Journal of Applied Mathematics and Computing
https://doi.org/10.1007/s12190-024-02219-9

ORIGINAL RESEARCH

Extension of the subgradient extragradient algorithm for solving variational inequalities without monotonicity

Jiaxin Chen[1] · Zunjie Huang[1] · Yongle Zhang[1]

Received: 31 May 2024 / Revised: 31 July 2024 / Accepted: 13 August 2024
© The Author(s) under exclusive licence to Korean Society for Informatics and Computational Applied Mathematics 2024

Abstract

Two improved subgradient extragradient algorithms are proposed for solving nonmonotone variational inequalities under the nonempty assumption of the solution set of the dual variational inequalities. First, when the mapping is Lipschitz continuous, we propose an improved subgradient extragradient algorithm with self-adaptive step-size (ISEGS for short). In ISEGS, the next iteration point is obtained by projecting sequentially the current iteration point onto two different half-spaces, and only one projection onto the feasible set is required in the process of constructing the half-spaces per iteration. The self-adaptive technique allows us to determine the step-size without using the Lipschitz constant. Second, we extend our algorithm into the case where the mapping is merely continuous. The Armijo line search approach is used to handle the non-Lipschitz continuity of the mapping. The global convergence of both algorithms is established without monotonicity assumption of the mapping. The computational complexity of the two proposed algorithms is analyzed. Some numerical examples are given to show the efficiency of the new algorithms.

Keywords Nonmonotone variational inequality · Subgradient extragradient algorithm · Self-adaptive · Armijo line search · Global convergence

1 Introduction

Let $\mathbb{R}^n$ be an n-dimensional Euclidean space, $C \subset \mathbb{R}^n$ be a nonempty closed convex set and $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous mapping. We consider the variational inequalities

✉ Yongle Zhang
yongle-zhang@163.com

Jiaxin Chen
cjxmath2018@163.com

Zunjie Huang
hzjmath@163.com

1 School of Mathematical Sciences, Sichuan Normal University, Chengdu 610066, Sichuan, People's Republic of China

Published online: 13 September 2024

(VI(F, C) for short), which is to find a vector $x^* \in C$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \ \forall x \in C, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^n$. The dual variational inequalities of (1.1), is to find a vector $x^* \in C$ such that

$$\langle F(x), x - x^* \rangle \geq 0, \ \forall x \in C. \tag{1.2}$$

Denote by S the solution set of (1.1), i.e.,

$$S := \{x^* \in C : \langle F(x^*), x - x^* \rangle \geq 0, \ \forall x \in C\}, \tag{1.3}$$

and denote by S_D the solution set of (1.2), i.e.,

$$S_D := \{x^* \in C : \langle F(x), x - x^* \rangle \geq 0, \ \forall x \in C\}. \tag{1.4}$$

When the mapping F is continuous and the set C is convex, we have $S_D \subset S$.

For solving VI(F, C), different types of projection algorithms have been proposed in the past few decades, see, for example, Goldstein–Levitin–Polyak projection algorithms [1, 2], proximal point algorithms [3, 4], extragradient algorithms [5, 6], projection and contraction algorithm [7], double projection algorithms [8, 9], subgradient extragradient algorithm [10], projected reflected gradient algorithm [11] and golden ratio algorithm [12], etc. The convergence of all the above algorithms is based on the common condition that $S = S_D$, which is a direct consequence of pseudomonotonicity of F on C when F is continuous. Note that the relationship $S = S_D$ may fail if the mapping F is continuous and quasimonotone on C, see, for example, [13, Example 4.2].

In order to solve quasimonotone variational inequalities, interior proximal algorithms [14, 15] have been proposed. However, the global convergence of those algorithms was obtained under more assumptions than $S_D \neq \emptyset$. Recently, the gradient algorithms with the new step-sizes [16] and two relaxed inertial subgradient extragradient algorithms [17] were proposed for quasimonotone variational inequalities. The global convergence was proved under the assumptions $S_D \neq \emptyset$ and

$$\{z \in C : F(z) = 0\} \backslash S_D \text{ is a finite set.} \tag{1.5}$$

However, the assumption (1.5) may not suit for solving quasimonotone variational inequalities, see, for example, [18, Example 4.4]. Therefore, it is important to propose some algorithms for quasimonotone variational inequalities or the general nonmonotone variational inequalities without the condition (1.5).

Under the assumption of $S_D \neq \emptyset$, Konnov [19] proposed the first extrapolation algorithm with Armijo-type line search and proved the sequence generated by that algorithm has a subsequence converging to a solution. To obtain the global convergence of the generated sequence, Ye and He [13] proposed a double projection algorithm for solving variational inequalities without monotonicity. The global convergence was proved under the mapping F is continuous and $S_D \neq \emptyset$. Note that the next iteration point in [13] is generated by projecting the current iteration point onto the intersection of the feasible set C and $k+1$ half-spaces, where k represents the current number of iteration steps. Therefore, in order to calculate the next iteration point, the computational cost will increase as the number of iteration steps increasing. To overcome the difficulty, Lei and He [20] proposed an extragradient algorithm for variational inequalities without monotonicity, in which the next iteration point is generated by projecting the current iteration point onto the feasible set C. In 2022, a modified Solodov–Svaiter algorithm was introduced by Van, Manh and Thanh [21], in which the next iteration point is obtained by projecting the current iteration point onto the intersection of the feasible set C and one half-space. The global convergence of these algorithms in [20] and [21] is proved under the mapping F is continuous and $S_D \neq \emptyset$. Note that in order to obtain the next iteration point of these algorithms, in [13, 20, 21], one projection onto the feasible set C or the intersection of the feasible C and one (or more) half-space(s) is required.

Very recently, some new projection algorithms have been proposed for nonmonotone variational inequalities to reduce the projection onto the feasible set C or the intersection of C and half-space(s). First, based on the modified projection-type algorithm [22], an infeasible projection algorithm was proposed by Ye [18] and the global convergence was proved under the mapping F is continuous and $S_D \neq \emptyset$. In the algorithm, the next iteration point is generated by projecting the current iteration point onto only one half-space. Then, based on the projection and contraction algorithm [7] and Tseng-type extragradient algorithm [23], Huang, Zhang and He [24] proposed a class modified projection algorithms for solving variational inequalities without monotonicity. In this algorithm, the generation of the next iteration point is obtained by projecting the current iteration point onto the intersection of multiple half-spaces. The global convergence is established under the mapping F is continuous and $S_D \neq \emptyset$. More algorithms for nonmonotone variational inequalities can be seen [25–27].

Inspired by the above related works, in this paper, we propose two improved subgradient extragradient algorithms for solving VI(F, C) without monotonicity. First, an improved subgradient extragradient algorithm with self-adaptive step-size (ISEGS for short) is proposed. The generation of the next iteration point in the algorithm is obtained by projecting sequentially the current iteration point onto two different half-spaces, and only one projection onto the feasible set C is needed in the process of constructing the half-spaces per iteration. Under the mapping F is Lipschitz continuous and $S_D \neq \emptyset$, the global convergence of ISEGS is established without the prior knowledge of Lipschitz constant of the mapping. Then, to weaken the Lipschitz continuity of the mapping F, an improved subgradient extragradient algorithm with Armijo line search (ISEGA for short) is proposed. Under the mapping F is continuous and $S_D \neq \emptyset$, the global convergence of ISEGA is proved.

Our algorithms can be regarded as an extension of the subgradient extragradient algorithm in [10], the infeasible projection algorithm in [18] and the class modified projection algorithms in [24]. Compared with the classical subgradient extragradient algorithm [10], in our two algorithms, the next iteration point is generated by introducing a new half-space and then projecting the point generated by [10] onto the new half-space. In this way, by improving the generation method of the next iteration point, we generalize the classical subgradient extragradient algorithm [10] from a class of monotone (or pseudomonotone) VI(F, C) to a class of nonmonone VI(F, C). Compared with the infeasible projection algorithm [18], although our two algorithms need to project sequentially onto two different half-spaces, the projection onto the half-space has an explicit expression, so it is as easy to implement as the algorithm in [18]. Finally, compared with the class modified projection algorithms in [24], our two algorithms do not need to project the current iteration point onto the intersection of multiple half-spaces, but onto two different half-spaces in sequence. Some numerical experiments show that our two algorithms are much more efficient than the existing algorithms from the view point of the number of iterations and CPU time.

The paper is organized as follows. In Sect. 2, some basic definitions and preliminary materials are introduced. In Sect. 3, we analyze the global convergence and the computational complexity of the proposed algorithms. Finally, we perform some numerical experiments to illustrate the behavior of our proposed algorithms and compare their performance with other algorithms in Sect. 4.

2 Preliminaries

In this section, we recall some basic definitions and well-known lemmas, which are helpful for convergence analysis in Sect. 3. Let $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ denote the norm and inner product of $\mathbb{R}^n$, respectively.

Definition 2.1 Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a mapping. Then, the mapping

(i) F is Lipschitz continuous on $\mathbb{R}^n$ with constant $L > 0$, if

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$

Specifically, if $L = 1$ then F is said nonexpansive mapping.

(ii) F is monotone on $\mathbb{R}^n$, if

$$\langle F(y) - F(x), y - x\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^n.$$

(iii) F is pseudomonotone on $\mathbb{R}^n$, if

$$\langle F(x), y - x\rangle \geq 0 \Rightarrow \langle F(y), y - x\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^n.$$

(iv) F is quasimonotone on $\mathbb{R}^n$, if

$$\langle F(x), y - x\rangle > 0 \Rightarrow \langle F(y), y - x\rangle \geq 0, \quad \forall x, y \in \mathbb{R}^n.$$

From the above definitions, we see that (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv). But the reverse is not true.

