TAIWANESE JOURNAL OF MATHEMATICS
Vol. 28, No. 2, pp. 397–421, April 2024
DOI: 10.11650/tjm/231202

Inertial Algorithms for Solving Nonmonotone Variational Inequality Problems

Bien Thanh Tuyen, Hy Duc Manh and Bui Van Dinh*

Abstract. In this paper, we will construct an inertial algorithm without using the embedded projection method to find a solution of variational inequality problems in which the cost mapping is not required to be satisfied any pseudomonotonicity. The iterative sequences generated by algorithms under the main assumption $S_M \neq \emptyset$ are proved that they converge to a solution of the corresponding problems. In addition, numerical experiments are provided to show the effectiveness of the algorithm.

1. Introduction

Let $\mathbb{R}^n$ be the Euclidean space with the inner product $\langle \cdot , \cdot \rangle$ and the induced norm $\| \cdot \|$. Let C be a nonempty closed convex set in $\mathbb{R}^n$ and F be a continuous mapping from $\mathbb{R}^n$ into $\mathbb{R}^n$. The classical variational inequality problem associated with C and F (VIP(C,F) for short) consist of finding an element $x \in C$ such that

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

We denote the set of solution of VIP(C,F) by S. The dual problem of this problem which is called Minty variational inequality problem (shortly, MVIP(C,F)) can be stated as follows: find $x \in C$ such that

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

The solution set of MVIP(C,F) is denoted by S_M. Let $L(y) = \{x \in C : \langle F(y), y - x \rangle \geq 0\}$. If $x^* \in S_M$, then

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

This means that

$$x^* \in L(y) = \{x \in C : \langle F(y), y - x \rangle \geq 0\}, \quad \forall\, y \in C.$$

Hence, $x^* \in \bigcap_{y \in C} L(y)$.

Received June 26, 2023; Accepted December 3, 2023.
Communicated by Jein-Shan Chen.
2020 *Mathematics Subject Classification.* 47J20, 49J40, 90C33.
Key words and phrases. variational inequality, nonmonotone, inertial algorithm, projection.
*Corresponding author.

On the other hand, take $\overline{x} \in \bigcap_{y \in C} L(y)$. For any fixed $y \in C$, we have $\overline{x} \in L(y)$, i.e.,

$$\langle F(y), y - \overline{x} \rangle \geq 0.$$

It follows that

$$\langle F(y), y - \overline{x} \rangle \geq 0, \quad \forall\, y \in C.$$

Then $\overline{x} \in S_M$. Consequently, $S_M = \bigcap_{y \in C} L(y)$. This means that S_M is the intersection of the closed and convex halfspaces. Then S_M is a closed and convex set. It is well-known that $S_M \subset S$ if F is continuous on C, in addition, if F is pseudomonotone on C then $S \subset S_M$ (see [9, Lemma 2.1]).

Variational inequality theory, initiated by Stampacchia [14, 25], is an important tool in operation research and mathematical physics, economics, transportation, and so on. In fact, variational inequalities provide a unifying framework for the study of such diverse problems as boundary value problems, price equilibrium problems. Due to its vast applications, variational inequality theory has become a crucial field and has been extensively studied by many researchers [1, 7, 8, 12, 18, 19] and the references quoted therein.

One of the most interesting and important problems in variational inequality theory is to build effective algorithms for finding a solution of VIPs among them the projection methods play an important role [3, 11, 17, 21, 31]. It is well-know that [16] the projection method, in general, might not be convergent for the monotone variational inequality problems. Hence, Korpelevich [20] proposed the following extragradient algorithm

$$x^0 \in C, \quad y^k = P_C(x^k - \lambda F(x^k)), \quad x^{k+1} = P_C(x^k - \lambda F(y^k)),$$

where P_C is the metric projection onto C and $\lambda \in (0, 1/L)$ (L is the Lipschitz constant of mapping F). The sequence $\{x^k\}$ was proved to converge to a solution of VIPs under the main assumption that F is pseudomonotone and Lipschitz on C. However, when F is not pseudomonotonicity or is not Lipschitz continuous, the extragradient method may not be applied to solve VIPs directly.

To find a solution of nonmonotone variational inequality problem, Ye and He [33] proposed to use the following shrinking projection algorithm:

Algorithm YH. Choose $x^0 \in C$, $\sigma \in (0,1)$ and $\gamma \in (0,1)$. Set $k = 0$.

Step 1. Having x^k. Compute $z^k := P_C(x^k - F(x^k))$ and $r(x^k) = x^k - z^k$. If $r(x^k) = 0$, Stop. Otherwise, go to Step 2.

Step 2. Compute $z^k = x^k - \eta_k r(x^k)$, where $\eta_k = \gamma^{m_k}$ with m_k being the smallest non-negative integer satisfying

$$\langle F(x^k) - F(x^k - \gamma^m r(x^k)), r(x^k) \rangle \leq \sigma \|r(x^k)\|^2.$$

Step 3. Compute $x^{k+1} = P_{C\cap\widehat{H}_k}(x^k)$, where $\widehat{H}_k := \bigcap_{j=0}^{j=k} H_j$ with $H_j := \{v : \langle F(z^j), v - y^j\rangle \le 0\}$. Let $k = k+1$ and return Step 1.

It was showed in [33] that if $S_M \neq \emptyset$ then the sequence $\{x^k\}$ converges to a solution $x^* \in S$. This method has been extended for solving nonmonotone multivalued variational inequalities [6] and equilibrium problems [10, 26]. One advantage of this algorithm is that it can be used to find a solution of VIPs when F is a nonmonotone and non-Lipschitz continuous function. However, at each iteration, it requires to solve an optimization problem in which its constraint set is the intersection of the constraint set in the previous iteration and a halfspace. Hence, the computation cost may increase, especially when n is large.

To overcome this drawback, very recently, Ye [32] have proposed the following infeasible projection type algorithm:

Algorithm IPA. Choose $x^0 \in \mathbb{R}^n$, $0 < \alpha_{\min} < \alpha_{\max}$, $\eta, \sigma \in (0,1)$, $\epsilon > 0$ as a terminate criterion. Set $k = 0$.

Step 1. Having x^k. Take $\alpha_0^k \in [\alpha_{\min}, \alpha_{\max}]$, $\alpha_k := \alpha_k^0\eta^{m_k}$, where m_k is the smallest nonnegative integer m satisfying

$$\alpha_k^0\eta^m\|F(x^k) - F(P_C(x^k - \alpha_k^0\eta^m F(x^k)))\| \le \sigma\|x^k - P_C(x^k - \alpha_k^0\eta^m F(x^k))\|.$$

Step 2. Compute $z^k := P_C(x^k - F(x^k))$, $r(x^k) = x^k - z^k$. If $r(x^k) \le \epsilon 0$, then stop; Otherwise, go to Step 3.

Step 3. Set $H_k = \{v \in \mathbb{R}^n : h_k(v) \le 0\}$ with $h_k(v) := \langle x^k - z^k - \alpha_k(F(x^k) - F(z^k)), v - z^k\rangle$. Let $t_k \in \operatorname{argmax}\{\operatorname{dist}(x^k, H_j) : 0 \le j \le k\}$ and $\widehat{H}_k := H_{t_k}$, and compute

$$x^{k+1} = P_{C\cap\widehat{H}_k}(x^k).$$

Step 4. Let $k = k+1$ and return Step 1.

The author proved that sequence $\{x^k\}$ generated by this algorithm converge to a solution of VIPs if S_M is nonempty. One advantage of this algorithm is that it can be used to find a solution of nonmonotone and non-Lipschitz VIPs without using the embedded method as in [33].

On the other hand, an inertial type algorithm was first proposed by Poljak [24] as an acceleration process in solving a smooth convex minimization problem. An inertial-type algorithm can be consider as a two-step iterative method in which the next iterate is computed by making use of the previous two iterates. It is well known that incorporating an inertial term in an algorithm may speed up the rate of convergence of the algorithm. Consequently, there are a vast of researchers have been investigating inertial type algorithms to solve their problems such as [2, 4, 22, 28].

Motivated by this fact, in this paper, we introduce some new inertial type algorithms for finding a solution of the variational inequality problem in which the cost mapping F is nonmonotone in two cases F is Lipschitz or non-Lipschitz continuous, in addition, we do not use the embedded projection methods as in papers [6, 10, 26, 33]. The rest of the paper is organized as follows. Section 2 contains some preliminaries on the metric projection and the natural residual mapping of the variational inequality problem. The proposed algorithms and their convergence are presented in Section 231202S3. Section 4 is devoted to present some numerical examples to illustrate the efficient of proposed algorithms.

2. Preliminaries

This section contains some basic results that will be used in our subsequent analysis.

