J Glob Optim (2007) 39:155–169
DOI 10.1007/s10898-006-9129-6

ORIGINAL ARTICLE

# An improved general extra-gradient method with refined step size for nonlinear monotone variational inequalities

**M. H. Xu · X. M. Yuan · Q. L. Huang**

Received: 8 April 2006 / Accepted: 6 December 2006 / Published online: 26 January 2007


**Abstract** Extra-gradient method and its modified versions are direct methods for variational inequalities $VI(\Omega,\ F)$ that only need to use the value of function $F$ in the iterative processes. This property makes the type of extra-gradient methods very practical for some variational inequalities arising from the real-world, in which the function $F$ usually does not have any explicit expression and only its value can be observed and/or evaluated for given variable. Generally, such observation and/or evaluation may be obtained via some costly experiments. Based on this view of point, reducing the times of observing the value of function $F$ in those methods is meaningful in practice. In this paper, a new strategy for computing step size is proposed in general extra-gradient method. With the new step size strategy, the general extra-gradient method needs to cost a relatively minor amount of computation to obtain a new step size, and can achieve the purpose of saving the amount of computing the value of $F$ in solving $VI(\Omega,\ F)$. Further, the convergence analysis of the new algorithm and the properties related to the step size strategy are also discussed in this paper. Numerical experiments are given and show that the amount of computing the value of function $F$ in solving $VI(\Omega,\ F)$ can be saved about 12–25% by the new general extra-gradient method.

**Keywords** Nonlinear monotone variational inequality · Extra-gradient method · Prediction-correction method · Projection contraction method

M. H. Xu
Department of Mathematics, Nanjing University, Jiangsu Province 210093, China

X. M. Yuan
Department of Management Science, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200052, China

M. H. Xu (✉) · Q. L. Huang
Department of Information Science, Jiangsu Polytechnic University, Changzhou, Jiangsu Province, 213164, China
e-mail: xuminghua@jpu.edu.cn

## 1 Introduction

Consider the following variational inequality $VI(\Omega, F)$: Find $u^* \in \Omega$ such that

$$(u - u^*)^T F(u^*) \geq 0, \quad \forall\, u \in \Omega, \tag{1.1}$$

where $\Omega$ is a closed convex subset of $R^n$, $F: \Omega \mapsto R^n$ is monotone, i.e., for all $u, v \in R^n$

$$(u - v)^T (F(u) - F(v)) \geq 0.$$

It is well known that variational inequality $VI(\Omega, F)$ includes nonlinear complementarity problems (when $\Omega = R^n_+$) and system of nonlinear equations (when $\Omega = R^n$) [2,3], and thus it has many important applications in the real world [1,4,9,18,23]. Until now, a variety of methods for solving $VI(\Omega, F)$ have been proposed and investigated [5–8,12–14,17,20]. Among them, extra-gradient method and its modified versions [13,15,16] are direct methods for variational inequalities $VI(\Omega, F)$ that only need to use the value of function $F$ in the iterative processes. In order to easily understand that, we first briefly describe the extra-gradient method and the general extra-gradient method below.

Let $\beta_0 > 0$ and $u^k$ be the kth approximate solution of $VI(\Omega, F)$, then the extra-gradient method generates $u^{k+1}$ via the following projection-type prediction-correction process [13]:

$$\text{Prediction:} \quad \bar{u} = P_\Omega[u^k - \beta_k F(u^k)], \tag{1.2}$$

$$\text{Correction:} \quad u^{k+1} = P_\Omega[u^k - \beta_k F(\bar{u})], \tag{1.3}$$

where $\beta_k > 0$ satisfies the following assumption

$$\beta_k \|F(u^k) - F(\bar{u})\| \leq \nu \|u^k - \bar{u}\|, \;\; \nu \in (0,\ 1). \tag{1.4}$$

Based on the prediction-correction process (1.2)–(1.3), a general extra-gradient method was proposed in paper He et al.[13] by just introducing a parameter $\alpha$ in the correction process of the extra-gradient method. Thus, the general extra-gradient method obtains the next $u^{k+1}$ by the following prediction-correction process:

$$\text{Prediction:} \quad \bar{u} = P_\Omega[u^k - \beta_k F(u^k)], \tag{1.5}$$

$$\text{Correction:} \quad u^{k+1} := u^{k+1}(\alpha) = P_\Omega[u^k - \alpha\beta_k F(\bar{u})], \tag{1.6}$$

where

$$\alpha \in \left(0,\ \frac{2e(u^k,\ \beta_k)^T d(u^k,\ \beta_k)}{\|d(u^k,\ \beta_k)\|^2}\right)$$

and

$$e(u^k,\ \beta_k) = u^k - P_\Omega[u^k - \beta_k F(u^k)] = u^k - \bar{u},$$
$$d(u^k,\ \beta_k) = u^k - \bar{u} - \beta_k (F(u^k) - F(\bar{u})).$$

**Remark 1.1** It is clear that if $\|d(u^k,\ \beta_k)\| = 0$, then $u^k$ produced by general extra-gradient method is a solution of (1.1). Thus, we can assume that $\|d(u^k,\ \beta_k)\| \neq 0$ throughout our paper.

As we know that direct methods are very useful for many practical variational inequalities $VI(\Omega, F)$ in which we can just observe the value of $F$ at a given variable and can not write down the explicit expression of the function $F$. However, such observation may be obtained via some costly experiments. Based on this view of point, reducing the times of observing the value of function $F$ in those methods is meaningful in practice.

