Journal of Global Optimization **28:** 97–113, 2004.

Solving Variational Inequality Problems with Linear Constraints by a Proximal Decomposition Algorithm

DEREN HAN[1] and HONG K. LO[2]
[1]*School of Mathematics and Computer Sciences, Nanjing Normal University, Nanjing, 210097, China (e-mail: handr00@sina.com)* [2]*Department of Civil Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China (e-mail: cehklo@ust.hk)*

(Received 12 November 2001; accepted in revised form 22 April 2003)

Abstract. The alternating direction method solves large scale variational inequality problems with linear constraints via solving a series of small scale variational inequality problems with simple constraints. The algorithm is attractive if the subproblems can be solved efficiently and exactly. However, the subproblem is itself variational inequality problem, which is structurally also difficult to solve. In this paper, we develop a new decomposition algorithm, which, at each iteration, just solves a system of well-conditioned linear equations and performs a line search. We allow to solve the subproblem approximately and the accuracy criterion is the constructive one developed recently by Solodov and Svaiter. Under mild assumptions on the problem's data, the algorithm is proved to converge globally. Some preliminary computational results are also reported to illustrate the efficiency of the algorithm.

Key words: Decomposition algorithms, Global convergence, Inexact method, Monotone mappings, Variational inequality problems

1. Introduction

Consider the variational inequality problem, denoted by VI(f,S), which is to find a vector $x^* \in S$, such that

$$f(x^*)^\top (z-x^*) \geqslant 0, \quad \forall z \in S,$$

where $S \subset R^n$ is a nonempty closed convex subset of R^n and f is a continuous, monotone mapping from R^n into itself. This problem has several important applications in many fields, such as network economics, traffic assignment, game theoretic problems, etc. [1, 3, 24]. There are a substantial number of iterative methods including the projection method and its variant forms [1, 18–20], the linearized Jacobi method [17], Newton-type methods [17, 23, 27], etc.

In this paper, we will focus our attention on VI(f,S) where S has the following structure:

$$S = S_1 = \{x \in R^n \mid Bx = b, x \geqslant 0\}, \tag{1}$$

or

$$S = S_2 = \{x \in R^n \mid Bx \geqslant b, x \geqslant 0\}, \tag{2}$$

where $B \in R^{m\times n}$ is a given matrix and $b \in R^m$ is a given vector. Though it is a special case of VI(f, S), this problem finds many important applications in the fields such as traffic equilibrium and network equilibrium problems.

A typical method for solving the primal problem VI(f, S) with structure (1) is the following decomposition algorithm proposed by Gabay [8] and Gabay and Mercier [9], which is called *alternating direction method*:

Given $(x^k, y^k) \in R^n_+ \times R^m$, *find* $x^{k+1} \geqslant 0$, *such that*

$$(x' - x^{k+1})^\top \{f(x^{k+1}) - B^\top [y^k - (Bx^{k+1} - b)]\} \geqslant 0, \quad \forall x' \geqslant 0, \tag{3}$$

then update y via

$$y^{k+1} = y^k - (Bx^{k+1} - b).$$

Note that this algorithm can also be used to solve the variational inequality problem with $S = S_2$ by introducing a slack vector to the linear inequality constraint to transform S_2 to the same form as S_1,

$$S_2 = \{(x, z) \in R^n \times R^m \mid Bx - z = b, z \geqslant 0\}.$$

However, this will increase the dimension of the subproblem (3) from n to $n+m$.

Then, for solving the variational inequality problem VI(f, S) with $S = S_2$, another decomposition method was proposed [7, 8, 9, 10], which is called *method of multiplier:*

Given $(x^k, z^k, y^k) \in R^n_+ \times R^m_+ \times R^m$, *find* $x^{k+1} \geqslant 0$, *such that*

$$(x' - x^{k+1})^\top \{f(x^{k+1}) - B^\top [y^k - (Bx^{k+1} - z^k - b)]\} \geqslant 0, \quad \forall x' \geqslant 0, \tag{4}$$

then update z via

$$z^{k+1} = \max\{0, Bx^{k+1} - y^k - b\},$$

and update y via

$$y^{k+1} = y^k - (Bx^{k+1} - z^k - b).$$

These decompositions methods are attractive for large scale problems, since they decompose the original problems into a series of subproblems with lower scale. However, note that both (3) and (4) are still variational inequality problems, which are structurally difficult to solve.

Recently, Wang, Yang and He [28] proposed a new decomposition algorithm for solving VI(f, S) with $S = S_1$ or $S = S_2$ uniformly. At each iteration k, for a given (or obtained) point x^k, they first solve the following linear variational inequality problem to get y^k

$$(y' - y^k)^\top \{(Ax^k - a) - A(f(x^k) - A^\top y^k)\} \geqslant 0, \quad \forall y' \in Y, \tag{5}$$

and then solve the following system of nonlinear equations to get the next iteration x^{k+1}

$$x^{k+1} + f(x^{k+1}) = x^k + f(x^k) - \gamma(f(x^k) - A^\top y^k), \tag{6}$$

where $\gamma \in (0,2)$ is a given constant,

$$A = \begin{pmatrix} B \\ I \end{pmatrix}, \quad a = \begin{pmatrix} b \\ 0 \end{pmatrix},$$

and Y is a set in R^{m+n} with

$$Y = Y_I \times Y_{II}, \quad Y_I \subset R^m, \quad \text{and} \quad Y_{II} = R^n_+,$$

