Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

# Comparison of two approximal proximal point algorithms for monotone variational inequalities*

TAO Min
(*School of Applied Mathematics and Physics, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*)
E-mail: mathsgirl2002@yahoo.com.cn

Received Sept. 8, 2005; revision accepted Dec. 21, 2005

**Abstract:** Proximal point algorithms (PPA) are attractive methods for solving monotone variational inequalities (MVI). Since solving the sub-problem exactly in each iteration is costly or sometimes impossible, various approximate versions of PPA (APPA) are developed for practical applications. In this paper, we compare two APPA methods, both of which can be viewed as prediction-correction methods. The only difference is that they use different search directions in the correction-step. By extending the general forward-backward splitting methods, we obtain Algorithm I; in the same way, Algorithm II is proposed by spreading the general extra-gradient methods. Our analysis explains theoretically why Algorithm II usually outperforms Algorithm I.

For computation practice, we consider a class of MVI with a special structure, and choose the extending Algorithm II to implement, which is inspired by the idea of Gauss-Seidel iteration method making full use of information about the latest iteration. And in particular, self-adaptive techniques are adopted to adjust relevant parameters for faster convergence. Finally, some numerical experiments are reported on the separated MVI. Numerical results showed that the extending Algorithm II is feasible and easy to implement with relatively low computation load.

**Key words:** Projection and contraction methods, Proximal point algorithm (PPA), Approximate PPA (APPA), Monotone variational inequality (MVI), Prediction and correction
**doi:**10.1631/jzus.2007.A0969 **Document code:** A **CLC number:** O221.2

## INTRODUCTION

Generally a variational inequalities problem has the following mathematical form: Let $\Omega\subset\mathbb{R}^n$ be a nonempty closed convex set and $F$ be a continuous mapping from $\mathbb{R}^n$ into itself. Find $\boldsymbol{x}^*\in\Omega$, such that

$$\text{VI}(\Omega, F)\quad (\boldsymbol{x}-\boldsymbol{x}^*)^{\text{T}}F(\boldsymbol{x}^*)\geq 0,\ \ \forall \boldsymbol{x}\in\Omega. \tag{1}$$

The so-called monotone variational inequalities (MVI) problem means that the operator $F$ is monotone, i.e.,

$$(\boldsymbol{u}-\boldsymbol{v})^{\text{T}}[F(\boldsymbol{u})-F(\boldsymbol{v})]\geq 0,\ \ \forall \boldsymbol{u}, \boldsymbol{v}\in\Omega. \tag{2}$$

VI($\Omega$,$F$) problems include nonlinear complementarity problems (when $\Omega=\mathbb{R}^n_+$) and system of nonlinear equations (when $\Omega=\mathbb{R}^n$), and thus have many important applications (Harker and Pang, 1990). For any $\beta$>0, it is well known (Bertsekas and Tsitsiklis, 1989) that

$$\boldsymbol{u}^* \text{ is a solution of VI}(\Omega,F)\Leftrightarrow \boldsymbol{u}^*=P_\Omega[\boldsymbol{u}^*-\beta F(\boldsymbol{u}^*)], \tag{3}$$

where $P_\Omega(\cdot)$ denotes the projection on $\Omega$. Denote

$$e(\boldsymbol{u},\beta):=\boldsymbol{u}-P_\Omega[\boldsymbol{u}-\beta F(\boldsymbol{u})]. \tag{4}$$

A classical method for solving MVI is the proximal point algorithm (PPA) (Rockafellar, 1976). For given $\boldsymbol{u}^k\in\Omega$ and $\beta_k$>0, the new iteration $\boldsymbol{u}^{k+1}$ of the exact version of PPA is

$$\boldsymbol{u}^{k+1}=\boldsymbol{u}_*^{k+1},$$

* Project (No. 1027054) supported by the National Natural Science Foundation of China

where $u_*^{k+1}$ is the exact solution of the following variational inequality:

$$\text{(PPA)}\quad \boldsymbol{u}\in\Omega,\ (\boldsymbol{u}'-\boldsymbol{u})^{\mathrm{T}}F_k(\boldsymbol{u})\geq 0,\ \forall \boldsymbol{u}'\in\Omega, \tag{5}$$

with

$$F_k(\boldsymbol{u})=(\boldsymbol{u}-\boldsymbol{u}^k)+\beta_k F(\boldsymbol{u}). \tag{6}$$

An equivalent recursion form of the exact PPA is

$$\boldsymbol{u}^{k+1}=P_\Omega[\boldsymbol{u}^{k+1}-F_k(\boldsymbol{u}^{k+1})]. \tag{7}$$

The ideal form Eq.(7) of the method is often impractical since in many cases solving problem exactly is either impossible or expensive. In order to overcome such obstacles, some APPAs have been proposed in (He *et al*., 2004). In this paper, we compare two kinds of APPAs. Given $\boldsymbol{u}^k\in\Omega$ and $\beta_k>0$, let $\boldsymbol{v}^k\in\Omega$ be an approximate solution in the sense that

$$\boldsymbol{v}^k\approx P_\Omega[\boldsymbol{v}^k-F_k(\boldsymbol{v}^k)], \tag{8}$$

and define

$$\tilde{\boldsymbol{v}}^k:=P_\Omega[\boldsymbol{v}^k-F_k(\boldsymbol{v}^k)]. \tag{9}$$

We denote

$$\boldsymbol{\xi}^k:=\beta_k[F(\tilde{\boldsymbol{v}}^k)-F(\boldsymbol{v}^k)], \tag{10}$$

and define

$$\boldsymbol{d}^k=\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k+\boldsymbol{\xi}^k. \tag{11}$$

The general update forms of Algorithms I and II are

$$u^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k):=u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)=P_\Omega[\boldsymbol{u}^k-\alpha\boldsymbol{d}^k], \tag{12}$$

and

$$u^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k):=u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)=P_\Omega[\boldsymbol{u}^k-\alpha\beta_k F(\tilde{\boldsymbol{v}}^k)], \tag{13}$$

respectively. We say that such update forms are indirectly based on $\tilde{\boldsymbol{v}}^k$ and consider the following inexactness criterion:

$$\|(\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{\xi}^k\|\leq\nu\|\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k\|^2 \text{ and } \|\boldsymbol{\xi}^k\|\leq\mu\|\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k\|, \tag{14}$$
$$0<\nu<1\leq\mu.$$

Our interest in this work, however, is only to compare the efficiencies of Algorithm I and Algorithm II. For any solution point $\boldsymbol{u}^*\in\Omega^*$ [$\Omega^*$ denotes the solution set of VI($\Omega$, $F$), which is nonempty]. Let

$$\theta_{\mathrm{I}}:=\|\boldsymbol{u}^k-\boldsymbol{u}^*\|^2-\|u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-\boldsymbol{u}^*\|^2,$$

and

$$\theta_{\mathrm{II}}:=\|\boldsymbol{u}^k-\boldsymbol{u}^*\|^2-\|u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-\boldsymbol{u}^*\|^2.$$

We will prove that for two suitably introduced amounts $\Phi_{\mathrm{I}}$ and $\Phi_{\mathrm{II}}$,

$$\theta_{\mathrm{I}}\geq\Phi_{\mathrm{I}}:=\Phi+\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|, \tag{15}$$

$$\theta_{\mathrm{II}}\geq\Phi_{\mathrm{II}}:=\Phi+\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2, \tag{16}$$

and

$$\Phi_{\mathrm{II}}\geq\Phi_{\mathrm{I}}+\|u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2, \tag{17}$$

where

$$\Phi=2\alpha(\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{d}^k-\alpha^2\|\boldsymbol{d}^k\|^2.$$

Moreover, it will be shown by an example that both Eqs.(15) and (16) are tight. The main result Eq.(17) indicates that Algorithm II is likely better than Algorithm I.