In this paper, we are mainly concerned with the type of variational inequality $VI(\Omega,\ F)$ in which the cost of observing or computing the value of function $F$ is very expensive, and the projection of a vector on $\Omega$ is relatively easy to be computed. In this setting, the important task in improving the general extra-gradient method is to reduce the amount of observing or computing the value of function $F$ in solving this kind of variational inequality. We will see that this task can be achieved in general extra-gradient method by costing a relatively minor amount of computation used for obtaining projections of some vectors on $\Omega$. It is clear that the additional computation is worthy to cost in those practical problems.

In order to obtain the more efficient and practical algorithm for this kind of variational inequality, let

$$\Theta(\alpha) = \|u^k - u^*\|^2 - \|u^{k+1}(\alpha) - u^*\|^2, \tag{1.7}$$

$$\Phi(\alpha) = \|u^{k+1}(\alpha) - u^k\|^2 + 2\alpha\beta_k(u^{k+1}(\alpha) - \bar{u})^T F(\bar{u}), \tag{1.8}$$

$$\Psi(\alpha) = 2\alpha e(u^k,\ \beta_k)^T d(u^k,\ \beta_k) - \alpha^2\|d(u^k,\ \beta_k)\|^2, \tag{1.9}$$

where $u^* \in \Omega$ is a solution of problem (1.1).

In the next section, we study some properties of $\Phi(\alpha)$ and show that

$$\Theta(\alpha) \geq \Phi(\alpha) \geq \Psi(\alpha). \tag{1.10}$$

Following the inequalities (1.10), we will develop an improved general extra-gradient method for problem (1.1) and provide the convergence analysis of the new method. In Sect. 3, examples and the computational results are presented. Conclusions are presented in Sect. 4.

## 2 The improved general extra-gradient method

In this section, we first establish the inequalities (1.10) which is a little modification of the results of Theorem 2 in paper He et al.[13] and can be proved similarly. Following that, an improved algorithm for problem (1.1) is defined. Finally, we study the properties of function $\Phi(\alpha)$.

We first note that $\bar{u} = P_\Omega[u^k - \beta_k F(u^k)] \in \Omega$, it follows from (1.1) that

$$\beta_k F(u^*)^T(\bar{u} - u^*) \geq 0. \tag{2.1}$$

Under the assumption that $F$ is monotone, we have

$$(\beta_k F(\bar{u}) - \beta_k F(u^*))^T(\bar{u} - u^*) \geq 0. \tag{2.2}$$

Adding (2.1) and (2.2), we get

$$(u^k - u^*)^T F(\bar{u}) \geq (u^k - \bar{u})^T F(\bar{u}) \tag{2.3}$$

(2.3) is a basic inequality which will be applied in the following proposition.

**Proposition 2.1** *Let $\Theta(\alpha)$, $\Phi(\alpha)$ and $\Psi(\alpha)$ be defined by (1.7)–(1.9), respectively, and $F$ be monotone. We have*

$$\Theta(\alpha) \geq \Phi(\alpha) \geq \Psi(\alpha), \tag{2.4}$$

*where $\alpha \geq 0$.*

*Proof* Since

$$\|P_\Omega(v) - u\|^2 \leq \|v - u\|^2 - \|v - P_\Omega(v)\|^2, \quad \forall v \in R^n, \quad u \in \Omega,$$

we have

$$\|u^{k+1}(\alpha) - u^*\|^2 \leq \|u^k - \alpha\beta_k F(\bar{u}) - u^*\|^2 - \|u^k - \alpha\beta_k F(\bar{u}) - u^{k+1}(\alpha)\|^2.$$

It follows that

$$\Theta(\alpha) \geq \|u^k - u^*\|^2 - \|u^k - \alpha\beta_k F(\bar{u}) - u^*\|^2 + \|u^k - u^{k+1}(\alpha) - \alpha\beta_k F(\bar{u})\|^2.$$

By using (2.3) and a simple manipulation, we obtain

$$\Theta(\alpha) \geq \|u^k - u^{k+1}(\alpha)\|^2 + 2\alpha\beta_k(u^{k+1}(\alpha) - \bar{u})^T F(\bar{u}) = \Phi(\alpha). \tag{2.5}$$

Thus, we have proven the first part of the proposition.

Note that the $\Phi(\alpha)$ can be rewritten as

$$\Phi(\alpha) = \|u^k - u^{k+1}(\alpha)\|^2 + 2\alpha\beta_k e(u^k,\ \beta_k)^T F(\bar{u}) - 2\alpha\beta_k(u^k - u^{k+1}(\alpha))^T F(\bar{u}).$$

Using $\beta_k F(\bar{u}) = d(u^k,\ \beta_k) - [e(u^k,\ \beta_k) - \beta_k F(u^k)]$, it follows that

$$\begin{aligned}
\Phi(\alpha) &= 2\alpha\beta_k e(u^k,\ \beta_k)^T F(\bar{u}) + 2\alpha(u^k - u^{k+1}(\alpha))^T (e(u^k,\ \beta_k) - \beta_k F(u^k)) \\
&\quad - \alpha^2\|d(u^k,\ \beta_k)\|^2 + \|(u^k - u^{k+1}(\alpha)) - \alpha d(u^k,\ \beta_k)\|^2 \\
&= 2\alpha e(u^k,\ \beta_k)^T d(u^k,\ \beta_k) - \alpha^2\|d(u^k,\ \beta_k)\|^2 \\
&\quad + \|(u^k - u^{k+1}(\alpha)) - \alpha d(u^k,\ \beta_k)\|^2 \\
&\quad + 2\alpha(u^k - u^{k+1}(\alpha) - e(u^k,\ \beta_k))^T (e(u^k,\ \beta_k) - \beta_k F(u^k)) \\
&= \Psi(\alpha) + \|(u^k - u^{k+1}(\alpha)) - \alpha d(u^k,\ \beta_k)\|^2 \\
&\quad + 2\alpha(u^k - u^{k+1}(\alpha) - e(u^k,\ \beta_k))^T (e(u^k,\ \beta_k) - \beta_k F(u^k)).
\end{aligned} \tag{2.6}$$

