2 + \tau_n^2\|Fy_n - Fw_n\|^2 - 2\tau_n\langle y_n - p, Fy_n\rangle.\end{aligned} \tag{3.19}$$

Since p is the solution of VI(C,F), we have $\langle Fp, x-p\rangle \geq 0$ for all $x \in C$. By the strong pseudomontonicity of F on C we have $\langle Fx, x-p\rangle \geq \delta\|x-p\|^2$ for all $x \in C$.
Taking $x := y_n \in C$ we get

$$\langle Fy_n, p - y_n\rangle \leq -\delta\|y_n - p\|^2. \tag{3.20}$$

From (3.19) and (3.20) we obtain

$$\begin{aligned}\|x_{n+1} - p\|^2 \leq &\|w_n - p\|^2 - \|w_n - y_n\|^2 + \tau_n^2\|Fy_n - Fw_n\|^2 - 2\tau_n\langle y_n - p, Fy_n\rangle \\ \leq&\|w_n - p\|^2 - \|w_n - y_n\|^2 + \tau_n^2\|Fy_n - Fw_n\|^2 - 2\tau_n\delta\|y_n - p\|^2.\end{aligned} \tag{3.21}$$

Moreover, it is easy to see that by the definition of $\{\tau_n\}$ we have

$$\|Fw_n - Fy_n\| \leq \frac{\mu}{\tau_{n+1}}\|w_n - y_n\| \quad \forall n. \tag{3.22}$$

Combining (3.21) and (3.22), we obtain

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - (1 - \mu^2\frac{\tau_n^2}{\tau_{n+1}^2})\|y_n - w_n\|^2 - 2\tau_n\delta\|y_n - p\|^2.$$

Lemma 3.5. *Assume that $F : H \to H$ is L-Lipschitz continuous on H and δ-strongly pseudomonotone on C. Let p be the unique solution of $VI(C,F)$ and $\{x_n\}$ be generated by Algorithm 2. Then there exists $N \in \mathbb{N}$ and $\rho, \xi \in (0,1)$ such that*

$$\|x_{n+1} - p\|^2 \leq\rho\|w_n - p\|^2 - \xi\|x_{n+1} - w_n\|^2 \quad \forall n \geq N.$$

Proof: Indeed, thanks to Lemma 3.4, we have

$$\|x_{n+1} - p\|^2 \leq \|w_n - p\|^2 - (1 - \mu^2\frac{\tau_n^2}{\tau_{n+1}^2})\|y_n - w_n\|^2 - 2\tau_n\delta\|y_n - p\|^2.$$

Hence for any $\theta \in (0,1)$ we can deduce

$$\|x_{n+1}-p\|^2 \leq \|w_n-p\|^2-(1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2})(1-\theta)\|y_n-w_n\|^2-(1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2})\theta\|y_n-w_n\|^2-2\tau_n\delta\|y_n-p\|^2. \tag{3.23}$$

By the definition of x_{n+1} we have

$$\begin{aligned}\|x_{n+1} - y_n\| &= \|y_n - \tau_n(Fy_n - Fw_n) - y_n\| \\ &\leq \tau_n\|Fy_n - Fw_n\| \\ &\leq \mu\frac{\tau_n}{\tau_{n+1}}\|y_n - w_n\|.\end{aligned}$$

Therefore

$$\|x_{n+1}-w_n\| \le \|x_{n+1}-y_n\| + \|y_n-w_n\| \le (1+\mu\frac{\tau_n}{\tau_{n+1}})\|y_n-w_n\|.$$

This implies

$$\|y_n-w_n\| \ge \frac{1}{(1+\mu\frac{\tau_n}{\tau_{n+1}})}\|x_{n+1}-w_n\|. \tag{3.24}$$

From $\lim_{n\to\infty}\left(1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2}\right)=1-\mu^2>0$, thus, there exists $N_0\in\mathbb{N}$ such that

$$1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2} \ge \frac{1-\mu^2}{2} > 0 \quad \forall n \ge N_0.$$

