

RESEARCH

Open Access



Bounded perturbation resilience of extragradient-type methods and their applications

Q-L Dong¹, A Gibali^{2*}, D Jiang¹ and Y Tang³

*Correspondence:
avivg@braude.ac.il

²Department of Mathematics, ORT
Braude College, Karmiel, 2161002,
Israel

Full list of author information is
available at the end of the article

Abstract

In this paper we study the bounded perturbation resilience of the extragradient and the subgradient extragradient methods for solving a variational inequality (VI) problem in real Hilbert spaces. This is an important property of algorithms which guarantees the convergence of the scheme under summable errors, meaning that an inexact version of the methods can also be considered. Moreover, once an algorithm is proved to be bounded perturbation resilience, superiorization can be used, and this allows flexibility in choosing the bounded perturbations in order to obtain a superior solution, as well explained in the paper. We also discuss some inertial extragradient methods. Under mild and standard assumptions of monotonicity and Lipschitz continuity of the VI's associated mapping, convergence of the perturbed extragradient and subgradient extragradient methods is proved. In addition we show that the perturbed algorithms converge at the rate of $O(1/t)$. Numerical illustrations are given to demonstrate the performances of the algorithms.

MSC: 49J35; 58E35; 65K15; 90C47

Keywords: inertial-type method; bounded perturbation resilience; extragradient method; subgradient extragradient method; variational inequality

1 Introduction

In this paper we are concerned with the variational inequality (VI) problem of finding a point x^* such that

$$\langle F(x^*), x - x^* \rangle \geq 0 \quad \text{for all } x \in C, \quad (1.1)$$

where $C \subseteq \mathcal{H}$ is a nonempty, closed and convex set in a real Hilbert space \mathcal{H} , $\langle \cdot, \cdot \rangle$ denotes the inner product in \mathcal{H} , and $F: \mathcal{H} \rightarrow \mathcal{H}$ is a given mapping. This problem is a fundamental problem in optimization theory, and it captures various applications such as partial differential equations, optimal control and mathematical programming; for the theory and application of VIs and related problems, the reader is referred, for example, to the works of Ceng *et al.* [1], Zegeye *et al.* [2], the papers of Yao *et al.* [3–5] and the many references therein.

Many algorithms for solving VI (1.1) are projection algorithms that employ projections onto the feasible set C of VI (1.1), or onto some related set, in order to reach iteratively a

solution. Korpelevich [6] and Antipin [7] proposed an algorithm for solving (1.1), known as the *extragradient method*, see also Facchinei and Pang [8, Chapter 12]. In each iteration of the algorithm, in order to get the next iterate x^{k+1} , two orthogonal projections onto C are calculated according to the following iterative step. Given the current iterate x^k , calculate

$$\begin{cases} y^k = P_C(x^k - \gamma_k F(x^k)), \\ x^{k+1} = P_C(x^k - \gamma_k F(y^k)), \end{cases} \tag{1.2}$$

where $\gamma_k \in (0, 1/L)$, and L is the Lipschitz constant of F , or γ_k is updated by the following adaptive procedure:

$$\gamma_k \|F(x^k) - F(y^k)\| \leq \mu \|x^k - y^k\|, \quad \mu \in (0, 1). \tag{1.3}$$

In the extragradient method there is the need to calculate twice the orthogonal projection onto C in each iteration. In case that the set C is simple enough so that projections onto it can be easily computed, this method is particularly useful; but if C is a general closed and convex set, a minimal distance problem has to be solved (twice) in order to obtain the next iterate. This might seriously affect the efficiency of the extragradient method. Hence, Censor *et al.* in [9–11] presented a method called the *subgradient extragradient method*, in which the second projection (1.2) onto C is replaced by a specific subgradient projection which can be easily calculated. The iterative step has the following form:

$$\begin{cases} y^k = P_C(x^k - \gamma F(x^k)), \\ x^{k+1} = P_{T_k}(x^k - \gamma F(y^k)), \end{cases} \tag{1.4}$$

where T_k is the set defined as

$$T_k := \{w \in \mathcal{H} \mid \langle (x^k - \gamma F(x^k)) - y^k, w - y^k \rangle \leq 0\}, \tag{1.5}$$

and $\gamma \in (0, 1/L)$.

In this manuscript we prove that the above methods, the extragradient and the subgradient extragradient methods, are bounded perturbation resilient, and the perturbed methods have the convergence rate of $O(1/t)$. This means that that will show that an inexact version of the algorithms allows incorporating summable errors that also converge to a solution of VI (1.1) and, moreover, their superiorized version can be introduced by choosing the perturbations. In order to obtain a superior solution with respect to some new objective function, for example, by choosing the norm, we can obtain a solution to VI (1.1) which is closer to the origin.

Our paper is organized as follows. In Section 2 we present the preliminaries. In Section 3 we study the convergence of the extragradient method with outer perturbations. Later, in Section 4, the bounded perturbation resilience of the extragradient method as well as the construction of the inertial extragradient methods are presented.

In the spirit of the previous sections, in Section 5 we study the convergence of the subgradient extragradient method with outer perturbations, show its bounded perturbation resilience and the construction of the inertial subgradient extragradient methods. Finally, in Section 6 we present numerical examples in signal processing which demonstrate the performances of the perturbed algorithms.

2 Preliminaries

Let \mathcal{H} be a real Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and the induced norm $\| \cdot \|$, and let D be a nonempty, closed and convex subset of \mathcal{H} . We write $x^k \rightharpoonup x$ to indicate that the sequence $\{x^k\}_{k=0}^\infty$ converges weakly to x and $x^k \rightarrow x$ to indicate that the sequence $\{x^k\}_{k=0}^\infty$ converges strongly to x . Given a sequence $\{x^k\}_{k=0}^\infty$, denote by $\omega_w(x^k)$ its weak ω -limit set, that is, any $x \in \omega_w(x^k)$ such that there exists a subsequence $\{x^{k_j}\}_{j=0}^\infty$ of $\{x^k\}_{k=0}^\infty$ which converges weakly to x .

For each point $x \in \mathcal{H}$, there exists a unique nearest point in D denoted by $P_D(x)$. That is,

$$\|x - P_D(x)\| \leq \|x - y\| \quad \text{for all } y \in D. \tag{2.1}$$

The mapping $P_D : \mathcal{H} \rightarrow D$ is called the metric projection of \mathcal{H} onto D . It is well known that P_D is a *nonexpansive mapping* of \mathcal{H} onto D , i.e., and even *firmly nonexpansive mapping*. This is captured in the next lemma.

Lemma 2.1 *For any $x, y \in \mathcal{H}$ and $z \in D$, it holds*

- $\|P_D(x) - P_D(y)\|^2 \leq \|x - y\|^2$;
- $\|P_D(x) - z\|^2 \leq \|x - z\|^2 - \|P_D(x) - x\|^2$.

The characterization of the metric projection P_D [12, Section 3] is given by the following two properties in this lemma.

Lemma 2.2 *Given $x \in \mathcal{H}$ and $z \in D$. Then $z = P_D(x)$ if and only if*

$$P_D(x) \in D \tag{2.2}$$

and

$$\langle x - P_D(x), P_D(x) - y \rangle \geq 0 \quad \text{for all } x \in \mathcal{H}, y \in D. \tag{2.3}$$

Definition 2.3 The normal cone of D at $v \in D$ denoted by $N_D(v)$ is defined as

$$N_D(v) := \{d \in \mathcal{H} \mid \langle d, y - v \rangle \leq 0 \text{ for all } y \in D\}. \tag{2.4}$$

Definition 2.4 Let $B : \mathcal{H} \rightrightarrows 2^{\mathcal{H}}$ be a point-to-set operator defined on a real Hilbert space \mathcal{H} . The operator B is called a maximal monotone operator if B is monotone, i.e.,

$$\langle u - v, x - y \rangle \geq 0 \quad \text{for all } u \in B(x) \text{ and } v \in B(y), \tag{2.5}$$

and the graph $G(B)$ of B ,

$$G(B) := \{(x, u) \in \mathcal{H} \times \mathcal{H} \mid u \in B(x)\}, \tag{2.6}$$

is not properly contained in the graph of any other monotone operator.

Based on Rockafellar [13, Theorem 3], a monotone mapping B is maximal if and only if, for any $(x, u) \in \mathcal{H} \times \mathcal{H}$, if $\langle u - v, x - y \rangle \geq 0$ for all $(v, y) \in G(B)$, then it follows that $u \in B(x)$.

Definition 2.5 The subdifferential set of a convex function c at a point x is defined as

$$\partial c(x) := \{ \xi \in \mathcal{H} \mid c(y) \geq c(x) + \langle \xi, y - x \rangle \text{ for all } y \in \mathcal{H} \}. \tag{2.7}$$

For $z \in \mathcal{H}$, take any $\xi \in \partial c(z)$ and define

$$T(z) := \{ w \in \mathcal{H} \mid c(z) + \langle \xi, w - z \rangle \leq 0 \}. \tag{2.8}$$

This is a half-space, the bounding hyperplane of which separates the set D from the point z if $\xi \neq 0$; otherwise $T(z) = \mathcal{H}$; see, e.g., [14, Lemma 7.3].

Lemma 2.6 ([15]) *Let D be a nonempty, closed and convex subset of a Hilbert space \mathcal{H} . Let $\{x^k\}_{k=0}^\infty$ be a bounded sequence which satisfies the following properties:*

- every limit point of $\{x^k\}_{k=0}^\infty$ lies in D ;
- $\lim_{n \rightarrow \infty} \|x^k - x\|$ exists for every $x \in D$.

Then $\{x^k\}_{k=0}^\infty$ converges to a point in D .

Lemma 2.7 *Assume that $\{a_k\}_{k=0}^\infty$ is a sequence of nonnegative real numbers such that*

$$a_{k+1} \leq (1 + \gamma_k)a_k + \delta_k, \quad \forall k \geq 0, \tag{2.9}$$

where the nonnegative sequences $\{\gamma_k\}_{k=0}^\infty$ and $\{\delta_k\}_{k=0}^\infty$ satisfy $\sum_{k=0}^\infty \gamma_k < +\infty$ and $\sum_{k=0}^\infty \delta_k < +\infty$, respectively. Then $\lim_{k \rightarrow \infty} a_k$ exists.

3 The extragradient method with outer perturbations

In order to discuss the convergence of the extragradient method with outer perturbations, we make the following assumptions.

Condition 3.1 The solution set of (1.1), denoted by $SOL(C, F)$, is nonempty.

Condition 3.2 The mapping F is monotone on C , i.e.,

$$\langle F(x) - F(y), x - y \rangle \geq 0, \quad \forall x, y \in C. \tag{3.1}$$

Condition 3.3 The mapping F is Lipschitz continuous on C with the Lipschitz constant $L > 0$, i.e.,

$$\|F(x) - F(y)\| \leq L\|x - y\|, \quad \forall x, y \in C. \tag{3.2}$$

Observe that while Censor *et al.* in [10, Theorem 3.1] showed the weak convergence of the extragradient method (1.2) in Hilbert spaces for a fixed step size $\gamma_k = \gamma \in (0, 1/L)$, this can be easily improved in case that the adaptive rule (1.3) is used. The next theorem shows this, and its proof can easily be derived by following similar lines of the proof of [10, Theorem 3.1].

Theorem 3.4 *Assume that Conditions 3.1-3.3 hold. Then any sequence $\{x^k\}_{k=0}^\infty$ generated by the extragradient method (1.2) with the adaptive rule (1.3) weakly converges to a solution of the variational inequality (1.1).*

Denote $e_i^k := e_i(x^k)$, $i = 1, 2$. The sequences of perturbations $\{e_i^k\}_{k=0}^\infty$, $i = 1, 2$, are assumed to be summable, i.e.,

$$\sum_{k=0}^\infty \|e_i^k\| < +\infty, \quad i = 1, 2. \tag{3.3}$$

Now we consider the extragradient method with outer perturbations.

