On the $O(1/t)$ convergence rate of the projection and contraction methods for variational inequalities with Lipschitz continuous monotone operators

Bingsheng He[1]

Department of Mathematics, Nanjing University, Nanjing, 210093, P. R. China.

Abstract. Recently, Nemirovski's analysis indicates that the extragradient method has the $O(1/t)$ convergence rate for variational inequalities with Lipschitz continuous monotone operators. For the same problems, in the last decades, we have developed a class of Fejér monotone projection and contraction methods. Until now, only convergence results are available to these projection and contraction methods, though the numerical experiments indicate that they always outperform the extragradient method. The reason is that the former benefits from the 'optimal' step size in the contraction sense. In this paper, we prove the convergence rate under a unified conceptual framework, which includes the projection and contraction methods as special cases and thus perfects the theory of the existing projection and contraction methods. Preliminary numerical results demonstrate that the projection and contraction methods converge two times faster than the extragradient method.

Keywords. Variational inequality, projection and contraction method, convergence rate.

1   Introduction

Let Ω be a nonempty closed convex subset of $\Re^n$, F be a continuous mapping from $\Re^n$ to itself. The variational inequality problem, denoted by $\text{VI}(\Omega, F)$, is to find a vector $u^* \in \Omega$ such that

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

Notice that $\text{VI}(\Omega, F)$ is invariant when F is multiplied by some positive scalar $\beta > 0$. It is well known ([1], pp. 267) that, for any $\beta > 0$,

$$u^* \text{ is a solution of } \text{VI}(\Omega, F) \iff u^* = P_\Omega[u^* - \beta F(u^*)], \tag{1.2}$$

where $P_\Omega(\cdot)$ denotes the projection onto Ω with respect to the Euclidean norm, *i.e.*,

$$P_\Omega(v) = \text{argmin}\{\|u - v\| \,|\, u \in \Omega\}.$$

Throughout this paper we assume that the mapping F is monotone and Lipschitz continuous, *i.e.*,

$$(u - v)^T (F(u) - F(v)) \geq 0, \quad \forall\, u, v \in \Re^n,$$

and there is a constant $L > 0$ (not necessary known), such that

$$\|F(u) - F(v)\| \leq L\|u - v\|, \quad \forall\, u, v \in \Re^n.$$

Moreover, we assume that the solution set of $\text{VI}(\Omega, F)$, denoted by Ω^*, is nonempty. The nonempty assumption of the solution set, together with the monotonicity assumption of F, implies that Ω^* is closed and convex (see pp. 158 in [3]).

Among the algorithms for monotone variational inequalities, the extragradient (EG) method proposed by Korpelevich [11] is one of the attractive methods. In fact, each iteration of the extragradient method can be

---

divided into two steps. The k-th iteration of EG method begins with a given $u^k \in \Omega$, the first step produces a vector $\tilde{u}^k$ via a projection

$$\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)], \tag{1.3a}$$

where $\beta_k > 0$ is selected to satisfy

$$\beta_k \|F(u^k) - F(\tilde{u}^k)\| \leq \nu \|u^k - \tilde{u}^k\|, \quad \nu \in (0, 1). \tag{1.3b}$$

Since $\tilde{u}^k$ is not accepted as the new iterate, for designation convenience, we call it as a *predictor* and β_k is named the *prediction step size*. The second step (correction step) of the k-th iteration updates the new iterate u^{k+1} by

$$u^{k+1} = P_\Omega[u^k - \beta_k F(\tilde{u}^k)], \tag{1.4}$$

where β_k is called the *correction step size*. The sequence $\{u^k\}$ generated by the extragradient method is Fejér monotone with respect to the solution set, namely,

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - (1 - \nu^2)\|u^k - \tilde{u}^k\|^2. \tag{1.5}$$

For a proof of the above contraction property, the readers may consult [3] (see pp. 1115-1118 therein). Notice that, in the extragradient method, the step size of the prediction (1.3a) and that of the correction (1.4) are equal. Thus the two steps seem like 'symmetric'.

Because of its simple iterative forms, recently, the extragradient method has been applied to solve some large optimization problems in the area of information science, such as in machine learning [12, 13, 20, 21], optical network [15, 16] and speech recognition [17], etc. In addition, Nemirovski [14] and Tseng [22] proved the $O(1/t)$ convergence rate of the extragradient method. Both in the theoretical and practical aspects, the interest in the extragradient method becomes more active.

In the last decades, we devoted our effort to develop a class of projection and contraction (PC) methods for monotone variational inequalities [5, 6, 10, 18]. Similarly as in the extragradient method, each iteration of the PC methods consists of two steps. The prediction step of PC methods produces the predictor $\tilde{u}^k$ via (1.3) just as in the extragradient method. The PC methods exploit a pair of geminate directions [9, 10] offered by the predictor, namely, they are

$$d(u^k, \tilde{u}^k) = (u^k - \tilde{u}^k) - \beta_k(F(u^k) - F(\tilde{u}^k)) \qquad \text{and} \qquad \beta_k F(\tilde{u}^k). \tag{1.6}$$

Here, both the directions are ascent directions of the unknown distance function $\frac{1}{2}\|u - u^*\|^2$ at the point u^k. Based on such directions, the goal of the correction step is to generate a new iterate which is more closed to the solution set. It leads to choosing the 'optimal' step length

$$\varrho_k = \frac{(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k)}{\|d(u^k, \tilde{u}^k)\|^2}, \tag{1.7}$$

and a relaxation factor $\gamma \in (0, 2)$, the second step (*correction step*) of the PC methods updates the new iterate u^{k+1} by

$$u^{k+1} = u^k - \gamma \varrho_k d(u^k, \tilde{u}^k), \tag{1.8}$$

or

$$u^{k+1} = P_\Omega[u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k)]. \tag{1.9}$$

The PC methods (without line search) make one (or two) projection(s) on Ω at each iteration, and the distance of the iterates to the solution set monotonically converges to zero. According to the terminology in [2], these methods belong to the class of Fejér contraction methods. In fact, the only difference between the extragradient method and one of the PC methods is that they use different step sizes in the correction step (see (1.4) and (1.9)). According to our numerical experiments [6, 10], the PC methods always outperform the extragradient methods.

Stimulated by the complexity statement of the extragradient method, this paper shows the $O(1/t)$ convergence rate of the projection and contraction methods for monotone VIs. Recall that Ω^* can be characterized as (see (2.3.2) in pp. 159 of [3])

$$\Omega^* = \bigcap_{u \in \Omega} \{\tilde{u} \in \Omega : (u - \tilde{u})^T F(u) \geq 0\}.$$

This implies that $\tilde{u} \in \Omega$ is an approximate solution of $\text{VI}(\Omega, F)$ with the accuracy ϵ if it satisfies

$$\tilde{u} \in \Omega \quad \text{and} \quad \inf_{u \in \Omega} \{(u - \tilde{u})^T F(u)\} \geq -\epsilon.$$

In this paper, we show that, for given $\epsilon > 0$ and $\mathcal{D} \subset \Omega$, in $O(L/\epsilon)$ iterations the projection and contraction methods can find a $\tilde{u}$ such that

$$\tilde{u} \in \Omega \quad \text{and} \quad \sup_{u \in \mathcal{D}} \{(\tilde{u} - u)^T F(u)\} \leq \epsilon. \tag{1.10}$$

As a byproduct of the complexity analysis, we find why taking a suitable relaxation factor $\gamma \in (1, 2)$ in the correction steps (1.8) and (1.9) of the PC methods can achieve the faster convergence.