$Y_I = R^m$ when $S = S_1$ and $Y_I = R^m_+$ when $S = S_2$. Their algorithm avoids solving the variational inequality problems (3) or (4), which are structurally more difficult to solve than (6). Instead, they solve the linear variational inequality problem (5) and a system of nonlinear equations (6). Though the latter problem is structurally easier to solve than the variational inequality problems (3) and (4), the linear variational problem is sometimes time consuming. Moreover, in [28], the authors proved that $\{x^k\}$ converges to x^* and when the matrix A has full row rank, $\{y^k\}$ converges to y^*, where x^* is a solution of VI(f,S) and y^* is the corresponding Lagrange multiplier to the linear constraints $Bx = b$ ($Bx \geqslant b$) and the nonnegative constraints. From the structure of A, we can see that A has full row rank only for the special case that $B = 0$ and VI(f,S) reduces to the nonlinear complementarity problem of finding $x \in R^n$, such that

$$x \geqslant 0, \quad f(x) \geqslant 0, \quad x^\top f(x) = 0.$$

More recently, Han [16] also considered the variational inequality problem with (1) and (2) uniformly and proposed the following proximal decomposition algorithm:

Given the current iteration $(x^k, y^k) \in R^n \times Y$, solve the following system of nonlinear equations

$$c_k(f(\cdot) - A^\top y^k) + (\cdot - x^k) = r^k, \tag{7}$$

such that

$$\|r^k\| \leqslant \sigma \|x^k - \bar{x}^k\|. \tag{8}$$

Then, set

$$\bar{y}^k = P_Y[y^k - (A\bar{x}^k - a)] \tag{9}$$

and

$$g(u^k) = g(x^k, y^k) = \begin{pmatrix} f(\bar{x}^k) - A^\top \bar{y}^k \\ y^k - \bar{y}^k \end{pmatrix}.$$

Finally compute α_k *by*

$$\alpha_k = g(u^k)^\top (u^k - \bar{u}^k)/\|g(u^k)\|^2, \tag{10}$$

and get the new iteration $u^{k+1} = (x^{k+1}, y^{k+1})$ *via*

$$u^{k+1} = u^k - \alpha_k g(u^k). \tag{11}$$

Note that at each iteration, this algorithm needs only to solve a system of well conditioned nonlinear equations with the same structure as (6) and perform a projection to the simple set Y. Furthermore, it allows to solve the system of equations approximately and adopts a constructive accuracy criterion (8) developed recently by Solodov and Svaiter [26], which is more constructive than the classical one assuming the summability or the square summability of the sequence of the error tolerance parameters [25, 21, 15]. The computational results reported there are encouraging.

Inspired by these, this paper develops a new decomposition algorithm for solving variational inequality problems VI(f, S) with $S = S_1$ or $S = S_2$. At each iteration, instead of solving the structurally difficult problems (3) or (4), or the system of nonlinear equations (7), this algorithm just first solves a system of linear equations with respect to the variable x. Then, it performs an Armijo-type line search to get a suitable stepsize. It also allows to solve the equations approximately, and adopts the same accuracy criterion, which makes the algorithm more practical. We prove that under mild assumptions that the underlying mapping f is continuous and monotone and the solutions set is nonempty, the sequence generated by the algorithm converges to a solution globally.

The remainder of the paper is organized as follows. In the next section, we summarize some basic definitions and properties to be used in this paper. In Section 3, the new decomposition algorithm is described formally and its global convergence is proved in Section 4 under mild condition that the underlying mapping f is continuous and monotone. In Section 5, we report some preliminary computational results of the proposed method, and Section 6 gives some concluding remarks.

2. Preliminaries

In this section, we summarize some basic concepts and their properties that will be useful in the sequent sections.

First, we denote $\|x\| = \sqrt{x^\top x}$ as the Euclidean norm. Let K be a nonempty closed convex subset of R^n and let $P_K[\cdot]$ denote the projection mapping from R^n onto K. The following well known properties of the projection operator will be used bellow.

LEMMA 2.1. *Let K be a nonempty closed convex subset of R^n. For any $x, y \in R^n$ and any $z \in K$, the following properties hold.*

1. $(x - P_K[x])^\top (z - P_K[x]) \leqslant 0$.
2. $\|P_K[x] - P_K[y]\|^2 \leqslant \|x - y\|^2 - \|P_K[x] - x + y - P_K[y]\|^2$. □

For any positive real number β and $x \in K$, let $e(x, \beta)$ denote the residue function associated with the mapping f, i.e.,

$$e(x, \beta) = x - P_K[x - \beta f(x)].$$

LEMMA 2.2. *x^* is a solution of the VI(f, K) if and only if $e(x^*, \beta) = 0$ for any given positive real number β.*

Proof. It is clear that the solutions set of VI(f, K) is invariant under multiplication f by some positive scalar β. According to ([2], p. 267) (See also [5]), it is equivalent to $e(x^*, \beta) = 0$. □

Clearly, to solve VI(f, K) is equivalent to finding a zero point of the residue function $e(x, \beta)$ for any given positive β. The lemma also provides an important stopping criterion for designing a solution method.

We need the following definitions concerning the functions.

DEFINITION 2.3.

a). A mapping $f : R^n \to R^n$ is said to be monotone, if

$$(x - y)^\top (f(x) - f(y)) \geqslant 0, \ \forall x, y \in R^n.$$

b). A mapping $f : R^n \to R^n$ is said to be strongly monotone with modulus $\gamma > 0$, if

$$(x - y)^\top (f(x) - f(y)) \geqslant \gamma \|x - y\|^2, \ \forall x, y \in R^n.$$

If f is continuously differentiable and strongly monotone, then $\nabla f(x)$ is uniformly positive definite for all $x \in R^n$.

In the remainder of this paper, we always suppose that the underlying mapping f of the variational inequality problem under consideration is continuous and monotone.

