J. Nonlinear Var. Anal. 4 (2020), No. 3, pp. 469-482
Available online at http://jnva.biemdas.com
https://doi.org/10.23952/jnva.4.2020.3.10

A DESCENT SQP ALTERNATING DIRECTION METHOD FOR MINIMIZING THE SUM OF THREE CONVEX FUNCTIONS

ABDELLAH BNOUHACHEM

Equipe MAISI, Ibn Zohr University, ENSA, BP 1136, Agadir, Morocco

Abstract. In this paper, with the aid of square quadratic proximal (SQP) method and the self-adaptive adjustment rule, we propose an SQP alternating direction method for solving the linearly constrained separable convex programming with three separable operators. Under standard assumptions, the global convergence of the proposed method is proved. Its efficiency is also verified via some numerical experiments.
Keywords. Variational inequalities; Monotone operator; Square-quadratic proximal method; Projection method; Alternating direction method.

1. INTRODUCTION

We consider the constrained convex programming problem with the following separate structure:

$$\min\{\theta_1(x)+\theta_2(y)+\theta_3(z)|Ax+By+Cz=b, x\in\mathbb{R}^{n_1}_+, y\in\mathbb{R}^{n_2}_+, z\in\mathbb{R}^{n_3}_+\}, \tag{1.1}$$

where $\theta_1:\mathbb{R}^{n_1}_+\to\mathbb{R}$, $\theta_2:\mathbb{R}^{n_2}_+\to\mathbb{R}$ and $\theta_3:\mathbb{R}^{n_3}_+\to\mathbb{R}$ are closed proper convex functions, $A\in\mathbb{R}^{l\times n_1}$, $B\in\mathbb{R}^{l\times n_2}$ and $C\in\mathbb{R}^{l\times n_3}$ are given matrices, and $b\in\mathbb{R}^l$ is a given vector.

Various methods have been suggested for solving the constrained convex programming problem, where the objective function is the sum of two separable convex functions and the constraint set is also separable into two parts, i.e.,

$$\min\{\theta_1(x)+\theta_2(y)|Ax+By=b, x\in\mathbb{R}^{n_1}_+, y\in\mathbb{R}^{n_2}_+\}, \tag{1.2}$$

where $\theta_1:\mathbb{R}^{n_1}_+\to\mathbb{R}$ and $\theta_2:\mathbb{R}^{n_2}_+\to\mathbb{R}$ are closed proper convex functions, $A\in\mathbb{R}^{l\times n_1}$ and $B\in\mathbb{R}^{l\times n_2}$ are given matrices and $b\in\mathbb{R}^l$.

The alternating direction method (ADM), originally proposed in Glowinski and Marrocco [9], is one of the most popular methods for solving (1.2). We now have a variety of techniques to suggest and analyze various iterative ADMs. For more details, one is referred to [8, 10, 11, 12, 15, 16]. For the ADM with logarithmic-quadratic proximal regularization, we refer to [2, 3, 4, 13, 15, 17, 18] and the references therein.

Let $\partial(.)$ denote the sub-gradient operator of a convex function. Let $f(x)\in\partial\theta_1(x)$, $g(y)\in\partial\theta_2(y)$ and $h(z)\in\partial\theta_3(z)$ be the sub-gradient of $\theta_1(x)$, $\theta_2(y)$ and $\theta_3(z)$, respectively. By attaching a Lagrange multiplier vector $\lambda\in\mathbb{R}^l$ to the linear constraint $Ax+By+Cz=b$, problem (1.1) can be written in terms of finding $w\in\mathscr{W}$ such that

$$(w'-w)^TQ(w)\geq 0, \qquad \forall w'\in\mathscr{W}, \tag{1.3}$$

E-mail address: babedallah@yahoo.com.
Received November 20, 2019; Accepted February 8, 2020.

where

$$w=\begin{pmatrix} x\\ y\\ z\\ \lambda \end{pmatrix} \quad Q(w)=\begin{pmatrix} f(x)-A^T\lambda\\ g(y)-B^T\lambda\\ h(z)-C^Tz\\ Ax+By+Cz-b \end{pmatrix}, \quad \mathscr{W}=\mathbb{R}^{n_1}_+\times\mathbb{R}^{n_2}_+\times\mathbb{R}^{n_3}_+\times\mathbb{R}^l. \tag{1.4}$$

The problem (1.3)-(1.4) is referred to as SVI$_3$. Peng and Wu [14] formulated the constrained matrix optimization problem into SVI$_3$. By combining the ADM and parallel splitting augmented Lagrangian method, they proposed a partial parallel splitting augmented Lagrangian method for solving SVI$_3$. Recently, Cao, Han and Xu [7] proposed a new partial splitting augmented Lagrangian method for solving SVI$_3$. From a given $w^k=(x^k,y^k,z^k,\lambda^k)\in\mathscr{W}$, the predictor $\tilde{w}^k=(\tilde{x}^k,\tilde{y}^k,\tilde{z}^k,\tilde{\lambda}^k)$ was obtained via solving the following system:

$$(x-\tilde{x}^k)^\top(f(\tilde{x}^k)-A^\top[\lambda^k-\beta H(A\tilde{x}^k+By^k+Cz^k-b)])\geq 0, \quad \forall x\in\mathbb{R}^{n_1}_+, \tag{1.5a}$$

$$(y-\tilde{y}^k)^\top(g(\tilde{y}^k)-B^\top[\lambda^k-\beta H(A\tilde{x}^k+B\tilde{y}^k+Cz^k-b)])\geq 0, \quad \forall y\in\mathbb{R}^{n_2}_+, \tag{1.5b}$$

$$(z-\tilde{z}^k)^\top(h(\tilde{z}^k)-C^\top[\lambda^k-\beta H(A\tilde{x}^k+By^k+C\tilde{z}^k-b)])\geq 0, \quad \forall z\in\mathbb{R}^{n_3}_+, \tag{1.5c}$$

$$\tilde{\lambda}^k=\lambda^k-\beta H(A\tilde{x}^k+B\tilde{y}^k+C\tilde{z}^k-b). \tag{1.5d}$$

Very recently, Bnouhachem, Ansari and Al-Homidan [5] suggested that the complementarity subproblems arising in ADM (1.5a)-1.5c) could be regularized by the square quadratic proximal (SQP) regularization, and the SQP regularization forces the solutions of ADM subproblems to be interior points of $\mathbb{R}^{n_1}_+$, $\mathbb{R}^{n_2}_+$ and $\mathbb{R}^{n_3}_+$, respectively. More specifically, from a given $w^k=(x^k,y^k,z^k,\lambda^k)\in\mathscr{W}$, the predictor $\tilde{w}^k=(\tilde{x}^k,\tilde{y}^k,\tilde{z}^k,\tilde{\lambda}^k)$ was obtained via solving the following system:

$$f(x)-A^\top[\lambda^k-\beta H(Ax+By^k+Cz^k-b)]+R[\frac{1}{2}(x-x^k)+\mu(x^k-X_k(\sqrt{x})^{-1})]=0, \tag{1.6a}$$

$$g(y)-B^\top[\lambda^k-\beta H(Ax^k+By+Cz^k-b)]+S[\frac{1}{2}(y-y^k)+\mu(y^k-Y_k(\sqrt{y})^{-1})]=0, \tag{1.6b}$$

$$h(z)-C^\top[\lambda^k-\beta H(Ax^k+By^k+Cz-b)]+P[\frac{1}{2}(z-z^k)+\mu(z^k-Z_k(\sqrt{z})^{-1}))]=0, \tag{1.6c}$$