Substituting (3.24) into (3.23) we have for all $n \ge N_0$ that

$$\begin{aligned}
\|x_{n+1}-p\|^2 \le & \|w_n-p\|^2 - \frac{\left(1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2}\right)}{\left(1+\mu\frac{\tau_n}{\tau_{n+1}}\right)^2}(1-\theta)\|x_{n+1}-w_n\|^2 - (1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2})\theta\|y_n-w_n\|^2 - 2\tau_n\delta\|y_n-p\|^2 \\
= & \|w_n-p\|^2 - \frac{\left(1-\mu\frac{\tau_n}{\tau_{n+1}}\right)}{\left(1+\mu\frac{\tau_n}{\tau_{n+1}}\right)}(1-\theta)\|x_{n+1}-w_n\|^2 - \left(1-\mu^2\frac{\tau_n^2}{\tau_{n+1}^2}\right)\theta\|y_n-w_n\|^2 - 2\tau_n\delta\|y_n-p\|^2
\end{aligned} \tag{3.25}$$

Let $\beta := \min\left\{\frac{(1-\mu^2)\theta}{4}, \frac{\tau\delta}{2}\right\}$, where $\tau := \lim_{n\to\infty}\tau_n$, we have

$$\lim_{n\to\infty}\frac{\left(1-\mu\frac{\tau_n}{\tau_{n+1}}\right)}{\left(1+\mu\frac{\tau_n}{\tau_{n+1}}\right)}(1-\theta) = \frac{1-\mu}{1+\mu}(1-\theta) > \frac{1-\mu}{1+\mu}(1-\theta)\theta;$$

$$\lim_{n\to\infty}\left(1-\mu^2\frac{\tau_n}{\tau_{n+1}}\right)\theta = (1-\mu^2)\theta \ge 4\beta;$$

$$\lim_{n\to\infty}\tau_n\delta = \tau\delta \ge 2\beta.$$

Thus, there exists N_1 such that for all $n \ge N_1$ we have

$$\frac{\left(1-\mu\frac{\tau_n}{\tau_{n+1}}\right)}{\left(1+\mu\frac{\tau_n}{\tau_{n+1}}\right)}(1-\theta) \ge \frac{1-\mu}{1+\mu}(1-\theta)\theta; \quad \left(1-\mu^2\frac{\tau_n}{\tau_{n+1}}\right)\theta \ge 2\beta; \quad \tau_n\delta \ge \beta.$$

Let $N=\max\{N_0,N_1\}$, using (3.25) we get for all $n\geq N$ that

$$
\begin{aligned}
\|x_{n+1}-p\|^2 \leq& \|w_n-p\|^2-\frac{1-\mu}{1+\mu}(1-\theta)\theta\|x_{n+1}-w_n\|^2-2\beta(\|y_n-w_n\|^2+\|y_n-p\|^2)\\
\leq& \|w_n-p\|^2-\frac{1-\mu}{1+\mu}(1-\theta)\theta\|x_{n+1}-w_n\|^2-\beta\|w_n-p\|^2\\
\leq& (1-\beta)\|w_n-p\|^2-\frac{1-\mu}{1+\mu}(1-\theta)\theta\|x_{n+1}-w_n\|^2\\
=& \rho\|w_n-p\|^2-\xi\|x_{n+1}-w_n\|^2,
\end{aligned}
$$

where $\rho:=1-\beta\in(0,1)$ and $\xi:=\frac{1-\mu}{1+\mu}(1-\theta)\theta\in(0,1)$.