The outline of this paper is as follows. Section 2 recalls some basic concepts in the projection and contraction methods. In Section 3, we investigate the geminate descent directions of the distance function. Section 4 shows the contraction property of the PC methods. In Section 5, we carry out the complexity analysis, which results in an $O(1/t)$ convergence rate and suggests using the large relaxation factor in the correction step of the PC methods. In Section 6, we present some numerical results to indicate the efficiency of the PC methods in comparison with the extragradient method. Finally, some conclusion remarks are addressed in the last section.

Throughout the paper, the following notational conventions are used. We use u^* to denote a fixed but arbitrary point in the solution set Ω^*. A superscript such as in u^k refers to a specific vector and usually denotes an iteration index. For any real matrix M and vector v, we denote the transpose by M^T and v^T, respectively. The Euclidean norm will be denoted by $\|\cdot\|$.

2 Preliminaries

In this section, we summarize the basic concepts of the projection mapping and three fundamental inequalities for constructing the PC methods. Throughout this paper, we assume that the projection on Ω in the Euclidean-norm has a closed form and it is easy to be carried out. Since

$$P_\Omega(v) = \text{argmin}\{\frac{1}{2}\|u - v\|^2 \mid u \in \Omega\},$$

according to the optimal solution of the convex minimization problem, we have

$$(v - P_\Omega(v))^T(u - P_\Omega(v)) \leq 0, \quad \forall \, v \in \Re^n, \forall \, u \in \Omega. \tag{2.1}$$

Consequently, for any $u \in \Omega$, it follows from (2.1) that

$$
\begin{aligned}
\|u - v\|^2 &= \|(u - P_\Omega(v)) - (v - P_\Omega(v))\|^2 \\
&= \|u - P_\Omega(v)\|^2 - 2(v - P_\Omega(v))^T(u - P_\Omega(v)) + \|v - P_\Omega(v)\|^2 \\
&\geq \|u - P_\Omega(v)\|^2 + \|v - P_\Omega(v)\|^2.
\end{aligned}
$$

Therefore, we have

$$\|u - P_\Omega(v)\|^2 \leq \|u - v\|^2 - \|v - P_\Omega(v)\|^2, \quad \forall \, v \in \Re^n, \forall \, u \in \Omega. \tag{2.2}$$

For given u and $\beta > 0$, let $\tilde{u} = P_\Omega[u - \beta F(u)]$ be given via a projection. We say that $\tilde{u}$ is a test-vector of $\text{VI}(\Omega, F)$ because

$$u = \tilde{u} \quad \Leftrightarrow \quad u \in \Omega^*.$$

3

Since $\tilde{u} \in \Omega$, it follows from (1.1) that

$$(\textbf{FI-1}) \qquad (\tilde{u} - u^*)^T \beta F(u^*) \geq 0, \quad \forall\, u^* \in \Omega^*. \tag{2.3}$$

Setting $v = u - \beta F(u)$ and $u = u^*$ in the inequality (2.1), we obtain

$$(\textbf{FI-2}) \qquad (\tilde{u} - u^*)^T \big((u - \tilde{u}) - \beta F(u)\big) \geq 0, \quad \forall\, u^* \in \Omega^*. \tag{2.4}$$

Under the assumption that F is monotone we have

$$(\textbf{FI-3}) \qquad (\tilde{u} - u^*)^T \beta\big(F(\tilde{u}) - F(u^*)\big) \geq 0, \quad \forall\, u^* \in \Omega^*. \tag{2.5}$$

The inequalities (2.3), (2.4) and (2.5) play an important role in the projection and contraction methods. They were emphasized in [5] as *three fundamental inequalities* in the projection and contraction methods.

3 Predictor and the ascent directions

For given u^k, the predictor $\tilde{u}^k$ in the projection and contraction methods [5, 6, 10, 18] is produced by (1.3). Because the mapping F is Lipschitz continuous (even if the constant $L > 0$ is unknown), without loss of generality, we can assume that $\inf_{k \geq 0}\{\beta_k\} \geq \beta_L > 0$ and $\beta_L = O(1/L)$. In practical computation, we can make an initial guesses of $\beta = \nu/L$ and decrease β by a constant factor and repeat the procedure whenever (1.3b) is violated.

For any but fixed $u^* \in \Omega^*$, $(u - u^*)$ is the gradient of the unknown distance function $\frac{1}{2}\|u - u^*\|^2$ in the Euclidean-norm[2] at the point u. A direction d is called an ascent direction of $\frac{1}{2}\|u - u^*\|^2$ at u if and only if the inner-product $(u - u^*)^T d > 0$.

3.1 Ascent directions by adding the fundamental inequalities

Setting $u = u^k$, $\tilde{u} = \tilde{u}^k$ and $\beta = \beta_k$ in the fundamental inequalities (2.3), (2.4) and (2.5), and adding them, we get

$$(\tilde{u}^k - u^*)^T d(u^k, \tilde{u}^k) \geq 0, \quad \forall\, u^* \in \Omega^*, \tag{3.1}$$

where

$$d(u^k, \tilde{u}^k) = (u^k - \tilde{u}^k) - \beta_k\big(F(u^k) - F(\tilde{u}^k)\big), \tag{3.2}$$

which is the same $d(u^k, \tilde{u}^k)$ defined in (1.6). It follows from (3.1) that

$$(u^k - u^*)^T d(u^k, \tilde{u}^k) \geq (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k). \tag{3.3}$$

Note that, under the condition (1.3b), we have

$$\begin{aligned}
&2(u^k - \tilde{u}^k)d(u^k, \tilde{u}^k) - \|d(u^k, \tilde{u}^k)\|^2 \\
&= \; d(u^k, \tilde{u}^k)^T\{2(u^k - \tilde{u}^k) - d(u^k, \tilde{u}^k)\} \\
&= \; \{(u^k - \tilde{u}^k) - \beta_k(F(u^k) - F(\tilde{u}^k))\}^T\{(u^k - \tilde{u}^k) + \beta_k(F(u^k) - F(\tilde{u}^k))\} \\
&= \; \|u^k - \tilde{u}^k\|^2 - \beta_k^2\|F(u^k) - F(\tilde{u}^k)\|^2 \\
&\geq \; (1 - \nu^2)\|u^k - \tilde{u}^k\|^2.
\end{aligned} \tag{3.4}$$

Consequently, from (3.3) and (3.4) we have

$$(u^k - u^*)^T d(u^k, \tilde{u}^k) \geq \frac{1}{2}\big(\|d(u^k, \tilde{u}^k)\|^2 + (1 - \nu^2)\|u^k - \tilde{u}^k\|^2\big).$$

This means that $d(u^k, \tilde{u}^k)$ is an ascent direction of the unknown distance function $\frac{1}{2}\|u - u^*\|^2$ at the point u^k.

