Applicable Analysis and Discrete Mathematics
available online at http://pefmath.etf.rs

Appl. Anal. Discrete Math. **12** (2018), 224–243.
https://doi.org/10.2298/AADM1801224B

SQP ALTERNATING DIRECTION METHOD WITH A NEW OPTIMAL STEP SIZE FOR SOLVING VARIATIONAL INEQUALITY PROBLEMS WITH SEPARABLE STRUCTURE

*Abdellah Bnouhachem** *and Themistocles M.Rassias*

This paper is dedicated to Taha Bnouhachem

In this paper, we suggest and analyze a new alternating direction scheme for the separable constrained convex programming problem. The theme of this paper is twofold. First, we consider the square-quadratic proximal (SQP) method. Next, by combining the alternating direction method with SQP method, we propose a descent SQP alternating direction method by using the same descent direction as in [**6**] with a new step size α_k. Under appropriate conditions, the global convergence of the proposed method is proved. We show the $O(1/t)$ convergence rate for the SQP alternating direction method. Some preliminary computational results are given to illustrate the efficiency of the proposed method.

1. Introduction

Let $\mathcal{R}$ stand for the real axis; and $\mathcal{R}_+ = \{x \in \mathcal{R}; x \geq 0\}, \mathcal{R}_{++} = \{x \in \mathcal{R}; x > 0\}$, denote the positive half-axis and strict positive half-axis, respectively.
Further, given $n \in \mathcal{N}$, put

$$\mathcal{R}^n_+ = \{x = (x_1, \ldots, x_n)^\top; x_1, \ldots, x_n \in \mathcal{R}_+\}$$

*Corresponding author. Abdellah Bnouhachem
2010 Mathematics Subject Classification. 49J40, 65N30.
Keywords and Phrases. Variational inequalities, logarithmic-quadratic proximal method, square-quadratic proximal method, projection method, alternating direction method.

and

$$\mathcal{R}^n_{++} = \{x = (x_1, \dots, x_n)^\top; x_1, \dots, x_n \in \mathcal{R}_{++}\}$$

where $(.)^\top$ denotes the transpose.
This paper considers the constrained convex programming problem with the following separate structure:

$$\min\left\{\theta_1(x) + \theta_2(y) | Ax + By = b, x \in \mathcal{R}^n_+, y \in \mathcal{R}^m_+\right\}, \tag{1}$$

where $\theta_1 : \mathcal{R}^n_+ \to \mathcal{R}$ and $\theta_2 : \mathcal{R}^m_+ \to \mathcal{R}$ are closed proper convex functions not necessarily smooth, $A \in \mathcal{R}^{l\times n}$, $B \in \mathcal{R}^{l\times m}$ are given matrices and $b \in \mathcal{R}^l$ is a given vector.
The alternating directions method (ADM for short) [**17, 18**] is an attractive tool for solving (1). The main advantage of the ADM is that the method can decompose a high-dimensional and complicated variational inequality problem into a series of low-dimensional subproblems, sometimes much easier than the original one. The ADM has been studied extensively in the theoretical frameworks of both Lagrangian functions and maximal monotone operators. To make the ADM more efficient and practical some strategies have been studied, for more details, one can refer [**7, 8, 9, 11, 14, 22, 25, 26, 31, 34**].

Let $\partial(.)$ denote the sub-gradient operator of a convex function, and $f(x) \in \partial\theta_1(x)$ and $g(y) \in \partial\theta_2(y)$ are the sub-gradient of $\theta_1(x)$ and $\theta_2(y)$, respectively. By attaching a Lagrange multiplier vector $\lambda \in \mathcal{R}^l$ to the linear constraint $Ax+By = b$, problem (1) can be written in terms of finding $w \in \mathcal{W}$ such that

$$(w' - w)^\top Q(w) \geq 0, \quad \forall\, w' \in \mathcal{W}, \tag{2}$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix} \quad Q(w) = \begin{pmatrix} f(x) - A^\top\lambda \\ g(y) - B^\top\lambda \\ Ax + By - b \end{pmatrix}, \quad \mathcal{W} = \mathcal{R}^n_+ \times \mathcal{R}^m_+ \times \mathcal{R}^l. \tag{3}$$

Problem (2)–(3) is referred to as *structural variational inequality* (in short, SVI).
We now have a variety of techniques to suggest and analyze various ADMs with logarithmic-quadratic proximal (LQP) regularization [**2, 3, 4, 5, 6, 10, 29, 31, 35**] for solving SVI. Recently, Li [**29**] and Bnouhachem [**3**] developed the following LQP method by applying the LQP terms to regularize the ADM subproblems: for a given $w^k = (x^k, y^k, \lambda^k) \in \mathcal{R}^n_{++}\times\mathcal{R}^m_{++}\times\mathcal{R}^l$, and $\mu \in (0, 1)$, the predictor $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$ is obtained via solving the following system:

$$f(x) - A^\top\left[\lambda^k - H(Ax + By^k - b)\right] + R\left[(x - x^k) + \mu(x^k - X_k^2 x^{-1})\right] = 0, \tag{4}$$

$$g(y) - B^\top\left[\lambda^k - H(Ax + By - b)\right] + S\left[(y - y^k) + \mu(y^k - Y_k^2 y^{-1})\right] = 0, \tag{5}$$

$$\tilde{\lambda}^k = \lambda^k - H(Ax + By - b), \tag{6}$$

where $H \in \mathcal{R}^{l\times l}$, $R \in \mathcal{R}^{n\times n}$, and $S \in \mathcal{R}^{m\times m}$ are symmetric positive definite, $X_k := \mathrm{diag}(x_1^k, x_2^k, \cdots, x_n^k)$, $Y_k := \mathrm{diag}(y_1^k, y_2^k, \cdots, y_n^k)$, x^{-1} be an n-vector whose j-th element is $1/x_j$, y^{-1} be an n-vector whose j-th element is $1/y_j$.
The main disadvantage of the methods in [**3, 29**] is that solving equation (5) requires the solution of equation (4). Hence, the ADMs are not eligible for parallel computing in the sense that the solutions of (4) and (5) cannot be obtained simultaneously. To overcome this difficulty, Bnouhachem and Hamdi [**6**] proposed a parallel descent LQP-ADM for solving SVI. The main advantage of the method in [**6**] is that the predictor is obtained via solving a system of non-linear equations in a parallel wise and the new iterate $w^{k+1}(\alpha_k) = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ is given by:

$$w^{k+1}(\alpha_k) = (1-\sigma)w^k + \sigma P_{\mathcal{W}}[w^k - \alpha_{1k} G_1^{-1} d(w^k, \tilde{w}^k)], \qquad \sigma \in (0,1)$$

where

$$\alpha_{1k} = \frac{\varphi_{1k}}{\|w^k - \tilde{w}^k\|_{G_1}^2},$$

$$\varphi_{1k} := \|w^k - \tilde{w}^k\|_{M_1}^2 + (\lambda^k - \tilde{\lambda}^k)^\top (A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)),$$

$$d(w^k, \tilde{w}^k) = \begin{pmatrix} f(\tilde{x}^k) - A^\top \tilde{\lambda}^k + A^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ g(\tilde{y}^k) - B^\top \tilde{\lambda}^k + B^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix}$$

and

$$G_1 = \begin{pmatrix} (1+\mu)R + A^\top HA & 0 & 0 \\ 0 & (1+\mu)S + B^\top HB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix},$$

$$M_1 = \begin{pmatrix} R + A^\top HA & 0 & 0 \\ 0 & S + B^\top HB & 0 \\ 0 & 0 & H^{-1} \end{pmatrix}.$$