Algorithm 3.5 The extragradient method with outer perturbations:

Step 0: Select a starting point $x^0 \in C$ and set $k = 0$.

Step 1: Given the current iterate x^k , compute

$$y^k = P_C(x^k - \gamma_k F(x^k) + e_1(x^k)), \tag{3.4}$$

where $\gamma_k = \sigma \rho^{m_k}$, $\sigma > 0$, $\rho \in (0, 1)$ and m_k is the smallest nonnegative integer such that (see [16])

$$\gamma_k \|F(x^k) - F(y^k)\| \leq \mu \|x^k - y^k\|, \quad \mu \in (0, 1). \tag{3.5}$$

Calculate the next iterate

$$x^{k+1} = P_C(x^k - \gamma_k F(y^k) + e_2(x^k)). \tag{3.6}$$

Step 2: If $x^k = y^k$, then stop. Otherwise, set $k \leftarrow (k + 1)$ and return to *Step 1*.

3.1 Convergence analysis

Lemma 3.6 ([17]) *The Armijo-like search rule (3.5) is well defined. Besides, $\underline{\gamma} \leq \gamma_k \leq \sigma$, where $\underline{\gamma} = \min\{\sigma, \frac{\mu\rho}{L}\}$.*

Theorem 3.7 *Assume that Conditions 3.1-3.3 hold. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by Algorithm 3.5 converges weakly to a solution of the variational inequality (1.1).*

Proof Take $x^* \in SOL(C, F)$. From (3.6) and Lemma 2.1(ii), we have

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq \|x^k - \gamma_k F(y^k) + e_2^k - x^*\|^2 - \|x^k - \gamma_k F(y^k) + e_2^k - x^{k+1}\|^2 \\ &= \|x^k - x^*\|^2 - \|x^k - x^{k+1}\|^2 + 2\gamma_k \langle F(y^k), x^* - x^{k+1} \rangle \\ &\quad - 2\langle e_2^k, x^* - x^{k+1} \rangle. \end{aligned} \tag{3.7}$$

From the Cauchy-Schwarz inequality and the mean value inequality, it follows

$$\begin{aligned} -2\langle e_2^k, x^* - x^{k+1} \rangle &\leq 2\|e_2^k\| \|x^{k+1} - x^*\| \\ &\leq \|e_2^k\| + \|e_2^k\| \|x^{k+1} - x^*\|^2. \end{aligned} \tag{3.8}$$

Using $x^* \in SOL(C, F)$ and the monotone property of F , we have $\langle y^k - x^*, F(y^k) \rangle \geq 0$ and, consequently, we get

$$2\gamma_k \langle F(y^k), x^* - x^{k+1} \rangle \leq 2\gamma_k \langle F(y^k), y^k - x^{k+1} \rangle. \tag{3.9}$$

Thus, we have

$$\begin{aligned} & -\|x^k - x^{k+1}\|^2 + 2\gamma_k \langle F(y^k), x^* - x^{k+1} \rangle \\ & \leq -\|x^k - x^{k+1}\|^2 + 2\gamma_k \langle F(y^k), y^k - x^{k+1} \rangle \\ & = -\|x^k - y^k\|^2 - \|y^k - x^{k+1}\|^2 \\ & \quad + 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle, \end{aligned} \tag{3.10}$$

where the equality comes from

$$-\|x^k - x^{k+1}\|^2 = -\|x^k - y^k\|^2 - \|y^k - x^{k+1}\|^2 - 2\langle x^k - y^k, y^k - x^{k+1} \rangle. \tag{3.11}$$

Using $x^{k+1} \in C$, the definition of y^k and Lemma 2.2, we have

$$\langle y^k - x^k + \gamma_k F(x^k) - e_1^k, x^{k+1} - y^k \rangle \geq 0. \tag{3.12}$$

So, we obtain

$$\begin{aligned} & 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle \\ & \leq 2\gamma_k \langle F(x^k) - F(y^k), x^{k+1} - y^k \rangle - 2\langle e_1^k, x^{k+1} - y^k \rangle \\ & \leq 2\gamma_k \|F(x^k) - F(y^k)\| \|x^{k+1} - y^k\| + 2\|e_1^k\| \|x^{k+1} - y^k\| \\ & \leq 2\mu \|x^k - y^k\| \|x^{k+1} - y^k\| + \|e_1^k\| + \|e_1^k\| \|x^{k+1} - y^k\|^2 \\ & \leq \mu \|x^k - y^k\|^2 + \mu \|x^{k+1} - y^k\|^2 + \|e_1^k\| + \|e_1^k\| \|x^{k+1} - y^k\|^2 \\ & = \mu \|x^k - y^k\|^2 + (\mu + \|e_1^k\|) \|x^{k+1} - y^k\|^2 + \|e_1^k\|. \end{aligned} \tag{3.13}$$

From (3.3), it follows

$$\lim_{k \rightarrow \infty} \|e_i^k\| = 0, \quad i = 1, 2. \tag{3.14}$$

Therefore, we assume $\|e_1^k\| \in [0, 1 - \mu - \nu)$ and $\|e_2^k\| \in [0, 1/2)$, $k \geq 0$, where $\nu \in (0, 1 - \mu)$. So, using (3.13), we get

$$2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle \leq \mu \|x^k - y^k\|^2 + (1 - \nu) \|x^{k+1} - y^k\|^2 + \|e_1^k\|. \tag{3.15}$$

Combining (3.7)-(3.10) and (3.15), we obtain

$$\begin{aligned} \|x^{k+1} - x^*\|^2 & \leq \|x^k - x^*\|^2 - (1 - \mu) \|x^k - y^k\|^2 - \nu \|x^{k+1} - y^k\|^2 \\ & \quad + \|e^k\| + \|e_2^k\| \|x^{k+1} - x^*\|^2, \end{aligned} \tag{3.16}$$

where

$$\|e^k\| := \|e_1^k\| + \|e_2^k\|. \tag{3.17}$$

From (3.16), it follows

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq \frac{1}{1 - \|e_2^k\|} \|x^k - x^*\|^2 - \frac{1 - \mu}{1 - \|e_2^k\|} \|x^k - y^k\|^2 \\ &\quad - \frac{\nu}{1 - \|e_2^k\|} \|x^{k+1} - y^k\|^2 + \frac{\|e^k\|}{1 - \|e_2^k\|}. \end{aligned} \tag{3.18}$$

Since $\|e_2^k\| \in [0, 1/2)$, $k \geq 0$, we get

$$1 \leq \frac{1}{1 - \|e_2^k\|} \leq 1 + 2\|e_2^k\| < 2. \tag{3.19}$$

So, from (3.18), we have

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &\leq (1 + 2\|e_2^k\|) \|x^k - x^*\|^2 - (1 - \mu) \|x^k - y^k\|^2 \\ &\quad - \nu \|x^{k+1} - y^k\|^2 + 2\|e^k\| \\ &\leq (1 + 2\|e_2^k\|) \|x^k - x^*\|^2 + 2\|e^k\|. \end{aligned} \tag{3.20}$$

Using (3.3) and Lemma 2.7, we get the existence of $\lim_{k \rightarrow \infty} \|x^k - x^*\|^2$ and then the boundedness of $\{x^k\}_{k=0}^\infty$. From (3.20), it follows

$$(1 - \mu) \|x^k - y^k\|^2 + \nu \|x^{k+1} - y^k\|^2 \leq (1 + 2\|e_2^k\|) \|x^k - x^*\|^2 - \|x^{k+1} - x^*\|^2 + 2\|e^k\|, \tag{3.21}$$

which means that

$$\sum_{k=0}^\infty \|x^k - y^k\|^2 < +\infty \quad \text{and} \quad \sum_{k=0}^\infty \|x^{k+1} - y^k\|^2 < +\infty. \tag{3.22}$$

Thus, we obtain

$$\lim_{k \rightarrow \infty} \|x^k - y^k\| = 0 \quad \text{and} \quad \lim_{k \rightarrow \infty} \|x^{k+1} - y^k\| = 0, \tag{3.23}$$

and consequently,

$$\lim_{k \rightarrow \infty} \|x^{k+1} - x^k\| = 0. \tag{3.24}$$

Now, we are to show $\omega_w(x^k) \subseteq SOL(C, F)$. Due to the boundedness of $\{x^k\}_{k=0}^\infty$, it has at least one weak accumulation point. Let $\hat{x} \in \omega_w(x^k)$. Then there exists a subsequence $\{x^{k_i}\}_{i=0}^\infty$ of $\{x^k\}_{k=0}^\infty$ which converges weakly to \hat{x} . From (3.23), it follows that $\{y^{k_i}\}_{i=0}^\infty$ also converges weakly to \hat{x} .

We will show that \hat{x} is a solution of the variational inequality (1.1). Let

$$A(v) = \begin{cases} F(v) + N_C(v), & v \in C, \\ \emptyset, & v \notin C, \end{cases} \tag{3.25}$$

where $N_C(v)$ is the normal cone of C at $v \in C$. It is known that A is a maximal monotone operator and $A^{-1}(0) = SOL(C, F)$. If $(v, w) \in G(A)$, then we have $w - F(v) \in N_C(v)$ since $w \in A(v) = F(v) + N_C(v)$. Thus it follows that

$$\langle w - F(v), v - y \rangle \geq 0, \quad y \in C. \tag{3.26}$$

Since $y^{k_i} \in C$, we have

$$\langle w - F(v), v - y^{k_i} \rangle \geq 0. \tag{3.27}$$

On the other hand, by the definition of y^k and Lemma 2.2, it follows that

$$\langle x^k - \gamma_k F(x^k) + e_1^k - y^k, y^k - v \rangle \geq 0, \tag{3.28}$$

and consequently,

$$\left\langle \frac{y^k - x^k}{\gamma_k} + F(x^k), v - y^k \right\rangle - \frac{1}{\gamma_k} \langle e_1^k, v - y^k \rangle \geq 0. \tag{3.29}$$

Hence we have

$$\begin{aligned} \langle w, v - y^{k_i} \rangle &\geq \langle F(v), v - y^{k_i} \rangle \\ &\geq \langle F(v), v - y^{k_i} \rangle - \left\langle \frac{y^{k_i} - x^{k_i}}{\gamma_{k_i}} + F(x^{k_i}), v - y^{k_i} \right\rangle + \frac{1}{\gamma_{k_i}} \langle e_1^{k_i}, v - y^{k_i} \rangle \\ &= \langle F(v) - F(y^{k_i}), v - y^{k_i} \rangle + \langle F(y^{k_i}) - F(x^{k_i}), v - y^{k_i} \rangle \\ &\quad - \left\langle \frac{y^{k_i} - x^{k_i}}{\gamma_{k_i}}, v - y^{k_i} \right\rangle + \frac{1}{\gamma_{k_i}} \langle e_1^{k_i}, v - y^{k_i} \rangle \\ &\geq \langle F(y^{k_i}) - F(x^{k_i}), v - y^{k_i} \rangle - \left\langle \frac{y^{k_i} - x^{k_i}}{\gamma_{k_i}}, v - y^{k_i} \right\rangle + \frac{1}{\gamma_{k_i}} \langle e_1^{k_i}, v - y^{k_i} \rangle, \end{aligned} \tag{3.30}$$

which implies

$$\langle w, v - y^{k_i} \rangle \geq \langle F(y^{k_i}) - F(x^{k_i}), v - y^{k_i} \rangle - \left\langle \frac{y^{k_i} - x^{k_i}}{\gamma_{k_i}}, v - y^{k_i} \right\rangle + \frac{1}{\gamma_{k_i}} \langle e_1^{k_i}, v - y^{k_i} \rangle. \tag{3.31}$$

Taking the limit as $i \rightarrow \infty$ in the above inequality and using Lemma 3.6, we obtain

$$\langle w, v - \hat{x} \rangle \geq 0. \tag{3.32}$$

Since A is a maximal monotone operator, it follows that $\hat{x} \in A^{-1}(0) = SOL(C, F)$. So, $\omega_w(x^k) \subseteq SOL(C, F)$.

