

RESEARCH

Open Access



A projection descent method for solving variational inequalities

Abdellah Bnouhachem^{1,2}, Qamrul Hasan Ansari^{3,4} and Ching-Feng Wen^{5*}

*Correspondence:

cfwen@kmu.edu.tw

⁵Center for Fundamental Science,
Kaohsiung Medical University,
Kaohsiung, 807, Taiwan

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

Abstract

In this paper, we propose a descent direction method for solving variational inequalities. A new iterate is obtained by searching the optimal step size along a new descent direction which is obtained by the linear combination of two descent directions. Under suitable conditions, the global convergence of the proposed method is studied. Two numerical experiments are presented to illustrate the efficiency of the proposed method.

MSC: 49J40; 65N30

Keywords: variational inequalities; self-adaptive rules; pseudomonotone operators; projection method

1 Introduction

The theory of variational inequalities is a well-established subject in nonlinear analysis and optimization. It was started during early sixties by the pioneering work of Fichera [1] and Stampacchia [2] (see also [3]). Since then it has been extended and studied in several directions. One of the most important aspects of the theory of variational inequalities is the solution method. Several solution methods were proposed and analyzed in the literature (see, for example, [4–10] and the references therein). The fixed point theory plays an important role in developing various kinds of algorithms for solving variational inequalities. By using the projection operator technique, one can easily establish the equivalence between variational inequalities and fixed point problems. This alternative equivalent formulation provides an elegant solution method, known as projection gradient method, for solving variational inequalities. The convergence of this method requires the strong monotonicity and Lipschitz continuity of the underlying operator. Many applications do not have these strong conditions. As a result, the projection gradient method is not applicable for such problems. Korpelevich [11] modified the projection gradient method to overcome these difficulties and introduced the so-called extragradient method. It generates the iterates according to the following recursion:

$$u^{k+1} = P_K[u^k - \rho T(\bar{u}^k)],$$

where

$$\bar{u}^k = P_K[u^k - \rho T(u^k)],$$

and $\rho > 0$ is a fixed parameter. This method overcomes the difficulties arising in the project gradient method by performing an additional forward step and a projection at each iteration according to double projection. It was proved in [11] that the extragradient method is globally convergent if T is monotone and Lipschitz continuous on K and $0 < \rho < 1/L$, where L is the Lipschitz constant of T . When the operator T is not Lipschitz continuous, or it is Lipschitz but the constant L is not known, the fixed parameter ρ must be replaced by the step size computed through an Armijo-like line search procedure with a new projection needed for each trial point (see, e.g., [11, 12]), and this can be computationally very expensive. To overcome these difficulties, several modified projection and extragradient-type methods [12–22] have been suggested and developed for solving variational inequalities. It was shown in [14, 16] that three-step method performs better than two-step and one-step methods for solving variational inequalities.

The aim of this paper is to develop an algorithm, inspired by Fu [16], for solving variational inequalities. More precisely, in the first method, a new iterate is obtained by searching the optimal step size along the integrated descent direction from the linear combination of two descent directions, and another optimal step length is employed to reach substantial progress in each iteration for the second method. It is proved theoretically that the lower-bound of the progress is greater when obtained by the second method than by the first one. Under certain conditions, the global convergence of the proposed methods is proved. Our results can be viewed as significant extensions of the previously known results.

2 Preliminaries

Let $K \subset \mathbb{R}^n$ be a nonempty closed convex set and $T : K \rightarrow \mathbb{R}^n$ be an operator. A classical variational inequality problem, denoted by $VI(T, K)$, is to find a vector $u^* \in K$ such that

$$\langle T(u^*), v - u^* \rangle \geq 0, \quad \forall v \in K. \tag{1}$$

It is worth to mention that the solution set S^* of $VI(T, K)$ is nonempty if T is continuous and K is compact.

It is well known that if K is a closed convex cone, then $VI(T, K)$ is equivalent to the nonlinear complementarity problem of finding $u^* \in K$ such that

$$T(u^*) \in K^* \quad \text{and} \quad \langle T(u^*), u^* \rangle = 0, \tag{2}$$

where $K^* := \{y \in \mathbb{R}^n : \langle y, x \rangle \geq 0 \text{ for all } x \in K\}$. For further details on variational inequalities and complementarity problems, we refer to [4–8] and the references therein.

In what follows, we always assume that the underlying operator T is continuous and pseudomonotone, that is,

$$\langle T(u), v - u \rangle \geq 0 \quad \Rightarrow \quad \langle T(v), v - u \rangle \geq 0, \quad \forall u, v \in K,$$

and the solution set of problem (1), denoted by S^* , is nonempty.

The following results will be used in the sequel.

Lemma 1 [23] *Let $K \subset \mathbb{R}^n$ be a nonempty closed convex set and $P_K(\cdot)$ denote the projection on K under the Euclidean norm, that is, $P_K(z) = \arg \min\{\|z - x\| : x \in K\}$. Then the following statements hold:*

- (a) $\langle z - P_K(z), P_K(z) - v \rangle \geq 0, \forall z \in \mathbb{R}^n, v \in K$.

- (b) $\|P_K(z) - v\|^2 \leq \|z - v\|^2 - \|z - P_K(z)\|^2, \forall z \in \mathbb{R}^n, v \in K.$
- (c) $\|P_K(w) - P_K(v)\|^2 \leq \langle w - v, P_K(w) - P_K(v) \rangle, \forall v, w \in \mathbb{R}^n.$

Lemma 2 [7] u^* is a solution of problem (1) if and only if

$$u^* = P_K[u^* - \rho T(u^*)], \quad \text{where } \rho > 0. \tag{3}$$

From Lemma 2, it is clear that u is a solution of (1) if and only if u is a zero point of the function

$$e(u, \rho) := u - P_K[u - \rho T(u)].$$

The next lemma shows that $\|e(u, \rho)\|$ is a non-decreasing function, while $\frac{\|e(u, \rho)\|}{\rho}$ is a non-increasing one with respect to ρ .

