

AN ALTERNATE MINIMIZATION METHOD BEYOND POSITIVE DEFINITE PROXIMAL REGULARIZATION: CONVERGENCE AND COMPLEXITY

XIANFU WANG¹ XIAOQING OU², TAO ZHANG³, JIAWEI CHEN^{1,*}

¹*School of Mathematics and Statistics, Southwest University, Chongqing 400715, China*

²*College of Management, Chongqing College of Humanities, Science & Technology, Chongqing 401524, China*

³*School of Information and Mathematics, Yangtze University, Jingzhou 434023, China*

Abstract. In this paper, an alternate minimization method beyond positive definite proximal regularization is introduced for solving linearly constrained separable convex optimization problems. The proposed method can be interpreted as the prediction-correction method from the perspective of variational inequalities. The convergence of the proposed method is established without strong convexity. Finally, the iteration complexity of the proposed method is also derived in the ergodic sense.

Keywords. Alternate minimization method; Convex optimization; Complexity; Positive-indefinite proximal regularization; Variational inequality.

2010 Mathematics Subject Classification. 47H05, 90C33.

1. INTRODUCTION

In this paper, we consider the following linearly constrained convex optimization problem with separable structure:

$$\min \{ \theta_1(x) + \theta_2(y) \mid Ax + By = b, x \in \mathcal{X}, y \in \mathcal{Y} \}, \quad (1.1)$$

where $A \in \mathbb{R}^{m \times n_1}$, $B \in \mathbb{R}^{m \times n_2}$, $b \in \mathbb{R}^m$, $\mathcal{X} \subset \mathbb{R}^{n_1}$ and $\mathcal{Y} \subset \mathbb{R}^{n_2}$ are nonempty closed convex sets, $\theta_1(x) : \mathbb{R}^{n_1} \rightarrow \mathbb{R}$ and $\theta_2(y) : \mathbb{R}^{n_2} \rightarrow \mathbb{R}$ are two proper convex and lower semicontinuous (not necessarily smooth) functions. Throughout this paper, the solution set of problem (1.1) is assumed to be nonempty.

For solving (1.1), various numerical methods have been proposed, such as, proximal point method, incremental gradient projection method, extragradient method and alternate minimization method; see, e.g., [2, 3, 5, 6, 7, 8, 9, 16, 17, 19, 20, 21, 22, 24, 27, 28] and the references therein. Alternate minimization method is a powerful method for constrained convex optimization models with separable structures, which fully utilize the structure properties of the considered problems. The alternate minimization method contained many efficient numerical methods such as alternate direction method of multipliers (ADMM) and various proximal ADMMs. In particular, the proximal ADMM involves the positive definiteness of the corresponding proximal matrix. Nevertheless, the positive definiteness of the related

*Corresponding author.

E-mail address: J.W.Chen713@163.com (J. Chen).

Received May 16, 2019; Accepted July 10, 2019.

proximal matrix involved in proximal ADMM usually affect the overall convergence rate of the method as reported in the literature.

The Lagrangian function of (1.1) is defined by

$$L(x, y, \lambda) := \theta_1(x) + \theta_2(y) - \lambda^\top (Ax + By - b), \quad (x, y, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m,$$

where $\lambda \in \mathbb{R}^m$ the Lagrangian multiplier. The augmented Lagrangian function of problem (1.1) is defined by

$$\mathcal{L}_\beta(x, y, \lambda) := \theta_1(x) + \theta_2(y) - \lambda^\top (Ax + By - b) + \frac{\beta}{2} \|Ax + By - b\|^2, \quad (x, y, \lambda) \in \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m, \quad (1.2)$$

where $\lambda \in \mathbb{R}^m$ is the Lagrangian multiplier and $\beta > 0$ is a penalty parameter.

For a given (y^k, λ^k) , the iterative scheme of alternate minimization method based on augmented Lagrangian function for solving optimization problem (1.1) reads as follows:

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\ y^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x^{k+1}, y, \lambda^k) \mid y \in \mathcal{Y} \}, \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b), \end{cases} \quad (1.3)$$

where $\mathcal{L}_\beta(x, y, \lambda)$ is defined by (1.2).

The method (1.3) is also called as the alternating direction method of multipliers (ADMM). In (1.3), the main task is to solve the subproblem of (1.3) to get the iteration point (x^{k+1}, y^{k+1}) . An effective method is to regularize the subproblem by using quadratic proximal term. Without loss of generality, we only discuss the case where the y -subproblem of (1.3) is regularized as follows:

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\ y^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x^{k+1}, y, \lambda^k) + \frac{1}{2} \|y - y^k\|_D^2 \mid y \in \mathcal{Y} \}, \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b). \end{cases} \quad (1.4)$$

In (1.4), $\frac{1}{2} \|y - y^k\|_D^2$ is the quadratic proximal regularization term and $D \in \mathbb{R}^{n_2 \times n_2}$ is the proximal matrix, which is usually required to be positive definite in the literature; see, e.g., [23, 25] and the references therein. Observe that it is also probable to consider the case when the x -subproblem and y -subproblem are both proximally regularized; see, e.g., [10]. When $\mathcal{Y} = \mathbb{R}^{n_2}$, if the proximal matrix D in (1.4) is defined by

$$D = \gamma I_{n_2} - \beta B^\top B \text{ and } \gamma > \beta \|B^\top B\|, \quad (1.5)$$

then the y -subproblem in (1.4) can be replaced by

$$y^{k+1} = \operatorname{argmin} \left\{ \theta_2(y) + \frac{\gamma}{2} \|y - (y^k + \frac{1}{\gamma} q_k)\|^2 \mid y \in \mathbb{R}^{n_2} \right\},$$

where

$$q_k = B^\top \left[\lambda^k - \beta(Ax^{k+1} + By^k - b) \right]. \quad (1.6)$$

Hence, (1.4) with D given in (1.5) can be equivalently reformulated as

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\ y^{k+1} = \operatorname{argmin} \left\{ \theta_2(y) + \frac{\gamma}{2} \|y - (y^k + \frac{1}{\gamma} q_k)\|^2 \mid y \in \mathcal{Y} \right\}, \\ \lambda^{k+1} = \lambda^k - \beta(Ax^{k+1} + By^{k+1} - b), \end{cases} \quad (1.7)$$

where q_k is given in (1.6).

Recently, He, Ma and Yuan [14] proposed the possibility of relaxing the positive definiteness requirement of (1.4) and affirmatively gave that it was not necessary to require the positive definiteness of the proximal matrix D in (1.4). The parameter γ determines the step size for solving the y -subproblem in (1.4). If the choice of (1.5) is relaxed to

$$D = \tau\gamma I_{n_2} - \beta B^\top B, \quad (1.8)$$

where $\gamma > \beta\|B^\top B\|$ and $\tau \in (0, 1)$, then the quadratic proximal term $\frac{1}{2}\|y - y^k\|_D^2$ with D in (1.8) plays a lighter weight in the objective and y can be updated with a large step size. Finally, they suggested choosing an optimal choice that is $\tau = 0.8$.

Recently, it has been mentioned that the large step length may accelerate the convergence empirically in the literature. He, Ma and Yuan [15] investigated the step length for updating the variable in the augmented Lagrangian method (shortly, ALM), and proposed the positive-indefinite proximal ALM (shortly, PIDP-ALM) with a general step size for updating the dual variable:

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, \lambda^k) + \frac{1}{2}\|x - x^k\|_{D_0}^2 \mid x \in \mathcal{X} \}, \\ \lambda^{k+1} = \lambda^k - \gamma\beta(Ax^{k+1} - b), \gamma \in (0, 2), \end{cases} \quad (1.9)$$

where $D_0 = D - (1 - \tau)\beta A^\top A$ with an arbitrarily given positive-definite matrix $D \in \mathbb{R}^{n_1 \times n_1}$, $\tau \in (0, 1)$ and

$$\mathcal{L}_\beta(x, \lambda) = \theta(x) - \lambda^\top(Ax - b) + \frac{\beta}{2}\|Ax - b\|^2.$$

In order to ensure the convergence of PIDP-ALM (1.9), they deduced the relationship between the two step size parameters τ and γ , i.e., $\tau > \frac{2+\gamma}{4}$. Based on the Peaceman-Rachford splitting method (shortly, PRSM), He et al. [11] introduced a relaxed factor α for ensuring the strictly contraction of the Peaceman-Rachford splitting method, i.e.,

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b), \\ y^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x^{k+1}, y, \lambda^{k+\frac{1}{2}}) \mid y \in \mathcal{Y} \}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(Ax^{k+1} + By^{k+1} - b), \alpha \in (0, 1). \end{cases} \quad (1.10)$$

They also established the worst-case $\mathcal{O}(1/t)$ convergence rate of this strictly contractive PRSM in a nonergodic sense and gave some applications in statistical learning and image processing.

Motivated by the works [11, 14, 15], we develop the convergence and iteration complexity of the alternate minimization method with the positive-indefinite proximal regularization terms for solving problem (1.1):

$$\begin{cases} x^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x, y^k, \lambda^k) \mid x \in \mathcal{X} \}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b), \\ y^{k+1} = \operatorname{argmin} \{ \mathcal{L}_\beta(x^{k+1}, y, \lambda^{k+\frac{1}{2}}) + \frac{1}{2}\|y - y^k\|_{D_0}^2 \mid y \in \mathcal{Y} \}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b), \end{cases} \quad (1.11)$$

where $\alpha \in [0, 1)$, $\gamma \in (0, 2)$, D_0 is defined by

$$D_0 = D - (1 - \tau)\beta B^\top B, \quad (1.12)$$

with an arbitrarily given positive-definite matrix $D \in \mathbb{R}^{n_2 \times n_2}$ and $\tau \in (0, 1]$. Clearly, D_0 defined by (1.12) is not necessarily positive definite when $\tau \in (0, 1)$.

In (1.11), one of the primary purpose is to obtain the relationships among the step length parameters τ , α , γ and the penalty parameter β . Besides, the convergence and iteration complexity of the method (1.11) are established under suitable conditions. It is observed that method (1.11) is different from method (1.9) since it fully does exploit the separable structure of problem (1.1). Moreover, method (1.11) covers some existing algorithms as special cases. It is easy to see that if $\alpha = 0$, $\gamma = 1$ and $\tau = 1$ in (1.11), then (1.11) is reduced to positive definite linearized ADMM (1.4); if $\alpha = \gamma \in (0, 1)$ and τ satisfies $D_0 = 0$ for column full rank B in (1.11), then (1.11) is reduced to the strictly contractive PRSM (1.10); if $\alpha = 0$ and $\gamma = 1$ in (1.11), then (1.11) is reduced to the form of positive indefinite linearized ADMM in [14] with different proximal matrix D defined by (1.8).

The rest of this paper is organized as follows. In Section 2, we recall some preliminaries for the considered problem. The variational characterization of the proposed method (1.11) is presented in Section 3. In Section 4, we analyze the convergence properties of the proposed method (1.11) and the relationship among two step size parameters τ , α , γ and the penalty parameter β to ensure the convergence of the proposed method (1.11). In Section 5, the iterative complexity of the proposed method is established in the Ergodic sense.