Let C be a nonempty subset of n-dimensions Euclidean spaces $\mathbb{R}^n$.

Definition 2.1. [12] A mapping $F\colon C \to \mathbb{R}^n$ is said to be

(a) monotone on C if for all vectors x and y in C,

$$\langle F(x) - F(y), x - y\rangle \geq 0;$$

(b) pseudomonotone on C if for all vectors x and y in C,

$$\langle F(y), x - y\rangle \geq 0 \quad \Longrightarrow \quad \langle F(x), x - y\rangle \geq 0;$$

(c) quasimonotone on C if for all vectors x and y in C,

$$\langle F(y), x - y\rangle > 0 \quad \Longrightarrow \quad \langle F(x), x - y\rangle \geq 0.$$

The following implications are obvious from the above definition:

$$\text{(a)} \Rightarrow \text{(b)} \Rightarrow \text{(c)}.$$

Let $d(x, C)$ denote the Euclidean distance from a vector x to C, i.e.,

$$d(x, C) := \inf\{\|x - y\|, y \in C\}.$$

If C is a nonempty and closed set, then it can be seen that

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

In the rest of this paper, we consider the set C is a nonempty, closed and convex set. By P_C we denote the metric projection operator on C, that is,

$$P_C(x) \in C: \ \|x - P_C(x)\| \leq \|y - x\|, \ \ \forall\, y \in C.$$

For a fixed $x \in \mathbb{R}^n$ and $\alpha \geq 0$, we denote the natural residual mapping

$$r(x, \alpha) := x - P_C(x - \alpha F(x)). \tag{2.1}$$

The following lemmas are well-known for the projection operator onto a closed convex set and the natural residual mapping.

Lemma 2.2. [34] *Suppose that C is a nonempty closed convex subset in $\mathbb{R}^n$. Then, the following statements hold:*

(a) *$P_C(x)$ is singleton and well defined for every x;*

(b) *$z = P_C(x)$ if and only if $\langle x - z, y - z\rangle \leq 0$, $\forall\, y \in C$;*

(c) *$\|P_C(x) - P_C(y)\|^2 \leq \|x - y\|^2 - \|P_C(x) - x + y - P_C(y)\|^2$, $\forall\, x, y \in C$.*

Lemma 2.3. [15] *Let C be a nonempty closed and convex set in $\mathbb{R}^n$, $t(\cdot)$ be a real-valued function on $\mathbb{R}^n$. We define $\Omega := \{x \in C : t(x) \leq 0\}$. If Ω is nonempty and $t(\cdot)$ is Lipschitz continuous on C with modulus $L > 0$, then*

$$d(x, \Omega) \geq L^{-1} \max\{t(x), 0\}, \quad \forall\, x \in C.$$

Lemma 2.4. [29] *Let $C \subset \mathbb{R}^n$ be a nonempty closed and convex set, $r(x, \alpha)$ be defined in* (2.1)*. Then x is a solution of* $\mathrm{VIP}(C, F)$ *if and only if $\|r(x, \alpha)\| = 0$ for each $\alpha > 0$.*

Lemma 2.5. [13] *For any $x \in \mathbb{R}^n$ and $\alpha \geq 0$, let $r(x, \alpha)$ be defined in* (2.1)*. Then*

(a) *function $\alpha \mapsto \|r(x, \alpha)\|$ is nondecreasing whenever $\alpha > 0$;*

(b) *function $\alpha \mapsto \frac{\|r(x,\alpha)\|}{\alpha}$ is nonincreasing whenever $\alpha > 0$.*

Remark 2.6. From Lemma 2.5(b), we obtain that

$$\|r(x, 1)\| \begin{cases} \leq \frac{\|r(x,\alpha)\|}{\alpha} & \text{if } \alpha \leq 1, \\ \geq \frac{\|r(x,\alpha)\|}{\alpha} & \text{if } \alpha \geq 1. \end{cases}$$

This means that

$$\alpha\|r(x, 1)\| \begin{cases} \leq \|r(x, \alpha)\| & \text{if } \alpha \leq 1, \\ \geq \|r(x, \alpha)\| & \text{if } \alpha \geq 1. \end{cases}$$

Hence, we have the following inequality

$$\min\{1, \alpha\}\|r(x, 1)\| \leq \|r(x, \alpha)\| \leq \max\{1, \alpha\}\|r(x, 1)\| \quad \text{for any fixed } \alpha > 0. \tag{2.2}$$

Lemma 2.7. [30] *Let $\{a_k\}_{k=0}^{\infty}$ and $\{b_k\}_{k=0}^{\infty}$ be the nonnegative real number sequences satisfying*

$$\sum_{i=0}^{\infty} b_i < \infty \quad \textit{and} \quad a_{i+1} \le a_i + b_i, \quad \forall\, i.$$

Then, the sequence $\{a_k\}_{k=0}^{\infty}$ is convergent.

Lemma 2.8. [23] *Let $\{x^k\}_{k=0}^{\infty}$ be a sequence in $\mathbb{R}^n$ and let C be a nonempty subset of $\mathbb{R}^n$. Suppose that, for every $x \in C$, $\{\|x^k - x\|\}_{k=0}^{\infty}$ converges and that every sequential cluster point of $\{x^k\}_{k=0}^{\infty}$ belongs to C. Then $\{x^k\}_{k=0}^{\infty}$ converges to a point in C.*

3. Inertial algorithms for VIPs

Now, we assume that F is a continuous mapping from $\mathbb{R}^n$ and C is a nonempty, closed and convex subset of $\mathbb{R}^n$ such that F is not necessarily quasimonotone on C. The following algorithm give us a way to find a solution of $\mathrm{VIP}(C, F)$.

Algorithm 1.

Initialization. Let $\theta \in [0, 1)$, $\eta > 0$ and $\lambda, \delta \in (0, 1)$. Choose positive sequence $\{\mu_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} \mu_k < \infty$. Take $x^0, x^1 \in C$ and $k = 0$.

Iteration k ($k = 0, 1, 2, \ldots$). Having x^k do the following steps.

Step 1. Compute $w^k := x^k + \theta_k(x^k - x^{k-1})$, where $0 \le \theta_k \le \overline{\theta}_k$, and

$$\overline{\theta}_k = \begin{cases} \min\left\{\theta, \frac{\mu_k}{\|x^k - x^{k-1}\|}\right\} & \text{if } x^k \neq x^{k-1}, \\ \theta & \text{otherwise.} \end{cases}$$

Step 2. Find m_k, the smallest nonnegative integer m satisfying

$$\langle F(w^k) - F(y^{k,m}), w^k - y^{k,m} \rangle \le \delta \left(\frac{\|w^k - y^{k,m}\|}{\eta\lambda^m} \right)^2, \tag{3.1}$$

where

$$y^{k,m} := P_C(w^k - \eta^2\lambda^{2m} F(w^k)).$$

Set $\lambda_k := \eta^2\lambda^{2m_k}$, $z^k := y^{k,m_k}$, $r(w^k, \lambda_k) = w^k - z^k$.

Step 3. Set $T_k = \{x \in \mathbb{R}^n : t_k(x) \le 0\}$ where

$$t_k(x) := \langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), x - z^k \rangle. \tag{3.2}$$

Select

$$j_k \in \operatorname{argmax}\{d(w^k, T_j) : 1 \le j \le k\} \quad \text{and} \quad \widehat{T}_k := T_{j_k}, \tag{3.3}$$

and compute

$$x^{k+1} := P_{\widehat{T}_k}(w^k),$$

and go to Iteration k with k replaced by $k+1$.

Remark 3.1. (a) Observe that Algorithm 1 implies

$$\lim_{k\to\infty} \theta_k \|x^k - x^{k-1}\| = 0, \quad \text{and} \quad \sum_{k=0}^{\infty} \theta_k \|x^k - x^{k-1}\| < \infty.$$

(b) We remark also here that Step 1 in Algorithm 1 is easily implemented in numerical computation since $\|x^k - x^{k-1}\|$ can be computed easily to choose θ_k.

We start to analyze the convergence of the algorithm by proving the following lemmas.

Lemma 3.2. *Let F be a continuous function on $\mathbb{R}^n$. Then there exists a nonnegative integer m such that, the condition* (3.1) *is satisfied.*

Proof. If $w^k \in S$, then by Lemma 2.4, we have $\|r(w^k, \eta\lambda^m)\| = \|w^k - y^{k,m}\| = 0$. Therefore, (3.1) holds with $m = 0$. For $w^k \notin S$, we have $\|r(w^k, \eta^2\lambda^{2m})\| = \|w^k - y^{k,m}\| > 0$.