Note that

$$u^k - u^{k+1}(\alpha) - e(u^k,\ \beta_k) = P_\Omega[u^k - \beta_k F(u^k)] - u^{k+1}(\alpha).$$

Setting $v := u^k - \beta_k F(u^k)$ and $u := u^{k+1}(\alpha)$ in the following basic inequality of projection mapping:

$$(v - P_\Omega(v))^T (P_\Omega(v) - u) \geq 0, \quad \forall v \in R^n, \quad \forall\, u \in \Omega \tag{2.7}$$

we get

$$(e(u^k,\ \beta_k) - \beta_k F(u^k))^T (P_\Omega[u^k - \beta_k F(u^k)] - u^{k+1}(\alpha)) \geq 0$$

and therefore have

$$(u^k - u^{k+1}(\alpha) - e(u^k,\ \beta_k))^T (e(u^k,\ \beta_k) - \beta_k F(u^k)) \geq 0. \tag{2.8}$$

Substituting (2.8) into (2.6), it follows that

$$\Phi(\alpha) \geq \Psi(\alpha). \tag{2.9}$$

Following (2.5) and (2.9), the proof is complete. □

Note that $\Psi(\alpha)$ is a quadratic function of $\alpha$, it attains its maximum at

$$\alpha_k^* = \frac{e(u^k,\ \beta_k)^T d(u^k,\ \beta_k)}{\|d(u^k,\ \beta_k)\|^2}. \tag{2.10}$$

Based on the $\Theta(\alpha) \geq \Psi(\alpha)$ (see [13], Theorems 2 and 3 with $\gamma = 1$), we have the following convergence results for the general extra-gradient method:

$$\Theta(\alpha_k^*) \geq \Psi(\alpha_k^*)$$

and

$$\Psi(\alpha_k^*) \geq \frac{(1-\nu)}{2}\|e(u^k,\ \beta_k)\|^2. \tag{2.11}$$

Similar to He et al. [13] by maximizing $\Psi(\alpha)$ to obtain a proper step size in each iteration of general extra-gradient method , we can improve the general extra-gradient method by replacing the step size $\alpha_k^*$ used in the original extra-gradient method with a refined step size computed by maximizing $\Phi(\alpha)$.

Now, for the same $k$th approximate solution $u^k$, let

$$\alpha_2^*(k) = \arg\max_{\alpha}\{\Psi(\alpha)|\alpha \geq 0\} \tag{2.12}$$

and

$$\alpha_1^*(k) = \arg\max_{\alpha}\{\Phi(\alpha)|\alpha \geq 0\}.$$

In order to make $\alpha_1^*(k)$ be obtained easily, we approximately compute $\alpha_1^*(k)$ by solving the following simple optimization problem.

$$\alpha_1^*(k) = \arg\max_{\alpha}\{\Phi(\alpha)|0 \leq \alpha \leq m_1\alpha_2^*(k)\}, \tag{2.13}$$

where $m_1 \geq 1$.

Based on the assumption of $\|d(u^k,\ \beta_k)\| \neq 0$ and (2.10), it is clear that $\alpha_1^*(k)$ and $\alpha_2^*(k)$ can be obtained by (2.13) and (2.12), respectively.

**Remark 2.2** It is worth mentioning that when the value of $F$ is not easy to be obtained or observed and the projection of a vector on $\Omega$ can be computed relatively easily, such as in some practical application problems, the main amount of computation of approximately finding the point $\alpha_1^*(k)$ (at which the maximum value of $\Phi(\alpha)$ on $[0,\ m_1\alpha_2^*(k)]$ is attained) is to obtain or observe the value of function $F$ at $\bar{u}$. Note that the value of $F$ at $\bar{u}$ will be needed again in the correction process of the general extra-gradient method. Moreover, for a given $\alpha$, we can obtain the associated value of $\Phi(\alpha)$ by mainly computing the projection of $u^k - \alpha\beta_k F(\bar{u})$ on $\Omega$. Thus, if we use $\alpha_1^*(k)$ instead of $\alpha_2^*(k)$ at each prediction-correction process of the general extra-gradient method, we need not observe or compute additional value of function $F$, just cost a relatively minor amount of computation for obtaining projections of some vectors on $\Omega$ during computing the approximate value of $\alpha_1^*(k)$.

Based on the proposition 2.1, the definition of $\alpha_1^*(k)$ and $\alpha_2^*(k)$, and the inequality (2.11), the following convergence results can be proved easily.

**Proposition 2.2** *Let $\alpha_1^*(k)$ and $\alpha_2^*(k)$ be defined by* (2.13) *and* (2.12)*, respectively, F be monotone, then we have.*

(1) $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_1^*(k)) - u^*\|^2 \ge \Phi(\alpha_1^*(k))$,

(2) $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_2^*(k)) - u^*\|^2 \ge \Psi(\alpha_2^*(k))$,

(3) $\Phi(\alpha_1^*(k)) \ge \Psi(\alpha_2^*(k)) \ge \frac{(1-\nu)}{2}\|e(u^k,\ \beta_k)\|^2$.