2. PRELIMINARIES

In this section, we recall some definitions and basic results.

A pair $((x^*, y^*), \lambda^*)$ is called a saddle point of the Lagrangian function iff,

$$L(x^*, y^*, \lambda) \leq L(x^*, y^*, \lambda^*) \leq L(x, y, \lambda^*), \quad \forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \lambda \in \mathbb{R}^m. \quad (2.1)$$

It is easy to reformulate (2.1) as the variational inequalities:

$$\begin{cases} x^* \in \mathcal{X}, \theta_1(x) - \theta_1(x^*) + (x - x^*)^\top (-A^\top \lambda^*) \geq 0, & \forall x \in \mathcal{X}, \\ y^* \in \mathcal{Y}, \theta_2(y) - \theta_2(y^*) + (y - y^*)^\top (B^\top \lambda^*) \geq 0, & \forall y \in \mathcal{Y}, \\ \lambda^* \in \mathbb{R}^m, (\lambda - \lambda^*)^\top (Ax^* + By^* - b) \geq 0, & \forall \lambda \in \mathbb{R}^m. \end{cases}$$

The above variational characterization can be rewritten as the compactly form of finding $\omega^* \in \Omega$ such that

$$\theta(u) - \theta(u^*) + (\omega - \omega^*)^\top F(\omega^*) \geq 0, \quad \forall \omega \in \Omega, \quad (2.2)$$

where $\Omega := \mathcal{X} \times \mathcal{Y} \times \mathbb{R}^m$, $u := (x, y)^\top$, $\omega := (x, y, \lambda)^\top$, $\theta(u) := \theta_1(x) + \theta_2(y)$ and

$$F(\omega) := \begin{pmatrix} -A^\top \lambda \\ -B^\top \lambda \\ Ax + By - b \end{pmatrix}.$$

We denote by Ω^* the solution set of mixed variational inequality (2.2). For the convenience, we present the following notations:

$$v := \begin{pmatrix} y \\ \lambda \end{pmatrix}, \mathcal{V} := \mathcal{Y} \times \mathbb{R}^m, \mathcal{V}^* := \{(y^*, \lambda^*) | (x^*, y^*, \lambda^*) \in \Omega^*\}, \quad (2.3)$$

and the signum function of a real number z is defined as follows:

$$\text{sgn}(z) = \begin{cases} 1, & \text{if } z > 0, \\ 0, & \text{if } z = 0, \\ -1, & \text{if } z < 0. \end{cases}$$

It is easy to check that the mapping F is affine and $(\omega - \bar{\omega})^\top (F(\omega) - F(\bar{\omega})) = 0$ for all $\omega, \bar{\omega} \in \Omega$. So, Ω^* is convex because of the convexity of θ_1 and θ_2 . Therefore, $\omega^* \in \Omega$ is a solution of (2.2) if and only if it is a solution of the following mixed variational inequality:

$$\theta(u) - \theta(u^*) + (\omega - \omega^*)^\top F(\omega) \geq 0, \forall \omega \in \Omega.$$

It thus yields that

$$\Omega^* = \bigcap_{\omega \in \Omega} \left\{ \omega^* \in \Omega \mid \theta(u) - \theta(u^*) + (\omega - \omega^*)^\top F(\omega) \geq 0 \right\}.$$

In view of this, we next define the approximate solution of the variational inequality (2.2) with the accuracy $\varepsilon > 0$.

Definition 2.1. ([12]) $\tilde{\omega} \in \Omega$ is an approximation solution of (2.2) with the accuracy $\varepsilon > 0$ iff it satisfies

$$\theta(u) - \theta(\tilde{u}) + (\omega - \tilde{\omega})^\top F(\omega) \geq -\varepsilon, \forall \omega \in \mathcal{B}_{\tilde{\omega}},$$

where

$$\mathcal{B}_{\tilde{\omega}} := \{\omega \in \Omega \mid \|\omega - \tilde{\omega}\| \leq 1\}. \quad (2.4)$$

In the sequel, the following result will be frequently used.

Fact 2.1. ([14, Lemma 2.1]) *Let $\mathcal{X} \subset \mathbb{R}^n$ be a closed convex set, and let $\theta(x)$ and $f(x)$ be convex functions. Assume that f is differentiable and the solution set of problem $\min\{\theta(x) + f(x) \mid x \in \mathcal{X}\}$ is nonempty. Then*

$$x^* \in \operatorname{argmin}\{\theta(x) + f(x) \mid x \in \mathcal{X}\},$$

if and only if

$$x^* \in \mathcal{X}, \theta(x) - \theta(x^*) + (x - x^*)^\top \nabla f(x^*) \geq 0, \forall x \in \mathcal{X}.$$

3. VARIATIONAL CHARACTERIZATION OF ALGORITHM (1.11)

In this section, we will give a variational characterization of algorithm (1.11) from the perspective of variational inequalities, which can facilitate the understanding of the major difficulty in the convergence analysis of the proposed method.

Let $(x^{k+1}, y^{k+1}, \lambda^{k+1})$ be generated by the algorithm (1.11) with the given essential variables (y^k, λ^k) . For the convergence analysis of algorithm (1.11), we give some auxiliary variables: $\tilde{\omega}^k := (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$ and $\tilde{v}^k := (\tilde{y}^k, \tilde{\lambda}^k)$, where

$$\tilde{x}^k := x^{k+1}, \tilde{y}^k := y^{k+1}, \text{ and } \tilde{\lambda}^k := \lambda^k - \beta(Ax^{k+1} + By^k - b). \quad (3.1)$$

We first show the difference between the auxiliary variable $\tilde{\omega}^k$ and a solution of variational inequality (2.2)

Proposition 3.1. *Let ω^{k+1} be generated by (1.11) with the given v^k , and let $\tilde{\omega}^k$ be defined by (3.1). Then*

$$\tilde{\omega}^k \in \Omega, \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\tilde{\omega}^k) \geq (v - \tilde{v}^k)^\top Q(v^k - \tilde{v}^k), \forall \omega \in \Omega, \quad (3.2)$$

where

$$Q := \begin{pmatrix} \beta B^\top B + D_0 & -\alpha B^\top \\ -B & \frac{1}{\beta} I_m \end{pmatrix}. \quad (3.3)$$

Proof. It follows from Fact 2.1 that the subproblems of (1.11) can be respectively formulated as

$$x^{k+1} \in \mathcal{X}, \theta_1(x) - \theta_1(x^{k+1}) + (x - x^{k+1})^\top \left[-A^\top \lambda^k + \beta A^\top (Ax^{k+1} + By^k - b) \right] \geq 0, \forall x \in \mathcal{X} \quad (3.4)$$

and

$$\begin{aligned} y^{k+1} \in \mathcal{Y}, \theta_2(y) - \theta_2(y^{k+1}) + (y - y^{k+1})^\top \left[-B^\top \lambda^{k+\frac{1}{2}} + \beta B^\top (Ax^{k+1} + By^{k+1} - b) \right. \\ \left. + D_0(y^{k+1} - y^k) \right] \geq 0, \forall y \in \mathcal{Y}. \end{aligned} \quad (3.5)$$

Also, (3.5) is equivalent to the following:

$$\begin{aligned} y^{k+1} \in \mathcal{Y}, \theta_2(y) - \theta_2(y^{k+1}) + (y - y^{k+1})^\top \left[-B^\top \lambda^{k+\frac{1}{2}} + \beta B^\top (Ax^{k+1} + By^k - b) \right. \\ \left. + \beta B^\top B(y^{k+1} - y^k) + D_0(y^{k+1} - y^k) \right] \geq 0, \forall y \in \mathcal{Y}. \end{aligned} \quad (3.6)$$

Recalling $\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b) = \lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k) = \tilde{\lambda}^k + (\alpha - 1)(\tilde{\lambda}^k - \lambda^k)$, we have

$$-A^\top \lambda^k + \beta A^\top (Ax^{k+1} + By^k - b) \stackrel{(3.1)}{=} -A^\top \lambda^k + A^\top (\lambda^k - \tilde{\lambda}^k) = -A^\top \tilde{\lambda}^k$$

and

$$\begin{aligned} & -B^\top \lambda^{k+\frac{1}{2}} + \beta B^\top (Ax^{k+1} + By^{k+1} - b) + D_0(y^{k+1} - y^k) \\ &= -B^\top \left(\tilde{\lambda}^k + (\alpha - 1)(\tilde{\lambda}^k - \lambda^k) \right) + B^\top (\lambda^k - \tilde{\lambda}^k) + (\beta B^\top B + D_0)(y^{k+1} - y^k) \\ &= -B^\top \tilde{\lambda}^k + (\beta B^\top B + D_0)(\tilde{y}^k - y^k) - \alpha B^\top (\tilde{\lambda}^k - \lambda^k). \end{aligned}$$

Accordingly, the inequalities (3.4) and (3.6) can be equivalently reformulated respectively as

$$\tilde{x}^k \in \mathcal{X}, \theta_1(x) - \theta_1(\tilde{x}^k) + (x - \tilde{x}^k)^\top (-A^\top \tilde{\lambda}^k) \geq 0, \forall x \in \mathcal{X} \quad (3.7)$$

and

$$\begin{aligned} \tilde{y}^k \in \mathcal{Y}, \theta_2(y) - \theta_2(\tilde{y}^k) \\ + (y - \tilde{y}^k)^\top \left[-B^\top \tilde{\lambda}^k + (\beta B^\top B + D_0)(\tilde{y}^k - y^k) - \alpha B^\top (\tilde{\lambda}^k - \lambda^k) \right] \geq 0, \forall y \in \mathcal{Y}. \end{aligned} \quad (3.8)$$

By the definition of $\tilde{\lambda}^k$, we have

$$\tilde{\lambda}^k \in \mathbb{R}^m, (\lambda - \tilde{\lambda}^k)^\top \left[(A\tilde{x}^k + B\tilde{y}^k - b) - B(\tilde{y}^k - y^k) + \frac{1}{\beta}(\tilde{\lambda}^k - \lambda^k) \right] \geq 0, \forall \lambda \in \mathbb{R}^m. \quad (3.9)$$

Adding the left sides of the inequalities (3.7)-(3.9), we have

$$\begin{aligned} \tilde{\omega}^k \in \Omega, \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\tilde{\omega}^k) \\ - \begin{pmatrix} y - \tilde{y}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^\top \begin{pmatrix} \beta B^\top B + D_0 & -\alpha B^\top \\ -B & \frac{1}{\beta} I_m \end{pmatrix} \begin{pmatrix} y^k - \tilde{y}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix} \geq 0, \forall \omega \in \Omega, \end{aligned}$$

i.e.,

$$\begin{aligned} \tilde{\omega}^k \in \Omega, \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\tilde{\omega}^k) \\ \geq \begin{pmatrix} y - \tilde{y}^k \\ \lambda - \tilde{\lambda}^k \end{pmatrix}^\top \begin{pmatrix} \beta B^\top B + D_0 & -\alpha B^\top \\ -B & \frac{1}{\beta} I_m \end{pmatrix} \begin{pmatrix} y^k - \tilde{y}^k \\ \lambda^k - \tilde{\lambda}^k \end{pmatrix} \\ \geq (v - \tilde{v}^k)^\top Q(v^k - \tilde{v}^k), \forall \omega \in \Omega, \end{aligned}$$

as required. \square

Remark 3.1. It is noted that the matrix Q defined by (3.3) is not symmetric and strong positive (or, psd-plus matrix) because of the positive indefinite D_0 in particular, when $\tau \in (0, 1)$; the definition of strong positive/psd-plus matrix refer to the references [4, 18, 26]. So, the variational characterization (3.2) is different from the auxiliary principle in [4, 26].