This paper is organized as follows. In Section 2, we summarize some basic concepts and the consequent results. Sections 3 and 4 analyze the convergence behaviours of Algorithm I and Algorithm II (including the extending Algorithms II), repectively. Based on the analysis in Sections 3 and 4, the main theoretical result is given in Section 5. In Section 6, we provide the implementation details of the extending Algorithm II and give some numerical experiments. Finally, we give the conclusions.

Throughout this paper, we assume that $\{\beta_k\}\subset[\beta,+\infty)$ and $\beta>0$. The operator $F$ is monotone and continuous on $\Omega$, and the solutions set of VI($\Omega$,$F$), denoted by $\Omega^*$, is nonempty. We use $\boldsymbol{u}^*$ to denote any point in $\Omega^*$. In the case that $\Omega^*$ is not a singleton, we denote, for a given $\boldsymbol{u}$,

$$\|\boldsymbol{u}-\boldsymbol{u}^*\|:=\inf\{\|\boldsymbol{u}-\boldsymbol{u}^*\|\mid \boldsymbol{u}^*\in\Omega^*\}.$$

## PRELIMINARIES

In this section, we summarize some basic concepts and important preliminary results which will be used in the following analysis. Let $\Omega\subset\mathbb{R}^n$ be a nonempty closed convex set. Given $\boldsymbol{\omega}\subset\mathbb{R}^n$, the projection mapping under the Euclidean norm ($\|\cdot\|$), denoted by $P_\Omega(\boldsymbol{\omega})$, is defined as follows:

$$P_\Omega(\boldsymbol{\omega})=\arg\min\{\|\boldsymbol{\omega}-\boldsymbol{u}\| \mid \boldsymbol{u}\in\Omega\}.$$

Generally a complementarity problem (CP) has the following mathematical form:

$$\boldsymbol{x}\geq 0,\ F(\boldsymbol{x})\geq 0,\ \boldsymbol{x}^{\mathrm{T}}F(\boldsymbol{x})=0. \tag{18}$$

A basic property of the projection mapping on a closed convex set is

$$[\boldsymbol{v}-P_\Omega(\boldsymbol{v})]^{\mathrm{T}}[P_\Omega(\boldsymbol{v})-\boldsymbol{\omega}]\geq 0,\ \forall\boldsymbol{v}\in\mathbb{R}^n,\ \forall\boldsymbol{\omega}\in\Omega. \tag{19}$$

Consequently following Eq.(19), we have

$$\|P_\Omega(\boldsymbol{v})-P_\Omega(\boldsymbol{\omega})\|\leq\|\boldsymbol{v}-\boldsymbol{\omega}\|,\ \forall\boldsymbol{v},\boldsymbol{\omega}\in\mathbb{R}^n, \tag{20}$$

and

$$\|P_\Omega(\boldsymbol{v})-\boldsymbol{u}\|^2\leq\|\boldsymbol{v}-\boldsymbol{u}\|^2-\|\boldsymbol{v}-P_\Omega(\boldsymbol{v})\|^2,\quad \forall\boldsymbol{v}\in\mathbb{R}^n,\ \forall\boldsymbol{u}\in\Omega. \tag{21}$$

Note that $e(\boldsymbol{u},\beta)$ is a continuous function of $u$ because the projection mapping is non-expansive.

The next lemma states that $\|e(\boldsymbol{u},\beta)\|$ is a non-decreasing function for $\beta>0$.

**Lemma 1** (Zhu and Yu, 2004) $\forall\boldsymbol{u}\subset\mathbb{R}^n$ and $\tilde{\beta}\geq\beta>0$,

$$\|e(\boldsymbol{u},\tilde{\beta})\|\geq\|e(\boldsymbol{u},\beta)\|, \tag{22}$$

and

$$\frac{\|e(\boldsymbol{u},\beta)\|}{\beta}\geq\frac{\|e(\boldsymbol{u},\tilde{\beta})\|}{\tilde{\beta}}. \tag{23}$$

**Lemma 2** Let $\beta>0$ and $\{\boldsymbol{u}^k\}$ be a bounded sequence and $\lim_{k\to\infty}e(\boldsymbol{u}^k,\beta)=0$, then $\{\boldsymbol{u}^k\}$ has a subsequence $\{\boldsymbol{u}^{k_j}\}$ which converges to some $\boldsymbol{u}^\infty$ which is a solution point of VI($\Omega$, $F$).

**Proof** First, the bounded suquence $\{\boldsymbol{u}^k\}$ has a subsequence $\{\boldsymbol{u}^{k_j}\}$ converging to a point, say $\boldsymbol{u}^\infty$, note that $e(\boldsymbol{u},\beta)$ is a continuous function of $\boldsymbol{u}$ because the projection mapping is non-expansive. So $\boldsymbol{u}^\infty$ is a solution point of VI($\Omega$, $F$). Using the notation of $\xi^k$ in Eq.(10), from Eqs.(6), (9) and (10) we have

$$\tilde{\boldsymbol{v}}^k=P_\Omega[\boldsymbol{u}^k-\beta_kF(\tilde{\boldsymbol{v}}^k)+\xi^k]. \tag{24}$$

Since $\tilde{\boldsymbol{v}}^k\in\Omega$, it follows from Eq.(1) that

$$\text{(FI1)}\qquad \beta_kF(\boldsymbol{u}^*)^{\mathrm{T}}(\tilde{\boldsymbol{v}}^k-\boldsymbol{u}^*)\geq 0. \tag{25}$$

From Eqs.(19) and (24), we obtain

$$\text{(FI2)}\qquad \{[\boldsymbol{u}^k-\beta_kF(\tilde{\boldsymbol{v}}^k)+\xi^k]-\tilde{\boldsymbol{v}}^k\}^{\mathrm{T}}(\tilde{\boldsymbol{v}}^k-\boldsymbol{u}^*)\geq 0. \tag{26}$$

Under the monotonicity assumption we have

$$\text{(FI3)}\qquad [\beta_kF(\tilde{\boldsymbol{v}}^k)-\beta_kF(\boldsymbol{u}^*)]^{\mathrm{T}}(\tilde{\boldsymbol{v}}^k-\boldsymbol{u}^*)\geq 0. \tag{27}$$

The above three fundamental Eqs.(25)~(27) play a very important role in the convergence analysis of projection type methods (He *et al*., 2004). Adding the above three inequalities and using the notation of $\xi^k$ in Eq.(10), we get

$$(\tilde{\boldsymbol{v}}^k-\boldsymbol{u}^*)^{\mathrm{T}}(\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k+\xi^k)\geq 0. \tag{28}$$

Note that

$$\begin{aligned}&\|e(\tilde{\boldsymbol{v}}^k,\beta_k)\|\triangleq\|\tilde{\boldsymbol{v}}^k-P_\Omega[\tilde{\boldsymbol{v}}^k-\beta_kF(\tilde{\boldsymbol{v}}^k)]\|\\&=\|P_\Omega[\boldsymbol{u}^k-\beta_kF(\tilde{\boldsymbol{v}}^k)+\xi^k]-P_\Omega[\tilde{\boldsymbol{v}}^k-\beta_kF(\tilde{\boldsymbol{v}}^k)]\|\ \text{by Eq.(24)}\\&\leq\|\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k+\xi^k\|\leq(1+\mu)\|\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k\|.\end{aligned} \tag{29}$$

## PRELIMINARY ANALYSIS FOR ALGORITHM I

In this section we set up some preliminary results for Algorithm I.