**Remark 2.3** In a sense, $\alpha_2^*(k)$ is the original optimal step size used in the general extra-gradient method. According to proposition 2.2, $\alpha_1^*(k)$ defined by (2.13) can be taken as the more proper step size instead of $\alpha_2^*(k)$ and thus an improved general extra-gradient method may be obtained.

Following the proposition 2.2, we now show that the sequence $\{u^k\}$ obtained from (1.5) and (1.6) with $\alpha = \alpha_1^*(k)$ converges to a solution of the variational inequality (1.1). For this purpose, we need the following results, which can be found in Peng and Fukushima [19].

**Lemma 2.1** *For all $u \in R^n$ and $\bar{\beta} \ge \beta > 0$, it holds that*

$$\|e(u, \bar{\beta})\| \ge \|e(u, \beta)\|$$

*and*

$$\frac{\|e(u, \bar{\beta})\|}{\bar{\beta}} \le \frac{\|e(u, \beta)\|}{\beta}.$$

By using the technique of He [10], we have

**Theorem 2.1** *Let the sequence $\{u^k\}$ be generated by* (1.5) *and* (1.6) *with $\alpha = \alpha_1^*(k)$, then $\{u^k\}$ converges to a solution of* (1.1).

*Proof* Because the sequence $\{u^k\}$ generated by (1.5) and (1.6) with the conclusions of Proposition 2.2 is bounded and the mapping $F$ is continuous, it is possible to prove that while $\|e(u^k,\ \beta_k)\| \ge \varepsilon > 0$, there is a $\beta_{\min} > 0$ such that, for all $k$,

$$\beta_k \ge \beta_{\min}$$

and the inequality (1.4) holds [11].

Now, let $\hat{u}$ be a solution of (1.1). From Proposition 2.2, we get

$$\|u^{k+1} - \hat{u}\|^2 \le \|u^k - \hat{u}\|^2 - c_0\|e(u^k, \beta_k)\|^2, \tag{2.14}$$

where $c_0 = (1-\nu)/2$, and thus we have that the sequence $\{u^k\}$ is bounded and

$$\sum_{k=0}^{\infty} c_0\|e(u^k,\ \beta_k)\|^2 \le \|u^0 - \hat{u}\|^2$$

and it follows from Lemma 2.1 that

$$\lim_{k\to\infty} e(u^k,\ \beta_{\min}) = 0.$$

Further, since the sequence $\{u^k\}$ is bounded, let $\tilde{u}^*$ be a cluster point of $\{u^k\}$ and the subsequence $\{u^{k_j}\}$ converges to $\tilde{u}^*$. Because $e(u,\ \beta_{\min})$ is continuous, we have

$$e(\tilde{u}^*,\ \beta_{\min}) = \lim_{j\to\infty} e(u^{k_j},\ \beta_{\min}) = 0$$

and thus $\tilde{u}^*$ is a solution of (1.1). In the following, we prove that the sequence $\{u^k\}$ has exactly one cluster point. Assume that $\tilde{u}$ is another cluster point and denote

$$\delta = \|\tilde{u} - \tilde{u}^*\| > 0.$$

Because $\tilde{u}^*$ is a cluster point of sequence $\{u^k\}$, there is a $k_0 > 0$ such that

$$\|u^{k_0} - \tilde{u}^*\| \le \delta/2.$$

On other hand, since $\tilde{u}^*$ is a solution of (1.1), it follows from Proposition 2.2 that

$$\|u^k - \tilde{u}^*\| \le \|u^{k_0} - \tilde{u}^*\|, \quad \forall\, k\ \ge k_0.$$

And thus, we have

$$\|u^k - \tilde{u}\| \ge \|\tilde{u} - \tilde{u}^*\| - \|u^k - \tilde{u}^*\| \ge \delta/2, \quad \forall\, k\ \ge k_0.$$

This contradicts the assumption, thus the sequence $u^k$ converges to a solution $\tilde{u}^*$ of (1.1). □

Before we present the improved extra-gradient method, we first briefly describe the general extra-gradient method [13] below.

**Algorithm 1** The general extra-gradient method

**Step 1** Initialization:
pLet $\beta_0 > 0,\ \varepsilon > 0,\ 0 < \mu < \nu < 1, 0 < \gamma < 2,\ u^0 \in \Omega$ and set $k := 0$.

**Step 2** Prediction:
$\bar{u} := P_\Omega[u^k - \beta_k F(u^k)]$.

**Step 3** Verifying convergence:
Let $e(u^k,\ \beta_k) = u^k - \bar{u}$. If $\|e(u^k,\ \beta_k)\| < \varepsilon$ then stop, else go to Step 4.

**Step 4** Modifying $\beta_k$ and computing $\beta_{k+1}$:
If $r_k := \beta_k\|F(u^k) - F(\bar{u})\|/\|u^k - \bar{u}\| \le \nu$, then
  $\beta_{k+1} = \beta_k$.
 Else
  While $r_k > \nu$, do
    $\beta_k := \frac{3}{4}\beta_k \times \min\{1,\ \nu/r_k\}$,
    $\bar{u} := P_\Omega[u^k - \beta_k u^k]$,
    $r_k := \beta_k\|F(u^k) - F(\bar{u})\|/\|u^k - \bar{u}\|$.
  End While
  $\beta_{k+1} = \beta_k$.
 End If

**Step 5** Searching step size $\alpha_k^*$ :
Solve the following optimization problem
$\alpha_k^* = \arg\max_{\alpha}\{\Psi(\alpha) \ | \ \alpha \geq 0\}$,
where $\Psi(\alpha)$ is defined by (1.9).