3. The Decomposition Algorithm

Note that, by introducing a Lagrange multiplier y to the linear constraint and the nonnegative constraint, we can transform VI(f, S_1) and VI(f, S_2) to the uniform description of finding a vector $u^* \in \Omega$, such that

$$F(u^*)^\top (u - u^*) \geqslant 0, \quad \forall u \in \Omega, \tag{12}$$

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(u) = F(x,y) = \begin{pmatrix} f(x) - A^\top y \\ Ax - a \end{pmatrix}, \quad \Omega = R^n \times Y, \tag{13}$$

and

$$A = \begin{pmatrix} B \\ I \end{pmatrix}, \quad a = \begin{pmatrix} b \\ 0 \end{pmatrix},$$

Y is a set in R^{m+n} with $Y = Y_I \times Y_{II}$, $Y_I \subset R^m$ and $Y_{II} = R^n_+$. The only difference is $Y_I = R^m$ if $S = S_1$ and $Y_I = R^m_+$ if $S = S_2$. In this paper, we focus our attention to the structured variational inequality problem (12)–(13), denoted by VI(F, Ω). We assume that the solution set of VI(F, Ω) is nonempty.

We are now in the position to describe our method formally.

ALGORITHM 3.1. An inexact decomposition algorithm.

Step 0. Choose an arbitrary initial point $u^0 = (x^0, y^0) \in R^n \times Y$, and parameters $\epsilon > 0$, $\sigma \in (0,1)$, $\beta \in (0,1)$, $\lambda \in (0,1)$, $\bar{\mu} \in [\|A\|^2/(\lambda(1-\sigma)), \infty)$. Set $k := 0$.

Step 1. Choose $\mu_k \in [\|A\|^2/(\lambda(1-\sigma)), \bar{\mu}]$, and a positive semidefinite matrix $G_k \in R^{n\times n}$, then find $\tilde{x}^k \in R^n$ by solving the following system of linear equations

$$f(x^k) - A^\top y^k + (G_k + \mu_k I)(\tilde{x}^k - x^k) = r^k, \tag{14}$$

such that

$$\|r^k\| \leqslant \sigma \|x^k - \tilde{x}^k\|. \tag{15}$$

Step 2. Find $\bar{x}^k = x^k + t_k(\tilde{x}^k - x^k)$, such that

$$(f(\bar{x}^k) - A^\top y^k)^\top (x^k - \tilde{x}^k) \geqslant \lambda(1-\sigma)\mu_k \|x^k - \tilde{x}^k\|^2, \tag{16}$$

where $t_k = \beta^{m_k}$ and m_k is the smallest nonnegative integer such that (16) is satisfied.

Step 3. Set

$$\bar{y}^k = P_Y[y^k - (A\bar{x}^k - a)]. \tag{17}$$

If

$$\|x^k - \bar{x}^k\| + \|y^k - \bar{y}^k\| \leqslant \epsilon,$$

then stop. Otherwise, go to Step 4.

Step 4. Set

$$g(u^k) = g(x^k, y^k) = \begin{pmatrix} f(\bar{x}^k) - A^\top \bar{y}^k \\ y^k - \bar{y}^k \end{pmatrix}. \tag{18}$$

Then compute α_k by

$$\alpha_k = g(u^k)^\top (u^k - \bar{u}^k)/\|g(u^k)\|^2. \tag{19}$$

Step 5. Compute $u^{k+1}=(x^{k+1},y^{k+1})$ via

$$u^{k+1}=u^k-\alpha_k g(u^k). \tag{20}$$

Set $k:=k+1$ and goto Step 1.

Note that solving (14) is equivalent to solving a system of linear equations with the following structure

$$(G_k+\mu_k I)x=d^k.$$

Since G_k is positive semidefinite and $\mu_k \geqslant \|A\|^2/(\lambda(1-\sigma))>0$, the coefficient matrix of the above system is positive definite. For solving this system of linear equations, there are many iterative methods, see for example, [4, 6, 11]. Furthermore, we allow to solve it approximately with accuracy criterion (15), which is more constructive than the classical one assuming summability or square summability of the sequence of error tolerance, see [25, 15].

Note also that the system of linear equations is much easier to solve than variational inequality problems (3) and (4), and the system of nonlinear equations (6) and (7).

If $x^k=\tilde{x}^k$, then it follows from (15) that $r^k=0$. On the same time, if $y^k=\bar{y}^k$, then from (14) and (17), we have

$$f(x^k)-A^\top y^k=0,$$

and

$$y^k=P_Y[y^k-(Ax^k-a)].$$

Then, it follows from Lemma 2.2 that (x^k,y^k) is a solution of VI(F,Ω). On the other hand, if (x^k,y^k) is a solution of VI(F,Ω), then we have $x^k=\tilde{x}^k$, $y^k=\bar{y}^k$. That is

$$(x^k,y^k) \text{ is a solution of VI}(F,\Omega) \Leftrightarrow \|x^k-\tilde{x}^k\|=\|y^k-\bar{y}^k\|=0.$$

We thus can use $\|x^k-\tilde{x}^k\|+\|y^k-\bar{y}^k\|$ as a measure, which measures how much that (x^k,y^k) fails to be a solution of VI(F,Ω). The stopping criterion in Step 3 is thus reasonable.

The following lemma shows that the whole algorithm is well defined.