$$\tilde{\lambda}^k=\lambda^k-\beta H(A\tilde{x}^k+B\tilde{x}^k+C\tilde{z}^k-b),$$

where $\mu\in(0,1)$ and $\beta>0$ are given constants; $H\in\mathbb{R}^{l\times l}$, $R\in\mathbb{R}^{n_1\times n_1}$, $S\in\mathbb{R}^{n_2\times n_2}$ and $P\in\mathbb{R}^{n_3\times n_3}$ are positive definite diagonal matrices; X_k,Y_k and Z_k are positive definite diagonal matrices defined by

$$X_k=diag(x_1^k\sqrt{x_1^k},\ldots,u_n^x\sqrt{x_n^k}):=\begin{pmatrix} x_1^k\sqrt{x_1^k} & & \\ & \ddots & \\ & & x_n^k\sqrt{x_n^k} \end{pmatrix},$$

$$Y_k=diag(y_1^k\sqrt{y_1^k},\ldots,y_n^k\sqrt{y_n^k})$$

and

$$Z_k=diag(z_1^k\sqrt{z_1^k},\ldots,z_n^k\sqrt{z_n^k}),$$

$(\sqrt{x})^{-1}\in\mathscr{R}^{n_1}_{++}$ is a vector whose j-th element is $1/\sqrt{x_j}$, $(\sqrt{y})^{-1}\in\mathscr{R}^{n_2}_{++}$ is a vector whose j-th element is $1/\sqrt{y_j}$, and $(\sqrt{z})^{-1}\in\mathscr{R}^{n_3}_{++}$ is a vector whose j-th element is $1/\sqrt{z_j}$.

Note that the numerical experience significantly depends on the initial penalty parameter. The method converges quite quickly when a proper fixed penalty parameter is chosen. However, this proper penalty parameter is unknown beforehand. The main aim of this paper is twofold. First, by combining the ADM and SQP method, the predictor is obtained by solving the SQP system approximately. Second, since the self-adaptive adjustment rule is necessary in practice, we propose a self-adaptive method that adjusts the scalar parameter automatically. We also study the global convergence of the proposed method under certain conditions. The effectiveness and superiority of the proposed method is verified by the numerical experiments.

2. THE PROPOSED METHOD

We state some preliminaries that are useful in later analysis.

For any vector $u \in \mathbb{R}^n$, we denote $\|u\|^2 = u^\top u$, and $\|u\|_\infty = \max\{|u_1|\ldots,|u_n|\}$. Let $D \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. Let the operators $\lambda_l(D)$ and $\lambda_m(D)$ denote the largest eigenvalue and the smallest eigenvalue of D, respectively. We denote the D-norm of u by $\|u\|_D^2 = u^\top Du$.

Some important properties of projections are gathered as following.

Lemma 2.1. *Let Ω be a nonempty closed convex subset of $\mathbb{R}^l$, and let $P_\Omega[.]$ be the projection on Ω with respect to the Euclidean norm, that is,*

$$P_\Omega[v] = argmin\{\|v-u\| \,:\, u \in \Omega\}.$$

Then, we have the following inequalities:

$$(z - P_\Omega[z])^\top (P_\Omega[z] - v) \geq 0, \quad \forall\, z \in \mathbb{R}^l, v \in \Omega; \tag{2.1}$$

$$\|u - P_\Omega[z]\|^2 \leq \|z-u\|^2 - \|z - P_\Omega[z]\|^2, \quad \forall\, z \in \mathbb{R}^l, u \in \Omega. \tag{2.2}$$

Definition 2.1. *The mapping $T : \mathbb{R}^n \to \mathbb{R}^n$ is said to be*

(a) *monotone if*

$$(Tx - Ty)^\top (x-y) \geq 0, \qquad \forall x,y \in \mathbb{R}^n;$$

(b) *L-Lipschitz continuous if there exists a constant $L > 0$ such that*

$$\|Tx - Ty\| \leq L\|x-y\|, \qquad \forall x,y \in \mathbb{R}^n;$$

We next make the following standard assumptions:

Assumption A. f is monotone and continuous on $\mathbb{R}_+^{n_1}$, g is monotone and continuous on $\mathbb{R}_+^{n_2}$, and h is monotone and continuous on $\mathbb{R}_+^{n_3}$.

Assumption B. The solution set of SVI$_3$, denoted by $\mathscr{W}^*$, is nonempty.

Let $\beta_k > 0$, $r > 0, s > 0, p > 0$, $H \in \mathbb{R}^{l\times l}$, $R \in \mathbb{R}^{n_1\times n_1}$, $S \in \mathbb{R}^{n_2\times n_2}$ and $P \in \mathbb{R}^{n_3\times n_3}$ be positive definite diagonal matrices, where $R = rI_{n_1\times n_1}, S = sI_{n_2\times n_2}$ and $P = pI_{n_3\times n_3}$. We propose the following inexact SQP alternating direction method for solving SVI$_3$.

Algorithm 2.1.

Step 0. *The initial step.*
Given $\varepsilon > 0$, $\mu \in (0,1), \eta \in (0,1), \rho > 0$ and $w^0 = (x^0,y^0,z^0,\lambda^0) \in \mathbb{R}_{++}^{n_1} \times \mathbb{R}_{++}^{n_2} \times \mathbb{R}_{++}^{n_3} \times \mathbb{R}^l$, $\beta_0 = \min\left\{\frac{(1-\eta)\lambda_m(R)}{10\lambda_l(H)\|A\|^2}, \frac{(1-\eta)\lambda_m(S)}{10\lambda_l(H)\|B\|^2}, \frac{(1-\eta)\lambda_m(P)}{10\lambda_l(H)\|C\|^2}\right\}$, set $k=0$.

Step 1. *Prediction step:*
Compute $\tilde{w}^k = (\tilde{x}^k,\tilde{y}^k,\tilde{z}^k,\tilde{\lambda}^k) \in \mathbb{R}_{++}^{n_1} \times \mathbb{R}_{++}^{n_2} \times \mathbb{R}_{++}^{n_3} \times \mathbb{R}^l$ by solving the following system:

$$\beta_k\left(f(x) - A^\top[\lambda^k - H(Ax^k + By^k + Cz^k - b)]\right) + R[\frac{1}{2}(x-x^k) + \mu(x^k - X_k(\sqrt{x})^{-1})] =: \xi_x^k \approx 0, \tag{2.3a}$$

$$\beta_k\left(g(y) - B^\top[\lambda^k - H(A\tilde{x}^k + By^k + Cz^k - b)]\right) + S[\frac{1}{2}(y-y^k) + \mu(y^k - Y_k(\sqrt{y})^{-1})] =: \xi_y^k \approx 0, \tag{2.3b}$$

$$\beta_k\left(h(z) - C^\top[\lambda^k - H(A\tilde{x}^k + B\tilde{y}^k + Cz^k - b)]\right) + P[\frac{1}{2}(z-z^k) + \mu(z^k - Z_k(\sqrt{z})^{-1})] =: \xi_z^k \approx 0, \tag{2.3c}$$