In this paper, we suggest that the complementarity subproblems arising in ADM (4)-(5) could be regularized by the square quadratic proximal (SQP) regularization, from a given $w^k = (x^k, y^k, \lambda^k) \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$, $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$ is obtained via solving the following system:
(7a)
$$f(x) - A^\top[\lambda^k - H(\frac{1}{2}Ax^k + \frac{1}{2}Ax + By^k - b)] + R[\frac{1}{2}(x - x^k) + \mu(x^k - U_k(\sqrt{x})^{-1})] =: \xi_x^k \approx 0,$$

(7b)
$$g(y) - B^\top[\lambda^k - H(A\tilde{x}^k + \frac{1}{2}By + \frac{1}{2}By^k - b)] + S[\frac{1}{2}(y - v^k) + \mu(y^k - V_k(\sqrt{y})^{-1})] =: \xi_y^k \approx 0,$$

(7c)
$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b),$$

where

(8) $$|(w^k - \tilde{w}^k)^\top \xi^k| \le \eta \|w^k - \tilde{w}^k\|_M^2, \quad \eta \in (0,1),$$

$$M = \begin{pmatrix} \frac{1}{2}R & & \\ & \frac{1}{2}S & \\ & & \frac{1}{2}H^{-1} \end{pmatrix}, \tag{9}$$

$$\xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ 0 \end{pmatrix} = \begin{pmatrix} f(\tilde{x}^k) - f(x^k) - \frac{1}{2}A^\top HA(x^k - \tilde{x}^k) \\ g(\tilde{y}^k) - g(y^k) - \frac{1}{2}B^\top HB(y^k - \tilde{y}^k) \\ 0 \end{pmatrix} \tag{10}$$

U_k and V_k are positive definite diagonal matrices defined by

$$U_k = diag(x_1^k\sqrt{x_1^k}, ..., x_n^k\sqrt{x_n^k}) := \begin{pmatrix} x_1^k\sqrt{x_1^k} & & \\ & \ddots & \\ & & x_n^k\sqrt{x_n^k} \end{pmatrix}$$

and

$$V_k = diag(y_1^k\sqrt{y_1^k}, ..., y_n^k\sqrt{y_n^k}),$$

$(\sqrt{x})^{-1} \in \mathcal{R}^n_{++}$ is a vector whose j-th element is $1/\sqrt{x_j}$, $(\sqrt{y})^{-1} \in \mathcal{R}^m_{++}$ is a vector whose j-th element is $1/\sqrt{y_j}$.

Since (7a)–(7b) include both square and quadratic terms, the method is called the square-quadratic proximal (SQP) method, and (7a) and (7b) are called the SQP system of nonlinear equations (SQP system).

By combining the alternating direction method with SQP method, we propose a descent SQP alternating direction method for SVI. The predictor is obtained by solving (7a)–(7c) and the new iterate is obtained by using the same descent direction as in [**6**] with a new step size α_k. Global convergence of the proposed method is proved under certain assumptions. To illustrate the proposed method and demonstrate its efficiency, some applications and their numerical results are also provided. Our results can be viewed as significant extensions of the previously known results.

2. SQP alternating direction method

For any vector $u \in \mathcal{R}^n$, $\|u\|^2 = u^\top u$, $\|u\|_\infty = \max\{|u_1| \dots, |u_n|\}$. Let $D \in \mathcal{R}^{n\times n}$ be a symmetric positive definite matrix. We denote the D-norm of u by $\|u\|_D^2 = u^\top Du$.

The following lemma provides some basic properties of the projection operator onto a closed convex subset Ω of $\mathcal{R}^l$.

Lemma 1. Let Ω be a nonempty closed convex subset of $\mathcal{R}^l$. Denote by $P_\Omega(.)$ the projection on Ω with respect to the Euclidean norm, that is,

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

Then, we have the following inequalities.

$$(z - P_\Omega[z])^\top(P_\Omega[z] - v) \geq 0, \quad \forall\, z \in \mathcal{R}^l, v \in \Omega; \tag{11}$$

$$\|u - P_\Omega[z]\|^2 \le \|z-u\|^2 - \|z - P_\Omega[z]\|^2, \quad \forall\, z \in \mathcal{R}^l, u \in \Omega. \tag{12}$$

We introduce the following definitions which are useful in the sequel.

Definition 1. The mapping $T : \mathcal{R}^n \to \mathcal{R}^n$ is said to be

(a) monotone if
$$(Tx - Ty)^\top (x-y) \ge 0, \qquad \forall x, y \in \mathcal{R}^n;$$

(b) k-Lipschitz continuous if there exists a constant $k > 0$ such that
$$\|Tx - Ty\| \le k\|x-y\|, \qquad \forall x, y \in \mathcal{R}^n;$$

We make the following standard assumptions.

Assumption A. f is monotone and Lipschitz continuous on $\mathcal{R}^n_+$ with Lipschitz constant k_f and g is monotone and Lipschitz continuous on $\mathcal{R}^m_+$ with Lipschitz constant k_g.

Assumption B. The solution set of SVI, denoted by $\mathcal{W}^*$, is nonempty.

We propose the following SQP alternating direction method for solving SVI:

Algorithm 1.

Step 0. *The initial step:*
Given $\varepsilon > 0$, $\mu \in (0,1)$ and $w^0 = (x^0, y^0, \lambda^0) \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$. Set $k = 0$.

Step 1. *Prediction step:*
Compute $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k) \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$ by solving (7a)–(7c).

Step 2. *Convergence verification:*
If $\max\{\|x^k - \tilde{x}^k\|_\infty, \|y^k - \tilde{y}^k\|_\infty, \|\lambda^k - \tilde{\lambda}^k\|_\infty\} < \epsilon$, then stop.

Step 3. *Correction step:*
The new iterate $w^{k+1}(\alpha_k) = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ is given by:

$$w^{k+1}(\alpha_k) = (1-\sigma)w^k + \sigma P_{\mathcal{W}}[w^k - \alpha_k d_2(w^k, \tilde{w}^k)], \quad \sigma \in (0,1) \tag{13}$$

where

$$\alpha_k = \frac{\varphi(w^k, \tilde{w}^k)}{\|d_1(w^k, \tilde{w}^k)\|^2}, \tag{14}$$

$$\varphi(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|^2_M + (w^k - \tilde{w}^k)^\top G(w^k - \tilde{w}^k) + (w^k - \tilde{w}^k)^\top \xi^k, \tag{15}$$

$$d_1(w^k, \tilde{w}^k) = \begin{pmatrix} \frac{(1+\mu)}{2} R(x^k - \tilde{x}^k) - [f(x^k) - f(\tilde{x}^k)] \\ \frac{(1+\mu)}{2} S(y^k - \tilde{y}^k) - [g(y^k) - g(\tilde{y}^k)] + B^\top HA(x^k - \tilde{x}^k) \\ H^{-1}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix}$$

(16)

$$d_2(w^k, \tilde{w}^k) = \begin{pmatrix} f(\tilde{x}^k) - A^\top \tilde{\lambda}^k + A^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ g(\tilde{y}^k) - B^\top \tilde{\lambda}^k + B^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ A\tilde{x}^k + B\tilde{y}^k - b \end{pmatrix}$$

and

$$G = \begin{pmatrix} \frac{1}{2}A^\top HA & 0 & 0 \\ B^\top HA & \frac{1}{2}B^\top HB & 0 \\ A & B & \frac{1}{2}H^{-1} \end{pmatrix}.$$

Set $k := k+1$ and go to Step 1.

We need the following result to study the convergence analysis of the proposed method.

