Numerical Algorithms (2022) 89:1723–1742
https://doi.org/10.1007/s11075-021-01170-1

ORIGINAL PAPER

# An infeasible projection type algorithm for nonmonotone variational inequalities

**Minglu Ye**[1]

Received: 10 November 2020 / Accepted: 5 July 2021 / Published online: 3 August 2021


**Abstract**

It is well known that the monotonicity of the underlying mapping of variational inequalities plays a central role in the convergence analysis. In this paper, we propose an infeasible projection algorithm (IPA for short) for nonmonotone variational inequalities. The next iteration point of IPA is generated by projecting a vector onto a half-space. Hence, the computational cost of computing the next iteration point of IPA is much less than the algorithm of Ye and He (Comput. Optim. Appl. **60**, 141–150, 2015) (YH for short). Moreover, if the underlying mapping is Lipschitz continuous with its modulus is known, by taking suitable parameters, IPA requires only one projection onto the feasible set per iteration. The global convergence of IPA is obtained when the solution set of its dual variational inequalities is nonempty. Moreover, if in addition error bound holds, the convergence rate of IPA is Q-linear. IPA can be used for a class of quasimonotone variational inequality problems and a class of quasiconvex minimization problems. Comparing with YH and Algorithm 2 in Deng, Hu and Fang (Numer. Algor. **86**, 191–221, 2021) (DHF for short) by solving high-dimensional nonmonotone variational inequalities, numerical experiments show that IPA is much more efficient than YH and DHF from CPU time point of view. Moreover, IPA is less dependent on the initial value than YH and DHF.

**Keywords** Variational inequalities · Projection algorithm · Infeasible · Quasimonotone · Nonmonotone

**Mathematics Subject Classification (2010)** 47J20 · 49J40 · 90C33

✉ Minglu Ye
yml2002cn@aliyun.com

[1] School of Mathematics and Information, China West Normal University, Nanchong, Sichuan, China

## 1 Introduction

Let $\mathbb{R}^n$ be $n$ dimension Euclidean space, $C \subset \mathbb{R}^n$ be a nonempty closed and convex set and $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping. We consider the classical variational inequality problem of finding a vector $x \in C$ such that

$$\langle F(x), u - x\rangle \geq 0,\ \forall u \in C, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of $\mathbb{R}^n$. We let

$$S := \{x \in C : \langle F(x), u - x\rangle \geq 0,\ \forall u \in C\}$$

denote the solution set of variational inequalities and let

$$S_D := \{x \in C : \langle F(u), u - x\rangle \geq 0,\ \forall u \in C\}$$

denote the solution set of the **dual** variational inequalities of Problem (1).

If $C$ is nonempty closed and convex and $F$ is continuous or hemicontinuous on $C$, the relationship $S_D \subset S$ is known as Minty lemma. The relationship $S \subset S_D$ holds when $F$ is pseudomonotone on $C$ in the sense of Karamardian [1]; i.e., for all $x, y \in C$,

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

Hence, we obtain that

$$S = S_D, \tag{2}$$

whenever $C$ is nonempty closed and convex, $F$ is continuous and pseudomonotone on $C$.

From [2, Example 4.2], we see that $S = S_D$ may not satisfied when $F$ is quasimonotone in the sense of Karamardian [1] on $C$, i.e., for all $x, y \in C$,

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

Recently, an interesting work in [3] show that if $C \subset \mathbb{R}$ is closed and bounded, then $S_D \neq \emptyset$ if and only if $F$ is quasimonotone, see Lemma 3.1 and Proposition 3.1 therein. Moreover, from Theorem 3.1 and Example 3.1 in [3], we construct high-dimensional variational inequalities, which is not quasimonotone with $S_D \neq \emptyset$ whenever the number of dimensional $n \geq 2$, see Example 6.1.

The global convergence of many well-known algorithms are based on the condition that $S = S_D$; see, for example, Goldstein-Levitin-Polyak Projection algorithms [4, 5], proximal point algorithm [6], extragradient projection algorithms [7, 8] and its variant algorithms [9, 10], infeasible projection-based algorithm [11–13], self-adaptive projection-based algorithm [14] and projected reflected gradient algorithm [15] (because these algorithms need $F$ is at least pseudomonotone on $C$). The condition $S = S_D$ is also used for generalized monotone or nonmonotone variational inequalities or nonmonotone equilibrium problems to establish the global convergence, see, for example, [16–18] and [19]. Moreover, the assumption about monotonicity or pseudomonotonicity of the underlying mapping is used in variational inequality and fixed point problems, see, for example, [20–24]. However, if we generalize the pseudomonotone of $F$ to qusimonotone, there exists the case that $S_D \subsetneqq S$; see, for example, [2, Example 4.2]. So, assumption $S = S_D$ may not suit for quasimonotone or nonmonotone variational inequalities.

There still exist some algorithms for quasimonotone variational inequality problems; see, for example, [25] and [26]. However, the global convergence of these algorithms need extra assumptions than $S_D \neq \emptyset$. Recently, [27] presented an iterative algorithm for quasimonotone variational inequalities, the global convergence is established under the assumptions $S_D \neq \emptyset$ and

$$\{z \in C : F(z) = 0\} \backslash S_D \text{ is a finite set.} \tag{3}$$

However, from Example 4.4 below, we see that assumption (3) may not suit for quasimonotone variational inequalities.

In [2], we proposed a double projection algorithm for nonmonotone variational inequalities. The global convergence is obtained under assumptions that $S_D \neq \emptyset$[1], $C$ is closed and convex and $F$ is continuous on $C$ (without needing any monotonicity of $F$). This algorithm is generalized to solve nonmonotone set-valued variational inequalities and nonmonotone equilibrium problems; see, for example, [28–30] and [31]. However, the next iteration point $x^{k+1}$ in [2] is generated by projecting a vector onto the intersection of the feasible set $C$ and $k+1$ half-spaces. Hence, the computational cost of computing $x^{k+1}$ will increase as $k$ increase. Moreover, the next iterate point of algorithms in [28–30] and [31] are all similarly generated in [2]. Very recently, [32] proposed an extragradient method for solving variational inequalities without monotonicity.

Inspired by [11–13, 33, 34], we present an infeasible projection algorithm (IPA) for Problem (1) with $S_D \neq \emptyset$. In IPA, $x^{k+1}$ is generated by projecting $x^k$ onto only **one** half-space, which is selected from former $k+1$ half-spaces and has the largest distance from $x^k$. Moreover, the computational cost of computing $x^{k+1}$ can be ignored, see Remark 3.1 below. If $F$ is in addition $L$-Lipschitz continuous, by taking suitable parameters, IPA requires only **one** projection onto the feasible set per iteration (see Proposition 4.1 and Remark 4.3 below). Comparing with the algorithm in [27], the global convergence of IPA without needing extra assumption in (3). Moreover, if in addition the error bound condition holds and $F$ is Lipschitz continuity, the convergence rate of IPA is Q-linear (see Theorem 5.1 below). To compare with YH and DHF, we construct high-dimensional variational inequalities with $S_D \neq \emptyset$ and perform these algorithms in Matlab. Numerical experiments show that IPA is much more efficient than YH and DHF in terms of CPU time. Moreover, IPA is less dependent on the initial value than YH and DHF.

The rest of this paper is organized as follows. Some notations and preliminary materials are introduced in Section 2. IPA and its convergence analysis is introduced in Section 3. The simplified IPA is introduced in Section 4 when $F$ is in addition Lipschitz continuous on $\mathbb{R}^n$. The convergence rate of IPA is introduced in Section 5 under additional error bound assumption. Finally, numerical experiments are reported in Section 6.