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b), \tag{2.3d}$$

where β_k is a proper parameter satisfying

$$\|\xi_x^k\| \le \frac{\eta r}{2}\|x^k - \tilde{x}^k\|, \qquad \|\xi_y^k\| \le \frac{\eta s}{2}\|y^k - \tilde{y}^k\|, \qquad \|\xi_z^k\| \le \frac{\eta p}{2}\|z^k - \tilde{z}^k\| \tag{2.4}$$

and

$$\xi^k = \begin{pmatrix} \xi_x^k \\ \xi_y^k \\ \xi_z^k \\ 0 \end{pmatrix} = \beta_k \begin{pmatrix} f(\tilde{x}^k) - f(x^k) + \rho A^\top HA(x^k - \tilde{x}^k) \\ g(\tilde{y}^k) - g(y^k) + \rho B^\top HB(y^k - \tilde{y}^k) \\ h(\tilde{z}^k) - h(z^k) + \rho C^\top HC(z^k - \tilde{z}^k) \\ 0 \end{pmatrix}. \tag{2.5}$$

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

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

$$w^{k+1}(\alpha_k) = (1-\sigma)w^k + \sigma P_{\mathscr{W}}[w^k - \alpha_k d_2(w^k, \tilde{w}^k)], \quad \sigma \in (0,1), \tag{2.6}$$

where

$$\alpha_k = \frac{\varphi(w^k, \tilde{w}^k)}{\|d_1(w^k, \tilde{w}^k)\|^2}, \tag{2.7}$$

$$d_2(w^k, \tilde{w}^k) = \begin{pmatrix} \beta_k(f(\tilde{x}^k) - A^\top\tilde{\lambda}^k) + \beta_k A^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)) \\ \beta_k(g(\tilde{y}^k) - B^\top\tilde{\lambda}^k) + \beta_k B^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)) \\ \beta_k(h(\tilde{z}^k) - C^\top\tilde{\lambda}^k) + \beta_k C^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)) \\ \beta_k(A\tilde{x}^k + B\tilde{y}^k + C\tilde{z}^k - b) \end{pmatrix} \tag{2.8}$$

$$\begin{aligned} \varphi(w^k, \tilde{w}^k) = (w^k - \tilde{w}^k)^\top d_1(w^k, \tilde{w}^k) - \frac{\mu}{2}\|x^k - \tilde{x}^k\|_R^2 - \frac{\mu}{2}\|y^k - \tilde{y}^k\|_S^2 - \frac{\mu}{2}\|z^k - \tilde{z}^k\|_P^2 \\ + \beta_k(\lambda^k - \tilde{\lambda}^k)^\top(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)), \end{aligned} \tag{2.9}$$

$$d_1(w^k, \tilde{w}^k) = \begin{pmatrix} \frac{(1+\mu)}{2}R(x^k - \tilde{x}^k) - \beta_k[f(x^k) - f(\tilde{x}^k)] + \rho\beta_k A^\top HA(x^k - \tilde{x}^k) \\ \frac{(1+\mu)}{2}S(y^k - \tilde{y}^k) - \beta_k[g(y^k) - g(\tilde{y}^k)] + \beta_k B^\top HA(x^k - \tilde{x}^k) \\ +\rho\beta_k B^\top HB(y^k - \tilde{y}^k) \\ \frac{(1+\mu)}{2}P(z^k - \tilde{z}^k) - \beta_k[h(z^k) - h(\tilde{z}^k)] + \beta_k C^\top H(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k)) \\ +\rho\beta_k C^\top HC(z^k - \tilde{z}^k) \\ \beta_k H^{-1}(\lambda^k - \tilde{\lambda}^k) \end{pmatrix} \tag{2.10}$$

Step 4. *Adjusting.*
Adaptive rule of choosing a suitable β_{k+1} is as the start prediction step size for the next iteration

$$\beta_{k+1} := \begin{cases} \min\{\beta_0, \tau * \beta_k\} & \text{if } \max\{r_1, r_2, r_3\} \le 0.5, \\ \beta_k & \text{otherwise}, \end{cases} \tag{2.11}$$

where $r_1 = \frac{2\|\xi_x^k\|}{r\|x^k - \tilde{x}^k\|}$, $r_2 = \frac{2\|\xi_y^k\|}{s\|y^k - \tilde{y}^k\|}$ $r_3 = \frac{2\|\xi_z^k\|}{p\|z^k - \tilde{z}^k\|}$ and $\tau > 1$. Set $k := k+1$ and go to Step 1.

Remark 2.1. *In general, the prediction step is implementable. Sometimes, we can get the approximate solution of* (2.3a)-(2.3d) *directly by choosing a suitable* $\beta_k > 0$. *Observe* $R = rI_{n_1\times n_1}, S = sI_{n_2\times n_2}$ *and* $P = pI_{n_3\times n_3}$. *If* f *is Lipschitz continuous on* $\mathbb{R}_+^{n_1}$ *with Lipschitz constant* k_f, g *is Lipschitz continuous on* $\mathbb{R}_+^{n_2}$ *with Lipschitz constant* k_g, *and* h *is Lipschitz continuous on* $\mathbb{R}_+^{n_3}$ *with Lipschitz constant* k_h, *then criterion* (2.4) *is satisfied provided that* $\beta_k \le \min\{\frac{\eta r}{2(k_f + \rho\|A^\top HA\|)}, \frac{\eta s}{2(k_g + \rho\|B^\top HB\|)}, \frac{\eta p}{2(k_h + \rho\|C^\top HC\|)}\}$.

Remark 2.2. *Our method can be viewed as an extension and improvement of some known results.*

- *The proposed method obtains the predictors* $\tilde{x}^k$, $\tilde{y}^k$ *and* $\tilde{z}^k$ *by solving easier systems of nonlinear equations. In contrast, the predictors* $\tilde{x}^k$, $\tilde{y}^k$ *and* $\tilde{z}^k$ *in* [7, 14] *by solving a series of variational inequalities.*

- *The method proposed in [5] solved problem (2.3a)-(2.3b) exactly. It is more practical to find approximate solutions of problem (2.3a)-(2.3b) rather than the exact solutions due to the fact that in general this excludes some practical applications. Driven by the fact of eliminating this drawback, we solve problem (2.3a)-(2.3b) approximately.*
- *The initial penalty parameter in [5, 7, 14] is unknown, while the initial penalty parameter of the proposed method is defined by* $\beta_0 = \min\left\{\frac{(1-\eta)\lambda_m(R)}{10\lambda_l(H)\|A\|^2}, \frac{(1-\eta)\lambda_m(S)}{10\lambda_l(H)\|B\|^2}, \frac{(1-\eta)\lambda_m(P)}{10\lambda_l(H)\|C\|^2}\right\}$ *and therefore more precise. Moreover, since the self-adaptive adjustment rule is necessary in practice, we propose a self-adaptive method that adjusts the scalar parameter* β_k *automatically.*
- *Comparing the proposed method with the methods in [5, 7, 14], the new iterate is obtained by using a new direction with a new step size* α_k.

The following lemma plays a crucial role in our convergence analysis to be conducted.