We next present the connections between the essential variable v^k updated by the algorithm (1.11) and the auxiliary variable \tilde{v}^k .

Proposition 3.2. Let ω^{k+1} be generated by (1.11) with the given v^k , and let \tilde{v}^k be defined by (3.1). Then

$$v^{k+1} = v^k - M(v^k - \tilde{v}^k), \quad (3.10)$$

where

$$M = \begin{pmatrix} I_{n_2} & 0 \\ -\gamma\beta B & (\alpha + \gamma)I_m \end{pmatrix}. \quad (3.11)$$

Proof. The proof is similar to that of [13, Lemma 3.2]. So, it is omitted. \square

Remark 3.2. (1) According to Propositions 3.1 and 3.2, algorithm (1.11) can be regarded as a prediction-correction method, whose new iterate is generated by the correcting point \tilde{v}^k satisfying auxiliary variational inequality (3.2) and the prediction point v^k satisfying (3.10). From (1.11), we can see that only v^k is used to generate the next iteration and x^k plays an intermediate role in the iteration process. This is the reason that the convergence results of various ADMM were established in terms of only the variable v^k ; see, e.g., [1, 2, 8, 9, 10, 13, 14, 15] and the references therein. Also, the variables x and (y, λ) in the literature were called intermediate and essential variables, respectively. Here we do emphasize that the prediction-correction interpretation of the algorithm (1.11) is only assisted for the theoretical analysis and there is not necessary to adopt this prediction-correction process to implement the scheme (1.11) practically.

(2) It follows from (3.2) that if $(v - \tilde{v}^k)^\top Q(v^k - \tilde{v}^k) \geq 0$ for all $\omega \in \Omega$, then $\tilde{\omega}^k \in \Omega$ is a solution of (2.2). In view of this, we will focus on analyzing the properties of the term: $(v - \tilde{v}^k)^\top Q(v^k - \tilde{v}^k)$. Besides, by Proposition 3.2, we obtain that $M(v^k - \tilde{v}^k) = v^k - v^{k+1}$. Since M is nonsingular, we have

$$v^k - \tilde{v}^k = M^{-1}(v^k - v^{k+1}),$$

where M^{-1} is the inverse matrix of M . Hence, we have

$$(v - \tilde{v}^k)^\top Q(v^k - \tilde{v}^k) = (v - \tilde{v}^k)^\top QM^{-1}(v^k - v^{k+1}).$$

After calculation, one has

$$M^{-1} = \begin{pmatrix} I_{n_2} & 0 \\ \frac{\gamma\beta}{\alpha+\gamma}B & \frac{1}{\alpha+\gamma}I_m \end{pmatrix}$$

and

$$\begin{aligned} QM^{-1} &= \begin{pmatrix} \beta B^\top B + D_0 & -\alpha B^\top \\ -B & \frac{1}{\beta} I_m \end{pmatrix} \begin{pmatrix} I_{n_2} & 0 \\ \frac{\gamma\beta}{\alpha+\gamma} B & \frac{1}{\alpha+\gamma} I_m \end{pmatrix} \\ &= \begin{pmatrix} \left(1 - \frac{\alpha\gamma}{\alpha+\gamma}\right) \beta B^\top B + D_0 & -\frac{\alpha}{\alpha+\gamma} B^\top \\ -\frac{\alpha}{\alpha+\gamma} B & \frac{1}{(\alpha+\gamma)\beta} I_m \end{pmatrix}. \end{aligned}$$

For the sake of brevity, we set

$$H := QM^{-1} = \begin{pmatrix} \left(1 - \frac{\alpha\gamma}{\alpha+\gamma}\right) \beta B^\top B + D_0 & -\frac{\alpha}{\alpha+\gamma} B^\top \\ -\frac{\alpha}{\alpha+\gamma} B & \frac{1}{(\alpha+\gamma)\beta} I_m \end{pmatrix} \quad (3.12)$$

and

$$G := Q^\top + Q - M^\top H M,$$

where the matrices Q and M are defined respectively by (3.3) and (3.11). Then (3.2) is equivalent to the following:

$$\tilde{\omega}^k \in \Omega, \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\tilde{\omega}^k) \geq (v - \tilde{v}^k)^\top H(v^k - v^{k+1}), \forall \omega \in \Omega. \quad (3.13)$$

(3) Since $D_0 = D - (1 - \tau)\beta B^\top B$ (see (1.12)), one has

$$\begin{aligned} H &= \begin{pmatrix} \left(1 - \frac{\alpha\gamma}{\alpha+\gamma}\right) \beta B^\top B + D_0 & -\frac{\alpha}{\alpha+\gamma} B^\top \\ -\frac{\alpha}{\alpha+\gamma} B & \frac{1}{(\alpha+\gamma)\beta} I_m \end{pmatrix} \\ &= \begin{pmatrix} \left(\tau - \frac{\alpha\gamma}{\alpha+\gamma}\right) \beta B^\top B + D & -\frac{\alpha}{\alpha+\gamma} B^\top \\ -\frac{\alpha}{\alpha+\gamma} B & \frac{1}{(\alpha+\gamma)\beta} I_m \end{pmatrix}. \end{aligned}$$

It is easy to verify that H is a positive definite matrix for all $(\tau, \alpha, \gamma) \in \mathcal{P}$, where

$$\mathcal{P} := \{(\tau, \alpha, \gamma) \in \mathbb{R}_{++} \times \mathbb{R}_+ \times \mathbb{R}_+ \mid \tau > \alpha, \gamma \geq \alpha, \alpha + \gamma > 0\}.$$

By the definition of H , we have $HM = Q$ and so, $M^\top HM = M^\top Q$. Then

$$\begin{aligned} G &= Q^\top + Q - M^\top H M \\ &= \begin{pmatrix} 2(\beta B^\top B + D_0) & -(\alpha + 1)B^\top \\ -(\alpha + 1)B & \frac{2}{\beta} I_m \end{pmatrix} - \begin{pmatrix} (\gamma + 1)\beta B^\top B + D_0 & -(\alpha + \gamma)B^\top \\ -(\alpha + \gamma)B & \frac{\alpha + \gamma}{\beta} I_m \end{pmatrix} \\ &= \begin{pmatrix} (1 - \gamma)\beta B^\top B + D_0 & (\gamma - 1)B^\top \\ (\gamma - 1)B & \frac{2 - \alpha - \gamma}{\beta} I_m \end{pmatrix} \\ &= \begin{pmatrix} (\tau - \gamma)\beta B^\top B + D & (\gamma - 1)B^\top \\ (\gamma - 1)B & \frac{2 - \alpha - \gamma}{\beta} I_m \end{pmatrix}. \end{aligned} \quad (3.14)$$

Generally, the matrix G is not necessary positive definite for all $(\tau, \alpha, \gamma) \in \mathcal{P}$.

4. THE CONVERGENCE ANALYSIS

In this section, we focus on the convergence analysis of algorithm (1.11). In order to establish the convergence of algorithm (1.11), we give some useful preliminaries results.

Proposition 4.1. *Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and let $\tilde{\omega}^k$ be defined by (3.1). Then*

$$\begin{aligned} \tilde{\omega}^k \in \Omega, \quad & \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\omega) \\ & \geq \frac{1}{2}(\|v - v^{k+1}\|_H^2 - \|v - v^k\|_H^2) + \frac{1}{2}(v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k), \quad \forall \omega \in \Omega, \end{aligned} \quad (4.1)$$

where G is defined by (3.14).

Proof. The proof is similar to that of [14, Lemma 4.1] and so, it is omitted. \square

It is well-known that if the matrix G defined by (3.14) is positive definite, then inequality (4.1) essentially yields the convergence and its worst-case convergence rate; see, e.g., [2, 8, 13, 17, 22] and the references therein. Unfortunately, The G defined by (3.14) generally is not positive definite for $(\tau, \alpha, \gamma) \in \mathcal{P}$. So, inequality (4.1) can not be directly applied to obtain the convergence and convergence rate. This difficulty makes the convergence analysis for algorithm (1.11) more challenging than that for the existing linearized ADMM (1.7). For dealing with the trouble resulted from the positive-indefiniteness of the matrix G defined by (3.14), we intend to relax the crossing term $(v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k)$ such as

$$(v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k) \geq \psi(v^k, v^{k+1}) - \psi(v^{k-1}, v^k) + \phi(v^k, v^{k+1}), \quad (4.2)$$

where $\psi(\cdot, \cdot)$ and $\phi(\cdot, \cdot)$ are both nonnegative functions. The terms $\psi(v^k, v^{k+1})$ and $\psi(v^{k-1}, v^k)$ in the right side of (4.2) can be manipulated consecutively between iterates, and the term $\phi(v^k, v^{k+1})$ should be such an error bound that can measure how much ω^{k+1} fails to be a solution of the mixed variational inequality (2.2). If one gets such functions suiting (4.2), it follows from (4.1) that

$$\begin{aligned} & \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\omega) \\ & \geq \frac{1}{2}(\|v - v^{k+1}\|_H^2 + \psi(v^k, v^{k+1})) - \frac{1}{2}(\|v - v^k\|_H^2 + \psi(v^{k-1}, v^k)) + \frac{1}{2}\phi(v^k, v^{k+1}), \end{aligned} \quad (4.3)$$

for all $\omega \in \Omega$.

Let us first refine the crossing term $(v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k)$.