Lemma 2. Let $q(u) \in \mathcal{R}^n$ be a monotone mapping of u with respect to $\mathcal{R}^n_+$ and $R := \text{diag}(r_1, ..., r_n) \in \mathcal{R}^{n\times n}$ be a positive definite diagonal matrix. For a given $u^k > 0$, $\mu > 0$, if $U_k := \text{diag}(u_1^k\sqrt{u_1^k}, ..., u_n^k\sqrt{u_n^k})$, $\sqrt{u} = (\sqrt{u_1}, ..., \sqrt{u_n})$, and $(\sqrt{u})^{-1}$ be an n-vector whose j-th element is $1/\sqrt{u_j}$, then the equation

$$q(u) + R[\frac{1}{2}(u - u^k) + \mu(u^k - U_k(\sqrt{u})^{-1})] = 0 \tag{17}$$

has a unique positive solution u. Moreover, for any $v \geq 0$, we have

$$(v-u)^\top q(u) \geq \tfrac{1+\mu}{4}\left(\|u-v\|_R^2 - \|u^k - v\|_R^2\right) + \tfrac{1-\mu}{4}\|u^k - u\|_R^2. \tag{18}$$

Proof. The proof of the first assertion is similar with the one of Proposition 2 in [**1**]; hence it is omitted. We now prove the second assertion. For each $t > 0$, we have $\frac{1}{2}\left(1 - \frac{1}{t}\right) \leq 1 - \frac{1}{\sqrt{t}} \leq \frac{1}{2}(t-1)$, then we obtain after multiplication by $v_j u_j^k \geq 0$ for each $j = 1, ..., n$,

$$v_j u_j^k (1 - \frac{\sqrt{u_j^k}}{\sqrt{u_j}}) \leq v_j u_j^k \frac{1}{2}\left(\frac{u_j}{u_j^k} - 1\right) = \frac{1}{2} v_j (u_j - u_j^k)$$

and after multiplication by $u_j u_j^k \geq 0$ for each $j = 1, ..., n$,

$$-u_j u_j^k (1 - \frac{\sqrt{u_j^k}}{\sqrt{u_j}}) \leq u_j u_j^k \frac{1}{2}\left(\frac{u_j^k}{u_j} - 1\right) = \frac{1}{2} u_j^k (u_j^k - u_j),$$

adding the two inequalities, then we obtain

$$\begin{aligned}&(v_j - u_j)\left(\tfrac{1}{2}(u_j - u_j^k) + \mu\left(u_j^k - (\sqrt{u_j^k})^3(\sqrt{u_j})^{-1}\right)\right) \\ &\leq \tfrac{1}{2}\mu(v_j - u_j^k)(u_j - u_j^k) + \tfrac{1}{2}(u_j - u_j^k)(v_j - u_j).\end{aligned}$$

Using the identities

$$\frac{1}{2}(v_j - u_j^k)(u_j - u_j^k) = \frac{1}{4}\left((u_j - u_j^k)^2 - (u_j - v_j)^2 + (v_j - u_j^k)^2\right)$$

$$\frac{1}{2}(u_j - u_j^k)(v_j - u_j) = \frac{1}{4}\left((v_j - u_j^k)^2 - (v_j - u_j)^2 - (u_j - u_j^k)^2\right)$$

and recalling (17), we obtain

$$(u_j - v_j)(-q_j) \geq r_j \left(\frac{1+\mu}{4}\left((u_j - v_j)^2 - (u_j^k - v_j)^2\right) + \frac{1-\mu}{4}(u_j^k - u_j)^2\right).$$

Summing over $j = 1, ..., n$, gives (18). $\square$

In the next theorem, we show that α_k is lower bounded away from zero and it is useful for the convergence analysis.

Theorem 1. For given $w^k \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$, let $\tilde{w}^k$ be generated by (7a)-(7c), then there exist two constants $\alpha_1 > 0$ and $\alpha_2 > 0$ such that

$$\varphi(w^k, \tilde{w}^k) \geq \alpha_1 \|w^k - \tilde{w}^k\|^2 \tag{19}$$

and

$$\alpha_k \geq \frac{\alpha_1}{\alpha_2}. \tag{20}$$

Proof. For any vector $w := (x, y, \lambda)$, we have

$$\begin{aligned} w^\top G w &= \frac{1}{2}x^\top A^\top H A x + y^\top B^\top H A x + \frac{1}{2}y^\top B^\top H B y + \lambda^\top A x + \lambda^\top B y + \frac{1}{2}\lambda^\top H^{-1}\lambda \\ &= \frac{1}{2}\|Ax + By + H^{-1}\lambda\|_H^2 \\ &\geq 0. \end{aligned}$$

Then, from the definition of $\varphi(w^k, \tilde{w}^k)$, we have

$$\begin{aligned} \varphi(w^k, \tilde{w}^k) &\geq \|w^k - \tilde{w}^k\|_M^2 + (w^k - \tilde{w}^k)^\top \xi^k \\ &\geq \alpha_1 \|w^k - \tilde{w}^k\|^2, \end{aligned} \tag{21}$$

where $\alpha_1 > 0$ is a constant. It follows from (15) that

$$\begin{aligned} \|d_1(w^k, \tilde{w}^k)\| &\leq \frac{1+\mu}{2}\|R(x^k - \tilde{x}^k)\| + \|f(x^k) - f(\tilde{x}^k)\| \\ &\quad + \frac{1+\mu}{2}\|S(y^k - \tilde{y}^k)\| + \|g(y^k) - g(\tilde{y}^k)\| \\ &\quad + \|B^\top H A(x^k - \tilde{x}^k)\| + \|H^{-1}(\lambda^k - \tilde{\lambda}^k)\| \end{aligned} \tag{22}$$

$$\begin{aligned} &\leq \left(\frac{1+\mu}{2}\|R\| + k_f + \|B^\top H A\|\right)\|x^k - \tilde{x}^k\| \\ &\quad + \left(\frac{1+\mu}{2}\|S\| + k_g\right)\|y^k - \tilde{y}^k\| \\ &\quad + \|H^{-1}\|\|\lambda^k - \tilde{\lambda}^k\| \end{aligned} \tag{23}$$

$$\begin{aligned} &\leq \max\{\frac{1+\mu}{2}\|R\| + k_f + \|B^\top H A\|, \frac{1+\mu}{2}\|S\| + k_g, \|H^{-1}\|\} \\ &\quad \left(\|x^k - \tilde{x}^k\| + \|y^k - \tilde{y}^k\| + \|\lambda^k - \tilde{\lambda}^k\|\right) \\ &\leq \alpha_2 \|w^k - \tilde{w}^k\|^2, \end{aligned} \tag{24}$$

where $\alpha_2 > 0$ is a constant. Therefore, it follows from (14) and (19) that

$$\alpha_k \geq \frac{\alpha_1}{\alpha_2}$$

and this completes the proof. □

3. Basic results

In this section, we prove some basic properties, which will be used to establish the sufficient and necessary conditions for the convergence of the proposed method. The following results are due to applying Lemma 2 to the SQP systems in the prediction step of the proposed method.

Lemma 3. For given $w^k = (x^k, y^k, \lambda^k) \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$, let $\tilde{w}^k$ be generated by (7a)–(7c). Then for any $w = (x, y, \lambda) \in \mathcal{W}$, we have

(25)
$$(w - \tilde{w}^k)^\top d_2(w^k, \tilde{w}^k) \quad \geq \quad (w - \tilde{w}^k)^\top d_1(w^k, \tilde{w}^k) - \tfrac{\mu}{2}\|x^k - \tilde{x}^k\|^2_R - \tfrac{\mu}{2}\|y^k - \tilde{y}^k\|^2_S.$$

Proof. Applying Lemma 2 to (7a) by setting $u^k = x^k$, $u = \tilde{x}^k$, $v = x$ in (18) and

$$q(u) = f(\tilde{x}^k) - A^\top[\lambda^k - H(\frac{1}{2}Ax^k + \frac{1}{2}A\tilde{x}^k + By^k - b)] - \xi^k_x,$$

we get

$$\begin{aligned} &(x - \tilde{x}^k)^\top \Big\{ f(\tilde{x}^k) - A^\top[\lambda^k - H(\frac{1}{2}Ax^k + \frac{1}{2}A\tilde{x}^k + By^k - b)] - \xi^k_x \Big\} \\ \geq\ & \tfrac{1+\mu}{4}\Big(\|\tilde{x}^k - x\|^2_R - \|x^k - x\|^2_R\Big) + \tfrac{1-\mu}{4}\|x^k - \tilde{x}^k\|^2_R. \end{aligned} \tag{26}$$