In the following, we give the definition and some properties of metric projection.

The metric projection of $x \in \mathbb{R}^n$ onto C is denoted by

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

The distance function from $x \in \mathbb{R}^n$ to C is denoted by

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

Since C is a closed convex set and Π_C is a single-valued mapping, we can obtain that $\text{dist}(x, C) = \|x - \Pi_C(x)\|$.

Lemma 2.1 [28, 29] *Let $C \subset \mathbb{R}^n$ be a nonempty closed convex set and $x \in \mathbb{R}^n$ be any fixed point. Then, the following inequalities hold:*

(i) $\langle \Pi_C(x) - x, \Pi_C(x) - y\rangle \leq 0, \ \forall y \in C;$
(ii) $\|\Pi_C(x) - y\|^2 \leq \|x - y\|^2 - \|x - \Pi_C(x)\|^2, \ \forall y \in C;$
(iii) $\frac{\|x-\Pi_C(x-\alpha F(x))\|}{\alpha} \leq \frac{\|x-\Pi_C(x-\beta F(x))\|}{\beta}, \ \forall \alpha \geq \beta > 0.$

Lemma 2.2 [28] *Let $a \in \mathbb{R}$ be a fixed vector and $H := \{x \in \mathbb{R}^n : \langle w, x\rangle \leq a\}$ be a half-space. Then, for any $u \in \mathbb{R}^n$, we have*

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

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

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

Lemma 2.3 [28] *Let $\{a_k\}$ and $\{b_k\}$ be two nonnegative real number sequences satisfying*

$$a_{k+1} \leq a_k + b_k, \ \Sigma_{k=1}^{\infty} b_k < +\infty, \ \forall k \geq 0.$$

Then, the sequence $\{a_k\}$ is convergent.

3 Main results

In this section, in order to solve VI(F, C) with the mapping F is Lipschitz continuous or continuous (without any generalized monotonicity), two improved subgradient extragradient algorithms are proposed. The generation of the next iteration point of these algorithms does not need to project the current point onto the feasible set C. In order to obtain the convergence proof of those two algorithms, we always assume that the following condition holds:

Condition 3.1 *The solution set of the dual variational inequalities is nonempty, i.e., $S_D \neq \emptyset$.*

Remark 3.1 Some sufficient conditions for $S_D \neq \emptyset$ were given in [13, Proposition 2.1] and [30, Proposition 1].

3.1 ISEGS for nonmonotone VI(F,C) with Lipschitz continuous

In this subsection, based on the subgradient extragradient algorithm [10], we introduce a new half-space and self-adaptive method to propose an improved subgradient extragradient algorithm with self-adaptive step-size (ISEGS) for solving nonmonotone variational inequalities. To obtain the global convergence of ISEGS, we assume the following condition holds:

Condition 3.2 *The mapping $F:\mathbb{R}^n\to\mathbb{R}^n$ is Lipschitz continuous.*

Now, we give the iterative format of ISEGS as follows:

Algorithm 3.1

Step 1. Let $x^0\in\mathbb{R}^n$ be chosen arbitrarily. Choose parameters $\tau\in(0,1)$, $\lambda_0>0$ and a nonnegative real sequence $\{\varepsilon_k\}$ satisfies $\sum_{k=0}^{\infty}\varepsilon_k<+\infty$. Set $k:=0$.

Step 2. Given the current iterations x^k and λ_k, compute

$$y^k=\Pi_C(x^k-\lambda_k F(x^k)).$$

If $x^k=y^k$ or $F(y^k)=0$, then stop and y^k is a solution of VI(F,C). Otherwise, go to Step 3.

Step 3. Compute

$$z^k=\Pi_{T_k}(x^k-\lambda_k F(y^k)),$$

where

$$T_k:=\{x\in\mathbb{R}^n:\langle x^k-y^k-\lambda_k F(x^k),x-y^k\rangle\le 0\}. \tag{3.1}$$

Step 4. Let $H_k:=\{x\in\mathbb{R}^n:\langle F(y^k),x-y^k\rangle\le 0\}$ and choose

$$i_k:=\arg\max_{0\le j\le k}\frac{\langle F(y^j),z^k-y^j\rangle}{\|F(y^j)\|}. \tag{3.2}$$

Compute

$$x^{k+1}=\Pi_{\widetilde{H}_k}(z^k)\text{ with }\widetilde{H}_k:=H_{i_k} \tag{3.3}$$

and

$$\lambda_{k+1}=\begin{cases}\min\{\lambda_k+\varepsilon_k,\ \tau\frac{\|x^k-y^k\|}{\|F(x^k)-F(y^k)\|}\}, & \text{if } F(x^k)\ne F(y^k),\\ \lambda_k+\varepsilon_k, & \text{otherwise.}\end{cases} \tag{3.4}$$

Let $k:=k+1$ and return to Step 2.

Remark 3.2 In [10], the next iteration point is defined by $x^{k+1} = \Pi_{T_k}(x^k - \lambda_k F(y^k))$. In Algorithm 3.1, the generation of the next iteration point is based on the subgradient extragradient algorithm in [10], by introducing a new half-space H_k, the point $\Pi_{T_k}(x^k - \lambda_k F(y^k))$ generated by [10] is projected onto one half-space $\widetilde{H}_k$, i.e.,

$$x^{k+1} = \Pi_{\widetilde{H}_k}(\Pi_{T_k}(x^k - \lambda_k F(y^k))).$$

In this manner, it is unnecessary to assume the generalized monotonicity of the mapping F when proving the global convergence of the Algorithm 3.1.

According to Lemma 2.2, the projection onto the half-space has an explicit expression. Therefore, the calculation of the iteration points z^k and x^{k+1} is easy to implement.

Lemma 3.1 *Assume that the Condition 3.2 holds. Let $\{\lambda_k\}$ be the sequence generated by Algorithm 3.1. Then, there exists a constant $\lambda > 0$ such that $\lim\limits_{k\to\infty} \lambda_k = \lambda$.*

Proof Since the mapping F is Lipschitz continuous (Assume that the Lipschitz constant is $L > 0$). In this case that $F(x^k) \neq F(y^k)$, we have

$$\tau \frac{\|x^k - y^k\|}{\|F(x^k) - F(y^k)\|} \geq \tau \frac{\|x^k - y^k\|}{L\|x^k - y^k\|} = \frac{\tau}{L}. \tag{3.5}$$

According to the definition of λ_k in (3.4) and mathematical induction, it follows that the sequence $\{\lambda_k\}$ has a lower bound $\min\{\frac{\tau}{L}, \lambda_0\}$. Hence, $\lambda_k > 0$ holds for all k.

Next, we prove that the sequence $\{\lambda_k\}$ is convergent. By the definition of λ_k in (3.4), we know that

$$\lambda_{k+1} \leq \lambda_k + \varepsilon_k, \ \forall k \geq 0. \tag{3.6}$$

Since the sequence $\{\varepsilon_k\}$ satisfies $\sum\limits_{k=0}^{\infty} \varepsilon_k < +\infty$, use (3.6) and Lemma 2.3, we deduce that the sequence $\{\lambda_k\}$ is convergent, i.e., there exists a number $\lambda > 0$ such that $\lim\limits_{k\to\infty} \lambda_k = \lambda$. This completes the proof. □

In the following, we shall present three lemmas utilized in the global convergence proof of Algorithm 3.1.

Lemma 3.2 *Let $\{z^k\}$ be a sequence and H_k be a half-space generated by Algorithm 3.1 and $\widetilde{H}_k$ be a half-space defined in (3.3). Then, for all $k \geq 0$, we deduce that* $\operatorname{dist}(z^k, H_j) \leq \operatorname{dist}(z^k, \widetilde{H}_k)$ *holds for any* $j \in \{0, 1, \ldots, k\}$.

Proof Without loss of generality, let

$$I := \{j : z^k \in H_j, \ j \in \{0, 1, \ldots, k\}\},$$
$$J := \{j : z^k \notin H_j, \ j \in \{0, 1, \ldots, k\}\}.$$

Therefore, for all $j \in \{0, 1, \ldots, k\}$, the above equations can be rewritten as

$$\begin{cases} z^k \in H_j, & j \in I, \\ z^k \notin H_j, & j \in J. \end{cases} \tag{3.7}$$

For any fixed $\varsigma \in I$, according to (3.7), we have $z^k \in H_\varsigma$, then $\text{dist}(z^k, H_\varsigma) = 0$. For any fixed $l \in J$, according to (3.7) and the definition of H_j, we know that $\langle F(y^l), z^k - y^l \rangle > 0$. Further, by Lemma 2.2, we obtain that

$$\Pi_{H_l}(z^k) = z^k - \frac{\langle F(y^l), z^k - y^l \rangle}{\|F(y^l)\|^2} F(y^l). \tag{3.8}$$

Thus,

$$\begin{aligned} \text{dist}(z^k, H_l) &= \|z^k - \Pi_{H_l}(z^k)\| \\ &= \|z^k - z^k + \frac{\langle F(y^l), z^k - y^l \rangle}{\|F(y^l)\|^2} F(y^l)\| \\ &= \frac{\langle F(y^l), z^k - y^l \rangle}{\|F(y^l)\|} > 0. \end{aligned} \tag{3.9}$$

Hence, for all $j \in \{0, 1, \ldots, k\}$, we have

$$\text{dist}(z^k, H_j) = \begin{cases} 0, & j \in I, \\ \frac{\langle F(y^j), z^k - y^j \rangle}{\|F(y^j)\|} > 0, & j \in J. \end{cases} \tag{3.10}$$