LEMMA 3.1. *Algorithm* 3.1 *is well defined*

Proof. To show that Algorithm 3.1 is well-defined, we need only to show that the linesearch procedure is well-defined. If $\tilde{x}^k=x^k$ and $y^k=\bar{y}^k$, then the algorithm terminates with $u^k=(\tilde{x}^k,y^k)$ a solution of VI(F,Ω). Therefore, from now on, we assume that $\|\tilde{x}^k-x^k\|>0$. Since $\tilde{x}^k$ is a solution of (14), it follows that.

$$\begin{aligned}
&(f(x^k)-A^\top y^k)^\top(x^k-\tilde{x}^k)\\
&=(x^k-\tilde{x}^k)^\top(G_k+\mu_k I)(x^k-\tilde{x}^k)+(x^k-\tilde{x}^k)^\top r^k\\
&\geqslant \mu_k\|x^k-\tilde{x}^k\|^2-\|x^k-\tilde{x}^k\|\|r^k\|\\
&\geqslant(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2,
\end{aligned} \tag{21}$$

where the first inequality follows from the the Cauchy-Schwarz inequality and the last one follows from (15). Suppose that for some $k \geqslant 0$, (16) is not satisfied for any m, i.e.,

$$(f(x^k+\beta^m(\tilde{x}^k-x^k))-A^\top y^k)^\top(x^k-\tilde{x}^k)<\lambda(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2, \forall m.$$

Since $x^k+\beta^m(\tilde{x}^k-x^k)\to x^k$ as $m\to\infty$, and $f(\cdot)$ is continuous, taking limit as $m\to\infty$, we obtain

$$(f(x^k)-A^\top y^k)^\top(x^k-\tilde{x}^k)\leqslant\lambda(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2.$$

This together with (21) means that $\|x^k-\tilde{x}^k\|=0$ since $\lambda\in(0,1)$, which contradicts the assumption that $\|x^k-\tilde{x}^k\|>0$. Thus, (16) always terminates with a positive stepsize t_k. This complements the proof. □

4. Global Convergence

In this section, we analyze the global convergence of the proposed algorithm under the mild conditions that the underlying mapping f is continuous and monotone and the solution set of VI(F,Ω) (12)–(13), denoted by Ω^*, is nonempty.

If the algorithm stops at some iterative $k>0$, then (x^k,y^k) is an approximate solution of VI(F,Ω). We thus assume throughout this section that $\epsilon=0$ and the algorithm generates an infinite sequence $\{u^k\}=\{(x^k,y^k)\}$.

We now begin our analysis with two lemmas.

LEMMA 4.1. *If $u^k=(x^k,y^k)$ is not a solution of VI(F,Ω), then $-g(u^k)$ is a descent direction of the merit function $\frac{1}{2}\|u-u^*\|^2$, where $u^*\in\Omega^*$ is an arbitrary solution of VI(F,Ω).*

Proof. First, note that Y is a nonempty closed convex subset of R^{m+n}. Let $u^*=(x^*,y^*)\in\Omega^*$ be an arbitrary solution of VI(F,Ω). Then, from Lemma 2.1

$$\{y^k-(A\bar{x}^k-a)-P_Y[y^k-(A\bar{x}^k-a)]\}^\top\{P_Y[y^k-(A\bar{x}^k-a)]-y^*\}\geqslant 0.$$

Since u^* is a solution of VI(F,Ω) and $P_Y[\cdot]\in Y$, from (12), it follows that

$$(Ax^*-a)^\top(P_Y[y^k-(A\bar{x}^k-a)]-y^*)\geqslant 0.$$

Adding the above two inequalities,

$$\{y^k-\bar{y}^k-A(\bar{x}^k-x^*)\}^\top\{(y^k-y^*)-(y^k-\bar{y}^k)\}\geqslant 0, \tag{22}$$

which is equivalent to the inequality

$$\begin{aligned}&(x^k-x^*)^\top(A^\top(y^k-\bar{y}^k))+(y^k-y^*)^\top(y^k-\bar{y}^k)\\&\geqslant(y^k-y^*)^\top(A\bar{x}^k-Ax^*)+\|y^k-\bar{y}^k\|^2-(A\bar{x}^k-Ax^k)^\top(y^k-\bar{y}^k).\end{aligned} \tag{23}$$

Since u^* is a solution of VI(F,Ω), we obtain that

$$f(x^*) = A^\top y^*.$$

From the monotonicity of f,

$$\begin{aligned}
&(f(\bar{x}^k) - A^\top y^k)^\top (\bar{x}^k - x^*)\\
&= ((f(\bar{x}^k) - f(x^*)) - A^\top (y^k - y^*))^\top (\bar{x}^k - x^*)\\
&\geqslant -(y^k - y^*)^\top (A\bar{x}^k - Ax^*). \qquad (24)
\end{aligned}$$

Adding (23) and (24),

$$\begin{aligned}
&(x^k - x^*)^\top (f(\bar{x}^k) - A^\top \bar{y}^k) + (y^k - y^*)^\top (y^k - \bar{y}^k)\\
&\geqslant (x^k - \bar{x}^k)^\top (f(\bar{x}^k) - A^\top y^k) + \|y^k - \bar{y}^k\|^2 - (A\bar{x}^k - Ax^k)^\top (y^k - \bar{y}^k)\\
&\geqslant \lambda(1-\sigma)\mu_k t_k \|x^k - \tilde{x}^k\|^2 + \|y^k - \bar{y}^k\|^2 - \frac{1}{2} t_k^2 \|A\|^2 \|x^k - \tilde{x}^k\|^2 - \frac{1}{2}\|y^k - \bar{y}^k\|^2\\
&\geqslant \frac{\|A\|^2}{2} t_k \|x^k - \tilde{x}^k\|^2 + \frac{1}{2}\|y^k - \bar{y}^k\|^2, \qquad (25)
\end{aligned}$$

where the second inequality follows from (16) and the Cauchy-Schwarz inequality, and the last one follows from the choice of μ_k and the fact that $t_k \in (0,1]$. This completes the proof. □

The following lemma paves the way to prove the global convergence of the proposed algorithm.