Proposition 4.2. *Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and let $\tilde{\omega}^k$ be defined by (3.1). Then*

$$\begin{aligned} (v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k) = & \|y^k - y^{k+1}\|_D^2 + \left(\tau - \gamma + \frac{\gamma^2 - 2\gamma}{\alpha + \gamma} + \frac{2\gamma^2}{(\alpha + \gamma)^2} \right) \beta \|B(y^k - y^{k+1})\|^2 \\ & + \frac{2 - \alpha - \gamma}{(\alpha + \gamma)^2} \beta \|\lambda^k - \lambda^{k+1}\|^2 + \frac{2(\gamma - \alpha)}{(\alpha + \gamma)^2} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}). \end{aligned} \quad (4.4)$$

Proof. By the definition of G and $\tilde{y}^k = y^{k+1}$, we have

$$\begin{aligned} (v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k) \stackrel{(3.14)}{=} & (\tau - \gamma) \beta \|B(y^k - y^{k+1})\|^2 + \frac{2 - \alpha - \gamma}{\beta} \|\lambda^k - \tilde{\lambda}^k\|^2 \\ & + \|y^k - y^{k+1}\|_D^2 + 2(\gamma - 1)(\lambda^k - \tilde{\lambda}^k)^\top B(y^k - \tilde{y}^k). \end{aligned} \quad (4.5)$$

Due to $\tilde{\lambda}^k = \lambda^k - \beta(Ax^{k+1} + By^k - b)$ and $\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k)$, we have

$$\begin{aligned}\lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b) \\ &= \lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k) - \gamma\beta \left[(Ax^{k+1} + By^k - b) + B(y^{k+1} - y^k) \right] \\ &= \lambda^k - \alpha(\lambda^k - \tilde{\lambda}^k) - \gamma(\lambda^k - \tilde{\lambda}^k) - \gamma\beta B(y^{k+1} - y^k) \\ &= \lambda^k - (\alpha + \gamma)(\lambda^k - \tilde{\lambda}^k) + \gamma\beta B(y^k - y^{k+1})\end{aligned}$$

and then

$$\lambda^k - \tilde{\lambda}^k = \frac{1}{\alpha + \gamma} \left[(\lambda^k - \lambda^{k+1}) + \gamma\beta B(y^k - y^{k+1}) \right].$$

Hence, we have

$$\begin{aligned}& \frac{2 - \alpha - \gamma}{\beta} \|\lambda^k - \tilde{\lambda}^k\|^2 \\ &= \frac{2 - \alpha - \gamma}{\beta(\alpha + \gamma)^2} \|(\lambda^k - \lambda^{k+1}) + \gamma\beta B(y^k - y^{k+1})\|^2 \\ &= \frac{2 - \alpha - \gamma}{\beta(\alpha + \gamma)^2} \left[\|\lambda^k - \lambda^{k+1}\|^2 + 2\gamma\beta(\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) + \gamma^2\beta^2 \|B(y^k - y^{k+1})\|^2 \right] \\ &= \frac{2 - \alpha - \gamma}{\beta(\alpha + \gamma)^2} \|\lambda^k - \lambda^{k+1}\|^2 + \frac{2\gamma(2 - \alpha - \gamma)}{(\alpha + \gamma)^2} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) \\ &\quad + \frac{\gamma^2\beta(2 - \alpha - \gamma)}{(\alpha + \gamma)^2} \|B(y^k - y^{k+1})\|^2\end{aligned}\tag{4.6}$$

and

$$\begin{aligned}& 2(\gamma - 1)(\lambda^k - \tilde{\lambda}^k)^\top B(y^k - y^{k+1}) \\ &= \frac{2(\gamma - 1)}{\alpha + \gamma} \left[(\lambda^k - \lambda^{k+1}) + \gamma\beta B(y^k - y^{k+1}) \right]^\top B(y^k - y^{k+1}) \\ &= \frac{2(\gamma - 1)}{\alpha + \gamma} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) + \frac{2\gamma\beta(\gamma - 1)}{\alpha + \gamma} \|B(y^k - y^{k+1})\|^2.\end{aligned}\tag{4.7}$$

Therefore, (4.4) is derived by substituting (4.6) and (4.7) into (4.5). \square

We next continue to estimate the crossing term $\frac{2(\gamma - \alpha)}{(\alpha + \gamma)^2} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1})$ in (4.4).

Proposition 4.3. *Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and let $\tilde{\omega}^k$ be defined by (3.1). Then*

$$\begin{aligned}& \frac{2(\gamma - \alpha)}{(\alpha + \gamma)^2} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) \\ &\geq \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \|y^k - y^{k+1}\|_D^2 - \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \|y^{k-1} - y^k\|_D^2 \\ &\quad + \left[\frac{(\gamma - \alpha)(\alpha\beta + 1)}{(\alpha + 1)(\alpha + \gamma)^2} (\gamma - 1) \operatorname{sgn}(1 - \gamma) + \frac{2\alpha\beta(\gamma - \alpha)(1 - \gamma)}{(\alpha + 1)(\alpha + \gamma)^2} - \frac{3\beta(\gamma - \alpha)(1 - \tau)}{(\alpha + 1)(\alpha + \gamma)} \right] \|B(y^k - y^{k+1})\|^2 \\ &\quad + \left[\frac{\alpha\beta(\gamma - \alpha)}{(\alpha + 1)(\alpha + \gamma)^2} (\gamma - 1) \operatorname{sgn}(1 - \gamma) - \frac{\beta(\gamma - \alpha)(1 - \tau)}{(\alpha + 1)(\alpha + \gamma)} \right] \|B(y^{k-1} - y^k)\|^2 \\ &\quad + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)^2} (\gamma - 1) \operatorname{sgn}(1 - \gamma) \|\lambda^k - \lambda^{k-1}\|^2.\end{aligned}$$

Proof. Substituting $\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b)$ into the inequality (3.5), we have

$$\begin{aligned} y^{k+1} \in \mathcal{Y}, \theta_2(y) - \theta_2(y^{k+1}) + (y - y^{k+1})^\top \left[-B^\top \lambda^{k+\frac{1}{2}} + \beta B^\top (Ax^{k+1} + By^{k+1} - b) \right. \\ \left. + D_0(y^{k+1} - y^k) \right] \geq 0, \forall y \in \mathcal{Y}. \end{aligned} \quad (4.8)$$

Taking $y = y^k$ and $y = y^{k+1}$ in (4.8) for the $(k+1)$ -th and k -th iterations, respectively, and adding them, it follows that

$$\begin{aligned} (y^k - y^{k+1})^\top \left[-B^\top (\lambda^{k+\frac{1}{2}} - \lambda^{k-\frac{1}{2}}) + \beta B^\top (Ax^{k+1} + By^{k+1} - b) \right. \\ \left. - \beta B^\top (Ax^k + By^k - b) + D_0((y^{k+1} - y^k) - (y^k - y^{k-1})) \right] \geq 0. \end{aligned} \quad (4.9)$$

Taking into account that

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b),$$

we obtain

$$\begin{aligned} \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b) \\ &= \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b) - \gamma\beta(Ax^{k+1} + By^{k+1} - b) \\ &= \lambda^k - (\alpha + \gamma)\beta(Ax^{k+1} + By^{k+1} - b) - \alpha\beta B(y^k - y^{k+1}). \end{aligned}$$

By re-arranging the terms of the above formula and dividing by $\alpha + \gamma > 0$, we have

$$\beta(Ax^{k+1} + By^{k+1} - b) = \frac{1}{\alpha + \gamma} (\lambda^k - \lambda^{k+1}) - \frac{\alpha\beta}{\alpha + \gamma} B(y^k - y^{k+1}). \quad (4.10)$$

Again, from $\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1} + By^k - b)$, we deduce that

$$\begin{aligned} \lambda^{k-\frac{1}{2}} - \lambda^{k+\frac{1}{2}} &= \left[\lambda^{k-1} - \alpha\beta(Ax^k + By^{k-1} - b) \right] - \left[\lambda^k - \alpha\beta(Ax^{k+1} + By^k - b) \right] \\ &= (\lambda^{k-1} - \lambda^k) - \alpha\beta(Ax^k + By^k - b) + \alpha\beta(Ax^{k+1} + By^{k+1} - b) \\ &\quad - \alpha\beta B(y^{k-1} - y^k) + \alpha\beta B(y^k - y^{k+1}) \\ &\stackrel{(4.10)}{=} (\lambda^{k-1} - \lambda^k) - \alpha \left[\frac{1}{\alpha + \gamma} (\lambda^{k-1} - \lambda^k) - \frac{\alpha\beta}{\alpha + \gamma} B(y^{k-1} - y^k) \right] \\ &\quad + \alpha \left[\frac{1}{\alpha + \gamma} (\lambda^k - \lambda^{k+1}) - \frac{\alpha\beta}{\alpha + \gamma} B(y^k - y^{k+1}) \right] \\ &\quad - \alpha\beta B(y^{k-1} - y^k) + \alpha\beta B(y^k - y^{k+1}) \\ &= (\lambda^{k-1} - \lambda^k) - \frac{\alpha}{\alpha + \gamma} (\lambda^{k-1} - \lambda^k) + \frac{\alpha}{\alpha + \gamma} (\lambda^k - \lambda^{k+1}) \\ &\quad + \left(\frac{\alpha}{\alpha + \gamma} - 1 \right) \alpha\beta B(y^{k-1} - y^k) - \left(\frac{\alpha}{\alpha + \gamma} - 1 \right) \alpha\beta B(y^k - y^{k+1}). \end{aligned} \quad (4.11)$$

Note that

$$\begin{aligned}
& (y^k - y^{k+1})^\top \left[-B^\top (\lambda^{k+\frac{1}{2}} - \lambda^{k-\frac{1}{2}}) + \beta B^\top (Ax^{k+1} + By^{k+1} - b) - \beta B^\top (Ax^k + By^k - b) \right. \\
& \quad \left. + D_0((y^{k+1} - y^k) - (y^k - y^{k-1})) \right] \\
& \stackrel{(4.10)}{=} (y^k - y^{k+1})^\top B^\top (\lambda^{k-\frac{1}{2}} - \lambda^{k+\frac{1}{2}}) + \frac{1}{\alpha + \gamma} (y^k - y^{k+1})^\top B^\top \left[(\lambda^k - \lambda^{k+1}) - \alpha \beta B(y^k - y^{k+1}) \right] \\
& \quad - \frac{1}{\alpha + \gamma} (y^k - y^{k+1})^\top B^\top \left[(\lambda^{k-1} - \lambda^k) - \alpha \beta B(y^{k-1} - y^k) \right] \\
& \quad + (y^k - y^{k+1})^\top D_0 \left[(y^{k+1} - y^k) - (y^k - y^{k-1}) \right] \\
& \stackrel{(4.11)}{=} (y^k - y^{k+1})^\top B^\top \left\{ (\lambda^{k-1} - \lambda^k) - \frac{\alpha}{\alpha + \gamma} (\lambda^{k-1} - \lambda^k) + \frac{\alpha}{\alpha + \gamma} (\lambda^k - \lambda^{k+1}) \right. \\
& \quad \left. + \left(\frac{\alpha}{\alpha + \gamma} - 1 \right) \alpha \beta B(y^{k-1} - y^k) - \left(\frac{\alpha}{\alpha + \gamma} - 1 \right) \alpha \beta B(y^k - y^{k+1}) \right\} \\
& \quad + (y^k - y^{k+1})^\top B^\top \left[\frac{1}{\alpha + \gamma} (\lambda^k - \lambda^{k+1}) - \frac{\alpha \beta}{\alpha + \gamma} B(y^k - y^{k+1}) \right] \\
& \quad - (y^k - y^{k+1})^\top B^\top \left[\frac{1}{\alpha + \gamma} (\lambda^{k-1} - \lambda^k) - \frac{\alpha \beta}{\alpha + \gamma} B(y^{k-1} - y^k) \right] \\
& \quad + (y^k - y^{k+1})^\top D_0 \left[(y^{k+1} - y^k) - (y^k - y^{k-1}) \right].
\end{aligned}$$