Theorem 3.2. *Assume that $F:H\to H$ is L-Lipschitz continuous on H and δ-strongly pseudomonotone on C. Let $\theta,\gamma\in(0,1)$ and α be such that*

$$0\leq\alpha\leq\min\left\{1+\frac{1-\sqrt{1+4\xi}}{2\xi},\frac{\sqrt{(1+\gamma\xi)^2+4\gamma\xi}-(1+\gamma\xi)}{2},(1-\gamma)\left(1-\frac{(1-\mu)\theta}{2}\right)\right\}$$

where $\xi:=\frac{1-\mu}{1+\mu}(1-\theta)\theta$. Then the sequence $\{x_n\}$ is generated by Algorithm 2 converges in norm to the unique solution p of the problem (VI) *with an R-linear rate not worse than*

$$\rho=1-\min\left\{\frac{(1-\mu^2)\theta}{2},\tau\delta\right\}.$$

Proof: The proof is similar to that of Theorem 3.1, and therefore is omitted.

Remark 3.2. *Similar to Theorem 3.1, we see that the maximal value for the inertial α is attained when $\theta=0.5$. In addition, the upper bound of the linear rate obtained in* (**??**) *is slightly better than* (3.9)*.*

4 Numerical Examples

In this section, we provide some numerical examples to illustrate the linear convergence of the iterative sequence generated by Algorithm 1 and Algorithm 2. We focus on the class of strongly pseudomonotone but not (strongly) monotone VI. For numerical experiments with (strongly) monotone VI, we refer the readers to [28].

Example 4.1. *Let $H=\ell_2$, the real Hilbert space, whose elements are the square-summable sequences of real numbers, i.e., $H=\{x=(x_1,x_2,\ldots,x_i,\ldots):\sum_{i=1}^{\infty}|x_i|^2<+\infty\}$. The inner product and the norm on H are given by setting*

$$\langle x,y\rangle=\sum_{i=1}^{\infty}x_iy_i \quad \text{and} \quad \|x\|=\sqrt{\langle x,x\rangle}$$

for any $x=(x_1,x_2,\ldots,x_i,\ldots),y=(y_1,y_2,\ldots,y_i,\ldots)\in H$.
Let $a,b\in\mathbb{R}$ be such that $b>a>\frac{b}{2}>0$. Put

$$C:=\{x\in H:\|x\|\leq a\},\quad Fx:=(b-\|x\|)\,x.$$

Clearly, the $VI(C,F)$ has unique solution $x^*=0$. It was proved in [25] that F is strongly pseudomonotone and Lipschitz continuous but not monotone.

It follows from Section 3 that the iterations generated by the proposed algorithms converge linearly to the unique solution $x^* = 0$. In the following, we compare the performance of these algorithms with different inertial parameters. We choose $H = \mathbb{R}^{1000}, a = 5, b = 8$ and random starting points x_0, x_1. The stopping condition is $\|x_n\| \leq 10^{-10}$. It is clear that while iSEG and iFBF are comparable, their inertial effect versions are faster than the non-inertial ones.