Recall

$$\frac{1}{2}(x - \tilde{x}^k)^\top R(x^k - \tilde{x}^k) = \frac{1}{4}\left(\|\tilde{x}^k - x\|^2_R - \|x^k - x\|^2_R\right) + \frac{1}{4}\|x^k - \tilde{x}^k\|^2_R. \tag{27}$$

Adding (26) and (27), we obtain

$$\begin{aligned} &(x - \tilde{x}^k)^\top \Big\{ \tfrac{(1+\mu)}{2} R(x^k - \tilde{x}^k) - f(\tilde{x}^k) + A^\top\tilde{\lambda}^k + \tfrac{1}{2}A^\top HA(x^k - \tilde{x}^k) + \xi^k_x \\ &-A^\top H\left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\right)\Big\} \leq \tfrac{\mu}{2}\|x^k - \tilde{x}^k\|^2_R. \end{aligned} \tag{28}$$

Similarly, applying Lemma 2 to (7b), substituting $u^k = y^k$, $u = \tilde{y}^k$, $v = y$ and replacing R, n with S, m, respectively in (18) and

$$q(u) = g(\tilde{y}^k) - B^\top[\lambda^k - H(A\tilde{x}^k + \frac{1}{2}B\tilde{y}^k + \frac{1}{2}By^k - b)] - \xi^k_y,$$

we get

$$(y-\tilde{y}^k)^\top\{g(\tilde{y}^k)-B^\top[\lambda^k-H(A\tilde{x}^k+\frac{1}{2}B\tilde{y}^k+\frac{1}{2}By^k-b)]-\xi_y^k\} \geq \tfrac{1+\mu}{4}\Big(\|\tilde{y}^k-y\|_S^2-\|y^k-y\|_S^2\Big)+\tfrac{1-\mu}{4}\|y^k-\tilde{y}^k\|_S^2. \tag{29}$$

Recall

$$\frac{1}{2}(y-\tilde{y}^k)^\top S(y^k-\tilde{y}^k)=\frac{1}{4}\left(\|\tilde{y}^k-y\|_S^2-\|y^k-y\|_S^2\right)+\frac{1}{4}\|y^k-\tilde{y}^k\|_S^2. \tag{30}$$

Adding (29) and (30), we have

$$(y-\tilde{y}^k)^\top\Big\{\frac{(1+\mu)}{2}S(y^k-\tilde{y}^k)-g(\tilde{y}^k)+B^\top\tilde{\lambda}^k+\frac{1}{2}B^\top HB(y^k-\tilde{y}^k)+\xi_y^k -B^\top HB(y^k-\tilde{y}^k)\Big\}\leq\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2. \tag{31}$$

It follows from (28), (31), (7c) and (10) that

$$\begin{pmatrix} x-\tilde{x}^k \\ y-\tilde{y}^k \\ \lambda-\tilde{\lambda}^k \end{pmatrix}^\top \begin{pmatrix} \frac{(1+\mu)}{2}R(x^k-\tilde{x}^k)-f(x^k)+A^\top\tilde{\lambda}^k-A^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\right) \\ (\frac{(1+\mu)}{2}S(y^k-\tilde{y}^k)+B^\top HA(x^k-\tilde{x}^k)-g(y^k)+B^\top\tilde{\lambda}^k-B^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\right) \\ H^{-1}(\lambda^k-\tilde{\lambda}^k)-(A\tilde{x}^k+B\tilde{y}^k-b) \end{pmatrix}$$

$$\leq\frac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2+\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2,$$

which implies

$$(w-\tilde{w}^k)^\top(d_1(w^k,\tilde{w}^k)-d_2(w^k,\tilde{w}^k))-\frac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2-\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2\leq 0$$

and the assertion of this lemma is proved. □

Lemma 4. For given $w^k=(x^k,y^k,\lambda^k)\in\mathcal{R}^n_{++}\times\mathcal{R}^m_{++}\times\mathcal{R}^l$, let $\tilde{w}^k$ be generated by (7a)–(7c). Then for any $w^*=(x,y,\lambda)\in\mathcal{W}^*$, we have

$$(\tilde{w}^k-w^*)^\top d_2(w^k,\tilde{w}^k) \geq \varphi(w^k,\tilde{w}^k)-(w^k-\tilde{w}^k)^\top d_1(w^k,\tilde{w}^k) +\tfrac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2+\tfrac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2. \tag{32}$$

Proof. Recalling the definition in (15), we rewrite $\varphi(w^k,\tilde{w}^k)$ as

$$\varphi(w^k,\tilde{w}^k) = (w^k-\tilde{w}^k)^\top d_1(w^k,\tilde{w}^k)-\tfrac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2-\tfrac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2 +(\lambda^k-\tilde{\lambda}^k)^\top\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\right). \tag{33}$$

Using the monotonicity of f and g, we obtain

(34)

$$\begin{pmatrix} \tilde{x}^k-x^* \\ \tilde{y}^k-y^* \\ \tilde{\lambda}^k-\lambda^* \end{pmatrix}^\top \begin{pmatrix} f(\tilde{x}^k)-A^\top\tilde{\lambda}^k \\ g(\tilde{y}^k)-B^\top\tilde{\lambda}^k \\ A\tilde{x}^k+B\tilde{y}^k-b \end{pmatrix} \geq \begin{pmatrix} \tilde{x}^k-x^* \\ \tilde{y}^k-y^* \\ \tilde{\lambda}^k-\lambda^* \end{pmatrix}^\top \begin{pmatrix} f(x^*)-A^\top\lambda^* \\ g(y^*)-B^\top\lambda^* \\ Ax^*+By^*-b \end{pmatrix} \geq 0.$$

It follows from (34) that

$$\begin{aligned}(\tilde{w}^k - w^*)^\top d_2(w^k, \tilde{w}^k) &\geq (\tilde{w}^k - w^*)^\top \begin{pmatrix} A^\top H\left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\right) \\ B^\top H\left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\right) \\ 0 \end{pmatrix} \\ &= (A\tilde{x}^k + B\tilde{x}^k - b)^\top H\left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\right) \\ &= (\lambda^k - \tilde{\lambda}^k)^\top \left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)\right).\end{aligned}$$

Combining (33) and the above inequality, we can get the assertion of this lemma. □

The following theorem provides a unified framework for proving the convergence of the new algorithm.

Theorem 2. Let $w^* \in \mathcal{W}^*, w^{k+1}(\alpha_k)$ be defined by (13) and

$$\Theta(\alpha_k) := \|w^k - w^*\|_G^2 - \|w^{k+1}(\alpha_k) - w^*\|_G^2, \tag{35}$$

then

$$\Theta(\alpha_k) \geq \sigma(2\alpha_k \varphi(w^k, \tilde{w}^k) - \alpha_k^2 \|d_1(w^k, \tilde{w}^k)\|^2). \tag{36}$$

Proof. Since $w^* \in \mathcal{W}^*$ and

$$w_p^k := P_{\mathcal{W}}[w^k - \alpha_k d_2(w^k, \tilde{w}^k)], \tag{37}$$

it follows from (12) that

$$\|w_p^k - w^*\|^2 \leq \|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 - \|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2. \tag{38}$$

From (13), we get

$$\begin{aligned}&\|w^{k+1}(\alpha_k) - w^*\|^2 \\ =\ & \|(1-\sigma)(w^k - w^*) + \sigma(w_p^k - w^*)\|^2 \\ =\ & (1-\sigma)^2\|w^k - w^*\|^2 + \sigma^2\|w_p^k - w^*\|^2 + 2\sigma(1-\sigma)(w^k - w^*)^\top(w_p^k - w^*).\end{aligned}$$