Lemma 3 [24–26] For all $u \in \mathbb{R}^n$ and $\rho' > \rho > 0$,

$$\|e(u, \rho')\| \geq \|e(u, \rho)\| \tag{4}$$

and

$$\frac{\|e(u, \rho')\|}{\rho'} \leq \frac{\|e(u, \rho)\|}{\rho}. \tag{5}$$

3 Algorithm and convergence results

In this section, we suggest and analyze two new methods for solving variational inequalities (1). For given $u^k \in K$ and $\rho_k > 0$, each iteration of the first method consists of three steps, the first step offers \tilde{u}^k , the second step makes \bar{u}^k and the third step produces the new iterate u^{k+1} .

Algorithm 1

- Step 1. Given $u^0 \in K, \epsilon > 0, \rho_0 = 1, \nu > 1, \mu \in (0, \sqrt{2}), \sigma \in (0, 1), \zeta \in (0, 1), \eta_1 \in (0, \zeta), \eta_2 \in (\zeta, \nu)$ and let $k = 0$.
- Step 2. If $\|e(u^k, \rho_k)\| \leq \epsilon$, then stop. Otherwise, go to Step 3.
- Step 3. (1) For a given $u^k \in K$, calculate the two predictors

$$\tilde{u}^k = P_K[u^k - \rho_k T(u^k)], \tag{6a}$$

$$\bar{u}^k = P_K[\tilde{u}^k - \rho_k T(\tilde{u}^k)]. \tag{6b}$$

- (2) If $\|e(\tilde{u}^k, \rho_k)\| \leq \epsilon$, then stop. Otherwise, continue.
- (3) If ρ_k satisfies both

$$r_1 := \frac{|\rho_k[\langle \tilde{u}^k - \bar{u}^k, T(u^k) - T(\tilde{u}^k) \rangle - \langle u^k - \bar{u}^k, T(\tilde{u}^k) - T(\bar{u}^k) \rangle]|}{\|\tilde{u}^k - \bar{u}^k\|^2} \leq \mu^2 \tag{7}$$

and

$$r_2 := \frac{\|\rho_k(T(\tilde{u}^k) - T(\bar{u}^k))\|}{\|\tilde{u}^k - \bar{u}^k\|} \leq \nu, \tag{8}$$

then go to Step 4; otherwise, continue.

(4) Perform an Armijo-like line search via reducing ρ_k ,

$$\rho_k := \rho_k * \frac{\sigma}{\max(r_1, 1)} \tag{9}$$

and go to Step 3.

Step 4. Take the new iteration u^{k+1} by setting

$$u^{k+1}(\alpha_k) = P_K[u^k - \alpha_k d(\tilde{u}^k, \bar{u}^k)], \tag{10}$$

where

$$d(\tilde{u}^k, \bar{u}^k) = \beta_1 d_1(\tilde{u}^k, \bar{u}^k) + \beta_2 \rho_k F(\bar{u}^k) \tag{11}$$

and

$$d_1(\tilde{u}^k, \bar{u}^k) := (\tilde{u}^k - \bar{u}^k) - \rho_k (T(\tilde{u}^k) - T(\bar{u}^k)). \tag{12}$$

Step 5. Adaptive rule of choosing a suitable ρ_{k+1} as the start prediction step size for the next iteration.

(1) Prepare a proper ρ_{k+1} ,

$$\rho_{k+1} := \begin{cases} \rho_k * \zeta / r_2 & \text{if } r_2 \leq \eta_1, \\ \rho_k * \zeta / r_2 & \text{if } r_2 \geq \eta_2, \\ \rho_k & \text{otherwise.} \end{cases}$$

(2) Return to Step 2, with k replaced by $k + 1$.

Remark 1 (a) The proposed method can be viewed as a refinement and improvement of the method of He *et al.* [27] by performing an additional projection step at each iteration and another optimal step length is employed to reach substantial progress in each iteration.

(b) If $\beta_1 = 0$ and $\beta_2 = 1$, we obtain the method proposed in [16].

(c) If $\beta_1 = 1$ and $\beta_2 = 0$, we obtain a descent method, the new iterate is obtained along a descent direction $d_1(\tilde{u}^k, \bar{u}^k)$.

Remark 2 In (9), if $\sigma > \max(r_1, 1)$, it indicates that ρ_k will be too large for the next iteration and will increase the number of Armijo-like line searches. So, we choose ρ_k for the next iteration to be only modestly smaller than ρ_k to avoid expensive implementations in the next iteration.