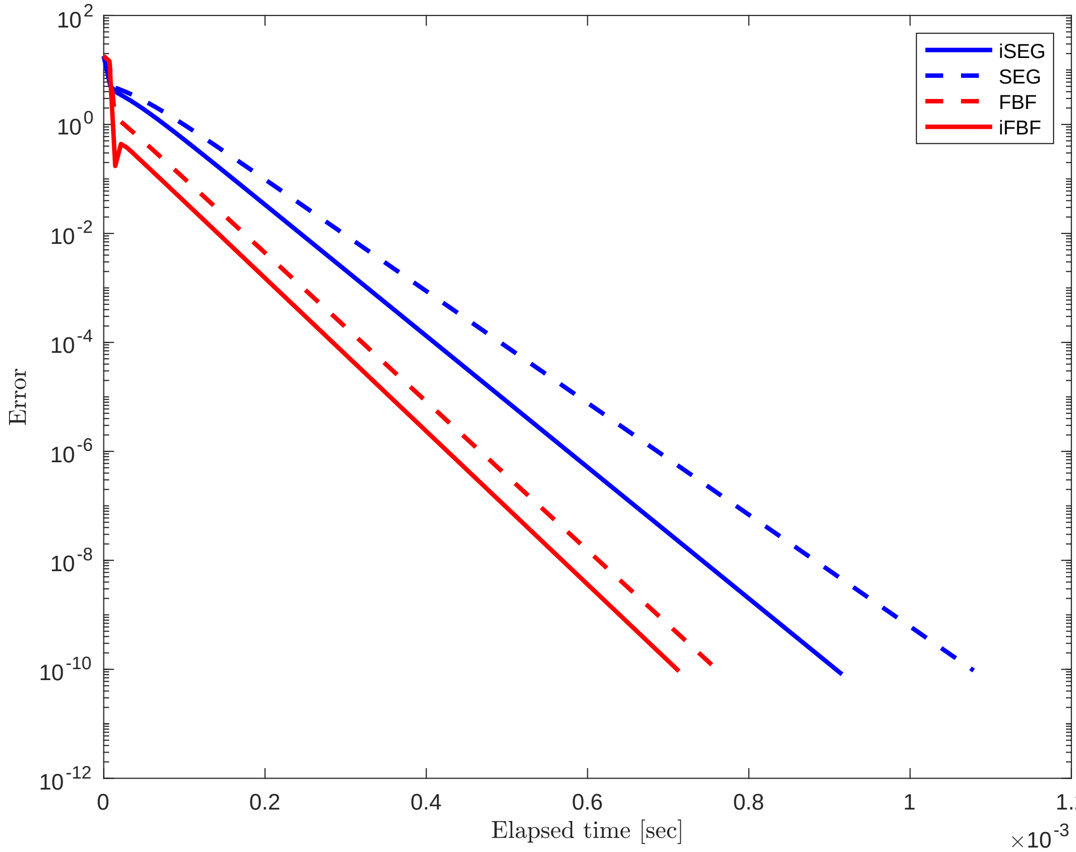

Figure 2: Comparison of Algorithm 1 (iSEG) and Algorithm 2 (iFBF) for Example 4.1 (SEG and FBF with $\alpha = 0$, iSEG and iFBF with $\alpha = 0.05$)

.

Example 4.2. *Let* $C = \left\{x \in [-5,5]^3 : x_1 + x_2 + x_3 = 0\right\} \subseteq \mathbb{R}^3$ *and* $F : \mathbb{R}^3 \to \mathbb{R}^3$ *be defined as*

$$Fx = \left(e^{-\|x\|^2} + q\right) Mx,$$

where $q = 0.2$ *and*

$$M = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1.5 & 0 \\ -1 & 0 & 2 \end{pmatrix}.$$

As proved in [7], F *is* γ*-strongly pseudo-monotone on* $\mathbb{R}^3$ *with constant* $\gamma := q \cdot \lambda_{min} \approx 0.0764$*, where* λ_{min} *is the smallest eigenvalue of* M*, and Lipschitz continuous with constant* $L \approx 5.0679$*.*

Moreover, F is not monotone since for $x = (-1,0,0)^T, y = (-2,0,0)^T \in \mathbb{R}^3$ we have

$$\langle F(x) - F(y), x - y\rangle = -0.1312 < 0.$$

We compare the performance of Algorithm 1 and Algorithm 2 with different inertial parameters. We choose $\mu = 0.5$, $\tau_1 = 0.1$ and random starting points x_0, x_1 for all tested algorithms. The stopping condition is $\|x_n - x^*\| \leq 10^{-10}$, where x^* is the unique solution obtained by running Algorithm 1 with $\alpha = 0$ for 1000 iterations. It is clear that the inertial effect speeds up the convergence rate.

