J. Nonlinear Var. Anal. 5 (2021), No. 4, pp. 627-644
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.5.2021.4.09

CONVERGENCE ANALYSIS OF AN INERTIAL TSENG'S EXTRAGRADIENT ALGORITHM FOR SOLVING PSEUDOMONOTONE VARIATIONAL INEQUALITIES AND APPLICATIONS

LIYA LIU[1], SUN YOUNG CHO[2,*], JEN-CHIH YAO[2]

[1]*School of Mathematics and Statistics, Southwest University, Chongqing 400715, China*
[2]*Research Center for Interneural Computing, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan*

Abstract. The purpose of this paper is to investigate the convergence analysis of an inertial Tseng's extragradient algorithm for solving pseudomonotone variational inequalities. Weak convergence theorems are established in Hilbert spaces. Simulation experiments show that our proposed algorithm outperforms some previously known algorithms.

Keywords. Inertial extrapolation; Pseudomonotone operator; Tseng's extragradient algorithm; Variational inequality; Weak convergence.

1. INTRODUCTION

Variational inequalities, which act as an efficient mathematical model, unify many significant concepts in the convex analysis, the nonsmooth analysis and the nonlinear functional analysis. They have many real applications in traffic network equilibrium modeling, economic equilibrium modeling, piece-wise-linear resistive circuits, signal processing, image recovery, pattern recognition and automatic control (see, e.g., [1, 2, 3, 4, 5]).

In recent years, much attention has been given to develop numerical methods for solving various variational inequality problems. Among them, projection-based methods have been widely studied and used in a number of situations [6, 7, 8, 9]. However, they are not efficient when the projections cannot be easily computed.

Recently, building stable and fast algorithms becomes popular and important. In 1964, Polyak [10] first proposed the inertial extrapolation, which can be viewed as an acceleration process. Since then, the inertial extrapolation has been employed to solve various convex minimization problems. Indeed, the inertial extrapolation is based on the heavy ball method of the two-order time dynamical system, and can be viewed as an efficient technique to deal with various iterative algorithms, in particular, the projection-based algorithms; see, e.g., [11, 12, 13, 14] and the references therein. Note that many convex optimization problems could be extended to pseudoconvex optimization problems. This kind of problems plays a crucial role in scientific

*Corresponding author.
E-mail addresses: liya42@qq.com (L. Liu), ooly61@hotmail.com (S.Y. Cho), yaojc@mail.cmu.edu.tw (J.C. Yao)

Received May 2, 2021; Accepted June 27, 2021.

and engineering applications, such as, financial, chemistry, corporate planning, computer vision and production planning. In addition, it was shown in [15] that a continuous operator is pseudoconvex if and only if its generalized gradient is a pseudomonotone operator. An interesting problem is how to establish an algorithm for solving pseudomonotone variational inequalites.

In this paper, we propose an efficient iterative algorithm with the inertial extrapolation for solving the variational inequality problem governed by the pseudomonotone operator. Our algorithm only needs one projection onto feasible sets at each iteration. Weak convergence theorems are established in the framework of real Hilbert spaces. Finally, we perform several numerical experiments to support the convergence of the algorithm presented in this paper. We further illustrate the computational performance of our proposed algorithm over some previously known algorithms in [16, 17, 18, 19, 20, 21, 22].

2. Preliminaries

From now on, we always assume that H is a Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\|\cdot\|$. Let Ω be a nonempty, convex, and closed set in H and let Id be the identity mapping on H. Let $\Phi : \Omega \to H$ be an operator. Recall that the following classical variational inequality problem is to find $z \in \Omega$ such that

$$\langle \Phi(z), x - z \rangle \geq 0, \quad \forall x \in \Omega. \tag{2.1}$$

In this paper, one denotes the solution set of the variational inequality by $VI(\Omega, \Phi)$.

Recall that the metric projection from H onto Ω is defined by

$$P_\Omega z := \{x \in \Omega : \|x - z\| = dist_\Omega(z)\}, \quad \forall z \in H,$$

where $dist_\Omega(z) := \inf_{x \in \Omega} \|x - z\|$. One knows that the metric projection has the following important and basic properties

(i) $P_\Omega x = z$ if and only if $\langle x - z, v - z \rangle \leq 0$, $\forall v \in \Omega$;
(ii) Given $x \in H$, $\|x - P_\Omega(x)\|^2 + \|v - P_\Omega(x)\|^2 \leq \|x - v\|^2, \quad \forall v \in \Omega$;
(iii) $\|v - v'\|^2 \geq \|(Id - P_\Omega)v - (Id - P_\Omega)v'\|^2 + \|P_\Omega v - P_\Omega v'\|^2, \quad \forall v, v' \in H$;
(iv) $\langle (Id - P_\Omega)v - (Id - P_\Omega)v', v - v' \rangle \geq \|(Id - P_\Omega)v - (Id - P_\Omega)v'\|^2, \forall v, v' \in H$;
(v) $\langle P_\Omega v - P_\Omega v', v - v' \rangle \geq \|P_\Omega v - P_\Omega v'\|^2, \quad \forall v, v' \in H$.

The projection operator plays a crucial role in the nonsmooth analysis, the convex programming and fixed point problems of nonexpansive-like mappings. One knows that z solves the problem $VI(\Omega, \Phi)$ if z satisfies the following nonlinear projection formulation

$$z = P_\Omega(z - \mu \Phi z), \quad \mu > 0. \tag{2.2}$$

Keep in mind that the well-known Brouwer's fixed point theorem guarantees that (2.2) has a solution if Ω is a bounded set, when H is a finite-dimensional space and Φ is continuous. When Ω is unbounded, some sufficient conditions for the existence of the solutions of nonlinear projection formulation (2.2) can be found in [23]. Now let us recall some related definitions and properties concerning the nonsmooth analysis, set-valued operators, and the convex analysis. They are needed for the theoretical analysis in this paper. We refer the readers to [24, 25] for more thorough discussion in details.

Definition 2.1. An operator $\Phi : H \to H$ is said to be sequentially weakly continuous iff, for each sequence $(z_k)_{k\in\mathbb{N}}$, $(z_k)_{k\in\mathbb{N}}$ converges weakly to z implies that $\Phi(z_k)$ converges weakly to $\Phi(z)$.

Definition 2.2. An operator $\Phi : H \to H$ is said to be Lipschitz continuous with a modulus $L > 0$ on the set Ω, if, for every pair of points $x, z \in \Omega$,

$$\|\Phi(x) - \Phi(z)\| \le L\|x - z\|. \tag{2.3}$$

Φ is said to be locally Lipschitz continuous on Ω if each point of Ω has a neighbourhood $C \subset \Omega$ such that (2.3) holds for every pair of points $x, z \in C$.

Definition 2.3. Suppose that $\Omega \subset H$ is a nonempty, closed, and convex set. An operator $\Phi : H \to H$ is said to be monotone on Ω if, for every pair of distinct points $x, z \in \Omega$,

$$0 \le \langle x - z, \Phi(x) - \Phi(z)\rangle.$$

Φ is said to be strictly monotone on Ω if the strict inequality holds whenever $x \neq z$. Furthermore, Φ is said to be strongly monotone on Ω with a modulus $\kappa > 0$ if, for every pair of distinct points $x, z \in \Omega$,

$$\kappa\|x - z\|^2 \le \langle x^{'} - z^{'}, x - z\rangle, \ \ \forall x^{'} \in \Phi(x),\ z^{'} \in \Phi(z).$$

Definition 2.4. [15] Let $\Omega \subset H$ be a nonempty, convex and closed set. An operator $\Phi : H \to H$ is said to be pseudomonotone on the set Ω if, for every pair of distinct points $x, z \in \Omega$,

$$\langle \Phi(x), x - z\rangle \le 0 \Rightarrow \langle \Phi(z), x - z\rangle \le 0.$$

3. Relation to the Previous Work

In 1976, Korpelevich [16] proposed an extragradient method for solving the monotone variational inequality problem in finite dimensional spaces. Let $\Phi : \Omega \to H$ be a monotone and L-Lipschitz continuous operator. Given the current iterate u_k, calculate the next iterate u_{k+1} via

$$\begin{cases} v_k = P_\Omega(Id - \alpha\Phi)u_k, \\ u_{k+1} = P_\Omega(Id - \alpha\Phi)v_k,\ k \ge 1, \end{cases}$$

where Id stands for the identity operator and $\alpha L \in (0,1)$. Korpelevich proved that $(u_n)_{n\in\mathbb{N}}$ converges weakly to a solution of variational inequality (2.1) (see [16] and the references therein). If the underlying operator is pseudomonotone, which is weaker than the monotonicity, it was shown in [26] that the extragradient method for solving problem (2.1) is also available. However, in order to obtain the next iterate u_{k+1}, there is still the need to calculate two orthogonal projections onto the feasible set Ω. Note that when the feasible set has a complex structure, it may be very expensive to calculate the orthogonal projection of one point onto Ω, which will further affects the efficiency of the extragradient method.