**Theorem 1** Given $\boldsymbol{u}^k\in\Omega$ and $\beta_k>0$, let $\boldsymbol{v}^k\in\Omega$ be an approximate solution of Eq.(5) in the sense of Eq.(8) and the new iteration $u^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)$ be given by the general forms of Algorithms I(12). Then for any $\alpha>0$ we have

$$\theta_{\mathrm{I}}(\alpha,\tilde{\boldsymbol{v}}^k)\geq\Phi_{\mathrm{I}}(\alpha), \tag{30}$$

where

$$\theta_{\mathrm{I}}(\alpha,\tilde{\boldsymbol{v}}^k):=\|\boldsymbol{u}^k-\boldsymbol{u}^*\|^2-\|u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-\boldsymbol{u}^*\|^2, \tag{31}$$

$$\Phi_{\mathrm{I}}(\alpha):=\Phi(\alpha)+\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2, \tag{32}$$

$$\Phi(\alpha)=2\alpha(\boldsymbol{u}^k-\tilde{\boldsymbol{v}}^k)^{\mathrm{T}}\boldsymbol{d}^k-\alpha^2\|\boldsymbol{d}^k\|^2. \tag{33}$$

**Proof** Since $u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)=P_\Omega[\boldsymbol{u}^k-\alpha\boldsymbol{d}^k]$ and $\boldsymbol{u}^*\in\Omega$, it follows from Eq.(21) that

$$\begin{aligned}&\|u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-\boldsymbol{u}^*\|^2\\&\leq\|\boldsymbol{u}^k-\boldsymbol{u}^*-\alpha\boldsymbol{d}^k\|^2-\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2.\end{aligned} \tag{34}$$

and thus

$$\theta_{\rm I}(\alpha,\tilde{\boldsymbol v}^k)=\|\boldsymbol u^k-\boldsymbol u^*\|^2-\|u_{\rm I}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\boldsymbol u^*\|^2$$
$$\geq\|\boldsymbol u^k-\boldsymbol u^*\|^2-\|\boldsymbol u^k-\boldsymbol u^*-\alpha\boldsymbol d^k\|^2+\|\boldsymbol u^k-\alpha\boldsymbol d^k-u_{\rm I}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2$$
$$=2\alpha(\boldsymbol u^k-\boldsymbol u^*)^{\rm T}\boldsymbol d^k-\alpha^2\|\boldsymbol d^k\|^2+\|\boldsymbol u^k-\alpha\boldsymbol d^k-u_{\rm I}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2. \quad (35)$$

Comparing the right side of Eq.(35) and $\Phi(\alpha)$ in Eq.(33), the remainder is to prove

$$(\boldsymbol u^k-\boldsymbol u^*)^{\rm T}\boldsymbol d^k\geq(\boldsymbol u^k-\tilde{\boldsymbol v}^k)^{\rm T}\boldsymbol d^k. \quad (36)$$

Using the notation of $\boldsymbol d^k$ in Eq.(11), Eq.(28) can be written as $(\tilde{\boldsymbol v}^k-\boldsymbol u^*)^{\rm T}\boldsymbol d^k\geq0$, and consequently Eq.(30) holds. This completes the proof.

**Remark 1** Algorithm I can be viewed as an extension of General Forward-backward Splitting methods, proposed in (Tseng, 2000). Obviously, if we let $\boldsymbol u^k=\boldsymbol v^k$ in Algorithm I, we will obtain General Foward-Backward Splitting method.

## PRELIMINARY ANALYSIS FOR ALGORITHM II

In this section we set up some preliminary results for Algorithm II (including the extending Algorithm II).

**Theorem 2** Given $\boldsymbol u^k\in\Omega$ and $\beta_k>0$, let $\boldsymbol v^k\in\Omega$ be an approximate solution of Eq.(5) in the sense of Eq.(8) and $\Phi(\alpha)$ be given by Eq.(33); the new iteration $u^{k+1}(\alpha,\tilde{\boldsymbol v}^k)$ be given by the general forms of Algorithms II(13). Then for any $\alpha>0$ we have

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k)\geq\Phi_{\rm II}(\alpha), \quad (37)$$

where

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k):=\|\boldsymbol u^k-\boldsymbol u^*\|^2-\|u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\boldsymbol u^*\|^2, \quad (38)$$
$$\Phi_{\rm II}(\alpha):=\Phi(\alpha)+\|\boldsymbol u^k-\alpha\boldsymbol d^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2. \quad (39)$$

**Proof** Since $u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)=P_\Omega[\boldsymbol u^k-\alpha\beta_kF(\tilde{\boldsymbol v}^k)]$ and $\boldsymbol u^*\in\Omega$, it follows from Eq.(21) that

$$\|u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\boldsymbol u^*\|^2\leq\|\boldsymbol u^k-\boldsymbol u^*-\alpha\beta_kF(\tilde{\boldsymbol v}^k)\|^2$$
$$-\|\boldsymbol u^k-\alpha\beta_kF(\tilde{\boldsymbol v}^k)-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2, \quad (40)$$

and thus

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k)=\|\boldsymbol u^k-\boldsymbol u^*\|^2-\|u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\boldsymbol u^*\|^2$$
$$\geq\|\boldsymbol u^k-\boldsymbol u^*\|^2-\|\boldsymbol u^k-\boldsymbol u^*-\alpha\beta_kF(\tilde{\boldsymbol v}^k)\|^2$$
$$+\|\boldsymbol u^k-\alpha\beta_kF(\tilde{\boldsymbol v}^k)-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2$$
$$=\|\boldsymbol u^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2+2\alpha\beta_k[u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)$$
$$-\boldsymbol u^k]^{\rm T}F(\tilde{\boldsymbol v}^k)+2\alpha\beta_k(\boldsymbol u^k-\boldsymbol u^*)^{\rm T}F(\tilde{\boldsymbol v}^k). \quad (41)$$

Since $\tilde{\boldsymbol v}^k\in\Omega$, using the monotonicity of $F$, we have

$$(\tilde{\boldsymbol v}^k-\boldsymbol u^*)^{\rm T}F(\tilde{\boldsymbol v}^k)\geq(\tilde{\boldsymbol v}^k-\boldsymbol u^*)^{\rm T}F(\boldsymbol u^*)\geq0,$$

and consequently

$$(\boldsymbol u^k-\boldsymbol u^*)^{\rm T}F(\tilde{\boldsymbol v}^k)\geq(\boldsymbol u^k-\tilde{\boldsymbol v}^k)^{\rm T}F(\tilde{\boldsymbol v}^k). \quad (42)$$

Applying Eq.(42) to the last term on the right side of Eq.(41), we obtain

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k)\geq\|\boldsymbol u^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2$$
$$+2\alpha\beta_k[u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\tilde{\boldsymbol v}^k]^{\rm T}F(\tilde{\boldsymbol v}^k). \quad (43)$$