We now consider the criteria for α_k , which ensures that $u^{k+1}(\alpha_k)$ is closer to the solution set than u^k . For this purpose, we define

$$\Theta_k(\alpha_k) := \|u^k - u^*\|^2 - \|u^{k+1}(\alpha_k) - u^*\|^2. \tag{13}$$

Theorem 1 Let $u^* \in S^*$. Then we have

$$\Theta(\alpha_k) \geq \Phi(\alpha_k) \geq \Psi(\alpha_k), \tag{14}$$

where

$$\Phi(\alpha_k) = \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k \langle u^{k+1}(\alpha_k) - \bar{u}^k, d(\tilde{u}^k, \bar{u}^k) \rangle, \tag{15}$$

and

$$\Psi(\alpha_k) = 2\alpha_k(\beta_1 + \beta_2) \langle u^k - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle - \alpha_k^2(\beta_1 + \beta_2)^2 \|d_1(\tilde{u}^k, \bar{u}^k)\|^2.$$

Proof Since $u^* \in K$, setting $v = u^*$ and $z = u^k - \alpha_k d(\tilde{u}^k, \bar{u}^k)$ in (b), we have

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

Using the definition of $\Theta(\alpha_k)$, we get

$$\begin{aligned} \Theta(\alpha_k) &\geq \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k \langle u^k - u^*, d(\tilde{u}^k, \bar{u}^k) \rangle \\ &\quad - 2\alpha_k \langle u^k - u^{k+1}(\alpha_k), d(\tilde{u}^k, \bar{u}^k) \rangle \\ &= \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k \langle u^{k+1}(\alpha_k) - u^*, d(\tilde{u}^k, \bar{u}^k) \rangle. \end{aligned} \tag{16}$$

For any solution $u^* \in S^*$ of problem (1), we have

$$\langle \rho_k T(u^*), \bar{u}^k - u^* \rangle \geq 0.$$

By the pseudomonotonicity of T , we obtain

$$\langle \rho_k T(\bar{u}^k), \bar{u}^k - u^* \rangle \geq 0. \tag{17}$$

Substituting $z = \tilde{u}^k - \rho_k T(\bar{u}^k)$ and $v = u^*$ into (a), we get

$$\langle \tilde{u}^k - \rho_k T(\bar{u}^k) - \bar{u}^k, \bar{u}^k - u^* \rangle \geq 0. \tag{18}$$

Adding (17) and (18), and using the definition of $d_1(\tilde{u}^k, \bar{u}^k)$, we have

$$\langle d_1(\tilde{u}^k, \bar{u}^k), \bar{u}^k - u^* \rangle \geq 0. \tag{19}$$

Multiplying (19) by $2\alpha_k\beta_1$ and (17) by $2\alpha_k\beta_2$, then adding the resultants with (16), we obtain

$$\begin{aligned} \Theta(\alpha_k) &\geq \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k \langle u^{k+1}(\alpha_k) - \bar{u}^k, d(\tilde{u}^k, \bar{u}^k) \rangle \\ &= \Phi(\alpha_k) \\ &= \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k\beta_1 \langle u^{k+1}(\alpha_k) - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &\quad + 2\alpha_k\beta_2 \langle u^{k+1}(\alpha_k) - \bar{u}^k, \rho_k T(\bar{u}^k) \rangle. \end{aligned} \tag{20}$$

Note that $\tilde{u}^k = P_K[\bar{u}^k - \rho_k T(\bar{u}^k)]$. We can apply (a) with $v = u^{k+1}$, to obtain

$$\langle u^{k+1}(\alpha_k) - \bar{u}^k, \tilde{u}^k - \rho_k T(\bar{u}^k) - \bar{u}^k \rangle \leq 0. \tag{21}$$

Multiplying (21) by $2\alpha_k\beta_2$ and then adding the resultant to (20) and using the definition of $d_1(\tilde{u}^k, \bar{u}^k)$, we have

$$\begin{aligned} \Phi(\alpha_k) &\geq \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k\beta_1\langle u^{k+1}(\alpha_k) - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &\quad + 2\alpha_k\beta_2\langle u^{k+1}(\alpha_k) - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &= \|u^k - u^{k+1}(\alpha_k)\|^2 + 2\alpha_k(\beta_1 + \beta_2)\langle u^{k+1}(\alpha_k) - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &= \|u^k - u^{k+1}(\alpha_k) - \alpha_k(\beta_1 + \beta_2)d_1(\tilde{u}^k, \bar{u}^k)\|^2 \\ &\quad + 2\alpha_k(\beta_1 + \beta_2)\langle u^k - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &\quad - \alpha_k^2(\beta_1 + \beta_2)^2\|d_1(\tilde{u}^k, \bar{u}^k)\|^2 \\ &\geq 2\alpha_k(\beta_1 + \beta_2)\langle u^k - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle - \alpha_k^2(\beta_1 + \beta_2)^2\|d_1(\tilde{u}^k, \bar{u}^k)\|^2 \\ &= \Psi(\alpha_k), \end{aligned}$$

and the theorem is proved. □

Proposition 1 *Assume that T is continuously differentiable, then we have*

- (i) $\Phi'(\alpha) = 2\langle u^{k+1}(\alpha) - \bar{u}^k, d(\tilde{u}^k, \bar{u}^k) \rangle$;
- (ii) $\Phi'(\alpha)$ is a non-increasing function with respect to $\alpha \geq 0$, and hence, $\Phi(\alpha)$ is concave.

Proof For given $u^k, \tilde{u}^k, \bar{u}^k \in K$, let

$$h(\alpha, y) = \|y - [u^k - \alpha d(\tilde{u}^k, \bar{u}^k)]\|^2 - \alpha^2\|d(\tilde{u}^k, \bar{u}^k)\|^2 - 2\alpha\langle \bar{u}^k - u^k, d(\tilde{u}^k, \bar{u}^k) \rangle. \tag{22}$$

It easy to see that the solution of the following problem

$$\min_y \{h(\alpha, y) : y \in K\}$$

is $y^* = P_K[u^k - \alpha d(\tilde{u}^k, \bar{u}^k)]$. By substituting y^* into (22) and simplifying it, we obtain

$$\Phi(\alpha) = h(\alpha, y)|_{y=P_K[u^k - \alpha d(\tilde{u}^k, \bar{u}^k)]} = \min_y \{h(\alpha, y) : y \in K\}.$$