[1]From Lemma 2.8 (b) below or [35, Theorem 3.5.4] , we see that the global minimizer on $C$ of a smooth quasiconvex function $f$ belongs to $S_D$ with $F = \nabla f$. Hence, our algorithm can be applied in a class of quasiconvex optimization problem.

## 2 Preliminaries

In this section, we introduce some properties about the projection mapping and some materials which will be used in future convergence analysis.

Let $\|\cdot\|$ denote Euclidean norm in $\mathbb{R}^n$ and let $\mathrm{dist}(x, C)$ denote the Euclidean distance from a vector $x$ to $C$; i.e.,

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

Let $P_C(x)$ denote the orthogonal projection of the vector $x$ onto $C$; i.e.,

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

From the fact that $C$ is closed and convex, we have $\mathrm{dist}(x, P_C(x)) = \inf\{\|x - y\| : y \in C\}$. For a fixed $x \in \mathbb{R}^n$ and a positive number $\mu$, we let $r(x, \mu)$ denote the natural residual mapping of problem (1); i.e.,

$$r(x, \mu) := x - P_C(x - \mu F(x)). \tag{4}$$

In the following, we recall some well-known properties about the projection mapping and the natural residual mapping.

**Lemma 2.1** [36] Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. Then, the following statements hold.

(a) for any fixed $x \in \mathbb{R}^n$, we have

$$z = P_C(x) \iff z \in C \text{ and } \langle z - x, y - z\rangle \geq 0 \text{ for all } y \in C. \tag{5}$$

(b) $P_C(\cdot)$ is nonexpansive; i.e.,

$$\|P_C(x) - P_C(z)\| \leq \|x - z\| \text{ for all } x, z \in \mathbb{R}^n.$$

**Lemma 2.2** [10] Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set and $x$ be any fixed point in $\mathbb{R}^n$. If $y = P_C(x)$, then we have

$$\|y - z\|^2 \leq \|x - z\|^2 - \|y - x\|^2 \text{ for all } z \in C.$$

**Lemma 2.3** [37] Let $v$ be a fixed vector and $H := \{x \in \mathbb{R}^n : \langle u, x\rangle \leq a\}$ be a half-space. Then

$$P_H(v) = v - \max\left\{\frac{\langle u, v\rangle - a}{\|u\|^2}, 0\right\} u.$$

Moreover, if in addition $v \notin H$, it follows that

$$P_H(v) = v - \frac{\langle u, v\rangle - a}{\|u\|^2} u.$$

**Lemma 2.4** [34, 38] Let $C$ be a nonempty closed and convex set of $\mathbb{R}^n$ and $r(x, \mu)$ be defined in (4). Then the following statements hold.

(a) $x$ is a solution of problems (1) if and only if $\|r(x, \mu)\| = 0$ for each fixed $\mu > 0$.

(b) If there exists a positive number $\mu > 0$ such that $\|r(x, \mu)\| = 0$, then $x$ is a solution of problems (1).

**Lemma 2.5** [39] Let $r(x, \mu)$ be defined in (4). Then, for any fixed point $x \in \mathbb{R}^n$, the following properties hold.

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

The following inequality is a direct conclusion from Lemma 2.5. For simplicity, we omit its proof.

$$\min(1, \mu)\|r(x, 1)\| \le \|r(x, \mu)\| \le \max(1, \mu)\|r(x, 1)\| \quad \text{for any fixed} \quad \mu > 0. \tag{6}$$

**Lemma 2.6** [10] Let $\Omega \subset \mathbb{R}^n$ be a closed convex set and $h$ be a real-valued function with $\mathrm{dom} h = \mathbb{R}^n$. Denote $\widetilde{\Omega} := \{x \in \Omega : h(x) \le 0\}$. If $\widetilde{\Omega}$ is nonempty and $h$ is Lipschitz continuous on $\Omega$ with modulus $\theta > 0$, then

$$\mathrm{dist}(x, \widetilde{\Omega}) \ge \theta^{-1} \max\{h(x), 0\} \quad \text{for all} \quad x \in \Omega.$$

**Lemma 2.7** [40, 41] Let $\{\gamma_k\}$ and $\{\beta_k\}$ be the nonnegative real number sequences satisfying $\sum_{k=0}^{\infty} \beta_k < \infty$ and $\gamma_{k+1} \le \gamma_k + \beta_k$ for all $k$. Then, the sequence $\{\gamma_k\}$ is convergent.

Before ending this section, we use the following lemma to show some sufficient condition about $S_D \neq \emptyset$, see, for example, [2, Proposition 2.1] and [42, Proposition 1].

**Lemma 2.8** If one of the following statements hold

(a) $F$ is pseudomonotone on $C$ and $S \neq \emptyset$;
(b) $F$ is the gradient of $f$, where $f$ is a differentiable quasiconvex function on an open set $K \supset C$ and can attains its global minimum on $C$;
(c) $F$ is quasimonotone on $C$, $F \neq 0$ and $C$ is bounded;
(d) $F$ is quasimonotone on $C$, $F \neq 0$ on $C$ and there exists a positive number $r$ such that, for every $x \in C$ with $\|x\| \ge r$, there exists $y \in C$ such that $\|y\| \le r$ and $\langle F(x), y - x\rangle \le 0$;
(e) $F$ is quasimonotone on $C$ and $S \setminus S_T \neq \emptyset$ with $S_T := \{x \in C | \langle F(x), y - x\rangle = 0, \text{ for all } y \in C\}$;
(f) $F$ is quasimonotone on $C$, $\mathrm{int}C$ is nonempty and there exists $x^* \in S$ such that $F(x^*) \neq 0$,

then $S_D$ is nonempty.

## 3 IPA and its convergence analysis

To solve nonmonotone Problem (1) with $S_D \neq \emptyset$, we see that the next iterate point in [2] is generated by projecting a vector onto the intersection of the feasible set $C$

and $k+1$ half-spaces. In this section, we present a new infeasible projection-based algorithm (IPA) for Problem (1) without monotonicity. The next iteration point of IPA is generated by projecting the current iteration point onto only **one** half-space, which is selected from the former $k+1$ half-spaces and has the largest distance from the current iteration point. In this section, we first present our IPA as Algorithm 1 below. Then, we will show the well-definedness and the global convergence of IPA.

**Algorithm 1**

Step 0. Choose parameters $0 < \alpha_{\min} < \alpha_{\max}$, $\eta \in (0, 1)$ and $\sigma \in (0, 1)$. Choose $x^0 \in \mathbb{R}^n$ as an initial point and $tol > 0$ as a terminate criterion. Set $k = 0$.

Step 1. Pick $\alpha_k^0 \in [\alpha_{\min}, \alpha_{\max}]$ and compute $\alpha_k := \alpha_k^0 \eta^{m_k}$, where $m_k$ is the smallest nonnegative integer $m$ satisfying

$$\alpha_k^0 \eta^m \left\| F(x^k) - F\left(P_C\left(x^k - \alpha_k^0 \eta^m F(x^k)\right)\right)\right\| \leq \sigma \|x^k - P_C(x^k - \alpha_k^0 \eta^m F(x^k))\|. \tag{7}$$

Step 2. Compute $r(x^k, \alpha_k) = x^k - z^k$ with $z^k := P_C(x^k - \alpha_k F(x^k))$. If $\|r(x^k, \alpha_k)\| \leq tol$, then stop; Otherwise, go to Step 3.

Step 3. Set $H_k = \{v \in \mathbb{R}^n : h_k(v) \leq 0\}$ with

$$h_k(v) := \langle x^k - z^k - \alpha_k(F(x^k) - F(z^k)), v - z^k \rangle. \tag{8}$$

Let

$$t_k \in \operatorname{Arg\,max}\{\operatorname{dist}(x^k, H_j) : 0 \leq j \leq k\} \text{ and } \hat{H}_k := H_{t_k}. \tag{9}$$

Compute

$$x^{k+1} := P_{\hat{H}_k}(x^k). \tag{10}$$

Step 4. Let $k = k + 1$ and go to Step 1.