Lemma 2.2. [1, 5] *Let* $q : \mathbb{R}^n_+ \to \mathbb{R}^n$ *be a monotone mapping and* $R_1 := diag(r_1, \dots, r_n) \in \mathbb{R}^{n\times n}$ *be a positive definite diagonal matrix. For given* $u^k > 0$, $\mu > 0$, *if* $U_k := diag\left(u_1^k\sqrt{u_1^k}, \dots, u_n^k\sqrt{u_n^k}\right)$, $\sqrt{u} = (\sqrt{u_1}, \dots, \sqrt{u_n})$, *and* $(\sqrt{u})^{-1}$ *is an n-vector whose j-th element is* $1/\sqrt{u_j}$, *then the equation*

$$q(u) + R_1\left[\frac{1}{2}(u-u^k) + \mu(u^k - U_k(\sqrt{u})^{-1})\right] = 0 \tag{2.12}$$

has a unique positive solution u. Moreover, for all $v \geq 0$,

$$(v-u)^\top q(u) \geq \tfrac{1+\mu}{4}\left(\|u-v\|_R^2 - \|u^k - v\|_R^2\right) + \tfrac{1-\mu}{4}\|u^k - u\|_R^2. \tag{2.13}$$

Now we are ready to present an inequality where a lower bound of $\varphi(w^k, \tilde{w}^k)$ is found for all $w^k \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++} \times \mathbb{R}^{n_3}_{++} \times \mathbb{R}^l$. And it is the key to the proof of the main convergence results.

Theorem 2.1. *Let* $w^k \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++} \times \mathbb{R}^{n_3}_{++} \times \mathbb{R}^l$ *be generated by (2.3a)-(2.3d). Then there exist two constants* $\alpha_1 > 0$ *and* $\alpha_2 > 0$ *such that*

$$\varphi(w^k, \tilde{w}^k) \geq \alpha_1\|w^k - \tilde{w}^k\|^2 \tag{2.14}$$

and

$$\alpha_k \geq \frac{\alpha_1}{\alpha_2}. \tag{2.15}$$

Proof. Observe that

$$(\lambda^k - \tilde{\lambda}^k)^\top A(x^k - \tilde{x}^k) \geq -\frac{1}{2}\left(\frac{1}{3}\|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2 + 3\|A(x^k - \tilde{x}^k)\|_H^2\right),$$

$$(\lambda^k - \tilde{\lambda}^k)^\top B(y^k - \tilde{y}^k) \geq -\frac{1}{2}\left(\frac{1}{3}\|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2 + 3\|B(y^k - \tilde{y}^k)\|_H^2\right),$$

$$(\lambda^k - \tilde{\lambda}^k)^\top C(z^k - \tilde{z}^k) \geq -\frac{1}{2}\left(\frac{1}{3}\|\lambda^k - \tilde{\lambda}^k\|_{H^{-1}}^2 + 3\|C(z^k - \tilde{z}^k)\|_H^2\right),$$

$$(A(x^k - \tilde{x}^k))^\top HB(y^k - \tilde{y}^k) \geq -\frac{1}{2}(\|A(x^k - \tilde{x}^k\|_H^2 + \|B(y^k - \tilde{y}^k)\|_H^2),$$

$$(C(z^k - \tilde{z}^k))^\top HA(x^k - \tilde{x}^k) \geq -\frac{1}{2}(\|C(z^k - \tilde{z}^k\|_H^2 + \|A(x^k - \tilde{x}^k)\|_H^2),$$

and

$$(C(z^k - \tilde{z}^k))^\top HB(y^k - \tilde{y}^k) \geq -\frac{1}{2}(\|C(z^k - \tilde{z}^k\|_H^2 + \|B(y^k - \tilde{y}^k)\|_H^2).$$

It follows from the definition of $\varphi(w^k,\tilde{w}^k)$ that

$$\begin{aligned}
\varphi(w^k,\tilde{w}^k) &= \frac{1}{2}\|x^k-\tilde{x}^k\|_R^2+\frac{1}{2}\|y^k-\tilde{y}^k\|_S^2+\frac{1}{2}\|z^k-\tilde{z}^k\|_P^2+\beta_k\|\lambda^k-\tilde{\lambda}^k\|_{H^{-1}}^2\\
&+\beta_k(A(x^k-\tilde{x}^k))^\top HB(y^k-\tilde{y}^k)-\beta_k(x^k-\tilde{x}^k)^\top[f(x^k)-f(\tilde{x}^k)\\
&-\rho A^\top HA(x^k-\tilde{x}^k)]-\beta_k(y^k-\tilde{y}^k)^\top[g(y^k)-g(\tilde{y}^k)-\rho B^\top HB(y^k-\tilde{y}^k)]\\
&-\beta_k(z^k-\tilde{z}^k)^\top[h(z^k)-h(\tilde{z}^k)-\rho C^\top HC(z^k-\tilde{z}^k)]\\
&+\beta_k(C(z^k-\tilde{z}^k))^\top H(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k))\\
&+\beta_k(\lambda^k-\tilde{\lambda}^k)^\top(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k))\\
&\geq \frac{1}{2}\|x^k-\tilde{x}^k\|_R^2+\frac{1}{2}\|y^k-\tilde{y}^k\|_S^2+\frac{1}{2}\|z^k-\tilde{z}^k\|_P^2+\frac{1}{2}\beta_k\|\lambda^k-\tilde{\lambda}^k\|_{H^{-1}}^2\\
&-\frac{5}{2}\beta_k\|A(x^k-\tilde{x}^k)\|_H^2-\frac{5}{2}\beta_k\|B(y^k-\tilde{y}^k)\|_H^2-\frac{5}{2}\beta_k\|C(z^k-\tilde{z}^k)\|_H^2\\
&-\beta_k(x^k-\tilde{x}^k)^\top[f(x^k)-f(\tilde{x}^k)-\rho A^\top HA(x^k-\tilde{x}^k)]\\
&-\beta_k(y^k-\tilde{y}^k)^\top[g(y^k)-g(\tilde{y}^k)-\rho B^\top HB(y^k-\tilde{y}^k)]\\
&-\beta_k(z^k-\tilde{z}^k)^\top[h(z^k)-h(\tilde{z}^k)-\rho C^\top HC(z^k-\tilde{z}^k)].
\end{aligned}$$

By using Cauchy-Schwarz inequality in (2.4), we have

$$(w^k-\tilde{w}^k)^\top\xi_k \geq -\eta\|w^k-\tilde{w}^k\|_M^2,$$

where

$$M=\begin{pmatrix} \frac{1}{2}R & & & \\ & \frac{1}{2}S & & \\ & & \frac{1}{2}P & \\ & & & \frac{1}{2}\beta_k H^{-1}\end{pmatrix}. \tag{2.16}$$

Then, we have

$$\begin{aligned}
&\varphi(w^k,\tilde{w}^k) \qquad (2.17)\\
\geq\ & \frac{1}{2}\|x^k-\tilde{x}^k\|_R^2+\frac{1}{2}\|y^k-\tilde{y}^k\|_S^2+\frac{1}{2}\|z^k-\tilde{z}^k\|_P^2+\frac{1}{2}\beta_k\|\lambda^k-\tilde{\lambda}^k\|_{H^{-1}}^2-\frac{5}{2}\beta_k\|A(x^k-\tilde{x}^k)\|_H^2\\
&-\frac{5}{2}\beta_k\|B(y^k-\tilde{y}^k)\|_H^2-\frac{5}{2}\beta_k\|C(z^k-\tilde{z}^k)\|_H^2-\eta\|w^k-\tilde{w}^k\|_M^2\\
\geq\ & \frac{(1-\eta)}{2}\|w^k-\tilde{w}^k\|_M^2+\left(\frac{(1-\eta)\lambda_m(R)}{4}-\frac{5}{2}\beta_k\lambda_l(H)\|A\|^2\right)\|x^k-\tilde{x}^k\|^2\\
&+\left(\frac{(1-\eta)\lambda_m(S)}{4}-\frac{5}{2}\beta_k\lambda_l(H)\|B\|^2\right)\|y^k-\tilde{y}^k\|^2\\
&+\left(\frac{(1-\eta)\lambda_m(P)}{4}-\frac{5}{2}\beta_k\lambda_l(H)\|C\|^2\right)\|z^k-\tilde{z}^k\|^2\\
\geq\ & \frac{(1-\eta)}{2}\|w^k-\tilde{w}^k\|_M^2+\left(\frac{(1-\eta)\lambda_m(R)}{4}-\frac{5}{2}\beta_0\lambda_l(H)\|A\|^2\right)\|x^k-\tilde{x}^k\|^2\\
&+\left(\frac{(1-\eta)\lambda_m(S)}{4}-\frac{5}{2}\beta_0\lambda_l(H)\|B\|^2\right)\|y^k-\tilde{y}^k\|^2\\
&+\left(\frac{(1-\eta)\lambda_m(P)}{4}-\frac{5}{2}\beta_0\lambda_l(H)\|C\|^2\right)\|z^k-\tilde{z}^k\|^2\\
\geq\ & \alpha_1\|w^k-\tilde{w}^k\|^2,
\end{aligned}$$