Using the following identity

$$2(a+b)^\top b = \|a+b\|^2 - \|a\|^2 + \|b\|^2$$

for $a = w^k - w_*^k$, $b = w_*^k - w^*$, and (38), we obtain

$$
\begin{aligned}
& \|w^{k+1}(\alpha_k) - w^*\|^2 \\
= \; & (1-\sigma)^2\|w^k - w^*\|^2 + \sigma^2\|w_p^k - w^*\|^2 + \sigma(1-\sigma)\{\|w^k - w^*\|^2 \\
& -\|w^k - w_p^k\|^2 + \|w_p^k - w^*\|^2\} \\
= \; & (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w_p^k - w^*\|^2 - \sigma(1-\sigma)\|w^k - w_p^k\|^2 \\
\le \; & (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 \\
& -\sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2 - \sigma(1-\sigma)\|w^k - w_p^k\|^2 \\
\le \; & (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 \\
& -\sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2.
\end{aligned} \tag{39}
$$

Using the definition of $\Theta(\alpha_k)$ and (39), we get

$$
\Theta(\alpha_k) \;\; \ge \;\; \sigma\|w^k - w_p^k\|^2 + 2\sigma\alpha_k(w_p^k - w^*)^\top d_2(w^k, \tilde{w}^k). \tag{40}
$$

Applying (25)(with $w = w_p^k$), we obtain

$$
(w_p^k - \tilde{w}^k)^\top d_2(w^k, \tilde{w}^k) \;\; \ge \;\; (w_p^k - \tilde{w}^k)^\top d_1(w^k, \tilde{w}^k) - \tfrac{\mu}{2}\|x^k - \tilde{x}^k\|_R^2 - \tfrac{\mu}{2}\|y^k - \tilde{y}^k\|_S^2. \tag{41}
$$

Adding (32) and (41), we get

$$
(w_p^k - w^*)^\top d_2(w^k, \tilde{w}^k) \;\; \ge \;\; (w_p^k - w^k)^\top d_1(w^k, w^k) + \varphi(w^k, \tilde{w}^k). \tag{42}
$$

Applying (42) to the last term on the right side of (40), we obtain

$$
\begin{aligned}
\Theta(\alpha_k) \;\; & \ge \;\; \sigma\|w^k - w_p^k\|^2 + 2\sigma\alpha_k(w_p^k - w^k)^\top d_1(w^k, \tilde{w}^k) + 2\sigma\alpha_k\varphi(w^k, \tilde{w}^k) \\
& = \;\; \sigma\{\|w^k - w_p^k - \alpha_k d_1(w^k, \tilde{w}^k)\|^2 - \alpha_k^2\|d_1(w^k, \tilde{w}^k)\|^2 + 2\alpha_k\varphi(w^k, \tilde{w}^k)\}
\end{aligned}
$$

and the theorem is proved. □

4. Convergence of the proposed method

In this section, we prove the global convergence of the proposed method. From the computational point of view, a relaxation factor $\gamma \in (0, 2)$ is preferable in the correction.

Theorem 3. Let $w^* \in \mathcal{W}^*$ be a solution of SVI and let $w^{k+1}(\gamma\alpha_k)$ be generated by (13). Then w^k and $\tilde{w}^k$ are bounded, and

$$
\|w^{k+1}(\gamma\alpha_k) - w^*\|^2 \le \|w^k - w^*\|^2 - c\|w^k - \tilde{w}^k\|^2, \tag{43}
$$

where

$$
c := \frac{\sigma\gamma(2-\gamma)\alpha_1^2}{\alpha_2} > 0.
$$

Proof. It follows from (36), (19) and (20) that

$$\begin{aligned} & \|w^{k+1}(\gamma\alpha_k) - w^*\|^2 \\ \le\ & \|w^k - w^*\|^2 - \sigma(2\gamma\alpha_k\varphi(w^k, \tilde{w}^k) - \gamma^2\alpha_k^2\|d_1(w^k, \tilde{w}^k)\|^2) \\ =\ & \|w^k - w^*\|^2 - \gamma(2-\gamma)\alpha_k\sigma\varphi(w^k, \tilde{w}^k) \\ \le\ & \|w^k - w^*\|^2 - \tfrac{\sigma\gamma(2-\gamma)\alpha_1^2}{\alpha_2}\|w^k - \tilde{w}^k\|^2. \end{aligned}$$

Since $\gamma \in (0, 2)$, we have

$$\|w^{k+1} - w^*\| \le \|w^k - w^*\| \le \cdots \le \|w^0 - w^*\|,$$

and thus, $\{w^k\}$ is a bounded sequence.

It follows from (43) that

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

which means that

$$\lim_{k\to\infty} \|w^k - \tilde{w}^k\| = 0. \tag{44}$$

Since $\{w^k\}$ is a bounded sequence, we conclude that $\{\tilde{w}^k\}$ is also bounded. □

Now, we are ready to prove the convergence of the proposed method.

Theorem 4. The sequence $\{w^k\}$ generated by the proposed method converges to some w^∞ which is a solution of SVI.

Proof. Since $\{w^k\}$ is bounded, it has at least one cluster point. Let w^∞ be a cluster point of $\{w^k\}$. By definition, there exists a subsequence $\{w^{k_j}\}$ converging to w^∞, and, since $\mathcal{W}$ is closed set, we have $w^\infty \in \mathcal{W}$. It follows from (44) that

$$\lim_{j\to\infty} d_1(w^{k_j}, \tilde{w}^{k_j}) = 0 \quad \text{and} \quad \lim_{j\to\infty} d_2(w^{k_j}, \tilde{w}^{k_j}) = Q(w^\infty). \tag{45}$$

Moreover, (45) and (25) imply that

$$\lim_{j\to\infty} (w - w^{k_j})^\top d_2(w^{k_j}, \tilde{w}^{k_j}) \ge 0, \qquad \forall w \in \mathcal{W}, \tag{46}$$

and consequently

$$(w - w^\infty)^\top Q(w^\infty) \ge 0, \qquad \forall w \in \mathcal{W},$$

which means that w^∞ is a solution of SVI.

Now we prove that the sequence $\{w^k\}$ converges to w^∞. Since

$$\lim_{k\to\infty} \|w^k - \tilde{w}^k\| = 0, \quad \text{and} \quad \{\tilde{w}^{k_j}\} \to w^\infty,$$

for any $\epsilon > 0$, there exists an $l > 0$ such that

$$\|\tilde{w}^{k_l} - w^{\infty}\| < \frac{\epsilon}{2} \quad \text{and} \quad \|w^{k_l} - \tilde{w}^{k_l}\| < \frac{\epsilon}{2}. \tag{47}$$

Therefore, for any $k \geq k_l$, it follows from (43) and (47) that

$$\|w^k - w^{\infty}\| \leq \|w^{k_l} - w^{\infty}\| \leq \|w^{k_l} - \tilde{w}^{k_l}\| + \|\tilde{w}^{k_l} - w^{\infty}\| < \epsilon.$$

This implies that the sequence $\{w^k\}$ converges to w^{∞} which is a solution of SVI. □

5. $O(1/t)$ Convergence Rate

In this section, we show that the proposed method has the $O(1/t)$ convergence rate. Recall that $\mathcal{W}^*$ can be characterized as (see (2.3.2) in pp. 159 of [**15**])

$$\mathcal{W}^* = \bigcap_{w \in \mathcal{W}} \{\hat{w} \in \mathcal{W} : (w - \hat{w})^{\top} Q(w) \geq 0\}.$$

This implies that $\hat{w}$ is an approximate solution of SVI with the accuracy $\epsilon > 0$ if it satisfies

$$\hat{w} \in \mathcal{W} \quad \text{and} \quad \sup_{w \in \mathcal{W}} \{(\hat{w} - w)^{\top} Q(w)\} \leq \epsilon. \tag{48}$$

