RAIRO-O per. Res. 54 (2020) 163–178
https://doi.org/10.1051/ro/2018115

RAIRO Operations Research
www.rairo-ro.org

AN INERTIAL MODIFIED ALGORITHM FOR SOLVING VARIATIONAL INEQUALITIES

Dang Van Hieu[1,*] and Pham Kim Quy[2]

Abstract. The paper deals with an inertial-like algorithm for solving a class of variational inequality problems involving Lipschitz continuous and strongly pseudomonotone operators in Hilbert spaces. The presented algorithm can be considered a combination of the modified subgradient extragradient-like algorithm and inertial effects. This is intended to speed up the convergence properties of the algorithm. The main feature of the new algorithm is that it is done without the prior knowledge of the Lipschitz constant and the modulus of strong pseudomonotonicity of the cost operator. Several experiments are performed to illustrate the convergence and computational performance of the new algorithm in comparison with others having similar features. The numerical results have confirmed that the proposed algorithm has a competitive advantage over the existing methods.

Mathematics Subject Classification. 65J15, 47H05, 47J25, 47J20, 91B50.

Received March 2, 2018. Accepted December 2, 2018.

1. Introduction

The variational inequality problem (VIP) [11, 13, 25–27] can be considered a fundamental problem in nonlinear analysis, especially in optimization theory. This problem comes from numerous problems in science and engineering including optimization problems, fixed point problems, transportation problems, financial equilibrium problems, migration equilibrium problems, see, *e.g.*, [3, 9, 10, 17]. This can be the main motivation for many authors who devoted their works to studying theory of variational analysis and after that introducing their iterative algorithms for approximating solutions of VIPs. Especially, problem (VIP) has been intensively exploited after appearing the monographs [11, 13, 26, 27]. Two notable solution approaches for solving problem (VIP) can be the regularized method and the projection method. For the first direction, it often includes monotone operator, and the regularized subproblem is thus strongly monotone. The regularized solution is unique and can be found more easily than solutions of the original problem. The sequence of regularized solutions can converge finitely or asymptotically to some solution of the original problem.

In this paper, we are interested in the second direction including projection methods. The methods of projection types are easier to solve numerically than the regularized methods. The simplest and oldest projection is the projected gradient method. After that many projection-like methods were proposed as the extragradient

Keywords. Modified extragradient method, inertial effect, variational inequality, pseudomonotone operator, Lipschitz continuity.

[1] Applied Analysis Research Group, Faculty of Mathematics and Statistics, Ton Duc Thang University, Ho Chi Minh City, Vietnam.

[2] Department of Mathematics, College of Air Force, Nha Trang City, Vietnam.

*Corresponding author: dangvanhieu@tdtu.edu.vn, dv.hieu83@gmail.com

method [28], the modified extragradient method [33], the subgradient extragradient method [6–8], the projected reflected gradient method [29] and others [14, 15, 23, 35, 37]. Recently, among kinds of generalized monotone operators [22], the group of strongly pseudomonotone operators has been widely and intensively investigated by several authors, for example, see [12, 38–40]. Note that the class of strongly pseumonotone operators contains properly the one of strongly monotone operators and the basic projection method is an efficient method for strongly monotone VIPs. Then, a natural question arises: Can the projection methods work well for strongly pseudomonotone VIPs? Several works proposed recently, for instance [18, 19, 24, 36, 38–41], have given positive answers for the aforementioned question.

It is emphasized here that almost aformentioned methods use stepsizes which depend on the Lipschitz constant or the modulus of strongly pseudomonotonicity. This can make several restrictions in applications because those constants are often unknown or difficult to approximate. Very recently, the authors in [19] have considered two extragradient − like algorithms for strongly pseudomonotone and Lipschitz VIPs. The main advantage of two algorithms in [19] is that they are done without the prior knownledge of the Lipschitz and strongly pseudomonotone constants. At this stage, it is worth mentioning that the second algorithm in Algorithm 2 of [19] has advantages over each computational step and the complexity of it is almost equivalent to the classical gradient method, *i.e.*, over each iteration it only requires to compute one projection and one value of operator at the current approximation.

Now, we wish to mention a kind of inertial-type algorithm which appears when we discretize a second order dissipative dynamical system in time [1, 2, 34]. The main feature of inetial algorithms is that the next iterate is constructed from the two previous iterates. Considering the algorithms with inertial efforts is intended to speed up the convergence properties. In recent years, many algorithms with inertial effects have been proposed for solving VIPs as well as some related problems, for instance, see in [4, 5, 30–32] and the references cited therein.

The main purpose of this paper is to speed up the convergence properties of the modified subgradient extragradient method (MSEGM) in Algorithm 2 of [19]. As aforementioned, this algorithm has notable advantages over each computational step. We will describe how to combine MSEGM with inertial effects where the complexity of the resulting algorithm is still equivalent to the original one. The new algorithm also works well for the class of Lipschitz and strongly pseudomonotone VIPs. The using a sequence of stepsizes, which is non-summable and diminishing, permits the algorithm to be done without previously knowing any information of Lipschitz and strongly pseudomonotone constants. Several numerical experiments are performed to support the theoretical results, and also to compare with others. Our numerical results have shown its effectiveness and fast convergence over some existing algorithms having the same features.

The organization of the rest of the paper is as follows: Section 2 recalls some definitions and preliminary results used throughout the paper. Section 3 deals with proposing the new algorithm and analyzing the convergence of it. Finally, in Section 4 we perform several experiments to show the numerical behavior of the new algorithm in comparisons with others.

2. Preliminaries

Let H denote a real Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$. Let C be a nonempty closed convex subset of H and $A : H \to H$ be an operator. The variational inequality problem (VIP) for A on C is to find $p \in C$ such that

$$\langle A(p), x - p \rangle \geq 0, \ \forall x \in C. \tag{VIP}$$

It is well-known that problem (VIP) is equivalent to the solving of the following fixed point problem,

$$x = P_C(x - \lambda Ax),$$

where λ is some positive number and P_C is the metric projection from H onto C. Recall that the projection $P_C : H \to C$ is defined, for each $x \in H$, by

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

Since C be a nonempty closed convex subset, $P_C(x)$ exists and is unique. Given $x \in H$ and $v \in H$, $v \neq 0$ and let $T = \{z \in H : \langle v, z - x\rangle \leq 0\}$ be a half-space. Then, for all $u \in H$, the projection $P_T(u)$ is defined by

$$P_T(u) = u - \max\left\{0, \frac{\langle v, u - x\rangle}{||v||^2}\right\} v, \tag{2.1}$$

which gives us an explicit formula to find the projection of any point onto a half-space.

From the definition, it is easy to show that P_C has the following characteristic properties, see [16] for more details.

Lemma 2.1. (i) $\langle P_C(x) - P_C(y), x - y\rangle \geq \|P_C(x) - P_C(y)\|^2$, $\forall x, y \in H$.
(ii) $\|x - P_C(y)\|^2 + \|P_C(y) - y\|^2 \leq \|x - y\|^2, \forall x \in C, y \in H$.
(iii) $z = P_C(x) \Leftrightarrow \langle x - z, y - z\rangle \leq 0, \quad \forall y \in C$.