$\Phi(\alpha)$ is differentiable and its derivative is given by

$$\begin{aligned} \Phi'(\alpha) &= \frac{\partial h(\alpha, y)}{\partial \alpha} \Big|_{y=P_K[u^k - \alpha d(\tilde{u}^k, \bar{u}^k)]} \\ &= 2\langle u^{k+1}(\alpha) - u^k + \alpha d(\tilde{u}^k, \bar{u}^k), d(\tilde{u}^k, \bar{u}^k) \rangle - 2\alpha\|d(\tilde{u}^k, \bar{u}^k)\|^2 \\ &\quad - 2\langle \bar{u}^k - u^k, d(\tilde{u}^k, \bar{u}^k) \rangle \\ &= 2\langle u^{k+1}(\alpha) - \bar{u}^k, d(\tilde{u}^k, \bar{u}^k) \rangle, \end{aligned}$$

and hence (i) is proved. We now establish the proof of the second assertion. Let $\bar{\alpha} > \alpha \geq 0$. We show that

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

that is,

$$\langle u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha), d(\bar{u}^k, \bar{u}^k) \rangle \leq 0. \tag{23}$$

By setting $z := u^k - \bar{\alpha}d(\bar{u}^k, \bar{u}^k)$, $v := u^{k+1}(\alpha)$ and $z := u^k - \alpha d(\bar{u}^k, \bar{u}^k)$, $v := u^{k+1}(\bar{\alpha})$ in (a), respectively, we get

$$\langle u^k - \bar{\alpha}d(\bar{u}^k, \bar{u}^k) - u^{k+1}(\bar{\alpha}), u^{k+1}(\alpha) - u^{k+1}(\bar{\alpha}) \rangle \leq 0 \tag{24}$$

and

$$\langle u^k - \alpha d(\bar{u}^k, \bar{u}^k) - u^{k+1}(\alpha), u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha) \rangle \leq 0. \tag{25}$$

By adding (24) and (25), we obtain

$$\langle u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha), u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha) + (\bar{\alpha} - \alpha)d(\bar{u}^k, \bar{u}^k) \rangle \leq 0,$$

that is,

$$\|u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)\|^2 + (\bar{\alpha} - \alpha)\langle u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha), d(\bar{u}^k, \bar{u}^k) \rangle \leq 0.$$

It follows that

$$\langle u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha), d(\bar{u}^k, \bar{u}^k) \rangle \leq \frac{-1}{(\bar{\alpha} - \alpha)} \|u^{k+1}(\bar{\alpha}) - u^{k+1}(\alpha)\|^2 \leq 0.$$

We obtain inequality (23) and complete the proof. □

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

$$\alpha_{k_1}^* = \arg \max_{\alpha} \{ \Psi(\alpha) \mid \alpha > 0 \}$$

and

$$\alpha_{k_2}^* = \arg \max_{\alpha} \{ \Phi(\alpha) \mid \alpha > 0 \}.$$

Since $\Psi(\alpha)$ is a quadratic function of α and it reaches its maximum at

$$\alpha_{k_1}^* = \frac{\langle u^k - \bar{u}^k, d_1(\bar{u}^k, \bar{u}^k) \rangle}{(\beta_1 + \beta_2)\|d_1(\bar{u}^k, \bar{u}^k)\|^2} \tag{26}$$

and

$$\Psi(\alpha_{k_1}^*) = \alpha_{k_1}^* (\beta_1 + \beta_2) \langle u^k - \bar{u}^k, d_1(\bar{u}^k, \bar{u}^k) \rangle. \tag{27}$$

In order to make $\alpha_{k_2}^*$ to be obtained more easily, we approximately compute $\alpha_{k_2}^*$ by solving the following simple optimization problem:

$$\alpha_{k_2}^* = \arg \max_{\alpha} \{ \Phi(\alpha) \mid 0 < \alpha \leq m_1 \alpha_{k_1}^* \}, \tag{28}$$

where $m_1 \geq 1$.

Based on Theorem 1 and Proposition 1, the following result can be proved easily.

Proposition 2 *Let $\alpha_{k_1}^*$ and $\alpha_{k_2}^*$ be defined by (26) and (28), respectively, and let T be pseudomonotone and continuously differentiable. Then*

- (i) $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_{k_2}^*) - u^*\|^2 \geq \Phi(\alpha_{k_2}^*)$;
- (ii) $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_{k_1}^*) - u^*\|^2 \geq \Psi(\alpha_{k_1}^*)$;
- (iii) $\Phi(\alpha_{k_2}^*) \geq \Psi(\alpha_{k_1}^*)$;
- (iv) *if $\Phi'(\alpha_{k_2}^*) = 0$, then $\|u^k - u^*\|^2 - \|u^{k+1}(\alpha_{k_2}^*) - u^*\|^2 \geq \|u^k - u^{k+1}(\alpha_{k_2}^*)\|^2$.*