By the definition of i_k in (3.2), we know that $i_k \in \{0, 1, \ldots, k\}$ and

$$\frac{\langle F(y^{i_k}), z^k - y^{i_k} \rangle}{\|F(y^{i_k})\|} \geq \frac{\langle F(y^j), z^k - y^j \rangle}{\|F(y^j)\|}, \ \forall j \in \{0, 1, \ldots, k\}. \tag{3.11}$$

Combining (3.10) and (3.11), it follows from the definition of $\widetilde{H}_k$ in (3.3) that

$$0 \leq \text{dist}(z^k, H_j) \leq \text{dist}(z^k, \widetilde{H}_k), \ \forall j \in \{0, 1, \ldots, k\}$$

holds for all k. This completes the proof. □

Lemma 3.3 *Assume that Condition 3.1 holds. Let H_k, $\widetilde{H}_k$ and T_k be the half-spaces generated by the Algorithm 3.1. Then, for all k, we have*

$$S_D \subseteq C \cap \bigcap_{k=0}^{\infty} H_k \subseteq C \cap \widetilde{H}_k \subseteq T_k \cap \widetilde{H}_k.$$

Proof For any fixed $u \in S_D$, by the definition of S_D in (1.4) and the fact $y^k \in C$, we have $\langle F(y^k), u - y^k \rangle \leq 0$. It follows from the definition of H_k, we deduce that $u \in H_k$ for all k, which means

$$u \in S_D \subseteq \bigcap_{k=0}^{\infty} H_k \subseteq \widetilde{H}_k. \tag{3.12}$$

By the definition of y^k and Lemma 2.1 (i), we have

$$\langle x^k - \lambda_k F(x^k) - y^k, y - y^k \rangle \leq 0, \ \forall y \in C.$$

Therefore, by the definition of T_k in (3.1), we know that $C \subseteq T_k$ holds for all k. Hence, by the definition of S_D and (3.12), we see that

$$S_D \subseteq C \cap \bigcap_{k=0}^{\infty} H_k \subseteq C \cap \widetilde{H}_k \subseteq T_k \cap \widetilde{H}_k. \tag{3.13}$$

This completes the proof. □

Lemma 3.4 *Assume that Conditions 3.1 and 3.2 hold. Let $\{x^k\}, \{y^k\}, \{z^k\}$ be the sequences and H_k be the half-space generated by Algorithm 3.1. Then, the following statements hold:*

(i) *For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k$, we have*

$$\|z^k - x\|^2 \leq \|x^k - x\|^2 - (1 - \tau \frac{\lambda_k}{\lambda_{k+1}})(\|x^k - y^k\|^2 + \|y^k - z^k\|^2). \tag{3.14}$$

(ii) *There exists $k_0 \in \mathbb{N}^+$ such that $1 - \tau\lambda_k/\lambda_{k+1} > 0$ holds for all $k \geq k_0$.*
(iii) $\lim_{k\to\infty} \|x^k - y^k\| = 0$, $\lim_{k\to\infty} \|y^k - z^k\| = 0$ *and* $\lim_{k\to\infty} \|x^k - z^k\| = 0$.

Proof (i) By Lemma 3.3, we know that $C \cap \bigcap_{k=0}^{\infty} H_k \neq \emptyset$. For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k \subseteq T_k \cap \bigcap_{k=0}^{\infty} H_k \subseteq T_k \cap \widetilde{H}_k$, by the definition of z^k and Lemma 2.1 (ii), we have

$$\begin{aligned} \|z^k - x\|^2 &\leq \|x^k - \lambda_k F(y^k) - x\|^2 - \|x^k - \lambda_k F(y^k) - z^k\|^2 \\ &= \|x^k - x\|^2 - \|x^k - z^k\|^2 + 2\lambda_k \langle F(y^k), x - z^k \rangle \\ &= \|x^k - x\|^2 - \|x^k - z^k\|^2 + 2\lambda_k \langle F(y^k), x - y^k \rangle + 2\lambda_k \langle F(y^k), y^k - z^k \rangle \end{aligned}$$

$$
\begin{aligned}
&\overset{(a)}{\leq} \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 - 2\langle x^k - y^k, y^k - z^k\rangle \\
&\quad + 2\lambda_k \langle F(y^k), y^k - z^k\rangle \qquad (3.15)\\
&= \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 - 2\langle x^k - y^k - \lambda_k F(x^k), y^k - z^k\rangle \\
&\quad + 2\lambda_k \langle F(y^k) - F(x^k), y^k - z^k\rangle \\
&\overset{(b)}{\leq} \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\lambda_k \langle F(y^k) - F(x^k), y^k - z^k\rangle \\
&\leq \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\lambda_k \|F(y^k) - F(x^k)\| \|y^k - z^k\|,
\end{aligned}
$$

where (a) holds from the facts that $\lambda_k > 0$ for all k, $x \in H_k$ and the definition of H_k, and (b) uses the facts that $z^k \in T_k$ and the definition of T_k in (3.1).
In addition, we claim that

$$
\lambda_k \|F(y^k) - F(x^k)\| \leq \tau \frac{\lambda_k}{\lambda_{k+1}} \|x^k - y^k\|. \qquad (3.16)
$$

Indeed, if $\|F(x^k) - F(y^k)\| = 0$, then (3.16) holds. Otherwise, by the definition of λ_k in (3.4), we have

$$
\lambda_{k+1} = \min\{\lambda_k + \varepsilon_k,\ \tau \frac{\|x^k - y^k\|}{\|F(x^k) - F(y^k)\|}\} \leq \tau \frac{\|x^k - y^k\|}{\|F(x^k) - F(y^k)\|}.
$$

It implies that

$$
\lambda_k \|F(x^k) - F(y^k)\| \leq \tau \frac{\lambda_k}{\lambda_{k+1}} \|x^k - y^k\|.
$$

Therefore, we have (3.16) holds.
Substituting (3.16) into (3.15), we have

$$
\begin{aligned}
\|z^k - x\|^2 &\leq \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\tau \frac{\lambda_k}{\lambda_{k+1}} \|x^k - y^k\| \|y^k - z^k\| \\
&\leq \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + \tau \frac{\lambda_k}{\lambda_{k+1}} (\|x^k - y^k\|^2 + \|y^k - z^k\|^2) \\
&\leq \|x^k - x\|^2 - (1 - \tau \frac{\lambda_k}{\lambda_{k+1}})(\|x^k - y^k\|^2 + \|y^k - z^k\|^2).
\end{aligned} \qquad (3.17)
$$

(ii) By the facts that $\lim_{k\to\infty} \lambda_k = \lambda > 0$ and $\tau \in (0, 1)$, we consider the limit

$$
\lim_{k\to\infty} (1 - \tau \frac{\lambda_k}{\lambda_{k+1}}) = 1 - \tau > 0.
$$

Thus, there exists $k_0 \in \mathbb{N}^+$ such that $1 - \tau\lambda_k/\lambda_{k+1} > 0$ holds for all $k \geq k_0$.

(iii) By the definition of x^{k+1} and the nonexpansive property of the projection, we obtain that

$$\begin{aligned}\|x^{k+1}-x\|^2 &= \|\Pi_{\widetilde{H}_k}(z^k)-\Pi_{\widetilde{H}_k}(x)\|^2\\ &\leq \|z^k-x\|^2\\ &\overset{(a)}{\leq} \|x^k-x\|^2-(1-\tau\frac{\lambda_k}{\lambda_{k+1}})(\|x^k-y^k\|^2+\|y^k-z^k\|^2),\end{aligned} \tag{3.18}$$

where (a) holds from (3.17). Combine the fact that $1-\tau\lambda_k/\lambda_{k+1} > 0$ for all $k \geq k_0$, we deduce that

$$\lim_{k\to\infty}\|x^k-y^k\| = 0 \text{ and } \lim_{k\to\infty}\|y^k-z^k\| = 0. \tag{3.19}$$

Moreover, together with the fact that $\|x^k-z^k\| \leq \|x^k-y^k\|+\|y^k-z^k\|$, we have

$$\lim_{k\to\infty}\|x^k-z^k\| = 0. \tag{3.20}$$

This completes the proof. □

Next, we give the global convergence analysis of Algorithm 3.1.

Theorem 3.1 *Assume that Conditions 3.1 and 3.2 hold. Let $\{x^k\}$, $\{y^k\}$, $\{z^k\}$ be the sequences and T_k, H_k, $\widetilde{H}_k$ be the half-spaces generated by Algorithm 3.1. Then, the infinite sequence $\{x^k\}$ converges globally to a solution of VI(F, C).*

Proof For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k \subseteq \widetilde{H}_k$, by the definition of x^{k+1} and Lemma 2.1 (ii), for all $k \geq k_0$, we have

$$\begin{aligned}\|x^{k+1}-x\|^2 &\leq \|z^k-x\|^2-\|x^{k+1}-z^k\|^2\\ &= \|z^k-x\|^2-\text{dist}^2(z^k,\widetilde{H}_k)\\ &\overset{(a)}{\leq} \|x^k-x\|^2-(1-\tau\frac{\lambda_k}{\lambda_{k+1}})(\|x^k-y^k\|^2+\|y^k-z^k\|^2)-\text{dist}^2(z^k,\widetilde{H}_k)\\ &\overset{(b)}{\leq} \|x^k-x\|^2-\text{dist}^2(z^k,\widetilde{H}_k),\end{aligned} \tag{3.21}$$

where (a) follows from Lemma 3.4 (i) and (b) follows from Lemma 3.4 (ii). From the above inequality, we deduce that the sequence $\{\|x^k-x\|\}$ converges and

$$\lim_{k\to\infty}\text{dist}(z^k,\widetilde{H}_k) = 0. \tag{3.22}$$

Hence, it implies that the sequence $\{x^k\}$ is bounded. Then, there exists a subsequence $\{x^{k_l}\}$ of $\{x^k\}$ and a vector $\overline{x} \in \mathbb{R}^n$ such that $\{x^{k_l}\}$ converges to $\overline{x}$. By the facts that

$\lim_{k\to\infty} \|x^k - y^k\| = 0$ and $\lim_{k\to\infty} \|x^k - z^k\| = 0$ in Lemma 3.4 (iii), we obtain that

$$\lim_{l\to\infty} y^{k_l} = \overline{x} \quad \text{and} \quad \lim_{l\to\infty} z^{k_l} = \overline{x}.$$

Moreover, by the definition of y^k, we know $y^{k_l} \in C$, combine the fact that C is a closed set, we have $\overline{x} \in C$.