Lemma 2.2. *For all $x, y \in H$ and $\alpha \in \Re$, the following equality holds,*

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

Next, we present some concepts of monotonicity of an operator. An operator $A : H \to H$ is said to be:

(i) *strongly monotone* on C if there exists $\gamma > 0$ such that

$$\langle A(x) - A(y), x - y\rangle \geq \gamma||x - y||^2, \quad \forall x, y \in C;$$

(ii) *monotone* on C if

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

(iii) *strongly pseudomonotone* on C if there exists $\gamma > 0$ such that

$$\langle A(x), y - x\rangle \geq 0 \Longrightarrow \langle A(y), y - x\rangle \geq \gamma||x - y||^2, \ \forall x, y \in C.$$

(iv) *pseudomonotone* on C if

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

(v) L – *Lipschitz continuous* on C if there exists $L > 0$ such that

$$||A(x) - A(y)|| \leq L||x - y||, \ \forall x, y \in C.$$

From the aforementioned definitions, it is easy to see that the following implications hold,

$$\text{(i)} \Longrightarrow \text{(ii)} \Longrightarrow \text{(iv) and (i)} \Longrightarrow \text{(iii)} \Longrightarrow \text{(iv)}.$$

3. Inertial modified algorithm

Throughout this section, we consider the operator $A : C \to H$ being strong pseudomonotone with some constant γ and Lipschitz continuous with some constant L, but these constants are not necessary to be known. The unique solution of problem (VIP) for A on C will be denoted by p. By some aforementioned advantages of the modified subgradient extragradient algorithm in Algorithm 2 of [19], we intend here to speed up the convergence properties of this algorithm. In that purpose, we reconsider it in incorporating with inertial effects. For designing the new algorithm, we take a sequence of stepsizes $\{\lambda_n\} \subset (0, +\infty)$ satisfying the following conditions,

$$\text{(C1)} : \lim_{n\to\infty} \lambda_n = 0, \qquad \text{(C2)} : \sum_{n=1}^{\infty} \lambda_n = +\infty,$$

and a decreasing parameter sequence $\{\theta_n\}$ in $[0,1]$ satisfying the condition,

$$(\text{C3}): 0 \le \theta_n \le \theta < \frac{1}{5}.$$

Unlike in Algorithm 2 of [19], the sequence of stepsize $\{\lambda_n\}$ is not required to be decreasing. This can give us a more flexibility in choosing a sequence $\{\lambda_n\}$ in numerical computations. The following is the algorithm in details.

Algorithm 3.1. (Inertial MSEGM for VIPs).

Initialization: Choose $x_{-1}, x_0 \in H$, $y_0 \in C$ and two sequence $\{\lambda_n\}$ and $\{\theta_n\}$ such that conditions (C1)–(C3) above hold. Set $w_0 = x_0 + \theta_0(x_0 - x_{-1})$ and compute

$$x_1 = P_C(w_0 - \lambda_0 A y_0), \quad w_1 = x_1 + \theta_1(x_1 - x_0), \quad y_1 = P_C(w_1 - \lambda_1 A y_0).$$

Iterative Steps: Assume that $x_{n-1},\ x_n,\ w_n \in H$ and $y_{n-1},\ y_n \in C$ are known, calculate x_{n+1}, y_{n+1}, w_{n+1} for each $n \ge 1$ as follows:

Step 1. Compute $x_{n+1} = P_{T^n}(w_n - \lambda_n A y_n)$,
where $T^n = \{x \in H : \langle w_n - \lambda_n A y_{n-1} - y_n, x - y_n \rangle \le 0\}$.
Step 2. Set $w_{n+1} = x_{n+1} + \theta_{n+1}(x_{n+1} - x_n)$ and compute

$$y_{n+1} = P_C(w_{n+1} - \lambda_{n+1} A y_n).$$

Stopping Criterion: If $y_{n+1} = w_{n+1} = y_n$ then stop and y_n is the solution of problem (VIP).

Before the formulation of convergence of Algorithm 3.1, we present the following remarks.

Remark 3.2. Since T^n is a half-space, the projection in Step 1 can be found explicitly by the formula (2.1). The main tasks of Algorithm 3.1 over each iteration are to find one projection for y_{n+1} on C and compute one value of operator A at y_n. Then, in that sense, the complexity of Algorithm 3.1 is almost equivalent to the one of the classical projected gradient algorithm.

Remark 3.3. In the case when $\theta_n = 0$ then Algorithm 3.1 becomes the modified subgradient extragradient algorithm in Algorithm 2 of [19]. The term $\theta_{n+1}(x_{n+1} - x_n)$ in Step 2 is called inertial effect. This term is incorporated in the algorithm and also intended to speed up the convergence property. Comparing with the inertial subgradient extragradient method (ISEGM) in Theorem 3.1 of [36], Algorithm 3.1 only requries to compute one value of operator A at y_n over each iteration, while the ISEGM needs to find two values of operator A. Moreover, from condition (C3), we see that the interval in which the inertial parameter θ_n can be chosen in Algorithm 3.1 is larger than that one in [36].

Remark 3.4. We remark that $C \subset T^n$ for all $n \ge 0$. Indeed, from the definition of y_n and Lemma 2.1(iii) we obtain $\langle w_n - \lambda_n A y_{n-1} - y_n, x - y_n \rangle \le 0$ for all $x \in C$. This together with the definition of T^n implies that $C \subset T^n$ for all $n \ge 0$.

Now, we will analyze the convergence of Algorithm 3.1. Since $0 \le \theta < \frac{1}{5}$, we get that $\frac{1-5\theta}{1+3\theta} > 0$. Throughout this section, let us take $\epsilon \in \left(0, \frac{1-5\theta}{1+3\theta}\right)$ being fixed and enough small, and set $\delta = 1 - \epsilon$. Thus, we have that $0 < \delta < 1$. Moreover, we also define A_n, M_n, N_n for each $n \ge 0$ by

$$\begin{cases} A_n = ||x_n - p||^2 - \theta_n ||x_{n-1} - p||^2 + 2\lambda_n L ||y_{n-1} - w_n||^2, \\ M_n = \frac{\delta}{2}(1 - \theta_n) - 2\lambda_{n+1} L \theta_{n+1}(1 + \theta_{n+1}), \\ N_n = \theta_n(1 + \theta_n) + \frac{\delta}{2}\theta_n(1 - \theta_n). \end{cases} \tag{3.1}$$

We begin with the following lemma which plays an important role in establishing the convergence of Algorithm 3.1.