LEMMA 4.2. *Suppose that f is continuous and monotone, the solution set Ω^* of VI(F,Ω) is nonempty. Then*

1. *The generated sequence $\{u^k\} = \{(x^k, y^k)\}$ is bounded.*
2. *The sequence $\{\bar{u}^k\} = \{(\bar{x}^k, \bar{y}^k)\}$ is bounded.*

Proof. It follows from (25) that

$$\begin{aligned}
&g(u^k)^\top (u^k - \bar{u}^k)\\
&= (f(\bar{x}^k) - A^\top y^k)^\top (x^k - \bar{x}^k) + \|y^k - \bar{y}^k\|^2 - (A\bar{x}^k - Ax^k)^\top (y^k - \bar{y}^k)\\
&\geqslant \frac{\|A\|^2}{2} t_k \|x^k - \tilde{x}^k\|^2 + \frac{1}{2}\|y^k - \bar{y}^k\|^2.
\end{aligned}$$

Since $f(x^*) = A^\top y^*$, it follows therefore,

$$\begin{aligned}
&g(u^k)^\top(\bar{u}^k - u^*)\\
&= \begin{pmatrix} f(\bar{x}^k) - A^\top \bar{y}^k \\ y^k - \bar{y}^k \end{pmatrix}^\top \begin{pmatrix} \bar{x}^k - x^* \\ \bar{y}^k - y^* \end{pmatrix}\\
&= \begin{pmatrix} (f(\bar{x}^k) - f(x^*)) - A^\top(\bar{y}^k - y^*) \\ y^k - \bar{y}^k \end{pmatrix}^\top \begin{pmatrix} \bar{x}^k - x^* \\ \bar{y}^k - y^* \end{pmatrix}\\
&= (\bar{x}^k - x^*)^\top(f(\bar{x}^k) - f(x^*)) + (y^k - \bar{y}^k - (A\bar{x}^k - Ax^*))^\top(\bar{y}^k - y^*)\\
&\geqslant (y^k - \bar{y}^k - (A\bar{x}^k - Ax^*))^\top(\bar{y}^k - y^*)\\
&\geqslant 0,
\end{aligned}$$

where the first inequality follows from the monotonicity of f and the last one follows from (22). We thus have that

$$\begin{aligned}
g(u^k)^\top(u^k - u^*) &= g(u^k)^\top(\bar{u}^k - u^*) + g(u^k)^\top(u^k - \bar{u}^k)\\
&\geqslant g(u^k)^\top(u^k - \bar{u}^k).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2\\
&\leqslant \|x^k - x^*\|^2 + \|y^k - y^*\|^2 - 2\alpha_k g(u^k)^\top(u^k - \bar{u}^k) + \alpha_k^2\|g(u^k)\|^2\\
&= \|x^k - x^*\|^2 + \|y^k - y^*\|^2 - \alpha_k g(u^k)^\top(u^k - \bar{u}^k)\\
&= \|x^k - x^*\|^2 + \|y^k - y^*\|^2 - \frac{(g(u^k)^\top(u^k - \bar{u}^k))^2}{\|g(u^k)\|^2}.
\end{aligned} \tag{26}$$

From the above inequality, we have

$$\|x^{k+1} - x^*\|^2 + \|y^{k+1} - y^*\|^2 \leqslant \cdots \leqslant \|x^0 - x^*\|^2 + \|y^0 - y^*\|^2.$$

Thus, the sequence $\{u^k\} = \{(x^k, y^k)\}$ is bounded.

It follows from (14) that

$$(G_k + \mu_k I)(\tilde{x}^k - x^k) - r^k = f(x^k) - A^\top y^k.$$

Since G_k is a positive semidefinite matrix and μ_k is bounded away from zero, the sequence $\{(x^k, y^k)\}$ is bounded and f is continuous, from (15) $\{\tilde{x}^k\}$ is bounded. Therefore, $\{\bar{x}^k\}$ is bounded and from (17), $\{\bar{y}^k\}$ is also bounded. □

We are now ready to prove the main result in this section.

THEOREM 4.3. *Let the assumptions in Lemma* 4.2 *hold. Then the whole sequence* $\{u^k\}$ *generated by Algorithm* 3.1 *converges to a solution of* $VI(F, \Omega)$ *globally.*

Proof. Lemma 4.2 shows that $\{u^k\} = \{(x^k, y^k)\}$ is bounded. It thus has at least one cluster point. Let $\tilde{u} = (\tilde{x}, \tilde{y})$ be a cluster point of $\{u^k\} = \{(x^k, y^k)\}$ and $\{u^{k_j}\} = \{(x^{k_j}, y^{k_j})\}$ be the corresponding subsequence converging to $\tilde{u}$. Since $\{u^k\}$

is bounded, from the continuity of g, there exists a positive constant M, such that $\|g(u^k)\| \leqslant M$, for all $k \geqslant 0$. Then, from (26)

$$\sum_{k=0}^{\infty}(g(u^k)^\top(u^k-\bar{u}^k))^2<+\infty.$$

Thus, from (25)

$$\lim_{k\to\infty} t_k\|x^k-\tilde{x}^k\|=\lim_{k\to\infty}\|y^k-\bar{y}^k\|=0. \tag{27}$$

We consider the two possible cases. Firstly, suppose that

$$\limsup_{k\to\infty} t_k>0.$$

From (29) we have that

$$\liminf_{k\to\infty}\|x^k-\tilde{x}^k\|=\lim_{k\to\infty}\|y^k-\bar{y}^k\|=0.$$