Now, we prove that $\overline{x} \in \bigcap_{k=0}^{\infty} H_k$. By the definition of $\widetilde{H}_k$ and Lemma 3.2, we know that

$$0 \leq \text{dist}(z^k, H_j) \leq \text{dist}(z^k, \widetilde{H}_k), \ \forall j \in \{0, 1, \dots, k\}. \tag{3.23}$$

Taking the limit as $k \to \infty$ in (3.23) and combining (3.22), for any fixed nonnegative integer j, we have

$$\lim_{k\to\infty} \text{dist}(z^k, H_j) = 0.$$

Hence,

$$\lim_{l\to\infty} \text{dist}(z^{k_l}, H_j) = 0.$$

Let $v^{k_l} := \Pi_{H_j}(z^{k_l})$, it is clear that

$$\lim_{l\to\infty} \|v^{k_l} - z^{k_l}\| = 0.$$

By the fact that $\lim_{l\to\infty} z^{k_l} = \overline{x}$, we have $\lim_{l\to\infty} v^{k_l} = \overline{x}$. From the facts that $\{v^{k_l}\} \subset H_j$ and H_j is closed set, we deduce that $\overline{x} \in H_j$ holds. By the arbitrariness of j, we obtain that $\overline{x} \in \bigcap_{k=0}^{\infty} H_k$.

Next, we prove that $\overline{x} \in S$. According to $y^{k_l} = \Pi_C(x^{k_l} - \lambda_{k_l} F(x^{k_l}))$ and Lemma 2.1 (i), we obtain that

$$\langle y^{k_l} - x^{k_l} + \lambda_{k_l} F(x^{k_l}), y - y^{k_l} \rangle \geq \ 0, \ \forall y \in C,$$

which implies that

$$\frac{\langle y^{k_l} - x^{k_l}, y - y^{k_l} \rangle}{\lambda_{k_l}} + \langle F(x^{k_l}), y - y^{k_l} \rangle \geq 0, \ \forall y \in C. \tag{3.24}$$

Taking the limit as $l \to \infty$ in (3.24), combine with the facts that $\lim_{l\to\infty} \lambda_{k_l} = \lambda > 0$ (see Lemma 3.1), $\lim_{l\to\infty} \|x^{k_l} - y^{k_l}\| = 0$ (see Lemma 3.4 (iii)), the continuity of the mapping F and $\lim_{l\to\infty} y^{k_l} = \overline{x}$, we obtain that

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

which implies $\overline{x} \in S$.

Finally, we prove that the entire sequence $\{x^k\}$ converges globally to $\overline{x}$. Since the fact

$$\overline{x} \in S \cap \bigcap_{k=0}^{\infty} H_k \subseteq C \cap \bigcap_{k=0}^{\infty} H_k.$$

By replacing x by $\overline{x}$ in (3.21), we see that

$$\lim_{k\to\infty} \|x^k - \overline{x}\| = \lim_{l\to\infty} \|x^{k_l} - \overline{x}\| = 0.$$

This completes the proof. □

Remark 3.3 According to Theorem 3.1, we know that the global convergence of Algorithm 3.1 is proved only need under conditions that the mapping F is Lipschitz continuous on $\mathbb{R}^n$ and $S_D \neq \emptyset$. This removes the condition that $S\backslash S_D$ is a finite set in [16].

Before ending this subsection, we give the computational complexity analysis of Algorithm 3.1.

Theorem 3.2 *Assume that Conditions 3.1 and 3.2 hold. For any $k \geq 0$, let x^k, y^k, z^k, T_k, H_k and $\widetilde{H}_k$ be generated by Algorithm 3.1. Denote $\bar{\lambda} = \lambda_0 + \bar{\varepsilon}$, where $\bar{\varepsilon} = \sum_{k=0}^{\infty} \varepsilon_k < +\infty$ and $\{\varepsilon_k\}$ is selected in Algorithm 3.1. If $\bar{\lambda} \in (0, \frac{1}{L})$ (where L is the Lipschitz constant of F), then for any $x^* \in S_D$ and for any $k \geq 0$,*

$$\min_{0\leq i\leq k} \|x^i - y^i\|^2 \leq \frac{1}{k}\frac{\|x^0 - x^*\|^2}{1 - \bar{\lambda}L}.$$

Proof By the definition of λ_{k+1} in (3.4) and mathematical induction, we deduce that the sequence $\{\lambda_k\}$ has upper bound $\bar{\lambda} = \lambda_0 + \bar{\varepsilon}$. Then, we know that

$$\lambda_k \leq \bar{\lambda}, \ \forall k \geq 0. \tag{3.25}$$

For any fixed $x^* \in S_D$, combine Lemma 3.3, we have $x^* \in T_k \cap \widetilde{H}_k \subseteq T_k$. By the definition of z^k and Lemma 2.1 (ii), for any $k \geq 0$, we have

$$\begin{aligned}\|z^k - x^*\|^2 &\leq \|x^k - \lambda_k F(y^k) - x^*\|^2 - \|x^k - \lambda_k F(y^k) - z^k\|^2\\ &\overset{(a)}{\leq} \|x^k - x^*\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\lambda_k\|F(y^k) - F(x^k)\|\|y^k - z^k\|\\ &\overset{(b)}{\leq} \|x^k - x^*\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\bar{\lambda}\|F(y^k) - F(x^k)\|\|y^k - z^k\|\end{aligned}$$

$$
\begin{aligned}
&\leq \|x^k - x^*\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\bar{\lambda}L\|y^k - x^k\|\|y^k - z^k\| \\
&\leq \|x^k - x^*\|^2 - (1 - \bar{\lambda}L)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2), \qquad (3.26)
\end{aligned}
$$

where (a) follows from (3.15) and (b) follows from (3.25).

By the fact $x^* \in T_k \cap \widetilde{H}_k \subseteq \widetilde{H}_k$, combine the definition of x^{k+1} and the nonexpansive property of the projection, for all $k \geq 0$, we deduce that

$$
\begin{aligned}
\|x^{k+1} - x^*\|^2 =\ & \|\Pi_{\widetilde{H}_k}(z^k) - \Pi_{\widetilde{H}_k}(x^*)\|^2 \\
\leq\ & \|z^k - x^*\|^2 \\
\overset{(a)}{\leq}\ & \|x^k - x^*\|^2 - (1 - \bar{\lambda}L)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2) \\
\overset{(b)}{\leq}\ & \|x^k - x^*\|^2 - (1 - \bar{\lambda}L)\|x^k - y^k\|^2,
\end{aligned}
$$

where (a) follows from (3.26) and (b) follows from the fact $\bar{\lambda} \in (0, \frac{1}{L})$. Then, rearranging the above formula, we have

$$
\|x^k - y^k\|^2 \leq \frac{\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2}{1 - \bar{\lambda}L}, \ \forall k \geq 0.
$$

Summing the above display on both sides from $i = 0$ to k, we obtain that

$$
\sum_{i=0}^{k} \|x^i - y^i\|^2 \leq \frac{\|x^0 - x^*\|^2 - \|x^{k+1} - x^*\|^2}{1 - \bar{\lambda}L} \leq \frac{\|x^0 - x^*\|^2}{1 - \bar{\lambda}L}.
$$

Then, we deduce that

$$
\min_{0 \leq i \leq k} \|x^i - y^i\|^2 \leq \frac{1}{k} \sum_{i=0}^{k} \|x^i - y^i\|^2 \leq \frac{1}{k} \frac{\|x^0 - x^*\|^2}{1 - \bar{\lambda}L}.
$$

This completes the proof. □

3.2 ISEGA for nonmonotone VI(F,C) with continuous

In this subsection, in order to further weaken the Lipschitz continuity of the mapping F, we introduce a line search method to search for the step-size and propose an improved subgradient extragradient algorithm with Armijo line search (ISEGA). To obtain the global convergence of ISEGA, we assume the following condition holds:

Condition 3.3 *The mapping $F : \mathbb{R}^n \to \mathbb{R}^n$ is continuous.*

Now, we give the iterative format of ISEGA as follows:

Algorithm 3.2

Step 1. Let $x^0 \in \mathbb{R}^n$ be chosen arbitrarily. Choose parameters $\eta, \sigma \in (0, 1)$ and $\mu > 0$. Set $k := 0$.

Step 2. Given the current iteration point x^k, compute

$$y^k = \Pi_C(x^k - \lambda_k F(x^k)),$$

where $\lambda_k := \mu\eta^{m_k}$ and m_k is the smallest nonnegative integer m satisfies

$$\mu\eta^m \|F(x^k) - F(\Pi_C(x^k - \mu\eta^m F(x^k)))\| \leq \sigma \|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\|. \tag{3.27}$$

If $x^k = y^k$ or $F(y^k) = 0$, then stop and y^k is a solution of VI(F, C). Otherwise, go to Step 3.