---

[2] For convenience, we only consider the distance function in the Euclidean-norm. All the results in this paper are easy to extended to the contraction of the distance function in G-norm where G is a positive definite matrix.

3.2 Geminate ascent directions

To the direction $d(u^k, \tilde{u}^k)$ defined in (3.2), there is a correlative ascent direction $\beta_k F(\tilde{u}^k)$. Use the notation of $d(u^k, \tilde{u}^k)$, the projection equation (1.3a) can be written as

$$\tilde{u}^k = P_\Omega\{\tilde{u}^k - [\beta_k F(\tilde{u}^k) - d(u^k, \tilde{u}^k)]\}. \tag{3.5a}$$

It follows that $\tilde{u}^k$ is a solution of $\mathrm{VI}(\Omega, F)$ if and only if $d(u^k, \tilde{u}^k) = 0$. Assume that there is a constant $c > 0$ such that

$$\varrho_k = \frac{(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k)}{\|d(u^k, \tilde{u}^k)\|^2} \geq c, \quad \forall k \geq 0. \tag{3.5b}$$

In this paper, we call (3.5) with $c > 0$ *the general conditions* and the forthcoming analysis is based of these conditions. For given u^k, there are different ways to construct $\tilde{u}^k$ and $d(u^k, \tilde{u}^k)$ which satisfy the conditions (3.5) (see [10] for an example). If β_k satisfies (1.3b) and $d(u^k, \tilde{u}^k)$ is given by (3.2), the general conditions (3.5) are satisfied with $c \geq \frac{1}{2}$ (see (3.4)). Note that an equivalent expression of (3.5a) is

$$\tilde{u}^k \in \Omega, \quad (u - \tilde{u}^k)^T\{\beta_k F(\tilde{u}^k) - d(u^k, \tilde{u}^k)\} \geq 0, \quad \forall u \in \Omega, \tag{3.6a}$$

and from (3.5b) we have

$$(u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) = \varrho_k \|d(u^k, \tilde{u}^k)\|^2. \tag{3.6b}$$

In [7, 8], $d(u^k, \tilde{u}^k)$ and $\beta_k F(\tilde{u}^k)$ in (3.5a) are called a pair of geminate directions and denoted by $d_1(u^k, \tilde{u}^k)$ and $d_2(u^k, \tilde{u}^k)$, respectively. In this paper, we restrict $d_2(u^k, \tilde{u}^k)$ to be $F(\tilde{u}^k)$ times a positive scalar β_k. If $d(u^k, \tilde{u}^k) = u^k - \tilde{u}^k$, then $\tilde{u}^k$ in (3.6a) is the solution of the subproblem in the k-th iteration when PPA applied to solve $\mathrm{VI}(\Omega, F)$. Hence, the projection and contraction methods considered in this paper belong to the *prox-like contraction methods* [7, 8].

The following lemmas tell us that both the direction $d(u^k, \tilde{u}^k)$ (for $u^k \in \Re^n$) and $F(\tilde{u}^k)$ (for $u^k \in \Omega$) are ascent directions of the function $\frac{1}{2}\|u - u^*\|^2$ whenever u^k is not a solution point. The proof is similar to those in [9], for completeness sake of this paper, we restate the short proofs.

Lemma 3.1 *Let the general conditions* (3.5) *be satisfied. Then we have*

$$(u^k - u^*)^T d(u^k, \tilde{u}^k) \geq \varrho_k \|d(u^k, \tilde{u}^k)\|^2, \quad \forall u^k \in \Re^n, u^* \in \Omega^*. \tag{3.7}$$

Proof. Note that $u^* \in \Omega$. By setting $u = u^*$ in (3.6a) (the equivalent expression of (3.5a)), we get

$$(\tilde{u}^k - u^*)^T d(u^k, \tilde{u}^k) \geq (\tilde{u}^k - u^*)^T \beta_k F(\tilde{u}^k) \geq 0, \quad \forall u^* \in \Omega^*.$$

The last inequality follows from the monotonicity of F and $(\tilde{u}^k - u^*)^T F(u^*) \geq 0$. Therefore,

$$(u^k - u^*)^T d(u^k, \tilde{u}^k) \geq (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k), \quad \forall u^* \in \Omega^*.$$

The assertion (3.7) is followed from the above inequality and (3.6b) directly. □

Lemma 3.2 *Let the general conditions* (3.5) *be satisfied. If* $u^k \in \Omega$, *then we have*

$$(u^k - u^*)^T \beta_k F(\tilde{u}^k) \geq \varrho_k \|d(u^k, \tilde{u}^k)\|^2, \quad \forall u^* \in \Omega^*. \tag{3.8}$$

Proof. Since $(\tilde{u}^k - u^*)^T \beta_k F(\tilde{u}^k) \geq 0$, we have

$$(u^k - u^*)^T \beta_k F(\tilde{u}^k) \geq (u^k - \tilde{u}^k)^T \beta_k F(\tilde{u}^k), \quad \forall u^* \in \Omega^*.$$

Note that because $u^k \in \Omega$, by setting $u = u^k$ in (3.6a), we get

$$(u^k - \tilde{u}^k)^T \beta_k F(\tilde{u}^k) \geq (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k).$$

From the above two inequalities follows that

$$(u^k - u^*)^T \beta_k F(\tilde{u}^k) \geq (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k), \quad \forall u^* \in \Omega^*.$$

The assertion (3.8) is followed from the above inequality and (3.6b) directly. □

Note that (3.7) holds for $u^k \in \Re^n$ while (3.8) is hold only for $u^k \in \Omega$.

4 Corrector and the convergence in the contraction sense

Based on the pair of geminate ascent directions in (3.5), namely, $d(u^k, \tilde{u}^k)$ and $\beta_k F(\tilde{u}^k)$, we use the one of the following corrector forms to update the new iterate u^{k+1}:

$$\text{(Correction of PC Method-I)} \qquad u_I^{k+1} = u^k - \gamma \varrho_k d(u^k, \tilde{u}^k), \tag{4.1a}$$

or

$$\text{(Correction of PC Method-II)} \qquad u_{II}^{k+1} = P_\Omega[u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k)], \tag{4.1b}$$

where $\gamma \in (0, 2)$ and ϱ_k is defined in (3.5b). Note that the same step size length is used in (4.1a) and (4.1b) even if the search directions are different. Recall that $\tilde{u}^k$ is obtained via a projection, by using the correction form (4.1b), we have to make an additional projection on Ω in the PC methods. Replacing $\gamma \varrho_k$ in (4.1b) by 1, it reduces to the update form of the extragradient method (see (1.4)).