Consequently,

$$\begin{aligned}
& (y^k - y^{k+1})^\top \left[-B^\top (\lambda^{k+\frac{1}{2}} - \lambda^{k-\frac{1}{2}}) + \beta B^\top (Ax^{k+1} + By^{k+1} - b) - \beta B^\top (Ax^k + By^k - b) \right. \\
& \quad \left. + D_0((y^{k+1} - y^k) - (y^k - y^{k-1})) \right] \\
& = \frac{\alpha + 1}{\alpha + \gamma} (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) + \frac{\gamma - 1}{\alpha + \gamma} (\lambda^{k-1} - \lambda^k)^\top B(y^k - y^{k+1}) \\
& \quad + \frac{\alpha \beta (\gamma - 1)}{\alpha + \gamma} \|B(y^k - y^{k+1})\|^2 - \frac{\alpha \beta (\gamma - 1)}{\alpha + \gamma} (y^{k-1} - y^k)^\top B^\top B(y^k - y^{k+1}) \\
& \quad + (y^k - y^{k+1})^\top D_0 \left[(y^{k+1} - y^k) - (y^k - y^{k-1}) \right],
\end{aligned}$$

which together with (4.9) obtains

$$\begin{aligned}
& (\lambda^k - \lambda^{k+1})^\top B(y^k - y^{k+1}) \\
& \geq \frac{\alpha + \gamma}{\alpha + 1} (y^k - y^{k+1})^\top D_0 \left[(y^k - y^{k+1}) - (y^{k-1} - y^k) \right] + \frac{1 - \gamma}{\alpha + 1} (\lambda^{k-1} - \lambda^k)^\top B(y^k - y^{k+1}) \\
& \quad + \frac{\alpha \beta (1 - \gamma)}{\alpha + 1} \|B(y^k - y^{k+1})\|^2 + \frac{\alpha \beta (\gamma - 1)}{\alpha + 1} (y^{k-1} - y^k)^\top B^\top B(y^k - y^{k+1}). \tag{4.12}
\end{aligned}$$

Recall that

$$\begin{cases} \gamma - 1 = (\gamma - 1) \operatorname{sgn}(1 - \gamma) \leq 0, & \text{if } \gamma \in (0, 1], \\ \gamma - 1 > 0 > (\gamma - 1) \operatorname{sgn}(1 - \gamma), & \text{if } \gamma \in (1, 2). \end{cases}$$

Using the Cauchy-Schwartz inequality, we conclude

$$\begin{aligned} & \frac{\alpha\beta(\gamma-1)}{\alpha+1} (y^{k-1} - y^k)^\top B^\top B(y^k - y^{k+1}) \\ & \geq \frac{\alpha\beta}{2(\alpha+1)} (\gamma-1) \operatorname{sgn}(1-\gamma) \left(\|B(y^{k-1} - y^k)\|^2 + \|B(y^k - y^{k+1})\|^2 \right) \end{aligned} \quad (4.13)$$

and

$$\begin{aligned} & \frac{1-\gamma}{\alpha+1} (\lambda^{k-1} - \lambda^k)^\top B(y^k - y^{k+1}) \\ & \geq \frac{1}{2(\alpha+1)} (\gamma-1) \operatorname{sgn}(1-\gamma) \left(\|\lambda^k - \lambda^{k-1}\|^2 + \|B(y^k - y^{k+1})\|^2 \right). \end{aligned} \quad (4.14)$$

Since $D_0 = D - (1-\tau)\beta B^\top B$ (see (1.12)), by the Cauchy-Schwartz inequality for $\tau \in (0, 1]$, we have

$$\begin{aligned} & 2(y^k - y^{k+1})^\top D_0 \left[(y^k - y^{k+1}) - (y^{k-1} - y^k) \right] \\ & = 2(y^k - y^{k+1})^\top \left(D - (1-\tau)\beta B^\top B \right) \left[(y^k - y^{k+1}) - (y^{k-1} - y^k) \right] \\ & = 2\|y^k - y^{k+1}\|_D^2 - 2(y^k - y^{k+1})^\top D(y^{k-1} - y^k) - 2(1-\tau)\beta \|B(y^k - y^{k+1})\|^2 \\ & \quad + 2(1-\tau)\beta (y^k - y^{k+1})^\top (B^\top B)(y^{k-1} - y^k) \\ & \geq 2\|y^k - y^{k+1}\|_D^2 - \left(\|y^k - y^{k+1}\|_D^2 + \|y^{k-1} - y^k\|_D^2 \right) - 2(1-\tau)\beta \|B(y^k - y^{k+1})\|^2 \\ & \quad - (1-\tau)\beta \left(\|B(y^k - y^{k+1})\|^2 + \|B(y^{k-1} - y^k)\|^2 \right) \\ & = \|y^k - y^{k+1}\|_D^2 - \|y^{k-1} - y^k\|_D^2 - 3(1-\tau)\beta \|B(y^k - y^{k+1})\|^2 - (1-\tau)\beta \|B(y^{k-1} - y^k)\|^2. \end{aligned} \quad (4.15)$$

The conclusion therefore follows directly from (4.12), (4.13), (4.14) and (4.15). \square

For the sake of brevity, we set $\Gamma := (1-\gamma) \operatorname{sgn}(1-\gamma)$ in the rest of this paper. The following result follows directly from Propositions 4.2 and 4.3.

Proposition 4.4. *Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and $\tilde{\omega}^k$ be defined by (3.1). Then*

$$\begin{aligned} & (v^k - \tilde{v}^k)^\top G(v^k - \tilde{v}^k) \\ & \geq \frac{\gamma-\alpha}{(\alpha+1)(\alpha+\gamma)} \left[\|y^k - y^{k+1}\|_D^2 + \left(1-\tau + \frac{\alpha\Gamma}{\alpha+\gamma} \right) \beta \|B(y^{k+1} - y^k)\|^2 + \frac{\Gamma}{\alpha+\gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \\ & \quad - \frac{\gamma-\alpha}{(\alpha+1)(\alpha+\gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1-\tau + \frac{\alpha\Gamma}{\alpha+\gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 + \frac{\Gamma}{\alpha+\gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \\ & \quad + \left[\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^{k+1} - y^k)\|^2 + c_2 \|\lambda^{k+1} - \lambda^k\|^2 \right], \end{aligned} \quad (4.16)$$

where

$$c_1 = \beta \left(\tau - \gamma + \frac{\gamma^2 - 2\gamma}{\alpha + \gamma} + \frac{2\gamma^2}{(\alpha + \gamma)^2} - \frac{2(\gamma - \alpha)(1 - \tau)(\alpha + 2\gamma)}{(\alpha + 1)(\alpha + \gamma)^2} \right) - \frac{\alpha\beta + (1 + \alpha\beta)(\gamma - \alpha)}{(\alpha + 1)(\alpha + \gamma)^2} \Gamma \quad (4.17)$$

and

$$c_2 = \frac{1}{(\alpha + \gamma)^2} \left(\frac{2 - \alpha - \gamma}{\beta} + \frac{(\alpha - \gamma)\Gamma}{\alpha + 1} \right). \quad (4.18)$$

We next give the statements on how to construct the functions $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$ in (4.2) and show the relations among parameters α, γ, τ and penalty parameter β ensuring the nonnegativity of $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$.

Remark 4.1. (1) (Constructing functions $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$) In (4.16), if we set

$$\begin{aligned} \psi(v^k, v^{k+1}) := & \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^{k+1} - y^k)\|^2 \right] \\ & + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \end{aligned} \quad (4.19)$$

and

$$\varphi(v^k, v^{k+1}) := \|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^{k+1} - y^k)\|^2 + c_2 \|\lambda^{k+1} - \lambda^k\|^2, \quad (4.20)$$

then inequality (4.16) is in the form of (4.2).

- (2) (Parameters α, β, τ and γ for ensuring the nonnegativity of $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$) Note that the functions $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$ are nonnegative in (4.2). Clearly, if the coefficients in (4.16) are all nonnegative, then the functions $\psi(\cdot, \cdot)$ and $\varphi(\cdot, \cdot)$ defined respectively by (4.19) and (4.20) are nonnegative. Observe that $\Gamma = (1 - \gamma) \operatorname{sgn}(1 - \gamma) \geq 0$ and $1 - \tau \geq 0$ for all $\gamma > 0$ and $\tau \in (0, 1]$. Hence $\psi(\cdot, \cdot)$ defined by (4.19) is nonnegative whenever $\gamma \geq \alpha$. So, we only require that $c_1 \geq 0$ and $c_2 \geq 0$ under the condition $\gamma \geq \alpha$.

If $c_1 \geq 0$, then

$$\begin{aligned} & \tau \left(1 + \frac{2(\gamma - \alpha)(\alpha + 2\gamma)}{(\alpha + 1)(\alpha + \gamma)^2} \right) \\ & \geq \gamma + \frac{2(\gamma - \alpha)(\alpha + 2\gamma)}{(\alpha + 1)(\alpha + \gamma)^2} + \frac{\alpha\beta + (1 + \alpha\beta)(\gamma - \alpha)}{\beta(\alpha + 1)(\alpha + \gamma)^2} \Gamma - \frac{\gamma(\gamma + 1)^2 - \gamma(2\alpha + 1)}{(\alpha + \gamma)^2}. \end{aligned}$$

It follows that

$$\begin{aligned} \tau & \geq \frac{\gamma(\alpha + 1)(\alpha^2 + 2\alpha\gamma + 2\alpha - 2\gamma) + 2(\gamma - \alpha)(\alpha + 2\gamma)}{(\alpha + 1)(\alpha + \gamma)^2 + 2(\gamma - \alpha)(\alpha + 2\gamma)} \\ & \quad + \frac{[\alpha\beta + (1 + \alpha\beta)(\gamma - \alpha)]\Gamma}{[(\alpha + 1)(\alpha + \gamma)^2 + 2(\gamma - \alpha)(\alpha + 2\gamma)]\beta} \\ & \geq \frac{\{\gamma(\alpha + 1)(\alpha^2 + 2\alpha\gamma) + 2(\gamma - \alpha)[\alpha + \gamma(1 - \alpha)]\}\beta + [\alpha\beta + (1 + \alpha\beta)(\gamma - \alpha)]\Gamma}{[(\alpha + 1)(\alpha + \gamma)^2 + 2(\gamma - \alpha)(\alpha + 2\gamma)]\beta}. \end{aligned} \quad (4.21)$$

If $c_2 \geq 0$, then

$$\frac{2 - \alpha - \gamma}{\beta} \geq \frac{(\gamma - \alpha)\Gamma}{\alpha + 1} \geq 0. \quad (4.22)$$

So, $2 \geq \alpha + \gamma$. We next split into two cases:

Case I: If $2 = \alpha + \gamma$, then $\gamma = \alpha = 1, \beta > 0$.