Since $\tilde{\boldsymbol v}^k=P_\Omega[\boldsymbol u^k-\beta_kF(\boldsymbol v^k)]$ and $u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\in\Omega$, it follows from Eq.(19) that for any $\alpha>0$,

$$0\geq2\alpha[u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\tilde{\boldsymbol v}^k]^{\rm T}\{[\boldsymbol u^k-\beta_kF(\boldsymbol v^k)]-\tilde{\boldsymbol v}^k\}. \quad (44)$$

Adding Eqs.(43) and (44), we obtain

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k)\geq\|\boldsymbol u^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2+2\alpha[u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\tilde{\boldsymbol v}^k]^{\rm T}\boldsymbol d^k.$$

By manipulation, we obtain

$$\theta_{\rm II}(\alpha,\tilde{\boldsymbol v}^k)$$
$$\geq\|\boldsymbol u^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2)+2\alpha\{[u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\boldsymbol u^k]+(\boldsymbol u^k-\tilde{\boldsymbol v}^k)\}^{\rm T}\boldsymbol d^k$$
$$=\|\boldsymbol u^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)-\alpha\boldsymbol d^k\|^2+2\alpha(\boldsymbol u^k-\tilde{\boldsymbol v}^k)^{\rm T}\boldsymbol d^k-\alpha^2\|\boldsymbol d^k\|^2$$
$$=\Phi(\alpha)+\|\boldsymbol u^k-\alpha\boldsymbol d^k-u_{\rm II}^{k+1}(\alpha,\tilde{\boldsymbol v}^k)\|^2=\Phi_{\rm II}(\alpha).$$

This completes the proof.

**Remark 2** Algorithm II can be viewed as an extension of General Extra-gradient methods proposed in (Korpelevich, 1976). Obviously, if we let $\boldsymbol u^k=\boldsymbol v^k$ in Algorithm II, we will obtain General Extra-gradient methods.

**Remark 3** We can make Algorithms II more general, if $\tilde{\boldsymbol v}^k$ is given by Eq.(24) and $\boldsymbol\xi^k$ can be defined in any form, only if the inexactness criterion Eq.(14) holds, we can also prove the convergence of the extending Algorithms II. The following theorem gives preliminary results for convergence analysis of the extending Algorithms II.

**Theorem 3** Given $\boldsymbol{u}^k \in \Omega$, let $\tilde{\boldsymbol{v}}^k$ be a predictor given by Eq.(24) and $\boldsymbol{u}^k$, $\tilde{\boldsymbol{v}}^k$ and $\boldsymbol{\xi}^k$ satisfy inexactness criterion Eq.(14). We let the new iteration $\boldsymbol{u}_\alpha^{k+1}$ be given by the general forms of Algorithms II. We denote $\boldsymbol{u}_\alpha^{k+1} := u_{\mathrm{II}}^{k+1}(\alpha, \tilde{\boldsymbol{v}}^k)$ given in Eq.(13),

$$\|\boldsymbol{u}_\alpha^{k+1} - \boldsymbol{u}^*\|^2 \le \|\boldsymbol{u}^k - \boldsymbol{u}^*\|^2 - \Phi(\alpha), \ \forall \alpha > 0, \tag{45}$$

and $\Phi(\alpha)$ was defined in Eq.(33).

**Proof** Denote $\Theta_k(\alpha) := \|\boldsymbol{u}^k - \boldsymbol{u}^*\|^2 - \|\boldsymbol{u}_\alpha^{k+1} - \boldsymbol{u}^*\|^2$. In the following we prove the equivalent assertion $\Theta_k(\alpha) \ge \Phi(\alpha)$. Since $\boldsymbol{u}^* \in \Omega$ and $\boldsymbol{u}_\alpha^{k+1} = P_\Omega[\boldsymbol{u}^k - \alpha\beta_k F(\tilde{\boldsymbol{v}}^k)]$, it follows from Eq.(21) that

$$\|\boldsymbol{u}_\alpha^{k+1} - \boldsymbol{u}^*\|^2 \le \|\boldsymbol{u}^k - \alpha\beta_k F(\tilde{\boldsymbol{v}}^k) - \boldsymbol{u}^*\|^2 - \|\boldsymbol{u}^k - \alpha\beta_k F(\tilde{\boldsymbol{v}}^k) - \boldsymbol{u}_\alpha^{k+1}\|^2. \tag{46}$$

Consequently, using the notation of $\Theta_k(\alpha)$, we get

$$\begin{aligned}\Theta_k(\alpha) &\ge \|\boldsymbol{u}^k - \boldsymbol{u}^*\|^2 + \|\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1} - \alpha\beta_k F(\tilde{\boldsymbol{v}}^k)\|^2 - \|\boldsymbol{u}^k - \boldsymbol{u}^* - \alpha\beta_k F(\tilde{\boldsymbol{v}}^k)\|^2 \\ &= \|\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1}\|^2 + 2\alpha\beta_k(\boldsymbol{u}_\alpha^{k+1} - \boldsymbol{u}^k)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k) + 2\alpha\beta_k(\boldsymbol{u}^k - \boldsymbol{u}^*)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k).\end{aligned}$$

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

$$(\tilde{\boldsymbol{v}}^k - \boldsymbol{u}^*)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k) \ge (\tilde{\boldsymbol{v}}^k - \boldsymbol{u}^*)^{\mathrm{T}} F(\boldsymbol{u}^*). \tag{47}$$

Since $\tilde{\boldsymbol{v}}^k \in \Omega$, we have $(\tilde{\boldsymbol{v}}^k - \boldsymbol{u})^{\mathrm{T}} F(\boldsymbol{u}^*) \ge 0$ and consequently it follows from Eq.(47) that

$$(\boldsymbol{u}^k - \boldsymbol{u}^*)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k) \ge (\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k). \tag{48}$$

Applying Eq.(48) to the last term on the right side of the above expression of $\Theta_k(\alpha)$, we obtain

$$\begin{aligned}\Theta_k(\alpha) &\ge \|\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1}\|^2 + 2\alpha\beta_k(\boldsymbol{u}_\alpha^{k+1} - \boldsymbol{u}^k)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k) + 2\alpha\beta_k(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k) \\ &= \|\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1}\|^2 + 2\alpha\beta_k(\boldsymbol{u}_\alpha^{k+1} - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} F(\tilde{\boldsymbol{v}}^k).\end{aligned} \tag{49}$$

Using $a^2 \ge 2ab - b^2$ and the notation of $\Phi(\alpha)$, we have

$$\begin{aligned}\|\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1}\|^2 &\ge 2\alpha(\boldsymbol{u}^k - \boldsymbol{u}_\alpha^{k+1})^{\mathrm{T}} \boldsymbol{d}^k - \alpha^2\|\boldsymbol{d}^k\|^2 \\ &= 2\alpha[(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k) + (\tilde{\boldsymbol{v}}^k - \boldsymbol{u}_\alpha^{k+1})]^{\mathrm{T}} \boldsymbol{d}^k - \alpha^2\|\boldsymbol{d}^k\|^2 \\ &= \Phi(\alpha) + 2\alpha(\tilde{\boldsymbol{v}}^k - \boldsymbol{u}_\alpha^{k+1})^{\mathrm{T}} \boldsymbol{d}^k.\end{aligned} \tag{50}$$

Substituting Eq.(50) into Eq.(49) we obtain

$$\Theta_k(\alpha) \ge \Phi(\alpha) + 2\alpha(\tilde{\boldsymbol{v}}^k - \boldsymbol{u}_\alpha^{k+1})^{\mathrm{T}}[\boldsymbol{d}^k - \beta_k F(\tilde{\boldsymbol{v}}^k)]. \tag{51}$$