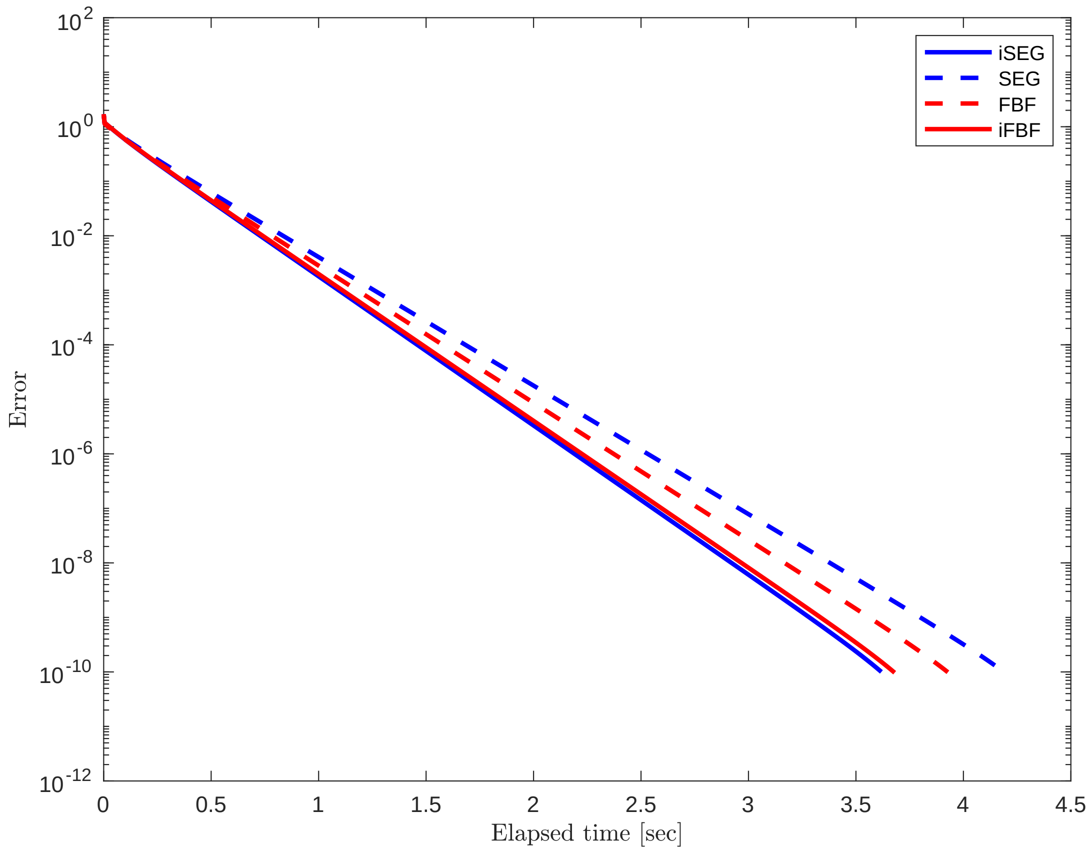

Figure 3: Comparison of Algorithm 1 (iSEG) and Algorithm 2 (iFBF) for Example 4.2 (SEG and FBF with $\alpha = 0$, iSEG and iFBF with $\alpha = 0.05$)

.

Acknowledgements

The authors are very thankful to both anonymous referees and the Principal Editor, Prof. Andre A. Keller, for their careful reading and constructive comments, which helped improving the presentation of the paper. This work was supported by the Vietnam National Foundation for Science and Technology Development (NAFOSTED) project 101.01-2019.320. A significant part of the paper was completed during a scientific stay of the authors at the Vietnam Institute for Advanced Study in Mathematics (VIASM), whose support and hospitality are gratefully appreciated.

References

[1] J. P. Aubin, I. Ekeland, Applied Nonlinear Analysis, Wiley, New York, 1984.

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

[3] F. Facchinei, J. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems, Volume I. Springer Series in Operations Research, Springer, New York, 2003.

[4] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and Their Applications, Academic Press, New York, 1980.

[5] I. V. Konnov, Combined Relaxation Methods for Variational Inequalities, Springer-Verlag, Berlin, 2001.

[6] M. Abbas, H. Iqbal, Two Inertial extragradient viscosity algorithms for solving variational inequality and fixed point problems, J. Nonlinear Var. Anal. 4 (2020) 377-398.

[7] R. I. Boţ, E. R. Csetnek, P.T. Vuong, The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces, European Journal of Operational Research 287 (2020), 49-60.