Lemma 3.5. *There exists* $n_0 \geq 0$ *such that*

$$A_{n+1} \leq A_n + N_n||x_n - x_{n-1}||^2 - M_n||x_{n+1} - x_n||^2 - 2\gamma\lambda_n||y_n - p||^2, \ \forall n \geq n_0.$$

Proof. From the definition of x_{n+1} and Lemma 2.1(iii), we obtain

$$\langle x - x_{n+1}, w_n - \lambda_n Ay_n - x_{n+1}\rangle \leq 0, \ \forall x \in T^n,$$

which follows that $2\langle p - x_{n+1}, w_n - x_{n+1}\rangle \leq 2\lambda_n \langle p - x_{n+1}, Ay_n\rangle$ because of $p \in C \subset T^n$. Combining this with the equality $2\langle a, b\rangle = ||a||^2 + ||b||^2 - ||a - b||^2$, we get

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

Moreover, since $x_{n+1} \in T^n$, it follows from the definition of T^n that

$$\langle w_n - \lambda_n Ay_{n-1} - y_n, x_{n+1} - y_n\rangle \leq 0,$$

which implies $2\langle w_n - y_n, x_{n+1} - y_n\rangle \leq 2\lambda_n \langle Ay_{n-1}, x_{n+1} - y_n\rangle$. Therefore, also by using the equality $2\langle a, b\rangle = ||a||^2 + ||b||^2 - ||a - b||^2$, we come to the following estimate,

$$||x_{n+1} - y_n||^2 + ||w_n - y_n||^2 - ||x_{n+1} - w_n||^2 \leq 2\lambda_n \langle x_{n+1} - y_n, Ay_{n-1}\rangle. \tag{3.3}$$

Adding both two sides of the inequalities (3.2) and (3.3), respectively, we get

$$\begin{aligned} ||x_{n+1} - p||^2 &\leq ||w_n - p||^2 - ||w_n - y_n||^2 - ||x_{n+1} - y_n||^2 \\ &\quad + 2\lambda_n [\langle p - x_{n+1}, Ay_n\rangle + \langle x_{n+1} - y_n, Ay_{n-1}\rangle] \\ &\leq ||w_n - p||^2 - ||w_n - y_n||^2 - ||x_{n+1} - y_n||^2 \\ &\quad + 2\lambda_n [\langle p - y_n, Ay_n\rangle + \langle y_n - x_{n+1}, Ay_n - Ay_{n-1}\rangle]. \end{aligned} \tag{3.4}$$

Since p is the solution of problem (VIP), $\langle Ap, y_n - p\rangle \geq 0$. Thus, from the pseudomonotonicity of A, we obtain that

$$\langle p - y_n, Ay_n\rangle \leq -\gamma||y_n - p||^2. \tag{3.5}$$

Now, we will estimate the term $\langle y_n - x_{n+1}, Ay_n - Ay_{n-1}\rangle$ in the righ-hand side of inequality (3.4). Using the Lipschitz property of A, the Cauchy–Schwarz inequality and the Cauchy inequality, we get that

$$\begin{aligned} \langle y_n - x_{n+1}, Ay_n - Ay_{n-1}\rangle &\leq ||y_n - x_{n+1}||||Ay_n - Ay_{n-1}|| \\ &\leq L||y_n - x_{n+1}||||y_n - y_{n-1}|| \\ &\leq \frac{L}{2}||y_n - x_{n+1}||^2 + \frac{L}{2}||y_n - y_{n-1}||^2 \\ &\leq \frac{L}{2}||y_n - x_{n+1}||^2 + \frac{L}{2}(2||y_n - w_n||^2 + 2||w_n - y_{n-1}||^2) \\ &= \frac{L}{2}||y_n - x_{n+1}||^2 + L||y_n - w_n||^2 + L||w_n - y_{n-1}||^2). \end{aligned} \tag{3.6}$$

Using relation (3.4) and taking into acount relations (3.5) and (3.6), we come to

$$\begin{aligned} ||x_{n+1} - p||^2 \leq ||w_n - p||^2 &- (1 - 2\lambda_n L)||w_n - y_n||^2 - (1 - \lambda_n L)||x_{n+1} - y_n||^2 \\ &+ 2\lambda_n L||w_n - y_{n-1}||^2 - 2\gamma\lambda_n||y_n - p||^2. \end{aligned} \tag{3.7}$$

Adding the term $2\lambda_{n+1}L||y_n - w_{n+1}||^2$ to both two sides of relation (3.7), we obtain

$$\begin{aligned} ||x_{n+1}-p||^2 + 2\lambda_{n+1}L||y_n - w_{n+1}||^2 \le{}& ||w_n - p||^2 - (1-2\lambda_n L)||w_n - y_n||^2 \\ &- (1-\lambda_n L)||x_{n+1}-y_n||^2 + 2\lambda_n L||w_n - y_{n-1}||^2 \\ &+ 2\lambda_{n+1}L||y_n - w_{n+1}||^2 - 2\gamma\lambda_n||y_n - p||^2. \end{aligned} \tag{3.8}$$

On the other hand, from the definiton of w_n, Lemma 2.2 and the fact $\theta_n \le \theta_{n+1}$, we obtain

$$\begin{aligned} ||w_n - p||^2 &= ||(1+\theta_n)(x_n - p) - \theta_n(x_{n-1}-p)||^2 \\ &= (1+\theta_n)||x_n - p||^2 - \theta_n||x_{n-1}-p||^2 + \theta_n(1+\theta_n)||x_n - x_{n-1}||^2 \\ &\le (1+\theta_{n+1})||x_n - p||^2 - \theta_n||x_{n-1}-p||^2 \\ &\quad +\theta_n(1+\theta_n)||x_n - x_{n-1}||^2. \end{aligned} \tag{3.9}$$

Similarly, we also have

$$\begin{aligned} ||w_{n+1} - y_n||^2 &= (1+\theta_{n+1})||x_{n+1}-y_n||^2 - \theta_{n+1}||x_n - y_n||^2 \\ &\quad +\theta_{n+1}(1+\theta_{n+1})||x_{n+1}-x_n||^2 \\ &\le (1+\theta_{n+1})||x_{n+1}-y_n||^2 + \theta_{n+1}(1+\theta_{n+1})||x_{n+1}-x_n||^2. \end{aligned}$$

Substituting this and relation (3.9) into the right-hand side of relation (3.8), we come to the following estimate

$$\begin{aligned} ||x_{n+1}-p||^2 + 2\lambda_{n+1}L||y_n - w_{n+1}||^2 \le{}& (1+\theta_{n+1})||x_n - p||^2 - \theta_n||x_{n-1}-p||^2 \\ &+ \theta_n(1+\theta_n)||x_n - x_{n-1}||^2 - (1-2\lambda_n L)||w_n - y_n||^2 - (1-\lambda_n L)||x_{n+1}-y_n||^2 \\ &+ 2\lambda_n L||w_n - y_{n-1}||^2 + 2\lambda_{n+1}L\left[(1+\theta_{n+1})||x_{n+1}-y_n||^2\right. \\ &\left.+\theta_{n+1}(1+\theta_{n+1})||x_{n+1}-x_n||^2\right] - 2\gamma\lambda_n||y_n - p||^2. \end{aligned}$$