In the rest, our purpose is to show that after t iterations of the proposed method, we can find a $\hat{w} \in \mathcal{W}$ such that (48) is satisfied with $\epsilon = O(1/t)$.
Our analysis needs a new sequence defined by

$$\hat{w}^k = \begin{pmatrix} \hat{x}^k \\ \hat{y}^k \\ \hat{\lambda}^k \end{pmatrix} = \begin{pmatrix} \tilde{x}^k \\ \tilde{y}^k \\ \lambda^k - H(Ax^k + By^k - b) \end{pmatrix}. \tag{49}$$

Based on (33) and (49), we easily have a relationship

$$\varphi(w^k, \tilde{w}^k) = (w^k - \hat{w}^k)^{\top} d_1(w^k, \tilde{w}^k) - \frac{\mu}{2}\|x^k - \hat{x}^k\|_R^2 - \frac{\mu}{2}\|y^k - \hat{y}^k\|_S^2. \tag{50}$$

Using (3), (16) and (49), we obtain

$$d_2(w^k, \tilde{w}^k) = Q(\hat{w}^k). \tag{51}$$

Lemma 5. Let $\hat{w}^k$ be defined by (49) and $w \in \mathcal{W}$, then, we have

$$(w - \hat{w}^k)^{\top}(Q(\hat{w}^k) - d_1(w^k, \tilde{w}^k)) \geq -\frac{\mu}{2}\|x^k - \hat{x}^k\|_R^2 - \frac{\mu}{2}\|y^k - \hat{y}^k\|_S^2. \tag{52}$$

Proof. It follows from (28) and (31) that

$$(x-\tilde{x}^k)^\top\Big\{\tfrac{(1+\mu)}{2}R(x^k-\tilde{x}^k)-f(x^k)+A^\top\tilde{\lambda}^k -A^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)\right)\Big\}\le \tfrac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2 \tag{53}$$

and

$$(y-\tilde{y}^k)^\top\Big\{\frac{(1+\mu)}{2}S(y^k-\tilde{y}^k)-g(y^k)+B^\top\tilde{\lambda}^k+B^\top HA(x^k-\tilde{x}^k) -B^\top H(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k))\Big\}\le\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2. \tag{54}$$

Then, by using the notation of $\hat{w}^k$ in (49), (53) and (54) can be written as

$$(x-\hat{x}^k)^\top\Big\{\tfrac{(1+\mu)}{2}R(x^k-\hat{x}^k)-f(x^k)+A^\top\hat{\lambda}^k\Big\}\le\tfrac{\mu}{2}\|x^k-\hat{x}^k\|_R^2 \tag{55}$$

and

$$(y-\hat{y}^k)^\top\Big\{\frac{(1+\mu)}{2}S(y^k-\hat{y}^k)-g(y^k)+B^\top\hat{\lambda}^k+B^\top HA(x^k-\hat{x}^k)\Big\} \le\frac{\mu}{2}\|y^k-\hat{y}^k\|_S^2. \tag{56}$$

In addition, it follows from (6) and (49) that

$$A\hat{x}^k+B\hat{y}^k-b-H^{-1}(\lambda^k-\tilde{\lambda}^k)=0. \tag{57}$$

Combining (55)–(57), we get

$$\begin{pmatrix} x-\hat{x}^k\\ y-\hat{y}^k\\ \lambda-\hat{\lambda}^k\end{pmatrix}^\top\begin{pmatrix} f(\hat{x}^k)-A^\top\hat{\lambda}^k-\left(\frac{(1+\mu)}{2}R(x^k-\hat{x}^k)-(f(x^k)-f(\hat{x}^k))\right)\\ g(\hat{y}^k)-B^\top\hat{\lambda}^k-\left(\frac{(1+\mu)}{2}S(y^k-\hat{y}^k)-(g(y^k)-g(\hat{y}^k))+B^\top HA(x^k-\hat{x}^k)\right)\\ A\hat{x}^k+B\hat{y}^k-b-H^{-1}(\lambda^k-\tilde{\lambda}^k)\end{pmatrix}$$
$$\ge-\frac{\mu}{2}\|x^k-\hat{x}^k\|_R^2-\frac{\mu}{2}\|y^k-\hat{y}^k\|_S^2.$$

Recall the definition of $d_1(w^k,\tilde{w}^k)$, we obtain the assertion (52). The proof is completed. □

Lemma 6. For given $w^k\in\mathcal{R}^n_{++}\times\mathcal{R}^m_{++}\times\mathcal{R}^l$ and let w_p^k be defined by (37), then we have

$$\gamma\alpha_k(w-\hat{w}^k)^\top Q(w)+\frac{1}{2}(\|w-w^k\|^2-\|w-w_p^k\|^2)\ge\frac{1}{2}\gamma(2-\gamma)\alpha_k^2\|w^k-\tilde{w}^k\|^2. \tag{58}$$

Proof. Since $w_p^k\in\mathcal{W}$, substituting $w=w_p^k$ in (52) and using (50), we get

$$\begin{aligned}
\gamma\alpha_k(w_p^k-\hat{w}^k)^\top Q(\hat{w}^k) &\ge \gamma\alpha_k(w_p^k-\hat{w}^k)^\top d_1(w^k,\tilde{w}^k)-\frac{\mu}{2}\gamma\alpha_k\|x^k-\hat{x}^k\|_R^2-\frac{\mu}{2}\gamma\alpha_k\|y^k-\hat{y}^k\|_S^2\\
&= \gamma\alpha_k(w^k-\hat{w}^k)^\top d_1(w^k,\tilde{w}^k)+\gamma\alpha_k(w_p^k-w^k)^\top d_1(w^k,\tilde{w}^k)\\
&\quad -\frac{\mu}{2}\gamma\alpha_k\|x^k-\hat{x}^k\|_R^2-\frac{\mu}{2}\gamma\alpha_k\|y^k-\hat{y}^k\|_S^2.\\
&= \gamma\alpha_k\varphi(w^k,\tilde{w}^k)+\gamma\alpha_k(w_p^k-w^k)^\top d_1(w^k,\tilde{w}^k)\\
&\ge \gamma\alpha_k\varphi(w^k,\tilde{w}^k)-\frac{1}{2}\|w^k-w_p^k\|^2-\frac{1}{2}\gamma^2\alpha_k^2\|d_1(w^k,\tilde{w}^k)\|^2\\
&= \frac{1}{2}\gamma(2-\gamma)\alpha_k^2\|d_1(w^k,\tilde{w}^k)\|^2-\frac{1}{2}\|w^k-w_p^k\|^2.
\end{aligned} \tag{59}$$

On the other hand, using (37) and (51), w_p^k is the projection of $w^k - \gamma\alpha_k Q(\hat{w}^k)$ on $\mathcal{W}$, it follows from (11) that

$$(w^k - \gamma\alpha_k Q(\hat{w}^k) - w_p^k)^\top (w - w_p^k) \le 0, \quad \forall w \in \mathcal{W}$$

and consequently

$$\gamma\alpha_k (w - w_p^k)^\top Q(\hat{w}^k) \ge (w^k - w_p^k)^\top (w - w_p^k).$$

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

$$\gamma\alpha_k (w - w_p^k)^\top Q(\hat{w}^k) \ge \frac{1}{2}\left(\|w - w_p^k\|^2 - \|w - w^k\|^2\right) + \frac{1}{2}\|w^k - w_p^k\|^2. \tag{60}$$