**Step 6** Extending the step size:
$\alpha_k = \gamma\alpha_k^*$.

**Step 7** Correction:
$u^{k+1} = P_\Omega[u^k - \alpha_k\beta_k F(\bar{u})]$.

**Step 8** Adjusting $\beta_{k+1}$:
If $r_k \leq \mu$
$\beta_{k+1} := \beta_{k+1}\nu/r_k$.
End If
$k := k + 1$, go to Step 2. □

Based on Propositions 2.1 and 2.2, we are now ready to present the improved general extra-gradient algorithms for (1.1). As we have known that the general extra-gradient method [13] was introduced by choosing a proper step size $\alpha$ in each iteration of extra-gradient method. And the proper step size $\alpha$ can be obtained by initially obtaining $\alpha_k^*$ by maximizing $\Psi(\alpha)$ and then extending $\alpha_k^*$ according to the fact that $\Psi(\alpha)$ is a quadratic function with respect to $\alpha$. This is achieved in the Steps 5 and 6 of the Algorithm 1, respectively. Similarly, the Proposition 2.2 motivates us that we can improve the general extra-gradient method by choosing a more proper step size $\alpha$ based on finding $\alpha_1^*(k)$ instead of $\alpha_2^*(k)$ in the Step 5 of Algorithm 1 and correspondingly extending the step size by solving the following subproblem:

$$\alpha_k = \max_{\alpha}\{\alpha_1^*(k) \leq \alpha \leq m_2\alpha_1^*(k) \mid \Phi(\alpha) \geq \rho\Phi(\alpha_1^*(k))\}, \tag{2.15}$$

where $\rho \in (0,\ 1)$ but closes to 0 and $m_2 \geq 2$.
It is clear that the $\alpha_k$ can be obtained. And the main work of solving the problem (2.15), as mentioned in Remark 2.2, is to obtain the projections of some vectors during computing the value of $\Phi(\alpha)$, and this is a relatively minor amount of computation compared with the cost of observing the value of function $F$ under the assumptions above.

Following the studies above, we now briefly describe the new algorithm below and call the new method the improved general extra-gradient method.

**Algorithm 2** The improved general extra-gradient method

**Step 1** Initialization:
Let $\beta_0 > 0,\ \varepsilon > 0,\ 0 < \mu < \nu < 1, 0 < \rho < 1, m_1 \geq 1, m_2 \geq 2,\ u^0 \in \Omega$ and set $k := 0$.

**Step 2** Prediction:
$\bar{u} := P_\Omega[u^k - \beta_k F(u^k)]$.

**Step 3** Verifying convergence:
Let $e(u^k,\ \beta_k) = u^k - \bar{u}$. If $\|e(u^k,\ \beta_k)\| < \varepsilon$ then stop, else go to Step 4.

**Step 4** Modifying $\beta_k$ and computing $\beta_{k+1}$:
If $r_k := \beta_k \|F(u^k) - F(\bar{u})\| / \|u^k - \bar{u}\| \le \nu$, then
$\beta_{k+1} = \beta_k$.
Else
While $r_k > \nu$, do
$\beta_k := \frac{3}{4}\beta_k \times \min\{1,\ \nu/r_k\}$,
$\bar{u} := P_\Omega[u^k - \beta_k u^k]$,
$r_k := \beta_k \|F(u^k) - F(\bar{u})\| / \|u^k - \bar{u}\|$.
End While
$\beta_{k+1} = \beta_k$.
End If

**Step 5** Searching step size $\alpha_k^*$ :
Let $\bar{\alpha}_k = \arg\max_\alpha\{\Psi(\alpha)\ \mid\ \alpha \ge 0\}$,
where $\Psi(\alpha)$ is defined by (1.9).
Solve the following optimization problem
$\alpha_k^* = \arg\max_\alpha\{\Phi(\alpha)\ \mid\ 0 \le \alpha \le m_1\bar{\alpha}_k\}$,
where $\Phi(\alpha)$ is defined by (1.8).

**Step 6** Extending the step size:
$\alpha_k = \max_\alpha\{\alpha_k^* \le \alpha \le m_2\alpha_k^* \mid \Phi(\alpha) \ge \rho\Phi(\alpha_k^*)\}$.

**Step 7** Correction:
$u^{k+1} = P_\Omega[u^k - \alpha_k\beta_k F(\bar{u})]$.

**Step 8** Adjusting $\beta_{k+1}$:
If $r_k \le \mu$
$\beta_{k+1} := \beta_{k+1}\nu/r_k$.
End If
$k := k+1$, go to Step 2. □

Further, we can obtain the following analytic properties of function $\Phi(\alpha)$.