Since

$$\langle F(w^k) - F(y^{k,m}), w^k - y^{k,m}\rangle \le \|F(w^k) - F(y^{k,m})\| \|w^k - y^{k,m}\|,$$

therefore, to prove (3.1) we will prove that there exists a nonnegative m such that

$$\begin{aligned}\langle F(w^k) - F(y^{k,m}), w^k - y^{k,m}\rangle &\le \|F(w^k) - F(y^{k,m})\| \|w^k - y^{k,m}\| \\ &\le \frac{\delta}{(\eta\lambda^m)^2} \|w^k - y^{k,m}\|^2.\end{aligned}$$

This means that

$$\|F(w^k) - F(P_C(w^k - \eta^2\lambda^{2m}F(w^k)))\| \le \frac{\delta}{(\eta\lambda^m)^2} \|w^k - P_C(w^k - \eta^2\lambda^{2m}F(w^k))\|. \tag{3.4}$$

We prove by contradiction. Assume that, (3.4) is not satisfied for any m, i.e.,

$$\begin{aligned}&\|F(w^k) - F(P_C(w^k - \eta^2\lambda^{2m}F(w^k)))\| \\ &> \frac{\delta}{(\eta\lambda^m)^2} \|w^k - P_C(w^k - \eta^2\lambda^{2m}F(w^k))\| \quad \text{for all } m.\end{aligned} \tag{3.5}$$

If $w^k \in C$, then $w^k = P_C(w^k)$. Using the continuity of $F(\cdot)$ and $P_C(\cdot)$ we have

$$\lim_{m\to\infty} \|F(w^k) - F(P_C(w^k - \eta^2\lambda^{2m}F(w^k)))\| = 0. \tag{3.6}$$

On the other hand, using Lemma 2.5, $\lambda \in (0,1)$ and $w^k \notin S$, we get

$$\delta \frac{\|w^k - P_C(w^k - \eta^2\lambda^{2m}F(w^k))\|}{(\eta\lambda^m)^2} = \delta \frac{\|r(w^k, \eta^2\lambda^{2m})\|}{(\eta\lambda^m)^2} \ge \delta \frac{\|r(w^k, 1)\|}{1} > 0. \tag{3.7}$$

Thus, (3.6) and (3.7) contradict (3.5).

For $w^k \notin C$, we have

$$\lim_{m\to\infty} \|w^k - y^{k,m}\| = \lim_{m\to\infty} \|w^k - P_C(w^k - \eta^2\lambda^{2m}F(w^k))\| = \|w^k - P_C(w^k)\| > 0$$

and

$$\begin{aligned}&\lim_{m\to\infty} \frac{\eta^2\lambda^{2m}}{\delta}\|F(w^k) - F(y^{k,m})\|\\ =&\lim_{m\to\infty} \frac{\eta^2\lambda^{2m}}{\delta}\|F(w^k) - F(P_C(w^k - \eta^2\lambda^{2m}F(w^k)))\| = 0.\end{aligned}$$

We get a contradiction with (3.5). Consequently, (3.4) holds and the linesearch (3.1) is well-defined. □

We proceed to prove the following lemma before proving the convergence of our Algorithm 1.

Lemma 3.3. *Let the solution set S_M of the Minty variational inequality problem be nonempty, $x^* \in S_M$ and $t_k(\cdot)$ be defined in* (3.2). *Then the following statements hold:*

(a) $t_k(x^*) \le 0$ *and* $S_M \subset \bigcap_{k=1}^{\infty} T_k$.

(b) $t_k(w^k) \ge (1-\delta)\|r(w^k, \lambda_k)\|^2$. *In particular, if* $w^k \ne z^k$, *then* $t_k(w^k) > 0$ *and* $w^k \notin T_k$, $\forall\, k$.

Proof. (a) Since $x^* \in S_M$, then $\langle F(z^k), z^k - x^*\rangle \ge 0$. From (3.2), we have

$$\begin{aligned}t_k(x^*) &= \langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), x^* - z^k\rangle\\ &= \langle w^k - z^k - \lambda_k F(w^k), x^* - z^k\rangle + \lambda_k\langle F(z^k), x^* - z^k\rangle\\ &\le \langle w^k - z^k - \lambda_k F(w^k), x^* - z^k\rangle\\ &= \langle w^k - \lambda_k F(w^k) - P_C(w^k - \lambda_k F(w^k)), x^* - P_C(w^k - \lambda_k F(w^k))\rangle \le 0.\end{aligned}$$

Therefore, $x^* \in T_k$ for all k. This implies that $S_M \subset T_k$, $\forall\, k$. This means that $S_M \subset \bigcap_{k=1}^{\infty} T_k$.

(b) Similarly, using (3.1), we get

$$\begin{aligned}t_k(w^k) &= \langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), w^k - z^k\rangle\\ &= \|w^k - z^k\|^2 - \lambda_k\langle F(w^k) - F(z^k), w^k - z^k\rangle\\ &\ge \|w^k - z^k\|^2 - \delta\|w^k - z^k\|^2\\ &\ge (1-\delta)\|r(w^k, \lambda_k)\|^2 \ge 0.\end{aligned}$$

If $w^k \ne z^k$ then $t_k(w^k) > 0$. Thus, $w^k \notin T_k$. □

The following theorem establishes the convergence of Algorithm 1.

Theorem 3.4. *Suppose that F is a continuous function on $\mathbb{R}^n$ and S_M is nonempty. Then $\{x^k\}$ is generated by Algorithm 1 converges to a solution of $\mathrm{VIP}(C, F)$.*

Proof. We break our proof into four steps.

Step (i). We first show that the sequence $\{x^k\}$ is bounded. Indeed, for any fixed point $p \in \bigcap_{i=1}^{\infty} T_i$, it is clear that $p \in \widehat{T}_k$. Therefore, we have

$$\begin{aligned}\|x^{k+1} - p\| &= \|P_{\widehat{T}_k}(w^k) - P_{\widehat{T}_k}(p)\| \leq \|w^k - p\| \\ &= \|x^k + \theta_k(x^k - x^{k-1}) - p\| \leq \|x^k - p\| + \theta_k\|x^k - x^{k-1}\|.\end{aligned} \tag{3.8}$$

Using Lemma 2.7 with $a_k = \|x^k - p\|$, $b_k = \theta_k\|x^k - x^{k-1}\|$, we obtain that, for any fixed $p \in \bigcap_{i=1}^{\infty} T_i$, the sequence $\{\|x^k - p\|\}$ is convergent. Thus, we can deduce that the sequence $\{x^k\}$ is bounded.

In addition, we have

$$\|w^k\| = \|x^k + \theta_k(x^k - x^{k-1})\| \leq \|x^k\| + \theta_k\|x^k - x^{k-1}\|.$$

Beside that $\lim_{k\to\infty} \theta_k\|x^k - x^{k-1}\| = 0$ by Remark 3.1. This implies that $\{w^k\}$ is bounded. Because F is continuous on $\mathbb{R}^n$, then $\{F(w^k)\}$ is bounded. Combining this with definitions of z^k and the continuity of $P_C(\cdot)$, we further obtain that $\{z^k\}$, $\{F(z^k)\}$ and $\{w^k - z^k - \lambda_k(F(w^k) - F(z^k))\}$ are also bounded.

Step (ii). We show that any cluster point x^* of the sequence $\{x^k\}$ belongs to T_k for all k. Since $w^k \notin T_k$ for all k, then $d(w^k, T_k) > 0$. Combining with Lemma 2.2, we get

$$\begin{aligned}d^2(w^k, \widehat{T}_k) &= \|w^k - P_{\widehat{T}_k}(w^k)\|^2 = \|w^k - x^{k+1}\|^2 \leq \|w^k - p\|^2 - \|x^{k+1} - p\|^2 \\ &= \|x^k + \theta_k(x^k - x^{k-1}) - p\|^2 - \|x^{k+1} - p\|^2 \\ &= \|x^k - p\|^2 - \|x^{k+1} - p\|^2 + \theta_k^2\|x^k - x^{k-1}\|^2 + 2\theta_k\langle x^k - x^{k-1}, x^k - p\rangle.\end{aligned}$$

Hence,

$$\begin{aligned}0 &< d^2(w^k, T_k) \leq d^2(w^k, \widehat{T}_k) \\ &\leq \|x^k - p\|^2 - \|x^{k+1} - p\|^2 + \theta_k^2\|x^k - x^{k-1}\|^2 + 2\theta_k\langle x^k - x^{k-1}, x^k - p\rangle.\end{aligned} \tag{3.9}$$

Combining with Remark 3.1, we get in the limit from (3.9) that

$$\lim_{k\to\infty} d(w^k, T_k) = \lim_{k\to\infty} d(w^k, \widehat{T}_k) = 0. \tag{3.10}$$

Assume that x^* is a cluster point of $\{x^k\}$. Then there exists $\{x^{k_l}\} \subset \{x^k\}$ such that $\lim_{l\to\infty} x^{k_l} = x^*$. From the definition of w^k, we get

$$\|w^k - x^k\| = \theta_k\|x^k - x^{k-1}\| \to 0 \quad \text{as } k \to \infty.$$

Since $x^{k_l} \to x^*$, we have $w^{k_l} \to x^*$. By (3.2) and (3.3), we see that

$$0 \le d(w^k, T_i) \le d(w^k, \widehat{T}_k) \quad \text{for all } i \le k.$$

This together with (3.10) implies that

$$\lim_{k\to\infty} d(w^k, T_i) = 0 \quad \text{for any fixed } i.$$

Using the continuity of $d(\,\cdot\,, T_i)$ on $\mathbb{R}^n$, we obtain that

$$d(x^*, T_i) = 0 \quad \text{for any fixed } i.$$

Therefore, $x^* \in \bigcap_{i=1}^\infty T_i$.