Remark 3 Let

$$u_I^{k+1}(\alpha_{k_1}^*) = P_K[u^k - \alpha_{k_1}^* d(u^k, \rho_k)]$$

and

$$u_{II}^{k+1}(\alpha_{k_2}^*) = P_K[u^k - \alpha_{k_2}^* d(u^k, \rho_k)]$$

represent the new iterates generated by the proposed method with $\alpha_k = \alpha_{k_1}^*$ and $\alpha_k = \alpha_{k_2}^*$, respectively. Let

$$\Theta_I(\alpha_{k_1}^*) = \|u^k - u^*\|^2 - \|u_I^{k+1}(\alpha_{k_1}^*) - u^*\|^2$$

and

$$\Theta_{II}(\alpha_{k_2}^*) = \|u^k - u^*\|^2 - \|u_{II}^{k+1}(\alpha_{k_2}^*) - u^*\|^2$$

measure the progresses made by the new iterates, respectively. By using Proposition 2, it is easy to show that

$$\begin{aligned} \Theta_I(\alpha_{k_1}^*) &\geq \Psi(\alpha_{k_1}^*), \\ \Theta_{II}(\alpha_{k_2}^*) &\geq \Phi(\alpha_{k_2}^*) \end{aligned}$$

and

$$\Phi(\alpha_{k_2}^*) \geq \Psi(\alpha_{k_1}^*).$$

The above inequalities show that the proposed method with $\alpha_k = \alpha_{k_2}^*$ is expected to make more progress than the proposed method with $\alpha_k = \alpha_{k_1}^*$ at each iteration, and so it explains theoretically that the proposed method with $\alpha_k = \alpha_{k_2}^*$ outperforms the proposed method with $\alpha_k = \alpha_{k_1}^*$.

In the next theorem, we show that $\alpha_{k_1}^*$ and $\Psi(\alpha_{k_1}^*)$ are lower bounded away from zero, and it is one of the keys to prove the global convergence results.

Theorem 2 *Let u^* be a solution of problem (1). For given $u^k \in K$, let \tilde{u}^k, \bar{u}^k be the predictors produced by (6a) and (6b). Then*

$$\alpha_{k_1}^* \geq \frac{2 - \mu^2}{(\beta_1 + \beta_2)(1 + \nu)^2} \tag{29}$$

and

$$\Psi(\alpha_{k_1}^*) \geq \frac{(2 - \mu^2)^2}{(1 + \nu)^2} \|\tilde{u}^k - \bar{u}^k\|^2. \tag{30}$$

Proof Note that $\tilde{u}^k = P_K[u^k - \rho_k T(u^k)]$, $\bar{u}^k = P_K[\tilde{u}^k - \rho_k T(\tilde{u}^k)]$. We can apply (c) with $\nu = u^k - \rho_k T(u^k)$, $w = \tilde{u}^k - \rho_k T(\tilde{u}^k)$ to obtain

$$\langle u^k - \rho_k T(u^k) - (\tilde{u}^k - \rho_k T(\tilde{u}^k)), \tilde{u}^k - \bar{u}^k \rangle \geq \|\tilde{u}^k - \bar{u}^k\|^2.$$

By some manipulations, we obtain

$$\langle u^k - \tilde{u}^k, \tilde{u}^k - \bar{u}^k \rangle \geq \|\tilde{u}^k - \bar{u}^k\|^2 + \rho_k \langle \tilde{u}^k - \bar{u}^k, T(u^k) - T(\tilde{u}^k) \rangle.$$

Then we have

$$\begin{aligned} \langle u^k - \tilde{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle &= \langle u^k - \tilde{u}^k, \tilde{u}^k - \bar{u}^k \rangle - \rho_k \langle u^k - \tilde{u}^k, T(\tilde{u}^k) - T(\bar{u}^k) \rangle \\ &\geq \|\tilde{u}^k - \bar{u}^k\|^2 + \rho_k \langle \tilde{u}^k - \bar{u}^k, T(u^k) - T(\tilde{u}^k) \rangle \\ &\quad - \rho_k \langle u^k - \tilde{u}^k, T(\tilde{u}^k) - T(\bar{u}^k) \rangle. \end{aligned} \tag{31}$$

By using (31), (7) and the definition of $d_1(\tilde{u}^k, \bar{u}^k)$, we get

$$\begin{aligned} \langle u^k - \tilde{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle &= \langle u^k - \tilde{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle + \langle \tilde{u}^k - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle \\ &\geq \|\tilde{u}^k - \bar{u}^k\|^2 + \rho_k \langle \tilde{u}^k - \bar{u}^k, T(u^k) - T(\tilde{u}^k) \rangle \\ &\quad - \rho_k \langle u^k - \tilde{u}^k, T(\tilde{u}^k) - T(\bar{u}^k) \rangle + \|\tilde{u}^k - \bar{u}^k\|^2 \\ &\quad - \rho_k \langle \tilde{u}^k - \bar{u}^k, T(\tilde{u}^k) - T(\bar{u}^k) \rangle \\ &\geq (2 - \mu^2) \|\tilde{u}^k - \bar{u}^k\|^2. \end{aligned} \tag{32}$$

Recalling the definition of $d_1(\tilde{u}^k, \bar{u}^k)$ (see (12)) and applying criterion (8), it is easy to see that

$$\|d_1(\tilde{u}^k, \bar{u}^k)\|^2 \leq (\|\tilde{u}^k - \bar{u}^k\| + \|\rho_k(T(\tilde{u}^k) - T(\bar{u}^k))\|)^2 \leq (1 + \nu)^2 \|\tilde{u}^k - \bar{u}^k\|^2. \tag{33}$$

Moreover, by using (32) together with (33), we get

$$\alpha_{k_1}^* = \frac{\langle u^k - \bar{u}^k, d_1(\tilde{u}^k, \bar{u}^k) \rangle}{(\beta_1 + \beta_2) \|d_1(\tilde{u}^k, \bar{u}^k)\|^2} \geq \frac{2 - \mu^2}{(\beta_1 + \beta_2)(1 + \nu)^2} > 0, \quad \mu \in (0, \sqrt{2}). \tag{34}$$

By substituting (34) in (27), we get the assertion of this theorem and the proof is completed. □

From the computational point of view, a relaxation factor $\gamma \in (0, 2)$ is preferable in the correction. We are now in the position to prove the contractive property of the iterative sequence.