Now we consider the last term on the right side of Eq.(51). Setting $\boldsymbol{v} = \boldsymbol{u}^k - \beta_k F(\tilde{\boldsymbol{v}}^k) + \boldsymbol{\xi}^k$ and $\boldsymbol{\omega} = \boldsymbol{u}_\alpha^{k+1}$ in the basic inequality Eq.(19), we get

$$\{\boldsymbol{u}^k - \beta_k F(\tilde{\boldsymbol{v}}^k) + \boldsymbol{\xi}^k - P_\Omega[\boldsymbol{u}^k - \beta_k F(\tilde{\boldsymbol{v}}^k) + \boldsymbol{\xi}^k]\}^{\mathrm{T}} \cdot \{P_\Omega[\boldsymbol{u}^k - \beta_k F(\tilde{\boldsymbol{v}}^k) + \boldsymbol{\xi}^k] - \boldsymbol{u}_\alpha^{k+1}\} \ge 0.$$

Since $\tilde{\boldsymbol{v}}^k = P_\Omega[\boldsymbol{u}^k - \beta_k F(\tilde{\boldsymbol{v}}^k) + \boldsymbol{\xi}^k]$ and $\boldsymbol{d}^k = \boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k + \boldsymbol{\xi}^k$, it follows from the above inequality that

$$[\boldsymbol{d}^k - \beta_k F(\tilde{\boldsymbol{v}}^k)]^{\mathrm{T}} (\tilde{\boldsymbol{v}}^k - \boldsymbol{u}_\alpha^{k+1}) \ge 0. \tag{52}$$

Substituting Eq.(52) into Eq.(51) we obtain $\Theta_k(\alpha) \ge \Phi(\alpha)$. This completes the proof.

## MAIN THEORETICAL RESULT

The assertions of Theorems 1~3 are similar. Since

$$\Phi(\alpha) = 2\alpha(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k - \alpha^2\|\boldsymbol{d}^k\|^2,$$

it follows from Theorems 1~3 that both Algorithm I and Algorithm II (including the extending Algorithm II) are contraction methods for any

$$\alpha \in \left(0, \ \frac{2(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k}{\|\boldsymbol{d}^k\|^2}\right). \tag{53}$$

Note that under inexactness restriction Eq.(14) we have

$$\begin{aligned}(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k &= (\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} (\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k + \boldsymbol{\xi}^k) \\ &= \|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|^2 + (\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{\xi}^k \\ &\ge (1-\nu)\|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|^2.\end{aligned} \tag{54}$$

Since $\Phi(\alpha)$ is a quadratic funtion of $\alpha$, it reaches its maximum at

$$\alpha_k^* = \frac{(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k}{\| \boldsymbol{d}^k \|^2}$$

In addition, since $0<\nu<1$, we have

$$\alpha_k^* \geq \frac{1-\nu}{1+\mu^2}. \tag{55}$$

The following theorem gives a common result for both Algorithms I and Algorithms II (including the extending Algorithms II).

**Theorem 4** Given $\boldsymbol{u}^k \in \Omega$ and $\beta_k \geq \beta > 0$, let $\boldsymbol{d}^k$ and $\tilde{\boldsymbol{v}}^k$ be given by Eqs.(11) and (24) respectively, if $\boldsymbol{u}^k$, $\tilde{\boldsymbol{v}}^k$ and $\boldsymbol{\xi}^k$ satisfy inexactness criterion Eq.(14), and

$$\alpha_k^* = \frac{(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k}{\| \boldsymbol{d}^k \|^2}, \ \ \alpha_k = \gamma_k \alpha_k^*, \ \gamma_k \in (0, 2), \tag{56}$$

whether the new iteration $\boldsymbol{u}^{k+1}$ is generated by

$$\boldsymbol{u}^{k+1} = P_{\Omega}(\boldsymbol{u}^k - \alpha_k \boldsymbol{d}^k)$$

or

$$\boldsymbol{u}^{k+1} = P_{\Omega}[\boldsymbol{u}^k - \alpha_k \beta_k F(\tilde{\boldsymbol{v}}^k)],$$

we have, $\forall \boldsymbol{u}^* \in \Omega$

$$\|\boldsymbol{u}^{k+1} - \boldsymbol{u}^*\|^2 \leq \|\boldsymbol{u}^k - \boldsymbol{u}^*\|^2 - \gamma_k(2-\gamma_k)\frac{(1-\nu)^2}{1+\mu^2}\|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|^2, \tag{57}$$

**Proof** From Theorems 1~3 we have

$$\|\boldsymbol{u}^k - \boldsymbol{u}^*\|^2 - \|\boldsymbol{u}^{k+1} - \boldsymbol{u}^*\|^2 \geq \Phi(\alpha_k).$$

Using Eqs.(33), (54)~(56) we obtain

$$\begin{aligned} \Phi(\alpha_k) &= 2\alpha_k(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k - (\gamma_k^2 \alpha_k^*)(\alpha_k^* \|\boldsymbol{d}^k\|^2) \\ &= (2\gamma_k \alpha_k^* - \gamma_k^2 \alpha_k^*)(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k \\ &= \gamma_k(2-\gamma_k)\alpha_k^*(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k \\ &\geq \gamma_k(2-\gamma_k)\frac{(1-\nu)^2}{1+\mu^2}\|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|^2, \end{aligned}$$

and the assertion is proved.

**Theorem 5** Given $\boldsymbol{u}^k \in \Omega$ and $\beta_k \geq \beta > 0$, let $\tilde{\boldsymbol{v}}^k$ be a predictor given by Eq.(24) and $\boldsymbol{u}^k$, $\tilde{\boldsymbol{v}}^k$ and $\boldsymbol{\xi}^k$ satisfy inexactness criterion Eq.(14), the sequence $\{\boldsymbol{u}^k\}$ generated by Algorithm I and Algorithm II (including the extending Algorithm II) converges to $\boldsymbol{u}^\infty \in \Omega^*$.

**Proof** It follows from Eq.(57) that $\{\boldsymbol{u}^k\}$ is bounded and $\lim_{k\to\infty}\|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|=0$. It follows from Eq.(29) that

$$\| e(\tilde{\boldsymbol{v}}^k, \beta)\|^2 \leq (1+\mu)\|\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k\|,$$

and thus

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

Then, it follows from Lemma 2 that $\{\boldsymbol{u}^k\}$ has a subsequence $\{\boldsymbol{u}^{k_j}\}$ which converges to a point $\boldsymbol{u}^\infty$, and $u^\infty$ is a solution point of VI($\Omega$, $F$). Since Eq.(57) is true for all solution points of VI($\Omega$, $F$), we have $\|\boldsymbol{u}^{k+1} - \boldsymbol{u}^\infty\| \leq \|\boldsymbol{u}^k - \boldsymbol{u}^\infty\|$, $\forall k \geq 0$, and it follows that the sequence $\{\boldsymbol{u}^k\}$ converges to $u^\infty$.

**Remark 4** From the above Theorems 1~5, we proved the convergence of Algorithm I and Algorithm II (including the extending Algorithm II).