Since $\lim_{k \rightarrow \infty} \|x^k - x^*\|$ exists and $\omega_w(x^k) \subseteq SOL(C, F)$, using Lemma 2.6, we conclude that $\{x^k\}_{k=0}^\infty$ weakly converges to a solution of the variational inequality (1.1). This completes the proof. \square

3.2 Convergence rate

Nemirovski [18] and Tseng [19] proved the $O(1/t)$ convergence rate of the extragradient method. In this subsection, we present the convergence rate of Algorithm 3.5.

Theorem 3.8 *Assume that Conditions 3.1-3.3 hold. Let the sequences $\{x^k\}_{k=0}^\infty$ and $\{y^k\}_{k=0}^\infty$ be generated by Algorithm 3.5. For any integer $t > 0$, we have $y_t \in C$ which satisfies*

$$\langle F(x), y_t - x \rangle \leq \frac{1}{2\Upsilon_t} (\|x - x^0\|^2 + M(x)), \quad \forall x \in C, \tag{3.33}$$

where

$$y_t = \frac{1}{\Upsilon_t} \sum_{k=0}^t \gamma_k y^k, \quad \Upsilon_t = \sum_{k=0}^t \gamma_k \tag{3.34}$$

and

$$M(x) = \sup_k \{ \max \{ \|x^{k+1} - y^k\|, \|x^{k+1} - x\| \} \} \sum_{k=0}^\infty \|e^k\|. \tag{3.35}$$

Proof Take arbitrarily $x \in C$. From Conditions 3.2 and 3.3, we have

$$\begin{aligned} & -\|x^k - x^{k+1}\|^2 + 2\gamma_k \langle F(y^k), x - x^{k+1} \rangle \\ &= -\|x^k - x^{k+1}\|^2 + 2\gamma_k [\langle F(y^k) - F(x), x - y^k \rangle + \langle F(x), x - y^k \rangle \\ &\quad + \langle F(y^k), y^k - x^{k+1} \rangle] \\ &\leq -\|x^k - x^{k+1}\|^2 + 2\gamma_k [\langle F(x), x - y^k \rangle + \langle F(y^k), y^k - x^{k+1} \rangle] \\ &= -\|x^k - y^k\|^2 - \|y^k - x^{k+1}\|^2 + 2\gamma_k \langle F(x), x - y^k \rangle \\ &\quad + 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle. \end{aligned} \tag{3.36}$$

By (3.6) and Lemma 2.2, we get

$$\begin{aligned} & 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle \\ &= 2\langle x^k - \gamma_k F(x^k) + e_1^k - y^k, x^{k+1} - y^k \rangle - 2\langle e_1^k, x^{k+1} - y^k \rangle \\ &\quad + 2\gamma_k \langle F(x^k) - F(y^k), x^{k+1} - y^k \rangle \\ &\leq -2\langle e_1^k, x^{k+1} - y^k \rangle + 2\gamma_k \langle F(x^k) - F(y^k), x^{k+1} - y^k \rangle \\ &\leq 2\|e_1^k\| \|x^{k+1} - y^k\| + 2\mu \|x^k - y^k\| \|x^{k+1} - y^k\| \\ &\leq 2\|e_1^k\| \|x^{k+1} - y^k\| + \mu^2 \|x^k - y^k\|^2 + \|x^{k+1} - y^k\|^2. \end{aligned} \tag{3.37}$$

Identifying x^* with x in (3.7) and (3.8), and combining (3.36) and (3.37), we get

$$\begin{aligned} & \|x^{k+1} - x\|^2 \\ & \leq \|x^k - x\|^2 + 2\|e_1^k\| \|x^{k+1} - y^k\| - (1 - \mu^2) \|x^k - y^k\|^2 \\ & \quad + 2\|e_2^k\| \|x^{k+1} - x\| + 2\gamma_k \langle F(x), x - y^k \rangle \\ & \leq \|x^k - x\|^2 + 2\|e_1^k\| \|x^{k+1} - y^k\| + 2\|e_2^k\| \|x^{k+1} - x\| \\ & \quad + 2\gamma_k \langle F(x), x - y^k \rangle. \end{aligned} \tag{3.38}$$

Thus, we have

$$\begin{aligned} & \gamma_k \langle F(x), y^k - x \rangle \\ & \leq \frac{1}{2} (\|x^k - x\|^2 - \|x^{k+1} - x\|^2) + \|e_1^k\| \|x^{k+1} - y^k\| + \|e_2^k\| \|x^{k+1} - x\| \\ & \leq \frac{1}{2} (\|x^k - x\|^2 - \|x^{k+1} - x\|^2) + M'(x) \|e^k\|, \end{aligned} \tag{3.39}$$

where $M'(x) = \sup_k \{\max\{\|x^{k+1} - y^k\|, \|x^{k+1} - x\|\}\} < +\infty$. Summing inequality (3.39) over $k = 0, \dots, t$, we obtain

$$\begin{aligned} \left\langle F(x), \sum_{k=0}^t \gamma_k y^k - \left(\sum_{k=0}^t \gamma_k \right) x \right\rangle & \leq \frac{1}{2} \|x^0 - x\|^2 + \frac{M'(x)}{2} \sum_{k=0}^t \|e^k\| \\ & = \frac{1}{2} \|x^0 - x\|^2 + \frac{1}{2} M(x). \end{aligned} \tag{3.40}$$

Using the notations of Υ_t and y^t in the above inequality, we derive

$$\langle F(x), y_t - x \rangle \leq \frac{1}{2\Upsilon_t} (\|x - x^0\|^2 + M(x)), \quad \forall x \in C. \tag{3.41}$$

The proof is complete. □

Remark 3.9 From Lemma 3.6, it follows

$$\Upsilon_t \geq (t + 1)\underline{\gamma}, \tag{3.42}$$

thus Algorithm 3.5 has $O(1/t)$ convergence rate. In fact, for any bounded subset $D \subset C$ and given accuracy $\epsilon > 0$, our algorithm achieves

$$\langle F(x), y_t - x \rangle \leq \epsilon, \quad \forall x \in D \tag{3.43}$$

in at most

$$t = \left\lceil \frac{m}{2\underline{\gamma}\epsilon} \right\rceil \tag{3.44}$$

iterations, where y_t is defined by (3.34) and $m = \sup\{\|x - x^0\|^2 + M(x) \mid x \in D\}$.

4 The bounded perturbation resilience of the extragradient method

In this section, we prove the bounded perturbation resilience (BPR) of the extragradient method. This property is fundamental for the application of the superiorization methodology (SM) to them.

The superiorization methodology of [20–22], which originates in the papers by Butnariu, Reich and Zaslavski [23–25], is intended for constrained minimization (CM) problems of the form

$$\min\{\phi(x) \mid x \in \Psi\}, \tag{4.1}$$

where $\phi : H \rightarrow \mathbb{R}$ is an objective function and $\Psi \subseteq H$ is the solution set of another problem. Here, we assume $\Psi \neq \emptyset$ throughout this paper. Assume that the set Ψ is a closed convex subset of a Hilbert space H , the minimization problem (4.1) becomes a standard CM problem. Here, we are interested in the case wherein Ψ is the solution set of another CM of the form

$$\min\{f(x) \mid x \in \Omega\}, \tag{4.2}$$

i.e., we wish to look at

$$\Psi := \{x^* \in \Omega \mid f(x^*) \leq f(x) \text{ for all } x \in \Omega\} \tag{4.3}$$

provided that Ψ is nonempty. If f is differentiable, and let $F = \nabla f$, then CM (4.2) equals the following variational inequality: to find a point $x^* \in C$ such that

$$\langle F(x^*), x - x^* \rangle \geq 0, \quad \forall x \in C. \tag{4.4}$$

The superiorization methodology (SM) strives not to solve (4.1) but rather the task is to find a point in Ψ which is superior, *i.e.*, has a lower, but not necessarily minimal, value of the objective function ϕ . This is done in the SM by first investigating the bounded perturbation resilience of an algorithm designed to solve (4.2) and then proactively using such permitted perturbations in order to steer the iterates of such an algorithm toward lower values of the ϕ objective function while not losing the overall convergence to a point in Ψ .

In this paper, we do not investigate superiorization of the extragradient method. We prepare for such an application by proving the bounded perturbation resilience that is needed in order to do superiorization.

Algorithm 4.1 The basic algorithm:

Initialization: $x^0 \in \Theta$ is arbitrary;

Iterative step: Given the current iterate vector x^k , calculate the next iterate x^{k+1} via

$$x^{k+1} = \mathbf{A}_\Psi(x^k). \tag{4.5}$$

The bounded perturbation resilience (henceforth abbreviated to BPR) of such a basic algorithm is defined next.

Definition 4.2 An algorithmic operator $\mathbf{A}_\Psi : H \rightarrow \Theta$ is said to be bounded perturbations resilient if the following is true. If Algorithm 4.5 generates sequences $\{x^k\}_{k=0}^\infty$ with $x^0 \in \Theta$ that converge to points in Ψ , then any sequence $\{y^k\}_{k=0}^\infty$ starting from any $y^0 \in \Theta$, generated by

$$y^{k+1} = \mathbf{A}_\Psi(y^k + \lambda_k v^k) \quad \text{for all } k \geq 0, \tag{4.6}$$

also converges to a point in Ψ , provided that (i) the sequence $\{v^k\}_{k=0}^\infty$ is bounded, and (ii) the scalars $\{\lambda_k\}_{k=0}^\infty$ are such that $\lambda_k \geq 0$ for all $k \geq 0$, and $\sum_{k=0}^\infty \lambda_k < +\infty$, and (iii) $y^k + \lambda_k v^k \in \Theta$ for all $k \geq 0$.

Definition 4.2 is nontrivial only if $\Theta \neq \mathcal{H}$, in which condition (iii) is enforced in the superiorized version of the basic algorithm, see step (xiv) in the ‘Superiorized Version of Algorithm P’ in ([26], p.5537) and step (14) in ‘Superiorized Version of the ML-EM Algorithm’ in ([27], Subsection II.B). This will be the case in the present work.

Treating the extragradient method as the Basic Algorithm \mathbf{A}_Ψ , our strategy is to first prove convergence of the iterative step (1.2) with bounded perturbations. We show next how the convergence of this yields BPR according to Definition 4.2.

A superiorized version of any Basic Algorithm employs the perturbed version of the Basic Algorithm as in (4.6). A certificate to do so in the superiorization method, see [28], is gained by showing that the Basic Algorithm is BPR. Therefore, proving the BPR of an algorithm is the first step toward superiorizing it. This is done for the extragradient method in the next subsection.

4.1 The BPR of the extragradient method

In this subsection, we investigate the bounded perturbation resilience of the extragradient method whose iterative step is given by (1.2).

To this end, we treat the right-hand side of (1.2) as the algorithmic operator \mathbf{A}_Ψ of Definition 4.2, namely, we define, for all $k \geq 0$,

$$\mathbf{A}_\Psi(x^k) = P_C(x^k - \gamma_k F(P_C(x^k - \gamma_k F(x^k)))) \tag{4.7}$$

and identify the solution set Ψ with the solution set of the variational inequality (1.1) and identify the additional set Θ with C .