Recently, various modifications of the extragradient method was extensively investigated. In 2000, Tseng [17] investigated the following algorithm. Given the current iterate u_k, calculate the next iterate u_{k+1} via

$$\begin{cases} v_k = P_\Omega(Id - \alpha\Phi)u_k, \\ u_{k+1} = P_X((Id - \alpha\Phi)v_k + \alpha\Phi(u_k)),\ k \ge 1, \end{cases} \tag{3.1}$$

where Id stands for the identity operator, Φ is a monotone and L-Lipchitz continuous operator, X is some convex and closed set and α is a constant in $\left(0, \frac{1}{L}\right)$. The generated sequence $(u_k)_{k\geq 0}$ is weakly convergent in the setting of infinite dimensional Hilbert spaces.

The main advantage of Tseng's extragradient method is that the projection onto the feasible set Ω only needs to be calculated once per iteration. In 2020, Boţ, Csetnek and Vuong [18] proved that Tseng's extragradient method is weakly convergent under the assumption that the underlying operator is pseudomonotone. In this paper, based on the Tseng's extragradient method and inertial method discussed above, we propose a modified Tseng's extragradient algorithm with the inertial extrapolation for solving pseudomonotone variational inequality problems in a real Hilbert space. Our algorithm needs to calculate the projection onto the feasible set only once in each iteration. Weak convergence theorems are established under suitable conditions. Several computational experiments are carried out to demonstrate the benefit and reliability of the suggested algorithm, in comparison with some existing algorithms.

4. The Algorithm and Its Convergence

We work in the following framework,

- The feasible set Ω is a nonempty, convex and closed set in a real Hilbert space H;
- the operator $\Phi : H \to H$ is pseudomonotone, L-Lipschitz, and sequentially weakly continuous with its solution set $VI(\Omega, \Phi) \neq \emptyset$;
- $(\theta_i)_{i\in\mathbb{N}} \subseteq (0, +\infty)$, and $(\gamma_i)_{i\in\mathbb{N}} \subseteq (0, +\infty)$.

Our algorithm is designed as follows.

Algorithm 1: The Inertial Tseng's Extragradient Algorithm

Set $k \leftarrow 1$;
while *not converged* **do**
  $q_k := p_k + \theta_k(p_k - p_{k-1})$;
  if $q_k := P_\Omega(Id - \gamma_k\Phi)q_k$ **then**
    Goto final;
  else
    $u_k := P_\Omega(Id - \gamma_k\Phi)q_k$;
    $p_{k+1} = (Id - \gamma_k\Phi)u_k + \gamma_k\Phi(q_k)$;
    Set $k \leftarrow k+1$;
final;
return $p = p_k$

Now, we are in a position to state and prove our main result in details.

Theorem 4.1. *Let Ω be a nonempty, convex, and closed set in Hilbert space H. Assume that the relaxation parameters $(\gamma_k)_{k\in\mathbb{N}} \in (0, \frac{1}{3L})$, $\underline{\lim}_{k\to\infty}\gamma_k > 0$, and $\Phi(q_k) \neq 0$, for all $k \in \mathbb{N}$. In addition, assume that the sequence of extrapolation factors $(\theta_k)_{k\in\mathbb{N}}$ is nondecreasing and belongs to $[0, \theta] \subseteq [0, \sqrt{5} - 2)$. Then, for initial points p_0, $p_1 \in H$, the sequence $(p_k)_{n\in\mathbb{N}}$ generated by Algorithm 1 converges weakly to an element of $VI(\Omega, \Phi)$.*

Before proceeding with the proof of Theorem 4.1, we establish two technical lemmas first, which play a crucial role in the proof of our main result.

Lemma 4.1. [27] *Let $(\alpha_k)_{k\in\mathbb{N}}$, $(\beta_k)_{k\in\mathbb{N}}$, and $(\gamma_k)_{k\in\mathbb{N}}$ be sequences in $[0,+\infty)$ such that*

$$\alpha_{k+1} \le \alpha_k + \beta_k(\alpha_k - \alpha_{k-1}) + \gamma_k,\ \forall k \ge 1,\ \sum_{k=1}^{+\infty} \gamma_k < +\infty.$$

If there exists a real constant β such that $\beta_k \in [0,\beta] \subseteq [0,1]$, $\forall k \in \mathbb{N}$, then

(i) $\sum_{k=1}^{\infty}[\alpha_{k+1} - \alpha_k]_+ < +\infty$*, where* $[s]_+ := \max\{0,s\}$*;*

(ii) there exists $\alpha \in [0,+\infty)$ such that $\lim_{k\to+\infty} \alpha_k = \alpha$.

Lemma 4.2. [28] *Let $\Phi : \Omega \to H$ be a pseudomonotone and continuous operator. Then, $\hat{x}$ is a solution of the $VI(\Omega,\Phi)$ if and only if $\langle \Phi(x), x - \hat{x}\rangle \ge 0$, $\forall x \in \Omega$.*

Next, we prove the main result of this section.

Proof. Let $z \in VI(\Omega,\Phi)$ be arbitrarily fixed. Hence, $\langle \Phi(z), p - z\rangle \ge 0$, $\forall p \in \Omega$. By replacing $p := u_k \in \Omega$, one asserts from the pseudomonotonicity of Φ on Ω that

$$\langle \Phi(u_k), z - u_k\rangle \le 0. \tag{4.1}$$

The definition of $(u_k)_{k\in\mathbb{N}}$ entails that $\langle u_k - (q_k - \gamma_k\Phi(q_k)), z - u_k\rangle \ge 0$, which boils down to

$$\langle u_k - q_k, z - u_k\rangle \ge \gamma_k\langle \Phi(q_k), u_k - z\rangle. \tag{4.2}$$

In view of (4.1) and (4.2), one successively finds that

$$\begin{aligned}
\|p_{k+1} - z\|^2 =& \|u_k - \gamma_k(\Phi(u_k) - \Phi(q_k)) - z\|^2\\
=& \|u_k - z\|^2 + \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2 - 2\gamma_k\langle u_k - z, \Phi(u_k) - \Phi(q_k)\rangle\\
=& \|u_k - q_k\|^2 + 2\langle u_k - q_k, q_k - z\rangle + \|q_k - z\|^2 + \gamma_k^2\|\Phi u_k - \Phi q_k\|^2\\
&- 2\gamma_k\langle u_k - z, \Phi(u_k) - \Phi(q_k)\rangle\\
=& \|u_k - q_k\|^2 + 2\langle u_k - q_k, q_k - u_k\rangle + 2\langle u_k - q_k, u_k - z\rangle + \|q_k - z\|^2\\
&+ \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2 - 2\gamma_k\langle u_k - z, \Phi(u_k) - \Phi(q_k)\rangle\\
\le& - \|u_k - q_k\|^2 - 2\gamma_k\langle \Phi q_k, u_k - z\rangle + \|q_k - z\|^2 + \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2\\
&- 2\gamma_k\langle u_k - z, \Phi(u_k) - \Phi(q_k)\rangle\\
\le& - \|u_k - q_k\|^2 + \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2 + \|q_k - z\|^2 - 2\gamma_k\langle u_k - z, \Phi(u_k)\rangle\\
\le& - \|u_k - q_k\|^2 + \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2 + \|q_k - z\|^2.
\end{aligned} \tag{4.3}$$

By the Lipschitz continuity of Φ, one has

$$\begin{aligned}
-\|u_k - q_k\|^2 + \gamma_k^2\|\Phi(u_k) - \Phi(q_k)\|^2 \le& - \|u_k - q_k\|^2 + \gamma_k^2 L^2\|u_k - q_k\|^2\\
\le& - (1 - \gamma_k^2 L^2)\|u_k - q_k\|^2.
\end{aligned} \tag{4.4}$$

Taking account into (4.3) and (4.4), one obtains that

$$\|p_{k+1} - z\|^2 \le - (1 - \gamma_k^2 L^2)\|u_k - q_k\|^2 + \|q_k - z\|^2. \tag{4.5}$$

On the other hand, one has

$$\begin{aligned}
\|p_{k+1} - q_k\| \le& \|p_{k+1} - u_k\| + \|u_k - q_k\|\\
=& \|u_k - \gamma_k(\Phi(u_k) - \Phi(q_k)) - u_k\| + \|u_k - q_k\|\\
\le& (1 + \gamma_k L)\|u_k - q_k\|,
\end{aligned} \tag{4.6}$$

which further yields that

$$-\|u_k - q_k\|^2 \le -\frac{1}{(1+\gamma_k L)^2}\|p_{k+1} - q_k\|^2. \tag{4.7}$$