where $\alpha_1 > 0$ is a constant. The third inequality holds because $\beta_k \le \beta_0$ for any k. The fourth inequality is obtained from the definition of β_0. Recalling the definition in (2.10), we rewrite $d_1(w^k, \tilde{w}^k)$ as

$$d_1(w^k, \tilde{w}^k) = \xi_k + G(w^k - \tilde{w}^k),$$

where

$$G = \begin{pmatrix} \frac{(1+\mu)}{2}R & 0 & 0 & 0 \\ \beta_k B^\top HA & \frac{(1+\mu)}{2}S & 0 & 0 \\ \beta_k C^\top HA & \beta_k C^\top HB & \frac{(1+\mu)}{2}P & 0 \\ 0 & 0 & 0 & \beta_k H^{-1} \end{pmatrix}.$$

Note that, for any $a, b \in \mathbb{R}^n$, $\|a+b\|^2 \le 2\|a\|^2 + 2\|b\|^2$, which implies that

$$\begin{aligned} & \|d_1(w^k, \tilde{w}^k)\|^2 \\ \le\ & 2\|\beta_k(f(\tilde{x}^k) - f(x^k)) + \rho\beta_k A^\top HA(x^k - \tilde{x}^k)\|^2 \\ & +2\|\beta_k(g(\tilde{y}^k) - g(y^k)) + \rho\beta_k B^\top HB(y^k - \tilde{y}^k)\|^2 \\ & +2\|\beta_k(h(\tilde{z}^k) - h(z^k)) + \rho\beta_k C^\top HC(z^k - \tilde{z}^k)\|^2 \\ & +2\|G(w^k - \tilde{w}^k)\|^2 \\ \le\ & \frac{\eta^2 r^2}{2}\|x^k - \tilde{x}^k\|^2 + \frac{\eta^2 s^2}{2}\|y^k - \tilde{y}^k\|^2 + \frac{\eta^2 p^2}{2}\|z^k - \tilde{z}^k\|^2 + 2\|G(w^k - \tilde{w}^k)\|^2 \\ \le\ & \frac{\eta^2}{2}\max\left(r^2, s^2, p^2\right)\|w^k - \tilde{w}^k\|^2 + 2\lambda_l(G^\top G)\|w^k - \tilde{w}^k\|^2 \\ \le\ & \alpha_2\|w^k - \tilde{w}^k\|^2, \end{aligned} \tag{2.18}$$

where $\alpha_2 > 0$ is a constant. Therefore, it follows from (2.7) and (2.14) that $\alpha_k \ge \frac{\alpha_1}{\alpha_2}$. This completes the proof.

3. Basic results

First, we give some useful lemmas before proving the global convergence of the proposed method.

Lemma 3.1. *For given $w^k = (x^k, y^k, z^k, \lambda^k) \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++} \times \mathbb{R}^{n_3}_{++} \times \mathbb{R}^l$, let $\tilde{w}^k$ be generated by (2.3a)–(2.3d). Then, for any $w = (x, y, z, \lambda) \in \mathscr{W}$,*

$$(w - \tilde{w}^k)^\top d_2(w^k, \tilde{w}^k) \ \ge\ (w - \tilde{w}^k)^\top d_1(w^k, \tilde{w}^k) - \tfrac{\mu}{2}\|x^k - \tilde{x}^k\|_R^2 - \tfrac{\mu}{2}\|y^k - \tilde{y}^k\|_S^2 - \tfrac{\mu}{2}\|z^k - \tilde{z}^k\|_P^2. \tag{3.1}$$

Proof. Applying Lemma 2.2 to (2.3a) with $u^k = x^k$, $u = \tilde{x}^k$, $v = x$ in (2.13) and

$$q(u) = \beta_k(f(\tilde{x}^k) - A^\top[\lambda^k - H(Ax^k + By^k + Cz^k - b)]) - \xi_x^k,$$

we get

$$\begin{aligned} & (x - \tilde{x}^k)^\top\left\{\beta_k(f(\tilde{x}^k) - A^\top[\lambda^k - H(Ax^k + By^k + Cz^k - b)]) - \xi_x^k\right\} \\ \ge\ & \tfrac{1+\mu}{4}\left(\|\tilde{x}^k - x\|_R^2 - \|x^k - x\|_R^2\right) + \tfrac{1-\mu}{4}\|x^k - \tilde{x}^k\|_R^2. \end{aligned} \tag{3.2}$$

Observe that

$$\frac{1}{2}(x - \tilde{x}^k)^\top R(x^k - \tilde{x}^k) = \frac{1}{4}\left(\|\tilde{x}^k - x\|_R^2 - \|x^k - x\|_R^2\right) + \frac{1}{4}\|x^k - \tilde{x}^k\|_R^2. \tag{3.3}$$

Adding (3.2) and (3.3), we obtain

$$\begin{aligned} & (x - \tilde{x}^k)^\top\Big\{\tfrac{(1+\mu)}{2}R(x^k - \tilde{x}^k) - \beta_k f(\tilde{x}^k) + \beta_k A^\top\tilde{\lambda}^k + \xi_x^k \\ & -\beta_k A^\top H\left(A(x^k - \tilde{x}^k) + B(y^k - \tilde{y}^k) + C(z^k - \tilde{z}^k)\right)\Big\} \le \tfrac{\mu}{2}\|x^k - \tilde{x}^k\|_R^2. \end{aligned} \tag{3.4}$$

Similarly, applying Lemma 2.2 to (2.3b) with $u^k = y^k$, $u = \tilde{y}^k$, $v = y$ in (2.13) and

$$q(u) = \beta_k(g(\tilde{y}^k) - B^\top[\lambda^k - H(A\tilde{x}^k + By^k + Cz^k - b)]) - \xi_y^k,$$

we get

$$\begin{array}{l}(y-\tilde{y}^k)^\top\{\beta_k(g(\tilde{y}^k)-B^\top[\lambda^k-H(A\tilde{x}^k+By^k+Cz^k-b)])-\xi_y^k\}\\ \geq \ \frac{1+\mu}{4}\left(\|\tilde{y}^k-y\|_S^2-\|y^k-y\|_S^2\right)+\frac{1-\mu}{4}\|y^k-\tilde{y}^k\|_S^2.\end{array} \tag{3.5}$$