*Remark 3.1* The line search procedure in (7) is inspired by [11]. The initial step-size $\alpha_k^0$ therein is suggested as $\alpha_{k-1}$. This together with the fact $\eta \in (0, 1)$ show that $\{\alpha_k\}$ is nonincreasing. In this paper, we relax the range of $\alpha_k^0$ on a bounded closed interval below away from zero. Thus, we can use Barzilai-Borwein step-size to accelerate algorithm (Barzilai-Borwein step-size is wildly used in convex and nonconvex optimization). Moreover, to avoid a smaller step-size, we enlarge the initial step-size $\alpha_k^0$ by using a self-adaptive strategy. So, our $\alpha_k^0$ is computed by (30), see also our former work [43].

The procedure to choose the half-space which has the largest distance from $x^k$ to half-spaces $\{H_0, \cdots, H_k\}$ needs $k+1$ projection from $x^k$ onto each half-space in $\{H_0, \cdots, H_k\}$. Fortunately, this procedure can be performed in a parallel way. Hence, together with Lemma 2.3, the computational cost of this procedure can be ignored.

Now, we show the well-definedness of line search of IPA whenever $F$ is continuous on $\mathbb{R}^n$.

**Lemma 3.2** Suppose that $F$ is continuous on $\mathbb{R}^n$. Let $\alpha \in [\alpha_{\min}, \alpha_{\max}]$ be a fixed number and $\eta \in (0, 1)$. Then there exists a nonnegative integer $m$ such that

$$\alpha\eta^m \|F(x) - F(P_C(x - \alpha\eta^m F(x)))\| \leq \sigma\|x - P_C(x - \alpha\eta^m F(x))\|. \quad (11)$$

*Proof* If $x \in S$, then, from Lemma 2.4 (a), we have $x = P_C(x - \alpha F(x))$. Hence, (11) holds at $m = 0$.

If $x \notin S$, from Lemma 2.4 (b), we see that $\|r(x, \nu)\| > 0$ for any $\nu > 0$. In this case, we claim that inequality (11) holds after finitely many iterations. Suppose to the contrary that

$$\alpha\eta^m \|F(x) - F(P_C(x - \alpha\eta^m F(x)))\| > \sigma\|x - P_C(x - \alpha\eta^m F(x))\| \text{ for all } m. \quad (12)$$

Then, we consider the following two cases.

Case A: If $x \in C$, then we have $P_C(x) = x$. This together with the continuity of both $P_C(\cdot)$ and $F(\cdot)$ on $\mathbb{R}^n$ and the fact $\eta \in (0, 1)$ gives

$$\|F(x) - F(P_C(x - \alpha\eta^m F(x)))\| \to 0 \text{ as } m \to \infty. \quad (13)$$

On the other hand, in view of the fact that $\eta \in (0, 1)$ and the second relation of Lemma 2.5, for sufficient large $m$, we see that

$$\sigma\frac{\|x - P_C(x - \alpha\eta^m F(x))\|}{\alpha\eta^m} = \sigma\frac{\|r(x, \alpha\eta^m)\|}{\alpha\eta^m} \geq \sigma\frac{\|r(x, 1)\|}{1} > 0, \quad (14)$$

where the last inequality holds due to the fact $x \notin S$ and Lemma 2.4(b). Consequently, (13) and (14) contradict (12).

Case B: If $x \notin C$, then we have $\sigma\|x - P_C(x)\| > 0$. Using this together with the continuity of $F(\cdot)$ and $P_C(\cdot)$, it follows that $\alpha\eta^m\|F(x) - F(P_C(x - \alpha\eta^m F(x)))\| \to 0$ and $\sigma\|x - P_C(x - \alpha\eta^m F(x))\| \to \sigma\|x - P_C(x)\| > 0$ as $m \to \infty$. However, this contradicts (12).

Hence, we see that (11) holds for some finite nonnegative integer $m$. □

Next, we use following lemma to show that the half-space $H_k$ defined by $h_k(v)$ in (8) can separate strictly the current iteration point $x^k$ from $S_D$.

**Lemma 3.3** Suppose that $F$ is continuous on $\mathbb{R}^n$ and $S_D \neq \emptyset$. Let $h_k$ be defined in (8) and $\{x^k\}$ be the infinite sequence generated by IPA. Then, for any fixed $x^* \in S_D$, we have

$$h_k(x^*) \leq 0 \text{ and } h_k(x^k) \geq (1 - \sigma)\|r(x^k, \alpha_k)\|^2 > 0 \text{ for all } k. \quad (15)$$

Moreover, we have $S_D \subseteq H_k$ and $x^k \notin H_k$ for all $k$.

*Proof* Let $x^* \in S_D$. Invoking the definition of $h_k$, we see that

$$\begin{aligned} h_k(x^*) &= \langle x^k - z^k - \alpha_k(F(x^k) - F(z^k)), x^* - z^k\rangle \\ &= \langle x^k - z^k - \alpha_k F(x^k), x^* - z^k\rangle + \alpha_k\langle F(z^k), x^* - z^k\rangle \\ &\leq \langle x^k - z^k - \alpha_k F(x^k), x^* - z^k\rangle \leq 0, \end{aligned}$$

where the first inequality holds from the definition of $S_D$ and the fact that $z^k \in C$, the last inequality holds from (5), the fact $z^k = P_C(x^k - \alpha_k F(x^k))$ and the fact $x^* \in C$. This completes the first relation of (15).

Next, from the definition of $h_k$, we see that

$$\begin{aligned} h_k(x^k) &= \langle x^k - z^k - \alpha_k(F(x^k) - F(z^k)), x^k - z^k\rangle = \|x^k - z^k\|^2 - \alpha_k\langle F(x^k) \\ &\quad - F(z^k), x^k - z^k\rangle \\ &= \|r(x^k, \alpha_k)\|^2 - \alpha_k\langle F(x^k) - F(z^k), r(x^k, \alpha_k)\rangle \geq \|r(x^k, \alpha_k)\|^2 - \sigma\|r(x^k, \alpha_k)\|^2 \\ &= (1-\sigma)\|r(x^k, \alpha_k)\|^2 > 0, \end{aligned}$$

where the first inequality holds from Cauchy-Schwartz inequality and (7), the last inequality holds from the fact $\|r(x^k, \alpha_k)\| > tol > 0$. This completes the second relation of (15).

Hence, from (15) and the definition of $H_k$, we further obtain that $x^* \in H_k$ and $x^k \notin H_k$ for all $k$. This completes the proof. $\square$

Before analyze the convergence of the sequence $\{x^k\}$ generated by IPA, we study the properties of the sequence $\{x^k, z^k, \alpha_k\}$ generated by IPA.

**Theorem 3.4** Suppose that $F$ is continuous on $\mathbb{R}^n$ and $S_D \neq \emptyset$. Let $\sigma$ and $\eta$ be given in IPA and $\{x^k, z^k, \alpha_k\}$ be the infinite sequence generated by IPA. Then the following statements hold.

(a) For any fixed $x \in \bigcap_{i=0}^{\infty} H_i$, the sequence $\{\|x^k - x\|\}$ is convergent. Moreover, the sequences $\{x^k\}$, $\{F(x^k)\}$, $\{z^k\}$ and $\{x^k - z^k - \alpha_k(F(x^k) - F(z^k))\}$ are all bounded. If $F$ is in addition Lipschitz continuous with modulus $L$, one has

$$\inf_{k>0} \alpha_k \geq \min\{\alpha_{min}, \eta\frac{\sigma}{L}\} > 0. \tag{16}$$

(b) For any fixed $x \in \bigcap_{i=0}^{\infty} H_i$, it holds that

$$0 < \mathrm{dist}^2(x^k, H_k) \leq \mathrm{dist}^2(x^k, \hat{H}_k) \leq \|x^k - x\|^2 - \|x^{k+1} - x\|^2 \text{ for all } k. \tag{17}$$

Moreover, we have $\lim_{k\to\infty} \mathrm{dist}(x^k, \hat{H}_k) = 0$.