Lemma 4 *Let u^* be a solution of problem (1) and let $u^{k+1}(\gamma\alpha_k)$ be the sequence generated by Algorithm 1 with $\alpha_k = \alpha_{k_1}^*$ or $\alpha_k = \alpha_{k_2}^*$. Then u^k is bounded and*

$$\|u^{k+1}(\gamma\alpha_k) - u^*\|^2 \leq \|u^k - u^*\|^2 - c\|u^k - \tilde{u}^k\|^2, \tag{35}$$

where

$$c := \frac{(2 - \mu^2)^2}{(1 + \nu)^2} > 0.$$

Proof If $\alpha_k = \alpha_{k_2}^*$. It follows from (14), (29) and (30) that

$$\begin{aligned} \|u^{k+1}(\gamma\alpha_{k_2}^*) - u^*\|^2 &\leq \|u^k - u^*\|^2 - \Phi(\alpha_{k_2}^*) \\ &\leq \|u^k - u^*\|^2 - \Psi(\alpha_{k_1}^*) \\ &\leq \|u^k - u^*\|^2 - \frac{(2 - \mu^2)^2}{(1 + \nu)^2} \|\tilde{u}^k - \bar{u}^k\|^2. \end{aligned}$$

If $\alpha_k = \alpha_{k_1}^*$, then we have

$$\begin{aligned} \|u^{k+1}(\gamma\alpha_{k_1}^*) - u^*\|^2 &\leq \|u^k - u^*\|^2 - \Psi(\alpha_{k_1}^*) \\ &\leq \|u^k - u^*\|^2 - \frac{(2 - \mu^2)^2}{(1 + \nu)^2} \|\tilde{u}^k - \bar{u}^k\|^2. \end{aligned}$$

Since $\mu \in (0, \sqrt{2})$, we have

$$\|u^{k+1}(\gamma\alpha_k) - u^*\| \leq \|u^k - u^*\| \leq \dots \leq \|u^0 - u^*\|.$$

From the above inequality, it is easy to verify that the sequence u^k is bounded. □

We now present the convergence result of Algorithm 1.

Theorem 3 *If $\inf_{k=0}^\infty \rho_k := \rho > 0$, then any cluster point of the sequence $\{\tilde{u}^k\}$ generated by Algorithm 1 is a solution of problem (1).*

Proof It follows from (35) that

$$\sum_{k=0}^\infty c\|u^k - \tilde{u}^k\|^2 < +\infty,$$

which means that

$$\lim_{k \rightarrow \infty} \|\tilde{u}^k - \bar{u}^k\| = 0.$$

Since the sequence u^k is bounded, $\{\bar{u}^k\}$ is too, and it has at least a cluster point. Let u^∞ be a cluster point of $\{\bar{u}^k\}$ and the subsequence $\{\tilde{u}^{k_j}\}$ converges to u^∞ . By the continuity of $e(u, \rho)$ and inequality (4), it follows that

$$\|e(u^\infty, \rho)\| = \lim_{k_j \rightarrow \infty} \|e(\tilde{u}^{k_j}, \rho)\| \leq \lim_{k_j \rightarrow \infty} \|e(\tilde{u}^{k_j}, \rho_{k_j})\| = \lim_{k_j \rightarrow \infty} \|\tilde{u}^{k_j} - \bar{u}^{k_j}\| = 0.$$

This means that u^∞ is a solution of problem (1).

We now prove that the sequence $\{u^k\}$ has exactly one cluster point. Assume that \tilde{u} is another cluster point and satisfies

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

Since u^∞ is a cluster point of the sequence $\{u^k\}$, there is $k_0 > 0$ such that

$$\|u^{k_0} - u^\infty\| \leq \frac{\delta}{2}.$$

On the other hand, since $u^\infty \in S^*$ and from (35), we have

$$\|u^k - u^\infty\| \leq \|u^{k_0} - u^\infty\| \quad \text{for all } k \geq k_0,$$

it follows that

$$\|u^k - \tilde{u}\| \geq \|\tilde{u} - u^\infty\| - \|u^k - u^\infty\| \geq \frac{\delta}{2} \quad \forall k \geq k_0.$$

This contradicts the assumption that \tilde{u} is a cluster point of $\{u^k\}$. Thus, the sequence $\{u^k\}$ converges to $u^\infty \in S^*$. □

4 Numerical experiments

In order to verify the theoretical assertions, we consider the nonlinear complementarity problems

$$u \geq 0, \quad T(u) \geq 0, \quad u^\top T(u) = 0, \tag{36}$$

where $T(u) = D(u) + Mu + q$, $D(u)$ and $Mu + q$ are the nonlinear part and linear part of $T(u)$, respectively. Problem (36) is equivalent to problem (1) by taking $K = \mathbb{R}_+^n$.

Example 1 We form the linear part in the test problems as

$$M = \begin{pmatrix} 1 & 2 & 2 & \cdots & 2 \\ 0 & 1 & 2 & \cdots & 2 \\ 0 & 0 & 1 & \cdots & 2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}_{n \times n} \quad \text{and} \quad q = \begin{pmatrix} -1 \\ -1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}_{n \times 1}.$$

In $D(u)$, the nonlinear part of $T(u)$, the components are chosen to be $D_j(u) = d_j * \arctan(u_j)$, where d_j is a random variable in $(0, 1)$.