Case II: If $2 > \alpha + \gamma$, then we have from (4.22) that

(a) $\gamma = \alpha, \beta > 0$ or, $\gamma = 1, \beta > 0$;

(b) $\gamma > \alpha, \gamma \neq 1$ and

$$\beta \leq \frac{(\alpha + 1)(2 - \alpha - \gamma)}{(\gamma - \alpha)\Gamma}.$$

In turn, since $\tau \leq 1$, it follows from (4.21) and simple calculation that

$$\{(\alpha + 1)[(\alpha^2 + 2\alpha\gamma)(1 - \gamma) + \gamma(3\gamma - 2\alpha) + \alpha\Gamma] - \alpha(\gamma + 2)\Gamma\} \beta \geq (\gamma - \alpha)\Gamma \geq 0. \quad (4.23)$$

Clearly, if $\gamma = 1, \beta > 0$, (4.23) holds. Besides, if $\gamma = \alpha < 1$, $3\alpha^3 - \alpha^2 - 5\alpha + 1 \leq 0$ and $\beta > 0$, then (4.23) holds. If $\gamma > \alpha, \gamma \neq 1$ and $2 > \alpha + \gamma$, then (4.23) holds whenever parameters (α, β, γ) satisfy the following inequalities:

$$(\alpha + 1)[(\alpha^2 + 2\alpha\gamma)(1 - \gamma) + \gamma(3\gamma - 2\alpha) + \alpha\Gamma] - \alpha(\gamma + 2)\Gamma > 0$$

and

$$\beta \geq \frac{(\gamma - \alpha)\Gamma}{(\alpha + 1)[(\alpha^2 + 2\alpha\gamma)(1 - \gamma) + \gamma(3\gamma - 2\alpha) + \alpha\Gamma] - \alpha(\gamma + 2)\Gamma} > 0.$$

Set

$$\mathcal{S} := \frac{\{\gamma(\alpha + 1)(\alpha^2 + 2\alpha\gamma) + 2(\gamma - \alpha)[\alpha + \gamma(1 - \alpha)]\} \beta + [\alpha\beta + (1 + \alpha\beta)(\gamma - \alpha)] \Gamma}{[(\alpha + 1)(\alpha + \gamma)^2 + 2(\gamma - \alpha)(\alpha + 2\gamma)] \beta}$$

and

$$\mathcal{L} := (\alpha + 1)[(\alpha^2 + 2\alpha\gamma)(1 - \gamma) + \gamma(3\gamma - 2\alpha) + \alpha\Gamma] - \alpha(\gamma + 2)\Gamma.$$

Altogether, if parameters $(\alpha, \beta, \gamma, \tau)$ belong to the following set:

$$\begin{aligned} \mathcal{S} := & \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \gamma = 1, \alpha \leq 1\} \\ & \cup \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha = \gamma < 1, 3\alpha^3 - \alpha^2 - 5\alpha + 1 \leq 0\} \\ & \cup \left\{ (\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \gamma \neq 1, \alpha + \gamma < 2, \mathcal{L} > 0, \right. \\ & \quad \left. \frac{(\gamma - \alpha)\Gamma}{\mathcal{L}} \leq \beta \leq \frac{(\alpha + 1)(2 - \alpha - \gamma)}{(\gamma - \alpha)\Gamma} \right\}, \end{aligned}$$

then $c_1 \geq 0$ and $c_2 \geq 0$. This together with Remark 3.2 3 yields that if parameters $(\alpha, \beta, \gamma, \tau) \in \mathcal{S}_1$, where

$$\begin{aligned} \mathcal{S}_1 := & \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \gamma = 1, \alpha \leq 1\} \\ & \cup \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \alpha = \gamma < 1, 3\alpha^3 - \alpha^2 - 5\alpha + 1 \leq 0\} \\ & \cup \left\{ (\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \alpha < \gamma \neq 1, \alpha + \gamma < 2, \mathcal{L} > 0, \right. \\ & \quad \left. \frac{(\gamma - \alpha)\Gamma}{\mathcal{L}} \leq \beta \leq \frac{(\alpha + 1)(2 - \alpha - \gamma)}{(\gamma - \alpha)\Gamma} \right\}, \end{aligned} \quad (4.24)$$

then H defined by (3.12) is positive definite matrix, $c_1 \geq 0$ and $c_2 \geq 0$.

Particularly, if parameters $(\alpha, \beta, \gamma, \tau) \in \mathcal{S}_2$, where

$$\begin{aligned} \mathcal{S}_2 := & \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \gamma = 1, \alpha < 1\} \\ & \cup \{(\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \alpha = \gamma < 1, 3\alpha^3 - \alpha^2 - 5\alpha + 1 \leq 0\} \\ & \cup \left\{ (\alpha, \beta, \gamma, \tau) \in \mathbb{R}_+ \times \mathbb{R}_{++}^3 \mid \mathcal{S} \leq \tau \leq 1, \alpha < \tau, \alpha < \gamma \neq 1, \alpha + \gamma < 2, \mathcal{L} > 0, \right. \\ & \quad \left. \frac{(\gamma - \alpha)\Gamma}{\mathcal{L}} \leq \beta < \frac{(\alpha + 1)(2 - \alpha - \gamma)}{(\gamma - \alpha)\Gamma} \right\}, \end{aligned} \quad (4.25)$$

then H defined by (3.12) is positive definite matrix, $c_1 \geq 0$ and $c_2 > 0$.

Theorem 4.1. Let $(\alpha, \beta, \gamma, \tau)$ satisfy (4.24). Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and let $\tilde{\omega}^k$ be defined by (3.1). Then

$$\begin{aligned}
& \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\omega) \\
& \geq \frac{1}{2} \left\{ \|v - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^{k+1} - y^k)\|^2 \right. \right. \\
& \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\} \\
& \quad - \frac{1}{2} \left\{ \|v - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\
& \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\} \\
& \quad + \frac{1}{2} \left[\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^k - y^{k+1})\|^2 + c_2 \|\lambda^k - \lambda^{k+1}\|^2 \right], \quad \forall \omega \in \Omega, \tag{4.26}
\end{aligned}$$

where c_1 is defined by (4.17) and c_2 is defined by (4.18).

Proof. Combining Propositions 4.1 and 4.4 with Remark 4.12, we can obtain the desired theorem. \square

Proposition 4.5. Let $(\alpha, \beta, \gamma, \tau)$ satisfy (4.24). Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11), and let $v^* \in \mathcal{V}^*$. Then

$$\begin{aligned}
& \left\{ \|v^* - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 \right. \right. \\
& \quad \left. \left. + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^{k+1} - y^k)\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\} \\
& \leq \left\{ \|v^* - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\
& \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\} - \left[\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^k - y^{k+1})\|^2 + c_2 \|\lambda^k - \lambda^{k+1}\|^2 \right],
\end{aligned}$$

where c_1 is defined by (4.17) and c_2 is defined by (4.18).

Proof. Letting $\omega := \omega^* = (x^*, v^*) \in \Omega^*$ in (4.26) and rearranging terms, we have

$$\begin{aligned}
& \left[\theta(\tilde{u}^k) - \theta(u^*) + (\tilde{\omega}^k - \omega^*)^\top F(\omega^*) \right] + \frac{1}{2} \left\{ \|v^* - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 \right. \right. \\
& \quad \left. \left. + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^{k+1} - y^k)\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\} \\
& \leq \frac{1}{2} \left\{ \|v^* - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\
& \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\} - \frac{1}{2} \left[\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^k - y^{k+1})\|^2 + c_2 \|\lambda^k - \lambda^{k+1}\|^2 \right]. \tag{4.27}
\end{aligned}$$

Since ω^* is a solution of variational inequality (2.2), we obtain

$$\theta(\tilde{u}^k) - \theta(u^*) + (\tilde{\omega}^k - \omega^*)^\top F(\omega^*) \geq 0.$$

This together with (4.27) yields the desired result. \square

We now present the main result on the convergence of the proposed method (1.11).

Theorem 4.2. *Let $(\alpha, \beta, \gamma, \tau)$ satisfy (4.25) and let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11). Then the sequence $(v^k)_{k \in \mathbb{N}}$ converges to an element $\hat{v} \in \mathcal{V}^*$.*

Proof. Let us first show that the sequence $(v^k)_{k \in \mathbb{N}}$ is bounded. Let $v^* \in \mathcal{V}^*$. By Proposition 4.5, we have

$$\begin{aligned} & \|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^k - y^{k+1})\|^2 + c_2 \|\lambda^k - \lambda^{k+1}\|^2 \\ & \leq \left\{ \|v^* - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\} - \left\{ \|v^* - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 \right. \right. \\ & \quad \left. \left. + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^{k+1} - y^k)\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\}. \end{aligned} \quad (4.28)$$

Thus,

$$\begin{aligned} & \|v^* - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^{k+1} - y^k)\|^2 \right. \\ & \quad \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \\ & \leq \|v^* - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^k - y^{k-1})\|^2 \right. \\ & \quad \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right]. \end{aligned} \quad (4.29)$$

Since $(\alpha, \beta, \gamma, \tau)$ satisfies (4.25), we have

$$\begin{aligned} & \|v^* - v^{k+n}\|_H^2 \\ & \leq \|v^* - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^k - y^{k-1})\|^2 \right. \\ & \quad \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \\ & \leq \|v^* - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^1 - y^0)\|^2 \right. \\ & \quad \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right], \quad \forall k, n \in \mathbb{N}, \end{aligned} \quad (4.30)$$

where $v^0 = (y^0, \lambda^0)$ and $v^1 = (y^1, \lambda^1)$ are the initial and the first generated iterates, respectively. Therefore, the sequence $(v^k)_{k \in \mathbb{N}}$ is bounded.