**Proposition 2.3** *Assume that $\Phi(\alpha)$ is defined by (1.8), $F$ is monotone and continuously differentiable, then we have*

(1) $\Phi'(\alpha) = 2\beta_k(u^{k+1}(\alpha) - \bar{u})^T F(\bar{u})$,

(2) *$\Phi'(\alpha)$ is a decreasing function with respect to $\alpha \geq 0$, i.e., when $\alpha \geq 0$, $\Phi(\alpha)$ is concave.*
*Furthermore, if $\Phi'(\alpha_1^*(k)) = 0$, we have*

(3) $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_1^*(k)) - u^*\|^2 \geq \|u^k - u^{k+1}(\alpha_1^*(k))\|^2.$

*Proof* For given $\beta_k,\ u^k,\ \bar{u},\ F(\bar{u})$, let

$$h(\alpha,\ y) = \|y - [u^k - \alpha\beta_k F(\bar{u})]\|^2 - \alpha^2\beta_k^2\|F(\bar{u})\|^2 - 2\alpha\beta_k(\bar{u} - u^k)^T F(\bar{u}). \quad (2.16)$$

It is easy to see that the solution of the following problem

$$\min_y\{h(\alpha,\ y)|y \in \Omega\}$$

is $y^* = P_\Omega[u^k - \alpha\beta_k F(\bar{u})]$. Substituting $y^*$ into (2.16) and simplifying it, we have

$$\Phi(\alpha) = h(\alpha,\ y)|_{y=P_\Omega[u^k-\alpha\beta_k F(\bar{u})]},$$

i.e.,

$$\Phi(\alpha) = \min_y\{h(\alpha,\ y)|y \in \Omega\}. \quad (2.17)$$

Due to the identity (2.17), it follows from Auslender [1] (Chapter 4, Theorem 1.7) that $\Phi(\alpha)$ is differentiable and its derivative is given by

$$\begin{aligned}\Phi'(\alpha) &= \frac{\partial h(\alpha,\ y)}{\partial\alpha}|_{y=P_\Omega[u^k-\alpha\beta_k F(\bar{u})]}\\ &= 2\beta_k(u^{k+1}(\alpha) - u^k + \alpha\beta_k F(\bar{u}))^T F(\bar{u}) - 2\alpha\beta_k^2\|F(\bar{u})\|^2 - 2\beta_k(\bar{u} - u^k)^T F(\bar{u})\\ &= 2\beta_k(u^{k+1}(\alpha) - \bar{u})^T F(\bar{u}).\end{aligned}$$

Thus, the first conclusion is proved. We now establish the proof of the second assertion. Let $\bar{\alpha} > \alpha \geq 0$, we will prove that

$$\Phi'(\bar{\alpha}) \leq \Phi'(\alpha),$$

i.e.,

$$(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha))^T F(\bar{u}) \leq 0. \quad (2.18)$$

Setting $v := u^k - \bar{\alpha}\beta_k F(\bar{u})$, $u := u^{k+1}(\alpha)$ and $v := u^k - \alpha\beta_k F(\bar{u})$, $u := u^{k+1}(\bar{\alpha})$ in the basic inequality (2.7) of projection mapping, respectively, we have

$$(u^k - \bar{\alpha}\beta_k F(\bar{u}) - u^{k+1}(\bar{\alpha}))^T(u^{k+1}(\alpha) - u^{k+1}(\bar{\alpha})) \leq 0, \quad (2.19)$$

$$(u^k - \alpha\beta_k F(\bar{u}) - u^{k+1}(\alpha))^T(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)) \leq 0. \quad (2.20)$$

Adding (2.19) and (2.20), we obtain

$$(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha))^T\{(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)) + (\bar{\alpha} - \alpha)\beta_k F(\bar{u})\} \leq 0,$$

i.e.,

$$\|u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)\|^2 + (\bar{\alpha} - \alpha)\beta_k(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha))^T F(\bar{u}) \leq 0.$$

It follows that

$$(u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha))^T F(\bar{u}) \leq \frac{-1}{\beta_k(\bar{\alpha} - \alpha)}\|u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)\|^2 \leq 0.$$

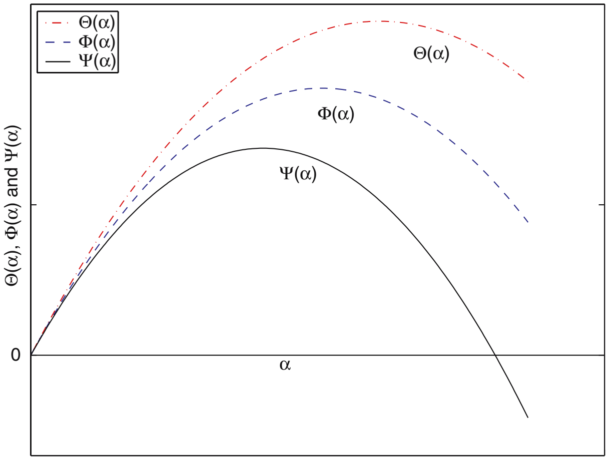


**Fig. 1** Sketch map of the relationship among the three functions: $\Theta(\alpha)$, $\Phi(\alpha)$ and $\Psi(\alpha)$

Thus, we obtain the inequality (2.18).
Finally, the third part of the proposition is easy to see from the Proposition 2.2 and the proof is completed. □

Figure 1 intuitively shows the results described in the Propositions 2.1 and 2.3.

## 3 Numerical experiments

In this section, we present some numerical experiments in the aim of comparing the improved general extra-gradient method with the original general extra-gradient method and a variant of the extra-gradient method recently proposed by Wang et al. [22]. All programs are coded in MATLAB and the programs are run on a IBM notebook PC R51.