Thus, from (15)

$$\liminf_{k\to\infty} r^k=0.$$

Since f and the projection operator P_Y are continuous, taking limit along the subsequence, it follows from (14) that

$$f(\tilde{x})-A^\top\tilde{y}=0,$$

and

$$\tilde{y}=P_Y[\tilde{y}-(A\tilde{x}-a)],$$

which mean that $\tilde{u}$ is a solution of VI(F,Ω). Since u^* is an arbitrary solution, we can just take $u^*=\tilde{u}$ in the above analysis and

$$\|u^{k+1}-\tilde{u}\|\leqslant\|u^k-\tilde{u}\|.$$

Using the same argument as given in [25], one can show easily that the whole sequence $\{u^k\}$ converges to $\tilde{u}$, a solution of VI(F,Ω).

Now, we consider the other possible case that

$$\lim_{k\to\infty} t_k=0.$$

We will show that in this case, $\|x^k-\tilde{x}^k\|$ also tends to 0. Suppose that

$$\|x^{k_l}-\tilde{x}^{k_l}\|>d>0$$

for an infinite subsequence $\{k_l\}$, and (without loss of generality) that $\{x^{k_l}\}$ and $\{\tilde{x}^{k_l}\}$ are convergent subsequences. By the choice of t_k we know that (16) was not satisfied for m_k-1. That is,

$$(f(x^k+\beta^{m_k-1}(\tilde{x}^k-x^k))-A^\top y^k))^\top(x^k-\tilde{x}^k)<\lambda(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2,$$

which means, together with (21), that (for k large enough such that $m_k > 1$)

$$(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2+(f(x^k+\beta^{m_k-1}(\tilde{x}^k-x^k))- \\ f(x^k))^\top(x^k-\tilde{x}^k)<\lambda(1-\sigma)\mu_k\|x^k-\tilde{x}^k\|^2. \tag{28}$$

Then, taking limit along the subsequence $\{k_l\}$ and using the continuity of f again, we have

$$(1-\sigma)\|d\|^2\leqslant\lambda(1-\sigma)\|d\|^2.$$

Since $\lambda\in(0,1)$, we have that $\lim_{k\to\infty}\|x^k-\tilde{x}^k\|=0$. By a similar analysis as the first case, we can show that $\tilde{u}=(\tilde{x},\tilde{y})$ is a solution of VI(F,Ω) and the whole sequence $\{u^k\}$ converges to $\tilde{u}$, a solution of VI(F,Ω). This completes the proof.

□

5. Numerical Results

To test the ability of the proposed algorithm, in this section, we implement it in MATLAB to solve variational inequality problems with linear constraints. The examples used here are taken from the test problems of Taji, Fukushima and Ibaraki [27], which are modifications of the test problems of Marcotte and Dussault [23]. The constraint set S and the mapping f are taken respectively as

$$S=S_2=\left\{x\in R^5\,\Big|\,\sum_{i=1}^{5}x_i\geqslant 10,\ x_i\geqslant 0,\ i=1,2,\cdots,5\right\}$$

and

$$f(x)=Mx+\rho C(x)+q,$$

where M is a 5×5 asymmetric positive definite matrix and $C_i(x)=\arctan(x_i-2)$, $i=1,2,\cdots,5$. The parameter ρ is used to vary the degree of asymmetry and nonlinearity. The data of this example are given as follows.

$$f(x)=\begin{pmatrix} 0.726 & -0.949 & 0.266 & -1.193 & -0.504\\ 1.645 & 0.678 & 0.333 & -0.217 & -1.443\\ -1.016 & -0.225 & 0.769 & 0.934 & 1.007\\ 1.063 & 0.567 & -1.144 & 0.550 & -0.548\\ -0.259 & 1.453 & -1.073 & 0.509 & 1.026 \end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}$$

$$+\rho\begin{pmatrix}\arctan(x_1-2)\\ \arctan(x_2-2)\\ \arctan(x_3-2)\\ \arctan(x_4-2)\\ \arctan(x_5-2)\end{pmatrix}+\begin{pmatrix}5.308\\0.008\\-0.938\\1.024\\-1.312\end{pmatrix}.$$

Thus,

$$B=(1,1,1,1,1),\qquad \text{and}\qquad b=10.$$

Table 1. Numerical results for $\rho = 10$

Starting pint	Method	Num. of Iter.	CPU(sec.)	$\|x^k - x^*\|$
(25,0,0,0,0)	NDM	27	0.11	1.34×10^{-7}
	GCNM	7	0.50	4.58×10^{-7}
(10,0,10,0,10)	NDM	29	0.16	2.55×10^{-7}
	GCNM	6	0.44	1.16×10^{-7}
(10,0,0,0,0)	NDM	25	0.11	7.73×10^{-7}
	GCNM	7	0.49	4.56×10^{-7}
(0,2.5,2.5,2.5,2.5)	NDM	21	0.06	2.93×10^{-7}
	GCNM	6	0.44	1.01×10^{-7}
(0,0,0,0,0)	NDM	19	0.06	1.85×10^{-7}
	GCNM	7	0.55	2.54×10^{-7}
(1,1,1,1,1)	NDM	22	0.06	5.73×10^{-7}
	GCNM	5	0.28	9.48×10^{-7}

In our formulation (12)–(13),

$$A = \begin{pmatrix} 1\,1\,1\,1\,1 \\ 1\,0\,0\,0\,0 \\ 0\,1\,0\,0\,0 \\ 0\,0\,1\,0\,0 \\ 0\,0\,0\,1\,0 \\ 0\,0\,0\,0\,1 \end{pmatrix}, \qquad a = \begin{pmatrix} 10 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad Y = R^6_+.$$

The problem has a unique solution $x^* = (2,2,2,2,2)^\top$ and the Lagrange multiplier is $y^* = (2,0,0,0,0,0)^\top$. Since for any given positive parameter $\delta > 0$, VI(F, Ω) is equivalent to VI($\delta F, \Omega$), we scale the mapping F by a factor $\delta = 0.15$.