Thus

$$\begin{aligned} ||x_{n+1}-p||^2 - \theta_{n+1}||x_n - p||^2 + 2\lambda_{n+1}L||y_n - w_{n+1}||^2 \le{}& ||x_n - p||^2 \\ &- \theta_n||x_{n-1}-p||^2 + 2\lambda_n L||w_n - y_{n-1}||^2 + \theta_n(1+\theta_n)||x_n - x_{n-1}||^2 \\ &- (1-2\lambda_n L)||w_n - y_n||^2 - (1-\lambda_n L - 2\lambda_{n+1}L(1+\theta_{n+1}))||x_{n+1}-y_n||^2 \\ &+ 2\lambda_{n+1}L\theta_{n+1}(1+\theta_{n+1})||x_{n+1}-x_n||^2 - 2\gamma\lambda_n||y_n - p||^2. \end{aligned} \tag{3.10}$$

Since $\lambda_n \to 0$ and $0 < \delta < 1$, there exists $n_0 \ge 0$ such that

$$1 - 2\lambda_n L \ge \delta > 0 \text{ and } 1 - \lambda_n L - 2\lambda_{n+1}L(1+\theta_{n+1}) \ge \delta > 0. \tag{3.11}$$

These together with relation (3.10) and the definition of A_n in (3.1), we obtain

$$\begin{aligned} A_{n+1} \le{}& A_n - \delta(||w_n - y_n||^2 + ||x_{n+1}-y_n||^2) + \theta_n(1+\theta_n)||x_n - x_{n-1}||^2 \\ &+ 2\lambda_{n+1}L\theta_{n+1}(1+\theta_{n+1})||x_{n+1}-x_n||^2 - 2\gamma\lambda_n||y_n - p||^2. \end{aligned} \tag{3.12}$$

Moreover, we have

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

This together with relation (3.12) and the definitions of M_n, N_n in (3.1) implies that

$$A_{n+1} \leq A_n + N_n||x_n - x_{n-1}||^2 - M_n||x_{n+1} - x_n||^2 - 2\gamma\lambda_n||y_n - p||^2, \ \forall n \geq n_0.$$

This completes the proof of Lemma 3.5. □

Next, we define φ_n for each $n \geq 0$ by $\varphi_n = A_n + N_n||x_n - x_{n-1}||^2$, *i.e.*,

$$\varphi_n := ||x_n - p||^2 - \theta_n||x_{n-1} - p||^2 + 2\lambda_n L||y_{n-1} - w_n||^2 + N_n||x_n - x_{n-1}||^2. \tag{3.13}$$

We have the following lemma.

Lemma 3.6. *The following estimates hold for all $n \geq n_0$,*

(i) $\varphi_{n+1} - \varphi_n \leq -\sigma||x_{n+1} - x_n||^2 - 2\gamma\lambda_n||y_n - p||^2$,
(ii) $\varphi_n \geq \tau||x_n - p||^2$, *where $\tau > 0$, $\sigma > 0$ are some positive real numbers.*

Proof. (i) From the definition of φ_n and Lemma 3.5, we have

$$\begin{aligned}
\varphi_{n+1} - \varphi_n &= A_{n+1} + N_{n+1}||x_{n+1} - x_n||^2 - A_n - N_n||x_n - x_{n-1}||^2 \\
&\leq -M_n||x_{n+1} - x_n||^2 + N_{n+1}||x_{n+1} - x_n||^2 - 2\gamma\lambda_n||y_n - p||^2 \\
&= -(M_n - N_{n+1})||x_{n+1} - x_n||^2 - 2\gamma\lambda_n||y_n - p||^2.
\end{aligned} \tag{3.14}$$

Note that $\theta_n \leq \theta_{n+1}$, $\theta_{n+1}^2 \leq \theta_{n+1}$, and the fact $L\lambda_{n+1} \leq \frac{1-\delta}{2}$ for all $n \geq n_0$ (see, relation (3.11)). Thus, it follows from the definition of M_n and N_{n+1} that

$$\begin{aligned}
M_n - N_{n+1} &= \frac{\delta}{2}(1-\theta_n) - 2\lambda_{n+1}L\theta_{n+1}(1+\theta_{n+1}) - \theta_{n+1}(1+\theta_{n+1}) \\
&\quad -\frac{\delta}{2}\theta_{n+1}(1-\theta_{n+1}) \\
&\geq \frac{\delta}{2}(1-\theta_{n+1}) - 2\lambda_{n+1}L\theta_{n+1}(1+\theta_{n+1}) - \theta_{n+1}(1+\theta_{n+1}) \\
&\quad -\frac{\delta}{2}\theta_{n+1}(1-\theta_{n+1}) \\
&= \frac{\delta}{2} - (1 + 2\lambda_{n+1}L + \delta)\theta_{n+1} - (1 + 2\lambda_{n+1}L - \frac{\delta}{2})\theta_{n+1}^2 \\
&\geq \frac{\delta}{2} - (1 + 2\lambda_{n+1}L + \delta)\theta_{n+1} - (1 + 2\lambda_{n+1}L - \frac{\delta}{2})\theta_{n+1}
\end{aligned}$$

$$= \frac{\delta}{2} - (2 + 4\lambda_{n+1}L + \frac{\delta}{2})\theta_{n+1}$$
$$\geq \frac{\delta}{2} - (2 + 4\frac{1-\delta}{2} + \frac{\delta}{2})\theta$$
$$= \frac{\delta - (8-3\delta)\theta}{2} =: \sigma.$$

Thus, since $\delta = 1 - \epsilon$,

$$M_n - N_{n+1} \geq \sigma = \frac{1 - \epsilon - (8 - 3(1-\epsilon))\theta}{2} = \frac{(1-5\theta) - \epsilon(1+3\theta)}{2} > 0,$$

due to $\epsilon \in \left(0, \frac{1-5\theta}{1+3\theta}\right)$. This together with relation (3.14) implies conclusion (i).