Step 3. Compute

$$z^k = \Pi_{T_k}(x^k - \lambda_k F(y^k)),$$

where

$$T_k := \{x \in \mathbb{R}^n : \langle x^k - y^k - \lambda_k F(x^k), x - y^k\rangle \leq 0\}. \tag{3.28}$$

Step 4. Let $H_k := \{x \in \mathbb{R}^n : \langle F(y^k), x - y^k\rangle \leq 0\}$ and choose

$$i_k := \arg\max_{0\leq j\leq k} \frac{\langle F(y^j), z^k - y^j\rangle}{\|F(y^j)\|}. \tag{3.29}$$

Compute

$$x^{k+1} = \Pi_{\widetilde{H}_k}(z^k) \text{ with } \widetilde{H}_k := H_{i_k}. \tag{3.30}$$

Let $k := k + 1$ and go to Step 2.

In the following lemma, we show that the Armijo line search rule (3.27) in Algorithm 3.2 is well defined.

Lemma 3.5 *Assume that Condition 3.3 holds. Let x^k be generated at the beginning of the k-th iteration of Algorithm 3.2 for any $k \geq 0$. Then, there exists a nonnegative integer m such that (3.27) holds.*

Proof If $x^k \in S$, then $x^k = \Pi_C(x^k - \mu F(x^k))$. Hence, (3.27) holds with $m = 0$.

If $x^k \notin S$, suppose to the contrary that for all m, we have

$$\mu\eta^m \|F(x^k) - F(\Pi_C(x^k - \mu\eta^m F(x^k)))\| > \sigma \|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\|. \tag{3.31}$$

Then,

$$\|F(x^k) - F(\Pi_C(x^k - \mu\eta^m F(x^k)))\| > \sigma \frac{\|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\|}{\mu\eta^m}. \tag{3.32}$$

Now, we consider the following two possibilities of x^k:

i) If $x^k \in C$, by the continuity of the metric projection Π_C, we have

$$\lim_{m\to\infty} \|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\| = 0. \tag{3.33}$$

The continuity of the mapping F implies that

$$\lim_{m\to\infty} \|F(x^k) - F(\Pi_C(x^k - \mu\eta^m F(x^k)))\| = 0.$$

Combining this with (3.32), we have

$$\lim_{m\to\infty} \frac{\|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\|}{\mu\eta^m} = 0. \tag{3.34}$$

Let $y_m^k := \Pi_C(x^k - \mu\eta^m F(x^k))$. By the Lemma 2.1 (i), we deduce that

$$\langle y_m^k - x^k + \mu\eta^m F(x^k), y - y_m^k\rangle \geq 0, \ \forall y \in C,$$

or equivalently,

$$\langle \frac{y_m^k - x^k}{\mu\eta^m}, y - y_m^k\rangle + \langle F(x^k), y - y_m^k\rangle \geq 0, \ \forall y \in C. \tag{3.35}$$

Taking the limit as $m \to \infty$ in (3.35), combining (3.33) with (3.34), we obtain that

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

which implies that it contradicts $x^k \notin S$.

ii) If $x^k \notin C$, then

$$\lim_{m\to\infty} \|x^k - \Pi_C(x^k - \mu\eta^m F(x^k))\| = \|x^k - \Pi_C(x^k)\| > 0 \tag{3.36}$$

and

$$\lim_{m\to\infty} \mu\eta^m \|F(x^k) - F(\Pi_C(x^k - \mu\eta^m F(x^k)))\| = 0. \tag{3.37}$$

Combining (3.31), (3.36) and (3.37), we obtain a contradiction. Therefore, there exists a nonnegative integer m such that (3.27) holds. This completes the proof. □

Lemma 3.6 *Assume that Conditions 3.1 and 3.3 hold. Let $\{x^k\}$, $\{y^k\}$, $\{z^k\}$ be the sequences and T_k, H_k, $\widetilde{H}_k$ be the half-spaces generated by Algorithm 3.2. Then, the following statements hold:*

(i) *For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k$, we have*

$$\|z^k - x\|^2 \le \|x^k - x\|^2 - (1-\sigma)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2), \ \forall k \ge 0. \tag{3.38}$$

(ii) $\lim_{k\to\infty} \|x^k - y^k\| = 0$, $\lim_{k\to\infty} \|y^k - z^k\| = 0$ *and* $\lim_{k\to\infty} \|x^k - z^k\| = 0$.

Proof (i) For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k \subseteq C \cap \widetilde{H}_k \subseteq T_k \cap \widetilde{H}_k$, by the definition of z^k and Lemma 2.1 (ii), we have

$$\begin{aligned}
\|z^k - x\|^2 &\le \|x^k - \lambda_k F(y^k) - x\|^2 - \|x^k - \lambda_k F(y^k) - z^k\|^2 \\
&= \|x^k - x\|^2 - \|x^k - z^k\|^2 + 2\lambda_k\langle F(y^k), x - y^k\rangle + 2\lambda_k\langle F(y^k), y^k - z^k\rangle \\
&\le \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 - 2\langle x^k - y^k, y^k - z^k\rangle \\
&\quad + 2\lambda_k\langle F(y^k), y^k - z^k\rangle \\
&= \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 - 2\langle x^k - y^k - \lambda_k F(x^k), y^k - z^k\rangle \\
&\quad + 2\lambda_k\langle F(y^k) - F(x^k), y^k - z^k\rangle \\
&\overset{(a)}{\le} \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\lambda_k\langle F(y^k) - F(x^k), y^k - z^k\rangle \\
&\overset{(b)}{\le} \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\lambda_k\|F(y^k) - F(x^k)\|\|y^k - z^k\| \\
&\overset{(c)}{\le} \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + 2\sigma\|y^k - x^k\|\|y^k - z^k\| \\
&\le \|x^k - x\|^2 - \|x^k - y^k\|^2 - \|y^k - z^k\|^2 + \sigma(\|x^k - y^k\|^2 + \|y^k - z^k\|^2) \\
&= \|x^k - x\|^2 - (1-\sigma)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2),
\end{aligned} \tag{3.39}$$

where (a) follows from the fact $z^k \in T_k$ and the definition of T_k in (3.28), (b) follows from the fact that $\lambda_k > 0$ for all k and Cauchy–Schwarz inequality, and (c) follows from (3.27).

(ii) By the definition of x^{k+1}, the nonexpansive property of the projection and (3.39), we obtain that

$$\begin{aligned}
\|x^{k+1} - x\|^2 &= \|\Pi_{\widetilde{H}_k}(z^k) - \Pi_{\widetilde{H}_k}(x)\|^2 \\
&\le \|z^k - x\|^2 \\
&\le \|x^k - x\|^2 - (1-\sigma)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2).
\end{aligned} \tag{3.40}$$

Since the fact $\sigma \in (0, 1)$, we deduce that

$$\lim_{k\to\infty} \|x^k - y^k\| = 0 \ \text{ and } \ \lim_{k\to\infty} \|y^k - z^k\| = 0. \tag{3.41}$$

Furthermore, together with the fact $\|x^k - z^k\| \leq \|x^k - y^k\| + \|y^k - z^k\|$, we have

$$\lim_{k\to\infty} \|x^k - z^k\| = 0.$$

This completes the proof. □

In the following, we give the global convergence analysis of Algorithm 3.2.

Theorem 3.3 *Assume that Conditions 3.1 and 3.3 hold. Let $\{x^k\}$, $\{y^k\}$, $\{z^k\}$ be the sequences and T_k, H_k, $\widetilde{H}_k$ be the half-spaces generated by Algorithm 3.2. Then, the infinite sequence $\{x^k\}$ converges globally to a solution of VI(F, C).*

Proof For any fixed $x \in C \cap \bigcap_{k=0}^{\infty} H_k \subseteq \widetilde{H}_k$, by the definition of x^{k+1} and Lemma 2.1 (ii), we have

$$\begin{aligned} \|x^{k+1} - x\|^2 \leq & \ \|z^k - x\|^2 - \|x^{k+1} - z^k\|^2 \\ = & \ \|z^k - x\|^2 - \mathrm{dist}^2(z^k, \widetilde{H}_k) \\ \overset{(a)}{\leq} & \ \|x^k - x\|^2 - (1-\sigma)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2) - \mathrm{dist}^2(z^k, \widetilde{H}_k) \\ \overset{(b)}{\leq} & \ \|x^k - x\|^2 - \mathrm{dist}^2(z^k, \widetilde{H}_k), \end{aligned} \tag{3.42}$$

where (a) follows from Lemma 3.6 (i), (b) follows from $\sigma \in (0, 1)$. From the above inequality, we deduce that the sequence $\{\|x^k - x\|\}$ converges and

$$\lim_{k\to\infty} \mathrm{dist}(z^k, \widetilde{H}_k) = 0. \tag{3.43}$$

Hence, it implies that the sequence $\{x^k\}$ is bounded. Then, there exists a subsequence $\{x^{k_n}\}$ of $\{x^k\}$ and $\widetilde{x} \in \mathbb{R}^n$ such that the subsequence $\{x^{k_n}\}$ converges to $\widetilde{x}$. By the facts that $\lim_{k\to\infty} \|x^k - y^k\| = 0$ and $\lim_{k\to\infty} \|x^k - z^k\| = 0$ in Lemma 3.6 (ii), we know that

$$\lim_{n\to\infty} y^{k_n} = \widetilde{x} \ \text{ and } \ \lim_{n\to\infty} z^{k_n} = \widetilde{x}.$$

Moreover, by the definition of y^k, we know $y^{k_n} \in C$, combine the fact that C is a closed set, we have $\widetilde{x} \in C$.