(c) For any cluster point $\bar{x}$ of the sequence $\{x^k\}$, it holds that $\bar{x} \in \bigcap_{i=0}^{\infty} H_i$.

(d) **(Global subsequential convergence)** Any cluster point of the infinite sequence $\{x^k\}$ is a solution of Problem (1).

*Proof* (a) From the fact $S_D \neq \emptyset$ and Lemma 3.3 that $S_D \subset \bigcap_{i=0}^{\infty} H_i$, we have $\bigcap_{i=0}^{\infty} H_i \neq \emptyset$. Moreover, together with the fact $\bigcap_{i=0}^{\infty} H_i \subset \hat{H}_k$ for any $k$ and (10), we see that, for any fixed $x \in \bigcap_{i=0}^{\infty} H_i$,

$$\|x^{k+1} - x\| = \|P_{\hat{H}_k}(x^k) - x\| = \|P_{\hat{H}_k}(x^k) - P_{\hat{H}_k}(x)\| \leq \|x^k - x\| \text{ for any } k. \tag{18}$$

Using this together with Lemma 2.7, we obtain that, for any fixed $x \in \bigcap_{i=0}^{\infty} H_i$, the sequence $\{\|x^{k+1} - x\|\}$ is convergent. Moreover, $\{x^k\}$ is bounded. Hence, by using the continuity of $F$, we see that $\{F(x^k)\}$ is bounded. Combine this with the fact $z^k = P_C(x^k - \alpha_k F(x^k))$, $0 < \alpha_k < \alpha_{\max}$ and the continuity of

$P_C(\cdot)$, we further obtain that $\{z^k\}$ is bounded. similarly, we have $\{F(z^k)\}$, and $\{x^k - z^k - \alpha_k(F(x^k) - F(z^k))\}$ are all bounded.

Now, we show that $\inf\limits_{k>0} \alpha_k > 0$ when $F$ is Lipschitz continuous. Let $L > 0$ be the Lipschitz modulus of $F$. Then, we see that (7) holds when $\alpha \leq \frac{\sigma}{L}$. Hence, from the definition of $\alpha_k$, we deduces that either $\alpha_k = \alpha_k^0$ (if $\alpha_k^0 \leq \frac{\sigma}{L}$) or $\alpha_k > \eta\frac{\sigma}{L}$. Using this together with the fact $0 < \alpha_{min} \leq \alpha_k^0$, we have

$$\alpha_k \geq \min\{\alpha_{min}, \eta\frac{\sigma}{L}\} > 0 \ \text{ for all } \ k.$$

(b) The first inequality holds from the fact $x^k \notin H_k$ in Lemma 3.3, the second inequality holds from (9) and the third inequality holds from the fact $\bigcap_{i=0}^{\infty} H_i \subseteq \hat{H}_k$ for any $k$, Lemma 2.2 and (10). This completes the proof of (17). Note that the right hand of (17) is summable. Then, we have $\lim\limits_{k\to\infty} \operatorname{dist}(x^k, \hat{H}_k) = 0$.

(c) From the definition of $\hat{H}_k$, we have

$$0 \leq \operatorname{dist}(x^k, H_i) \leq \operatorname{dist}(x^k, \hat{H}_k) \text{ for all } i \leq k.$$

This together with Theorem 3.4(b) implies that

$$\lim_{k\to\infty} \operatorname{dist}(x^k, H_i) = 0 \text{ for any fixed } i.$$

Using this together with the fact $\operatorname{dist}(\cdot, H_i)$ is continuous on $\mathbb{R}^n$ and the fact $\bar{x}$ is a cluster point of $\{x^k\}$, we further obtain that $\operatorname{dist}(\bar{x}, H_i) = 0$ for any fixed $i$. This completes the proof.

(d) Let $\bar{x}$ be a cluster point of the sequence $\{x^k\}$ and $I \subset \mathbb{N}$ be an index set such that $\lim_{k\to\infty, k\in I} x^k = \bar{x}$.

We first show that $\lim_{k\to\infty} \|r(x^k, \alpha_k)\| = 0$. To this end, we recall from (a) that there exists $M_1 > 0$ such that $\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\| \leq M_1$ for all $k$. Using this together with the definition of $h_k$ in (8), we see that $M_1$ is a Lipschitz modulus of $h_k$ for all $k$. Moreover, from Lemma 2.6 and (15), we have

$$\operatorname{dist}(x^k, H_k) \geq M_1^{-1} h_k(x^k) \geq M_1^{-1}(1-\sigma)\|r(x^k, \alpha_k)\|^2 > 0.$$

This together with (b) gives

$$\lim_{k\to\infty} \|r(x^k, \alpha_k)\| = \lim_{k\to\infty} \|x^k - z^k\| = 0. \tag{19}$$

Next, we show that there exists an index set $J \subseteq I$ such that

$$\lim_{k\to\infty, k\in J} \|r(x^k, 1)\| = 0. \tag{20}$$

To this end, we denote $\widetilde{\alpha} := \inf_{i\in I}\{\alpha_i\}$ and consider the following two cases.

Case 1: If $\widetilde{\alpha} > 0$, then $\widetilde{\alpha} \leq \alpha_i$ for all $i$. Using this together with (6), we see that

$$0 \leq \|r(x^i, 1)\| \leq \frac{\|r(x^i, \alpha_i)\|}{\min\{\alpha_i, 1\}} \leq \frac{\|r(x^i, \alpha_i)\|}{\min\{\widetilde{\alpha}, 1\}}, \forall i \in I.$$

Hence, together with (19), we have $\lim_{i\to\infty, i\in I} \|r(x^i, 1)\| = 0$. By taking $J = I$, we see that (20) holds.

Case 2: If $\widetilde{\alpha} = 0$, then there exists an index set $J \subseteq I$ such that $\lim_{i\to\infty, i\in J} \alpha_i = 0$. This together with (7) implies that, for sufficiently large $i \in J$,

$$\alpha_i \eta^{-1} \left\| F(x^i) - F\left(P_C\left(x^i - \alpha_i \eta^{-1} F(x^i)\right)\right)\right\| > \sigma \|r(x^i, \alpha_i \eta^{-1})\|.$$

Hence, for sufficiently large $i \in J$, it follows that

$$\begin{aligned} \left\| F(x^i) - F\left(P_C\left(x^i - \alpha_i \eta^{-1} F(x^i)\right)\right)\right\| &> \frac{\sigma \|r(x^i, \alpha_i \eta^{-1})\|}{\alpha_i \eta^{-1}} \\ &\geq \frac{\sigma \|r(x^i, 1)\|}{1} > 0, \end{aligned} \tag{21}$$

where the second inequality holds from the second relation of Lemma 2.5 and the fact that $\lim_{i\to\infty, i\in J} \alpha_i = 0$, the third inequality holds because $x^i \notin S$.