For any solution point $u^* \in \Omega^*$, we define

$$\vartheta_I(\gamma) = \|u^k - u^*\|^2 - \|u_I^{k+1} - u^*\|^2 \tag{4.2a}$$

and

$$\vartheta_{II}(\gamma) = \|u^k - u^*\|^2 - \|u_{II}^{k+1} - u^*\|^2, \tag{4.2b}$$

which measure the profit in the k-th iteration. The following theorem gives a lower bound of the profit function, the similar results were established in [6, 9, 10].

Theorem 4.1 *For given u^k, let the general conditions (3.5) be satisfied. If the corrector is updated by (4.1a) or (4.1b), then for any $u^* \in \Omega^*$ and $\gamma > 0$, we have*

$$\vartheta_I(\gamma) \geq q(\gamma), \tag{4.3}$$

and

$$\vartheta_{II}(\gamma) \geq q(\gamma) + \|u_I^{k+1} - u_{II}^{k+1}\|^2, \tag{4.4}$$

respectively, where

$$q(\gamma) = \gamma(2 - \gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2. \tag{4.5}$$

Proof. Using the definition of $\vartheta_I(\gamma)$ and u_I^{k+1} (see (4.1a)), we have

$$\begin{aligned} \vartheta_I(\gamma) &= \|u^k - u^*\|^2 - \|u^k - u^* - \gamma \varrho_k d(u^k, \tilde{u}^k)\|^2 \\ &= 2\gamma \varrho_k (u^k - u^*)^T d(u^k, \tilde{u}^k) - \gamma^2 \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2. \end{aligned} \tag{4.6}$$

Recalling (3.7), we obtain

$$2\gamma \varrho_k (u^k - u^*)^T d(u^k, \tilde{u}^k) \geq 2\gamma \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2.$$

Substituting it in (4.6) and using the definition of $q(\gamma)$, we get $\vartheta_I(\gamma) \geq q(\gamma)$ and the first assertion is proved. Now, we turn to show the second assertion. Because

$$u_{II}^{k+1} = P_\Omega[u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k)],$$

and $u^* \in \Omega$, by setting $u = u^*$ and $v = u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k)$ in (2.2), we have

$$\|u^* - u_{II}^{k+1}\|^2 \leq \|u^* - (u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k))\|^2 - \|u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k) - u_{II}^{k+1}\|^2. \tag{4.7}$$

Thus,

$$\begin{aligned} \vartheta_{II}(\gamma) &= \|u^k - u^*\|^2 - \|u_{II}^{k+1} - u^*\|^2 \\ &\geq \|u^k - u^*\|^2 - \|(u^k - u^*) - \gamma \varrho_k \beta_k F(\tilde{u}^k)\|^2 + \|(u^k - u_{II}^{k+1}) - \gamma \varrho_k \beta_k F(\tilde{u}^k)\|^2 \\ &= \|u^k - u_{II}^{k+1}\|^2 + 2\gamma \varrho_k \beta_k (u_{II}^{k+1} - u^*)^T F(\tilde{u}^k) \\ &\geq \|u^k - u_{II}^{k+1}\|^2 + 2\gamma \varrho_k \beta_k (u_{II}^{k+1} - \tilde{u}^k)^T F(\tilde{u}^k). \end{aligned} \tag{4.8}$$

The last inequality in (4.8) follows from $(\tilde{u}^k - u^*)^T F(\tilde{u}^k) \geq 0$. Since $u_{II}^{k+1} \in \Omega$, by setting $u = u_{II}^{k+1}$ in (3.6a), we get

$$(u_{II}^{k+1} - \tilde{u}^k)^T \{\beta_k F(\tilde{u}^k) - d(u^k, \tilde{u}^k)\} \geq 0,$$

and consequently, substituting it in the right hand side of (4.8), we obtain

$$
\begin{aligned}
\vartheta_{II}(\gamma) &\geq \|u^k - u_{II}^{k+1}\|^2 + 2\gamma \varrho_k (u_{II}^{k+1} - \tilde{u}^k)^T d(u^k, \tilde{u}^k) \\
&= \|u^k - u_{II}^{k+1}\|^2 + 2\gamma \varrho_k (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) - 2\gamma \varrho_k (u^k - u_{II}^{k+1})^T d(u^k, \tilde{u}^k). \quad (4.9)
\end{aligned}
$$

To the two crossed term in the right hand side of (4.9), we have (by using (3.6b))

$$2\gamma \varrho_k (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) = 2\gamma \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2,$$

and

$$-2\gamma \varrho_k (u^k - u_{II}^{k+1})^T d(u^k, \tilde{u}^k) = \|(u^k - u_{II}^{k+1}) - \gamma \varrho_k d(u^k, \tilde{u}^k)\|^2 - \|u^k - u_{II}^{k+1}\|^2 - \gamma^2 \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2,$$

respectively. Substituting them in the right hand side of (4.9) and using $u^k - \gamma \varrho_k d(u^k, \tilde{u}^k) = u_I^{k+1}$, we obtain

$$\vartheta_{II}(\gamma) \geq \gamma(2-\gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2 + \|u_I^{k+1} - u_{II}^{k+1}\|^2 = q(\gamma) + \|u_I^{k+1} - u_{II}^{k+1}\|^2, \quad (4.10)$$

and the proof is complete. □

Note that $q(\gamma)$ is a quadratic function of γ, it reaches its maximum at $\gamma^* = 1$. In practice, ϱ_k is the 'optimal' step size in (4.1) and γ is a relaxation factor. Because $q(\gamma)$ is a lower bound of $\vartheta_I(\gamma)$ (resp. $\vartheta_{II}(\gamma)$), the desirable new iterate is updated by (4.1) with $\gamma \in [1, 2)$ and the reason is interpreted in the following figure.
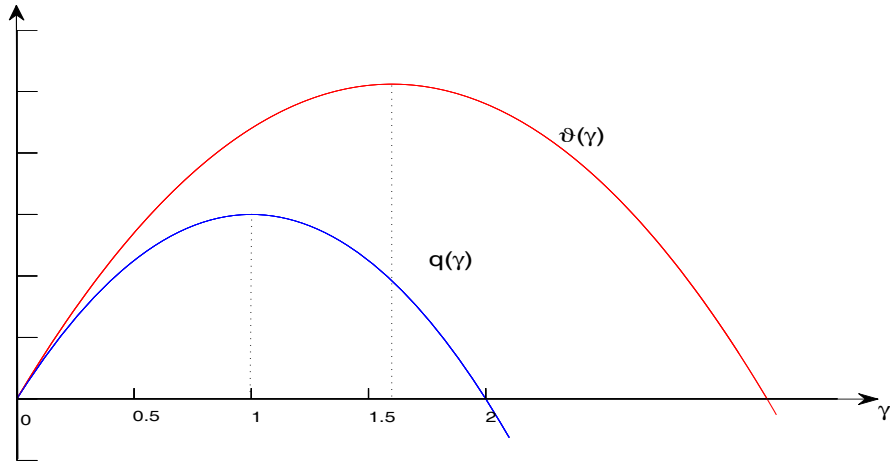
The interpreting of $\gamma \in [1, 2)$ for more profit in each iteration