Now, we prove that $\widetilde{x} \in \bigcap_{k=0}^{\infty} H_k$. We omit this part of the proof and refer to the corresponding part of the proof in Theorem 3.1.

Next, we prove that $\widetilde{x} \in S$. According to $y^{k_n} = \Pi_C(x^{k_n} - \lambda_{k_n} F(x^{k_n}))$ and Lemma 2.1 (i), we have

$$\langle y^{k_n} - x^{k_n} + \lambda_{k_n} F(x^{k_n}), y - y^{k_n} \rangle \geq 0, \ \forall y \in C,$$

or equivalently,

$$\frac{\langle y^{k_n} - x^{k_n}, y - y^{k_n} \rangle}{\lambda_{k_n}} + \langle F(x^{k_n}), y - y^{k_n} \rangle \geq 0, \ \forall y \in C. \tag{3.44}$$

We consider the following two cases:

i) If $\lim_{n\to\infty} \lambda_{k_n} > 0$. Taking the limit as $n \to \infty$ in (3.44), by the fact $\lim_{n\to\infty} \|x^{k_n} - y^{k_n}\| = 0$ in Lemma 3.6 (ii), the continuity of the mapping F and $\lim_{n\to\infty} y^{k_n} = \widetilde{x}$, we obtain that

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

which means $\widetilde{x} \in S$.

ii) If $\lim_{n\to\infty} \lambda_{k_n} = 0$. Assume $w^{k_n} := \Pi_C(x^{k_n} - \lambda_{k_n}\eta^{-1}F(x^{k_n}))$. By the definition of λ_k, we know that $\lambda_{k_n} > 0$ for all n. Hence, $\lambda_{k_n}\eta^{-1} > \lambda_{k_n}$ (since $\eta \in (0, 1)$). Applying Lemma 2.1 (iii), we obtain that

$$\|x^{k_n} - w^{k_n}\| \leq \frac{1}{\eta}\|x^{k_n} - y^{k_n}\|.$$

By the fact that $\lim_{n\to\infty} \|x^{k_n} - y^{k_n}\| = 0$ in Lemma 3.6 (ii), we have

$$\lim_{n\to\infty} \|x^{k_n} - w^{k_n}\| = 0,$$

which implies that $\lim_{n\to\infty} w^{k_n} = \widetilde{x}$. Since the mapping F is continuous, we have

$$\lim_{n\to\infty} \|F(x^{k_n}) - F(w^{k_n})\| = 0. \tag{3.45}$$

By the definition of λ_k in Step 2 of Algorithm 3.2 and (3.27), we have

$$\lambda_{k_n}\eta^{-1}\|F(x^{k_n}) - F(\Pi_C(x^{k_n} - \lambda_{k_n}\eta^{-1}F(x^{k_n})))\| > \sigma\|x^{k_n} - \Pi_C(x^{k_n} - \lambda_{k_n}\eta^{-1}F(x^{k_n}))\|.$$

Reformulating the above inequality, we obtain that

$$\frac{1}{\sigma}\|F(x^{k_n}) - F(w^{k_n})\| > \frac{\|x^{k_n} - w^{k_n}\|}{\lambda_{k_n}\eta^{-1}}. \tag{3.46}$$

Combining (3.45) and (3.46), we deduce that

$$\lim_{n\to\infty} \frac{\|x^{k_n} - w^{k_n}\|}{\lambda_{k_n}\eta^{-1}} = 0. \tag{3.47}$$

Furthermore, by the definition of w^{k_n} and Lemma 2.1 (i), we have

$$\langle w^{k_n} - x^{k_n} + \lambda_{k_n}\eta^{-1}F(x^{k_n}), y - w^{k_n}\rangle \geq 0, \ \forall y \in C.$$

The above inequality can be rewritten as

$$\frac{\langle w^{k_n} - x^{k_n}, y - w^{k_n}\rangle}{\lambda_{k_n}\eta^{-1}} + \langle F(x^{k_n}), y - w^{k_n}\rangle \geq 0, \ \forall y \in C. \tag{3.48}$$

Taking the limit as $n \to \infty$ in (3.48), according to (3.47), the mapping F is continuous and $\lim_{n\to\infty} w^{k_n} = \widetilde{x}$, we have

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

which means $\widetilde{x} \in S$.

Finally, we prove that the entire sequence $\{x^k\}$ converges globally to $\widetilde{x}$. Since the fact that

$$\widetilde{x} \in S \cap \bigcap_{k=0}^{\infty} H_k \subseteq C \cap \bigcap_{k=0}^{\infty} H_k.$$

By replacing x by $\widetilde{x}$ in (3.42), we have

$$\lim_{k\to\infty} \|x^k - \widetilde{x}\| = \lim_{n\to\infty} \|x^{k_n} - \widetilde{x}\| = 0.$$

This completes the proof. □

Before ending this subsection, we give the computational complexity analysis of Algorithm 3.2.

Theorem 3.4 *Assume that Conditions 3.1 and 3.3 hold. For any $k \geq 0$, let x^k, y^k, z^k T_k, H_k and$\widetilde{H}_k$ be generated by Algorithm 3.2. If $\sigma \in (0, 1)$, then, for any $x^* \in S_D$ and $k \geq 0$, we have*

$$\min_{0\leq i\leq k} \|x^i - y^i\|^2 \leq \frac{1}{k}\frac{\|x^0 - x^*\|^2}{1-\sigma}.$$

Proof For any fixed $x^* \in S_D$, combine Lemma 3.3, we have $x^* \in T_k \cap \widetilde{H}_k \subseteq \widetilde{H}_k$. By the definition of x^{k+1} and the nonexpansive property of the projection, for any $k \geq 0$, we have

$$\begin{aligned}\|x^{k+1} - x^*\|^2 =& \ \|\Pi_{\widetilde{H}_k}(z^k) - \Pi_{\widetilde{H}_k}(x^*)\|^2 \\ \leq& \ \|z^k - x^*\|^2\end{aligned}$$

$$\overset{(a)}{\leq} \|x^k - x^*\|^2 - (1-\sigma)(\|x^k - y^k\|^2 + \|y^k - z^k\|^2)$$
$$\overset{(b)}{\leq} \|x^k - x^*\|^2 - (1-\sigma)\|x^k - y^k\|^2,$$

where (a) follows from Lemma 3.6 (i) and (b) follows from $\sigma \in (0, 1)$. Then, rearranging the above formula, we deduce that

$$\|x^k - y^k\|^2 \leq \frac{\|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2}{1-\sigma}, \ \forall k \geq 0.$$

By adding up the two sides of the above inequality, we obtain that

$$\sum_{i=0}^{k} \|x^i - y^i\|^2 \leq \frac{\|x^0 - x^*\|^2 - \|x^{k+1} - x^*\|^2}{1-\sigma} \leq \frac{\|x^0 - x^*\|^2}{1-\sigma}.$$

Then, we deduce that

$$\min_{0 \leq i \leq k} \|x^i - y^i\|^2 \leq \frac{1}{k}\sum_{i=0}^{k} \|x^i - y^i\|^2 \leq \frac{1}{k}\frac{\|x^0 - x^*\|^2}{1-\sigma}.$$

This completes the proof. □

4 Numerical experiments

In this section, the effectiveness of our Algorithm 3.1 (Algorithm 3.1 for short) and Algorithm 3.2 (Algorithm 3.2 for short) will be illustrated through numerical experiments and compared with the existing algorithms. All programs are written in Matlab R2017a and executed on an Intel(R) Core(TM) i5-5350U CPU @ 1.80GHz 1.80GHz.

Let $\|x^k - y^k\| \leq Err$ be the stopping criterion of these algorithms. We report that Iter. represents the number of iterations, np represents the number of projections onto the feasible set C and CPU(s) represents the time the program runs (in seconds).

Example 4.1 Considering the following variational inequality problem VI(F, C) with

$$F : (x_1, x_2, \ldots, x_n) \mapsto (\sin x_1, \sin x_2, \ldots, \sin x_n) \ \text{ and } \ C = [0, \pi]^n.$$

One can see that the mapping F is not quasimonotone on the feasible set C. Indeed, taking $\bar{x} = (\frac{\pi}{6}, \pi, \pi, \ldots, \pi)$ and $\bar{y} = (\frac{\pi}{4}, \frac{5\pi}{6}, \pi, \ldots, \pi)$, we have

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

Table 1 Results for Example 4.1

n	Err	[18, Alg.1]			Algorithm 3.1			Algorithm 3.2		
		Iter.	np	CPU(s)	Iter.	np	CPU(s)	Iter.	np	CPU(s)
500	10^{-4}	16	32	1.5554	8	8	0.3918	9	17	0.8589
	10^{-5}	18	36	1.7880	10	10	0.5138	11	21	1.0854
	10^{-6}	20	40	1.9528	12	12	0.6170	12	23	1.1555
3000	10^{-4}	17	34	114.8506	9	9	31.9254	10	19	70.2992
	10^{-5}	19	38	129.7722	11	11	39.6853	12	23	93.0657
	10^{-6}	21	42	186.6146	12	12	58.0784	13	25	116.2408

In addition, it is clear that the mapping F is Lipschitz continuous on $\mathbb{R}^n$. Moreover, we have

$$S = \{(x_1, x_2, \ldots, x_n) : \ x_i \in \{0, \pi\}\} \ \text{and} \ S_D = \{(0, 0, \ldots, 0)\}.$$

Let $x^0 := (\frac{\pi}{2}, \frac{\pi}{2}, \ldots, \frac{\pi}{2})$ be the initial point. In this example, we compare [18, Alg.1], Algorithms 3.1 and 3.2 by changing the dimension n and the value of Err. We use the same parameter settings as in [18]. That is, we choose $\eta = 0.5, \ \sigma = 0.99$ for [18, Alg.1]. For Algorithm 3.1, we choose $\lambda_0 = 1.5, \ \mu = 0.6, \varepsilon_k = \frac{100}{(k+1)^2}$. For Algorithm 3.2, we choose $\mu = 1.55, \ \eta = 0.4, \ \sigma = 0.99$. The computational results are reported in Table 1.