On the other hand, from the triangle inequality and the fact that $P_C(\cdot)$ is nonexpansive, it holds that

$$\begin{aligned} &\|x^i - P_C\left(x^i - \alpha_i \eta^{-1} F(x^i)\right)\| \\ &\leq \|r(x^i, \alpha_i)\| + \|P_C\left(x^i - \alpha_i F(x^i)\right) - P_C\left(x^i - \alpha_i \eta^{-1} F(x^i)\right)\| \\ &\leq \|r(x^i, \alpha_i)\| + \|x^i - \alpha_i F(x^i) - (x^i - \alpha_i \eta^{-1} F(x^i))\| \\ &= \|r(x^i, \alpha_i)\| + \alpha_i(\eta^{-1} - 1)\|F(x^i)\| \to 0 \text{ (as } i \to \infty, i \in J), \end{aligned} \tag{22}$$

where the limit holds from the fact $J \subseteq I$, (19), the fact $\lim_{i\to\infty, i\in J} \alpha_i = 0$ and the fact $\{F(x^i)\}$ is bounded. Using this together with the fact $\lim_{k\to\infty, k\in I} x^k = \bar{x}$, the continuity of $F$ and (22), it follows that

$$\lim_{i\to\infty, i\in J} \left\| F(x^i) - F\left(P_C\left(x^i - \alpha_i \eta^{-1} F(x^i)\right)\right)\right\| = 0.$$

This together with (21) gives (20).

Finally, by using the continuity of $\|r(\cdot, 1)\|$ and passing to the limit in (20) along the index set $J$, we see that $\|r(\bar{x}, 1)\| = 0$. This together with Lemma 2.4 completes the proof. □

Now, we are ready to show the global convergence of IPA.

**Theorem 3.5** [**Global sequential convergence**] Suppose that $F$ is continuous on $\mathbb{R}^n$ and $S_D \neq \emptyset$. Let $\{x^k\}$ be the infinite sequence generated by IPA. Then $\{x^k\}$ is globally convergent to a solution of Problem (1).

*Proof* Let $x^*$ be any fixed cluster point of $\{x^k\}$. Then, from Theorem 3.4(c), we see that $x^* \in \bigcap_{i=0}^{\infty} H_i$. Hence, by replacing $x$ by $x^*$ in (18), we see that

$$\|x^{k+1} - x^*\| \leq \|x^k - x^*\| \text{ for all } k.$$

Moreover, from Lemma 2.7, it follows that $\{\|x^k - x^*\|\}$ is convergent. This together with the fact that $x^*$ is a cluster point of $\{x^k\}$ gives

$$\lim_{k\to\infty} x^k = x^*.$$

This together with Theorem 3.4(d) gets the desired conclusion. □

## 4 Simplified IPA when underlying mapping is Lipschitz continuous

In Section 3, we present IPA for nonmonotone Problem (1). The global convergence is established under the assumption that $S_D \neq \emptyset$. In this section, we aims to simply IPA by omitting its line search procedure[2] whenever $F$ is $L$-Lipschitz continuous. We call this simplified IPA as IPA$_L$

Now, we show the simplified IPA as Algorithm 2 below.

**Algorithm 2**

Step 0. Choose parameter $0 < \lambda < \frac{1}{L}$. Choose $x^0 \in \mathbb{R}^n$ as an initial point and $tol > 0$ as a terminate criterion. Set $k = 0$.

Step 1. Compute $r(x^k, \lambda) = x^k - z^k$ with $z^k := P_C(x^k - \lambda F(x^k))$. If $\|r(x^k, \lambda)\| \leq tol$, then stop; Otherwise, go to Step 2.

Step 2. Set $H_k = \{v \in \mathbb{R}^n : \langle x^k - z^k - \lambda(F(x^k) - F(z^k)), v - z^k\rangle \leq 0\}$. Choose

$$t_k \in \operatorname{Arg\,max}\{\operatorname{dist}(x^k, H_j) : 0 \leq j \leq k\} \text{ and } \hat{H}_k := H_{t_k}.$$

Compute

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

Step 3. Let $k = k + 1$ and go to Step 1.

We use the following Proposition to show that IPA can reduce to IPA$_L$ whenever $F$ is $L$-Lipschitz continuous.

**Proposition 4.1** Suppose that $F$ is $L$-Lipschitz continuous on $\mathbb{R}^n$. Let $\alpha_k^0 = \lambda$ with $\lambda \in (0, 1/L)$ and $\sigma = \lambda L$. Then, IPA reduces to IPA$_L$.

*Proof* It suffices to show that, for each $k$, $\alpha_k = \lambda$ whenever $\alpha_k^0 = \lambda$ with $\lambda \in (0, 1/L)$. Invoking the definition of Lipschitz continuous, we see that

$$\lambda \left\| F(x^k) - F\left(P_C\left(x^k - \lambda F(x^k)\right)\right)\right\| \leq \lambda L\|x^k - P_C(x^k - \lambda F(x^k))\|.$$

Using this together with the fact that $\alpha_k^0 = \lambda$ and $\sigma = \lambda L$, we see that (7) holds at $m_k = 0$ for any $k$. Hence, we obtain that $\alpha_k = \lambda$ for any $k$. This completes the proof. □

From Proposition 4.1 and Theorem 3.5, we have the following corollary.

[2] For each $k$, we take step-size $\alpha_k$ as a fixed positive number to avoid computing $\alpha_k$ by (7).

**Corollary 4.2** Suppose that $F$ is $L$-Lipschitz continuous on $\mathbb{R}^n$ and $S_D \neq \emptyset$. Let $\{x^k\}$ be an infinite sequence generated by IPA$_L$. Then $\{x^k\}$ is globally convergent to a solution of Problem (1).

*Remark 4.3* If the Lipschitz modulus of $F$ is known, from Proposition 4.1, we see that, by taking suitable parameters, IPA requires only **one** projection onto the feasible set in each iteration. Moreover, from Lemma 2.3, the computational cost of the next iteration point in IPA is much less than the algorithm in [2]. Comparing with algorithm in [27], the global convergence of simplified IPA without needing extra assumption $\{z \in C : F(z) = 0\}\backslash S_D$ is a finite set. Moreover, from (23) in the following example, we see that this assumption may not suit for quasimonotone variational inequality problems.

Before ending this section, we use the following example to show the extra assumption in [27] (see also assumption (3)). Note that the label of formula (3) is (cond:notsuit) may not suit for quasimonotone variational inequality problems.

*Example 4.4* Let $C = [-1, 1] \times [-1, 1]$ and $F(x) = (x_1^2, 0)^T$ with $x = (x_1, x_2)^T \in \mathbb{R}^2$. Then, it is routine to check that $F$ is quasimonotone on $C$. Moreover, we see that

$$S = \{(x_1, x_2) : x_1 = -1 \text{ or } 0, x_2 \in [-1, 1]\} \text{ and } S_D = \{(-1, x_2) : x_2 \in [-1, 1]\}.$$

Moreover,

$$\{z \in C : F(z) = 0\}\backslash S_D = \{(0, x_2) : x_2 \in [-1, 1]\}. \tag{23}$$

## 5 Convergence rate of IPA

In this section, we analyze the rate of the sequence generated by IPA. Our analysis based on the following local error bound; there exist positive numbers $c_1$ and $c_2$ such that

$$\operatorname{dist}(x, S_D) \leq c_1 \|r(x, 1)\| \text{ for all } x \in \mathbb{R}^n \text{ with } \|r(x, 1)\| \leq c_2, \tag{24}$$

where $S_D$ is the solution set of dual variational inequality of Problem (1). This condition is different with the error bound condition in [11], see formula (3) therein. The difference is that we replace $S$ with $S_D$ in [11] therein. However, from (2), we see that (24) is coincide with the error bound in [11] if $F$ is in addition pseudomonotone on $C$. The error bound plays an central role in convergence rate analysis. Interested reader can refer [44–46] and the survey paper [47] for the sufficient conditions of the error bound.

**Theorem 5.1** Suppose that $F$ is Lipschitz continuous on $\mathbb{R}^n$ with modulus $L$ and $S_D \neq \emptyset$. Let $\{x^k\}$ be the infinite sequence generated by IPA. Then, the following statements hold.