In general, Eq.(30) in Theorem 1 [resp. Eq.(37) in Theorem 2] is tight. This can be seen from the following example. Let us consider a VI($\Omega$, $F$) with

$$\Omega = \mathbb{R}^2, \ F(\boldsymbol{u}) = \boldsymbol{M}\boldsymbol{u}, \ \text{and} \ \boldsymbol{M} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This variational inequality is monotone and has a unique solution $\boldsymbol{u}^*=0$. Note that

$$\boldsymbol{M}^2 = -\boldsymbol{I} \ \text{and} \ \boldsymbol{M}^{\mathrm{T}}\boldsymbol{M} = \boldsymbol{I}.$$

If we let any $\boldsymbol{v}^k \in \Omega$ be an approximate solution of Eq.(5) in the sense of Eq.(8) and $\beta \in (0,1)$ and let $\boldsymbol{u}^k = \boldsymbol{v}^k$, we have $\boldsymbol{\xi}^k = \beta^2 \boldsymbol{v}^k$ and $\boldsymbol{d}^k = \beta^2 \boldsymbol{v}^k + \beta \boldsymbol{M}\boldsymbol{v}^k$.

Using $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{u}=0$ and $\|\boldsymbol{M}\boldsymbol{u}\| = \|\boldsymbol{u}\|$ we get

$$(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k = \beta^2 \|\boldsymbol{v}^k\|^2$$

and

$$\|\boldsymbol{d}^k\|^2 = \beta^2(1+\beta^2)\|\boldsymbol{v}^k\|^2.$$

When the problem is solved by Algothms I with $\alpha \in (0, 2/(1+\beta^2))$, we have

$$\begin{aligned} \| \boldsymbol{u}_{\mathrm{I}}^{k+1}(\alpha, \tilde{\boldsymbol{v}}^k)\|^2 &= \|(1-\alpha\beta^2)\boldsymbol{v}^k - \alpha\beta \boldsymbol{M}\boldsymbol{v}^k\|^2 \\ &= [1 - 2\alpha\beta^2 + \alpha^2\beta^2(1+\beta^2)]\|\boldsymbol{v}^k\|^2. \end{aligned}$$

and

$$\begin{aligned} \Phi_{\mathrm{I}}(\alpha, \tilde{\boldsymbol{v}}^k) = \Phi(\alpha, \tilde{\boldsymbol{v}}^k) &= 2\alpha(\boldsymbol{u}^k - \tilde{\boldsymbol{v}}^k)^{\mathrm{T}} \boldsymbol{d}^k - \alpha^2\|\boldsymbol{d}^k\|^2 \\ &= 2\alpha\beta^2\|\boldsymbol{v}^k\|^2 - \alpha^2\beta^2(1+\beta^2)\|\boldsymbol{v}^k\|^2 \\ &= [2\alpha\beta^2 - \alpha^2\beta^2(1+\beta^2)]\|\boldsymbol{v}^k\|^2. \end{aligned}$$

so

$$\Theta_{\mathrm{I}}(\alpha, \tilde{\boldsymbol{v}}^k) = [2\alpha\beta^2 - \alpha^2\beta^2(1+\beta^2)]\|\boldsymbol{v}^k\|^2 = \Phi_{\mathrm{I}}(\alpha, \tilde{\boldsymbol{v}}^k).$$

For Algorithm II, since in this special example, we have

$$\beta F(\tilde{\boldsymbol{v}}^k)=\beta\boldsymbol{M}(\boldsymbol{I}-\beta\boldsymbol{M})\boldsymbol{v}^k=\beta\boldsymbol{M}\boldsymbol{v}^k+\beta^2\boldsymbol{v}^k=\boldsymbol{d}^k,$$

we also have $\Theta_{\mathrm{II}}(\alpha,\tilde{\boldsymbol{v}}^k)=\Phi_{\mathrm{II}}(\alpha,\tilde{\boldsymbol{v}}^k)$, which means that Eq.(37) is also tight in this example.

Nevertheless, the following theorem indicates that in each iterative step, we may expect Algorithm II to get more progress than Algorithm I.

**Theorem 6** Let $\Phi_{\mathrm{I}}(\alpha,\tilde{\boldsymbol{v}}^k)$ and $\Phi_{\mathrm{II}}(\alpha,\tilde{\boldsymbol{v}}^k)$ be defined as in Eqs.(32) and (39). We have

$$\Phi_{\mathrm{II}}(\alpha,\tilde{\boldsymbol{v}}^k)-\Phi_{\mathrm{I}}(\alpha,\tilde{\boldsymbol{v}}^k)\geq\|u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2. \quad (58)$$

**Proof** It follows from Eqs.(32) and (39) that

$$\Phi_{\mathrm{II}}(\alpha,\tilde{\boldsymbol{v}}^k)-\Phi_{\mathrm{I}}(\alpha,\tilde{\boldsymbol{v}}^k)=\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2 -\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2. \quad (59)$$

Note that $u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\in\Omega$, setting $\boldsymbol{v}=\boldsymbol{u}^k-\alpha\boldsymbol{d}^k$ and $\boldsymbol{u}=u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)$ in Eq.(21), we obtain

$$\|u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)-u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2\leq\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{II}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2 -\|\boldsymbol{u}^k-\alpha\boldsymbol{d}^k-u_{\mathrm{I}}^{k+1}(\alpha,\tilde{\boldsymbol{v}}^k)\|^2. \quad (60)$$

The assertion of this theorem follows directly from Eqs.(59) and (60).

## IMPLEMENTATION DETAILS AND NUMERICAL EXPERIMENTS

In this section, we consider a special class of variational inequalities problems in the following format

$$\exists\boldsymbol{x}\in\Omega,\ (\boldsymbol{x}-\boldsymbol{x}^*)^{\mathrm{T}}F(\boldsymbol{x}^*)\geq0,\ \forall\boldsymbol{x}\in\Omega, \quad (61)$$

where $F(\boldsymbol{x})=F_1(\boldsymbol{x})+F_2(\boldsymbol{x})$, $F_2(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{q}$, and $A\in\mathbb{R}^{n\times n}$. $F_1(\boldsymbol{x})$ is a nonlinear separated monotone function, that is $F_1(\boldsymbol{x})=(F_{1i}(\boldsymbol{x}_i))_{i=1}^n$, and $\partial F/\partial x\geq0$.

It is well known that when solving linear equation, Gauss-Seidel iteration method converges faster than Jacobi iteration method, because Gauss-Seidel iteration method makes full use of the latest information to improve convergence rate. Inspired by such idea, we adopt the extending Algorithm II to solve the above variational inequalities problems. In order to make it clearer, we give the prediction step form again:

$$\text{(Prediction)}\quad \tilde{\boldsymbol{x}}^k=P_\Omega[\boldsymbol{x}^k-\beta_kF(\tilde{\boldsymbol{x}}^k)+\boldsymbol{\xi}^k].$$

If we denote $\boldsymbol{\xi}^k=\beta_k\boldsymbol{A}(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)$, we will have

$$\tilde{\boldsymbol{x}}^k=P_\Omega\{\tilde{\boldsymbol{x}}^k-\{(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)+\beta_k[F_1(\tilde{\boldsymbol{x}}^k)+\boldsymbol{A}\boldsymbol{x}^k+\boldsymbol{q}]\}\}$$
$$\Leftrightarrow\exists\,\tilde{\boldsymbol{x}}^k\in\Omega,\ \forall\boldsymbol{x}\in\Omega,$$
$$(\boldsymbol{x}-\tilde{\boldsymbol{x}}^k)^{\mathrm{T}}\{(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)+\beta_k[F_1(\tilde{\boldsymbol{x}}^k)+\boldsymbol{A}\boldsymbol{x}^k+\boldsymbol{q}]\}\geq0. \quad (62)$$

Because $F_1(\boldsymbol{x})$ is separated, so variational inequalities Eq.(62) can be decomposed into a series of 1D sub-variational inequalities problems, which can be easily solved.