From Theorem 4.1 we obtain

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \gamma(2-\gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2. \quad (4.11)$$

Convergence result follows from (4.11) directly. Due to the property (4.11) we call the methods which use different update forms in (4.1) PC Method-I and PC Method II, respectively. Note that the assertion (4.11) is derived from the general conditions (3.5). For the PC methods using correction form (1.8) or (1.9), because $\varrho_k > \frac{1}{2}$, by using (3.6b) and (1.3b), it follows from (4.11) that

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \frac{1}{2}\gamma(2-\gamma)(1-\nu)\|u^k - \tilde{u}^k\|^2. \quad (4.12)$$

5   Convergence rate of the PC methods

This section proves the convergence rate of the projection and contraction methods. Recall that the base of the complexity proof is (see (2.3.2) in pp. 159 of [3])

$$\Omega^* = \bigcap_{u \in \Omega} \{\tilde{u} \in \Omega : (u - \tilde{u})^T F(u) \geq 0\}. \tag{5.1}$$

In the sequel, for given $\epsilon > 0$ and $\mathcal{D} \subset \Omega$, we focus our attention to find a $\tilde{u}$ such that

$$\tilde{u} \in \Omega \quad \text{and} \quad \sup_{u \in \mathcal{D}} (\tilde{u} - u)^T F(u) \leq \epsilon. \tag{5.2}$$

Although the PC Method I uses the update form (4.1a) and it does not guarantee that $\{u^k\}$ belongs to Ω, the sequence $\{\tilde{u}^k\} \subset \Omega$ in the PC methods with different corrector forms. Now, we prove the key inequality of the PC Method I for the complexity analysis.

Lemma 5.1 *For given $u^k \in \Re^n$, let the general conditions (3.5) be satisfied. If the new iterate u^{k+1} is updated by (4.1a) with any $\gamma > 0$, then we have*

$$(u - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) + \frac{1}{2} \big(\|u - u^k\|^2 - \|u - u^{k+1}\|^2 \big) \geq \frac{1}{2} q(\gamma), \ \forall u \in \Omega, \tag{5.3}$$

where $q(\gamma)$ is defined in (4.5).

Proof. Because (due to (3.6a))

$$(u - \tilde{u}^k)^T \beta_k F(\tilde{u}^k) \geq (u - \tilde{u}^k)^T d(u^k, \tilde{u}^k), \quad \forall u \in \Omega,$$

and (see (4.1a))

$$\gamma \varrho_k d(u^k, \tilde{u}^k) = u^k - u^{k+1},$$

we need only to show that

$$(u - \tilde{u}^k)^T (u^k - u^{k+1}) + \frac{1}{2} \big(\|u - u^k\|^2 - \|u - u^{k+1}\|^2 \big) \geq \frac{1}{2} q(\gamma), \ \forall u \in \Omega. \tag{5.4}$$

To the crossed term in the left hand side of (5.4), namely $(u - \tilde{u}^k)^T (u^k - u^{k+1})$, using an identity

$$(a - b)^T (c - d) = \frac{1}{2} \big(\|a - d\|^2 - \|a - c\|^2 \big) + \frac{1}{2} \big(\|c - b\|^2 - \|d - b\|^2 \big),$$

we obtain

$$(u - \tilde{u}^k)^T (u^k - u^{k+1}) = \frac{1}{2} \big(\|u - u^{k+1}\|^2 - \|u - u^k\|^2 \big) + \frac{1}{2} \big(\|u^k - \tilde{u}^k\|^2 - \|u^{k+1} - \tilde{u}^k\|^2 \big). \tag{5.5}$$

By using $u^{k+1} = u^k - \gamma \varrho_k d(u^k, \tilde{u}^k)$ and (3.6b), we get

$$
\begin{aligned}
\|u^k - \tilde{u}^k\|^2 - \|u^{k+1} - \tilde{u}^k\|^2 &= \|u^k - \tilde{u}^k\|^2 - \|(u^k - \tilde{u}^k) - \gamma \varrho_k d(u^k, \tilde{u}^k)\|^2 \\
&= 2\gamma \varrho_k (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) - \gamma^2 \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2 \\
&= \gamma(2 - \gamma) \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2.
\end{aligned}
$$

Substituting it in the right hand side of (5.5) and using the definition of $q(\gamma)$, we obtain (5.4) and the lemma is proved.   □

The both sequences $\{\tilde{u}^k\}$ and $\{u^k\}$ in the PC method $\mathbb{II}$ belong to Ω. In the following lemma we prove the same assertion for PC method $\mathbb{II}$ as in Lemma 5.1.

Lemma 5.2 *For given $u^k \in \Omega$, let the general conditions* (3.5) *be satisfied. If the new iterate u^{k+1} is updated by* (4.1b) *with any $\gamma > 0$, then we have*

$$(u - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) + \frac{1}{2}(\|u - u^k\|^2 - \|u - u^{k+1}\|^2) \geq \frac{1}{2}q(\gamma), \ \ \forall u \in \Omega, \tag{5.6}$$

where $q(\gamma)$ is defined in (4.5).

Proof. For investigating $(u - \tilde{u}^k)^T \beta_k F(\tilde{u}^k)$, we divide it in the terms

$$(u^{k+1} - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) \qquad \text{and} \qquad (u - u^{k+1})^T \gamma \varrho_k \beta_k F(\tilde{u}^k).$$

First, we deal with the term $(u^{k+1} - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k)$. Since $u^{k+1} \in \Omega$, substituting $u = u^{k+1}$ in (3.6a) we get

$$
\begin{aligned}
(u^{k+1} - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) &\geq \gamma \varrho_k (u^{k+1} - \tilde{u}^k)^T d(u^k, \tilde{u}^k) \\
&= \gamma \varrho_k (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) - \gamma \varrho_k (u^k - u^{k+1})^T d(u^k, \tilde{u}^k).
\end{aligned} \tag{5.7}
$$

To the first crossed term of the right hand side of (5.7), using (3.6b), we have

$$\gamma \varrho_k (u^k - \tilde{u}^k)^T d(u^k, \tilde{u}^k) = \gamma \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2.$$

To the second crossed term of the right hand side of (5.7), using the Cauchy-Schwarz Inequality, we get

$$-\gamma \varrho_k (u^k - u^{k+1})^T d(u^k, \tilde{u}^k) \geq -\frac{1}{2}\|u^k - u^{k+1}\|^2 - \frac{1}{2}\gamma^2 \varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2.$$

Substituting them in the right hand side of (5.7), we obtain

$$(u^{k+1} - \tilde{u}^k)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) \geq \frac{1}{2}\gamma(2 - \gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2 - \frac{1}{2}\|u^k - u^{k+1}\|^2. \tag{5.8}$$