(a) There exist $t_1 > 0$ and $t_2 > 0$ such that

$$t_1 \|r(x^k, 1)\| \leq \operatorname{dist}(x^k, H_k) \leq t_2 \|r(x^k, 1) \text{ for all } k. \tag{25}$$

(b) If in addition (24) holds, the convergence of the sequence generated by IPA is Q-linear; i.e., there exists a positive number $\tilde{c} \in [0, 1)$ such that

$$\text{dist}(x^{k+1}, S_D) \leq \tilde{c} \cdot \text{dist}(x^k, S_D) \text{ for sufficient large } k. \tag{26}$$

*Proof* (a) From Lemma 3.3, we have $x^k \notin H_k$ for all $k$. Using this together with Lemma 2.3, we see that

$$P_{H_k}(x^k) = x^k - \frac{\langle x^k - z^k - \alpha_k(F(x^k) - F(z^k))\rangle x^k - z^k}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|^2}(x^k - z^k - \alpha_k(F(x^k) \\ -F(z^k))). \tag{27}$$

Moreover, it follows that

$$\begin{aligned}
\text{dist}(x^k, H_k) &= \| \frac{\langle x^k - z^k - \alpha_k(F(x^k) - F(z^k)), x^k - z^k\rangle}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|^2} \\
&\quad (x^k - z^k - \alpha_k(F(x^k) - F(z^k)))\| \\
&= \frac{\left|\|x^k - z^k\|^2 - \alpha_k\langle F(x^k) - F(z^k), x^k - z^k\rangle\right|}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|} \\
&\geq \frac{\|x^k - z^k\|^2 - \alpha_k \left|\langle F(x^k) - F(z^k), x^k - z^k\rangle\right|}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|} \\
&\geq \frac{\|x^k - z^k\|^2 - \alpha_k\|F(x^k) - F(z^k)\|\|x^k - z^k\|}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|} \geq \frac{\|x^k - z^k\|^2 - \sigma\|x^k - z^k\|^2}{\|x^k - z^k - \alpha_k(F(x^k) - F(z^k))\|} \\
&\geq \frac{\|x^k - z^k\|^2 - \sigma\|x^k - z^k\|^2}{\|x^k - z^k\| + \alpha_k\|F(x^k) - F(z^k)\|} \geq \frac{\|x^k - z^k\|^2 - \sigma\|x^k - z^k\|^2}{\|x^k - z^k\| + \sigma\|x^k - z^k\|} \\
&= \frac{1-\sigma}{1+\sigma}\|x^k - z^k\| = \frac{1-\sigma}{1+\sigma}\|r(x^k, \alpha_k)\| \geq \frac{1-\sigma}{1+\sigma}\min\{1, \alpha_k\}\|r(x^k, 1)\| \\
&\geq \frac{1-\sigma}{1+\sigma}\min\{1, \alpha_{min}, \eta\frac{\sigma}{L}\}\|r(x^k, 1)\|,
\end{aligned}$$

where the first inequality holds from triangle inequality and the fact $\alpha_k > 0$ for all $k$, the second inequality holds from Cauchy-Schwartz inequality, the third inequality and the fifth inequality hold from (7) and the fact $\sigma < 1$, the sixth inequality holds from (6) and the last inequality holds from (16). Let $t_1 = \frac{1-\sigma}{1+\sigma}\min\{1, \alpha_{min}, \eta\frac{\sigma}{L}\}$. Then, the first inequality of (25) holds.

Next, from (27), we see that

$$\begin{aligned}
\text{dist}(x^k, H_k) &\leq \frac{\|x^k - z^k\|^2 + \sigma\|x^k - z^k\|^2}{\|x^k - z^k\| - \sigma\|x^k - z^k\|} = \frac{1+\sigma}{1-\sigma}\|r(x^k, \alpha_k)\| \\
&\leq \frac{1+\sigma}{1-\sigma}\max\{1, \alpha_k\}\|r(x^k, 1)\| \leq \frac{1+\sigma}{1-\sigma}\max\{1, \alpha_{\max}\}\|r(x^k, 1)\|,
\end{aligned}$$

where the first inequality holds from the fact $\alpha_k\|F(x^k) - F(z^k)\| \leq \sigma\|x^k - z^k\|$, the second inequality holds from (6) and the last inequality holds from the facts $0 < \alpha_k^0 \leq \alpha_{\max}$ and $0 < \eta < 1$. Let $t_2 = \frac{1+\sigma}{1-\sigma}\max\{1, \alpha_{\max}\}$. Then, the second inequality of (25) holds.

(b) Note the fact $S_D = \cap_{y\in C}\{x : \langle F(y), y - x\rangle \geq 0\}$ that $S_D$ is closed and convex. Using this together with the assumption $S_D \neq \emptyset$, we see that there exists only one vector $\bar{x}^k \in S_D$ such that

$$\bar{x}^k := P_{S_D}(x^k) \text{ for each } k. \tag{28}$$

We first claim that the following inequality holds

$$\mathrm{dist}^2(x^{k+1}, S_D) \leq \mathrm{dist}^2(x^k, S_D) - t_1^2\|r(x^k, 1)\|^2. \tag{29}$$

To this end, recall from Lemma 3.3 that $S_D \subset \bigcap_{k=0}^{\infty} H_k$ and (17), we see that

$$\|x^{k+1} - \bar{x}^k\|^2 \leq \|x^k - \bar{x}^k\|^2 - \mathrm{dist}^2(x^k, H_k).$$

Using this, we obtain that

$$\begin{aligned}&\mathrm{dist}^2(x^{k+1}, S_D) \leq \mathrm{dist}^2(x^{k+1}, \bar{x}^k) \leq \mathrm{dist}^2(x^k, \bar{x}^k) - \mathrm{dist}^2(x^k, H_k)\\ &= \mathrm{dist}^2(x^k, S_D) - \mathrm{dist}^2(x^k, H_k) \leq \mathrm{dist}^2(x^k, S_D) - t_1^2\|r(x^k, 1)\|^2,\end{aligned}$$

where the first equality holds from (28) and the last inequality holds from (25).

Now, we are ready to show our main conclusion.

From Theorem 3.4(b), we see that $\lim_{k\to\infty} \mathrm{dist}(x^k, H_k) = 0$. Using this together with the first relation of (25), it follows that $\lim_{k\to\infty} \|r(x^k, 1)\| = 0$. Hence, for any given $c_2 > 0$, there exists integer $N > 0$ such that $\|r(x^k, 1)\| \leq c_2$ for all $k \geq N$. This together with (29) and (24) gives

$$\mathrm{dist}^2\left(x^{k+1}, S_D\right) \leq \left(1 - \frac{t_1^2}{c_1^2}\right)\mathrm{dist}^2\left(x^k, S_D\right) \text{ for all } k \geq N.$$

Note the fact $\mathrm{dist}^2\left(x^{k+1}, S_D\right) \geq 0$ that $1 - \frac{t_1^2}{c_1^2} \geq 0$. Let $\tilde{c} = \sqrt{\left(1 - \frac{t_1^2}{c_1^2}\right)}$. Then (26) holds. □

## 6 Numerical experiments

In this section, we test IPA, Algorithm 2.1 in [2] (YH for short) and Algorithm 2 in [31] (DHF for short) for **high-dimensional** nonmonotone variational inequalities problem. All codes are written in Matlab and performed in Matlab2015a on a notebook with Intel(R) Core(TM) i7-5500U CPU (2.40GHZ 2.40 GHZ) and 8 GB of RAM.