(ii) From the definition of φ_n in (3.13), we obtain

$$\varphi_n \geq ||x_n - p||^2 - \theta_n||x_{n-1} - p||^2 + N_n||x_n - x_{n-1}||^2. \tag{3.15}$$

Set $k_n = \frac{2}{2\theta_n + \delta(1-\theta_n)} > 0$. This together with the definition of N_n implies that

$$N_n - \theta_n(1 + \frac{1}{k_n}) = 0. \tag{3.16}$$

Moreover, we have that

$$\begin{aligned}||x_{n-1} - p||^2 &= ||x_{n-1} - x_n||^2 + ||x_n - p||^2 + 2\langle x_{n-1} - x_n, x_n - p\rangle \\ &\leq ||x_{n-1} - x_n||^2 + ||x_n - p||^2 + \frac{1}{k_n}||x_{n-1} - x_n||^2 + k_n||x_n - p||^2 \\ &= (1 + \frac{1}{k_n})||x_{n-1} - x_n||^2 + (1 + k_n)||x_n - p||^2. \end{aligned} \tag{3.17}$$

It follows from relations (3.15), (3.16) and (3.17) that

$$\begin{aligned}\varphi_n &\geq (1 - \theta_n(1+k_n))||x_n - p||^2 + (N_n - \theta_n(1 + \frac{1}{k_n}))||x_n - x_{n-1}||^2 \\ &= (1 - \theta_n(1+k_n))||x_n - p||^2. \end{aligned} \tag{3.18}$$

On the other hand, from the definition of k_n, we have

$$1 - \theta_n(1 + k_n) = 1 - \theta_n\left(1 + \frac{2}{2\theta_n + \delta(1-\theta_n)}\right) = -\theta_n - \frac{\delta\theta_n - \delta}{(2-\delta)\theta_n + \delta}. \tag{3.19}$$

We define the function $f(t) = -t - \frac{\delta t - \delta}{(2-\delta)t + \delta}$, $t \in [0, 1]$. We have that $f'(t) = -1 - \frac{2}{((2-\delta)t+\delta)^2} < 0$. Thus, $f(t)$ is decreasing monotone. Since $\theta_n \leq \theta$, we come to the following estimate,

$$1 - \theta_n(1 + k_n) = f(\theta_n) \geq f(\theta) = \frac{\delta(1-\theta)^2 - 2\theta^2}{(2-\delta)\theta + \delta} := \tau > 0, \tag{3.20}$$

in which the inequality $\tau > 0$ follows from the facts that $\theta < \frac{1}{5}$ and $\delta = 1 - \epsilon$ is enough near to 1. Combining relations (3.18) and (3.20), we obtain the conclusion (ii). Lemma 3.6 is proved.

□

Lemma 3.7. *The following statements hold:*

(i) *The sequences $\{x_n\}$ is bounded.*
(ii) $\lim_{n\to\infty} \inf ||y_n - p|| = 0$ *and* $\lim_{n\to\infty} ||x_n - x_{n+1}|| = 0$.
(iii) $\lim_{n\to\infty} ||x_n - y_n|| = \lim_{n\to\infty} ||w_n - y_{n-1}|| = 0$.

Proof. (i) It follows from Lemma 3.6(i) that $\{\varphi_n\}_{n\geq n_0}$ is non-increasing, *i.e.*, $\varphi_{n+1} \leq \varphi_n$ for all $n \geq n_0$. On the other hand, it follows from Lemma 3.6(ii) that $\varphi_n \geq 0$ for all $n \geq n_0$. Thus, the limit of $\{\varphi_n\}$ exists. This together with Lemma 3.6(ii) implies that the sequence $\{||x_n - p||^2\}$, and therefore $\{x_n\}$, are bounded. This completes the proof of (i).

(ii) In view of Lemma 3.6(i), we see that

$$0 \leq \sigma||x_{n+1} - x_n||^2 + 2\gamma\lambda_n||y_n - p||^2 \leq \varphi_n - \varphi_{n+1}, \ \forall n \geq n_0.$$

Thus, since $\lim_{n\to\infty} \varphi_{n+1} \in \Re$, we obtain

$$\sigma \sum_{n=n_0}^{\infty} ||x_n - x_{n+1}||^2 + 2\gamma \sum_{n=n_0}^{\infty} \lambda_n||y_n - p||^2 \leq \varphi_{n_0} - \lim_{n\to\infty} \varphi_{n+1} < +\infty. \tag{3.21}$$

Therefore

$$\text{(S1)}: \sum_{n=n_0}^{\infty} \lambda_n||y_n - p||^2 < +\infty, \qquad \text{(S2)}: \sum_{n=n_0}^{\infty} ||x_n - x_{n+1}||^2 < +\infty.$$

It follows from (S1) and the fact $\sum_{n=n_0}^{\infty} \lambda_n = +\infty$ that $\lim_{n\to\infty} \inf ||y_n - p|| = 0$. From (S2), we also obtain that $||x_n - x_{n+1}||^2 \to 0$.

(iii) It follows from Lemma 3.7(ii) that $N_n||x_n - x_{n-1}||^2 \to 0$ as $n \to \infty$ because of the boundedness of $\{N_n\}$. Also from the definition of φ_n, we see that $A_n = \varphi_n - N_n||x_n - x_{n-1}||^2$, and so the limit of $\{A_n\}$ also exists. Thus, from relation (3.12), we come to

$$\begin{aligned}\delta(||w_n - y_n||^2 + ||x_{n+1} - y_n||^2) \leq A_n - A_{n+1} + \theta_n(1 + \theta_n)||x_n - x_{n-1}||^2 \\ +2\lambda L\theta_{n+1}(1 + \theta_{n+1})||x_{n+1} - x_n||^2 \to 0\end{aligned}$$

as $n \to \infty$. This implies that $||w_n - y_n|| \to 0$ and $||x_{n+1} - y_n|| \to 0$. Thus,

$$||x_n - y_n|| \leq ||x_n - x_{n+1}|| + ||x_{n+1} - y_n|| \to 0.$$

It follows from the definition of w_n that

$$\begin{aligned}||w_n - y_{n-1}||^2 = (1 + \theta_n)||x_n - y_{n-1}||^2 - \theta_n||x_{n-1} - y_{n-1}||^2 \\ +\theta_n(1 + \theta_n)||x_n - x_{n-1}||^2,\end{aligned}$$

which follows $||w_n - y_{n-1}|| \to \infty$ as $n \to \infty$. This finishes the proof. □

Now, we obtain the following main result.

Theorem 3.8. *The sequences $\{x_n\}$, $\{y_n\}$, $\{w_n\}$ generated by Algorithm 3.1 converge strongly to the unique solution p of problem (VIP).*

Proof. From the definition of φ_n, we obtain

$$\begin{aligned}\varphi_n &= ||x_n - p||^2 - \theta_n||x_{n-1} - p||^2 + 2\lambda_n L||y_{n-1} - w_n||^2 + N_n||x_n - x_{n-1}||^2 \\ &= -\theta_n(||x_{n-1} - p||^2 - ||x_n - p||^2) + 2\lambda_n L||y_{n-1} - w_n||^2 + N_n||x_n - x_{n-1}||^2 \\ &\quad +(1-\theta_n)||x_n - p||^2. \end{aligned} \tag{3.22}$$

Since the sequence $\{x_n\}$ is bounded and $||x_n - x_{n-1}|| \to 0$, we get

$$\begin{aligned}\left|||x_{n-1} - p||^2 - ||x_n - p||^2\right| &= |||x_{n-1} - p|| - ||x_n - p||| \left(||x_{n-1} - p|| + ||x_n - p||\right) \\ &\leq |||x_{n-1} - x_n||| \left(||x_{n-1} - p|| + ||x_n - p||\right) \to 0.\end{aligned}$$