From the fact that $\gamma_k \in \left(0, \frac{1}{3L}\right)$, we arrive at

$$\frac{1-\gamma_k L}{1+\gamma_k L} \ge \frac{1}{2}. \tag{4.8}$$

By substituting (4.7) and (4.8) into (4.5), we find that

$$\|p_{k+1} - z\|^2 \le \|q_k - z\|^2 - \frac{1-\gamma_k L}{1+\gamma_k L}\|p_{k+1} - q_k\|^2 \le \|q_k - z\|^2 - \frac{1}{2}\|p_{k+1} - q_k\|^2. \tag{4.9}$$

Using the definition of $(q_k)_{k\in\mathbb{N}}$, we infer that

$$\begin{aligned}\|q_k - z\|^2 =& \|(p_k + \theta_k(p_k - p_{k-1})) - z\|^2 \\ =& \|(1+\theta_k)(p_k - z) - \theta_k(p_{k-1} - z)\|^2 \\ =& (1+\theta_k)\|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 + (1+\theta_k)\theta_k\|p_k - p_{k-1}\|^2,\end{aligned} \tag{4.10}$$

and

$$\begin{aligned}\|p_{k+1} - q_k\|^2 =& \|p_{k+1} - (p_k + \theta_k(p_k - p_{k-1}))\|^2 \\ \ge& \|p_{k+1} - p_k\|^2 - 2\theta_k\|p_{k+1} - p_k\|\|p_k - p_{k-1}\| + \theta_k^2\|p_k - p_{k-1}\|^2 \\ \ge& (1-\theta_k)\|p_{k+1} - p_k\|^2 + (\theta_k^2 - \theta_k)\|p_k - p_{k-1}\|^2.\end{aligned} \tag{4.11}$$

Plugging (4.10), (4.11) into (4.9) and reorganizing, we immediately obtain that

$$\begin{aligned}\|p_{k+1} - z\|^2 \le& (1+\theta_k)\|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 + (1+\theta_k)\theta_k\|p_k - p_{k-1}\|^2 \\ & - \frac{1}{2}((1-\theta_k)\|p_{k+1} - p_k\|^2 + (\theta_k^2 - \theta_k)\|p_k - p_{k-1}\|^2) \\ \le& (1+\theta_k)\|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 \\ & - \frac{1}{2}(1-\theta_k)\|p_{k+1} - p_k\|^2 + \left((1+\theta_k)\theta_k - \frac{1}{2}(\theta_k^2 - \theta_k)\right)\|p_k - p_{k-1}\|^2 \\ \le& (1+\theta_k)\|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 \\ & - \frac{1}{2}(1-\theta_k)\|p_{k+1} - p_k\|^2 + \left(\frac{1}{2}\theta_k^2 + \frac{3}{2}\theta_k\right)\|p_k - p_{k-1}\|^2 \\ \le& (1+\theta_k)\|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 - \alpha_k\|p_{k+1} - p_k\|^2 \\ & + \beta_k\|p_k - p_{k-1}\|^2,\end{aligned} \tag{4.12}$$

where $\alpha_k = \frac{1}{2}(1-\theta_k)$ and $\beta_k = \frac{1}{2}\theta_k^2 + \frac{3}{2}\theta_k$. Set

$$\Upsilon_k = \|p_k - z\|^2 - \theta_k\|p_{k-1} - z\|^2 + \beta_k\|p_k - p_{k-1}\|^2. \tag{4.13}$$

Since $(\theta_k)_{k\in\mathbb{N}}$ is nondecreasing, we find from (4.12) that

$$\begin{aligned}\Upsilon_{k+1}-\Upsilon_k &= \|p_{k+1}-z\|^2-(1+\theta_{k+1})\|p_k-z\|^2+\theta_k\|p_{k-1}-z\|^2\\ &\quad+\beta_{k+1}\|p_{k+1}-p_k\|^2-\beta_k\|p_k-p_{k-1}\|^2\\ &\le \|p_{k+1}-z\|^2-(1+\theta_k)\|p_k-z\|^2+\theta_k\|p_{k-1}-z\|^2\\ &\quad+\beta_{k+1}\|p_{k+1}-p_k\|^2-\beta_k\|p_k-p_{k-1}\|^2\\ &\le -(\alpha_k-\beta_{k+1})\|p_{k+1}-p_k\|^2.\end{aligned} \tag{4.14}$$

Keeping in mind that $(\theta_k)_{k\in\mathbb{N}}$ is nondecreasing and $(\theta_k)_{k\in\mathbb{N}}\in[0,\theta]\subseteq[0,\sqrt{5}-2)$, we obtain

$$\begin{aligned}\alpha_k-\beta_{k+1} &= \frac{1}{2}(1-\theta_k)-\left(\frac{1}{2}\theta_{k+1}^2+\frac{3}{2}\theta_{k+1}\right)\\ &\ge \frac{1}{2}(1-\theta_{k+1})-\frac{1}{2}\theta_{k+1}^2-\frac{3}{2}\theta_{k+1}\\ &= \frac{1}{2}-2\theta_{k+1}-\frac{1}{2}\theta_{k+1}^2\\ &\ge \frac{1-4\theta-\theta^2}{2}>0.\end{aligned} \tag{4.15}$$

Let $\lambda=\frac{1-4\theta-\theta^2}{2}$. (4.14) and (4.15) send us to

$$0\le\lambda\|p_{k+1}-p_k\|^2\le\Upsilon_k-\Upsilon_{k+1}, \tag{4.16}$$

which implies that $\Upsilon_k-\Upsilon_{k+1}\ge 0$. Hence, $(\Upsilon_k)_{k=k_0}^{\infty}$ is nonincreasing. Coming back to (4.13), an immediate recurrence shows that, for every $k\ge 1$,

$$\begin{aligned}\|p_k-z\|^2 &=\theta_k\|p_{k-1}-z\|^2-\beta_k\|p_k-p_{k-1}\|^2+\Upsilon_k\\ &\le\theta_k\|p_{k-1}-z\|^2+\Upsilon_k\\ &\le\theta\|p_{k-1}-z\|^2+\Upsilon_{k_0}\\ &\le\cdots\le\theta^{k-k_0}\|p_{k_0}-z\|^2+\Upsilon_{k_0}(1+\cdots+\theta^{k-k_0-1})\\ &\le\theta^{k-k_0}\|p_{k_0}-z\|^2+\frac{\Upsilon_{k_0}}{1-\theta},\end{aligned} \tag{4.17}$$

where the integer k_0 belongs to $(0,k]$. Combining (4.13) with (4.17), we observe that

$$\begin{aligned}-\Upsilon_{k+1} &\le\theta_{k+1}\|p_k-z\|^2\le\theta\|p_k-z\|^2\\ &\le\theta^{k-k_0+1}\|p_{k_0}-z\|^2+\frac{\theta\Upsilon_{k_0}}{1-\theta}\le\|p_{k_0}-z\|^2+\frac{\theta\Upsilon_{k_0}}{1-\theta}.\end{aligned} \tag{4.18}$$

By summing the above inequality (4.16), together with (4.18), we obtain that

$$\lambda\sum_{i=k_0}^{k}\|p_{i+1}-p_i\|^2\le\Upsilon_{k_0}-\Upsilon_{k+1}\le\|p_{k_0}-z\|^2+\frac{\Upsilon_{k_0}}{1-\theta},$$

which guarantees the summability condition

$$\sum_{k=1}^{\infty}\|p_{k+1}-p_k\|^2<+\infty,$$

and hence

$$\lim_{k\to\infty}\|p_{k+1}-p_k\|=0. \tag{4.19}$$

The definition of $(q_k)_{k\in\mathbb{N}}$ ensures that

$$\|p_{k+1}-q_k\|\le\|p_{k+1}-p_k\|+\theta_k^2\|p_k-p_{k-1}\|^2-2\theta_k\langle p_{k+1}-p_k,p_k-p_{k-1}\rangle,$$

which amounts from (4.19) to

$$\lim_{k\to\infty}\|p_{k+1}-q_k\|=0. \tag{4.20}$$

Taking into (4.19) and (4.20) account, we immediately obtain that

$$\lim_{k\to\infty}\|p_k-q_k\|\le\lim_{k\to\infty}\|p_k-p_{k+1}\|+\lim_{k\to\infty}\|p_{k+1}-q_k\|=0. \tag{4.21}$$

Using (4.9) and (4.10) and $(\theta_k)_{k\in\mathbb{N}}\in[0,\theta]\subseteq[0,\sqrt{5}-2)$, we find that

$$\begin{aligned}\|p_{k+1}-z\|^2\le{}&\|p_k-z\|^2+\theta_k(\|p_k-z\|^2-\|p_{k-1}-z\|^2)\\&+(1+\theta_k)\theta_k\|p_k-p_{k-1}\|^2\\\le{}&\|p_k-z\|^2+\theta_k(\|p_k-z\|^2-\|p_{k-1}-z\|^2)+2\theta\|p_k-p_{k-1}\|^2.\end{aligned}$$