In IPA, we take $\alpha_{\min} = 10^{-10}$, $\alpha_{\max} = 10^{10}$, $\eta = 0.5$ and $\sigma = 0.99$. Inspired by the renowned Barzilai-Borwein step-size and self-adaptive step-size, we take the parameter $\alpha_k^0$ as follows: set $\alpha_0^0 = 1$ and take, for $k \geq 1$,

$$\alpha_k^0 = \begin{cases} P_{[10^{-10},10^{10}]}\left(\frac{\|x^k - x^{k-1}\|^2}{\langle x^k - x^{k-1}, F(x^k) - F(x^{k-1})\rangle}\right) & \text{if } \langle x^k - x^{k-1}, F(x^k) - F(x^{k-1})\rangle > 10^{-12},\\ P_{[10^{-10},10^{10}]}(1.5\alpha_{k-1}) & \text{otherwise,}\end{cases} \tag{30}$$

which is similar as in our former work [43]. In YH, we take $\gamma = 0.4$ and $\sigma = 0.99$. In DHF, we take the parameters are what the corresponding reference proposed in [31, Table 5], i.e., $\gamma_k = \frac{k+0.1}{30(k+1)}$, $\theta = 0.98$ and $\rho = 0.5$.

*Example 6.1* Let

$$C = [-1, 1]^n \text{ and } F : (x_1, \ldots, x_n) \mapsto (x_1^2, \cdots, x_n^2).$$

Then, it is routine to check that $S_D = \{(-1, \ldots, -1)\}$ and $S = \{(x_1, \ldots, x_n) : x_i \in \{-1, 0\}, \forall i\}$. Moreover, from the fact that $x \to x^2$ is quasimonotone on $[-1, 1]$, [3, Theorem 3.1] and [3, example 3.1], we see that $F$ is not quasimonotone on $C$ whenever $n \geq 2$.

In Example 6.1, the terminated criterion for YH and DHF are both $tol = 10^{-5}$ and the terminated criterion for IPA is $tol = 10^{-6}$, i.e., the procedure of YH and DHF terminated whenever $\|r(x, 1)\| \leq 10^{-5}$ and the procedure of IPA terminated whenever $\|r(x, \alpha)\| \leq 10^{-6}$. We first generate 9 initial points $x^0$ by setting

$$x^0 := 0.1 * i * ones(n, 1), i \in \{1, 2, 3, \ldots, 9\},$$

where $ones(n, 1) \in \mathbb{R}^{n\times 1}$ is a vector with each component is 1. Next, for a fixed dimensional number $n$, we repeatedly use YH, DHF and IPA to solve Example 6.1 with different initial point 9 times, respectively. Finally, we report the number of iterations (iter), the number of projections (np), the CPU time (in seconds) and dis$xS$[3] in Table 1, averaged over the 9 initial points.

*Example 6.2* In this example, we set

$$C = [0, 1]^n \text{ and } G : (x_1, \ldots, x_n) \mapsto (x_1^2 - x_1, \cdots, x_n^2 - x_n).$$

Then, we can check that $S_D = \{(1, \ldots, 1)\}$ and $S = \{(x_1, \ldots, x_n) : x_i \in \{0, 1\}, \forall i\}$. Let $I$ be the identity mapping. Then, from Example 6.1, we see that $G$ is the difference of $F$ and $I$.

In Example 6.2, the terminated criterion for YH and DHF are both $tol = 10^{-4}$ and the terminated criterion for IPA is $tol = 10^{-6}$. We first use YH, DHF and IPA to solve Example 6.2 with $n = 1$ and initial point $x^0 \in \{0.1, 0.2, 0.3, \ldots, 0.9\}$. We let \ denote the procedure of algorithm lose to find the approximate solution when CPU time less than 2 min. We report iter, np, the CPU time (in seconds) and dis$xS$ in Table 2.

Based on Table 2, we test DHF and IPA by solving high-dimension Example 6.2 with the initial point belongs to the following set

$$\{0.1 * i * ones(n, 1) : i \in \{4, 5, 6, \ldots, 9\}\}. \tag{31}$$

We report iter, np, the CPU time (in seconds) and dis$xS$ in Table 3, averaged over the 6 different initial points.

[3]Here, dis$xS$ is used to denote the distance of the output point to $S$.

**Table 1** Results for Example 6.1

| $n$ | iter | | | np | | | CPU | | | dis$x.S$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | YH | DHF | IPA | YH | DHF | IPA | YH | DHF | IPA | YH | DHF | IPA |
| 50 | 167 | 139 | 30 | 335 | 280 | 31 | 1.90 | 1.51 | 0.00 | 0.008 | 0.007 | 2e-06 |
| 100 | 327 | 302 | 61 | 656 | 606 | 63 | 5.38 | 5.38 | 0.01 | 0.02 | 0.01 | 5e-06 |
| 500 | 496 | 455 | 93 | 996 | 913 | 96 | 22.34 | 21.13 | 0.02 | 0.03 | 0.03 | 7e-06 |
| 1000 | 672 | 619 | 126 | 1348 | 1242 | 130 | 59.96 | 58.56 | 0.04 | 0.05 | 0.04 | 1e-05 |
| 3000 | 839 | 775 | 161 | 1683 | 1556 | 166 | 181.20 | 174.75 | 0.10 | 0.07 | 0.06 | 1e-05 |

**Table 2** Results for Example 6.2 with $n = 1$

| $x^0$ | iter | | | np | | | CPU | | | dis$x\,S$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | YH | DHF | IPA | YH | DHF | IPA | YH | DHF | IPA | YH | DHF | IPA |
| 0.9 | 2 | 6 | 14 | 5 | 13 | 29 | 0.37 | 0.06 | 0.01 | 0.0001 | 3e-05 | 1e-06 |
| 0.8 | \ | 6 | 15 | \ | 13 | 31 | \ | 0.39 | 0.01 | \ | 0.0001 | 2e-06 |
| 0.7 | \ | 7 | 16 | \ | 15 | 33 | \ | 0.40 | 0.01 | \ | 4e-05 | 1e-06 |
| 0.6 | \ | 7 | 14 | \ | 15 | 29 | \ | 0.40 | 0.01 | \ | 9e-05 | 1e-06 |
| 0.5 | \ | 8 | 17 | \ | 17 | 36 | \ | 0.40 | 0.01 | \ | 5e-05 | 1e-06 |
| 0.4 | \ | 8 | 16 | \ | 17 | 37 | \ | 0.40 | 0.01 | \ | 5e-05 | 2e-06 |
| 0.3 | \ | \ | 18 | \ | \ | 37 | \ | \ | 0.01 | \ | \ | 1e-06 |
| 0.2 | \ | \ | 18 | \ | \ | 39 | \ | \ | 0.01 | \ | \ | 1e-06 |
| 0.1 | \ | \ | 17 | \ | \ | 34 | \ | \ | 0.01 | \ | \ | 1e-06 |

**Table 3** Results for Example 6.2 with initial point choose in (31)

| $n$ | iter | | np | | CPU | | dis$xS$ | |
|---|---|---|---|---|---|---|---|---|
| | DHF | IPA | DHF | IPA | DHF | IPA | DHF | IPA |
| 1000 | 25 | 20 | 50 | 43 | 6.12 | 0.01 | 4e-05 | 1e-06 |
| 3000 | 34 | 41 | 70 | 86 | 10.83 | 0.04 | 8e-05 | 3e-06 |
| 5000 | 43 | 63 | 90 | 131 | 20.65 | 0.08 | 0.0001 | 4e-06 |

*Remark 6.3* From Tables 1, 2 and 3, we see that YH, DHF and IPA are all have the ability to find a solution of nonmonotone VI. Obviously, IPA is much more efficient than YH and DHF from the terms of CPU time. Moreover, from Table 2, we see that IPA is less dependent on the initial value than YH and DHF.