According to Definition 4.2, we need to show the convergence of the sequence $\{x^k\}_{k=0}^\infty$ that, starting from any $x^0 \in C$, is generated by

$$x^{k+1} = P_C((x^k + \lambda_k v^k) - \gamma_k F(P_C((x^k + \lambda_k v^k) - \gamma_k F(x^k + \lambda_k v^k))))), \tag{4.8}$$

which can be rewritten as

$$\begin{cases} y^k = P_C((x^k + \lambda_k v^k) - \gamma_k F(x^k + \lambda_k v^k)), \\ x^{k+1} = P_C((x^k + \lambda_k v^k) - \gamma_k F(y^k)), \end{cases} \tag{4.9}$$

where $\gamma_k = \sigma \rho^{m_k}$, $\sigma > 0$, $\rho \in (0, 1)$ and m_k is the smallest nonnegative integer such that

$$\gamma_k \|F(x^k + \lambda_k v^k) - F(y^k)\| \leq \mu (\|x^k - y^k\| + \lambda_k \|v^k\|), \quad \mu \in (0, 1). \tag{4.10}$$

The sequences $\{v^k\}_{k=0}^\infty$ and $\{\lambda_k\}_{k=0}^\infty$ obey conditions (i) and (ii) in Definition 4.2, respectively, and also (iii) in Definition 4.2 is satisfied.

The next theorem establishes the bounded perturbation resilience of the extragradient method. The proof idea is to build a relationship between BPR and the convergence of the iterative step (1.2).

Theorem 4.3 *Assume that Conditions 3.1-3.3 hold. Assume the sequence $\{v^k\}_{k=0}^\infty$ is bounded, and the scalars $\{\lambda_k\}_{k=0}^\infty$ are such that $\lambda_k \geq 0$ for all $k \geq 0$, and $\sum_{k=0}^\infty \lambda_k < +\infty$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by (4.9) and (4.10) converges weakly to a solution of the variational inequality (1.1).*

Proof Take $x^* \in SOL(C, F)$. From $\sum_{k=0}^\infty \lambda_k < +\infty$ and that $\{v^k\}_{k=0}^\infty$ is bounded, we have

$$\sum_{k=0}^\infty \lambda_k \|v^k\| < +\infty, \tag{4.11}$$

which means

$$\lim_{k \rightarrow \infty} \lambda_k \|v^k\| = 0. \tag{4.12}$$

So, we assume $\lambda_k \|v^k\| \in [0, (1 - \mu - \nu)/2]$, where $\nu \in [0, 1 - \mu]$. Identifying e_2^k with $\lambda_k v^k$ in (3.7) and (3.8) and using (3.10), we get

$$\begin{aligned} \|x^{k+1} - x^*\|^2 &= \|x^k - x^*\|^2 + \lambda_k \|v^k\| + \lambda_k \|v^k\| \|x^{k+1} - x^*\|^2 - \|x^k - y^k\|^2 \\ &\quad - \|y^k - x^{k+1}\|^2 + 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle. \end{aligned} \tag{4.13}$$

From $x^{k+1} \in C$, the definition of y^k and Lemma 2.2, we have

$$\langle y^k - x^k - \lambda_k v^k + \gamma_k F(x^k + \lambda_k v^k), x^{k+1} - y^k \rangle \geq 0.$$

So, we obtain

$$\begin{aligned} &2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle \\ &\leq 2\gamma_k \langle F(x^k + \lambda_k v^k) - F(y^k), x^{k+1} - y^k \rangle - 2\lambda_k \langle v^k, x^{k+1} - y^k \rangle. \end{aligned} \tag{4.14}$$

We have

$$\begin{aligned} &2\gamma_k \langle F(x^k + \lambda_k v^k) - F(y^k), x^{k+1} - y^k \rangle \\ &\leq 2\gamma_k \|F(x^k + \lambda_k v^k) - F(y^k)\| \|x^{k+1} - y^k\| \\ &\leq 2\mu \|x^k + \lambda_k v^k - y^k\| \|x^{k+1} - y^k\| \\ &\leq 2\mu (\|x^k - y^k\| + \lambda_k \|v^k\|) \|x^{k+1} - y^k\| \\ &\leq 2\mu \|x^k - y^k\| \|x^{k+1} - y^k\| + 2\mu \lambda_k \|v^k\| \|x^{k+1} - y^k\| \\ &\leq \mu \|x^k - y^k\|^2 + (\mu + \lambda_k \|v^k\|) \|x^{k+1} - y^k\|^2 + \mu^2 \lambda_k \|v^k\|. \end{aligned} \tag{4.15}$$

Similar to (3.8), we can show

$$-2\lambda_k \langle v^k, x^{k+1} - y^k \rangle \leq \lambda_k \|v^k\| + \lambda_k \|v^k\| \|x^{k+1} - y^k\|^2. \tag{4.16}$$

Combining (4.14)-(4.16), we get

$$\begin{aligned} & 2\langle x^k - \gamma_k F(y^k) - y^k, x^{k+1} - y^k \rangle \\ & \leq \mu \|x^k - y^k\|^2 + (\mu + 2\lambda_k \|v^k\|) \|x^{k+1} - y^k\|^2 + (1 + \mu^2) \lambda_k \|v^k\| \\ & \leq \mu \|x^k - y^k\|^2 + (1 - \nu) \|x^{k+1} - y^k\|^2 + 2\lambda_k \|v^k\|, \end{aligned} \tag{4.17}$$

where the last inequality comes from $\lambda_k \|v^k\| < (1 - \mu)/2$ and $\mu < 1$. Substituting (4.17) into (4.13), we get

$$\begin{aligned} \|x^{k+1} - x^*\|^2 & \leq \|x^k - x^*\|^2 - (1 - \mu) \|x^k - y^k\|^2 - \nu \|x^{k+1} - y^k\|^2 + 3\lambda_k \|v^k\| \\ & \quad + \|x^{k+1} - x^*\|^2. \end{aligned} \tag{4.18}$$

Following the proof line of Theorem 3.7, we get $\{x^k\}_{k=0}^\infty$ weakly converges to a solution of the variational equality (1.1). □

By using Theorems 3.8 and 4.3, we obtain the convergence rate of the extragradient method with BP.

Theorem 4.4 *Assume that Conditions 3.1-3.3 hold. Assume the sequence $\{v^k\}_{k=0}^\infty$ is bounded, and the scalars $\{\lambda_k\}_{k=0}^\infty$ are such that $\lambda_k \geq 0$ for all $k \geq 0$, and $\sum_{k=0}^\infty \lambda_k < +\infty$. Let the sequences $\{x^k\}_{k=0}^\infty$ and $\{y^k\}_{k=0}^\infty$ be generated by (4.9) and (4.10). For any integer $t > 0$, we have $y_t \in C$ which satisfies*

$$\langle F(x), y_t - x \rangle \leq \frac{1}{2\Upsilon_t} (\|x - x^0\|^2 + M(x)), \quad \forall x \in C, \tag{4.19}$$

where

$$y_t = \frac{1}{\Upsilon_t} \sum_{k=0}^t \gamma_k y^k, \quad \Upsilon_t = \sum_{k=0}^t \gamma_k, \tag{4.20}$$

and

$$M(x) = \sup_k \{ \max \{ \|x^{k+1} - y^k\|, 3 \|x^{k+1} - x\|^2 \} \} \sum_{k=0}^\infty \lambda_k \|v^k\|. \tag{4.21}$$

4.2 Construction of the inertial extragradient methods by BPR

In this subsection, we construct two classes of inertial extragradient methods by using BPR, *i.e.*, identifying $e_i^k, k = 1, 2$, and λ_k, v^k with special values.

Polyak [29, 30] first introduced the inertial-type algorithms by using the heavy ball method of the second-order dynamical systems in time. Since the inertial-type algorithms

speed up the original algorithms without the inertial effects, recently there has been increasing interest in studying inertial-type algorithms (see, e.g., [31–34]). The authors [35] introduced an inertial extragradient method as follows:

$$\begin{cases} w^k = x^k + \alpha_k(x^k - x^{k-1}), \\ y^k = P_C(w^k - \gamma F(w^k)), \\ x^{k+1} = (1 - \lambda_k)w^k + \lambda_k P_C(w^k - \gamma F(y^k)) \end{cases} \tag{4.22}$$

for each $k \geq 1$, where $\gamma \in (0, 1/L)$, $\{\alpha_k\}$ is nondecreasing with $\alpha_1 = 0$ and $0 \leq \alpha_k \leq \alpha < 1$ for each $k \geq 1$ and $\lambda, \sigma, \delta > 0$ are such that

$$\delta > \frac{\alpha[(1 + \gamma L)^2 \alpha(1 + \alpha) + (1 - \gamma^2 L^2) \alpha \sigma + \sigma(1 + \gamma L)^2]}{1 - \gamma^2 L^2} \tag{4.23}$$

and

$$0 < \lambda \leq \lambda_k \leq \frac{\delta(1 - \gamma^2 L^2) - \alpha[(1 + \gamma L)^2 \alpha(1 + \alpha) + (1 - \gamma^2 L^2) \alpha \sigma + \sigma(1 + \gamma L)^2]}{\delta[(1 + \gamma L)^2 \alpha(1 + \alpha) + (1 - \gamma^2 L^2) \alpha \sigma + \sigma(1 + \gamma L)^2]},$$

where L is the Lipschitz constant of F .

Based on the iterative step (1.2), we construct the following inertial extragradient method:

$$\begin{cases} y^k = P_C(x^k - \gamma_k F(x^k) + \alpha_k^{(1)}(x^k - x^{k-1})), \\ x^{k+1} = P_C(x^k - \gamma_k F(y^k) + \alpha_k^{(2)}(x^k - x^{k-1})), \end{cases} \tag{4.24}$$

where

$$\alpha_k^{(i)} = \begin{cases} \frac{\beta_k^{(i)}}{\|x^k - x^{k-1}\|}, & \text{if } \|x^k - x^{k-1}\| > 1, i = 1, 2, \\ \beta_k^{(i)}, & \text{if } \|x^k - x^{k-1}\| \leq 1. \end{cases} \tag{4.25}$$

Theorem 4.5 *Assume that Conditions 3.1-3.3 hold. Assume that the sequences $\{\beta_k^{(i)}\}_{k=0}^\infty$, $i = 1, 2$, satisfy $\sum_{k=1}^\infty \beta_k^{(i)} < \infty$, $i = 1, 2$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by the inertial extragradient method (4.24) converges weakly to a solution of the variational inequality (1.1).*

Proof Let $e_i^k = \beta_k^{(i)} v^k$, $i = 1, 2$, where

$$v^k = \begin{cases} \frac{x^k - x^{k-1}}{\|x^k - x^{k-1}\|}, & \text{if } \|x^k - x^{k-1}\| > 1, i = 1, 2, \\ x^k - x^{k-1}, & \text{if } \|x^k - x^{k-1}\| \leq 1. \end{cases} \tag{4.26}$$

It is obvious that $\|v^k\| \leq 1$. So, it follows that $\{e_i^k\}$, $i = 1, 2$, satisfy (3.3) from the condition on $\{\beta_k^{(i)}\}$. Using Theorem 3.7, we complete the proof. □

Remark 4.6 From (3.24), we have $\|x^k - x^{k-1}\| \leq 1$ for big enough k , that is, $\alpha_k^{(i)} = \beta_k^{(i)}$.

Using the extragradient method with bounded perturbations (4.9), we construct the following inertial extragradient method:

$$\begin{cases} y^k = P_C(x^k + \alpha_k(x^k - x^{k-1}) - \gamma_k F(x^k + \alpha_k(x^k - x^{k-1}))), \\ x^{k+1} = P_C(x^k + \alpha_k(x^k - x^{k-1}) - \gamma_k F(y^k)), \end{cases} \tag{4.27}$$

where

$$\alpha_k = \begin{cases} \frac{\beta_k}{\|x^k - x^{k-1}\|}, & \text{if } \|x^k - x^{k-1}\| > 1, i = 1, 2, \\ \beta_k, & \text{if } \|x^k - x^{k-1}\| \leq 1. \end{cases} \tag{4.28}$$

We extend Theorem 4.3 to the convergence of the inertial extragradient method (4.27).