Thus $||x_{n-1} - p||^2 - ||x_n - p||^2 \to 0$ as $n \to \infty$. Combining this and Lemma 3.7(ii, iii) with relation (3.22), we obtain

$$\begin{aligned}\lim_{n\to\infty} \inf \varphi_n &= \lim_{n\to\infty} \inf(1-\theta_n)||x_n - p||^2 \leq \lim_{n\to\infty} \inf ||x_n - p||^2 \\ &= \lim_{n\to\infty} \inf ||y_n - p||^2 = 0.\end{aligned}$$

Thus $\lim_{n\to\infty} \inf \varphi_n = 0$. Since the limit of $\{\varphi\}$ exists, $\lim_{n\to\infty} \varphi_n = 0$. This together with Lemma 3.6(ii) implies that $||x_n - p||^2 \to 0$, *i.e.*, the sequence $\{x_n\}$ converges strongly to the solution p of problem (VIP). The convergence of the sequences $\{y_n\}$, $\{w_n\}$ to p follows from Lemma 3.7(iii). Theorem 3.8 is proved. □

Remark 3.9. Algorithm 3.1 has combined the positive features of the Popov's extragradient algorithm in [33], the subgradient extragradient method in [6–8], and inertial effects. The analyses in this paper seem more complex and technical than the ones in Algorithm 2 of [19]. This is due to the inertial term in the algorithm. We remark additionally that by combining the Popov's extragradient algorithm in [33] with inertial effects, we also obtain the following inertial algorithm.

Algorithm 3.10. (Inertial Modified Extragradient Algorithm for VIPs).

Initialization: Choose $x_{-1},\ x_0,\ y_0 \in C$ and two sequences $\{\lambda_n\} \subset (0, +\infty)$, $\{\theta_n\} \subset [0, +\infty)$ such that conditions (C1)–(C3) above hold. Set $w_0 = x_0 + \theta_0(x_0 - x_{-1})$.

Iterative steps: Assume that $x_n,\ y_n \in C$ and $w_n \in H$ are known. Calculate x_{n+1}, y_{n+1} and w_{n+1} as follows:

$$\begin{cases} x_{n+1} = P_C(w_n - \lambda_n A y_n), \\ w_{n+1} = x_{n+1} + \theta_{n+1}(x_{n+1} - x_n), \\ y_{n+1} = P_C(w_{n+1} - \lambda_{n+1} A y_n). \end{cases}$$

Stopping criterion: If $y_{n+1} = w_{n+1} = y_n$ then stop and y_{n+1} is the solution of problem (VIP).

As the previous algorithm, Algorithm 3.10 also only requires to compute a value of operator A at y_n. However, it needs to compute two projections on the feasible set C. The proof of convergence of Algorithm 3.10 is almost similar to the one of Algorithm 3.1, but it is slightly different from obtaining relations (3.2) and (3.3) in the proof of Lemma 3.5. We leave the proof for the reader to verify. Finally, we have the following result.

Theorem 3.11. *The conclusion of Theorem 3.8 remains true for Algorithm 3.10.*

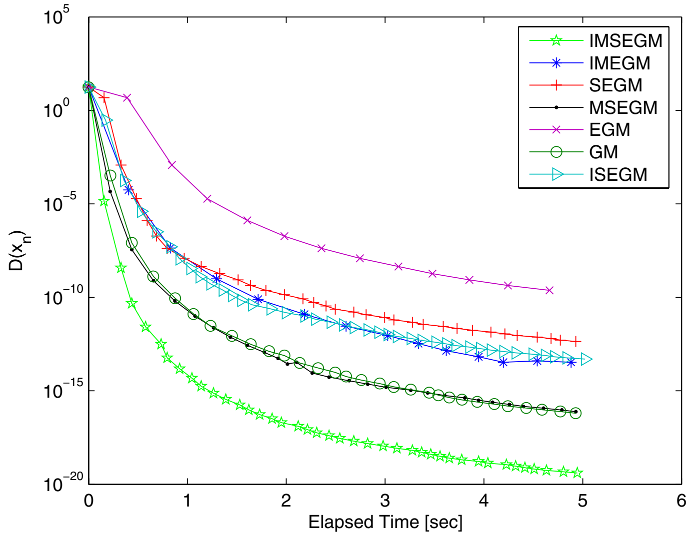

FIGURE 1. Example 1 in $\Re^{100}$. Numbers of iterations respectively are 243, 82, 203, 170, 77, 178, 234.

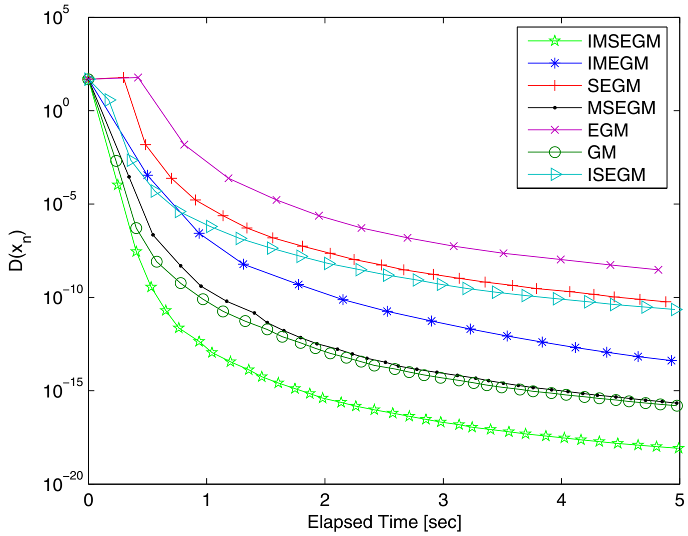

FIGURE 2. Example 1 in $\Re^{200}$. Numbers of iterations respectively are 205, 86, 136, 193, 74, 206, 127.

4. Computational experiments

In this section, we report some numerical results to illustrate the convergence of Algorithm 3.1 (shortly, IMSEGM) and Algorithm 3.10 (IMEGM), and also to compare with other algorithms having the same features. Five algorithms used to compare here are the subgradient extragradient method (SEGM) ([19], Algorithm 1), the modified subgradient extragradient method (MSEGM) ([19], Algorithm 2), the extragradient method (EGM)

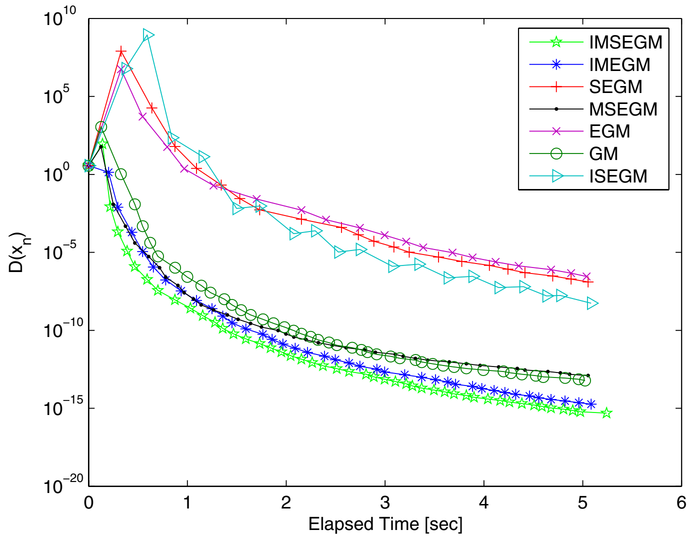

FIGURE 3. Example 2 in $\Re^{50}$. Numbers of iterations respectively are 47, 40, 22, 40, 20, 45, 20.