Observe that

$$\frac{1}{2}(y-\tilde{y}^k)^\top S(y^k-\tilde{y}^k)=\frac{1}{4}\left(\|\tilde{y}^k-y\|_S^2-\|y^k-y\|_S^2\right)+\frac{1}{4}\|y^k-\tilde{y}^k\|_S^2. \tag{3.6}$$

Adding (3.5) and (3.6), we have

$$\begin{array}{l}(y-\tilde{y}^k)^\top\Big\{\frac{(1+\mu)}{2}S(y^k-\tilde{y}^k)-\beta_k g(\tilde{y}^k)+\beta_k B^\top\tilde{\lambda}^k+\xi_y^k\\ -\beta_k B^\top H(B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k))\Big\}\leq\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2.\end{array} \tag{3.7}$$

Similarly, we have

$$\begin{array}{l}(z-\tilde{z}^k)^\top\Big\{\frac{(1+\mu)}{2}P(z^k-\tilde{z}^k)-\beta_k h(\tilde{z}^k)+\beta_k C^\top\tilde{\lambda}^k+\xi_z^k-\beta_k C^\top HC(z^k-\tilde{z}^k)\Big\}\\ \leq\frac{\mu}{2}\|z^k-\tilde{z}^k\|_P^2.\end{array} \tag{3.8}$$

It follows from (3.4), (3.7), (3.8), (2.3d) and (2.5) that

$$(w-\tilde{w}^k)^\top(d_1(w^k,\tilde{w}^k)-d_2(w^k,\tilde{w}^k))-\frac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2-\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2-\frac{\mu}{2}\|z^k-\tilde{z}^k\|_P^2\leq 0. \tag{3.9}$$

This completes the proof.

Lemma 3.2. *For given $w^k=(x^k,y^k,z^k,\lambda^k)\in\mathbb{R}^{n_1}_{++}\times\mathbb{R}^{n_2}_{++}\times\mathbb{R}^{n_3}_{++}\times\mathbb{R}^l$, let $\tilde{w}^k$ be generated by (2.3a)–(2.3d). Then, for any $w^*=(x,y,z,\lambda)\in\mathscr{W}^*$,*

$$\begin{array}{rcl}(\tilde{w}^k-w^*)^\top d_2(w^k,\tilde{w}^k) & \geq & \varphi(w^k,\tilde{w}^k)-(w^k-\tilde{w}^k)^\top d_1(w^k,\tilde{w}^k)\\ & & +\frac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2+\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2+\frac{\mu}{2}\|z^k-\tilde{z}^k\|_P^2.\end{array} \tag{3.10}$$

Proof. Recalling the definition in (2.9), we have

$$\begin{array}{rcl}\varphi(w^k,\tilde{w}^k) & = & (w^k-\tilde{w}^k)^\top d_1(w^k,\tilde{w}^k)-\frac{\mu}{2}\|x^k-\tilde{x}^k\|_R^2-\frac{\mu}{2}\|y^k-\tilde{y}^k\|_S^2-\frac{\mu}{2}\|z^k-\tilde{z}^k\|_P^2\\ & & +\beta_k(\lambda^k-\tilde{\lambda}^k)^\top\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right).\end{array} \tag{3.11}$$

Using the monotonicity of f,g and h, we obtain

$$\begin{pmatrix}\tilde{x}^k-x^*\\ \tilde{y}^k-y^*\\ \tilde{z}^k-z^*\\ \tilde{\lambda}^k-\lambda^*\end{pmatrix}^\top\begin{pmatrix}f(\tilde{x}^k)-A^\top\tilde{\lambda}^k\\ g(\tilde{y}^k)-B^\top\tilde{\lambda}^k\\ h(\tilde{z}^k)-C^\top\tilde{\lambda}^k\\ A\tilde{x}^k+B\tilde{y}^k+C\tilde{z}^k-b\end{pmatrix}\geq\begin{pmatrix}\tilde{x}^k-x^*\\ \tilde{y}^k-y^*\\ \tilde{z}^k-z^*\\ \tilde{\lambda}^k-\lambda^*\end{pmatrix}^\top\begin{pmatrix}f(x^*)-A^\top\lambda^*\\ g(y^*)-B^\top\lambda^*\\ h(z^*)-C^\top\lambda^*\\ Ax^*+By^*+Cz^*-b\end{pmatrix}\geq 0. \tag{3.12}$$

It follows from (3.12) that

$$\begin{array}{rl} & (\tilde{w}^k-w^*)^\top d_2(w^k,\tilde{w}^k)\\ \geq & \beta_k(\tilde{w}^k-w^*)^\top\begin{pmatrix}A^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right)\\ B^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right)\\ C^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right)\\ 0\end{pmatrix}\\ = & \beta_k(A\tilde{x}^k+B\tilde{y}^k+C\tilde{z}^k-b)^\top H\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right)\\ = & \beta_k(\lambda^k-\tilde{\lambda}^k)^\top\left(A(x^k-\tilde{x}^k)+B(y^k-\tilde{y}^k)+C(z^k-\tilde{z}^k)\right).\end{array} \tag{3.13}$$

Combining (3.11) and the above inequality, we can get the assertion of this lemma immediately.

Theorem 3.1. *Let $w^* \in \mathscr{W}^*, w^{k+1}(\alpha_k)$ be defined by (2.6) and*

$$\Theta(\alpha_k) := \|w^k - w^*\|^2 - \|w^{k+1}(\alpha_k) - w^*\|^2. \tag{3.14}$$

Then

$$\Theta(\alpha_k) \quad \geq \quad \sigma(2\alpha_k \varphi(w^k, \tilde{w}^k) - \alpha_k^2 \|d_1(w^k, \tilde{w}^k)\|^2), \tag{3.15}$$

where

$$w_p^k := P_{\mathscr{W}}[w^k - \alpha_k d_2(w^k, \tilde{w}^k)]. \tag{3.16}$$

Proof. Since $w^* \in \mathscr{W}^*$, we conclude from (2.2) that

$$\|w_p^k - w^*\|^2 \leq \|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 - \|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2. \tag{3.17}$$

From (2.6), we get

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

Using the equality that $2(a+b)^\top b = \|a+b\|^2 - \|a\|^2 + \|b\|^2$ with $a = w^k - w_p^k$, and $b = w_p^k - w^*$, and (3.17), we obtain

$$\begin{aligned}
\|w^{k+1}(\alpha_k) - w^*\|^2 \quad &= \quad (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w_p^k - w^*\|^2 - \sigma(1-\sigma)\|w^k - w_p^k\|^2 \\
&\leq \quad (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 \\
&\quad\quad -\sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2 - \sigma(1-\sigma)\|w^k - w_p^k\|^2 \\
&\leq \quad (1-\sigma)\|w^k - w^*\|^2 + \sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w^*\|^2 \\
&\quad\quad -\sigma\|w^k - \alpha_k d_2(w^k, \tilde{w}^k) - w_p^k\|^2.
\end{aligned} \tag{3.18}$$

Using the definition of $\Theta(\alpha_k)$ and (3.18), we get

$$\Theta(\alpha_k) \quad \geq \quad \sigma\|w^k - w_p^k\|^2 + 2\sigma\alpha_k (w_p^k - w^*)^\top d_2(w^k, \tilde{w}^k). \tag{3.19}$$

Applying (3.1) with $w = w_p^k$, we obtain

$$(w_p^k - \tilde{w}^k)^\top d_2(w^k, \tilde{w}^k) \geq (w_p^k - \tilde{w}^k)^\top d_1(w^k, \tilde{w}^k) - \frac{\mu}{2}\|x^k - \tilde{x}^k\|_R^2 - \frac{\mu}{2}\|y^k - \tilde{y}^k\|_S^2 - \frac{\mu}{2}\|z^k - \tilde{z}^k\|_P^2. \tag{3.20}$$

Adding (3.10) and (3.20), we get

$$(w_p^k - w^*)^\top d_2(w^k, \tilde{w}^k) \quad \geq \quad (w_p^k - w^k)^\top d_1(w^k, w^k) + \varphi(w^k, \tilde{w}^k). \tag{3.21}$$

Applying (3.21) to the last term on the right side of (3.19), we obtain

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

This theorem is proved.