In all tests we take $\rho_0 = 1, \zeta = 0.7, \eta_1 = 0.3, \eta_2 = 0.95, \mu = 0.95, \nu = 1.95, \gamma = 1.9, \beta_1 = 0.5$, and $\beta_2 = 1.5$. We employ $\|e(u^k, \rho_k)\| \leq 10^{-7}$ as the stopping criterion and choose $u^0 = 0$ as the initial iterative points. All codes were written in Matlab. We compare Algorithm 1 with $\alpha_k = \alpha_{k_2}^*$, with Algorithm 1 with $\alpha_k = \alpha_{k_1}^*$ and the method in [14]. The test results for problem (36) with different dimensions are reported in Tables 1 and 2. k is the number of iterations and l denotes the number of evaluations of mapping T .

Table 1 Numerical results for problem (36) with $\sigma = 0.8$ (see (9))

Dimension of the problem	The method in [14]			Algorithm 1 with $\alpha_{k_1}^*$			Algorithm 1 with $\alpha_{k_2}^*$		
	k	l	CPU (sec.)	k	l	CPU (sec.)	k	l	CPU (sec.)
$n = 100$	228	465	0.024	138	447	0.023	73	240	0.06
$n = 300$	259	540	0.04	185	580	0.03	103	332	0.09
$n = 500$	531	1,109	0.15	227	700	0.10	128	403	0.16
$n = 600$	520	1,160	0.22	225	701	0.13	135	419	0.20
$n = 800$	568	1,236	0.51	244	779	0.34	157	520	0.40

Table 2 Numerical results for problem (36) with $\sigma = 0.4$ (see (9))

Dimension of the problem	The method in [14]			Algorithm 1 with $\alpha_{k_1}^*$			Algorithm 1 with $\alpha_{k_2}^*$		
	k	l	CPU (sec.)	k	l	CPU (sec.)	k	l	CPU (sec.)
$n = 100$	228	465	0.02	142	439	0.01	99	309	0.06
$n = 300$	259	540	0.04	180	553	0.03	97	302	0.08
$n = 500$	531	1,109	0.15	226	699	0.09	189	579	0.23
$n = 600$	520	1,160	0.23	251	768	0.14	129	400	0.18
$n = 800$	568	1,236	0.48	246	754	0.31	197	603	0.48

Table 3 Numerical results for problem (36) with $\sigma = 0.8$ (see (9))

Dimension of the problem	The method in [14]			Algorithm 1 with $\alpha_{k_1}^*$			Algorithm 1 with $\alpha_{k_2}^*$		
	k	l	CPU (sec.)	k	l	CPU (sec.)	k	l	CPU (sec.)
$n = 100$	318	676	0.03	93	312	0.01	68	235	0.05
$n = 300$	435	936	0.07	127	404	0.03	111	356	0.09
$n = 500$	489	1,035	0.15	146	491	0.07	129	416	0.17
$n = 600$	406	877	0.18	117	378	0.08	92	299	0.15
$n = 800$	386	832	0.64	110	359	0.29	76	249	0.28

Table 4 Numerical results for problem (36) with $\sigma = 0.4$ (see (9))

Dimension of the problem	The method in [14]			Algorithm 1 with $\alpha_{k_1}^*$			Algorithm 1 with $\alpha_{k_2}^*$		
	k	l	CPU (sec.)	k	l	CPU (sec.)	k	l	CPU (sec.)
$n = 100$	318	676	0.02	157	793	0.02	84	298	0.06
$n = 300$	435	995	0.06	199	936	0.07	164	613	0.16
$n = 500$	489	1,035	0.14	190	769	0.12	155	550	0.22
$n = 600$	406	877	0.20	129	402	0.08	89	300	0.14
$n = 800$	386	832	0.35	169	714	0.32	89	309	0.26

Example 2 We form the linear part in the test problems similarly as in Harker and Pang [28]. The matrix $M = A^T A + B$, where A is an $n \times n$ matrix whose entries are randomly generated in the interval $(-5, +5)$, and a skew-symmetric matrix B is generated in the same way. The vector q is generated from a uniform distribution in the interval $(-500, 500)$. In $D(u)$, the nonlinear part of $T(u)$, the components are chosen to be $D_j(u) = d_j * \arctan(u_j)$, where d_j is a random variable in $(0, 1)$. A similar type of problem was tested in [29] and [30]. In all tests we take $\rho_0 = 1, \zeta = 0.7, \eta_1 = 0.3, \eta_2 = 0.95, \mu = 0.95, \nu = 1.95, \gamma = 1.9, \beta_1 = 0.5, \beta_2 = 1.5$. We employ $\|e(u^k, \rho_k)\| \leq 10^{-7}$ as the stopping criterion and choose $u^0 = 0$ as the initial iterative points. The test results for problem (36) with different dimensions are reported in Tables 3 and 4.

Tables 1-4 show that Algorithm 1 with $\alpha_k = \alpha_{k_2}^*$ is very efficient even for large-scale classical nonlinear complementarity problems. Moreover, it demonstrates computationally

that Algorithm 1 with $\alpha_k = \alpha_{k_2}^*$ is more effective than the method presented in [14] and Algorithm 1 with $\alpha_k = \alpha_{k_1}^*$ in the sense that Algorithm 1 with $\alpha_k = \alpha_{k_2}^*$ needs fewer iterations and less evaluation numbers of T , which clearly illustrates its efficiency.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have equal contribution. All authors read and approved the final manuscript.

Author details

¹School of Management Science and Engineering, Nanjing University, Nanjing, 210093, P.R. China. ²ENSA, Ibn Zohr University, BP 1136, Agadir, Morocco. ³Department of Mathematics, Aligarh Muslim University, Aligarh, 202 002, India.