([20], Cor. 3.3), the gradient method (GM) ([24], Algorithm 3.1 and Thm. 5.1) and the inertial subgradient extragradient method (ISEGM) ([36], Thm. 3.1).

We use the function $D(x) = ||x - P_C(x - \lambda Ax)||^2$ for some $\lambda > 0$ to compare the algorithms. Note that if $D(x) = 0$ then x is the solution of problem (VIP). The convergence of $D(x_n)$ to 0 implies that the sequence $\{x_n\}$ generated by each algorithm converges to the solution of the problem. The starting points are $x_{-1} = x_0 = y_0 = (1, 1, \dots, 1) \in \Re^m$, and the sequence of stepsizes is $\lambda_n = \frac{1}{\log^{15}(n+2)}$, and the inertial parameter is $\theta_n = 0.1$. The feasible set is a polyhedral convex set given by

$$C = \left\{x \in \Re^m_+ : Ex \leq f\right\},$$

where E is a random matrix of size $l \times m$ with $l = 10$, and $f \in \Re^l_+$ such that $y_0 \in C$. All the programs are written on Matlab 7.0 and computed on a PC Desktop Intel(R) Core(TM) i5-3210M CPU @ 2.50 GHz, RAM 2.00 GB.

Example 4.1. We first consider problem (VIP) for a linear operator $A : \Re^m \to \Re^m$ $(m = 100,\ 200)$ of the form $A(x) = Mx + q$ [19], where

$$M = NN^T + S + D,$$

q is a vector in $\Re^m$, N is a $m \times m$ matrix, S is a $m \times m$ skew-symmetric matrix with their entries being generated in $(-2, 2)$ and D is a $m \times m$ diagonal matrix, whose diagonal entries are positive in $(0, 2)$ (so M is positive symmetric definite). It is clear that A is strongly pseudomonotone and Lipschitz continuous. The numerical results for this example are shown in Figures 1 and 2.

Example 4.2. Next, we consider our problem for the nonlinear operator $A : \Re^m \to \Re^m$ $(m = 50,\ 100)$ of the form $A(x) = Mx + F(x) + q$ (see, [21]) and M is a $m \times m$ symmetric semidefinite matrix and $F(x)$ is the proximal mapping of the function $g(x) = \frac{1}{4}||x||^4$, *i.e.*,

$$F(x) = \arg\min\left\{\frac{||y||^4}{4} + \frac{1}{2}||y - x||^2 : y \in \Re^m\right\}.$$

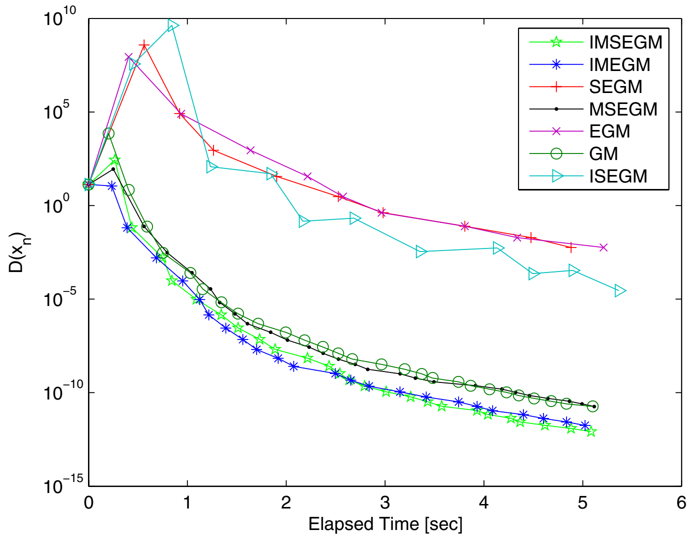

FIGURE 4. Example 2 in $\Re^{100}$. Numbers of iterations respectively are 26, 24, 10, 27, 10, 28, 12.

In this case, A is strongly pseudomonotone and Lipschitz continuous. For experiment, all the entries of M and q are also generated randomly as in the previous example. The numerical results are described in Figures 3 and 4. It is clear that, for this example, the computation of value of operator is expensive, and the numbers of iterations are thus smaller than the ones in the previous example.

Example 4.3. Finally, we consider the nonlinear operator $A : \Re^{2m} \to \Re^{2m}$ $(m = 50,\ 100)$ defined by

$$Ax = \begin{pmatrix} x_1 + x_2 + \sin x_1 \\ -x_1 + x_2 + \sin x_2 \\ x_3 + x_4 + \sin x_3 \\ -x_3 + x_4 + \sin x_4 \\ \dots \\ x_{2m-1} + x_{2m} + \sin x_{2m-1} \\ -x_{2m-1} + x_{2m} + \sin x_{2m} \end{pmatrix},\ x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_{2m-1} \\ x_{2m} \end{pmatrix}.$$

The behaviors of the sequences $D(x_n)$ generated by all the algorithms are described in Figures 5 and 6.

From the results are presented above, we see that Algorithm 3.1 works well. It is not surprising that Algorithm 3.10 (IMEGM) is worse than Algorithm 3.1 (IMSEGM) because it requires to compute two projections on the feasible set over each iteration. Algorithm 3.1 is developed from MSEGM in Algorithm 2 of [19] with incorporating inertial terms, and the numerical results here have illustrated that the algorithm with inertial efforts is better than the one without inertial terms.

5. Conclusions

In this paper, a new algorithm for solving a class of variational inequality problems in Hilbert spaces has been presented. The paper has described how to incorporate inertial terms into the modified subgradient extragradient method introduced recently in Algorithm 2 of [19]. The aim is to speed up the convergence property of the original algorithm. The complexity of the resulting algorithm is almost equivalent to the classical gradient projection method. It has been proved that the algorithm works well for Lipschitz and strongly pseudomonotone VIPs. One advantage of the new algorithm is that it does not assume the prior knowledge of the Lipschitz constant

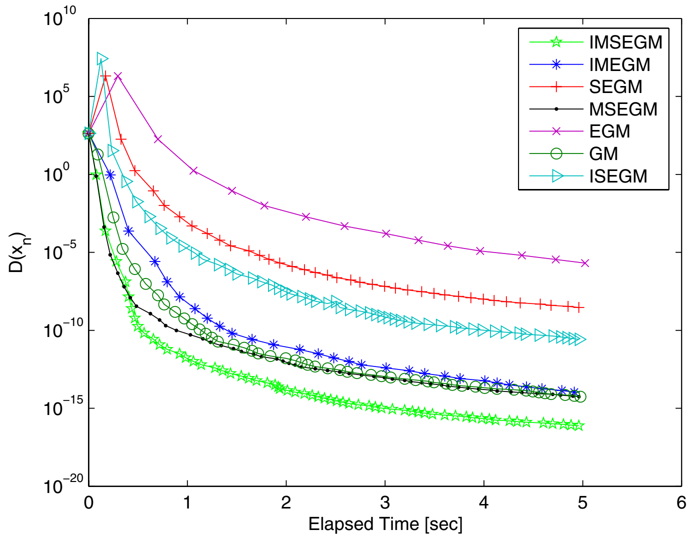

FIGURE 5. Example 3 in $\Re^{100}$. Numbers of iterations respectively are 207, 53, 134, 108, 35, 106, 145.