Next, we show that $v^k - v^{k+1} \rightarrow 0$ as $k \rightarrow +\infty$. Adding both sides of the inequality (4.28) over all k , we have

$$\begin{aligned} & \sum_{k=1}^{\infty} \left(\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^{k+1} - y^k)\|^2 + c_2 \|\lambda^{k+1} - \lambda^k\|^2 \right) \\ & \leq \left\{ \|v^* - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^1 - y^0)\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right] \right\}. \end{aligned}$$

Since D and H are positive definite, the above inequality implies that

$$\sum_{k=1}^{\infty} \left(\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^{k+1} - y^k)\|^2 + c_2 \|\lambda^{k+1} - \lambda^k\|^2 \right) < +\infty.$$

So,

$$\lim_{k \rightarrow +\infty} \left(\|y^k - y^{k+1}\|_D^2 + c_1 \|B(y^{k+1} - y^k)\|^2 + c_2 \|\lambda^{k+1} - \lambda^k\|^2 \right) = 0. \quad (4.31)$$

Using (4.25) yields that $c_1 \geq 0$ and $c_2 > 0$. Hence, $\|y^k - y^{k+1}\| \rightarrow 0$ and $\|\lambda^k - \lambda^{k+1}\| \rightarrow 0$ by (4.31), i.e.,

$$\lim_{k \rightarrow \infty} \|v^k - v^{k+1}\| = \lim_{k \rightarrow \infty} \left\| \begin{pmatrix} y^k - y^{k+1} \\ \lambda^k - \lambda^{k+1} \end{pmatrix} \right\| = 0. \quad (4.32)$$

Finally, we show that the sequence $(v^k)_{k \in \mathbb{N}}$ converges to some point \mathcal{V}^* . It follows from (4.32) and Proposition 3.2 that

$$\lim_{k \rightarrow \infty} \|M(v^k - \tilde{v}^k)\| = \lim_{k \rightarrow \infty} \|v^k - v^{k+1}\| = 0. \quad (4.33)$$

Since M is non-singular, (4.33) implies that $\|v^k - \tilde{v}^k\| \rightarrow 0$. According to the boundedness of $(v^k)_{k \in \mathbb{N}}$, $(\tilde{v}^k)_{k \in \mathbb{N}}$ is also bounded. Let \hat{v} be a cluster point of $(\tilde{v}^k)_{k \in \mathbb{N}}$ and let \hat{x} be the vector accompanied with $\hat{v} = (\hat{y}, \hat{\lambda}) \in \mathcal{V}$. Then there exists a subsequence $(\tilde{v}^{k_j})_{j \in \mathbb{N}}$ of $(\tilde{v}^k)_{k \in \mathbb{N}}$ such that $\tilde{v}^{k_j} \rightarrow \hat{v}$. Hence, using (4.26) yields that

$$\hat{\omega} = (\hat{x}, \hat{v}) \in \Omega, \quad \theta(u) - \theta(\hat{u}) + (\omega - \hat{\omega})^\top F(\hat{\omega}) \geq 0, \quad \forall \omega \in \Omega,$$

namely, $\hat{\omega}$ is a solution of variational inequality (2.2) and its essential part $\hat{v} \in \mathcal{V}^*$. Taking into account that $\|v^k - \tilde{v}^k\| \rightarrow 0$, there exists a subsequence $(v^{k_j})_{j \in \mathbb{N}}$ of $(v^k)_{k \in \mathbb{N}}$ corresponding to the subsequence $(\tilde{v}^{k_j})_{j \in \mathbb{N}}$ such that $v^{k_j} \rightarrow \hat{v}$. Setting $v^* := \hat{v} \in \mathcal{V}^*$ in (4.29) and (4.30), we deduce that the sequence

$$\left(\|\hat{v} - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right)_{k \in \mathbb{N}}$$

is convergent. Observe that

$$\begin{aligned} & \lim_{k \rightarrow +\infty} \left(\|\hat{v} - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^k - y^{k-1})\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right) \\ & \stackrel{(4.32)}{=} \lim_{k \rightarrow +\infty} \|\hat{v} - v^k\|_H^2. \end{aligned}$$

Therefore, the sequence $(\|\hat{v} - v^k\|_H^2)_{k \in \mathbb{N}}$ is convergent. Since H is positive definite, the sequence

$$(\|\hat{v} - v^k\|)_{k \in \mathbb{N}}$$

is convergent too. Accordingly, one has

$$\lim_{k \rightarrow +\infty} \|\hat{v} - v^k\| = \lim_{k \rightarrow +\infty} \|\hat{v} - v^{k_j}\| = 0,$$

i.e., the sequence $(v^k)_{k \in \mathbb{N}}$ converges to \hat{v} . □

We next give an example to show that the sequence the sequence $(v^k)_{k \in \mathbb{N}}$ generated by the proposed method (1.11) diverges when $\alpha + \gamma \geq 2$. Throughout this paper, all codes are written in Matlab and all experiments are performed in Matlab R2012b on a workstation with an Intel(R) Pentium(R) CPU 2117U (1.80 GHz) and 4GB RAM.

Example 4.1. Let $n_1 = n_2 = m := 1$, $b := 1$, $\mathcal{X} = \mathcal{Y} := \mathbb{R}_+$ and let $\theta_1(x) := 0$ and $\theta_2(y) := y$ for all $x, y \in \mathbb{R}$. Take $\beta := 2$, $\tau := \frac{1}{2}$, $A = B := (1)$ and $D := (\frac{1}{2})$. Then $D_0 = (-\frac{1}{2})$ is negative definite. The numerical simulation curves of sequences $(y^k)_{k \in \mathbb{N}}$, $(\lambda^k)_{k \in \mathbb{N}}$ and $(v^k)_{k \in \mathbb{N}} = (y^k, \lambda^k)_{k \in \mathbb{N}}$ generated by the proposed method (1.11) with initial point $(y^0, \lambda^0) = (1, 1)$ see Figures 1, 2 and 3, which also illustrate the divergence of $(v^k)_{k \in \mathbb{N}}$ and $(\lambda^k)_{k \in \mathbb{N}}$ with different α, γ and $\alpha + \gamma \geq 2$. However, if α and γ with the given $\beta = 2$ and $\tau = \frac{1}{2}$ satisfy (4.25), then the sequence $(\omega^k)_{k \in \mathbb{N}}$ generated by (1.11) indeed is convergent to some element of \mathcal{V}^* . The numerical computational results are reported in Table 1 with the terminate condition $\|v^{k+1} - v^k\| < 10^{-6}$. From Table 1, we can see that $(v^k)_{k \in \mathbb{N}} = (y^k, \lambda^k)_{k \in \mathbb{N}}$ and $(\omega^k)_{k \in \mathbb{N}} = (x^k, y^k, \lambda^k)_{k \in \mathbb{N}}$ are convergent to $(1, 0)$ and $(1, 0, 0)$, respectively. Actually, it is easy to verify that $\mathcal{V}^* = \{(1, 0)\}$ and $(x, y) = (1, 0)$ is the unique solution of the problem (1.1).

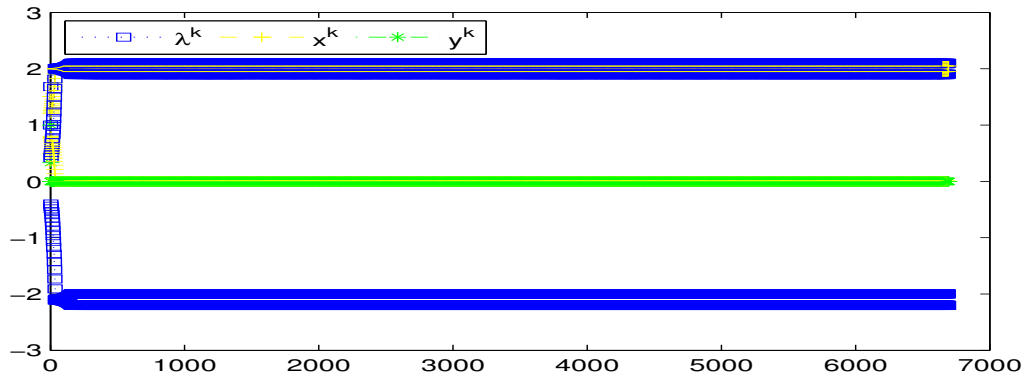


FIGURE 1. The simulation curves of $(x^k)_{k \in \mathbb{N}}$, $(y^k)_{k \in \mathbb{N}}$ and $(\lambda^k)_{k \in \mathbb{N}}$ with $\alpha = 0, \gamma = 2.05$.

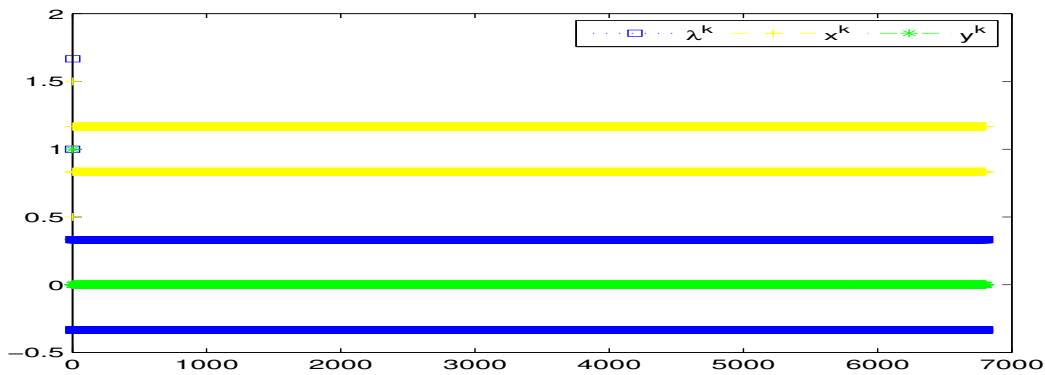


FIGURE 2. The simulation curves of $(x^k)_{k \in \mathbb{N}}$, $(y^k)_{k \in \mathbb{N}}$ and $(\lambda^k)_{k \in \mathbb{N}}$ with $\alpha = 0, \gamma = 2$.

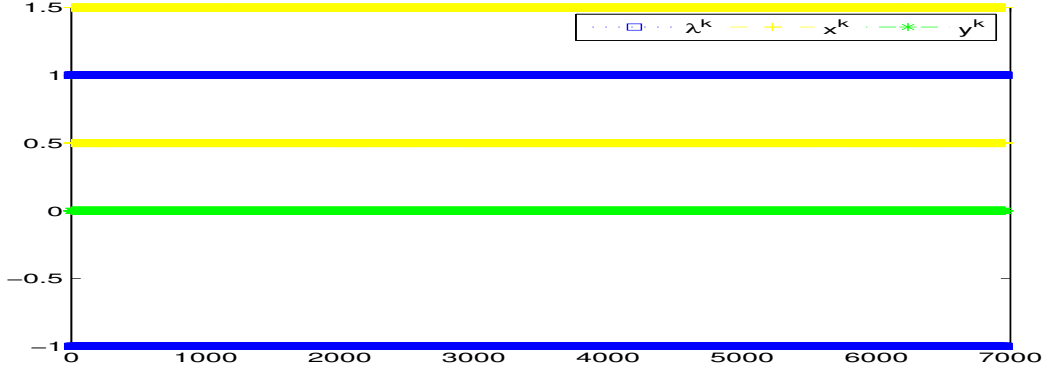


FIGURE 3. The simulation curves of $(x^k)_{k \in \mathbb{N}}$, $(y^k)_{k \in \mathbb{N}}$ and $(\lambda^k)_{k \in \mathbb{N}}$ with $\alpha = 1, \gamma = 1$.