In light of Lemma 4.1, we have that there exists $w>0$ such that

$$\lim_{k\to\infty}\|p_k-z\|=w. \tag{4.22}$$

Combining (4.10) with (4.19), we obtain that

$$\lim_{k\to\infty}\|q_k-z\|=w. \tag{4.23}$$

Taking into account (4.5), we have

$$(1-\gamma_k^2L^2)\|u_k-q_k\|^2\le\|q_k-z\|^2-\|p_{k+1}-z\|^2,$$

which together with (4.22) and (4.23) yields

$$\lim_{k\to\infty}\|u_k-q_k\|=0. \tag{4.24}$$

From (4.23), one asserts that $(q_k)_{k\in\mathbb{N}}$ is bounded, which further implies that there exists a subsequence $(q_{k_i})_{i\in\mathbb{N}}$ of $(q_k)_{k\in\mathbb{N}}$ converges weakly to an element p^*. Combining (4.21) and (4.24), we have that $(p_{k_i})_{i\in\mathbb{N}}$ and $(u_{k_i})_{i\in\mathbb{N}}$ also converge weakly to p^*. In light of $(u_{k_i})_{i\in\mathbb{N}}\in\Omega$, one implies $p^*\in\Omega$. Now we are in a position to prove that $p^*\in VI(\Omega,F)$. Let $z\in\Omega$ be fixed. For every $i\ge0$, one has

$$\begin{aligned}&\langle u_{k_i}-(q_{k_i}-\gamma_{k_i}\Phi(q_{k_i})),z-u_{k_i}\rangle\ge0\\\Leftrightarrow&\langle\Phi(q_{k_i}),z-q_{k_i}\rangle\ge\frac{1}{\gamma_{k_i}}\langle q_{k_i}-u_{k_i},z-u_{k_i}\rangle+\langle\Phi(q_{k_i}),u_{k_i}-q_{k_i}\rangle.\end{aligned} \tag{4.25}$$

Thanks to (4.24) and (4.25), one obtains that

$$\liminf_{i\to\infty}\langle\Phi(q_{k_i}),z-q_{k_i}\rangle\ge0.$$

Now we choose a sequence $(\zeta_i)_{i\in\mathbb{N}}$ of positive numbers decreasing and tending to 0. For each ζ_i, we denote by k_{j_i} the smallest positive integer from $(k_j)_{j\in\mathbb{N}}$ such that

$$\langle\Phi(q_{k_{j_i}}),z-q_{k_{j_i}}\rangle+\zeta_i\ge0,\ \ \forall i\ge0. \tag{4.26}$$

Since $(\zeta_i)_{i\in\mathbb{N}}$ is decreasing, one finds that $(k_{j_i})_{i\in\mathbb{N}}$ is increasing. The rest of the proof will be divided into two parts

Case 1. Suppose that there exists a subsequence of positive integers $(\phi(i))_{i\in\mathbb{N}}$ such that

$$\phi(i)=\min\{l\in\mathbb{N}\,|\,l\geq i \text{ and } \Phi(q_{k_{j_l}})\neq 0\}.$$

It is obviously seen that $(\phi(i))_{i\in\mathbb{N}}$ is increasing. Recalling that $(k_{j_i})_{i\in\mathbb{N}}$ is increasing, which together with the definition of $(\phi(i))_{i\in\mathbb{N}}$ yields that $(k_{j_{\phi(i)}})_{i\in\mathbb{N}}$ is also increasing. Set

$$\tau_{k_{j_{\phi(i)}}}=\frac{\Phi(q_{k_{j_{\phi(i)}}})}{\|\Phi(q_{k_{j_{\phi(i)}}})\|^2}.$$

As a classical result, we check that $\langle\Phi(q_{k_{j_{\phi(i)}}}),q_{k_{j_{\phi(i)}}}\rangle=1$ for each i. Hence (4.26) asserts that, for each i,

$$\langle\Phi(q_{k_{j_{\phi(i)}}}),z+\zeta_i\tau_{k_{j_{\phi(i)}}}-q_{k_{j_{\phi(i)}}}\rangle\geq 0. \tag{4.27}$$

Returning to (4.27), one obtains from the fact that Φ is pseudomonotone that

$$\langle\Phi(z+\zeta_i\tau_{k_{j_{\phi(i)}}}),z+\zeta_i\tau_{k_{j_{\phi(i)}}}-q_{k_{j_{\phi(i)}}}\rangle\geq 0. \tag{4.28}$$

On the other hand, we have that $(q_{k_j})_{j\in\mathbb{N}}$ converges weakly to p^* as $j\to\infty$. We deduce that $(q_{k_{j_{\phi(i)}}})_{i\in\mathbb{N}}$ converges weakly to p^* as $i\to\infty$. Applying the fact that Φ is sequentially weakly continuous on Ω, one sees that $(\Phi(q_{k_{j_{\phi(i)}}}))_{i\in\mathbb{N}}$ converges weakly to $\Phi(p^*)$. We assume that $\Phi(p^*)\neq 0$ (otherwise, p^* is a solution). Since the norm mapping is sequentially weakly lower semicontinuous, one concludes

$$\|\Phi(p^*)\|\leq\liminf_{i\to\infty}\|\Phi(q_{k_{j_{\phi(i)}}})\|.$$

In view of $\zeta_i\to 0$ as $i\to\infty$, one immediately obtains

$$0\leq\lim_{i\to\infty}\|\zeta_i\tau_{k_{j_{\phi(i)}}}\|=\liminf_{i\to\infty}\frac{\zeta_i}{\|\Phi(q_{k_{j_{\phi(i)}}})\|}\leq\frac{0}{\|\Phi(p^*)\|}=0. \tag{4.29}$$

Combining (4.28) with (4.29), one has

$$\langle\Phi(z),z-p^*\rangle\geq 0.$$

It follows from Lemma 4.2 that $p^*\in VI(\Omega,\Phi)$.

Finally, it only remains to prove that the sequence $(p_k)_{k\in\mathbb{N}}$ converges weakly to p^*. To this end, it is sufficient to show that $(p_k)_{k\in\mathbb{N}}$ cannot have two distinct weak sequential cluster points in $VI(\Omega,\Phi)$. Let $(p_{k_m})_{m\in\mathbb{N}}$ be another subsequence of $(p_k)_{k\in\mathbb{N}}$ converging weakly to $\hat{p}$ as m goes to $+\infty$. Next, one focuses on the proof of $p^*=\hat{p}$. As it has been proven above, one infers that $\hat{p}\in VI(\Omega,\Phi)$, which together with (4.22) yields that, for all $n\in\mathbb{N}$,

$$2\langle p_k,\hat{p}-p^*\rangle=\|p_k-p^*\|^2-\|p_k-\hat{p}\|^2+\|\hat{p}\|^2-\|p^*\|^2.$$

Hence, we deduce that the sequence $(\langle p_k,\hat{p}-p^*\rangle)_{k\in\mathbb{N}}$ also converges. Setting $\lim_{k\to\infty}\langle p_k,\hat{p}-p^*\rangle=\sigma$, and passing to the limit along $(p_{k_i})_{i\in\mathbb{N}}$ and $(p_{k_j})_{j\in\mathbb{N}}$, one concludes that $\langle p^*,\hat{p}-p^*\rangle=\langle\hat{p},\hat{p}-p^*\rangle=\sigma$. So, $\|\hat{p}-p^*\|^2=0$, that is, $\hat{p}=p^*$.

Case 2. Otherwise, we have $\lim_{i\to\infty}\Phi(p_{k_i})=0$, which together with $p_{k_i}\rightharpoonup p^*$, ensures that

$$\Phi(p^*)=\lim_{i\to\infty}\Phi(p_{k_i}).$$

Hence $p^* \in \Phi^{-1}(0)$, which further implies $p^* \in VI(\Omega, \Phi)$. This completes the proof. □

Let us now state a particular case of Theorem 4.1 as an immediate consequence of the monotonicity of Φ.

Remark 4.1. When working with a monotone operator Φ, it is not necessary to impose the sequential weak continuity on Φ. In view of this, the monotonicity of Φ, which together with (4.2), implies that

$$\begin{aligned}\gamma_{k_i}\langle \Phi(z), z - q_{k_i}\rangle \geq& \gamma_{k_i}\langle \Phi(q_{k_i}), z - q_{k_i}\rangle \\ \geq& \langle u_{k_i} - q_{k_i}, u_{k_i} - z\rangle + \gamma_{k_i}\langle \Phi(q_{k_i}), u_{k_i} - q_{k_i}\rangle.\end{aligned}$$

Letting i tend to $+\infty$ in (4.28) and keeping $\lim_{k\to\infty}\|u_k - q_k\| = 0$ and $\underline{\lim}_{k\to\infty}\gamma_k > 0$ in mind, one has

$$\langle \Phi(z), z - p^*\rangle \geq 0, \quad \forall z \in \Omega,$$

which leads to the desired conclusion.