Now, we turn to treat of the term $(u - u^{k+1})^T \gamma \varrho_k \beta_k F(\tilde{u}^k)$. Since u^{k+1} is updated by (4.1b), u^{k+1} is the projection of $(u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k))$ on Ω, it follows from (2.1) that

$$\left\{ \left(u^k - \gamma \varrho_k \beta_k F(\tilde{u}^k) \right) - u^{k+1} \right\}^T \left(u - u^{k+1} \right) \leq 0, \quad \forall u \in \Omega,$$

and consequently

$$\left(u - u^{k+1} \right)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) \geq \left(u - u^{k+1} \right)^T \left(u^k - u^{k+1} \right), \quad \forall u \in \Omega.$$

Using the identity $a^T b = \frac{1}{2}\{\|a\|^2 - \|a - b\|^2 + \|b\|^2\}$ to the right hand side of the last inequality, we obtain

$$\left(u - u^{k+1} \right)^T \gamma \varrho_k \beta_k F(\tilde{u}^k) \geq \frac{1}{2}(\|u - u^{k+1}\|^2 - \|u - u^k\|^2) + \frac{1}{2}\|u^k - u^{k+1}\|^2. \tag{5.9}$$

Adding (5.8) and (5.9) and using the definition of $q(\gamma)$, we get (5.6) and the proof is complete. □

For the different projection and contraction methods, we have the same key inequality which is shown in Lemma 5.1 and Lemma 5.2, respectively. By setting $u = u^*$ in (5.3) and (5.6), we get

$$\|u^k - u^*\|^2 - \|u^{k+1} - u^*\|^2 \geq 2\gamma \varrho_k \beta_k (\tilde{u}^k - u^*)^T F(\tilde{u}^k) + q(\gamma).$$

Because $(\tilde{u}^k - u^*)^T F(\tilde{u}^k) \geq (\tilde{u}^k - u^*)^T F(u^*) \geq 0$ and $q(\gamma) = \gamma(2 - \gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2$, it follows from the last inequality that

$$\|u^{k+1} - u^*\|^2 \leq \|u^k - u^*\|^2 - \gamma(2 - \gamma)\varrho_k^2 \|d(u^k, \tilde{u}^k)\|^2.$$

This is just the form (4.11) in Section 4. In other words, the contraction property (4.11) of PC methods is the consequent result of Lemma 5.1 and Lemma 5.2, respectively.

For the convergence rate proof, we allow $\gamma \in (0, 2]$. In this case, we still have $q(\gamma) \geq 0$. By using the monotonicity of F, from (5.3) and (5.6) we get

$$(u - \tilde{u}^k)^T \varrho_k \beta_k F(u) + \frac{1}{2\gamma}\|u - u^k\|^2 \geq \frac{1}{2\gamma}\|u - u^{k+1}\|^2, \ \ \forall u \in \Omega. \tag{5.10}$$

This inequality is essential for the convergence rate proofs.

Theorem 5.1 *For any integer $t > 0$, we have a $\tilde{u}_t \in \Omega$ which satisfies*

$$(\tilde{u}_t - u)^T F(u) \leq \frac{1}{2\gamma \Upsilon_t} \|u - u^0\|^2, \quad \forall u \in \Omega, \tag{5.11}$$

where

$$\tilde{u}_t = \frac{1}{\Upsilon_t} \sum_{k=0}^{t} \varrho_k \beta_k \tilde{u}^k \qquad and \qquad \Upsilon_t = \sum_{k=0}^{t} \varrho_k \beta_k. \tag{5.12}$$

Proof. Summing the inequality (5.10) over $k = 0, \ldots, t$, we obtain

$$\left(\left(\sum_{k=0}^{t} \varrho_k \beta_k \right) u - \sum_{k=0}^{t} \varrho_k \beta_k \tilde{u}^k \right)^T F(u) + \frac{1}{2\gamma} \|u - u^0\|^2 \geq 0, \quad \forall u \in \Omega.$$

Using the notations of Υ_t and $\tilde{u}_t$ in the above inequality, we derive

$$(\tilde{u}_t - u)^T F(u) \leq \frac{\|u - u^0\|^2}{2\gamma \Upsilon_t}, \quad \forall u \in \Omega.$$

Indeed, $\tilde{u}_t \in \Omega$ because it is a convex combination of $\tilde{u}^0, \tilde{u}^1, \ldots, \tilde{u}^t$. The proof is complete. □

For given u^k, the predictor $\tilde{u}^k$ is given by (1.3a) and the prediction step size β_k satisfies the condition (1.3b). Thus, the general conditions (3.5) are satisfied with $\varrho_k \geq c = \frac{1}{2}$. We choose (4.1a) (for the case that u^k is not necessary in Ω) or (4.1b) (for the case that $u^k \in \Omega$) to generate the new iterate u^{k+1}. Because $\varrho_k \geq \frac{1}{2}$, $\inf_{k \geq 0}\{\beta_k\} \geq \beta_L$ and $\beta_L = O(1/L)$, it follows from (5.12) that

$$\Upsilon_t \geq \frac{t+1}{2} \beta_L,$$

and thus the PC methods have $O(1/t)$ convergence rate. For any substantial set $\mathcal{D} \subset \Omega$, the PC methods reach

$$(\tilde{u}_t - u)^T F(u) \leq \epsilon, \quad \forall u \in \mathcal{D}, \qquad \text{in at most} \qquad t = \left\lceil \frac{D^2}{\gamma \beta_L \epsilon} \right\rceil$$

iterations, where $\tilde{u}_t$ is defined in (5.12) and $D = \sup\{\|u - u^0\| \mid u \in \mathcal{D}\}$. This convergence rate is in the ergodic sense, the statement (5.11) suggests us to take a larger parameter $\gamma \in (0, 2]$ in the correction steps of the PC methods.

6 Numerical experiments

This section is devoted to test the efficiency of the PC methods in comparison with the extragradient method [11]. Under the condition (1.3b), we have $\varrho_k > 1/2$. If we dynamically take $\gamma_k = 1/\varrho_k$ in (4.1b), then it becomes

$$u^{k+1} = P_\Omega[u^k - \beta_k F(\tilde{u}^k)], \tag{6.1}$$

which is the update form of the extragradient method [11]. Because $\gamma_k \varrho_k \equiv 1$, it follows from (5.10) that

$$(u - \tilde{u}^k)^T \beta_k F(u) + \frac{1}{2} \|u - u^k\|^2 \geq \frac{1}{2} \|u - u^{k+1}\|^2, \quad \forall u \in \Omega. \tag{6.2}$$

The results in Theorem 5.1 becomes

$$\tilde{u}_t = \frac{1}{\sum_{k=0}^{t} \beta_k} \sum_{k=0}^{t} \beta_k \tilde{u}^k \in \Omega, \qquad (\tilde{u}_t - u)^T F(u) \leq \frac{\|u - u^0\|^2}{2\left(\sum_{k=0}^{t} \beta_k\right)}, \qquad \forall u \in \Omega. \tag{6.3}$$

The $O(1/t)$ convergence rate follows from the above inequality directly. It should be mentioned that the projection-type method for VI(Ω, F) in [18] is a contraction method in the sense of P-norm, where P is a positive definite matrix. In the Euclidean-norm, its update form is (4.1a).