Step (iii). We show that $x^* \in S$. The boundedness of the sequence $\{w^k - z^k - \lambda_k(F(w^k) - F(z^k))\}$ implies that there exists $M > 0$ such that

$$\|w^k - z^k - \lambda_k(F(w^k) - F(z^k))\| \le M, \quad \forall\, k \in \mathbb{N}.$$

On the other hand, we have

$$\begin{aligned} \|t_k(x) - t_k(y)\| &= \|\langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), x - y\rangle\| \\ &\le \|w^k - z^k - \lambda_k(F(w^k) - F(z^k))\| \|x - y\| \\ &\le M\|x - y\|. \end{aligned}$$

Using Lemmas 2.3 and 3.3, we get

$$d(w^k, T_k) \ge M^{-1} t_k(w^k) \ge M^{-1}(1-\delta)\|r(w^k, \lambda_k)\|^2 > 0.$$

This with (3.10) establishes $\lim_{k\to\infty} \|r(w^k, \lambda_k)\| = 0$.

We next show that $\lim_{k\to\infty} \|r(w^k, 1)\| = 0$. To do so, let us consider two distinct cases.

Case 1. There exists $\widetilde{\lambda} > 0$ such that $\widetilde{\lambda} \le \lambda_i$ for all i. Using (2.2), we get

$$0 \le \|r(w^i, 1)\| \le \frac{\|r(w^i, \lambda_i)\|}{\min\{\lambda_i, 1\}} \le \frac{\|r(w^i, \lambda_i)\|}{\min\{\widetilde{\lambda}, 1\}}.$$

Hence, $\lim_{i\to\infty} \|r(w^i, \lambda_i)\| = 0$. Thus, $\lim_{i\to\infty} \|r(w^i, 1)\| = 0$. Consequently,

$$\lim_{l\to\infty} \|r(w^{k_l}, 1)\| = 0.$$

Case 2. $\lim_{l\to\infty} \lambda_{k_l} = 0$. This means that $\lim_{l\to\infty} \eta^2 \lambda^{2m_{k_l}} = 0$. By the linesearch rule (3.1), for $m_{k_l} - 1$, we have

$$\langle F(w^{k_l}) - F(y^{k_l, m_{k_l}-1}), w^{k_l} - y^{k_l, m_{k_l}-1}\rangle > \delta \left(\frac{\|w^{k_l} - y^{k_l, m_{k_l}-1}\|}{\eta \lambda^{m_{k_l}-1}} \right)^2.$$

Using the Cauchy–Schwartz inequality, we have

$$\|F(w^{k_l}) - F(y^{k_l,m_{k_l}-1})\| > \delta \frac{\|w^{k_l} - y^{k_l,m_{k_l}-1}\|}{(\eta\lambda^{m_{k_l}-1})^2}.$$

Combining with Lemma 2.5, we get

$$\|F(w^{k_l}) - F(y^{k_l,m_{k_l}-1})\| > \delta \frac{\|w^{k_l} - y^{k_l,m_{k_l}-1}\|}{(\eta\lambda^{m_{k_l}-1})^2} \geq \delta \frac{\|r(w^{k_l},1)\|}{1} > 0. \tag{3.11}$$

Furthermore, $P_C(\cdot)$ is nonexpansive, then

$$\begin{aligned}
&\|w^{k_l} - y^{k_l,m_{k_l}-1}\| \\
&\leq \|w^{k_l} - y^{k_l,m_{k_l}}\| + \|y^{k_l,m_{k_l}} - y^{k_l,m_{k_l}-1}\| \\
&= \|r(w^{k_l},\lambda_{k_l})\| + \|P_C(w^{k_l} - \eta^2\lambda^{2m_{k_l}}F(w^{k_l})) - P_C(w^{k_l} - \eta^2\lambda^{2(m_{k_l}-1)}F(w^{k_l}))\| \\
&\leq \|r(w^{k_l},\lambda_{k_l})\| + \eta^2\lambda^{2m_{k_l}}(\lambda^{-2} - 1)\|F(w^{k_l})\| \to 0 \quad \text{as } k \to \infty.
\end{aligned}$$

Using the above inequality and the fact that $\lim_{l\to\infty} w^{k_l} = x^*$, we get $\lim_{l\to\infty} y^{k_l,m_{k_l}-1} = x^*$. Combining this with continuity of F, we obtain that

$$\lim_{l\to\infty} \|F(w^{k_l}) - F(y^{k_l,m_{k_l}-1})\| = 0.$$

This together with (3.11) yields $\lim_{l\to\infty} \|r(w^{k_l},1)\| = 0$. Using the continuity of $\|r(\,\cdot\,,1)\|$, we have $\|r(x^*,1)\| = 0$. Combining with Lemma 2.4, it is immediate that $\{x^*\}$ is a solution of $\mathrm{VIP}(C,F)$.

Step (iv). We prove that $\{x^k\}$ is globally convergent to a solution of $\mathrm{VIP}(C,F)$. Since x^* is a cluster point of $\{x^k\}$, we have $x^* \in \bigcap_{i=1}^{\infty} T_i$. From (3.8), it is clear that

$$\|x^{k+1} - x^*\| \leq \|x^k - x^*\| + \theta_k\|x^k - x^{k-1}\|.$$

Using Lemma 2.7, we get that the sequence $\{\|x^k - x^*\|\}$ is convergent. Hence, there exits $a \geq 0$ such that $\lim_{k\to\infty} \|x^k - x^*\| = a$.

By the definition of x^*, there exists a subsequence $\{x^{k_l}\} \subset \{x^k\}$ such that $\lim_{l\to\infty} \|x^{k_l} - x^*\| = 0$. Using Lemma 2.8, we have

$$a = 0 \quad \text{and} \quad \lim_{k\to\infty} \|x^k - x^*\| = \lim_{l\to\infty} \|x^{k_l} - x^*\| = 0,$$

i.e., $\lim_{k\to\infty} x^k = x^*$. Consequently, $\{x^k\}$ is globally convergent to a solution of VIPs. $\square$

Next, we propose a modification of Algorithm 1 with another linesearch.

Algorithm 2.

Initialization. Choose $\theta \in [0,1)$, $\eta > 0$ and $\lambda, \delta \in (0,1)$. Choose positive sequence $\{\mu_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} \mu_k < \infty$. Take $x^0, x^1 \in C$ and $k = 0$.

Iteration k ($k = 0, 1, 2, \ldots$). Having x^k do the following steps.

Step 1. Compute $w^k := x^k + \theta_k(x^k - x^{k-1})$, where $0 \le \theta_k \le \overline{\theta}_k$, and

$$\overline{\theta}_k = \begin{cases} \min\left\{\theta, \frac{\mu_k}{\|x^k - x^{k-1}\|}\right\} & \text{if } x^k \neq x^{k-1}, \\ \theta & \text{otherwise.} \end{cases}$$

Step 2. Find m_k, the smallest nonnegative integer m satisfying

$$\eta\lambda^m \|F(w^k) - F(y^{k,m})\| \le \delta\|w^k - y^{k,m}\|, \tag{3.12}$$

where

$$y^{k,m} := P_C(w^k - \eta\lambda^m F(w^k)).$$

Set $\lambda_k := \eta\lambda^{m_k}$, $z^k := y^{k,m_k}$, $r(w^k, \lambda_k) = w^k - z^k$.

Step 3. Set $T_k = \{x \in \mathbb{R}^n : t_k(x) \le 0\}$ where

$$t_k(x) := \langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), x - z^k\rangle.$$

Select

$$j_k \in \operatorname{argmax}\{d(w^k, T_j) : 1 \le j \le k\} \quad \text{and} \quad \widehat{T}_k := T_{j_k},$$

and compute

$$x^{k+1} := P_{\widehat{T}_k}(w^k),$$

and go to Iteration k with k replaced by $k + 1$.