The shortcoming of Algorithm 1 is a requirement to estimate the Lipschitz constant more or less precisely. While the estimation is often quite conservative. Of course, this is not practical in most cases of interest. For this reason, we give a prediction of a stepsize with its further correction along a feasible direction. Now let us state the result below.

Theorem 4.2. *Suppose that (an upper bounded of) the Lipschitz constant of Φ is unknown. We deal with the relaxation parameters $(\gamma_k)_{k\in\mathbb{N}}$ in Algorithm 1 by using the following adaptive stepsize strategy*

$$\gamma_{k+1} = \begin{cases} \gamma_k, & \text{if } \Phi(u_k) - \Phi(q_k) = 0; \\ \min\left\{\frac{\eta\|u_k - q_k\|}{\|\Phi(u_k) - \Phi(q_k)\|}, \gamma_k\right\}, & \text{otherwise}, \end{cases} \tag{4.30}$$

where $\gamma_0 > 0$ and $\eta \in \left(0, \frac{1}{3}\right]$. Then the conclusion of Theorem 4.1 still remains valid by using adaptive stepsize (4.30).

Proof. In such a case, we need to proceed in much the same way as in Theorem 4.1. To avoid the possible repetition, we restrict our attention to the place where arguments differ. In the upcoming statement, the assumption of the relaxation parameters in Theorem 4.1 will be removed and replaced with the adaptive stepsize (4.30) in inequalities (4.4)-(4.9). From the definition, we obtain that $(\gamma_k)_{k\in\mathbb{N}}$ is nonincreasing. Additionally, if $\Phi(p_k) - \Phi(q_k) \neq 0$, where $k \geq 0$, then

$$\frac{\eta\|u_k - q_k\|}{\|\Phi(u_k) - \Phi(q_k)\|} \geq \frac{\eta\|u_k - q_k\|}{L\|u_k - q_k\|} = \frac{\eta}{L},$$

which shows from $\min\{\gamma_0, \frac{\eta}{L}\} > 0$ that $(\gamma_k)_{k\in\mathbb{N}}$ is bounded from below. It ensues that

$$(\gamma_k)_{k\in\mathbb{N}} \in \left(0, \min\{\gamma_0, \frac{\eta}{L}\}\right).$$

Notice that, if $\gamma_0 \leq \frac{\eta}{L}$, then $(\gamma_k)_{k\in\mathbb{N}}$ is a constant sequence, which leads to a fixed stepsize strategy. As a classical result, the limit $\lim_{k\to\infty}\gamma_k$ exists and it is a positive real number. Letting

$\chi_k = \frac{\gamma_k}{\gamma_{k+1}}$, it is immediate to check that $\lim_{k\to\infty}\chi_k = 1$. Let us rewrite inequality (4.4) as

$$\begin{aligned} &-\|u_k-q_k\|^2+\gamma_k^2\|\Phi(u_k)-\Phi(q_k)\|^2 \\ =&-\|u_k-q_k\|^2+\chi_k^2\gamma_{k+1}^2\|\Phi(u_k)-\Phi(q_k)\|^2 \\ =&-\|u_k-q_k\|^2+\left(\frac{\eta\|u_k-q_k\|}{\|\Phi(u_k)-\Phi(q_k)\|}\right)^2\chi_k^2\|\Phi(u_k)-\Phi(q_k)\|^2 \\ =&-(1-\eta^2\chi_k^2)\|u_k-q_k\|^2. \end{aligned} \tag{4.31}$$

Coming back to (4.6), which together with (4.30), infers that

$$\begin{aligned} \|p_{k+1}-q_k\| \le& \|p_{k+1}-u_k\|+\|u_k-q_k\| \\ =&\|u_k-\gamma_k(\Phi(u_k)-\Phi(q_k))-u_k\|+\|u_k-q_k\| \\ =&\gamma_k\|\Phi(u_k)-\Phi(q_k)\|+\|u_k-q_k\| \\ =&\chi_k\gamma_{k+1}\|\Phi(u_k)-\Phi(q_k)\|+\|u_k-q_k\| \\ =&\chi_k\left(\frac{\eta\|u_k-q_k\|}{\|\Phi(u_k)-\Phi(q_k)\|}\right)\|\Phi(u_k)-\Phi(q_k)\|+\|u_k-q_k\| \\ =&\eta\chi_k\|u_k-q_k\|+\|u_k-q_k\| \\ =&(1+\eta\chi_k)\|u_k-q_k\|. \end{aligned}$$

This further guarantees

$$-\|u_k-q_k\|^2 \le -\frac{1}{(1+\eta\chi_k)^2}\|p_{k+1}-q_k\|^2. \tag{4.32}$$

Thanks to $\lim_{k\to\infty}\chi_k = 1$, one deduces that $\lim_{k\to+\infty}(1-\eta^2\chi_k^2) = 1-\eta^2 > 0$. On the other hand, the condition $\gamma_k \in \left(0, \frac{1}{3L}\right]$ is guaranteed by the fact that $\eta \in \left(0, \frac{1}{3}\right]$. Hence $\frac{1-\eta\chi_k}{1+\eta\chi_k} \ge \frac{1}{2}$. By substituting (4.31), (4.32) into (4.3), one finds that

$$\begin{aligned} \|p_{k+1}-z\|^2 \le& \|q_k-z\|^2-\frac{1-\eta^2\chi_k^2}{(1+\eta\chi_k)^2}\|p_{k+1}-q_k\|^2 \\ =&\|q_k-z\|^2-\frac{1-\eta\chi_k}{1+\eta\chi_k}\|p_{k+1}-q_k\|^2 \\ \le&\|q_k-z\|^2-\frac{1}{2}\|p_{k+1}-q_k\|^2. \end{aligned}$$

The proof of the weak convergence of the sequence $(p_k)_{k\in\mathbb{N}}$ follows the same lines as in Theorem 4.1. So, it is omitted here. □

5. Applications

In this section, we present several illustrative numerical examples to demonstrate the performance, efficiency and applicability of the proposed algorithm. In all tests, we use our proposed algorithm to solve problems by letting the extrapolation factors $\theta_k = 0.23$ $(k \ge 0)$ and the relaxation parameters $\gamma_k = 0.01$ $(k \ge 0)$. All the experiments are performed on a PC with Intel (R) Core (TM) i5-8250U CPU @1.60GHz, under the Matlab computing environment.

Now we consider the following general nonsmooth optimization problem

$$\text{minimize } \phi(p), \quad \text{subject to } \Theta p = b, \; p \in \Omega, \tag{5.1}$$

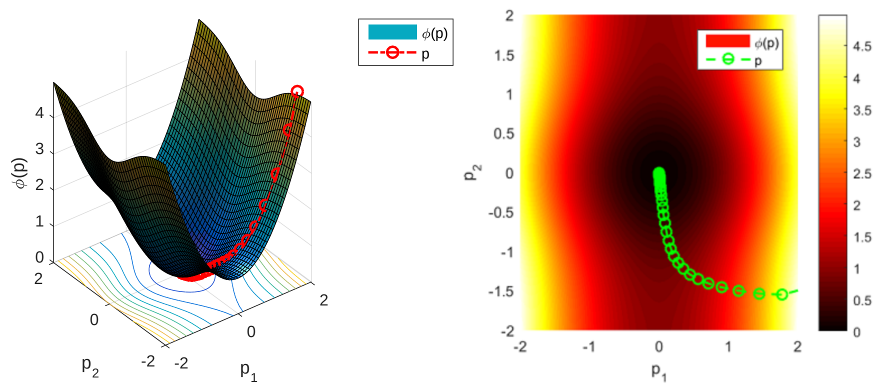

FIGURE 1. Isometric view of the objective function with the behavior of $(p_1, p_2, \phi(p))^T$, and require 50 iterations in Example 5.1.

where $\phi : \mathbb{R}^n \to \mathbb{R}$ is an objective function, $p = (p_1, p_2, \cdots, p_n)^T \in \mathbb{R}^n$, $b = (b_1, b_2, \cdots, b_m)^T \in \mathbb{R}^m$, $\Theta \in \mathbb{R}^{m \times n}$ is a full row-rank matrix (i.e., $rank(\Theta) = m \leq n$) and Ω is a convex and closed set in $\mathbb{R}^n$. In our experiments, the objective function of problem (5.1) is not necessarily convex everywhere. Indeed, it only needs to be pseudoconvex on a set defined by the constraints. We remark here that many functions in nature are pseudoconvex, such as fractional functions, Butterworth filter functions and some density functions in probability theory. As for problem (5.1), our proposed algorithm is effective and applicable in Example 5.1 and Example 5.2.