**Example 1** In the first test example, we mainly compare the improved general extra-gradient method with the original general extra-gradient method. We form our test problem (1.1) by taking

$$F(u) = D(u) + Mu + q,$$

where vector $D(u)$ and $Mu + q$ are the nonlinear part and the linear part of $F(u)$, respectively. We construct the linear part $Mu + q$ similarly as in Harker and Pang [8] and He et al.[13]. The matrix $M = A^T A + B$, where $A$ is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, +5)$ and a skew-symmetric matrix $B$ is generated in the same way. The vector $q$ is generated from a uniform distribution in the interval $(-500, 500)$ or $(-500, 0)$. In vector $D(u)$, the nonlinear part of $F(u)$, the components are $D_j(u) = a_j \times \arctan(u_j)$ and $a_j$ is a random variable in $(0, 1)$. Now, we solve this problem by improved general extra-gradient method and general extra-gradient method. Both methods start with $\beta_0 = 1$, $\mu = 0.3$, $\nu = 0.9$, $\varepsilon = 10^{-7}$ and with the same initial vector generated randomly in the interval $(0, 1)$. And let $\gamma = 1.8$

**Table 1** Numerical results of example 1 with $q \in (-500,\ 500)$

| Problem size | Algorithm 1 | | Algorithm 2 | |
|---|---|---|---|---|
| | No. of Iter. | CPU time(s) | No. of Iter. | CPU time(s) |
| 100 | 290 | 0.13 | 229 | 0.22 |
| 200 | 403 | 0.190 | 339 | 0.231 |
| 300 | 409 | 0.291 | 332 | 0.32 |
| 500 | 448 | 1.031 | 363 | 1.182 |
| 600 | 391 | 2.143 | 294 | 1.882 |
| 700 | 417 | 2.714 | 334 | 2.463 |
| 800 | 355 | 3.585 | 286 | 3.265 |
| 1,000 | 378 | 5.838 | 320 | 4.967 |
| 1,100 | 438 | 6.88 | 351 | 6.038 |

**Table 2** Numerical results of example 1 with $q \in (-500,\ 0)$

| Problem size | Algorithm 1 | | Algorithm 2 | |
|---|---|---|---|---|
| | No. of Iter. | CPU time(s) | No. of Iter. | CPU time(s) |
| 100 | 534 | 0.19 | 461 | 0.24 |
| 200 | 753 | 0.401 | 639 | 0.45 |
| 300 | 825 | 0.44 | 708 | 0.681 |
| 500 | 1,037 | 2.414 | 897 | 3.064 |
| 600 | 1,000 | 5.558 | 859 | 5.428 |
| 700 | 968 | 6.399 | 832 | 6.159 |
| 800 | 865 | 8.743 | 745 | 8.141 |
| 1,000 | 1,009 | 15.512 | 889 | 14.371 |
| 1,100 | 1,244 | 19.588 | 1,056 | 17.866 |

in the general extra-gradient method, $\rho = 0.05,\ m_1 = 3,\ m_2 = 4$ in the improved general extra-gradient method, respectively. The stopping test is $e(u^k,\ 1) \le \varepsilon$.

Tables 1 and 2 report the iteration numbers and CPU time for both methods. Numerical results show that the improved general extra-gradient method can save about 12–25 % of the number of iterations. This means that the amount of computing the value of function $F$ in solving $VI(\Omega,\ F)$ can be saved about 12–25 % by the new general extra-gradient method. Saving the amount of computing the value of function $F$ is very important for some practical problems in which to obtain the value of function $F$ is not easy, and thus is the main purpose of our algorithm. From Tables 1 and 2, we also see that while the problem size $n \ge 600$, the running CPU time can be saved by the algorithm 2 comparing with Algorithm 1.

**Example 2** In this test example, we mainly compare the improved general extra-gradient method with a variant of the extra-gradient method recently proposed by Wang et al. [22]. The test problem (1.1) was considered in Sun[21], and Wang et al. [22] where $F(u) = Mu + q$, $M$ is a nonsymmetric matrix of the form

$$\begin{bmatrix} 4 & -2 & & & & \\ 1 & 4 & -2 & & & \\ & 1 & \ddots & \ddots & & \\ & & \ddots & \ddots & -2 & \\ & & & 1 & 4 & -2 \\ & & & & 1 & 4 \end{bmatrix},$$

**Table 3** Numerical results of example 2

| Problem | Number of Iterations | |
|---|---|---|
| size | Algorithm NVE | Algorithm 2 |
| 10 | 13 | 9 |
| 50 | 13 | 10 |
| 100 | 13 | 10 |
| 200 | 13 | 10 |
| 500 | 13 | 10 |

**Table 4** Numerical results of example 3

| Problem | Number of Iterations | |
|---|---|---|
| size | Algorithm NVE | Algorithm 2 |
| 10 | 12 | 8 |
| 20 | 12 | 8 |
| 50 | 12 | 8 |
| 100 | 11 | 8 |

where $q = [-1, -1, \cdots, -1]^T$ is a vector. The best numerical results of the algorithm nVE in Wang et al.[22] is given in Table 3.

Now, we solve this problem by improved general extra-gradient method. As in paper Wang et al. [22], we start our algorithm with the initial vector $u^0 = [0, 0, \ldots, 0]^T$, and take $\|e(u^k, 1)\|^2 \le n10^{-14}$ as the termination criterion, where $n$ is the dimension of the problem. And let $\beta_0 = 1,\ \mu = 0.5,\ \nu = 0.6,\ \rho = 0.05,\ m_1 = 3,\ m_2 = 4$. Table 3 reports the iteration numbers of the improved general extra-gradient method for this test problem.