4. THE CONVERGENCE OF THE PROPOSED METHOD

In this section, we investigate the convergence analysis of the proposed method. From the computational point of view, a relaxation factor $\gamma \in (0,2)$ is preferable in the correction. The following result shows the contraction of the sequence generated by the proposed method.

Theorem 4.1. *Let $w^* \in \mathscr{W}^*$ be a solution of SVI$_3$ and let $w^{k+1}(\gamma\alpha_k)$ be generated by (2.6). Then w^k and $\tilde{w}^k$ are bounded, and*

$$\|w^{k+1}(\gamma\alpha_k) - w^*\|^2 \leq \|w^k - w^*\|^2 - c\|w^k - \tilde{w}^k\|^2, \tag{4.1}$$

where

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

Proof. It follows from (3.15), (2.14) and (2.15) 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

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

Thus, $\{w^k\}$ is a bounded sequence. It follows from (4.1) that

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

which means that

$$\lim_{k\to\infty}\|w^k-\tilde{w}^k\|=0. \tag{4.2}$$

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

Using Lemma 3.1 and Theorem 4.1, we are in a position to prove the convergence of the proposed method. The following result can be proved by using the technique of Theorem 4.2 in [6].

Theorem 4.2. *The sequence $\{w^k\}$ generated by the proposed method converges to some w^∞, which is a solution of SVI$_3$.*

Proof. Since $\{w^k\}$ is bounded, it has at least one cluster point. Let w^∞ be a cluster point of $\{w^k\}$ and the subsequence $\{w^{k_j}\}$ converges to w^∞. Since $\mathscr{W}$ is a closed set, we have $w^\infty\in\mathscr{W}$. By the construction of β_k, we have that $0<\beta_k\le\beta_0,\forall k$. It follows from (4.2) that

$$\lim_{j\to\infty}d_1(w^{k_j},\tilde{w}^{k_j})=0\quad\text{and}\quad\lim_{j\to\infty}\frac{d_2(w^{k_j},\tilde{w}^{k_j})}{\beta_{k_j}}=Q(w^\infty). \tag{4.3}$$

Moreover, (4.3) and (3.1) imply that

$$\lim_{j\to\infty}\frac{(w-w^{k_j})^\top d_2(w^{k_j},\tilde{w}^{k_j})}{\beta_{k_j}}\ge 0,\qquad\forall w\in\mathscr{W}. \tag{4.4}$$

Consequently,

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

which means that w^∞ is a solution of SVI$_3$.

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 $\varepsilon>0$, we have that there exists an $l>0$ such that

$$\|\tilde{w}^{k_l}-w^\infty\|<\frac{\varepsilon}{2}\quad\text{and}\quad\|w^{k_l}-\tilde{w}^{k_l}\|<\frac{\varepsilon}{2}. \tag{4.5}$$

Therefore, for any $k\ge k_l$, it follows from (4.1) and (4.5) that

$$\|w^k-w^\infty\|\le\|w^{k_l}-w^\infty\|\le\|w^{k_l}-\tilde{w}^{k_l}\|+\|\tilde{w}^{k_l}-w^\infty\|<\varepsilon.$$

This implies that the sequence $\{w^k\}$ converges to w^∞, which is a solution of SVI$_3$.

5. Preliminary computational results

In this section, we present some numerical experiments to illustrate our algorithm and convergence result.

We denote by $0_{n\times n} \in \mathbb{R}^{n\times n}$ the null matrix, and by $I_{n\times n} \in \mathbb{R}^{n\times n}$ the identity matrix. Let $S^n = \{X \in \mathbb{R}^{n\times n} : X^\top = X\}$, $S^n_+ = \{X \in S^n : X \succeq 0_{n\times n}\}$, $\mathbb{B} = \{X \in S^n : H_v \le X \le H_u\}$, and $H_v, H_u \in S^n$ be given proper matrices. The matrix inequality $S \preceq T$ means that T-S is a positive semi-definite matrix, while $S \le T$ means that $S_{ij} \le T_{ij}$ $(\forall i, j \in I = \{1,2,\dots,n\})$. For $C \in \mathbb{R}^{n\times n}$, we denote by $\|C\|_F$ the matrix Fröbenis norm of C, i.e., $\|C\|_F = (\sum_{i=1}^n \sum_{j=1}^n |C_{ij}|^2)^{1/2}$. Note that the matrix Fröbenis norm is induced by the inner product

$$\langle A, B\rangle = \operatorname{trace}(A^\top B).$$

We consider the following optimization problem with matrix variables, which was studied in [7] and [14]:

$$\min\left\{\frac{1}{2}\|X-Q\|_F^2 : 0_{n\times n} \preceq X \preceq M, X \in \mathbb{B}\right\}, \tag{5.1}$$

where $Q, M \in S^n$ are given proper matrices. Note that problem (5.1) can be reformulated into the following separable form:

$$\min\left\{\tfrac{1}{2}\|X-Q\|_F^2 + \tfrac{1}{2}\|Y+Q-M\|_F^2 + \tfrac{1}{2}\|Z-Q\|_F^2\right\} \tag{5.2}$$

$$\text{such that } X+Y=M, \tag{5.3}$$

$$Y-Z=0_{n\times n}, \tag{5.4}$$

$$Y+Z=M, \tag{5.5}$$

where $X, Y \in S^n_+$, $Z \in \mathbb{B}$. Then, the problem is equivalent to the following structured variational inequality problem, which consists of finding $w^* = (X^*, Y^*, Z^*, \lambda^*) \in \Omega := S^n_+ \times S^n_+ \times \mathbb{B} \times \mathbb{R}^{3n\times n}$ such that

$$\begin{cases} \langle X-X^*, f(X^*) - A^\top\lambda^*\rangle \ge 0, \\ \langle Y-Y^*, g(Y^*) - B^\top\lambda^*\rangle \ge 0, \\ \langle Z-Z^*, h(Z^*) - C^\top\lambda^*\rangle \ge 0, \qquad \forall w = (X,Y,Z,\lambda) \in \Omega, \\ AX^* + BY^* + CZ^* - b = 0, \end{cases} \tag{5.6}$$

where

$$A = \begin{pmatrix} I_{n\times n} \\ I_{n\times n} \\ 0_{n\times n} \end{pmatrix}, \quad B = \begin{pmatrix} I_{n\times n} \\ 0_{n\times n} \\ I_{n\times n} \end{pmatrix}, \quad C = \begin{pmatrix} 0_{n\times n} \\ -I_{n\times n} \\ I_{n\times n} \end{pmatrix}, \quad b = \begin{pmatrix} M \\ 0_{n\times n} \\ M \end{pmatrix}$$

and

$$f(X) = X - Q, \quad g(Y) = Y + Q - M, \quad h(Z) = Z - Q.$$

The entries of Q are randomly with the restriction that $Q_{ii} \in (0,2)$ and $Q_{ij} \in (-1,1)$. The matrices H_v and H_u are given by

$$(H_u)_{jj} = (H_v)_{jj} = 1 \text{ and } (H_u)_{ij} = -(H_v)_{ij} = 0.1, \quad \forall i \ne j, i, j = 1,2,\dots,n.$$