Theorem 4.7 *Assume that Conditions 3.1-3.3 hold. Assume that the sequence $\{\beta_k\}_{k=0}^\infty$ satisfies $\sum_{k=1}^\infty \beta_k < \infty$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by the inertial extragradient method (4.27) converges weakly to a solution of the variational inequality (1.1).*

Remark 4.8 The inertial parameter α_k in the inertial extragradient method (4.24) is bigger than that of the inertial extragradient method (4.27). The inertial extragradient method (4.24) becomes the inertial extragradient method (4.27) when $\lambda_k = 1$.

5 The extension to the subgradient extragradient method

In this section, we generalize the results of extragradient method proposed in the previous sections to the subgradient extragradient method.

Censor *et al.* [9] presented the subgradient extragradient method (1.4). In their method the step size is fixed $\gamma \in (0, 1/L)$, where L is a Lipschitz constant of F . So, in order to determine the stepsize γ_k , one needs first calculate (or estimate) L , which might be difficult or even impossible in general. So, in order to overcome this, the Armijo-like search rule can be used:

$$\gamma_k \|F(x^k) - F(y^k)\| \leq \mu \|x^k - y^k\|, \quad \mu \in (0, 1). \tag{5.1}$$

To discuss the convergence of the subgradient extragradient method, we make the following assumptions.

Condition 5.1 The mapping F is monotone on \mathcal{H} , *i.e.*,

$$\langle f(x) - f(y), x - y \rangle \geq 0, \quad \forall x, y \in \mathcal{H}. \tag{5.2}$$

Condition 5.2 The mapping F is Lipschitz continuous on \mathcal{H} with the Lipschitz constant $L > 0$, *i.e.*,

$$\|F(x) - F(y)\| \leq L \|x - y\|, \quad \forall x, y \in \mathcal{H}. \tag{5.3}$$

As before, Censor *et al.*'s subgradient extragradient method [10, Theorem 3.1] can be easily generalized by using some adaptive step rule, for example, (5.1). This result is captured in the next theorem.

Theorem 5.3 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by the subgradient extragradient methods (1.4) and (5.1) weakly converges to a solution of the variational inequality (1.1).*

5.1 The subgradient extragradient method with outer perturbations

In this subsection, we present the subgradient extragradient method with outer perturbations.

Algorithm 5.4 The subgradient extragradient method with outer perturbations:

Step 0: Select a starting point $x^0 \in \mathcal{H}$ and set $k = 0$.

Step 1: Given the current iterate x^k , compute

$$y^k = P_C(x^k - \gamma_k F(x^k) + e_1(x^k)), \tag{5.4}$$

where $\gamma_k = \sigma \rho^{m_k}$, $\sigma > 0$, $\rho \in (0, 1)$ and m_k is the smallest nonnegative integer such that (see [16])

$$\gamma_k \|F(x^k) - F(y^k)\| \leq \mu \|x^k - y^k\|, \quad \mu \in (0, 1). \tag{5.5}$$

Construct the set

$$T_k := \{w \in \mathcal{H} \mid \langle (x^k - \gamma_k F(x^k) + e_1(x^k)) - y^k, w - y^k \rangle \leq 0\}, \tag{5.6}$$

and calculate

$$x^{k+1} = P_{T_k}(x^k - \gamma_k F(y^k) + e_2(x^k)). \tag{5.7}$$

Step 2: If $x^k = y^k$, then stop. Otherwise, set $k \leftarrow (k + 1)$ and return to *Step 1*.

Denote $e_i^k := e_i(x^k)$, $i = 1, 2$. The sequences of perturbations $\{e_i^k\}_{k=0}^\infty$, $i = 1, 2$, are assumed to be summable, i.e.,

$$\sum_{k=0}^\infty \|e_i^k\| < +\infty, \quad i = 1, 2. \tag{5.8}$$

Following the proof of Theorems 3.7 and 3.8, we get the convergence analysis and convergence rate of Algorithm 5.4.

Theorem 5.5 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by Algorithm 5.4 converges weakly to a solution of the variational inequality (1.1).*

Theorem 5.6 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Let the sequences $\{x^k\}_{k=0}^\infty$ and $\{y^k\}_{k=0}^\infty$ be generated by Algorithm 5.4. For any integer $t > 0$, we have $y_t \in C$ which satisfies*

$$\langle F(x), y_t - x \rangle \leq \frac{1}{2\Upsilon_t} (\|x - x^0\|^2 + M(x)), \quad \forall x \in C, \tag{5.9}$$

where

$$y_t = \frac{1}{\Upsilon_t} \sum_{k=0}^t \gamma_k y^k, \quad \Upsilon_t = \sum_{k=0}^t \gamma_k, \tag{5.10}$$

and

$$M(x) = \sup_k \{ \max \{ \|x^{k+1} - y^k\|, \|x^{k+1} - x\| \} \} \sum_{k=0}^{\infty} \|e^k\|. \tag{5.11}$$

5.2 The BPR of the subgradient extragradient method

In this subsection, we investigate the bounded perturbation resilience of the subgradient extragradient method (1.4).

To this end, we treat the right-hand side of (1.4) as the algorithmic operator \mathbf{A}_Ψ of Definition 4.2, namely, we define, for all $k \geq 0$,

$$\mathbf{A}_\Psi(x^k) = P_{T(x^k)}(x^k - \gamma_k F(P_C(x^k - \gamma_k F(x^k))))), \tag{5.12}$$

where γ_k satisfies (5.1) and

$$T(x^k) = \{w \in \mathcal{H} \mid \langle (x^k - \gamma_k F(x^k)) - y^k, w - y^k \rangle \leq 0\}. \tag{5.13}$$

Identify the solution set Ψ with the solution set of the variational inequality (1.1) and identify the additional set Θ with C .

According to Definition 4.2, we need to show the convergence of the sequence $\{x^k\}_{k=0}^\infty$ that, starting from any $x^0 \in \mathcal{H}$, is generated by

$$x^{k+1} = P_{T(x^k + \lambda_k v^k)}((x^k + \lambda_k v^k) - \gamma_k F(P_C((x^k + \lambda_k v^k) - \gamma_k F(x^k + \lambda_k v^k))))), \tag{5.14}$$

which can be rewritten as

$$\begin{cases} y^k = P_C((x^k + \lambda_k v^k) - \gamma_k F((x^k + \lambda_k v^k))), \\ T(x^k + \lambda_k v^k) = \{w \in \mathcal{H} \mid \langle (x^k + \lambda_k v^k) - \gamma_k F(x^k + \lambda_k v^k) - y^k, w - y^k \rangle \leq 0\}, \\ x^{k+1} = P_{T(x^k + \lambda_k v^k)}((x^k + \lambda_k v^k) - \gamma_k F(y^k)), \end{cases} \tag{5.15}$$

where $\gamma_k = \sigma \rho^{m_k}$, $\sigma > 0$, $\rho \in (0, 1)$ and m_k is the smallest nonnegative integer such that

$$\gamma_k \|F(x^k + \lambda_k v^k) - F(y^k)\| \leq \mu (\|x^k - y^k\| + \lambda_k \|v^k\|), \quad \mu \in (0, 1). \tag{5.16}$$

The sequences $\{v^k\}_{k=0}^\infty$ and $\{\lambda_k\}_{k=0}^\infty$ obey conditions (i) and (ii) in Definition 4.2, respectively, and also (iii) in Definition 4.2 is satisfied.

The next theorem establishes the bounded perturbation resilience of the subgradient extragradient method. Since its proof is similar to that of Theorem 4.3, we omit it.

Theorem 5.7 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Assume the sequence $\{v^k\}_{k=0}^\infty$ is bounded, and the scalars $\{\lambda_k\}_{k=0}^\infty$ are such that $\lambda_k \geq 0$ for all $k \geq 0$, and $\sum_{k=0}^\infty \lambda_k < +\infty$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by (5.15) and (5.16) converges weakly to a solution of the variational inequality (1.1).*

We also get the convergence rate of the subgradient extragradient methods with BP (5.15) and (5.16).

Theorem 5.8 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Assume the sequence $\{v^k\}_{k=0}^\infty$ is bounded, and the scalars $\{\lambda_k\}_{k=0}^\infty$ are such that $\lambda_k \geq 0$ for all $k \geq 0$, and $\sum_{k=0}^\infty \lambda_k < +\infty$. Let the sequences $\{x^k\}_{k=0}^\infty$ and $\{y^k\}_{k=0}^\infty$ be generated by (5.15) and (5.16). For any integer $t > 0$, we have $y_t \in C$ which satisfies*

$$\langle F(x), y_t - x \rangle \leq \frac{1}{2\Upsilon_t} (\|x - x^0\|^2 + M(x)), \quad \forall x \in C, \tag{5.17}$$

where

$$y_t = \frac{1}{\Upsilon_t} \sum_{k=0}^t \gamma_k y^k, \quad \Upsilon_t = \sum_{k=0}^t \gamma_k, \tag{5.18}$$

and

$$M(x) = \sup_k \{ \max \{ \|x^{k+1} - y^k\|, 3\|x^{k+1} - x\|^2 \} \} \sum_{k=0}^\infty \lambda_k \|v^k\|. \tag{5.19}$$

5.3 Construction of the inertial subgradient extragradient methods by BPR

In this subsection, we construct two classes of inertial subgradient extragradient methods by using BPR, *i.e.*, identifying e_i^k , $k = 1, 2$, and λ_k, v^k with special values.

Based on Algorithm 5.4, we construct the following inertial subgradient extragradient method:

$$\begin{cases} y^k = P_C(x^k - \gamma_k F(x^k) + \alpha_k^{(1)}(x^k - x^{k-1})), \\ T_k := \{w \in \mathcal{H} \mid \langle (x^k - \gamma_k F(x^k) + \alpha_k^{(1)}(x^k - x^{k-1})) - y^k, w - y^k \rangle \leq 0\}, \\ x^{k+1} = P_{T_k}(x^k - \gamma_k F(y^k) + \alpha_k^{(2)}(x^k - x^{k-1})), \end{cases} \tag{5.20}$$

where γ_k satisfies (5.16) and

$$\alpha_k^{(i)} = \begin{cases} \frac{\beta_k^{(i)}}{\|x^k - x^{k-1}\|}, & \text{if } \|x^k - x^{k-1}\| > 1, i = 1, 2, \\ \beta_k^{(i)}, & \text{if } \|x^k - x^{k-1}\| \leq 1. \end{cases} \tag{5.21}$$

Similar to the proof of Theorem 4.5, we get the convergence of the inertial subgradient extragradient method (5.20).

Theorem 5.9 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Assume that the sequences $\{\beta_k^{(i)}\}_{k=0}^\infty$, $i = 1, 2$, satisfy $\sum_{k=1}^\infty \beta_k^{(i)} < \infty$, $i = 1, 2$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by the inertial subgradient extragradient method (5.20) converges weakly to a solution of the variational inequality (1.1).*

Using the subgradient extragradient method with bounded perturbations (5.15), we construct the following inertial subgradient extragradient method:

$$\begin{cases} w^k = x^k + \alpha_k(x^k - x^{k-1}), \\ y^k = P_C(w^k - \gamma_k F(w^k)), \\ T_k := \{w \in \mathcal{H} \mid \langle (w^k - \gamma_k F(w^k)) - y^k, w - y^k \rangle \leq 0\}, \\ x^{k+1} = P_{T_k}(w^k - \gamma_k F(y^k)), \end{cases} \tag{5.22}$$

where $\gamma_k = \sigma \rho^{m_k}$, $\sigma > 0$, $\rho \in (0, 1)$ and m_k is the smallest nonnegative integer such that

$$\gamma_k \|F(w^k) - F(y^k)\| \leq \mu \|w^k - y^k\|, \quad \mu \in (0, 1), \tag{5.23}$$

and

$$\alpha_k = \begin{cases} \frac{\beta_k}{\|x^k - x^{k-1}\|}, & \text{if } \|x^k - x^{k-1}\| > 1, i = 1, 2, \\ \beta_k, & \text{if } \|x^k - x^{k-1}\| \leq 1. \end{cases} \tag{5.24}$$

We extend Theorem 4.3 to the convergence of the inertial subgradient extragradient method (5.22).