Remark 4.1 From Table 1, one can see that Algorithm 3.1 needs only to calculate one projection onto the feasible set C per iteration. Moreover, from the view point of the number of iterations and computational time, we have Algorithm 3.1 is more efficient than [18, Alg.1] and Algorithm 3.2 for nonmonotone VI(F, C) with Lipschitz continuous on $\mathbb{R}^n$.

In the following three examples, it is easy to verify that the Lipschitz continuity of the mapping F on $\mathbb{R}^n$ is not true, therefore it does not satisfy the Condition 3.2 of Algorithm 3.1. Then, we only test the effectiveness of Algorithm 3.2 for these three examples.

Example 4.2 The well-known asymmetric transportation network equilibrium example was proposed in [31, Example 7.5]. The network has 25 nodes and 37 directed links (the set of directed links is denoted by L), and the network topology is shown in Fig. 1.

Let W be a set of the origin/destination (O/D) pairs of this network, P_w be a set of all paths corresponding to the O/D pair $\omega \in W$, and P be a set of all paths in the network. Then, we have the path set $P = \bigcup_{\omega \in W} P_\omega$. The O/D pairs and the number of paths of each O/D pair for this example are given in Table 2.

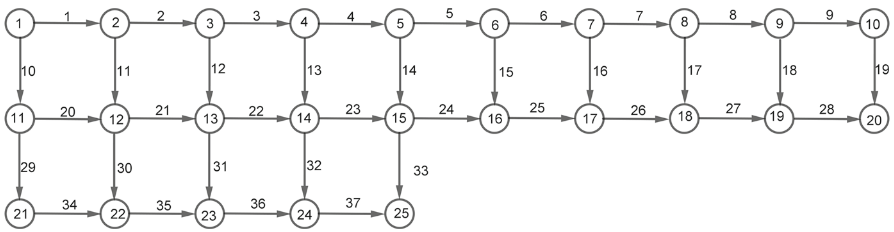

Fig. 1 A transportation network topology

Table 2 The O/D pairs and the number of paths of each O/D pair

No.pairs	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
(O,D)	(1,20)	(1,25)	(2,20)	(3,25)	(1,24)	(11,25)
No.paths	10	15	9	6	10	5

Denote by A the path-link incidence matrix, i.e.,

$$A = (\delta_{pl})_{55\times 37}, \quad \text{where } \delta_{pl} = \begin{cases} 1, & \text{if the link } l \in p \text{ (where } l \in L,\ p \in P), \\ 0, & \text{otherwise.} \end{cases}$$

Denote by B the path-O/D pair incidence matrix, i.e.,

$$B = (v_{p\omega})_{55\times 6}, \quad \text{where } v_{p\omega} = \begin{cases} 1, & \text{if the path } p \in P_\omega \text{ (where } p \in P,\ \omega \in W), \\ 0, & \text{otherwise.} \end{cases}$$

We will use the symbol x_p to represent the traffic flow on path p, then $x = (x_p)_{p\in P} \in \mathbb{R}^{55\times 1}$. Hence, the traffic flow f on link is $A^T x$ and the traffic flow d between the O/D pair w is $B^T x$.

For the link flow, there is a link travel cost function $t(u)$ and a travel disutility function $\lambda(v)$, which are given in [32, Example 5.1]. Because both the travel cost and the travel disutility are functions of the path flow x, we can describe the traffic equilibrium example with link capacity bounds as the VI(F, C), i.e.,

$$F(x) = At(A^T x) - B\lambda(B^T x), \quad C := \{x \in \mathbb{R}^{55}_+ : A^T x \le b\},$$

and $b = q \cdot (1, 1, \ldots, 1)^T \in \mathbb{R}^{37\times 1}$ is the vector indicating the capacity on links with $q \in \mathbb{R}^n_+$.

Since the mapping F is monotone, the relationship $S_D = S \neq \emptyset$ clearly holds. In this example, we compare [18, Alg.1], [24, Alg.3.2] and Algorithm 3.2 by changing the values of Err and q. We use the same parameter settings as in [18, 24]. That is, we choose $\eta = 0.5,\ \sigma = 0.99$ for [18, Alg.1] and $\mu = 0.5,\ \eta = 0.4,\ \sigma = 0.99$ for [24, Alg.3.2]. For Algorithm 3.2, we choose $\mu = 0.14,\ \eta = 0.13,\ \sigma = 0.99$. We take the initial point $x^0 = (1, 1, \ldots, 1)^T$. The results are reported in Table 3.

Table 3 Results for Example 4.2

Err	*q*	[18, Alg.1]			[24, Alg.3.2]			Algorithm 3.2		
		Iter.	np	CPU(s)	Iter.	np	CPU(s)	Iter.	np	CPU(s)
10^{-3}	30	308	1130	8.6682	108	489	4.9672	115	261	2.1267
	35	313	1163	8.8352	112	531	5.5052	112	259	1.9328
	40	448	1759	13.4821	152	750	8.7453	141	335	2.5296
10^{-4}	25	376	1359	10.6346	132	596	6.9278	155	349	2.7610
	30	415	1506	11.8484	134	612	7.2465	152	345	2.7457
	35	439	1683	13.6813	147	696	8.6390	141	326	2.5656

Remark 4.2 From Table 3, one can see that Algorithm 3.2 is more efficient than [18, Alg.1] in these aspects of the number of iterations and computational time. Moreover, it can be found from Table 3 that the computational time of Algorithm 3.2 is less than that of [24, Alg.3.2] even though the number of iterations of Algorithm 3.2 is more than [24, Alg.3.2].

Example 4.3 The following quasiconvex example was presented in [33, 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 $C = \{x \in \mathbb{R}^n : f_i(x) \leq 0,\ i = 1, 2, \ldots, m,\ Ax = b\}$, where $f_0, f_1, \ldots, f_m$ are convex functions, A is a matrix of order $m \times n$, $b \in \mathbb{R}^m$. This is a quasiconvex optimization problem.

In this example, we choose

$$g(x) = \frac{\frac{1}{2}x^T Hx + q^T x + r}{\Sigma_{i=1}^n x_i},$$

where $q = (-1, -1, \ldots, -1)^T \in \mathbb{R}^n$, $r = 1 \in \mathbb{R}$, the positive diagonal matrix $H = hI$ with $h \in (0.1, 1.6)$ and I is a identity matrix. The feasible set defined by

$$C := \{x \in \mathbb{R}^n : x_i \geq 0,\ i = 1, 2, \ldots, n,\ \Sigma_{i=1}^n x_i = a\},\ a > 0.$$

It is clear that g is a smooth quasiconvex and can attain its minimum value on C (see [33, Exercise 4.7]), Hence, its derivative $F(x) = (F_1(x), F_2(x), \ldots, F_n(x))^T$ with

$$F_i(x) = \frac{hx_i \Sigma_{i=1}^n x_i - \frac{1}{2}h\Sigma_{i=1}^n {x_i}^2 - 1}{(\Sigma_{i=1}^n x_i)^2}$$

is quasimonotone on C. Moreover, it is routine to check that $S_D = \{(\frac{1}{n}a, \frac{1}{n}a, \ldots, \frac{1}{n}a)^T\}$. Hence, solving this quasiconvex problem can be converted to solve quasimonotone VI(F, C).

Table 4 Results for Example 4.3

Err	n	[13, Alg.2.1]			[21, Alg.1]			Algorithm 3.2		
		Iter.	np	CPU(s)	Iter.	np	CPU(s)	Iter.	np	CPU(s)
10^{-4}	80	439	877	29.8849	439	877	5.7914	186	186	1.2018
	100	547	1093	61.7641	548	1095	7.2515	232	232	1.5955
	120	639	1278	100.9105	640	1279	8.6622	270	270	1.9221
10^{-5}	60	380	758	17.9917	450	899	5.2114	185	185	1.1197
	80	498	995	41.9185	595	1188	7.2073	243	243	1.5377
	100	618	1235	86.8684	741	1482	9.9367	301	301	2.0839

Since this example was solved in [13, 21]. In this example, we test [13, Alg.2.1], [21, Alg.1] and Algorithm 3.2 by changing the value of Err and the dimension n. Let $h = 1.2$. We use the same parameter settings as in [13, 21]. That is, we choose $\gamma = 0.4,\ \sigma = 0.99$ for [13, Alg.2.1] and $\eta = 0.99,\ \sigma = 0.4$ for [21, Alg.1]. For Algorithm 3.2, we choose $\mu = 2.8,\ \eta = 0.4,\ \sigma = 0.99$.

We take $a = n$ and the initial point

$$x^0 := \frac{a * rand(n, 1)}{sum(rand(n, 1))},$$

where $rand(n, 1) \in \mathbb{R}^{n\times 1}$ is a random vector with each component selected from $(0, 1)$. We generate 5 random initial points as described above. We present the computational results in Table 4, averaged over the 5 random initial points.