At each iteration, we choose $G_k = \nabla f(x^k)$ and use the function inv.m from MATLAB to find the inverse matrix of $G_k + \mu_k I$. In this sense, the subproblem (14) is solved exactly and $\sigma = 0$. The other parameters used in the algorithm are set as $\lambda = 0.95$, $\beta = 0.6$ and $\mu_k \equiv 2.5$ for all k. The stop parameter is set to be $\epsilon = 10^{-6}$. For comparison, we also code the globally convergent Newton method (GCNM) of Taji, Fukushima and Ibaraki [27]. We use the quadratic-program solver quadprog.m from the MATLAB optimization toolbox to perform the projection to the feasible set S. We rewrite the subproblem in [27] as a linear complementarity problem (LCP) and solve it by Lemke's complementarity pivoting method [5], which finds a solution of LCP in a finite number of steps. The parameters used in their algorithm are set the same as those in [27].

Table 5.1 and 5.2 report the computational results for $\rho = 10$ and 20, respectively. For simplicity, we denote the proposed method by NDM.

Table 2. Numerical results for $\rho = 20$

Starting pint	Method	Num. of Iter.	CPU(sec.)	$\|x^k - x^*\|$
(25,0,0,0,0)	NDM	39	0.16	7.17×10^{-7}
	GCNM	7	0.55	4.58×10^{-7}
(10,0,10,0,10)	NDM	41	0.22	8.12×10^{-7}
	GCNM	6	0.54	1.16×10^{-7}
(10,0,0,0,0)	NDM	36	0.16	5.75×10^{-7}
	GCNM	7	0.48	9.74×10^{-7}
(0,2.5,2.5,2.5,2.5)	NDM	31	0.11	6.86×10^{-7}
	GCNM	6	0.45	6.31×10^{-7}
(0,0,0,0,0)	NDM	34	0.11	5.71×10^{-7}
	GCNM	7	0.50	8.27×10^{-7}
(1,1,1,1,1)	NDM	32	0.11	6.39×10^{-7}
	GCNM	5	0.33	6.58×10^{-7}

The results in Table 5.1 and 5.2 indicate that the new decomposition algorithm is quite efficient. Though the iterative number is larger than Newton-type method [27], the total CPU time is smaller. Especially, the computational cost at each iteration is much smaller, since, at each iteration, the Newton-type method [27] needs to make some projections to the feasible set S, which is more difficult than making projections to the nonnegative orthant of R^{n+m} and, one needs to solve a linear variational inequality problem at each iteration, which is also time consuming from the computational point of view.

The same problem with $\rho = 10$ was also considered in [28]. At each iteration, their algorithm solves a system of nonlinear equations with the structure as (6), and this subproblem has to been solved exactly. Additional to this, one has also to solve a linear variational inequality problem (5) to get y^k. Though this problem can be solved by the standard Lemke Algorithm [5], it is still time consuming, see Table 1 in [28].

To show the advantage of this decomposition method for large scale problems, we implement it to a set of spatial price equilibrium problems. The details of these problems follow from [22, 12], as in the following:

$$\begin{array}{lll} \min & \sum_{i=1}^{m}\sum_{j=1}^{n}\left(c_{ij}x_{ij}+\frac{1}{2}h_{ij}x_{ij}^2\right), & \\ \text{s.t.} & \sum_{j=1}^{n}x_{ij}=s_i, & i=1,\ldots,m, \\ & \sum_{i=1}^{m}x_{ij}=d_j, & j=1,\ldots,n, \\ & x_{ij}\geqslant 0, & \end{array}$$

where

Table 3. Number of iterations for different scale and precisions.

m	n	mn	$\epsilon=0.1$	$\epsilon=10^{-2}$	$\epsilon=10^{-3}$	$\epsilon=10^{-4}$
5	5	25	6	14	71	92
5	10	50	18	84	132	173
5	20	100	13	25	96	125
10	10	100	9	31	151	203
10	20	200	9	23	160	194
20	30	600	14	36	170	470
30	40	1200	17	41	210	376
40	50	2000	20	63	271	767
50	60	3000	28	57	514	774

– s_i = the supply amount on the i^{th} supply market, $i = 1, \dots, m$, and
– d_j = the demand amount on the j^{th} demand market, $j = 1, \dots, n$.

We use the same cost function as in He and Zhou [22]:

$$c_{ij} \in (0, 100) \quad \text{and} \quad h_{ij} \in (0.005, 0.01).$$

The parameters s_i and d_j are generated randomly in $(0, 100)$ for all $i=1,\dots,m$ and $j=1,\dots,n$. We set $G_k \equiv H$ and $\mu \equiv 15$ and the other parameters are set as the first example. Thus, the algorithm avoids solving the system of linear equations and each iteration is very simple to evaluate. The calculations were started with u^0 generated randomly in $(0, 100)$ and stopped for some prescribed $\epsilon > 0$. The computational results are given in Table 5.3 for some m and n.

The results in Table 5.3 show that the required iterative numbers are relatively small as compared with the size of problems. As this decomposition method only requires function evaluations per iteration, it is attractive from a computational point of view.