Theorem 5.10 *Assume that Conditions 3.1, 5.1 and 5.2 hold. Assume that the sequence $\{\beta_k\}_{k=0}^\infty$ satisfies $\sum_{k=1}^\infty \beta_k < \infty$. Then the sequence $\{x^k\}_{k=0}^\infty$ generated by the inertial subgradient extragradient method (5.22) converges weakly to a solution of the variational inequality (1.1).*

6 Numerical experiments

In this section, we provide three examples to compare the inertial extragradient method (4.22) (iEG1), the inertial extragradient method (4.24) (iEG2), the inertial extragradient method (4.27) (iEG), the extragradient method (1.2), the inertial subgradient extragradient method (5.20) (iSEG1), the inertial subgradient extragradient method (5.22) (iSEG2) and the subgradient extragradient method (1.4).

In the first example, we consider a typical sparse signal recovery problem. We choose the following set of parameters. Take $\sigma = 5$, $\rho = 0.9$ and $\mu = 0.7$. Set

$$\alpha_k = \alpha_k^{(i)} = \frac{1}{k^2} \quad \text{if } \|x^k - x^{k-1}\| \leq 1 \tag{6.1}$$

in the inertial extragradient methods (4.22) and (4.24), and the inertial subgradient extragradient methods (5.20) and (5.22). Choose $\alpha_k = 0.35$ and $\lambda_k = 0.8$ in the inertial extragradient method (4.24).

Example 6.1 Let $x_0 \in R^n$ be a K -sparse signal, $K \ll n$. The sampling matrix $A \in R^{m \times n}$ ($m < n$) is stimulated by the standard Gaussian distribution and a vector $b = Ax_0 + e$, where e is additive noise. When $e = 0$, it means that there is no noise to the observed data. Our task is to recover the signal x_0 from the data b .

It is well known that the sparse signal x_0 can be recovered by solving the following LASSO problem [36]:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & \frac{1}{2} \|Ax - b\|_2^2 \\ \text{s.t.} & \|x\|_1 \leq t, \end{aligned} \tag{6.2}$$

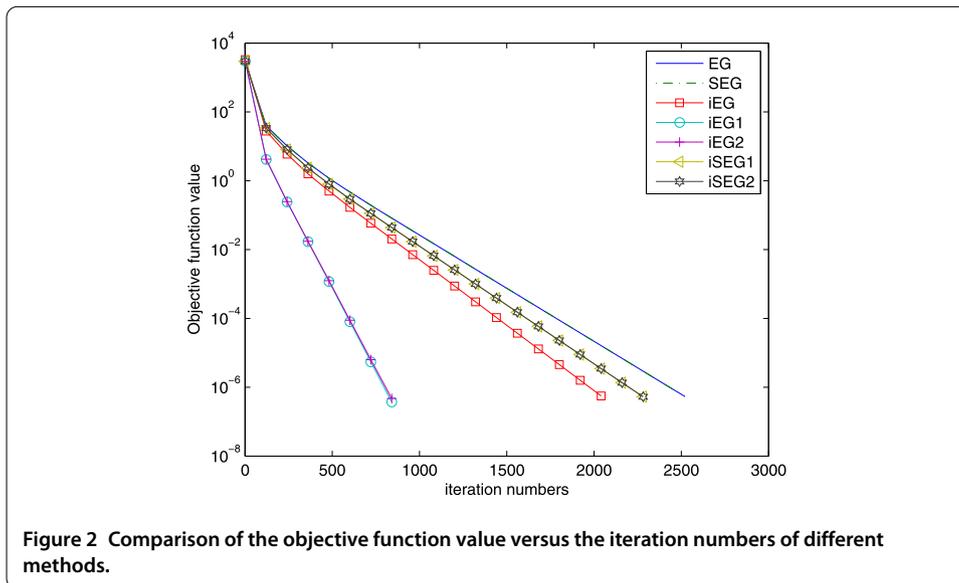
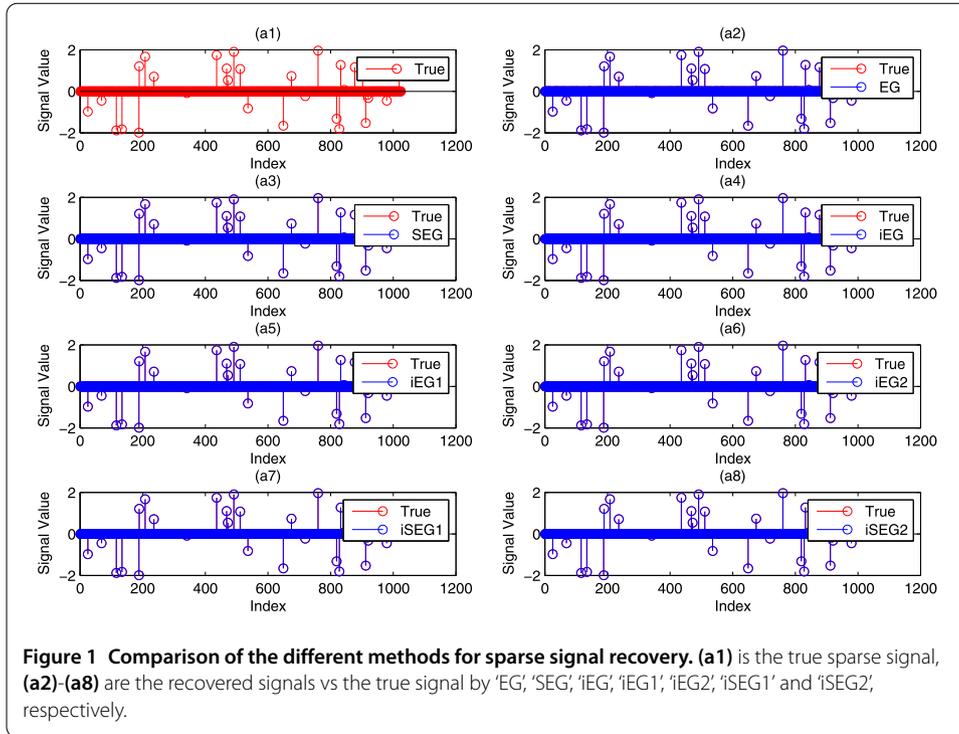
where $t > 0$. It is easy to see that the optimization problem (6.1) is a special case of the variational inequality problem (1.1), where $F(x) = A^T(Ax - b)$ and $C = \{x \mid \|x\|_1 \leq t\}$. We can use the proposed iterative algorithms to solve the optimization problem (6.1). Although the orthogonal projection onto the closed convex set C does not have a closed-form solution, the projection operator P_C can be precisely computed in a polynomial time. We include the details of computing P_C in Appendix. We conduct plenty of simulations to compare the performances of the proposed iterative algorithms. The following inequality was defined as the stopping criterion:

$$\|x^{k+1} - x^k\| \leq \epsilon,$$

where $\epsilon > 0$ is a given small constant. ‘*Iter*’ denotes the iteration numbers. ‘*Obj*’ represents the objective function value and ‘*Err*’ is the 2-norm error between the recovered signal and the true K -sparse signal. We divide the experiments into two parts. One task is to recover the sparse signal x_0 from noise observation vector b , and the other is to recover the sparse signal from noiseless data b . For the noiseless case, the obtained numerical results are reported in Table 1. To visually view the results, Figure 1 shows the recovered signal compared with the true signal x_0 when $K = 30$. We can see from Figure 1 that the

Table 1 Numerical results obtained by the proposed iterative algorithms when $m = 240$, $n = 1,024$ in the noiseless case

K-sparse signal	Methods	$\epsilon = 10^{-4}$			$\epsilon = 10^{-6}$		
		<i>Iter</i>	<i>Obj</i>	<i>Err</i>	<i>Iter</i>	<i>Obj</i>	<i>Err</i>
K = 20	EG	444	9.7346e-4	0.0080	817	9.6625e-8	7.9856e-5
	SEG	444	9.7272e-4	0.0080	817	9.6555e-8	7.9827e-5
	iEG	374	6.2389e-4	0.0064	675	6.3456e-8	6.4715e-5
	iEG1	159	7.0799e-5	0.0021	263	7.4280e-9	2.2041e-5
	iEG2	158	8.3897e-5	0.0023	273	1.0889e-8	2.6809e-5
	iSEG1	415	8.9563e-4	0.0076	787	2.3571e-7	5.2470e-5
	iSEG2	414	9.2167e-4	0.0077	760	9.1586e-8	7.7275e-5
	K = 30	EG	1,285	0.0035	0.0281	2,583	3.4535e-7
SEG		1,285	0.0035	0.0281	2,583	3.4534e-7	2.8035e-4
iEG		1,091	0.0023	0.0227	2,144	2.2732e-7	2.2745e-4
iEG1		532	3.7493e-4	0.0092	944	3.7522e-8	9.2287e-5
iEG2		535	3.7961e-4	0.0093	956	4.3181e-8	9.3120e-5
iSEG1		1,176	0.0031	0.0266	2,351	3.1038e-7	2.6137e-4
iSEG2		1,176	0.0031	0.0266	2,346	3.1635e-7	2.6784e-4
K = 40		EG	1,729	0.0050	0.0405	3,599	5.0237e-7
	SEG	1,729	0.0050	0.0405	3,599	5.0228e-7	4.0484e-4
	iEG	1,473	0.0033	0.0328	2,990	3.3182e-7	3.2905e-4
	iEG1	744	5.4838e-4	0.0134	1,361	5.5456e-8	1.3440e-4
	iEG2	745	5.4807e-4	0.0134	1,355	6.4785e-8	1.4191e-4
	iSEG1	1,570	0.0045	0.0384	3,246	4.5079e-7	3.8146e-4
	iSEG2	1,572	0.0045	0.0382	3,244	4.5389e-7	3.8435e-4

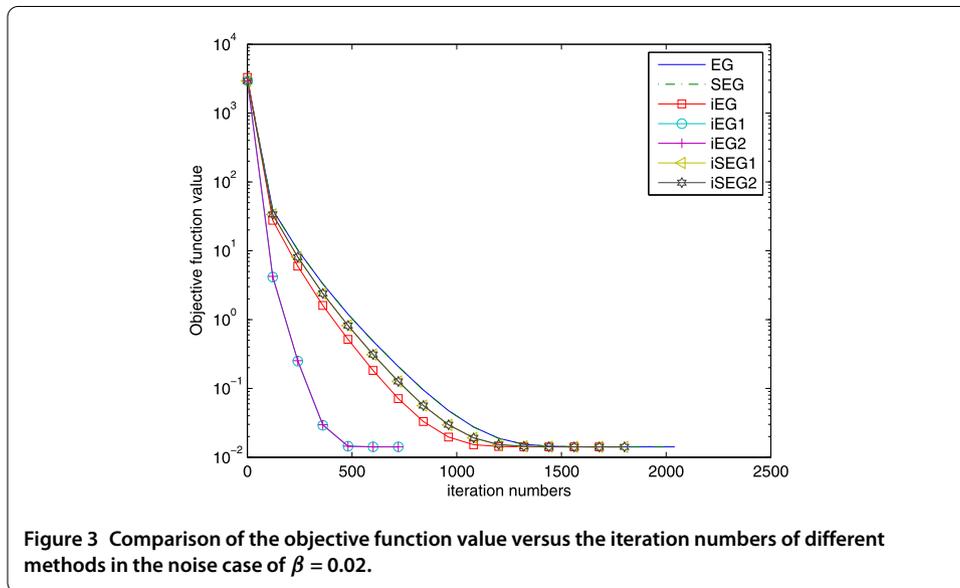


recovered signal is the same as the true signal. Further, Figure 2 presents the objective function value versus the iteration numbers.