Example 5.1. Consider the following nonlinear optimization problem via

$$\text{minimize } \phi(p) = 1 + p_1^2 - e^{-p_2^2}, \text{ subject to } p \in \Omega, \tag{5.2}$$

where $p = (p_1, p_2)^T$ and $\Omega = \{p \in \mathbb{R}^2 | -2 \leq p_1, p_2 \leq 2\}$. It is easy to check that the objective function $\phi(p)$ is convex on Ω. Hence, the generalized gradient $\Phi(p) := \nabla \phi(p) = (2p_1, 2p_2 e^{-p_2^2})^T$, $\forall p = (p_1, p_2)^T$ is pseudomonotone. In the following experiment, we choose the starting data in the range of $(0,1)^2$ randomly and take the iteration number $k = 50$ as the termination criterion.

Figure 1 depicts the isometric view of $\phi(p)$ respectively in a 3-D space and a 2-D space. Obviously, $(0,0,0)^T$ is a convergent point of this convex optimization problem, which is also the global minimum solution of the objective function in the whole space. The CPU time to compute the 50-th value is about 1.963199 s.

In the second experiment, we randomly choose many starting points in the range of $(0,1)^2$ and take the iteration number $k = 80$ as the termination criterion. In Figure 2, the left figure plots the changing processes of $(p_1, p_2)^T$ with the number of iterations k (x-axis), the right figure plots the changing processes of $(p_1, p_2)^T$ when the execution time in second elapses (x-axis).

To illustrate the computational performance, Figure 3 depicts the isometric view of $\phi(p)$ in a 2-D space with 20 initial points at star symbol and the unique convergent point at dot symbol. Obviously, the convergence of the values of $(p_1, p_2)^T$ to $(0,0)^T$, means that the convergent point $(0,0)^T$ is the optimal solution of problem (5.2), when $\phi(p) = 0$.

Example 5.2. Consider minimizing the condition number of a nonzero matrix A as follows

$$\begin{aligned} &\text{minimize } \chi(\Theta) \\ &\text{subject to } \Theta \in \Omega, \end{aligned}$$

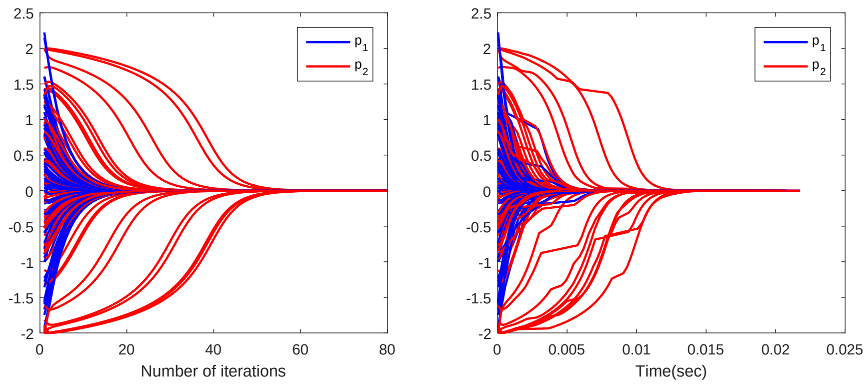

FIGURE 2. Behaviors of p_1, p_2 with many initial points. Numerical results for Algorithm 1.

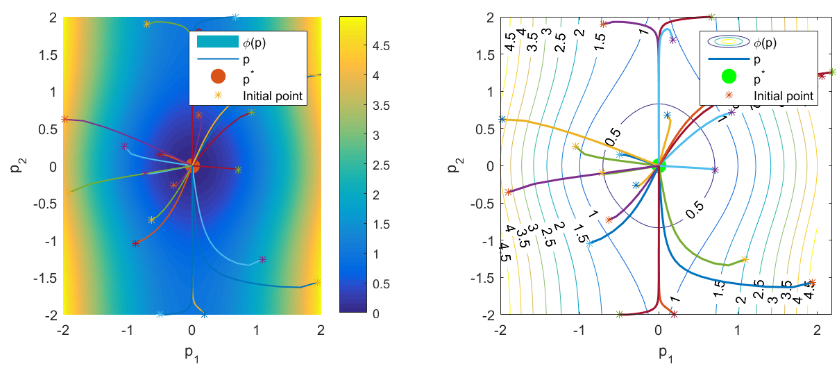

FIGURE 3. Behaviors of $(p_1, p_2, \phi(p))^T$ in terms of 20 random initial points, and require $k = 80$ iterations in Example 5.1.

where Ω is a compact convex set in $\mathbb{R}^{m\times m}$, $\Theta \in \mathbb{R}^{m\times m}$ is a symmetric matrix and the condition number $\chi(\Theta)$ is defined as $\chi(\Theta) = \frac{\delta_{\max(\Theta)}}{\delta_{\min(\Theta)}}$, in which $\delta_{\min(\Theta)}$ and $\delta_{\max(\Theta)}$ denote the minimum and maximum eigenvalues of the matrix Θ, respectively. When $m = 4$, we consider the diagonal matrix Θ defined by

$$\Theta = \begin{pmatrix} \rho^T p + \rho_0 & 0 \\ 0 & \sigma^T p + \sigma_0 \end{pmatrix},$$

where $p = (p_1, p_2, p_3, p_4)^T \in \mathbb{R}^4$, $\Omega = \{p \in \mathbb{R}^4 | 0 \leq p_i \leq 1,\ i = 1,2,3,4\}$, $\rho = (-2,-1,2,0)^T$, $\sigma = (1,-1,2,1)^T$, $\rho_0 = 4$ and $\sigma_0 = 2$. One has that the condition number $\chi(\Theta)$ is pseudoconvex with respect to $p \in \Omega$, see [29], which can be expressed as

$$\chi(\Theta) = \begin{cases} \frac{\rho^T p + \rho_0}{\sigma^T p + \sigma_0}, & \text{if } \sigma^T p + \sigma_0 \leq \rho^T p + \rho_0, \\ \frac{\sigma^T p + \sigma_0}{\rho^T p + \rho_0}, & \text{if } \sigma^T p + \sigma_0 > \rho^T p + \rho_0. \end{cases}$$

Then its gradient $\Phi(p) := \nabla \chi(\Theta)$ is pseudomonotone with respect to $p \in \Omega$. Thus we can use the condition number $\chi(\Theta)$ to study the convergence and computational performance of Algorithm 1. We study this problem for many different choices of starting points randomly generated in the range of $(0,1)^4$. Meanwhile, we take the number of iterations $k = 70$ as the

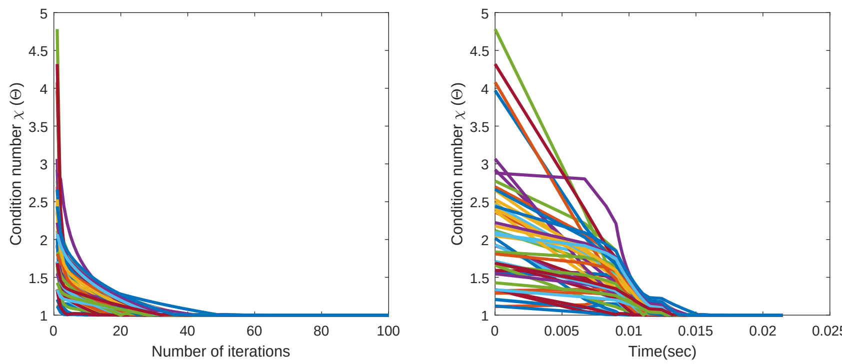

FIGURE 4. Behavior of condition number $\chi(\Theta)$ with the number of iterations (resp.) $k = 100$. Numerical results for Algorithms 1.

stopping criterion. From the results reported in Figure 4, we find that the convergent point of the condition number $\chi(\Theta)$ is 1.

Example 5.3. Consider the variational inequality with the following property

$$\Phi(p) = \Upsilon(p) + \Theta p + \omega, \tag{5.3}$$

where

$$\Upsilon(p) = \begin{pmatrix} \arctan(p_1) \\ \arctan(p_2) \\ \vdots \\ \arctan(p_m) \end{pmatrix}_{m\times 1}, \Theta = \begin{pmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & \vdots \\ 0 & -1 & 2 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 2 \end{pmatrix}_{m\times m},$$

and

$$\omega = \begin{pmatrix} -\frac{m}{2}+1 \\ -\frac{m}{2}+2 \\ \vdots \\ -\frac{m}{2}+m \end{pmatrix}_{m\times 1}.$$

Note that $\Phi(\cdot)$ is a strongly monotone and Lipschitz continuous operator, which means that it is a pseudomonotone operator on $\Omega = \{p \in \mathbb{R}^m : p \geq 0\}$. The problem is to find a point $p^* \in \Omega$ such that $\langle \Phi p^*, p - p^* \rangle \geq 0$, $\forall p \in \Omega$. Let us take $m = 8$ as a special case and denote $p = (p_1, p_2, \cdots, p_8)^T$.