Remark 4.3 From Table 4, one can see that Algorithm 3.2 performs better than [13, Alg.2.1] and [21, Alg.1] in the number of iterations and computational time, and Algorithm 3.2 is less affected by changing the dimension n at the same stopping criterion accuracy.

Example 4.4 The nonmonotone variational inequality problem is considered in [13, Example 4.2], where the mapping F is defined by

$$F : (x_1, x_2, \dots, x_n) \mapsto (x_1^2, x_2^2, \dots, x_n^2),$$

and the feasible set C is defined by $C = [-1, 1]^n$. It is easy to verify that the mapping F is quasimonotone on the feasible set C whenever $n = 1$ and the mapping F is not quasimonotone on the feasible set C whenever $n \geq 2$. Moreover, $S = \{(x_1, x_2, \dots, x_n)^T : x_i \in \{-1, 0\}, \forall i\},\ S_D = \{(-1, -1, \dots, -1)^T\}$.

Since this example was solved in [13, 20, 21]. In this example, we compare [13, Alg.2.1], [20, Alg.1], [21, Alg.1] and Algorithm 3.2 by changing the value of Err and the dimension n. Let $x^0 := (1, 1, \dots, 1)^T$ be the initial point. We use the same parameter settings as in [13, 20, 21]. That is, we choose $\gamma = 0.4,\ \sigma = 0.99$ for [13, Alg.2.1], $\gamma = 0.4,\ \sigma = 0.99,\ \alpha_k = \alpha_k^0$ for [20, Alg.1] and $\eta = 0.99,\ \sigma = 0.4$ for [21, Alg.1]. For Algorithm 3.2, we choose $\mu = 1.5,\ \eta = 0.9,\ \sigma = 0.7$. We present the computational results in Table 5.

Table 5 Results for Example 4.4

Err	n	[13, Alg.2.1]			[20, Alg.1]			[21, Alg.1]			Algorithm 3.2		
		Iter.	np	CPU(s)	Iter.	np	CPU(s)	Iter.	np	CPU(s)	Iter.	np	CPU(s)
10^{-4}	500	476	951	110.8641	467	933	0.9281	470	939	13.2511	382	403	0.6215
	1000	562	1123	491.8196	557	1113	2.6866	559	1117	63.4761	455	476	1.8015
	1500	619	1237	1082.0482	617	1233	6.5520	619	1237	206.8873	504	525	4.2003
10^{-5}	500	806	1611	433.1075	1489	2977	12.4591	1493	2985	61.0481	1216	1237	8.3413
	1000	971	1941	1483.5254	1771	3541	28.3333	1776	3551	219.3534	1447	1468	19.2496
	1500	1042	2083	2845.8303	1961	3921	58.2767	1965	3929	568.3477	1602	1623	38.6542

Remark 4.4 From Table 5, one can see that the number of iterations and computational time of Algorithm 3.2 performs better than [20, Alg.1] and [21, Alg.1]. Moreover, it can be found from Table 5 that the computational time and the number of projections onto C of Algorithm 3.2 are less than that of [13, Alg.2.1] even though the number of iterations of Algorithm 3.2 is more than [13, Alg.2.1] when $Err = 10^{-5}$.

5 Conclusion

In this paper, two improved subgradient extragradient algorithms are proposed for solving variational inequalities without monotonicity in a finite dimensional space $\mathbb{R}^n$. In the global convergence proof of the two new algorithms, we do not need to assume the generalized monotonicity of the mapping F. Under the mapping F is Lipschitz continuous (without the prior knowledge of Lipschitz constant) or continuous and $S_D \neq \emptyset$, the infinite sequences generated by Algorithms 3.1 and 3.2 converge globally to a solution of $\mathrm{VI}(F, C)$. Under some appropriate assumptions, we give the computational complexity analysis of the two algorithms. Finally, some numerical experiments are given to illustrate the efficiency of the two algorithms.

Acknowledgements The authors thank Professor Haisen Zhang for his valuable suggestions. The authors appreciate the valuable comments of anonymous referees which helped to improve the quality of this paper.

Funding The first two authors are supported by the National Natural Science Foundation of China (No. 12071324).

Data availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability The codes that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Consent for publication Not applicable.

Ethics approval Not applicable.

References

1. Goldstein, A.A.: Convex programming in Hilbert space. Bull. Amer. Math. Soc. **70**, 709–710 (1964)
2. Levitin, E.S., Polyak, B.T.: Constrained minimization methods. USSR Comput. Math. Math. Phys. **6**, 1–50 (1966)
3. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control. Optim. **14**, 877–898 (1976)
4. Burachik, R.S., Iusem, A.N.: A generalized proximal point algorithm for the variational inequality problem in a Hilbert space. SIAM J. Optim. **8**, 197–216 (1998)
5. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Matecon. **12**, 747–756 (1976)

6. Wang, Y.J., Xiu, N.H., Wang, C.Y.: Unified framework of extragradient-type methods for pseudomonotone variational inequalities. J. Optim. Theory Appl. **111**, 641–656 (2001)
7. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optim. **35**, 69–76 (1997)
8. Solodov, M.V., Svaiter, B.F.: A new projection method for variational inequality problems. SIAM J. Control. Optim. **37**, 765–776 (1999)
9. He, Y.R.: A new double projection algorithm for variational inequalities. J. Comput. Appl. Math. **185**, 166–173 (2006)
10. Censor, Y., Gibali, A., Reich, S.: The subgradient extragradient method for solving variational inequalites in Hilbert space. J. Optim. Theory Appl. **148**, 318–335 (2011)
11. Malitsky, Y.: Projected reflected gradient methods for monotone variational inequalities. SIAM J. Optim. **25**, 502–520 (2015)
12. Malitsky, Y.: Golden ratio algorithms for variational inequalities. Math. Program. **184**, 383–410 (2020)
13. Ye, M.L., He, Y.R.: A double projection method for solving variational inequalities without monotonicity. Comput. Optim. Appl. **60**, 141–150 (2015)
14. Brito, A.S., da Cruz Neto, J.X., Lopes, J.O., Oliveira, P.R.: Interior proximal algorithm for quasiconvex programming problems and variational inequalities with linear constraints. J. Optim. Theory Appl. **154**, 217–234 (2012)
15. Langenberg, N.: An interior proximal method for a class of quasimonotone variational inequalities. J. Optim. Theory Appl. **155**, 902–922 (2012)
16. Liu, H., Yang, J.: Weak convergence of iterative methods for solving quasimonotone variational inequalities. Comput. Optim. Appl. **77**, 491–508 (2020)
17. Ogwo, G.N., Izuchukwu, C., Shehu, Y., Mewomo, O.T.: Convergence of relaxed inertial subgradient extragradient methods for quasimonotone variational inequality problems. J. Sci. Comput. **90**, 1–35 (2022)
18. Ye, M.L.: An infeasible projection type algorithm for nonmonotone variational inequalities. Numer. Algorithms. **89**, 1723–1742 (2022)
19. Konnov, I.V.: Combined relaxation methods for finding equilibrium points and solving related problems. Russ. Math. (Iz VUZ) **37**, 44–51 (1993)
20. Lei, M., He, Y.R.: An extragradient method for solving variational inequalities without monotonicity. J. Optim. Theory Appl. **188**, 432–446 (2021)
21. Van Dinh, B., Manh, H.D., Thanh, T.T.H.: A modified Solodov-Svaiter method for solving nonmonotone variational inequality problems. Numer. Algorithms. **90**, 1715–1734 (2022)
22. Thong, D.V., Shehu, Y., Iyiola, O.S.: Weak and strong convergence theorems for solving pseudomonotone variational inequalities with non-Lipschitz mappings. Numer. Algorithms. **84**, 795–823 (2020)
23. Tseng, P.: A modified forward-backward splitting method for maximal monotone mappings. SIAM J. Control. Optim. **38**, 431–446 (2000)
24. Huang, Z.J., Zhang, Y.L., He, Y.R.: A class modified projection algorithms for nonmonotone variational inequalities with continuity. Optimization (2024). https://doi.org/10.1080/02331934.2024.2371041
25. Huang, K., Zhang, S.: Beyond monotone variational inequalities: Solution methods and iteration complexities. arXiv:2304.04153 (2023)
26. Lin, T., Jordan, M.I.: Perseus: A simple and optimal high-order method for variational inequalities. Math. Program. (2024). https://doi.org/10.1007/s10107-024-02075-2
27. Tuyen, B.T., Manh, H.D., Van Dinh, B.: Inertial algorithms for solving nonmonotone variational inequality problems. Taiwan. J. Math. **28**, 397–421 (2024)
28. Bauschke, H.H., Combettes, P.L.: Convex Analysis and Monotone Operator Theory in Hilbert Spaces, 2nd edn. Springer, New York (2017)
29. Denisov, S.V., Semenov, V.V., Chabak, L.M.: Convergence of the modified extragradient method for variational inequalities with non-Lipschitz operators. Cybernet. Syst. Anal. **51**, 757–765 (2015)
30. Marcotte, P., Zhu, D.L.: A cutting plane method for solving quasimonotone variational inequalities. Comput. Optim. Appl. **20**, 317–324 (2001)
31. Nagurney, A., Zhang, D.: Projected Dynamical Systems and Variational Inequalities with Applications. Springer, New York (1996)
32. Tu, K., Zhang, H.B., Xia, F.Q.: A new alternating projection-based prediction-correction method for structured variational inequalities. Optim. Method. Softw. **34**, 707–730 (2019)

33. Boyd, S., Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.