⁴Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

⁵Center for Fundamental Science, Kaohsiung Medical University, Kaohsiung, 807, Taiwan.

Acknowledgements

The research part of second author was done during his visit to King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia. In this research, third author was partially supported by the grant MOST 103-2115-M-037-001.

Received: 14 February 2015 Accepted: 14 April 2015 Published online: 23 April 2015

References

- Fichera, G: Problemi elettrostatici con vincoli unilaterali: il problema di Signorini con ambigue condizioni al contorno. *Atti Accad. Naz. Lincei, Mem. Cl. Sci. Fis. Mat. Nat., Sez. I* **7**, 91-140 (1964)
- Stampacchia, G: Formes bilineaires coercitives sur les ensembles convexes. *C. R. Math. Acad. Sci. Paris* **258**, 4413-4416 (1964)
- Stampacchia, G: Variational inequalities. In: Ghizzetti, A (ed.) *Theory and Applications of Monotone Operators*. Proc. NATO Adv. Study Institute, Venice, Oderisi, Gubbio (1968)
- Ansari, QH, Lalitha, CS, Mehta, M: *Generalized Convexity, Nonsmooth Variational Inequalities and Nonsmooth Optimization*. CRC Press, Boca Raton (2014)
- Facchinei, F, Pang, JS: *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer, Berlin (2003)
- Glowinski, R, Lions, JL, Trémolieres, R: *Numerical Analysis of Variational Inequalities*. North-Holland, Amsterdam (1981)
- Kinderlehrer, D, Stampacchia, G: *An Introduction to Variational Inequalities and Their Applications*. Academic Press, New York (1980)
- Konnov, IV: *Combined Relaxation Methods for Variational Inequalities*. Springer, Berlin (2001)
- Khobotov, EN: Modification of the extragradient method for solving variational inequalities and certain optimization problems. *USSR Comput. Math. Math. Phys.* **27**, 120-127 (1987)
- Nagurney, A, Zhang, D: *Projected Dynamical Systems and Variational Inequalities with Applications*. Kluwer Academic, Dordrecht (1996)
- Korpelevich, GM: The extragradient method for finding saddle points and other problems. *Matecon* **12**, 747-756 (1976)
- Iusem, AN, Svaiter, BF: A variant of Korpelevich's method for variational inequalities with a new strategy. *Optimization* **42**, 309-321 (1997)
- Auslender, A, Teboulle, M: Interior projection-like methods for monotone variational inequalities. *Math. Program., Ser. A* **104**, 39-68 (2005)
- Bnouhachem, A, Noor, MA, Rassias, TR: Three-steps iterative algorithms for mixed variational inequalities. *Appl. Math. Comput.* **183**, 436-446 (2006)
- Bnouhachem, A, Fu, X, Xu, MH, Zhaohan, S: Modified extragradient methods for solving variational inequalities. *Comput. Math. Appl.* **57**, 230-239 (2009)
- Fu, X: A two-stage prediction-correction method for solving variational inequalities. *J. Comput. Appl. Math.* **214**, 345-355 (2008)
- He, BS, Liao, LZ: Improvement of some projection methods for monotone variational inequalities. *J. Optim. Theory Appl.* **112**, 111-128 (2002)
- He, BS, Yang, H, Meng, Q, Han, DR: Modified Goldstein-Levitin-Polyak projection method for asymmetric strongly monotone variational inequalities. *J. Optim. Theory Appl.* **112**, 129-143 (2002)
- Wang, YJ, Xiu, NH, Wang, CY: A new version of extragradient method for variational inequalities problems. *Comput. Math. Appl.* **42**, 969-979 (2001)
- Xiu, NH, Zhang, JZ: Some recent advances in projection-type methods for variational inequalities. *J. Comput. Appl. Math.* **152**, 559-585 (2003)
- Bnouhachem, A: A self-adaptive method for solving general mixed variational inequalities. *J. Math. Anal. Appl.* **309**, 136-150 (2005)
- He, BS, Yang, ZH, Yuan, XM: An approximate proximal-extragradient type method for monotone variational inequalities. *J. Math. Anal. Appl.* **300**, 362-374 (2004)
- Zarantonello, EH: Projections on convex sets in Hilbert space and spectral theory. In: Zarantonello, EH (ed.) *Contributions to Nonlinear Functional Analysis*. Academic Press, New York (1971)
- Calamai, PH, Moré, JJ: Projected gradient methods for linearly constrained problems. *Math. Program.* **39**, 93-116 (1987)

25. Gafni, EM, Bertsekas, DP: Two-metric projection methods for constrained optimization. *SIAM J. Control Optim.* **22**, 936-964 (1984)
26. Peng, JM, Fukushima, M: A hybrid Newton method for solving the variational inequality problem via the D-gap function. *Math. Program.* **86**, 367-386 (1999)
27. He, BS, Yuan, XM, Zhang, JZ: Comparison of two kinds of prediction-correction methods for monotone variational inequalities. *Comput. Optim. Appl.* **27**, 247-267 (2004)
28. Harker, PT, Pang, JS: A damped-Newton method for the linear complementarity problem. *Lect. Appl. Math.* **26**, 265-284 (1990)
29. Marcotte, P, Dussault, JP: A note on a globally convergent Newton method for solving variational inequalities. *Oper. Res. Lett.* **6**, 35-42 (1987)
30. Taji, K, Fukushima, M, Ibaraki, T: A globally convergent Newton method for solving strongly monotone variational inequalities. *Math. Program.* **58**, 369-383 (1993)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ springeropen.com