We make a comparison of the behavior of the term $(\|p_k\|)_{k\in\mathbb{N}}$ (the y-axis) with respect to different θ, where p_k is generated by Algorithm 1. In this experiment, we fix the same initial data chosen randomly in the range of $(0,1)^8$ and we take the number of iterations $k = 600$ as the stopping criterion. We observed from the test result reported in Figure 5 and Table 1 that the bigger θ is, the less the required computer time becomes, the faster the convergence rate becomes. Moreover, it can be observed from the plots that the changing processes in all cases with nonzero extrapolation factors outperform the case without the inertial term as $\theta = 0$, and hence meeting the result drawn from the theoretical analysis. The effect of the acceleration

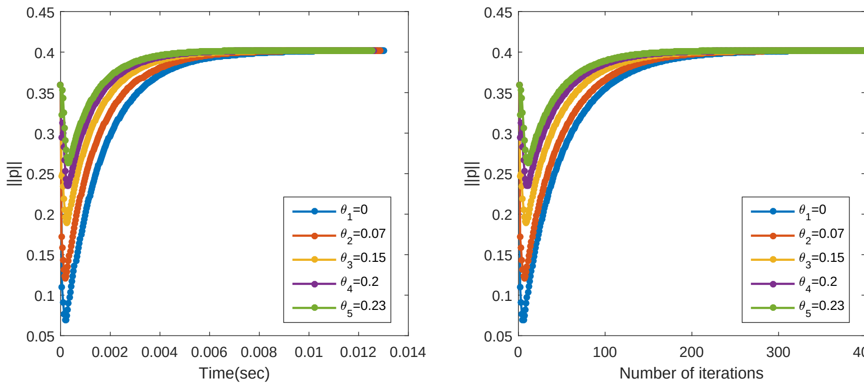

FIGURE 5. Behaviors of $(\|p_k\|)_{k\in\mathbb{N}}$ with the number of iterations (resp.) $k = 400$. Numerical results with different $\theta := 0, 0.07, 0.15, 0.2, 0.23$ for Algorithm 1. The convergence of p_k to $p^* = (0.3815, 0.1274, 0, 0)^T$.

TABLE 1. Comparison results between proposed Algorithm 1, EGM, HSEM, MEGM, Y-EAPM, L-EAPM, and SEGM.

Algorithm 1	$\theta = 0$		$\theta = 0.07$		$\theta = 0.15$		$\theta = 0.2$		$\theta = 0.23$	
Iter.	Time	$\|p_k\|$	Time	$\|p_k\|$	Time	$\|p_k\|$	Time	$\|p_k\|$	Time	$\|p_k\|$
50	**0.0017**	**0.2739**	0.0015	0.2966	0.0015	0.3254	0.0015	0.3438	0.0015	0.3549
100	**0.0033**	**0.3549**	0.0031	0.3659	0.0031	0.3784	0.0031	0.3854	0.0031	0.3893
200	**0.0066**	**0.03953**	0.0064	0.3976	0.0063	0.3997	0.0062	0.4007	0.0063	0.4011
300	**0.0099**	**0.4012**	0.0095	0.4016	0.0094	0.4019	0.0094	0.4020	0.0094	0.4021
400	**0.0130**	**0.4022**	0.0129	0.4022	0.0125	0.4022	0.0126	0.4022	0.0125	0.4022

of the inertial extrapolation to Algorithm 1 is obvious. And hence, the inertial extrapolation of Algorithm 1 plays a key role in the acceleration. Meanwhile, Figure 5 and Table 1 display that the iterative sequence $(\|p_k\|)_{k\in\mathbb{N}}$ (y-axis) converges uniquely to 0.4022, with the number of iterations (resp.) $k = 400$ (x-axis).

Let us recall some previously known algorithms. The extragradient method (EGM) was proposed by Korpelevich [16]. An alternative to the extragradient method or its modification (MEGM) was proposed by Tseng in [17] and was extended in [18]. By combining the extragradient method with the approximate method, the algorithms (Y-EAPM, L-EAPM) were proposed in [19, 20]. By combining the Popov extragradient method with the subgradient extragradient method, the algorithm (SEGM) was presented by Malitsky and Semenov in [21]. By combining with positive features of the Halpern method (HSEM), a modification of the subgradient extragradient method was proposed in [22]. Next, we will present the numerical results to illustrate the practicability and the competitive performance of our proposed algorithm, in comparison with algorithms mentioned above.

Example 5.4. Consider the linear variational inequality problem with the operator $\Phi(x) = \Theta p + \varsigma$, where ς is a vector in $\mathbb{R}^m$ with every entry uniformly generated from $(0, 2)$ and Θ is randomly generated as suggested in [30], that is,

$$\Theta = \Upsilon\Upsilon^T + \Psi + \Gamma,$$

TABLE 2. Comparison results between proposed Algorithm 1, EGM, HSEM, MEGM, Y-EAPM, L-EAPM, and SEGM.

Method	Algorithm 1		EGM		HSEM		MEGM		Y-EAPM		L-EAPM		SEGM	
Iter.	Time	E_k	Time	E_k	Time	E_k	Time	E_k	Time	E_k	Time	E_k	Time	E_k
30	**5.9051e-04**	**0.0176**	7.9701e-04	0.0192	1.5866e-03	0.0241	5.8937e-04	0.0192	1.4740e-03	0.0207	1.4564e-03	0.0229	2.2175e-03	0.0201
40	**7.8734e-04**	**0.0129**	1.0234e-03	0.0163	2.0838e-03	0.0209	7.7653e-04	0.0163	1.9183e-03	0.0182	1.8961e-03	0.0201	2.9127e-03	0.0176
50	**9.6085e-04**	**0.0103**	1.2453e-03	0.0138	2.5685e-03	0.0189	9.5687e-04	0.0138	2.3609e-03	0.0166	2.3319e-03	0.0182	3.6267e-03	0.0152
60	**1.1344e-03**	**0.0082**	1.4649e-03	0.0114	3.0527e-03	0.0175	1.1383e-03	0.0114	2.8075e-03	0.0156	2.7756e-03	0.0167	4.3315 e-03	0.0126
70	**1.3096e-03**	**0.0063**	1.6879e-03	0.0091	3.5885e-03	0.0163	1.3204e-03	0.0091	3.2557e-03	0.0150	3.2114e-03	0.0155	5.0267e-03	0.0101
80	**1.4860e-03**	**0.0048**	1.9183e-03	0.0070	4.0835e-03	0.0155	1.4996e-03	0.0070	3.7040e-03	0.0145	3.6796e-03	0.0146	5.7515e-03	0.0078
90	**1.6663e-03**	**0.0036**	2.1549e-03	0.0053	4.5824e-03	0.0150	1.6850e-03	0.0053	4.1540e-03	0.0143	4.1267e-03	0.0140	6.4728e-03	0.0058
100	**1.8466e-03**	**0.0028**	2.3882e-03	0.0039	5.0716e-03	0.0146	1.8688e-03	0.0039	4.5926e-03	0.0142	4.5631e-03	0.0136	7.1902e-03	0.0044

where Υ is an $m \times m$ matrix and Ψ is an $m \times m$ skew-symmetric matrix with their entries uniformly generated from $(-10, 10)$. Since every diagonal entry of the $m \times m$ diagonal Γ is uniformly generated in $(0, 2)$, thus Θ is positive symmetric definite. The feasible set $\Omega \subset \mathbb{R}^m$ is a convex closed set defined by

$$\Omega = \{p \in \mathbb{R}^m | -4 \leq p_i \leq 4, i = 1, 2, \cdots, m\}.$$

One sees that Φ is strongly pseudomonotone and Lipschitz continuous with modulus $L = \|\Theta\|$. Therefore, it is reasonable to assert that Φ is pseudomonotone and Lipschitz continuous. Our problem here is to find an element $p^* \in \Omega$ such that $\langle \Phi p^*, p - p^* \rangle \geq 0$, $\forall p \in \Omega$.