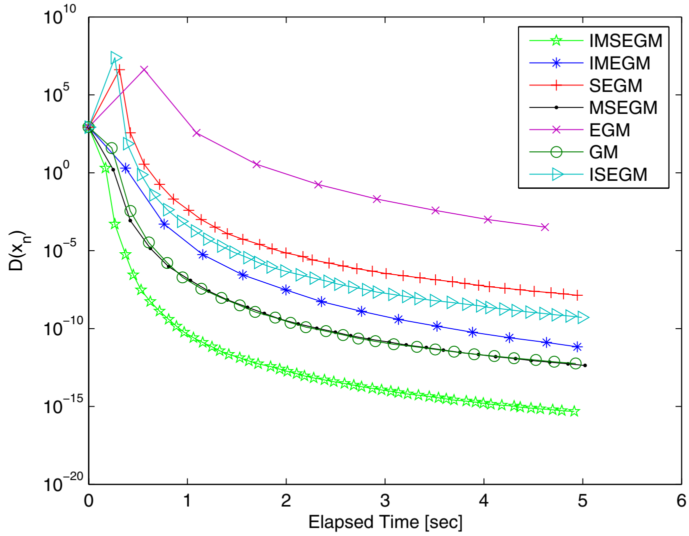

FIGURE 6. Example 3 in $\Re^{200}$. Numbers of iterations respectively are 227, 82, 234, 205, 67, 218.

and the modulus of strong pseudomonotonicity of the operator. Theoretical results have been confirmed by several numerical experiments.

Acknowledgements. We would like to thank the Associate Editor and the referees for their useful suggestions and comments that helped us very much in improving the original version of the paper.

References

[1] F. Alvarez and 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.

[2] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. *SIAM J. Optim.* **14** (2004) 773–782.

[3] E. Blum and W. Oettli, From optimization and variational inequalities to equilibrium problems. *Math. Student.* **63** (1994) 123–145.

[4] R.I. Bot, E.R. Csetnek and S.C. Laszlo, An inertial forward-backward algorithm for the minimization of the sum of two nonconvex functions. *Euro J. Comput. Optim.* **4** (2016) 3–25.

[5] R.I. Bot, E.R. Csetnek and C. Hendrich, Inertial Douglas-Rachford splitting for monotone inclusion problems. *Appl. Math. Comput.* **256** (2015) 472–487.

[6] 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.

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

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

[9] S.C. Dafermos and S.C. McKelvey, Partitionable variational inequalities with applications to network and economic equilibria. *J. Optim. Theory Appl.* **73** (1992) 243–268.

[10] S. Dafermos, Traffic equilibria and variational inequalities. *Transp. Sci.* **14** (1980) 42–54.

[11] F. Facchinei and J.S. Pang, Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, Berlin (2003).

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

[13] F. Giannessi, A. Maugeri and P.M. Pardalos, Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models. Kluwer, Dordrecht (2001).

[14] A. Gibali, S. Reich and R. Zalas, Iterative methods for solving variational inequalities in Euclidean spaces. *J. Fixed Point Theory Appl.* **17** (2015) 775–811.

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

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

[17] P. Hartman and G. Stampacchia, On some non-linear elliptic diferential-functional equations. *Acta Math.* **115** (1966) 271–310.

[18] D.V. Hieu, Convergence analysis of a new algorithm for strongly pseudomontone equilibrium problems. *Numer. Algor.* **77** (2018) 983–1001.

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

[20] D.V. Hieu, New extragradient method for a class of equilibrium problems in Hilbert spaces, *Appl. Anal.* **97** (2018) 811–824.

[21] D.V. Hieu, An inertial-like proximal algorithm for equilibrium problems. *Math. Meth. Oper. Res.* **88** (2018) 399–415.

[22] S. Karamardian and S. Schaible, Seven kinds of monotone maps. *J. Optim. Theory Appl.* **66** (1990) 37–46.

[23] G. Kassay, S. Reich and S. Sabach, Iterative methods for solving systems of variational inequalities in reflexive Banach spaces. *SIAM J. Optim.* **21** (2011) 1319–1344.

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

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

[26] I.V. Konnov, Combined Relaxation Methods for Variational Inequalities. Springer, Berlin (2000).

[27] I.V. Konnov, Equilibrium Models and Variational Inequalities. Elsevier, Amsterdam (2007).

[28] G.M. Korpelevich, The extragradient method for finding saddle points and other problems. *Ekonomikai Matematicheskie Metody.* **12** (1976) 747–756.

[29] Y.V. Malitsky, Projected reflected gradient methods for monotone variational inequalities. *SIAM J. Optim.* **25** (2015) 502–520.

[30] P.E. Maingé, Inertial iterative process for fixed points of certain quasi-nonexpansive mappings. *Set Valued Anal.* **15** (2007) 67–79.

[31] P.E. Maingé, Convergence theorems for inertial KM-type algorithms. *J. Comput.Appl. Math.* **219** (2008) 223–236.

[32] A. Moudafi, Second-order differential proximal methods for equilibrium problems. *J. Inequal. Pure Appl. Math.* **4** (2003) 18.

[33] L.D. Popov, A modification of the Arrow–Hurwicz method for searching for saddle points. *Mat. Zametki* **28** (1980) 777–784.

[34] B.T. Polyak, Some methods of speeding up the convergence of iterative methods. *Zh.Vychisl. Mat. Mat. Fiz.* **4** (1964) 1–17.

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

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

[37] P. Tseng, A modified Forward-Backward splitting method for maximal monotone mappings. S*IAM J. Control Optim.* **38** (2000) 431–446.
[38] R.U. Verma, Variational inequalities involving strongly pseudomonotone hemicontinuous mappings in nonreflexive Banach spaces. *Appl. Math. Lett.* **11** (1998) 41–43.
[39] R.U. Verma, Generalized strongly pseudomonotone nonlinear variational inequalities and general proximal point methods. *Math. Sci. Res. J.* **6** (2002) 417–427.
[40] R.U. Verma, General system of strongly pseudomonotone nonlinear variational inequalities based on projection systems. *J. Inequal. Pure Appl. Math.* **8** (2007) 6.
[41] 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.