Test examples of nonlinear complementarity problems.

We take nonlinear complementarity problems (NCP) as the test examples. The mapping $F(u)$ in the tested NCP is given by

$$F(u) = D(u) + Mu + q, \tag{6.4}$$

where $D(u) : \Re^n \to \Re^n$ is the nonlinear part, M is an $n \times n$ matrix, and $q \in \Re^n$ is a vector.

- In $D(u)$, the nonlinear part of $F(u)$, the components are

$$D_j(u) = d_j \cdot \arctan(a_j \cdot u_j),$$

where a and d are random vectors[3] whose elements are in $(0, 1)$.

- The matrix M in the linear part is given by $M = A^T A + B$. A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, +5)$, and B is an $n \times n$ skew-symmetric random matrix ($B^T = -B$) whose entries[4] are in the interval $(-5, +5)$.

It is clear that the mapping composed in this way is monotone. We construct the following 3 sets of test examples by choosing different vector q in (6.4).

1. In the first set of test examples, the elements of vector q is generated from a uniform distribution in the interval $(-500, 500)$.

2. The second set[5] of test examples is similar to the first set. Instead of $q \in (-500, 500)$, the vector q is generated from a uniform distribution in the interval $(-500, 0)$.

3. The third set of test examples has a known solution $u^* \in \Re^n_+$. Let vector p be generated from a uniform distribution in the interval $(-10, 10)$ and

$$u^* = \max(p, 0). \tag{6.5}$$

By setting

$$w = \max(-p, 0) \quad \text{and} \quad q = w - (D(u^*) + Mu^*),$$

we have $F(u^*) = D(u^*) + Mu^* + q = w = \max(-p, 0)$. Thus,

$$(u^*)^T F(u^*) = \big(\max(p, 0) \big)^T \big(\max(-p, 0) \big) = 0.$$

In this way we constructed a test NCP with a known solution u^* described in (6.5).

Implementation details.

For given u^k, we use (1.3) to produce $\tilde{u}^k$ with $\nu = 0.9$ in (1.3b). If $r_k := \beta_k \| F(u^k) - F(\tilde{u}^k) \| / \| u^k - \tilde{u}^k \|$ is too small, it will lead slow convergence. Therefore, if $r_k \leq \mu = 0.3$, the trial parameter β_k will be enlarged for the next iteration. These 'refined' strategies are necessary for fast convergence. The following is the implementation details.

Step 0. Set $\beta_0 = 1$, $u^0 \in \Omega$ and $k = 0$.
Step 1. $\tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]$,
$\qquad r_k := \dfrac{\beta_k \| F(u^k) - F(\tilde{u}^k) \|}{\| u^k - \tilde{u}^k \|}$,
$\qquad$ **while** $\quad r_k > \nu$
$\qquad\qquad \beta_k := 0.7 * \beta_k * \min\{1, \frac{1}{r_k}\}, \quad \tilde{u}^k = P_\Omega[u^k - \beta_k F(u^k)]$
$\qquad$ **end(while)**
$\qquad$ Use different forms ((6.1), (4.1a) or (4.1b)) to update u^{k+1}.
$\qquad$ **If** $\quad r_k \leq \mu \quad$ then $\quad \beta_k := \beta_k * \nu * 0.9 / r_k, \quad$ **end(if)**
Step 2. $\quad \beta_{k+1} = \beta_k \quad$ and $\quad k = k + 1, \quad$ go to Step 1.

---

[3]A similar type of (small) problems was tested in [19] where the components of the nonlinear mapping $D(u)$ are $D_j(u) = c \cdot \arctan(u_j)$.

[4]In the paper by Harker and Pang [4], the matrix $M = A^T A + B + D$, where A and B are the same matrices as what we use here, and D is a diagonal matrix with uniformly distributed random entries $d_{jj} \in (0.0, 0.3)$.

[5]In [4], the similar problems in the first set are called easy problems while the 2-nd set problems are called hard problems.

The iterations begin with $u^0 = 0$, $\beta_0 = 1$ and stop as soon as

$$\frac{\|u^k - P_\Omega[u^k - F(u^k)]\|_\infty}{\|u^0 - P_\Omega[u^0 - F(u^0)]\|_\infty} \le 10^{-6}. \tag{6.6}$$

Since both $F(u^k)$ and $F(\tilde{u}^k)$ are involved in those methods recursions, each iteration of the test methods needs at least 2 times of evaluations of the mapping F. We use No. It and No. F to denote the numbers of iterations and the evaluations of the mapping F, respectively. The size of the tested problems is from 500 to 2000. All codes are written in Matlab and run on a Lenovo X200 Computer with 2.53 GHz.

Comparison beteeen the extragradient method and the PC method II.

As mentioned in Section 4, replacing $\gamma\varrho_k$ in (4.1b) by 1, the PC method II becomes the extragradient method. According to the assertion in Theorem 4.1 and Theorem 5.1, we take the relaxation factor $\gamma = 2$ in the PC method II. The test results for the 3 sets of NCP are given in Tables 1-3, respectively.

Table 1. Numerical results of the first set examples

Problem	Extra-Gradient Method $u^{k+1} = P_\Omega[u^k - \beta_k F(\tilde{u}^k)]$			PC Method II $(\gamma = 2)$ $u^{k+1} = P_\Omega[u^k - 2\varrho_k\beta_k F(\tilde{u}^k)]$		
size n	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	496	1032	0.1626	224	490	0.0792
1000	439	917	1.5416	196	430	0.7285
2000	592	1236	7.8440	262	574	3.7305

Table 2. Numerical results of the second set examples

Problem	Extra-Gradient Method $u^{k+1} = P_\Omega[u^k - \beta_k F(\tilde{u}^k)]$			PC Method II $(\gamma = 2)$ $u^{k+1} = P_\Omega[u^k - 2\varrho_k\beta_k F(\tilde{u}^k)]$		
size n	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	1157	2412	0.3921	510	1113	0.1938
1000	1197	2475	4.1946	533	1162	1.9350
2000	1487	3099	19.6591	669	1452	9.3591

Table 3. Numerical results of the third set examples

Problem	Extra-Gradient Method $u^{k+1} = P_\Omega[u^k - \beta_k F(\tilde{u}^k)]$			PC Method II $(\gamma = 2)$ $u^{k+1} = P_\Omega[u^k - 2\varrho_k\beta_k F(\tilde{u}^k)]$		
size n	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	633	1318	0.2109	279	610	0.0988
1000	700	1458	2.4544	308	673	1.1272
2000	789	1643	10.4436	346	756	4.8455

In the third test examples, as the stop criterium is satisfied, we have $\|u^k - u^*\|_\infty \approx 2 \times 10^{-4}$ by using the both test methods. The PC Method II and the extragradient method use the same direction but different step size in the correction step. The numerical results show that the PC method II is much efficient than the extragradient method. Even if the PC methods need to calculate the step size ϱ_k in each iteration, while the computational load required by the additional effort is significantly less than the dominating task (the evaluations of $F(u^k)$ and $F(\tilde{u}^k)$). It is observed that

$$\frac{\text{Computational load of PC Method II}}{\text{Computational load of the extragradient method}} < 50\%.$$

Comparison between PC method I and PC method II.