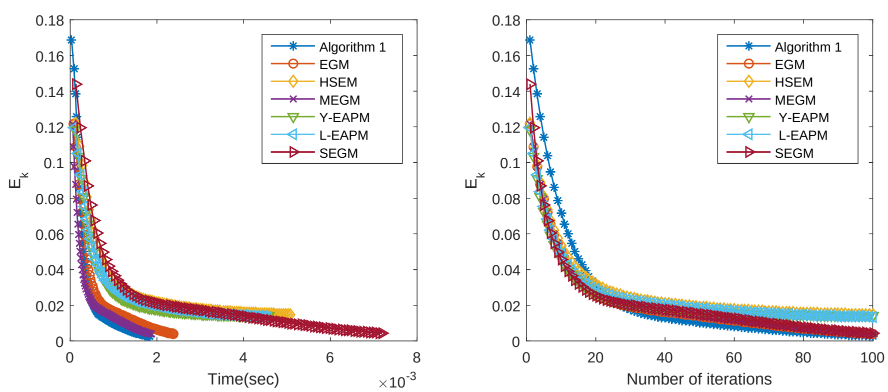

FIGURE 6. Comparison of the convergence behaviors of the error $(E_k)_{k\in\mathbb{N}}$ with the number of iterations (resp.) $k = 100$. Numerical results for different algorithms.

Since we do not know the exact solution of the problem, we use the sequence $E_k = \|p_k - P_\Omega(p_k - \beta_k \Phi p_k)\|$, $\forall k = 0, 1, 2, 3 \cdots$, to measure the error of the k-th iteration of Algorithm 1. Based on the metric projection, if the error distance $\|E_k\| < \varepsilon$, then p_k can be considered as an ε-solution of the problem, which also serves as the role of checking whether or not the proposed algorithm converges to the solution. In this experiment, we set $m = 7$. One randomly chooses the initial element in the range of $(0, 1)^m$ and takes the iteration number $k = 100$ as the stopping criterion in the following experiment.

For comparison on the convergence speed between those algorithms, we use the same random initial elements and the same number of iterations. To illustrate the convergence and computational performance of all the algorithms, one has shown that the values of $(E_k)_{k\in\mathbb{N}}$ (y-axis) with the number of iterations (resp.) $k = 100$ (x-axis) in Figure 6. One can check that methods (Algorithm 1, EGM, MEGM) have a similar convergence behaviour, which outperform the convergence behaviours of methods (HSEM, SEGM, Y-EAPM, L-EAPM). This can be explained by the fact that the calculation of the projection from one point onto a feasible set at each iteration really matters. Indeed, our proposed algorithm is more efficient in CPU-Time, compared with the extragradient method (EGM) and Tseng's extragradient method (MEGM), with respect to the convergence behavior of the l_2 norm error sequence $(E_k)_{k\in\mathbb{N}}$. That is, methods (EGM, MEGM) have a less convergence speed than Algorithm 1, which can be explained by the fact that the presence of the initial extrapolation term per iteration plays the key role of the acceleration process.

The test results are summarized in Table 2. From the changing processes of the values of $(E_k)_{k\in\mathbb{N}}$, one finds that Algorithm 1 has a better behavior, in contrast to that of other algorithms. It achieves a more stable and higher precision with the number of iterations. Besides, the convergence of $(E_k)_{k\in\mathbb{N}}$ to 0 means that the iterative sequence converges to the solution of the variational inequality problem. Above all, Algorithm 1 has substantially a better performance, compared with other algorithms. We also note that, in this experiment, since the computations of the operator Φ and projections are cheap, the running time is quite small. For this reason, our results cannot achieve a drastic contrast of algorithms' performance.

Acknowledgments

The research of J.C. Yao was supported by the Grant MOST (108- 2115-M-039-005-MY3).

References

[1] M. Eshaghnezhad, S. Effati, A. Mansoori, A neurodynamic model to solve nonlinear pseudo-monotone projection equation and its applications, IEEE Trans. Cybernetics, 47 (2016), 3050-3062.
[2] N.T. An, N.M. Nam, X. Qin, Solving k-center problems involving sets based on optimization techniques, J. Global Optim. 76 (2020), 189–209.
[3] L.V. Nguyen, X. Qin, The minimal time function associated with a collection of sets, ESAIM Control Optim. Calc. Var. 26 (2020), 93.
[4] X. He, T. Huang, J. Yu, C. Li, C. Li, An inertial projection neural network for solving variational inequalities, IEEE Trans. Cybernetics, 47 (2017), 809-814.
[5] L.V. Nguyen, Q.H. Ansari, X. Qin, Weak sharpness and finite convergence for solutions of nonsmooth variational inequalities in Hilbert spaces, Appl. Math. Optim. 84 (201), 807-828.
[6] Y. Xia, H. Leung, H. Wang, A projection neural network and its application to constrained optimization problems, IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications, 49 (2002), 447-458.
[7] X. Hu, J. Wang, Solving pseudomonotone variational inequalities and pseudoconvex optimization problems using the projection neural network, IEEE Trans. Neural Networks, 17 (2006), 1487-1499.
[8] X. Qin, N.T. An, Smoothing algorithms for computing the projection onto a Minkowski sum of convex sets, Comput. Optim. Appl. 74 (2019), 821-850.
[9] B. Tan, J. Fan, S. Li, Self-adaptive inertial extragradient algorithms for solving variational inequality problems, Comput. Appl. Math. 40 (2021), Article ID 19.
[10] B.T. Polyak, Some methods of speeding up the convergence of iteration methods, USSR Comput. Math. Math. Phys. 4 (1964), 1-17.
[11] Y. Shehu, O.S. Iyiola, Projection methods with alternating inertial steps for variational inequalities: Weak and linear convergence, Appl. Numer. Math. 157 (2020), 315-337.
[12] Y. Shehu, O.S. Iyiola, F.U. Ogbuisi, Iterative method with inertial terms for nonexpansive mappings: applications to compressed sensing, Numer. Algo. 83 (2020), 1321-1347.

[13] Z. Zhou, B. Tan, S. Li, An accelerated hybrid projection method with a self-adaptive step-size sequence for solving split common fixed point problems, Math. Meth. Appl. Sci. 44 (2021), 7294-7303.
[14] Y. Shehu, A. Gibali, New inertial relaxed method for solving split feasibilities, Optimization, in press. doi: 10.1007/s11590-020-01603-1.
[15] J.P. Penot, P.H. Quang, Generalized convexity of functions and generalized monotonicity of set-valued maps, J. Optim. Theory Appl. 92 (1997), 343-356.
[16] G.M. Korpelevich, The extragradient method for finding saddle points and other problems, Matecon, 12 (1976), 747-756.
[17] P. Tseng, A modified forward-backward splitting method for maximal monotone mappings, SIAM J. Control Optim. 38 (2000), 431-446.
[18] R.I. Boţ, E.R. Csetnek, P.T. Vuong, The forward-backward-forward method from continuous and discrete perspective for pseudo-monotone variational inequalities in Hilbert spaces, Euro. J. Oper. Res. 287 (2020), 49-60.
[19] Y. Yao, M. Postolache, Iterative methods for pseudomonotone variational inequalities and fixed-point problems, J. Optim. Theory Appl. 155 (2012), 273-287.
[20] L.C. Ceng, M. Teboulle, J.C. Yao, Weak convergence of an iterative method for pseudomonotone variational inequalities and fixed-point problems, J. Optim. Theory Appl. 146 (2010), 19-31.
[21] Y.V. Malitsky, V.V. Semenov, An extragradient algorithm for monotone variational inequalities, Cybernetics and Systems Analysis, 50 (2014), 271-277.
[22] R. Kraikaew, S. Saejung, Strong convergence of the Halpern subgradient extragradient method for solving variational inequalities in Hilbert spaces, J. Optim. Theory Appl. 163 (2014), 399-412.
[23] D. Kinderlehrer, G. Stampacchia, An Introduction to Variational Inequalities and their Applications, Academic Press, New York, 1980.
[24] A.A. Khan, E. Kobies, C. Tammer, Variational Analysis and Set Optimization, Taylor & Francis, Berlin, 2019.
[25] A. Filippov, Differential Equations with Discontinuous Righthand Sides, Mathematics and its applications (Soviet Series). Kluwer Academic Publishers, Boston, 1988.
[26] P.T. Vuong, On the weak convergence of the extragradient method for solving pseudo-monotone variational inequalities, J. Optim. Theory Appl. 176 (2018), 399-409.
[27] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, Set-Valued Anal. 9 (2001), 3-11.
[28] R.W. Cottle, J.C. Yao, Pseudo-monotone complementarity problems in Hilbert space, J. Optim. Theory Appl. 75 (1992), 281-295.
[29] P. Maréchal, J.J. Ye, Optimizing condition numbers. SIAM J. Optim. 20 (2009), 935-947.
[30] P.T. Harker, J.S. Pang, A damped-Newton method for the linear complementarity problem, Lectures in Applied Mathematics, 26 (1990), 265-284.