Theorem 3.5. *Suppose that F is a continuous function on $\mathbb{R}^n$ and $S_M \neq \emptyset$. Then $\{x^k\}$, generated by Algorithm* 2*, converges to a solution of* VIP(C, F)*.*

Proof. From (3.12), we have

$$\|F(w^k) - F(y^{k,m})\| \le \frac{\delta}{\eta\lambda^m}\|w^k - y^{k,m}\|.$$

Using this inequality and the same idea as in the proof of Lemma 3.2, we can prove that the linesearch rule (3.12) is well-defined.

Taking $x^* \in S_M$, similar to Lemma 3.3, we have $t_k(x^*) \le 0$ and $S_M \subset \bigcap_{k=1}^{\infty} T_k$. Moreover, using (3.12) we obtain

$$\begin{aligned} t_k(w^k) &= \langle w^k - z^k - \lambda_k(F(w^k) - F(z^k)), w^k - z^k\rangle \\ &= \|w^k - z^k\|^2 - \lambda_k\langle F(w^k) - F(z^k), w^k - z^k\rangle \\ &\ge \|w^k - z^k\|^2 - \lambda_k\|F(w^k) - F(z^k)\|\|w^k - z^k\| \\ &\ge \|w^k - z^k\|^2 - \delta\|w^k - z^k\|^2 \\ &\ge (1 - \delta)\|r(w^k, \lambda_k)\|^2 > 0. \end{aligned}$$

Therefore, $w^k \notin T_k$ for all k.

The proof of convergence can be done by the similar way as in Theorem 3.4, so we omit it. For instance, instead of using (3.11), we use the following inequality

$$\eta\lambda^{m_{k_l}-1}\|F(w^{k_l}) - F(y^{k_l,m_{k_l}-1})\| > \delta\|w^{k_l} - y^{k_l,m_{k_l}-1}\|.$$

It follows that

$$\|F(w^{k_l}) - F(y^{k_l,m_{k_l}-1})\| > \frac{\delta\|w^{k_l} - y^{k_l,m_{k_l}-1}\|}{\eta\lambda^{m_{k_l}-1}} \geq \frac{\delta\|r(w^{k_l},1)\|}{1} > 0.$$

The proof is completed. □

Remark 3.6. The main difference between Algorithms 1 and 2 is that in Step 2 of Algorithm 2, the linesearch (3.1) of Algorithm 1 is replaced by the linesearch (3.12). Indeed, in Algorithm 1 we need to calculate the scalar products, while, in Algorithm 2 we need to calculate the norms. In general, the cost of scalar product computation may be cheaper than the cost of norm computation.

Moreover, using Cauchy–Schwarz inequality, from (3.12) we have

$$\begin{aligned}\langle F(w^k) - F(y^{k,m_k}), w^k - y^{k,m_k}\rangle &\leq \|F(w^k) - F(y^{k,m_k})\|\|w^k - y^{k,m_k}\| \\ &\leq \delta\frac{\|w^k - y^{k,m_k}\|^2}{\eta\lambda^{m_k}} = \delta\frac{\|w^k - y^{k,m_k}\|^2}{\lambda_k}.\end{aligned}$$

Hence,

$$\langle F(w^k) - F(y^{k,m_k}), w^k - y^{k,m_k}\rangle \leq \delta\frac{\|w^k - y^{k,m_k}\|^2}{\lambda_k}.$$

Therefore, if we choose η, λ in Algorithm 2 equal to the square of η, λ respectively in Algorithm 1, then m_k that satisfies linesearch (3.12) also satisfies linesearch (3.1). This implies that m_k in Algorithm 2 is smaller than that m_k in Algorithm 1.

One advantage of Algorithms 1 and 2 is that they can be used to find a solution of nonmonotone and non-Lipschitz variational inequality problem without using embedding methods. However, in Step 2 of these algorithms, we have to do a linesearch procedure to determine the step length. In some cases, this may lead to the large cost of computation, especially, when n is large and F has a complicated form. The algorithm below shows that if F satisfies the Lipschitz condition, then the linesearch procedure in Step 2 of Algorithm 1 and Algorithm 2 can be removed.

Algorithm 3.

Initialization. Choose $\theta \in [0,1)$, $0 < \alpha < 1/L$. Choose positive sequence $\{\mu_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty}\mu_k < \infty$. Take $x^0, x^1 \in C$ and $k = 0$.

Iteration k ($k = 0, 1, 2, \ldots$). Having x^k do the following steps.

Step 1. Compute $z^k := P_C(w^k - \alpha F(w^k))$, where $w^k := x^k + \theta_k(x^k - x^{k-1})$, $0 \le \theta_k \le \overline{\theta}_k$, and

$$\overline{\theta}_k = \begin{cases} \min\left\{\theta, \frac{\mu_k}{\|x^k - x^{k-1}\|}\right\} & \text{if } x^k \neq x^{k-1}, \\ \theta & \text{otherwise.} \end{cases}$$

Step 2. Set $T_k = \{x \in \mathbb{R}^n : t_k(x) \le 0\}$, where

$$t_k(x) := \langle w^k - z^k - \alpha(F(w^k) - F(z^k)), x - z^k \rangle.$$

Select

$$j_k \in \operatorname{argmax}\{d(w^k, T_j) : 1 \le j \le k\} \quad \text{and} \quad \widehat{T}_k := T_{j_k},$$

and compute

$$x^{k+1} := P_{\widehat{T}_k}(w^k),$$

and go to Iteration k with k replaced by $k + 1$.

In this next theorem, we establish the convergence analysis of the sequence of iterates generated by Algorithm 3 to a solution of VIP(C, F).

Theorem 3.7. *Let F be L-Lipschitz continuous function on $\mathbb{R}^n$ and $S_M \neq \emptyset$. Then the sequence $\{x^k\}$ is generated by Algorithm* 3 *converges to a solution of* VIP(C, F).

Proof. Following the same reasoning as in Lemma 3.3, we obtain that $t_k(x^*) \le 0$, $\forall x^* \in S_M$ and $S_M \subset T_k$ for all k. Furthermore, since F is L-Lipschitz continuous on $\mathbb{R}^n$ and $0 < \alpha < 1/L$, we have

$$\begin{aligned} t_k(w^k) &= \langle w^k - z^k - \alpha(F(w^k) - F(z^k)), w^k - z^k \rangle \\ &= \|w^k - z^k\|^2 - \alpha\langle F(w^k) - F(z^k), w^k - z^k \rangle \\ &\ge \|w^k - z^k\|^2 - \alpha L \|w^k - z^k\|^2 \\ &\ge (1 - \alpha L)\|r(w^k, \alpha)\|^2 > 0. \end{aligned}$$

Hence, $w^k \notin T_k$ for all k. The rest of the proof is similar to that of Theorem 3.4 with $\delta = \alpha L$. Therefore, $\{x^k\}$ converges to a solution x^* of VIP(C, F). $\square$

Remark 3.8. One advantage of Algorithm 3 is that it can be applied for finding a solution of the nonmonotone and Lipschitz variational inequality problem without doing the line-search procedure. This can reduce the computation cost, especially, F has a complicated form. However, Algorithm 3 requires knowing the Lipschitz constant L of F.

4. Numerical examples

In this section, we provide five numerical examples to show the practicability and the advantage of proposed algorithms by comparing them with the embedded projection Algorithm of Ye and He [33] and Algorithm IPA of Ye [32].

All the programs are written in Mathlab R2015 and performed on a Laptop DELL Intel (R), Core (TM) i7-9700 CPU, 3.00 Ghz, Ram 16.0 GB.

In IPA (see [33]), they take $\alpha_{\min} = 10^{-10}$, $\alpha_{\max} = 10^{10}$. Inspired by the renowned Barzilai–Borwein step-size and self-adaptive step-size, they take parameter α_k^0 as follows: set $\alpha_0^0 = 1$ and take, for $k \geq 1$,

$$\alpha_k^0 = \begin{cases} P_{[10^{-10},10^{10}]}\left(\frac{\|x^k-x^{k-1}\|^2}{\langle x^k-x^{k-1}, F(x^k)-F(x^{k-1})\rangle}\right) & \text{if } \langle x^k - x^{k-1}, F(x^k) - F(x^{k-1})\rangle > 10^{-12}, \\ P_{[10^{-10},10^{10}]}(1.5\alpha_{k-1}) & \text{otherwise.} \end{cases}$$

Example 4.1. Consider a problem of the form (see [5, Exercise 4.7])

$$\min\left\{g(x) = \frac{f_0(x)}{cx+d} : x \in C\right\}$$

with $c \in \mathbb{R}^n$, $d \in \mathbb{R}$ and $C = \{x \in \mathbb{R}^n : f_i(x) \leq 0, i = 1, \ldots, m, Ax = b\}$, where $f_0, f_1, \ldots, f_m$ are convex functions, A is a matrix of order $m \times n$, $b \in \mathbb{R}^m$.