**Funding information** This work is partially supported by the National Science Foundation of China (11871359, 11871059 and 11801455) and is supported by scientific research projects of China West Normal University (20A024).

## References

1. Karamardian, S.: Complementarity problems over cones with monotone and pseudomonotone maps. J. Optim. Theory Appl. **18**, 445–454 (1976)
2. Ye, M., He, Y.: A double projection method for solving variational inequalities without monotonicity. Comput. Optim. Appl. **60**, 141–150 (2015)
3. He, Y.: Solvability of the minty variational inequality. J. Optim. Theory. Appl. **174**, 686–692 (2017)
4. Goldstein, A.A.: Convex programming in Hilbert space. Bull. Amer. Math. Soc. **70**, 709–710 (1964)
5. Levitin, E.S., Polyak, B.T.: Constrained minimization problems. USSR Comput. Math. Math. Phys. **6**, 1–50 (1966)
6. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**(5), 877–898 (1976)
7. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Matecon **17**(10), 747–756 (1976)
8. Khobotov, E.N.: Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR Comput. Math. Math. Phys. **27**, 120–127 (1987)
9. Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequality problems. SIAM J. Control Optim. **37**(3), 765–776 (1999)
10. He, Y.R.: A new double projection algorithm for variational inequalities. J. Comput. Appl. Math. **185**(1), 166–173 (2006)
11. Solodov, M.V., Tseng, P.: Modified projection-type methods for monotone variational inequalities. SIAM J. Control Optim. **34**(5), 1814–1830 (1996)
12. He, B.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35**, 69–76 (1997)
13. Censor, Y., Gibali, A., Reich, S.: Extensions of Korpelevich's extragradient method for the variational inequality problem in Euclidean space. Optimization **61**, 1119–1132 (2012)
14. Han, D.R., Lo, H.K.: Two new self-adaptive projection methods for variational inequality problems. Comp. Math. Appl. **43**, 1529–1537 (2002)
15. Malitsky, Y.V.: Projected reflected gradient methods for monotone variational inequalities. SIAM J. Optim. **25**(1), 502–520 (2015)
16. Konnov, I.V.: On the rate of convergence of combined relaxation methods. Izv. Vyssh. Uchebn. Zaved. Mat. **37**(12), 89–92 (1993)

17. Konnov, I.V.: Combined Relaxation Methods for Variational Inequalities. Springer, Berlin (2001)
18. Konnov, I.V.: Combined Relaxation Methods for Generalized Monotone Variational Inequalities. In: Generalized Convexity and Related Topics. Lecture Notes in Economics and Mathematical Systems, vol. 583, pp. 3–31 (2006)
19. Tran, D.Q., Dung, M.L., Nguyen, V.H.: Extragradient algorithms extended to equilibrium problems. Optim **57**(6), 749–776 (2008)
20. Alakoya, T.O., Jolaoso, L.O., Mewomo, O.T.: Modified inertial subgradient extragradient method with self adaptive stepsize for solving monotone variational inequality and fixed point problems. Optim **70**, 545–574 (2021)
21. Alakoya, T.O., Taiwo, A., Mewomo, O.T., et al.: An iterative algorithm for solving variational inequality, generalized mixed equilibrium, convex minimization and zeros problems for a class of nonexpansive-type mappings. Ann. Univ Ferrara (2021)
22. Gibali, A., Jolaoso, L.O., Mewomo, O.T., et al.: Fast and simple bregman projection methods for solving variational inequalities and related problems in banach spaces. Res. Math. **75**, 179 (2020)
23. Jolaoso, L.O., Taiwo, A., Alakoya, T.O., et al.: A Strong Convergence Theorem for Solving Pseudo-monotone Variational Inequalities Using Projection Methods. J. Optim. Theory Appl. **185**, 744–766 (2020)
24. Ogwo, G.N., Izuchukwu, C., Mewomo, O.T.: Inertial methods for finding minimum-norm solutions of the split variational inequality problem beyond monotonicity. Numer Algor. https://doi.org/10.1007/s11075-021-01081-11 (2021)
25. Brito, A.S., da Cruz Neto, J.X., Lopes, J.O., Oliveira, P.R.: Interior proximal algorithm for quasi-convex programming problems and variational inequalities with linear constraints. J. Optim. Theory Appl. **154**, 217–234 (2012)
26. Langenberg, N.: An interior proximal method for a class of quasimonotone variational inequalities. J. Optim. Theory Appl. **155**, 902–922 (2012)
27. Liu, H., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Compu. Optim. Appl. https://doi.org/10.1007/s10589-020-00217-8 (2020)
28. Burachik, R.S., Millan, R.D.: A projection algorithm for Non-Monotone variational inequalities. Set-valued Var. Anal. **28**(1), 149–166 (2020)
29. Strodiot, J.J., Vuong. P. T., Nguyen, T.T.: A class of shrinking projection extragradient methods for solving non-monotone equilibrium problems in Hilbert spaces. J. Glob. Optim. **64**(1), 159–178 (2016)
30. Van Dinh, B., Kim, D.S.: Projection algorithms for solving nonmonotone equilibrium problems in Hilbert space. J. Comput. Appl. Math. **302**, 106–117 (2016)
31. Deng, L., Hu, R., Fang, Y.: Projection extragradient algorithms for solving nonmonotone and non-Lipschitzian equilibrium problems in Hilbert spaces. Numer. Algor. **86**, 191–221 (2021)
32. Lei, M., He, Y.R.: An extragradient method for solving variational inequalities without monotonicity. J. Optim. Theory Appl. **188**, 432–446 (2021)
33. Ye, M.L.: A half-space projection method for solving generalized Nash equilibrium problems. Optim. **66**(7), 1119–1134 (2017)
34. Ye, M.L.: An improved projection method for solving generalized variational inequality problems. Optim. **67**(9), 1523–1533 (2018)
35. Bazaraa, M.S., Sherali, H.D., Shetty, C.M. Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, New Jersey (2006)
36. Zarantonello, E.H.: Projections on convex sets in Hilbert space and spectral theory, contributions to nonlinear functional analysis. Academic Press, New York (1971)
37. Boyd, S., Vandenberghe, L.: Convex optimization. Cambridge University Press (2004)
38. Xiu, N.H., Zhang, J.Z.: Some recent advances in projection-type methods for variational inequalities. J. Comput. Appl. Math. **152**, 559–585 (2003)
39. Gafni, E.M., Bertsekas, D.P.: Two-metric projection problems and descent methods for asymmetric variational inequality problems. Math. Program. **53**, 99–110 (1984)
40. Polyak, B.T.: Introduction to optimization. Optimization Software Inc., New York (1987)
41. Bauschke, H.H., Combettes, P.L.: Convex analysis and monotone operator theory in Hilbert spaces, 2nd edn. Springer International Publishing, AG (2017)
42. Marcotte, P., Zhu, D.L.: A cutting plane method for solving quasimonotone variational inequalities. Comput. Optim. Appl. **20**, 317–324 (2001)
43. Ye, M.L., Pong, T.K.: A subgradient-based approach for finding the maximum feasible subsystem with respect to a set. SIAM J. Optim. **30**(2), 1274–1299 (2020)

44. Luo, Z.Q., Tseng, P.: On the linear convergence of descent methods for convex essentially smooth minimization. SIAM J. Control Optim. **30**(2), 408–425 (1992)
45. Tseng, P.: On linear convergence of iterative methods for the variational inequality problem. J. Comput. Appl. Math. **60**(1-2), 237–252 (1995)
46. Solodov, M.V.: Convergence rate analysis of iteractive algorithms for solving variational inequality problems. Math. Program. **96**(3), 513–528 (2003)
47. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**(1-3), 299–332 (1997)