Adding (59) and (60), we get

$$\gamma\alpha_k (w - \hat{w}^k)^\top Q(\hat{w}^k) + \frac{1}{2}(\|w - w^k\|^2 - \|w - w_p^k\|^2) \ge \frac{1}{2}\gamma(2-\gamma)\alpha_k^2 \|d_1(w^k, \tilde{w}^k)\|^2$$

and by using the monotonicity of Q, we obtain (58) and the proof is completed. □

Lemma 7. For given $w^k \in \mathcal{R}^n_{++} \times \mathcal{R}^m_{++} \times \mathcal{R}^l$ and let $w^{k+1}(\gamma\alpha_k)$ be generated by (13), then we have

$$\gamma\sigma\alpha_k (w - \hat{w}^k)^\top Q(w) + \frac{1}{2}(\|w - w^k\|^2 - \|w - w^{k+1}(\gamma\alpha_k)\|^2 \ge \frac{1}{2}\sigma\gamma(2-\gamma)\alpha_k^2 \|d_1(w^k, \tilde{w}^k)\|^2. \tag{61}$$

Proof.

$$\begin{aligned} \|w - w^k\|^2 - \|w - w^{k+1}(\gamma\alpha_k)\|^2 &= \|w^k - w\|^2 - \|w^k - \sigma(w^k - w_p^k) - w\|^2 \\ &= 2\sigma(w^k - w)^\top (w^k - w_p^k) - \sigma^2\|w^k - w_p^k\|^2 \\ &= 2\sigma\left(\|w^k - w_p^k\|^2 - (w - w_p^k)^\top (w^k - w_p^k)\right) - \sigma^2\|w^k - w_p^k\|^2. \end{aligned} \tag{62}$$

Using the following identity

$$(w - w_p^k)^\top (w^k - w_p^k) = \frac{1}{2}\left(\|w_p^k - w\|^2 - \|w^k - w\|^2\right) + \frac{1}{2}\|w^k - w_p^k\|^2,$$

we get

$$\|w^k - w_p^k\|^2 - 2(w - w_p^k)^\top (w^k - w_p^k) = \|w^k - w\|^2 - \|w_p^k - w\|^2. \tag{63}$$

Substituting (63) into (62), we obtain

$$\begin{aligned} \|w - w^k\|^2 - \|w - w^{k+1}(\gamma\alpha_k)\|^2 &= \sigma(\|w - w^k\|^2 - \|w - w_p^k\|^2) + \sigma(1-\sigma)\|w^k - w_p^k\|^2 \\ &\ge \sigma(\|w - w^k\|^2 - \|w - w_p^k\|^2). \end{aligned} \tag{64}$$

Substituting (64) into (58), we obtain (61), the required result. □

Now, we are ready to present the $O(1/t)$ convergence rate of the proposed method.

Theorem 5. For any integer $t > 0$, we have a $\hat{w}_t \in \mathcal{W}$ which satisfies

$$(\hat{w}_t - w)^\top Q(w) \leq \frac{1}{2\gamma\sigma\Upsilon_t}\|w - w^0\|^2, \quad \forall w \in \mathcal{W},$$

where

$$\hat{w}_t = \frac{1}{\Upsilon_t}\sum_{k=0}^{t}\alpha_k\hat{w}^k \quad \textit{and} \quad \Upsilon_t = \sum_{k=0}^{t}\alpha_k.$$

Proof. Summing the inequality (61) over $k = 0, \cdots, t$, we obtain

$$\left(\left(\sum_{k=0}^{t}\gamma\sigma\alpha_k\right)w - \sum_{k=0}^{t}\gamma\sigma\alpha_k\hat{w}^k\right)^\top Q(w) + \frac{1}{2}\|w - w^0\|^2 \geq 0.$$

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

$$(\hat{w}_t - w)^\top Q(w) \leq \frac{1}{2\gamma\sigma\Upsilon_t}\|w - w^0\|^2, \quad \forall w \in \mathcal{W}.$$

Indeed, $\hat{w}_t \in \mathcal{W}$ because it is a convex combination of $\hat{w}^0, \hat{w}^1, \cdots, \hat{w}^\top$. The proof is completed. □

It follows from (20) that

$$\Upsilon_t \geq \frac{\alpha_1}{\alpha_2}(t+1).$$

Suppose that for any compact set $\mathcal{D} \subset \mathcal{W}$, let $d = \sup\{\|w - w^0\|_G | w \in \mathcal{D}\}$. For any given $\epsilon > 0$, after most

$$t = \left[\frac{\alpha_2 d^2}{2\alpha_1\gamma\sigma\epsilon} - 1\right]$$

iterations, we have

$$(\hat{w}_t - w)^\top Q(w) \leq \epsilon, \forall w \in \mathcal{D}.$$

That is, the $O(1/t)$ convergence rate is established in an ergodic sense.

6. Preliminary computational results

In order to verify the theoretical assertions, we consider the following optimization problem with matrix variables:

$$\min\left\{\frac{1}{2}\|X - C\|_F^2 : X \in S_+^n\right\}, \tag{65}$$

where $\|\cdot\|_F$ is the matrix Fröbenius norm, that is,

$$\|C\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |C_{ij}|^2 \right)^{1/2},$$

$$S^n_+ = \left\{ H \in \mathcal{R}^{n\times n} : H^\top = H,\ H \succeq 0 \right\}.$$

Note that the matrix Fröbenius norm is induced by the inner product

$$\langle A, B\rangle = \text{Trace}(A^\top B).$$

Note that the problem (65) is equivalent to the following:

$$\begin{aligned} &\min \tfrac{1}{2}\|X-C\|^2 + \tfrac{1}{2}\|Y-C\|^2 \\ &\text{s.t.} \quad X - Y = 0, \\ &X, Y \in S^n_+, \end{aligned} \tag{66}$$

by attaching a Lagrange multiplier $Z \in \mathcal{R}^{n\times n}$ to the linear constraint $X - Y = 0$, the Lagrange function of (66) is

$$L(X,Y,Z) = \frac{1}{2}\|X-C\|^2 + \frac{1}{2}\|Y-C\|^2 - \langle Z, X-Y\rangle,$$

which is defined on $S^n_+ \times S^n_+ \times \mathcal{R}^{n\times n}$. If $(X^*, Y^*, Z^*) \in S^n_+ \times S^n_+ \times \mathcal{R}^{n\times n}$ is a KKT point of (66), then (66) can be converted to the following variational inequality: find $u^* = (X^*, Y^*, Z^*) \in \mathcal{W} = S^n_+ \times S^n_+ \times \mathcal{R}^{n\times n}$ such that

$$\left\{ \begin{array}{l} \langle X - X^*, (X^* - C) - Z^*\rangle \ge 0, \\ \langle Y - Y^*, (Y^* - C) + Z^*\rangle \ge 0, \quad \forall\, u = (X, Y, Z) \in \mathcal{W}, \\ X^* - Y^* = 0. \end{array} \right. \tag{67}$$

Problem (67) is a special case of (2)–(3) with matrix variables where $A = I_{n\times n}$, $B = -I_{n\times n}$, $b = 0$, $f(X) = X - C$, $g(Y) = Y - C$ and $\mathcal{W} = S^n_+ \times S^n_+ \times \mathcal{R}^{n\times n}$.

For simplification, we take $R = rI_{n\times n}, S = sI_{n\times n}$ and $H = I_{n\times n}$ where $r > 0$ and $s > 0$ are scalars. In all tests we take $\gamma = 1.98$, $\mu = 0.1$, $\sigma = 0.95$, $C = \text{rand}(n)$ and $(X^0, Y^0, Z^0) = (I_{n\times n}, I_{n\times n}, 0_{n\times n})$ as the initial point in the test. The iteration is stopped as soon as

$$\max\left\{ \|X^k - \tilde{X}^k\|, \|Y^k - \tilde{Y}^k\|, \|Z^k - \tilde{Z}^k\| \right\} \le 10^{-5}.$$

All codes were written in Matlab; we compare the proposed method with that in [**29**]. The iteration numbers, denoted by k, and the computational time for problem (65) with different dimensions are given in tables 5.1-5.2.