**Example 3** This example was also considered in Sun [21] and Wang et al. [22], where

$$\begin{aligned}
&F(u) = F_1(u) + F_2(u),\\
&F_1(u) = [f_1(u), f_2(u), \ldots, f_n(u)]^T,\\
&F_2(u) = Mu + q,\\
&f_i(u) = u_{i-1}^2 + u_i^2 + u_{i-1}u_i + u_i u_{i+1}, \quad i = 1, 2, \ldots, n,\\
&u_0 = u_{n+1} = 0,
\end{aligned}$$

where $M$ and $q$ are the same as those in Example 2. With the same assumption as Example 2, Table 4 reports the iterations numbers of the improved general extra-gradient method and the algorithm NVE in Wang et al. [22], respectively.

Compared with a variant of extra-gradient algorithm proposed in Wang et al.[22], Tables 3 and 4 show that the improved general extra-gradient method also has good behavior.

## 4 Conclusions

In this paper, a new strategy based on Proposition 2.1–2.2 for computing step size in general extra-gradient method for nonlinear monotone variational inequalities $VI(\Omega,\ F)$ is introduced. In order to obtain the new step size, the new strategy just needs to additionally compute the projections of some vectors on $\Omega$ and doesn't need to compute additionally the value of function $F$. This is very important especially in some practical problems in which the cost of computing or observing the value of function $F$ is very expensive and the work of obtaining the projection of a vector on $\Omega$ is relatively easy. Furthermore, numerical experiments show that the amount of computing the value of function $F$ in solving $VI(\Omega, F)$ can be saved about 12–25% by the improved general extra-gradient method, thus the new method is more competitive than the original general extra-gradient method in solving those practical problems.

**Acknowledgments** We thank the anonymous referees for their kindly valuable suggestions. This research is financially supported by a research grant from the Research Grants Council of the People's Republic of China (Project No.10571083) and MOEC grant 20060284001. X.M. Yuan was supported by an internal grant of Antai College of Economics and Management, Shanghai Jiao Tong University.

## References

1. Auslender, A.: Optimisation Méthodes Numériques. Masson, Paris (1976)
2. Fancchinei, F., Pang, J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems. vol. 1. Springer, Berlin, Heidelberg New York (2003)
3. Fancchinei, F., Pang, J.S.: Finite-dimensional Variational Inequalities and Complementarity Problems. vol. 2. Springer, Berlin, Heidelberg New York (2003)
4. Ferris, M.C., Pang, J.S.: Engineering and economic applications of complementarity problems. SIAM Rev. **39**, 669–713 (1997)
5. Fletcher, R.: Practical Methods of Optimization. Wiley, New York (1985)
6. Fukushima, M.: The primal Douglas-Rachford splitting algorithm for a class of monotone mappings with application to the traffic equilibrium problem. Math. Program. **72**, 1–15 (1996)
7. Goldstein, A. A.: Convex programming in Hilbert space. Bull. Am. Math. Soc **70**, 709–710 (1964)
8. Harker, P.T., Pang, J.S.: A damped newton method for the linear complementarity problem. Lect. Appl. Math. **26**, 265–284 (1990)
9. Harker, P.T., Pang, J.S.: Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory algorithms and applications. Math. Program. **48**, 161–220 (1990)
10. He, B.S.: A new method for a class of linear variational inequalities. Math. Program. **66**, 137–144 (1994)
11. He, B.S.: A class of projection and contraction methods for monotone variational inequalities. Appl. Math. Optimization **35**, 69–76 (1997)
12. He, B.S., Yang, H., Zhang, C.S.: A modified augmented Lagrangian method for a class of monotone variational inequalities. Eur. J. Oper. Res. **159**, 35–51 (2004)
13. He, B.S., Yuan, X.M., Zhang, J.J.Z.: Comparison of two kinds of prediction-correction methods for monotone variational inequalities. Comput. Optimization Appl. **27**, 247–267 (2004)
14. Kanzow, C., Fukushima, M.: Solving box constrained variational inequalities by using the natural residual with D-gap function globalization. Oper. Res. Lett. **23**, 45–51 (1998)
15. Khobotov, E.N.: Modification of the extragradient method for solving variational inequalities and certain optimization problems. USSR. Comput. Math. Phys. **27**, 120–127 (1987)
16. Korpelevich, G.M.: The extragradient method for finding saddle points and other problems. Ekon. Matematchskie Metody **12**, 747–56 (1976)
17. Levitin, E.S., Polyak, B.T.: Constrained minimization problems. USSR Comput. Math. Math. Phys. **6**, 1–50 (1966)
18. Nagurney, A.: Network economics: A Variational Inequality Approach. Revised 2nd edn. Kluwer Academic Publishers Dordrecht, (1999)
19. Peng, J.M., Fukushima, M.: A hybrid newton method for solving the variational inequality problem via the D-gap function. Math. Program. **86**, 367–386 (1999)
20. Qi, L., Sun, D.: Smoothing functions and a smoothing newton method for complementarity and variational inequality problems. J. Optimization Theory appl. **113**, 121–148 (2002)

21. Sun, D.: A projection and constraction method for the nonlinear complementarity problem and its extensions. Math. Num. Sin. **16**, 183–194 (1994)
22. Wang, Y.J., Xiu, N.H., Wang, C.Y.: A new version of extragradient method for variational inequality problems. Compu. Math. Appl. **42**, 969–979 (2001)
23. Zangwill, W.I.: Nonlinear programming: A united approach. Prentice Hall, Englewood Cliffs, NJ (1969)