6. Concluding remarks

In this paper, we proposed a new decomposition algorithm for solving variational inequality problems with linear equality constraints or inequality constraints in a uniform framework. At each iteration, the algorithm solves a system of linear equations with respect to x, the primal variable, and then performs a line search step to get a suitable step size. Furthermore, we allow to solve the subproblem approximately with a constructive accuracy criterion. The algorithm is thus well comparable to the original decomposition algorithms, which solve a series of variational inequality problems, a class of problems that are structurally much more difficult to solve than system of equations. The proposed algorithm is also well comparable to [16], which solves VI(f, S_1) and VI(f, S_2) by solving a series of system of nonlinear equations, and [28], which solves this class of variational inequality problems by solving a series of system of nonlinear equations, as well as a series of linear variational inequality problems.

Note that if we take $G_k \equiv G$ and $\mu_k \equiv \mu$, then at each iteration, the algorithm will avoid solving the system of linear equations, and the cost at each iteration is small as those in [12, 13, 14, 22]. However, this will increase the total iterative number and the line search steps. Therefore, their is a trade-off between the cost of solving the system of linear equations and the total cost of the algorithm.

Acknowledgment

The authors would like to thank the two anonymous referees for their constructive suggestions which led to improvements of the paper. The first author is grateful to Prof. B.S. He from Nanjing University and Prof. W.Y. Sun from Nanjing Normal University for their constant help.

References

1. Bertsekas, D.P. and Gafni, E.M. (1982), Projection method for variational inequalities with applications to the traffic assignment problem, *Mathematical Programming Study* 17, 139–159.
2. Bertsekas, D.P. and Tsitsiklis, J.N. (1989), *Parallel Distributed Computation, Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ.
3. Dafermos, S. (1980), Traffic equilibrium and variational inequalities, *Transportation Science* 14, 42–54.
4. Dennis, J.E. and Schnabel, R.B. (1983), *Numerical Methods for unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ.
5. Eaves, B.C. (1978), Computing stationary points, *Mathematical Programming Study* 7, 1–14.
6. Fletcher, R. (1985) *Practical Methods of Optimization*, Wiley, New York.
7. Fortin, M and Glowinski, R. (1983), *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Valued Problems*, North-Holland, Amsterdam.
8. Gabay, D. (1983), Applications of the method of multipliers to variational inequalities. In: Fortin, M. and Glowinski R. (eds), *Augmented Lagrangian Methods: Applications to the Solution of Boundary-Valued Problems*, North-Holland, Amsterdam, 299–331.
9. Gabay, D. and Mercier, B. (1976), A dual algorithm for the solution of nonlinear variational problems via finite-element approximations, *Computers and Mathematics with Applications* 2(1), 17–40.
10. Glowinski, R. (1984), *Numerical Methods for Nonlinear Variational Problems*, Springer, New York.
11. Golub G. and Van Loan C. (1989), *Matrix Computations*, 2nd ed., Johns-Hopkins, Baltimore, MD.
12. Han, D.R. and Lo, H.K. (2002), A new alternating direction method for a class of nonlinear variational inequality problems, *Journal of Optimization Theory and Applications* 112, 549–560.

13. Han, D.R. and Lo, H.K. (2002), A new stepsize rule in He and Zhou's alternating direction method, *Applied Mathematics Letters* 15, 181–185.
14. Han, D.R. (2002), A modified alternating direction method for Variational Inequality Problems, *Applied Mathematics and Optimization* 45, 630-74.
15. Han, D.R. and He, B.S. (2001), A new accuracy criterion for approximate proximal point methods, *Journal of Mathematical Analysis and Applications* 263, 241–254.
16. Han, D.R. (2002), A proximal decomposition algorithm for variational inequality problems. Manuscript.
17. Harker, P.T. and Pang, P.S. (1990), Finite-dimensional variational inequality and complementarity problems: a survey of theory, algorithms and applications, *Mathematical Programming* 48, 161–220.
18. He, B.S. (1994), A new method for a class of linear variational inequalities, *Mathematical Programming* 66, 137–144.
19. He, B.S. (1994), Solving a class of linear projection equations, *Numerische Mathematik* 68, 71–80.
20. He, B.S. (1997), A class of new methods for monotone variational inequalities, *Applied Mathematics and Optimization* 35, 69–76.
21. He, B.S. (1999), Inexact implicit methods for monotone general variational inequalities, *Mathematical Programming* 86, 199–217.
22. He, B.S. and Zhou, J. (2000), A modified alternation direction method for convex minimization problems, *Applied Mathematics Letters* 13, 122–130.
23. Marcotte, P. and Dussault, J.P. (1987), A note on a globally convergent Newton method for solving variational inequalities, *Operations Research Letters*, 6, 35–42.
24. Nagurney, A. and Ramanujam, P. (1996), Transportation network policy modeling with goal targets and generalized penalty functions, *Transportation Science* 30, 3–13.
25. Rockafellar, R.T. (1976), Monotone operators and the proximal point algorithm, *SIAM, Journal on Control and Optimization* 14, 877–898.
26. Solodov, M.V. and Svaiter, B.F. (1999), A hybrid approximate extragradient-proximal point algorithm using the enlargement of a maximal monotone operator, *Set Valued Analysis* 7, 323–345.
27. Taji, K., Fukushima, M. and Ibaraki, T. (1993), A globally convergent Newton method for solving strongly monotone variational inequalities, *Mathematical Programming* 58, 369–383.
28. Wang, S.L., Yang, H. and He, B.S. (2000), Solving a class of asymmetric variational inequalities via a new alternating direction method, *Computers and Mathematics with Applications* 40, 927–937.