For the noise observation b , we assume that the vector e is corrupted by Gaussian noise with zero mean and β variances. The system matrix A is the same as in the noiseless case, and the sparsity level $K = 30$. We list the numerical results for different noise level β in Table 2. When the noise $\beta = 0.02$, Figure 3 shows the objective function value versus the iteration numbers. Figure 4 shows the recovered signal vs the true signal in the noise case.

Table 2 Numerical results for the proposed iterative algorithms with different noise value β

Variances	Methods	$\epsilon = 10^{-4}$			$\epsilon = 10^{-6}$		
		Iter	Obj	Err	Iter	Obj	Err
$\beta = 0.01$	EG	1,264	0.0092	0.0317	2,192	0.0061	0.0131
	SEG	1,264	0.0092	0.0317	2,192	0.0061	0.0131
	iEG	1,070	0.0081	0.0272	1,812	0.0061	0.0131
	iEG1	519	0.0063	0.0164	788	0.0061	0.0130
	iEG2	516	0.0063	0.0166	786	0.0061	0.0130
	iSEG1	1,156	0.0089	0.0305	1,995	0.0061	0.0131
	iSEG2	1,157	0.0089	0.0304	1,990	0.0061	0.0131
$\beta = 0.02$	EG	1,274	0.0163	0.0387	2,086	0.0142	0.0272
	SEG	1,274	0.0163	0.0387	2,086	0.0142	0.0272
	iEG	1,070	0.0154	0.0356	1,728	0.0142	0.0272
	iEG1	492	0.0144	0.0300	756	0.0142	0.0272
	iEG2	495	0.0143	0.0300	759	0.0142	0.0272
	iSEG1	1,163	0.0161	0.0378	1,899	0.0142	0.0272
	iSEG2	1,161	0.0161	0.0380	1,895	0.0142	0.0272
$\beta = 0.05$	EG	1,190	0.1012	0.0749	1,869	0.0991	0.0651
	SEG	1,190	0.1012	0.0749	1,869	0.0991	0.0651
	iEG	996	0.1005	0.0727	1,542	0.0991	0.0650
	iEG1	460	0.0993	0.0677	670	0.0991	0.0650
	iEG2	461	0.0993	0.0675	665	0.0991	0.0650
	iSEG1	1,084	0.1010	0.0742	1,704	0.0991	0.0651
	iSEG2	1,084	0.1010	0.0742	1,704	0.0991	0.0651

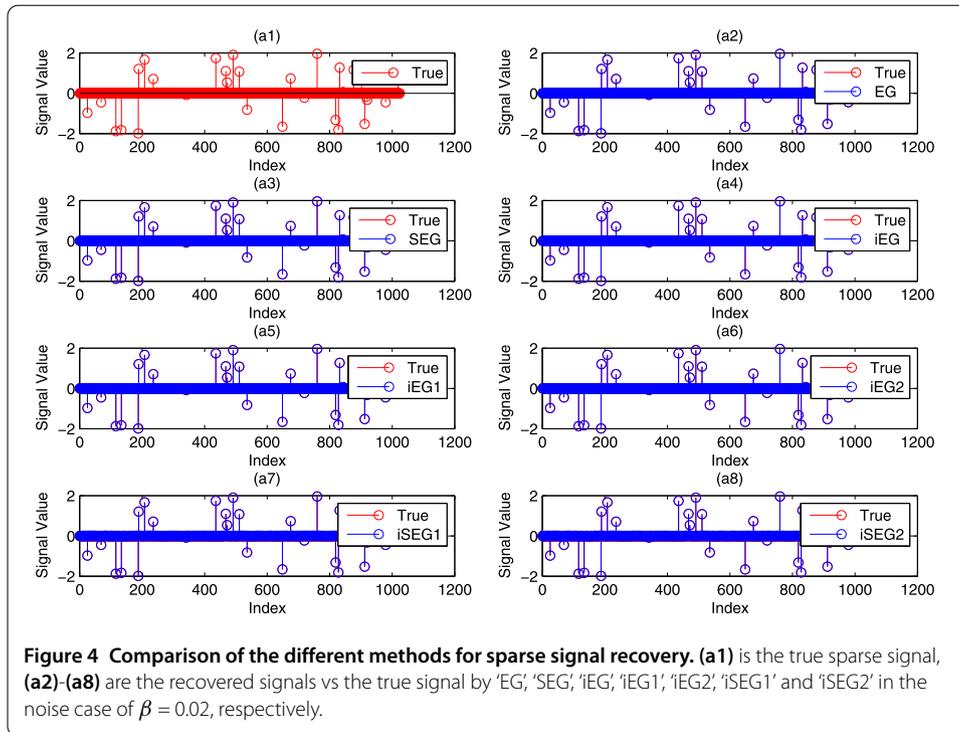


Example 6.2 Let $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be defined by

$$F(x, y) = (2x + 2y + \sin(x), -2x + 2y + \sin(y)), \quad \forall x, y \in \mathbb{R}. \tag{6.3}$$

The authors [37] proved that F is Lipschitz continuous with $L = \sqrt{26}$ and 1-strongly monotone. Therefore the variational inequality (1.1) has a unique solution, and $(0, 0)$ is its solution.

Let $C = \{x \in \mathbb{R}^2 \mid e_1 \leq x \leq e_2\}$, where $e_1 = (-10, -10)$ and $e_2 = (100, 100)$. Take the initial point $x_0 = (-100, 10) \in \mathbb{R}^2$. Since $(0, 0)$ is the unique solution of the variational inequality (1.1), denote by $D_k := \|x^k\| \leq 10^{-5}$ the stopping criterion.



Example 6.3 Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by $F(x) = Ax + b$, where $A = Z^T Z$, $Z = (z_{ij})_{n \times n}$ and $b = (b_i) \in \mathbb{R}^n$, where $z_{ij} \in (0, 1)$ and $b_i \in (0, 1)$ are generated randomly.

It is easy to verify that F is L -Lipschitz continuous and η -strongly monotone with $L = \max(\text{eig}(A))$ and $\eta = \min(\text{eig}(A))$.

Let $C := \{x \in \mathbb{R}^n \mid \|x - d\| \leq r\}$, where the center

$$d \in [(-10, -10, \dots, -10), (10, 10, \dots, 10)] \subset \mathbb{R}^n \tag{6.4}$$

and radius $r \in (0, 10)$ are randomly chosen. Take the initial point $x_0 = (c_i) \in \mathbb{R}^n$, where $c_i \in [0, 2]$ is generated randomly. Set $n = 100$. Take $\rho = 0.4$ and other parameters are set the same values as in Example 6.2. Although the variational inequality (1.1) has a unique solution, it is difficult to get the exact solution. So, denote by $D_k := \|x^{k+1} - x^k\| \leq 10^{-5}$ the stopping criterion.

From Figures 5 and 6, we conclude: (i) the inertial-type algorithms improve the original algorithms; (ii) the performances of the inertial extragradient methods (4.22) and (4.24) are almost the same; (iii) the inertial subgradient extragradient method (5.20) performs better than the inertial subgradient extragradient method (5.22) for Example 6.1, while they are almost the same for Example 6.2; (iv) the (inertial) extragradient methods behave better than the (inertial) subgradient extragradient methods since the sets C in Examples 6.2 and 6.3 are simple, and hence the computation load of the projection onto it is small; (v) the inertial extragradient method (4.22) has an advantage over the inertial extragradient methods (4.22) and (4.24). The reason may be that it takes a bigger inertial parameter α_k .

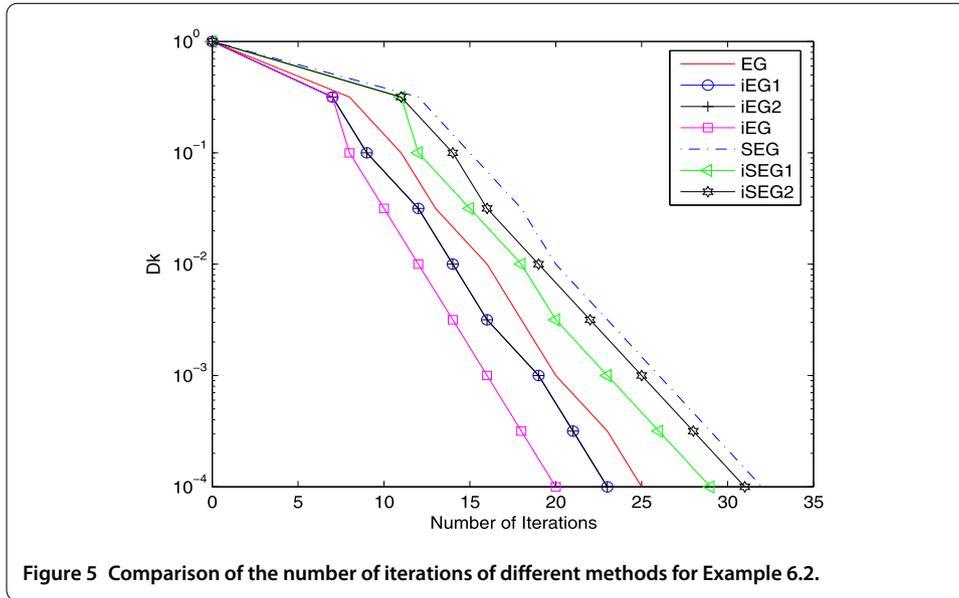


Figure 5 Comparison of the number of iterations of different methods for Example 6.2.

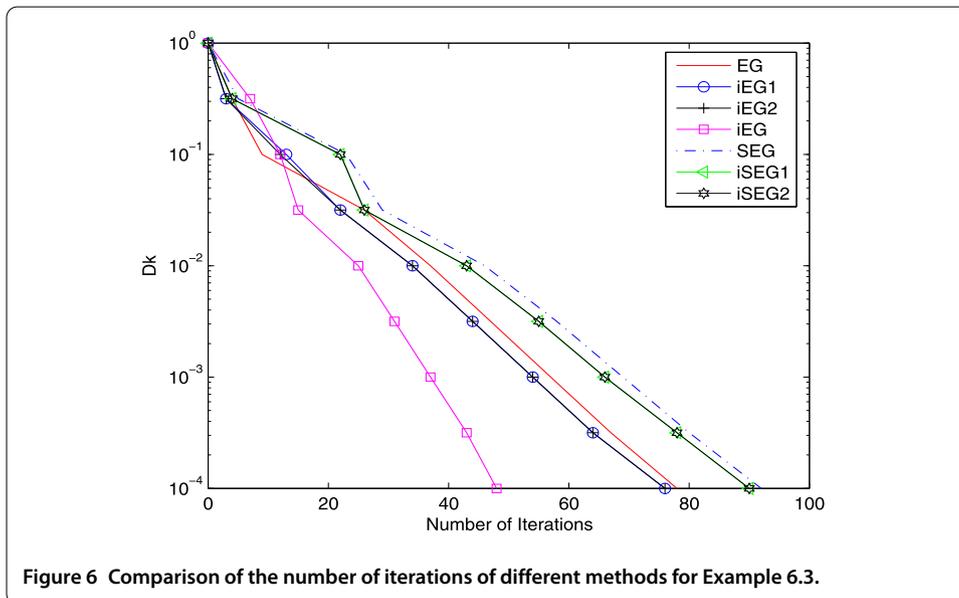


Figure 6 Comparison of the number of iterations of different methods for Example 6.3.

7 Conclusions

In this research article we study an important property of iterative algorithms for solving variational inequality (VI) problems which is called bounded perturbation resilience. In particular, we focus on extragradient-type methods. This enables us to develop inexact versions of the methods as well as apply the superiorization methodology in order to obtain a ‘superior’ solution to the original problem. In addition, some inertial extragradient methods are also derived. All the presented methods converge at the rate of $O(1/t)$, and three numerical examples illustrate, demonstrate and compare the performances of all the algorithms.

Appendix

In this part, we present the details of computing a vector $y \in R^n$ onto the ℓ_1 -norm ball constraint. For convenience, we consider projection onto the unit ℓ_1 -norm ball first. Then we extend it to the general ℓ_1 -norm ball constraint.