Tables 5.1-5.2 show the efficiency of the proposed method and its superiority to the methods of [**3**], [**6**] and [**29**] in terms of number of iteration and CPU time.

Dimension of the problem	The proposed method		The method in [6]		The method in [3]		The method in [29]	
	k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)
n=300	11	3.9	78	6.47	98	7.43	100	8.79
n=500	12	5.34	82	18.22	104	22.04	107	24.32
n=700	12	8.81	84	48.82	106	50.61	110	60.91
n=800	12	10.85	85	69.72	110	77.22	111	92.24

Table 5.1: Numerical results for the problem (65) with $r = s = 5$

Dimension of the problem	The proposed method		The method in [6]		The method in [3]		The method in [29]	
	k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)
n=300	10	3.24	139	9.27	161	10.32	173	12.61
n=500	11	4.8	146	34.15	170	42.52	170	43.65
n=700	12	7.51	151	70.46	178	79.75	183	81.85
n=800	12	10.71	159	110.41	182	104.14	183	129.62
multicolumn 9 c								

Table 5.2: Numerical results for the problem (65) with $r = s = 10$

7.Conclusions

In this paper, by combining the alternating direction method with SQP method, we proposed a descent SQP alternating direction method for solving variational inequality problems with separable structure. The main contribution of this paper, we used the same direction as as in [**6**], we presented a new step size α_k, we proved that the method is global convergence. The numerical results showed that our algorithm works well for the problem tested.

REFERENCES

1. A. Auslender A., M. Teboulle and S. Ben-Tiba, *A logarithmic-quadratic proximal method for variational inequalities*, Comput. Optim. Appl. **12** (1999), 31-40.

2. A. Bnouhachem, H. Benazza and M. Khalfaoui, *An inexact alternating direction method for solving a class of structured variational inequalities*, Appl. Math. Comput. **219** (2013), 7837-7846.

3. A. Bnouhachem, *On LQP alternating direction method for solving variational*, J. Inequal. Appl. **2014(80)** (2014), 1-15.

4. A. Bnouhachem and M.H. Xu, *An inexact LQP alternating direction method for solving a class of structured variational inequalities*, Comput. Math. Appl. **67** (2014) 671-680.

5. A. Bnouhachem and Q.H. Ansari, *A descent LQP alternating direction method for solving variational inequality problems with separable structure*, Appl. Math. Comput. **246** (2014), 519-532.

6. A. Bnouhachem and A. Hamdi, *Parallel LQP alternating direction method for solving variational inequality problems with separable structure*, J. Inequal. Appl. **2014(392)** (2014), 1-14.

7. A. Alhomaidan, A. Bnouhachem and A. Latif, *An LQP-SQP alternating direction for solving vriational inequality problems with separable structure*, J. Nonlinear Sci. Appl. **10(12)** (2017), 6246-6261.

8. A. Bnouhachem and T. M. Rassias, *An inexact alternating direction method with SQP regularization for the structured variational inequalities*, Inter. J. Nonlinear Anal. Appl. **8(1)** (2017), 269-289.
9. A. Bnouhachem, Q.H. Ansari and J.C. Yao, *SQP alternating direction for structured vriational inequality*, J. Nonlinear Convex Anal. In Press.
10. A. Bnouhachem and T. M. Rassias, *A new descent alternating direction method with LQP regularization for the structured variational inequalities*, Optim. Letters. In Press.
11. G. Chen and M. Teboulle, *A proximal-based decomposition method for convex minimization problems*, Math. Program. **64** (1994), 81-101.
12. J. Eckstein and D.P. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Program.**55** (1992), 293-318.
13. J. Eckstein, *Some saddle-function splitting methods for convex programming*, Optimization Methods and Software **4** (1994), 75-83.
14. J. Eckstein and M. Fukushima, *Some reformulation and applications of the alternating directions method of multipliers, Large Scale Optimization: State of the Art(W. W. Hager et al, Eds.)*, pp. 115-134, Kluwer Acad. Publ., 1994.
15. F. Facchinei and J.S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, vol. I and II. Springer Series in Operations Research. Springer, New York, 2003.
16. M. Fortin and R. Glowinski, Editors, *Augmented Lagrangian methods: Applications to the solution of boundary-valued problems*, North-Holland, Amsterdam, Holland, 1983.
17. D. Gabay and B. Mercier, *A dual algorithm for the solution of nonlinear variational problems via finite-element approximations*, Comput. Math. Appl. **2** (1976), 17-40.
18. D. Gabay, *Applications of the method of multipliers to variational inequalities, in Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems, M. Fortin and R. Glowinski, Eds.*, NorthHolland, Amsterdam, 299-331, 1983.
19. R. Glowinski, *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, 1984.
20. R. Glowinski and P. Le Tallec, *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, SIAM Studies in Applied mathematics, Philadelphia, PA, 1989.
21. B.S He, H. Yang and S.L. Wang, *Alternating directions method with self-adaptive penalty parameters for monotone variational inequalities*, J. Optim. Theory Appl. **106** (2000), 349-368.
22. B.S. He, L.Z. Liao, D.R. Han and H. Yang,*A new inexact alternating directions method for monotone variational inequalities*, Math. Program. **92** (2002), 103-118
23. B.S. He, *Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities*, Comput. Optim. Appl. **42** (2009), 195-212.
24. B.S. He,, M. Tao M., X.M. Yuan, *Alternating direction method with Gaussian back substitution for separable convex programming*, SIAM J. Optim. **22** (2012), 313-340.
25. Z.K. Jiang and A. Bnouhachem, *A projection-based prediction-correction method for structured monotone variational inequalities*, Appl. Math. Comput. **202** (2008), 747-759.

26. Z.K. Jiang and X.M. Yuan, *New parallel descent-like method for sloving a class of variational inequalities*, J. Optim. Theory Appl. **145** (2010), 311-323.

27. S. Kontogiorgis and R.R. Meyer, *A variable-penalty alternating directions method for convex optimization*, Math. Program. **83** (1998), 29-53.

28. L. S.Hou, *On the $O(1/t)$ convergence rate of the parallel descent-like method and parallel splitting augmented Lagrangian method for solving a class of variational inequalities*, Appl. Math. Comput. **219** (2013), 5862-5869.

29. M. Li, *A hybrid LQP-based method for structured variational inequalities*, Int. J. Comput. Math. **89(10)** (2012), 1412-1425.

30. A. Nagurney and P. Ramanujam, *Transportation network policy modeling with goal targets and generalized penalty functions*, Transportation Science **30** (1996), 3-13.

31. M. Tao M. and X.M. Yuan, *On the $O(1/t)$ convergence rate of alternating direction method with logarithmic-quadratic proximal regularization*, SIAM J. Optim. **22(4)** (2012), 1431-1448.

32. M. Teboulle, *Convergence of proximal-like algorithms*, SIAM J. Optim. **7** (1997), 1069-1083.

33. P .Tseng, *Alternating projection-proximal methods for convex programming and variational inequalities*, SIAM J. Optim. **7** (1997), 951-965.

34. K. Wang, L.L.Xu and D.R. Han, *A new parallel splitting descent method for structured variational inequalities*, J. Ind. Manag. Optim. **10(2)** (2014), 461-476.

35. X.M. Yuan and M. Li, *An LQP-based decomposition method for solving a class of variational inequalities*, SIAM J. Optim. **21(4)** (2011), 1309-1318.

(Received 24.09.2017)
(Revised 08.01.2018.)

Abdellah Bnouhachem
School of Management Science and Engineering,
Nanjing University,
Nanjing, 210093, P.R. China.
Ibn Zohr University,
ENSA, BP 1136,
Agadir, Morocco.
E-mail: *babedallah@yahoo.com*

Themistocles M.Rassias
Department of Mathematics,
National Technical University of
Athens Zografou Campus 157 80,
Athens, Greece.
E-mail: *trassias@math.ntua.gr*