The matrix M has the following form:

$$M = U\Sigma U, \quad U = I_{n\times n} - 2uu^\top, \quad \Sigma = \operatorname{diag}(e_1, e_2, \dots, e_n),$$

where u is a random unit vector, and e_i $(i = 1,2,\dots,n)$ is a given eigenvalue of the matrix M. For simplification, we take $R = rI_{n\times n}, S = sI_{n\times n}$, $P = pI_{n\times n}$ and $H = I_{n\times n}$, where $r > 0$, $s > 0$ and $p > 0$ are scalars. In all tests, for the proposed method and the method in [5], we take $\mu = 0.01$, $\gamma = 1.9$, $\sigma =$

$0.1, r = 0.1, s = 5, p = 10$ and for the method in [7], we take $\tau_1 = \tau_1 = \frac{1}{2}$, and $\gamma = 1.9$, $(X^0, Y^0, Z^0, \lambda^0) = (I_{n\times n}, I_{n\times n}, I_{n\times n}, 0_{3n\times n})$ as the initial point in the test. The iteration is stopped if

$$\frac{\max(abs(w^k - \tilde{w}^k))}{\max(abs(w^0 - \tilde{w}^0))} \leq 10^{-5}.$$

$abs(D)$ is the absolute value of matrix D, that is, if $D = [d_{ij}]$, where $d_{ij} \in \mathbb{R}, i = 1, \cdots n, j = 1, \cdots n$. Then, $abs(D) = [|d_{ij}|]$. All codes were written in Matlab, and we compare the proposed method with those in [5, 7, 14]. The iteration numbers denoted by k, and the computational time for the problem (5.1) with different dimensions are given in Tables 1–3.

TABLE 1. Numerical results for problem (5.1) with $e_i \in (1.8, 2)$.

Dimension of the problem	Values of β_0	The method in [5]		The method in [7]		The method in [14]	
		k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)
100	0.5	627	32.89	742	21.17	781	23.35
	1	261	13.05	345	10.15	358	11.39
	100	17	0.98	13	0.53	14	0.54
200	0.5	542	139.67	784	114.53	789	148.57
	1	270	87.15	361	61.04	362	61.21
	100	17	5.02	15	2.73	14	3.07
300	0.5	497	411.74	720	416.22	729	425.18
	1	244	185.63	329	156.61	330	186.11
	100	17	13.32	14	6.65	14	8.11

Dimension of the problem	The proposed method	
	k	CPU(Sec.)
n=100	9	0.75
n=200	9	4.52
n=300	9	15.27

TABLE 2. Numerical results for problem (5.1) with $e_i \in (2, 3)$.

Dimension of the problem	Values of β_0	The method in [5]		The method in [7]		The method in [14]	
		k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)
100	0.5	670	37.52	842	25.74	790	25.04
	1	258	14.63	340	10.84	343	11.46
	100	17	1.13	12	0.47	9	0.29
200	0.5	652	168.62	837	118.97	845	129.25
	1	295	78.44	393	58.34	393	68.99
	100	17	5.77	13	2.59	10	2.36
300	0.5	582	482.25	797	330.16	797	347.96
	1	281	195.84	374	149.84	380	167.42
	100	17	15.65	14	7.41	12	6.54

Dimension of the problem	The proposed method	
	k	CPU(Sec.)
n=100	7	0.65
n=200	7	4.17
n=300	8	14.12

TABLE 3. Numerical results for problem (5.1) with $e_i \in (10, 12)$.

Dimension of the problem	Values of β_0	The method in [5]		The method in [7]		The method in [14]	
		k	CPU(Sec.)	k	CPU(Sec.)	k	CPU(Sec.)
100	0.5	663	37.66	629	25.43	723	30.85
	1	174	10.15	213	11.04	220	7.94
	100	18	1.09	14	0.46	14	0.46
200	0.5	628	186.42	709	169.65	748	175.76
	1	260	88.38	325	64.83	342	68.38
	100	18	6.29	14	2.52	14	2.58
300	0.5	673	669.01	837	547.66	837	616.85
	1	305	349.46	377	243.76	398	214.48
	100	18	20.84	12	7.02	14	10.16

Dimension of the problem	The proposed method	
	k	CPU(Sec.)
n=100	8	0.66
n=200	8	4.18
n=300	9	13.14

Tables 1–3 report the comparison between the methods of [5, 7, 14] and the proposed method. The number of iteration has great advantage and a faster convergence speed. From tables 1–3, we could see that the methods proposed in [5, 7, 14] work well when β_0 is large. If parameter β_0 is small, the iteration numbers and the computational time can increase significantly.

6. THE CONCLUSION

In this paper, by applying the SQP regularization to the subproblems decomposed by ADM, we proposed an SQP alternating direction method for solving the linearly constrained separable convex programming with three separable operators. We used a self-adaptive method that adjusts the scalar parameter automatically. Under standard assumptions, the global convergence of the proposed method was proved. The numerical results show the high efficiency and robustness of the proposed method.

REFERENCES

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

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

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

[5] A. Bnouhachem, Q.H. Ansari, S. Al-Homidan, SQP alternating direction for structured vriational inequality, J. Nonlinear Convex Anal. 19 (2018), 461-476.

[6] A. Bnouhachem, T.M. Rassias, On descent alternating direction method with LQP regularization for the structured variational inequalities, Optim. Lett. 14 (2020), 1353–1369.

[7] C. Cao, D.R. Han, L.L. Xu, A new partial splitting augmented Lagrangian method for minimizing the sum of three convex functions, Appl. Math. Computat. 219 (2013), 5449-5457.

[8] S.Y. Cho, X. Qin, L. Wang, Strong convergence of a splitting algorithm for treating monotone operators, Fixed Point Theory Appl. 2014 (2014), 94.

[9] R. Glowinski, A. Marrocco, Sur lpproximation par éléments finis d'ordre un et la résolution par pénalisation-dualité d'une classe de problémes de Dirichlet non linéaires, Revue Fr. Autom. Inform. Rech. Oper. Anal. Numer. 2 (1975), 41-76

[10] B.S. He, Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities, Comput. Optim. Appl. 42 (2009), 195-212.

[11] Z.K. Jiang, A. Bnouhachem, A projection-based prediction-correction method for structured monotone variational inequalities, Appl. Math. Comput. 202 (2008), 747-759.

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

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

[14] Z. Peng, D.H. Wu, A partial splitting augmented Lagrangian method for constrained matrix optimization problem, Comput. Math. Appl. 60 (2010), 1515-1524.

[15] M. Tao, X.M. Yuan, On the $O(1/t)$ convergence rate of alternating direction method with Logarithmic-quadratic proximal regularization, SIAM J. Optim. 22 (2012), 1431-1448.

[16] K. Wang, L. Xu, D. Han, A new parallel splitting descent method for structured variational inequalities, J. Industrial Management Optim. 10 (2014), 461-476.

[17] C. Wu, Y. Xiao, P. Li, Semi-proximal augmented Lagrangian method for sparse estimation of high-dimensional inverse covariance matrices, J. Appl. Numer. Optim. 2 (2020), 155-169.

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