Table 4.1: Experiment for Example 4.1.

x^0	a	Alg. YH		Alg. IPA		Alg. 1		Alg. 2	
		CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.
$(0,0,0,0,5)^T$	5	0.55	35	15.47	24	0.41	32	0.44	32
$(2,1,0,0,2)^T$	-	0.45	31	11.31	30	0.31	30	0.36	30
$(1.5,1.2,1.3,0.3,0.7)^T$	-	0.44	29	10.20	19	0.3	28	0.29	28
$(5,0,0,0,5)^T$	10	0.88	70	13.19	23	0.59	66	0.51	66
$(1,3,2,3,1)^T$	-	0.72	57	11.72	21	0.50	58	0.46	58
$(1.7,1.8,1.9,3.5,1.1)^T$	-	0.72	56	10.86	20	0.50	57	0.47	57

This is a quasiconvex optimization problem. Choose $g(x) = \frac{\frac{1}{2}x^T Hx + q^T x + r}{\sum_{i=1}^5 x_i}$ and $C = \left\{x = (x_1, x_2, \ldots, x_5)^T \in \mathbb{R}^5 : x_i \geq 0, i = 1, \ldots, 5, \sum_{i=1}^5 x_i = a\right\}$, where $a > 0$. Then g is a smooth quasiconvex function and can attain its minimum value on C, where $q = (-1, \ldots, -1)^T$, $r = 1$ and $H = hI$ is a positive diagonal matrix with $h \in (0.1, 1.6)$. Let $F(x) = (F_1(x), \ldots, F_5(x))^T$ be the derivative of $g(x)$ with $F_i(x) = \frac{hx_i \sum_{j=1}^5 x_j - \frac{1}{2}h\sum_{j=1}^5 x_j^2 - 1}{\left(\sum_{j=1}^5 x_j\right)^2}$. Then the problem VIP$(C, F)$ is a quasimonotone variational inequality with $S_M = \big\{\big(\frac{1}{5}a,$

$\ldots, \frac{1}{5}a)^T\}$ (see [33]). We choose $h = 1.2$, $\sigma = 0.4$, $\gamma = 0.99$ in Algorithm YH, $\sigma = 0.4$, $\eta = 0.99$ in Algorithm IPA, $\theta = 0.1$, $\eta = 0.99$, $\lambda = 0.99$, $\delta = 0.8$, $\mu_k = \frac{1}{(k+1)^{1.8}}$ in Algorithms 1 and 2. The initial point is x^0 in Algorithm YH, Algorithm IPA, and $x^0 = x^1$ in Algorithms 1 and 2. We make comparison of algorithms with $\epsilon = \|r(x^k)\| \leq 10^{-4}$ and report the result in Table 4.1 and Figure 4.1.

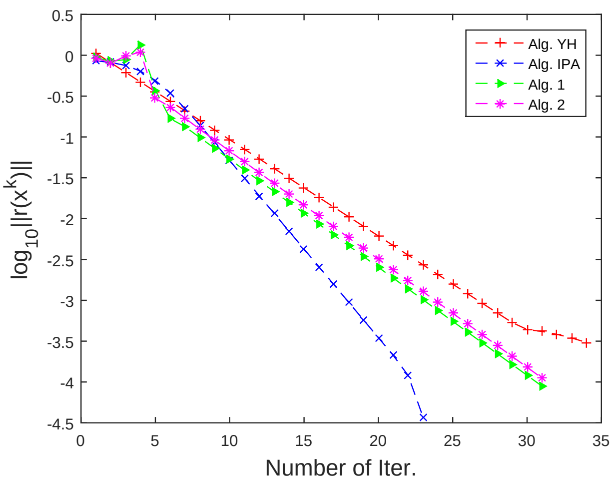

Figure 4.1: Number of iterations in Example 4.1.

Example 4.2. In this example, we consider the affine variational inequality problem

$$C = [0,1]^n \quad \text{and} \quad F(x) = Mx + d,$$

where M is an $n \times n$ tridiagonal matrix given by

$$M = \begin{bmatrix} 4 & -2 & & & \\ 1 & 4 & -2 & & \\ & 1 & 4 & -2 & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ & & & 1 & 4 \end{bmatrix}$$

and $d = (-1, \ldots, -1)^T$ (see [1, 27, 33]).

We take $\sigma = 0.4$, $\gamma = 0.1$ in Algorithm YH, $\sigma = 0.5$, $\eta = 0.99$ in Algorithm IPA. And take $\theta = 0.5$, $\lambda = 0.6$, $\delta = 0.4$, $\eta = 0.9$, $\mu_k = \frac{1}{(k+2)^{1.8}}$ in Algorithm 1, $\lambda = 0.1$, $\delta = 0.5$, $\theta = 0.2$, $\eta = 0.99$, $\mu_k = \frac{1}{(k+1)^{1.5}}$ in Algorithm 2, $\mu_k = \frac{1}{(k+2)^{1.5}}$, $L = \sqrt{(21n-5)}$, $\theta = 0.5$, $\alpha = \frac{1-\sigma}{L}$, where $\sigma = 0.01$ in Algorithm 3. The starting point is $x^0 = (0, \ldots, 0)^T$ in Algorithm YH, Algorithm IPA, $x^0 = x^1 = (0, \ldots, 0)^T$ in Algorithms 1, 2 and 3. We terminate the iteration if $\epsilon = \|r(x^k)\| \le 10^{-4}$. The results are shown in Table 4.2 and Figure 4.2.

Table 4.2: Experiment for Example 4.2.

n	Alg. YH		Alg. IPA		Alg. 1		Alg. 2		Alg. 3	
	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.
50	24.95	456	9.69	18	0.17	22	0.09	21	0.45	44
100	659.41	894	10.80	18	0.23	23	0.16	22	0.75	71
150	3556.47	1773	52.09	18	1.61	23	0.87	22	3.98	94
200	3986.81	1628	63.66	18	1.95	23	1.09	23	9.02	113
500	50530.97	3889	269.20	18	6.95	21	10.25	33	67.5	202

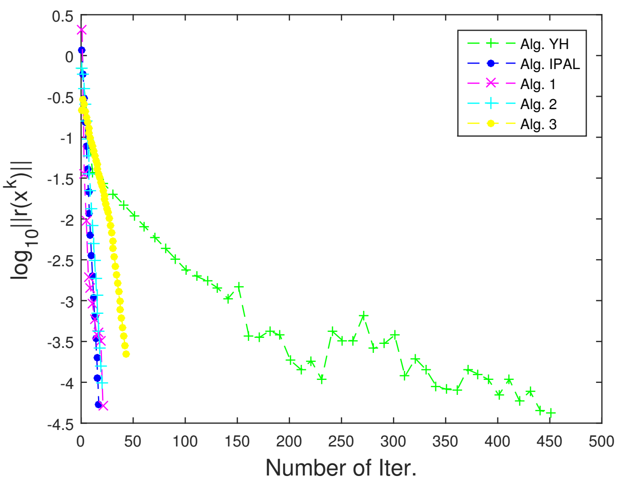

Figure 4.2: Number of iterations of Algorithms in Example 4.2 ($n = 50$, $x^0 = x^1 = (0, \dots, 0)^T$).

Example 4.3. In this example, we let

$$C = [-1, 1]^n \quad \text{and} \quad F(x) = (x_1^2, \dots, x_n^2)^T.$$

Here, the mapping is not quasimonotone but it is easy to check that $S_M = \{(-1, \dots, -1)^T\}$ (see [33]). Furthermore, it can be see that F satisfies the Lipschitz condition with $L = 2\sqrt{n}$. We use Algorithm YH with parameters $\sigma = 0.5$, $\gamma = 0.99$, Algorithm IPA with $\sigma = 0.5$, $\eta = 0.99$, Algorithm 1 with $\theta = 0.8$, $\eta = 0.99$, $\lambda = 0.99$, $\delta = 0.4$, $\mu_k = \frac{1}{(k+2)^{1.3}}$ and Algorithm 2 with $\theta = 0.5$, $\eta = 0.99$, $\lambda = 0.99$, $\delta = 0.4$, $\mu_k = \frac{1}{(k+2)^{1.3}}$.

The starting point is $x_0 = (-3/4, \ldots, -3/4)^T$ in Algorithm YH, Algorithm IPA and $x^0 = x^1 = (-3/4, \ldots, -3/4)^T$ in Algorithms 1 and 2. The numerical result is shown in Table 4.3 and Figure 4.3. The stopping criterion used is $\epsilon = \|r(x^k)\| \leq 10^{-4}$.