The different PC methods use the one of the geminate directions but the same step size in their correction forms. In order to ensure $\vartheta_I(\gamma) > 0$, we take $\gamma = 1.9$ in (4.1) for the both update forms. The test results for the 3 sets of NCP are given in Tables 4-6, respectively.

Table 4. Numerical results of the first set examples

Problem size n	PC Method I $(\gamma = 1.9)$ $u^{k+1} = u^k - \gamma\varrho_k d(u^k, \tilde{u}^k)$			PC Method II $(\gamma = 1.9)$ $u^{k+1} = P_\Omega[u^k - \gamma\varrho_k\beta_k F(\tilde{u}^k)]$		
	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	294	625	0.1060	233	507	0.0885
1000	253	546	0.9451	204	445	0.7714
2000	334	704	4.5035	271	591	3.7896

Table 5. Numerical results of the second set examples

Problem size n	PC Method I $(\gamma = 1.9)$ $u^{k+1} = u^k - \gamma\varrho_k d(u^k, \tilde{u}^k)$			PC Method II $(\gamma = 1.9)$ $u^{k+1} = P_\Omega[u^k - \gamma\varrho_k\beta_k F(\tilde{u}^k)]$		
	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	594	1273	0.2192	539	1170	0.2014
1000	635	1345	2.3151	559	1213	2.0908
2000	772	1641	10.4909	701	1518	9.7359

Table 6. Numerical results of the third set examples

Problem size n	PC Method I $(\gamma = 1.9)$ $u^{k+1} = u^k - \gamma\varrho_k d(u^k, \tilde{u}^k)$			PC Method II $(\gamma = 1.9)$ $u^{k+1} = P_\Omega[u^k - \gamma\varrho_k\beta_k F(\tilde{u}^k)]$		
	No. It	No. F	CPU Sec	No. It	No. F	CPU Sec
500	348	741	0.1328	295	642	0.1162
1000	368	782	1.3584	328	713	1.2441
2000	423	900	5.7394	370	803	5.1350

Between the PC methods, PC method II needs fewer iterations than PC method I, this evidence coincides with the assertions in Theorem 4.1 (see (4.3) and (4.4)). Thus, we suggest to use PC method II when the projection on Ω is easy to be carried out. Otherwise (when the projection is the dominating task in the iteration), we use PC method I because its update form (4.1a) does not contain the projection.

7   Conclusions

In a unified framework, we proved the $O(1/t)$ convergence rate of the projection and contraction methods for monotone variational inequalities. The convergence rate is the same as that for the extragradient method. In fact, our convergence rate include the extragradient method as a special case. The complexity analysis in this paper is based on the general conditions (3.5) and thus can be extended to a broaden class of similar contraction methods. Preliminary numerical results indicate that the PC methods do outperform the extragradient method.

References

[1] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation, Numerical Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1989.

[2] E. Blum and W. Oettli, *Mathematische Optimierung: Grundlagen und Verfahren. Ökonometrie und Unternehmensforschung*. Berlin-Heidelberg-New York, Springer-Verlag, 1975.

[3] F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Vol. I and II*. Springer Series in Operations Research. Springer Verlag, New York, 2003.

[4] P.T. Harker and J.S. Pang, A damped-Newton method for the linear complementarity problem, *Lectures in Applied Mathematics* **26**, pp. 265–284, 1990.

[5] B.S. He, A class of projection and contraction methods for monotone variational inequalities, *Applied Mathematics and optimization* **35**, pp. 69–76, 1997.

[6] B.S. He and L.-Z Liao, Improvements of some projection methods for monotone nonlinear variational inequalities, *Journal of Optimization Theory and Applications* **112**, pp. 111-128, 2002.

[7] B.S. He, L.-Z. Liao and X. Wang, Proximal-like contraction methods for monotone variational inequalities in a unified framework I: effective quadruplet and primary methods, *Computational Optimization and Applications*, 2011.

[8] B.S. He, L.-Z. Liao and X. Wang, Proximal-like contraction methods for monotone variational inequalities in a unified framework II: General methods and numerical experiments, *Computational Optimization and Applications*, 2011

[9] B.S. He and M.-H. Xu, A general framework of contraction methods for monotone variational inequalities, *Pacific J. Optimization* **4**, pp. 195-212, 2008.

[10] B.S. He, X.M. Yuan and J.J.Z. Zhang, Comparison of two kinds of prediction-correction methods for monotone variational inequalities, *Computational Optimization and Applications* **27**, pp. 247-267, 2004

[11] G.M. Korpelevich, The extragradient method for finding saddle points and other problems, *Ekonomika i Matematicheskie Metody* **12**, pp. 747-756, 1976.

[12] A.G. Howard, Large Margin, Transformation Learning, PhD Thesis, Graduate School of Arts and Science. Columbia University, 2009.

[13] S. Lacoste-Julien, Discriminative Machine Learning with Structure, PhD Thesis, Computer Science. University of California, Berkeley, 2009.

[14] A. Nemirovski, 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. Optimization* **15**, pp. 229-251, 2005.

[15] Y. Pan, A game theoretical approach to constrained OSNR optimization problems in optical network, PhD Thesis, Electrical and Computer Engineering. University of Toronto, 2009.

[16] Y. Pan and L. Pavel, Games with coupled propagated constraints in optical networks with multi-link topologies, *Automatica* **45**, pp. 871-880, 2009.

[17] F. Sha, Large Margin Training of Acoustic Models for Speech Recognition, PhD Thesis, Computer and Information Science, University of Pennsylvania, 2007.

[18] M.V. Solodov and P. Tseng, Modified projection-type methods for monotone variational inequalities, *SIAM J. Control and Optimization* **34**, pp. 1814–1830, 1996.

[19] K. Taji, M. Fukushima and I. Ibaraki, A globally convergent Newton mmethod for solving strongly monotone variational inequalities. *Mathematical Programming* **58**, pp. 369-383, 1993.

[20] B. Taskar, S. Lacoste-Julien and M. I. Jordan, Structured prediction, dual extragradient and Bregman projections, *Journal of Machine Learning Research* **7**, pp. 1627-1653, 2006.

[21] B. Taskar, S. Lacoste-Julien and M. I. Jordan. Structured prediction via Extragradient method, In Y. Weiss and B. Schoelkopf and J. Platt (Eds.), *Advances in Neural Information Processing Systems (NIPS)* **18**, 2006.

[22] P. Tseng, On accelerated proximal gradient methods for convex-concave optimization, Department of Mathematics, University of Washington, Seattle, WA 98195, USA, 2008.