According to our computational experience, we take the relaxation factor $\gamma$ in an open interval (0, 2), which is close to 2, say 1.8, to ensure faster convergence. The parameter $\beta$ should not only satisfy the condition Eq.(14), but also be adjusted appropriately and not be too small. Otherwise slow convergence may result. In our practical algorithm, we prefer to use the following self-adaptive technique:

$$\beta_{k+1}=\begin{cases}0.9\nu\beta_k/\gamma_k, & \text{if } \gamma_k<\mu,\\ \beta_k, & \text{otherwise,}\end{cases}$$

where

$$r_k=\frac{\beta_k\|\boldsymbol{A}(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)\|}{\|\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k\|}.$$

Now we present the following practical algorithm:

Step 0: Let $\beta_0>0$, $\boldsymbol{x}^0\in\Omega$, $0<\mu<1<\nu$, $\gamma=1.8$, $\varepsilon=10^{-8}$ and $k=0$.

Step 1: Find $\tilde{\boldsymbol{x}}^k$ satisfying: $\forall\boldsymbol{x}\in\Omega$,

$$(\boldsymbol{x}-\tilde{\boldsymbol{x}}^k)^{\mathrm{T}}\{(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)+\beta_k[F_1(\tilde{\boldsymbol{x}}^k)+\boldsymbol{A}\boldsymbol{x}^k+\boldsymbol{q}]\}\geq0, \quad (63)$$

and

$$\|(\boldsymbol{x}^k-\tilde{\boldsymbol{x}}^k)^{\mathrm{T}}\boldsymbol{\xi}^k\|\leq\nu\|\boldsymbol{x}^k-\tilde{\boldsymbol{x}}^k\|^2,\ \|\boldsymbol{\xi}^k\|\leq\mu\|\boldsymbol{x}^k-\tilde{\boldsymbol{x}}^k\|,\ \nu<1\leq\mu.$$

Step 2: Calculate $r_k=\beta_k\|\boldsymbol{A}(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)\|/\|\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k\|$. If $r_k>\nu$, reduce the value of $\beta_k$, $\beta_k=\beta_k\nu/r_k$, go to Step 1.

Step 3: Correction step: $\boldsymbol{x}^{k+1}=P_\Omega[\boldsymbol{x}^k-\alpha_k\beta_kF(\tilde{\boldsymbol{x}}^k)]$,

$$\alpha_k=\gamma\alpha^*,\ \alpha^*=(\boldsymbol{x}^k-\tilde{\boldsymbol{x}}^k)^{\mathrm{T}}\boldsymbol{d}^k/\|\boldsymbol{d}^k\|^2,$$
$$\boldsymbol{d}^k=\boldsymbol{x}^k-\tilde{\boldsymbol{x}}^k+\boldsymbol{\xi}^k,\ \boldsymbol{\xi}^k=\beta_k\boldsymbol{A}(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k).$$

Step 4: If $\|e(\boldsymbol{x}^{k+1})\|>\varepsilon$, set $\beta_{k+1}=\beta_k$, $k=k+1$, go to Step 1.

To show the ability of the extending Algorithm II for solving variational inequalities problems Eq.(61), we take two examples from (Li and Zeng, 2003).

**Example 1** In this problem, the constraint set $\Omega$ and the mapping $F_1(\boldsymbol{x})$ are taken respectively as

$$\Omega=\mathbb{R}_+^{N^2},$$
$$F_1(\boldsymbol{x})=\arctan(\boldsymbol{x})=(\arctan x_1, \arctan x_2, \arctan x_3, \ldots, \arctan x_{N^2})^{\mathrm{T}}.$$

That is

$$\exists \boldsymbol{x}\geq 0,\ (\boldsymbol{x}'-\boldsymbol{x})^{\mathrm{T}}F(\boldsymbol{x})\geq 0,\ \forall \boldsymbol{x}'\geq 0, \tag{64}$$

where $F(\boldsymbol{x})=F_1(\boldsymbol{x})+\boldsymbol{Ax}+\boldsymbol{q}$, $\boldsymbol{A}$ is a block $N^2\times N^2$ matrix and $\boldsymbol{B}$ is an $N\times N$ matrix:

$$\boldsymbol{A}=\begin{pmatrix} \boldsymbol{B} & -\boldsymbol{I} & \boldsymbol{0} & \cdots & \cdots & \boldsymbol{0} \\ -\boldsymbol{I} & \boldsymbol{B} & -\boldsymbol{I} & \ddots & & \vdots \\ \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{B} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \boldsymbol{0} \\ \vdots & & \ddots & -\boldsymbol{I} & \boldsymbol{B} & -\boldsymbol{I} \\ \boldsymbol{0} & \cdots & \cdots & \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{B} \end{pmatrix}, \tag{65}$$

$$\boldsymbol{B}=\begin{pmatrix} 4 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 4 & -1 & \ddots & & \vdots \\ 0 & -1 & 4 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & 4 & -1 \\ 0 & \cdots & \cdots & 0 & -1 & 4 \end{pmatrix}. \tag{66}$$

It is well known that $\boldsymbol{A}$ is positive definite. In this example, variational inequalities problem Eq.(63) in Step 1 of the above practical algorithm can be converted into the following complementarity problems:

$$\tilde{\boldsymbol{x}}^k\geq 0,\ \tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k+\beta_k[F_1(\tilde{\boldsymbol{x}}^k)+\boldsymbol{Ax}^k+\boldsymbol{q}]\geq 0,$$
$$(\tilde{\boldsymbol{x}}^k)^{\mathrm{T}}\{(\tilde{\boldsymbol{x}}^k-\boldsymbol{x}^k)+\beta_k[F_1(\tilde{\boldsymbol{x}}^k)+\boldsymbol{Ax}^k+\boldsymbol{q}]\}=0.$$

We denote

$$T(\boldsymbol{x}):=\boldsymbol{x}-\boldsymbol{x}^k+\beta_k[F_1(\boldsymbol{x})+\boldsymbol{Ax}^k+\boldsymbol{q}]. \tag{67}$$

Note that $T(\boldsymbol{x})$ is a monotone function, so Step 1 can be implemented in this way [$T_i(\boldsymbol{x})$ denotes the $i$th component of $T(\boldsymbol{x})$]: if $T_i(\boldsymbol{0})\geq 0$, then $\tilde{x}_i^k=0$; else find out the solution of equation $T_i(\boldsymbol{x})=0$ by Newton method.

To investigate the convergence behaviour of the extending Algorithm II, we form random test problems Eq.(64) with matrices Eqs.(65) and (66) as follows: First, we choose $\boldsymbol{v}\in\mathbb{R}^{N^2}$, and $v_i$ ($i$=1, …, $N^2$) randomly in (−5, 5). Then let

$$x_i^*=\max(0, v_i),\ f_i=\max(0,-v_i),\ i=1, \ldots, N^2$$

and

$$\boldsymbol{q}=\boldsymbol{f}-\boldsymbol{Ax}^*-\arctan \boldsymbol{x}^*.$$

In this way we get a random test problem with a given solution $\boldsymbol{x}^*$. Numerical results for this example solved by the extending Algorithm II are shown in Table 1. Here $n=N^2$ is the number of variables. The code is written in Matlab on a TOSHIBA notebook (Pentium M 1.6 GHz).