Table 4.3: Experiment for Example 4.3.

n	Alg. YH		Alg. IPA		Alg. 1		Alg. 2	
	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.
100	0.27	5	2.17	4	0.25	4	0.15	4
500	1.11	5	30.20	5	0.59	4	0.50	4
1000	1.78	5	61.88	5	0.86	4	0.72	4
5000	10.92	5	399.41	5	5.72	5	5.50	5
10000	33.77	5	1180.38	5	12.81	4	12.16	4

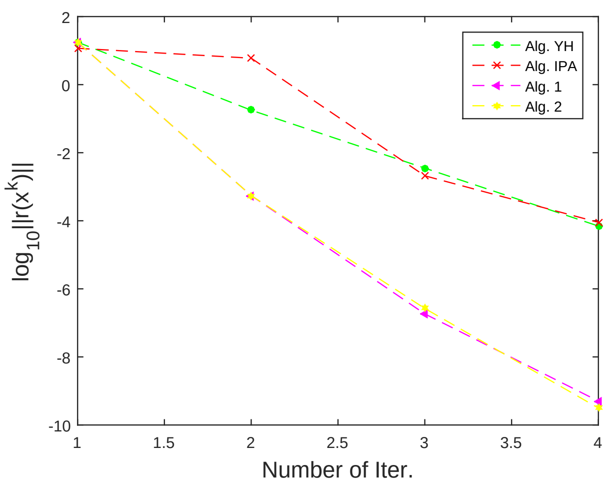

Figure 4.3: Number of iterations of Algorithms in Example 4.3 ($n = 5000$, $x^0 = x^1 = (-3/4, -3/4, -3/4, \ldots, -3/4)^T$).

Example 4.4. Let $C = [0,1]^n$ and $F(x) = (x_1^2 - x_1, \ldots, x_n^2 - x_n)^T$. Here, the mapping is not quasimonotone and $S_M = \{(1, \ldots, 1)^T\}$ (see [33]). We take $\sigma = 0.95$, $\gamma = 0.99$ in Algorithm YH, $\sigma = 0.95$, $\eta = 0.99$ in Algorithm IPA, $\theta = 0.1$, $\eta = 0.99$, $\lambda = 0.99$, $\delta = 0.99$, $\mu_k = \frac{1}{(k+2)^{1.7}}$ in Algorithm 1 and $\theta = 0.9$, $\eta = 0.8$, $\lambda = 0.9$, $\delta = 0.9$, $\mu_k = \frac{1}{(k+1)^3}$ in Algorithm 2. Let $x^0 = (1/6, \ldots, 1/6)^T$ be initial point in Algorithm YH, Algorithm IPA, $x^0 = x^1 = (1/6, \ldots, 1/6)^T$ be initial points in Algorithms 1 and 2. To terminate the Algorithms, we use the stopping criteria $\epsilon = \|r(x^k)\| \leq 10^{-4}$. The results are listed in

Table 4.4 and the corresponding figures are displayed in Figure 4.4.

Table 4.4: Experiment for Example 4.4.

n	Alg. YH		Alg. IPA		Alg. 1		Alg. 2	
	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.
100	0.31	9	1.62	7	0.23	9	0.23	12
500	2.59	10	19.07	7	1.37	10	1.82	13
1000	4.05	9	39.13	7	2.34	10	5.06	14
5000	27.82	10	253.15	8	13.25	10	77.12	17
10000	82.00	10	705.14	8	36.73	10	316.98	19

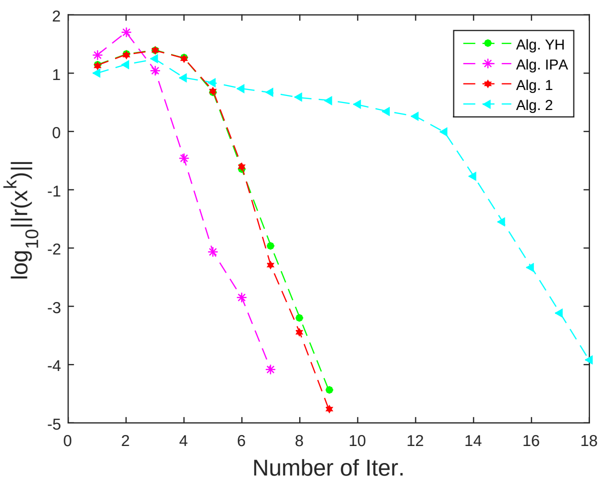

Figure 4.4: Number of iterations of Algorithms in Example 4.4 ($n = 10000$, $x^0 = (1/6, 1/6, 1/6, \ldots, 1/6)^T$).

In the last example, we present to numerical experiments to illustrate the performance of algorithms when F is L-Lipschitz continuous mapping. We compare the convergence behavior of Algorithm YH and Algorithm 1 with Algorithm 3 and the following algorithm (denoted by Algorithm IPAL) which is proposed by Ye [32].

Algorithm IPAL.

Initialization. Choose $0 < \lambda < 1/L$, $\epsilon > 0$ and $x^0 \in \mathbb{R}^n$.

Iteration k $(k = 0, 1, 2, \ldots)$. Having x^k do the following steps.

Step 1. Compute $z^k := P_C(x^k - \lambda F(x^k))$.

Step 2. Set $H_k = \{v \in \mathbb{R}^n : h_k(v) \leq 0\}$ with $h_k(v) := \langle x^k - z^k - \lambda(F(x^k) - F(z^k)), v - z^k\rangle$. Select $t_k \in \operatorname{argmax}\{\operatorname{dist}(x^k, H_j) : 0 \leq j \leq k\}$ and $\widehat{H}_k := H_{t_k}$, and compute

$$x^{k+1} := P_{\widehat{H}_k}(x^k),$$

and go to Iteration k with k replaced by $k+1$.

Example 4.5. Consider $C = [-n\pi/2, n\pi/2]^n$. Let $F\colon C \to \mathbb{R}^n$ be defined by

$$F(x) = \Big(\cos\frac{x_1}{n}, \cos\frac{x_2}{n}, \ldots, \cos\frac{x_n}{n}\Big)^T.$$

It is clear that F is not quasimonotone on C. Since, for $x = (-n\pi/3, n\pi/2, n\pi/2, \ldots, n\pi/2)^T$ and $y = (-n\pi/4, n\pi/3, n\pi/2, \ldots, n\pi/2)^T$, we have

$$\langle F(x), y - x\rangle = -\frac{n\pi}{24} < 0 \quad \text{and} \quad \langle F(y), y - x\rangle = \frac{(\sqrt{2}-1)n\pi}{12\sqrt{2}} > 0.$$

In addition,

$$\|F(x) - F(y)\| \leq \frac{1}{\sqrt{n}}\|x - y\|, \quad \forall\, x, y \in C,$$

so F is Lipschitz continuous with constant $L = \frac{1}{\sqrt{n}}$. On the other hand,

$$S = \big\{(x_1, x_2, \ldots, x_n)^T : x_i \in \{-n\pi/2; n\pi/2\}\big\}$$

and

$$S_M = \big\{(-n\pi/2, -n\pi/2, \ldots, -n\pi/2)^T\big\}.$$

We use Algorithm YH with $\sigma = 0.3$, $\gamma = 0.99$, Algorithm IPAL with parameters $L = \frac{1}{\sqrt{n}}$, $\lambda = \frac{1-\sigma}{L}$, where $\sigma = 0.01$, Algorithm 1 with $\mu_k = \frac{1}{(k+1)^{1.5}}$, $\theta = 0.99$, $\eta = 0.99$, $\lambda = 0.8$, $\delta = 0.8$ and Algorithm 3 with parameters $\mu_k = \frac{1}{(k+3)^{1.5}}$, $L = \frac{1}{\sqrt{n}}$, $\theta = 0.01$, $\alpha = \frac{1-\sigma}{L}$, where $\sigma = 0.01$, to find a solution of VIP(C, F). The starting point is $x^0 = (-n\pi/8, -n\pi/8, -n\pi/8, \ldots, -n\pi/8)^T$ in all algorithms and $x^1 = (-n\pi/8, -n\pi/8, -n\pi/8, \ldots, -n\pi/8)^T$ in Algorithm 1, $x^1 = (-n\pi/16, -n\pi/16, -n\pi/16, \ldots, -n\pi/16)^T$ in Algorithm 3. To terminate the algorithms, we use the stopping criteria $\epsilon = \|r(x^k)\| \leq 10^{-4}$. The numerical results are described in Figure 4.5, which shows that Algorithm IPAL and Algorithm 3 behave better then Algorithm YH, Algorithm 1. Table 4.5 shows that Algorithm IPAL, Algorithm 3 are better than Algorithm YH, Algorithm 1 in number of steps and CPU time.