Table 1. Numerical results for Example 4.1 with $\beta = 2$ and $\tau = \frac{1}{2}$

(α, γ)	Initial point (y^0, λ^0)	Val. x^k	Val. y^k	Val. λ^k	Iter.	CPU(s)
$(1/3, 1)$	$(1, 1)$	1.0000	0.0000	$1.6261e^{-07}$	15	4.8744
$(3/8, 1)$	$(10, 1)$	1.0000	0.0000	$-1.5293e^{-07}$	18	4.3455
$(2/5, 1)$	$(10, 10)$	1.0000	0.0000	$-1.3194e^{-07}$	20	4.9843
$(1/3, 1/3)$	$(100, 100)$	1.0000	0.0000	$2.6844e^{-07}$	18	4.2794
$(3/8, 3/8)$	$(100, 100)$	1.0000	0.0000	$9.8720e^{-07}$	15	3.3756
$(2/5, 2/5)$	$(100, 100)$	1.0000	0.0000	$8.8474e^{-08}$	13	2.9403
$(2/5, 1.2)$	$(1, 1)$	1.0000	0.0000	$3.0361e^{-07}$	31	7.5740
$(2/5, 1.1)$	$(1, 1)$	1.0000	0.0000	$3.0041e^{-07}$	23	6.1340
$(2/5, 0.8)$	$(1, 1)$	1.0000	0.0000	$7.7824e^{-08}$	11	3.0367

5. COMPLEXITY

In this section, we concentrate on the iteration complexity of the objective data for the proposed method (1.11) in the ergodic sense.

Theorem 5.1. *Let $(\alpha, \beta, \gamma, \tau)$ satisfy (4.24). Let $(\omega^k)_{k \in \mathbb{N}}$ be the sequence generated by (1.11) and let $\tilde{\omega}^k$ be defined by (3.1). Then after t iterations of method (1.11), $\tilde{\omega}_t$ is an approximation solution of (2.2) with the accuracy $\mathcal{O}(1/t)$, i.e., $\tilde{\omega}_t \in \Omega$ and*

$$\theta(u) - \theta(\tilde{u}_t) + (\omega - \tilde{\omega}_t)^\top F(\omega) \geq -\frac{d}{2t}, \quad \forall \omega \in \mathcal{B}_{\tilde{\omega}_t},$$

where $\mathcal{B}_{\tilde{\omega}_t}$ defined by (2.4),

$$\tilde{\omega}_t := \frac{1}{t} \sum_{k=1}^t \tilde{\omega}^k$$

and

$$d := \sup_{\omega \in \mathcal{B}_{\tilde{\omega}_t}} \left\{ \|v - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^1 - y^0)\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right] \right\}. \quad (5.1)$$

Proof. Using (2.4) yields that $\mathcal{B}_{\tilde{\omega}_k}$ is a nonempty compact and convex subset of Ω . Therefore, d defined by (5.1) is well-defined and finite-valued. It follows from Theorem 4.1 that

$$\begin{aligned} & \theta(u) - \theta(\tilde{u}^k) + (\omega - \tilde{\omega}^k)^\top F(\omega) \\ & \geq \frac{1}{2} \left\{ \|v - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^{k+1} - y^k)\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\} \\ & \quad - \frac{1}{2} \left\{ \|v - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\}, \forall \omega \in \Omega. \end{aligned}$$

By rearranging the terms in the above inequality, it shows that

$$\begin{aligned} & \theta(\tilde{u}^k) - \theta(u) + (\tilde{\omega}^k - \omega)^\top F(\omega) + \frac{1}{2} \left\{ \|v - v^{k+1}\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^k - y^{k+1}\|_D^2 \right. \right. \\ & \quad \left. \left. + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^{k+1} - y^k)\|^2 + \frac{\Gamma}{\alpha + \gamma} \|\lambda^{k+1} - \lambda^k\|^2 \right] \right\} \\ & \leq \frac{1}{2} \left\{ \|v - v^k\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^{k-1} - y^k\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^k - y^{k-1})\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^k - \lambda^{k-1}\|^2 \right] \right\}. \end{aligned} \quad (5.2)$$

Adding both sides of the inequality (5.2) over all k , we have

$$\begin{aligned} & \sum_{k=1}^t \theta(\tilde{u}^k) - t\theta(u) + \left(\sum_{k=1}^t \tilde{\omega}^k - t\omega \right)^\top F(\omega) \\ & \leq \frac{1}{2} \left\{ \|v - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^1 - y^0)\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right] \right\}. \end{aligned}$$

Hence,

$$\begin{aligned} & \frac{1}{t} \left(\sum_{k=1}^t \theta(\tilde{u}^k) \right) - \theta(u) + (\tilde{\omega}_t - \omega)^\top F(\omega) \\ & \leq \frac{1}{2t} \left\{ \|v - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma}\right) \beta \|B(y^1 - y^0)\|^2 \right. \right. \\ & \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right] \right\}. \end{aligned} \quad (5.3)$$

Since $\theta(u)$ is convex and

$$\tilde{u}_t = \frac{1}{t} \sum_{k=1}^t \tilde{u}^k,$$

we obtain

$$\theta(\tilde{u}_t) = \theta\left(\frac{1}{t} \sum_{k=1}^t \tilde{u}^k\right) \leq \frac{1}{t} \sum_{k=1}^t \theta(\tilde{u}^k).$$

This and (5.3) yield

$$\begin{aligned}
& \theta(\tilde{u}_t) - \theta(u) + (\tilde{\omega}_t - \omega)^\top F(\omega) \\
& \leq \frac{1}{2t} \left\{ \|v - v^1\|_H^2 + \frac{\gamma - \alpha}{(\alpha + 1)(\alpha + \gamma)} \left[\|y^0 - y^1\|_D^2 + \left(1 - \tau + \frac{\alpha\Gamma}{\alpha + \gamma} \right) \beta \|B(y^1 - y^0)\|^2 \right. \right. \\
& \quad \left. \left. + \frac{\Gamma}{\alpha + \gamma} \|\lambda^1 - \lambda^0\|^2 \right] \right\} \\
& \leq \frac{d}{2t}, \forall \omega \in \mathcal{B}_{\tilde{\omega}}.
\end{aligned}$$

This completes the proof. \square

Acknowledgments

This work was partially supported by the National Science Foundation of China (61673006, 11771058, 11871383), the Young Project of Hubei Provincial Department of Education (Q20141304), and the Basic and Advanced Research Project of Chongqing (cstc2016jcyjA0239).

REFERENCES

- [1] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends Mach. Learn.* 3 (2011), 1-122.
- [2] R. H. Chan, J. F. Yang, X. M. Yuan, Alternating direction method for image inpainting in wavelet domains, *SIAM J. Imag. Sci.* 4 (2011), 807-826.
- [3] T. H. Chang, W. C. Liao, M. Y. Hong, X. F. Wang, Asynchronous distributed ADMM for large-scale optimization-Part II: Linear convergence analysis and numerical performance, *IEEE Trans. Signal Process* 64 (2016), 3131-3144.
- [4] X. P. Ding, Auxiliary principle and algorithm for mixed equilibrium problems and bilevel mixed equilibrium problems in Banach spaces, *J. Optim. Theory Appl.* 146 (2010), 347-357.
- [5] M. Fortin, R. Glowinski, *Augmented Lagrangian Methods: Applications to The Numerical Solution of Boundary-Value Problems*, Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co, 1983.
- [6] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Comput. Math. Appl.* 2 (1976), 17-40.
- [7] R. Glowinski, A. Marrocco, Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires, *Revue française d'automatique, Informatique Recherche Opérationnelle, Analyse Numérique*, 9 (1975), 41-76.
- [8] D. Han, X. M. Yuan, A note on the alternating direction method of multipliers, *J. Optim. Theory Appl.* 155 (2012), 227-238.
- [9] D. Han, X. M. Yuan, W. Zhang, X. Cai, An ADM-based splitting method for separable convex programming, *Comput. Optim. Appl.* 54 (2013), 343-369.
- [10] B. S. He, L. Z. Liao, D. Han, H. Yang, A new inexact alternating directions method for monotone variational inequalities, *Math. Program.* 92 (2002), 103-118.
- [11] B. S. He, H. Liu, Z. R. Wang, X. M. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, *SIAM J. Optim.* 24 (2014), 1011-1040.
- [12] B. S. He, X. M. Yuan, On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method, *SIAM J. Numer. Anal.* 50 (2012), 700-709.
- [13] B. S. He, F. Ma, X. M. Yuan, Convergence study on the symmetric version of ADMM with larger step sizes, *SIAM J. Imag. Sci.* 9 (2016), 1467-1501.
- [14] B. S. He, F. Ma, X. M. Yuan, Linearized alternating direction method of multipliers via positive-indefinite proximal regularization for convex programming, *Optimization-online*, July, 2016, http://www.optimization-online.org/DB_HTML/2016/07/5569.html.

- [15] B. S. He, F. Ma, X. M. Yuan, Positive-indefinite proximal augmented Lagrangian method and its application to full Jacobian splitting for multi-block separable convex minimization problems, *Optimization-online*, September, 2016, http://www.optimization-online.org/DB_HTML/2016/09/5631.html.
- [16] M. Hestenes, Multiplier and gradient methods, *J. Optim. Theory Appl.* 4 (1969), 303-320.
- [17] Z. Lin, R. Liu, H. Li, Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning, *Mach. Learn.* 99 (2015), 287-325.
- [18] Z. Luo, P. Tseng, A decomposition property for a class of square matrices, *Appl. Math. Lett.* 4 (1991), 67-69.
- [19] M. K Ng, P. Weiss, X. M. Yuan, Solving constrained total-variation image restoration and reconstruction problems via alternating direction methods, *SIAM J. Sci. Comput.* 32 (2010), 2710-2736.
- [20] M. J. D. Powell, A method for nonlinear constraints in minimization problems, in *Optimization*, R. Fletcher, ed., pp. 283-298, Academic Press, 1969.
- [21] R. T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.* 1 (1976), 97-116.
- [22] Y. Wang, J. F. Yang, W. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, *SIAM J. Imag. Sci.* 1 (2008), 248-272.
- [23] J. F. Yang, X. M. Yuan, Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization, *Math. Comput.* 82 (2013), 301-329.
- [24] J. F. Yang, Y. Zhang and W. Yin, A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data, *IEEE J. Sel. Topics Signal Process.* 4 (2010), 288-297.
- [25] X. Q. Zhang, M. Burger, S. Osher, A unified primal-dual algorithm framework based on Bergman iteration, *J. Sci. Comput.* 46 (2010), 20-46.
- [26] D. L. Zhu, P. Marcotte, Co-coercivity and its role in the convergence of iterative schemes for solving variational inequalities, *SIAM J. Optim.* 6 (1996), 714-726.
- [27] L. Liu, X. Qin, Strong convergence of an extragradient-like algorithm involving pseudo-monotone mappings, *Numer. Algor.* (2019). doi: 10.1007/s11075-019-00737-3
- [28] L. Zhang, H. Zhao, Y. Lv, A modified inertial projection and contraction algorithms for quasi-variational inequalities, *Appl. Set-Val. Anal. Optim.* 1 (2019), 63-76.