[8] L. C. Ceng, Two inertial linesearch extragradient algorithms for the bilevel split pseudomonotone variational inequality with constraints, J. Appl. Numer. Optim. 2 (2020), 213-233.

[9] 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.

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

[11] Y. Censor, A. Gibali, S. Reich, Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space, Optimization 61 (2012), 1119-1132.

[12] A. Gibali, S. Reich, R. Zalas, Outer approximation methods for solving variational inequalities in Hilbert space, Optimization 66 (2017) 417-437.

[13] M. Tian, G. Xu, Inertial modified Tseng's extragradient algorithms for solving monotone variational inequalities and fixed point problems, J. Nonlinear Funct. Anal. (2020) 35.

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

[15] P. T. Vuong, Y. Shehu, Convergence of an extragradient-type method for variational inequality with applications to optimal control problems, Numer. Algorithms 81 (2019) 269-291.

[16] J. Yang, H. Liu, Strong convergence result for solving monotone variational inequalities in Hilbert space, Numer. Algorithms 80 (2019), 741-752.

[17] J. Yang, H. Liu, A modified projected gradient method for monotone variational inequalities, J. Optim. Theory Appl. 179 (2018), 197-211.

[18] G.M. Korpelevich, The extragradient method for finding saddle points and other problems, Ekon. Mat. Metody 12 (1976) 747-756 (in Russian).

[19] A. S. Antipin, On a method for convex programs using a symmetrical modification of the Lagrange function. Ekon. Mat. Metody 12 (1976), 1164-1173 (in Russian).

[20] L.D. Popov, A modification of the Arrow-Hurwicz method for search of saddle points, Mat. Zametski 28 (5) (1980) 777-784 (in Russian), Math. Notes 28 (5) (1980) 845-848 (in English).

[21] Q. L. Dong, Y. Y. Lu, J. F. Yang, The extragradient algorithm with inertial effects for solving the variational inequality, Optimization 65 (2016), 2217-2226.

[22] N. El Farouq, Pseudomonotone variational inequalities: convergence of the auxiliary problem method, J. Optim. Theory Appl. 111 (2001), 305-326.

[23] T. N. Hai, On gradient projection methods for strongly pseudomonotone variational inequalities without Lipschitz continuity, Optim. Lett. 14 (2019), 1177-1191.

[24] D. V. Hieu, D.V. Thong, New extragradient-like algorithms for strongly pseudomonotone variational inequalities, J. Glob. Optim. 70 (2018), 385-399.

[25] P. D. Khanh, P. T. Vuong, Modified projection method for strongly pseudomonotone variational inequalities, J. Glob. Optim. 58 (2014), 341-350.

[26] P. D. Khanh, A new extragradient method for strongly pseudomonotone variational inequalities, Numer. Funct. Anal. Optim. 37 (2016), 1131-1143.

[27] D. S. Kim, P. T. Vuong, P. D. Khanh, Qualitative properties of strongly pseudomonotone variational inequalities, Optim. Lett. 10 (2016), 1669-1679.

[28] D. V. Thong, D.V. Hieu, Inertial extragradient algorithms for strongly pseudomonotone variational inequalities, J. Comput. Appl. Math. 341 (2018), 80-98.

[29] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (2001), 3-11.

[30] Q. L. Dong, H. B. Yuan, Y. J. Cho, T. M. Rassias, Modified inertial Mann algorithm and inertial CQ-algorithm for nonexpansive mappings, Optim. Lett. 12 (2018), 87-102.

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

[32] D. V. Thong, D.V. Hieu, Modified subgradient extragradient method for inequality variational problems, Numer. Algorithms 79 (2018), 597-610.

[33] R. I. Boţ, M. Sedlmayer, P. T. Vuong, A Relaxed Inertial Forward-Backward-Forward Algorithm for Solving Monotone Inclusions with Application to GANs, arXiv:2003.07886 (2020).

[34] K. Goebel, S. Reich, Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings, Marcel Dekker, New York, 1984.

[35] J. M. Ortega, W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York (1970).