Table 4.5: Experiment for Example 4.5.

n	Alg. YH		Alg. IPAL		Alg. 1		Alg. 3	
	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.	CPU(s)	Iter.
10	1.35	92	0.53	31	0.81	100	0.34	31
50	62.66	462	4.03	75	7.27	620	0.77	78
100	635.09	952	19.03	114	19.61	1224	1.38	117
150	2265.48	1353	44.34	138	206.42	1988	10.23	143
200	5493.66	1835	92.83	175	306.91	2619	12.81	177

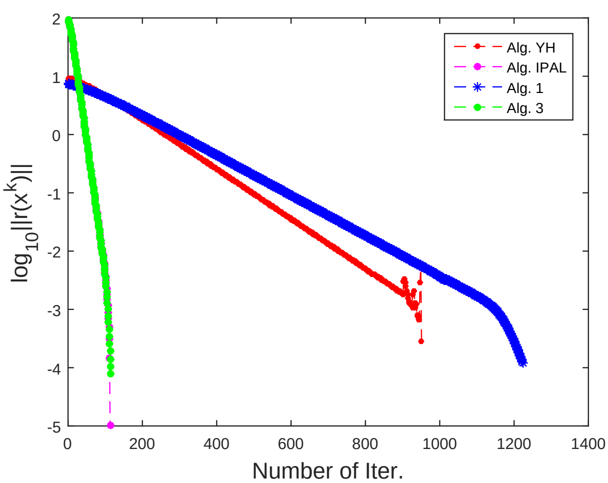

Figure 4.5: Number of iterations of Algorithms in Example 4.5 ($n = 100$, $x^0 = (-n\pi/8, -n\pi/8, -n\pi/8, \ldots, -n\pi/8)^T$).

Remark 4.6. It can be seen from the results of Example 4.5 that the inertial Algorithm 3 outperformed the non-inertial Algorithm IPAL in terms of time taken for computation.

Remark 4.7. From numerical results, we can see that if the cost mapping F is non-Lipschitz continuous and n is small, then the computation cost of Algorithms 1 and 2 are quite the same. On the other hand, if n is large, then Algorithm 1 is better than Algorithm 2. The cause may be that the computation cost in Step 2 of Algorithm 1 is cheaper than those of Algorithm 2.

Specially, these three algorithms can be used to solve VIP(C, F) when F is Lipschitz continuous. However, if the Lipschitz constant L is big, Algorithm 3 may be less efficient than Algorithms 1 and 2. This might be due to the reason that the stepsize in this algorithm is inversely proportional to L.

Conclusion. We have introduced some inertial type algorithms for finding a solution of a nonmonotone variational inequality problem with or without linesearch. All the proposed algorithms do not use the embedding projections and their convergences are obtained. Some numerical examples are reported to illustrate the convergence and also to show the advantage of the new algorithms over the existing method for solving these problems.

Acknowledgments

A part of this article was written while the third author was visiting Vietnam Institute for Advanced Study in Mathematics (VIASM). He would like to thank the institute for warm hospitality and partial support.

References

[1] B. H. Ahn, *Iterative methods for linear complementarity problems with upperbounds on primary variables*, Math. Programming **26** (1983), no. 3, 295–315.

[2] 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), no. 1-2, 3–11.

[3] P. K. Anh and N. T. Vinh, *Self-adaptive gradient projection algorithms for variational inequalities involving non-Lipschitz continuous operators*, Numer. Algorithms **81** (2019), no. 3, 983–1001.

[4] R. I. Boţ and E. R. Csetnek, *An inertial forward-backward-forward primal-dual splitting algorithm for solving monotone inclusion problems*, Numer. Algorithms **71** (2016), no. 3, 519–540.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004.

[6] R. S. Burachik and R. D. Millán, *A projection algorithm for non-monotone variational inequalities*, Set-Valued Var. Anal. **28** (2020), no. 1, 149–166.

[7] R. S. Burachik and S. Scheimberg, *A proximal point method for the variational inequality problem in Banach spaces*, SIAM J. Control Optim. **39** (2000), no. 5, 1633–1649.

[8] G. Cai, A. Gibali, O. S. Iyiola and Y. Shehu, *A new double-projection method for solving variational inequalities in Banach spaces*, J. Optim. Theory Appl. **178** (2018), no. 1, 219–239.

[9] G. P. Crespi, A. Guerraggio and M. Rocca, *Minty variational inequality and optimization: Scalar and vector case*, in: *Generalized Convexity, Generalized Monotonicity and Applications*, 193–211, Nonconvex Optim. Appl. **77**, Springer-Verlag, New York, 2005.

[10] B. V. Dinh and D. S. Kim, *Projection algorithms for solving nonmonotone equilibrium problems in Hilbert space*, J. Comput. Appl. Math. **302** (2016), 106–117.

[11] B. V. Dinh and L. D. Muu, *Algorithms for a class of bilevel programs involving pseudomonotone variational inequalities*, Acta Math. Vietnam. **38** (2013), no. 4, 529–540.

[12] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems: I and II*, Springer Ser. Oper. Res., Springer-Verlag, New York, 2003.

[13] E. M. Gafni and D. P. Bertsekas, *Two-metric projection problems and descent methods for asymmetric variational inequality problems*, Math. Programming **53** (1984), 99–110.

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

[15] Y. He, *A new double projection algorithm for variational inequalities*, J. Comput. Appl. Math. **185** (2006), no. 1, 166–173.

[16] A. N. Iusem and B. F. Svaiter, *A variant of Korpelevich's method for variational inequalities with a new search strategy*, Optimization **42** (1997), no. 4, 309–321.

[17] P. Q. Khanh, D. V. Thong and N. T. Vinh, *Versions of the subgradient extragradient method for pseudomonotone variational inequalities*, Acta Appl. Math. **170** (2020), 319–345.

[18] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, Pure Appl. Math. **88**, Academic Press, New York, 1980.

[19] I. Konnov, *Combined Relaxation Methods for Variational Inequalities*, Lecture Notes in Econom. and Math. Systems **495**, Springer-Verlag, Berlin, 2001.

[20] G. M. Korpelevič, *An extragradient method for finding saddle points and for other problems*, Èkonom. i Mat. Metody **12** (1976), no. 4, 747–756.

[21] Y. Malitsky, *Projected reflected gradient methods for monotone variational inequalities*, SIAM J. Optim. **25** (2015), no. 1, 502–520.

[22] P. Ochs, Y. Chen, T. Brox and T. Pock, *iPiano: Inertial proximal algorithm for nonconvex optimization*, SIAM J. Imaging Sci. **7** (2014), no. 2, 1388–1419.

[23] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc. **73** (1967), 591–597.

[24] B. T. Poljak, *Some methods of speeding up the convergence of iterative methods*, Ž. Vyčisl. Mat. i Mat. Fiz. **4** (1964), 791–803.

[25] G. Stampacchia, *Formes bilinéaires coercitives sur les ensembles convexes*, C. R. Acad. Sci. Paris **258** (1964), 4413–4416.

[26] J. J. Strodiot, P. T. Vuong and T. T. V. Nguyen, *A class of shrinking projection extragradient methods for solving non-monotone equilibrium problems in Hilbert spaces*, J. Global Optim. **64** (2016), no. 1, 159–178.

[27] D. Sun, *A new step-size skill for solving a class of nonlinear projection equations*, J. Comput. Math. **13** (1995), no. 4, 357–368.

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

[29] N. Xiu and J. Zhang, *Some recent advances in projection-type methods for variational inequalities*, J. Comput. Appl. Math. **152** (2003), no. 1-2, 559–585.

[30] H.-K. Xu, *Iterative algorithms for nonlinear operators*, J. London Math. Soc. (2) **66** (2002), no. 1, 240–256.

[31] M. Ye, *An improved projection method for solving generalized variational inequality problems*, Optimization **67** (2018), no. 9, 1523–1533.

[32] ———, *An infeasible projection type algorithm for nonmonotone variational inequalities*, Numer. Algorithms **89** (2022), no. 4, 1723–1742.

[33] M. Ye and Y. He, *A double projection method for solving variational inequalities without monotonicity*, Comput. Optim. Appl. **60** (2015), no. 1, 141–150.

[34] E. H. Zarantonello, *Projections on convex sets in Hilbert space and spectral theory I: Projections on convex sets*, in: *Contributions to Nonlinear Functional Analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971)*, 237–424, Academic Press, New York, 1971.

Bien Thanh Tuyen, Hy Duc Manh and Bui Van Dinh
Department of Mathematics, Faculty of Information Technology, Le Quy Don Technical University, Hanoi, Vietnam
E-mail addresses: thanhtuyenb@lqdtu.edu.vn, ducmanhh@lqdtu.edu.vn, vandinhb@lqdtu.edu.vn, vandinhb@gmail.com