The projection onto the unit ℓ_1 -norm ball is to solve the optimization problem

$$\begin{aligned} \min_{x \in R^n} & \frac{1}{2} \|x - y\|_2^2 \\ \text{s.t.} & \|x\|_1 \leq 1. \end{aligned}$$

The above optimization problem is a typical constrained optimization problem, we consider solving it based on the Lagrangian method. Define the Lagrangian function $L(x, \lambda)$ as

$$L(x, \lambda) = \frac{1}{2} \|x - y\|_2^2 + \lambda (\|x\|_1 - 1).$$

Let (x^*, λ^*) be the optimal primal and dual pair. It satisfies the KKT conditions of

$$\begin{aligned} 0 & \in (x^* - y) + \lambda^* \partial(\|x^*\|_1), \\ \lambda^* (\|x^*\|_1 - 1) & = 0, \\ \lambda^* & \geq 0. \end{aligned}$$

It is easy to check that if $\|y\|_1 \leq 1$, then $x^* = y$ and $\lambda^* = 0$. In the following, we assume $\|y\|_1 > 1$. Based on the KKT conditions, we obtain $\lambda^* > 0$ and $\|x^*\|_1 = 1$. From the first order optimality, we have $x^* = \max\{|y| - \lambda^*, 0\} \otimes \text{Sign}(y)$, where \otimes represents element-wise multiplication and $\text{Sign}(\cdot)$ denotes the symbol function, *i.e.*, $\text{Sign}(y_i) = 1$ if $y_i \geq 0$; otherwise $\text{Sign}(y_i) = -1$.

Define a function $f(\lambda) = \|x(\lambda)\|_1$, where $x(\lambda) = S_\lambda(y) = \max\{|y| - \lambda, 0\} \otimes \text{Sign}(y)$. We prove the following lemma.

Lemma A.1 *For the function $f(\lambda)$, there must exist $\lambda^* > 0$ such that $f(\lambda^*) = 1$.*

Proof Since $f(0) = \|S_0(y)\|_1 = \|y\|_1 > 1$. Let $\lambda^+ = \max_{1 \leq i \leq n} \{|y_i|\}$, then $f(\lambda^+) = 0 < 1$. Notice that $f(\lambda)$ is decreasing and convex. Therefore, by the intermediate value theorem, there exists $\lambda^* > 0$ such that $f(\lambda^*) = 1$. □

To find λ^* such that $f(\lambda^*) = 1$, we follow the following steps.

Step 1. Define a vector \bar{y} with the same element as $|y|$, which was sorted in descending order. That is, $\bar{y}_1 \geq \bar{y}_2 \geq \dots \geq \bar{y}_n \geq 0$.

Step 2. For every $k = 1, 2, \dots, n$, solve the equation $\sum_{i=1}^k \bar{y}_i - k\lambda = 1$. Stop search until the solution λ^* belongs to the interval $[\bar{y}_{k+1}, \bar{y}_k]$.

In conclusion, the optimal x^* can be computed by $x^* = \max\{|y| - \lambda^*, 0\} \otimes \text{Sign}(y)$. The next lemma extends the projection onto the unit ℓ_1 -norm ball to the general ℓ_1 -norm ball constraint.

Lemma A.2 *Let $C_1 = \{x \mid \|x\|_1 \leq 1\}$. For any $t > 0$, define a general ℓ_1 -norm ball constraint set $C = \{x \mid \|x\|_1 \leq t\}$. Then, for any vector $y \in R^n$, we have*

$$P_C(y) = tP_{C_1}\left(\frac{y}{t}\right).$$

Proof To compute the projection $P_C(y)$ is to solve the optimization problem

$$P_C(y) = \arg \min_{x \in R^n} \frac{1}{2} \|x - y\|_2^2$$

s.t. $\|x\|_1 \leq t$.

For any $x \in C$, let $\bar{x} = \frac{x}{t}$, it follows that $\bar{x} \in C_1$. The optimal solution x^* of the above optimization problem satisfies $x^* = P_C(y) = t\bar{x}^*$, where \bar{x}^* is the optimal solution of the optimization problem of

$$\bar{x}^* = \arg \min_{\bar{x} \in R^n} \frac{1}{2} \left\| \bar{x} - \frac{y}{t} \right\|_2^2$$

s.t. $\|\bar{x}\|_1 \leq 1$.

It is observed that \bar{x}^* is an exact projection onto the closed convex set C_1 . That is, $\bar{x}^* = P_{C_1}\left(\frac{y}{t}\right)$. This completes the proof. □

Acknowledgements

We wish to thank the anonymous referees for their thorough analysis and review. All their comments and suggestions helped tremendously in improving the quality of this paper and made it suitable for publication.

Funding

The first author is supported by the National Natural Science Foundation of China (No. 61379102) and the Open Fund of Tianjin Key Lab for Advanced Signal Processing (No. 2016ASP-TJ01). The third author is supported by the EU FP7 IRSES program STREVCOMS, grant no. PIRSES-GA-2013-612669. The fourth author is supported by Visiting Scholarship of Academy of Mathematics and Systems Science, Chinese Academy of Sciences (AM201622C04) and the National Natural Science Foundation of China (11401293, 11661056), the Natural Science Foundation of Jiangxi Province (20151BAB211010).

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors contributed equally to the writing of this paper. All authors read and approved the final manuscript.

Author details

¹Tianjin Key Laboratory for Advanced Signal Processing, College of Science, Civil Aviation University of China, Tianjin, 300300, P.R. China. ²Department of Mathematics, ORT Braude College, Karmiel, 2161002, Israel. ³Department of Mathematics, NanChang University, Nanchang, 330031, P.R. China.

Publisher's Note

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

Received: 23 August 2017 Accepted: 31 October 2017 Published online: 10 November 2017

References

1. Ceng, LC, Liou, YC, Yao, JC, Yao, YH: Well-posedness for systems of time-dependent hemivariational inequalities in Banach spaces. *J. Nonlinear Sci. Appl.* **10**, 4318-4336 (2017)
2. Zegeye, H, Shahzad, N, Yao, YH: Minimum-norm solution of variational inequality and fixed point problem in Banach spaces. *Optimization* **64**, 453-471 (2015)
3. Yao, YH, Liou, YC, Yao, JC: Iterative algorithms for the split variational inequality and fixed point problems under nonlinear transformations. *J. Nonlinear Sci. Appl.* **10**, 843-854 (2017)

4. Yao, YH, Noor, MA, Liou, YC, Kang, SM: Iterative algorithms for general multi-valued variational inequalities. *Abstr. Appl. Anal.* **2012**, Article ID 768272 (2012)
5. Yao, YH, Postolache, M, Liou, YC, Yao, Z-S: Construction algorithms for a class of monotone variational inequalities. *Optim. Lett.* **10**, 1519-1528 (2016)
6. Korpelevich, GM: The extragradient method for finding saddle points and other problems. *Èkon. Mat. Metody* **12**, 747-756 (1976)
7. Antipin, AS: On a method for convex programs using a symmetrical modification of the Lagrange function. *Èkon. Mat. Metody* **12**, 1164-1173 (1976)
8. Facchinei, F, Pang, JS: *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. II. Springer, New York (2003)
9. Censor, Y, Gibali, A, Reich, S: The subgradient extragradient method for solving variational inequalities in Hilbert space. *J. Optim. Theory Appl.* **148**, 318-335 (2011)
10. Censor, Y, Gibali, A, Reich, S: Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. *Optim. Methods Softw.* **6**, 827-845 (2011)
11. Censor, Y, Gibali, A, Reich, S: Extensions of Korpelevich's extragradient method for solving the variational inequality problem in Euclidean space. *Optimization* **61**, 1119-1132 (2012)
12. Goebel, K, Reich, S: *Uniform Convexity, Hyperbolic Geometry, and Nonexpansive Mappings*. Marcel Dekker, New York (1984)
13. Rockafellar, RT: Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.* **14**(5), 877-898 (1976)
14. Bauschke, HH, Borwein, JM: On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38**, 367-426 (1996)
15. Bauschke, HH, Combettes, PL: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, Berlin (2011)
16. Khobotov, EN: Modification of the extragradient method for solving variational inequalities and certain optimization problems. *USSR Comput. Math. Math. Phys.* **27**, 120-127 (1987)
17. Zhao, J, Yang, Q: Self-adaptive projection methods for the multiple-sets split feasibility problem. *Inverse Probl.* **27**, Article ID 035009 (2011)
18. Nemirovski, A: Prox-method with rate of convergence $O(1/t)$ for variational inequality with Lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM J. Optim.* **15**, 229-251 (2005)
19. Tseng, P: On accelerated proximal gradient methods for convex-concave optimization. Department of Mathematics, University of Washington, Seattle, WA 98195, USA (2008)
20. Censor, Y, Zaslavski, AJ: Strict Fejér monotonicity by superiorization of feasibility-seeking projection methods. *J. Optim. Theory Appl.* **165**, 172-187 (2015)
21. Censor, Y, Davidi, R, Herman, GT: Perturbation resilience and superiorization of iterative algorithms. *Inverse Probl.* **26**, Article ID 065008 (2010)
22. Herman, GT, Davidi, R: Image reconstruction from a small number of projections. *Inverse Probl.* **24**, Article ID 045011 (2008)
23. Butnariu, D, Reich, S, Zaslavski, AJ: Convergence to fixed points of inexact orbits of Bregman-monotone and of nonexpansive operators in Banach spaces. In: *Fixed Point Theory and Its Applications*, pp. 11-32. Yokohama Publishers, Yokohama (2006)
24. Butnariu, D, Reich, S, Zaslavski, AJ: Asymptotic behavior of inexact orbits for a class of operators in complete metric spaces. *J. Appl. Anal.* **13**, 1-11 (2007)
25. Butnariu, D, Reich, S, Zaslavski, AJ: Stable convergence theorems for infinite products and powers of nonexpansive mappings. *Numer. Funct. Anal. Optim.* **29**, 304-323 (2008)
26. Herman, GT, Garduño, E, Davidi, R, Censor, Y: Superiorization: an optimization heuristic for medical physics. *Med. Phys.* **39**, 5532-5546 (2012)
27. Garduño, E, Herman, GT: Superiorization of the ML-EM algorithm. *IEEE Trans. Nucl. Sci.* **61**, 162-172 (2014)
28. Dong, QL, Lu, YY, Yang, J: The extragradient algorithm with inertial effects for solving the variational inequality. *Optimization* **65**, 2217-2226 (2016)
29. Polyak, BT: Some methods of speeding up the convergence of iteration methods. *USSR Comput. Math. Math. Phys.* **4**(5), 1-17 (1964)
30. Polyak, BT: *Introduction to Optimization*. Optimization Software, New York (1987)
31. Alvarez, F: Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space. *SIAM J. Optim.* **14**(3), 773-782 (2004)
32. Attouch, H, Peypouquet, J, Redont, P: A dynamical approach to an inertial forward-backward algorithm for convex minimization. *SIAM J. Optim.* **24**(1), 232-256 (2014)
33. Bot, RI, Csetnek, ER: A hybrid proximal-extragradient algorithm with inertial effects. *Numer. Funct. Anal. Optim.* **36**, 951-963 (2015)
34. Ochs, P, Brox, T, Pock, T: iPiasco: inertial proximal algorithm for strongly convex optimization. *J. Math. Imaging Vis.* **53**, 171-181 (2015)
35. Dong, QL, Yang, J, Yuan, HB: The projection and contraction algorithm for solving variational inequality problems in Hilbert spaces. *J. Nonlinear Convex Anal.* (accepted)
36. Tibshirani, R: Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267-288 (1996)
37. Dong, QL, Cho, YJ, Zhong, LL, Rassias, ThM: Inertial projection and contraction algorithms for variational inequalities. *J. Glob. Optim.* (accepted)