**Table 1 Numerical results for the first example**

| $N$ | $n$ | Number of iterations | CPU time (s) | $\Vert\boldsymbol{x}^k-\boldsymbol{x}^*\Vert_\infty$ ($\times10^{-9}$) |
|---|---|---|---|---|
| 10 | 100 | 102 | 7.7110 | 1.4 |
| 20 | 400 | 101 | 33.8380 | 1.3 |
| 30 | 900 | 79 | 59.5460 | 1.1 |
| 40 | 1600 | 100 | 217.4830 | 1.3 |
| 50 | 2500 | 98 | 623.3060 | 1.3 |

**Example 2** In this problem, the constraint set $\Omega$ and the mapping $F_1$ are taken respectively as

$$\Omega=\{\boldsymbol{x}\in\mathbb{R}^{N^2}\mid l_i\leq x_i\leq h_i,\ i=1, \ldots, N^2\},$$
$$F_1(\boldsymbol{x})=\arctan(\boldsymbol{x})=(\arctan x_1, \arctan x_2, \arctan x_3, \ldots, \arctan x_{N^2})^{\mathrm{T}}.$$

That is

$$\exists \boldsymbol{x}\in\Omega,\ (\boldsymbol{x}'-\boldsymbol{x})^{\mathrm{T}}F(\boldsymbol{x})\geq 0,\ \forall \boldsymbol{x}'\in\Omega, \tag{68}$$

where $F(\boldsymbol{x})=F_1(\boldsymbol{x})+\boldsymbol{Ax}+\boldsymbol{q}$, $\boldsymbol{A}$ and $\boldsymbol{B}$ are defined in Eq.(65) and Eq.(66) respectively. In this example, variational inequalities problem Eq.(63) in Step 1 of the above practical algorithm can be converted into the following complementarity problems:

$$\begin{cases} \tilde{x}_i^k=l_i, & T_i(\tilde{x}_i^k)\geq 0, \\ l_i<\tilde{x}_i^k<h_i, & T_i(\tilde{x}_i^k)=0, \quad i=1, \ldots, N^2 \\ \tilde{x}_i^k=h_i, & T_i(\tilde{x}_i^k)\leq 0, \end{cases}$$

where $T(\boldsymbol{x})$ was denoted in Eq.(67) and $T_i(\boldsymbol{x})$ denotes the $i$th component of $T(\boldsymbol{x})$. So Step 1 can be implemented in this way ($i$=1, …, $N^2$): if $T_i(l_i)\geq 0$, then $\tilde{x}_i^k = l_i$; else if $T_i(h_i)\leq 0$, then $\tilde{x}_i^k = h_i$; else find the zero point of $T_i(\tilde{x}_i^k)=0$.

To investigate the convergence behaviour of the extending Algorithm II, we form random test problems Eq.(68) with matrices (65) and (66) as follows: First, for convenience, we set $l_i$=0 ($i$=1, …, $N^2$), and choose $h_i$ ($i$=1, …, $N^2$) randomly in (10, 20) and $t_i$ ($i$=1, …, $N^2$) randomly in (0, 1), then let

$$x_i^* = \begin{cases} 0, & t_i \leq 0.25, \\ (2t_i - 0.5)h_i, & 0.25 < t_i \leq 0.75, \\ h_i, & \text{otherwise}, \end{cases}$$

and

$$f_i = \begin{cases} \text{randomly in } (0,10), & t_i \leq 0.25, \\ 0, & 0.25 < t_i \leq 0.75, \\ \text{randomly in } (-10,0), & \text{otherwise}, \end{cases}$$

and

$$\boldsymbol{q}=\boldsymbol{f}-\boldsymbol{A}\boldsymbol{x}^*-\arctan\boldsymbol{x}^*.$$

In this way we get a random test problem with a given solution $\boldsymbol{x}^*$. Numerical results for this example solved by the extending Algorithm II are shown in Table 2. Here $n=N^2$ is the number of variables. The code is written in Matlab on a TOSHIBA notebook (Pentium M 1.6 GHz).

**Table 2 Numerical results for the second example**

| $N$ | $n$ | Number of iterations | CPU time (s) | $\lVert\boldsymbol{x}^k-\boldsymbol{x}^*\rVert_\infty$ ($\times10^{-9}$) |
|---|---|---|---|---|
| 10 | 100 | 105 | 7.1500 | 1.2 |
| 20 | 400 | 95 | 21.7310 | 1.3 |
| 30 | 900 | 85 | 57.8030 | 1.1 |
| 40 | 1600 | 95 | 207.1270 | 1.0 |
| 50 | 2500 | 65 | 531.0630 | 1.0 |

## CONCLUSION

In this work, we investigated the relationship between Algorithm I and Algorithm II for solving monotone variational inequalities (MVI) problems. Given $\boldsymbol{u}^k$, $\boldsymbol{v}^k\in\Omega$, both methods take the same prediction step. Once the same prediction step size is determined, the two methods have the same range for the correction step size. The only difference is that they use different search directions in the correction step. The computation costs of the two methods in each iteration are almost equal. Theoretical analysis showed that in general, we can expect Algorithm II to have better performance than Algorithm I. Considering the special structure of the separated MVI problems completely, we adopted the extending Algorithm II to solve the separated MVI problems. It was shown that the extending Algorithm II is simple in implementation and not expensive in computation time. Of course, there still exists somewhere some not satisfactory enough aspects [for instance, whether or not we can find some examples to show that the left side of Eq.(58) is strictly greater than the right side of it when Eqs.(30) and (37) are both tight] which deserve further study.

## References

Bertsekas, D.P., Tsitsiklis, J.N., 1989. Parallel and Distributed Computation, Numerical Methods. Prentice-Hall, Englewood Clis, NJ, p.267-268.

Harker, P.T., Pang, J.S., 1990. Finite-dimensional variational inequality and nonlinear complementarity problem: a survey of theory, algorithms and application. *Math. Programming*, **48**:161-220. [doi:10.1007/BF01582255]

He, B.S., Qian, M.J., Wang, Y.M., 2004. Study on Some Approximate Proximal Point Algorithms for Monotone Variational Inequalities. *In*: Yuan, Y.X. (Ed.), Numerical Linear Algebra and Optimization. Proc. 2003 Int. Conf. Numerical Optimization and Numerical Linear Algebra. Science Press, Beijing/New York, p.15-41 (in Chinese).

Korpelevich, G.M., 1976. The extragradient method for finding saddle points and other problems. *Ekonomika i Matematchskie Metody*, **12**:747-756. [English Translation: *Matecon*, **13**(1977):35-49]

Li, C.L., Zeng, J.P., 2003. Much splitting and adding Schwarz iteration for nonlinear complementarity problems. *Num. Comput. Appl. Computer*, **25**:269-275 (in Chinese).

Rockafellar, R.T., 1976. Monotone operators and the proximal point algorithm. *SIAM J. Control & Optim.*, **14**:877-898. [doi:10.1137/0314056]

Tseng, P., 2000. A modified forward-backward splitting method for maximal monotone mappings. *SIAM J. Control & Optim.*, **38**:431-446. [doi:10.1137/S0363012998338806]

Zhu, T., Yu, Z.G., 2004. A simple proof for some important properties of the projection mapping. *Math. Ineq. Appl.*, **7**:453-456.