J. Nonlinear Funct. Anal. 2020 (2020), Article ID 50 https://doi.org/10.23952/jnfa.2020.50

Journal of Nonlinear Functional Analysis
Available online at http://jnfa.mathres.org

AN INERTIAL PROXIMAL PEACEMAN-RACHFORD SPLITTING METHOD WITH SQP REGULARIZATION FOR CONVEX PROGRAMMING

A. BNOUHACHEM[1], X. QIN[2,*]

[1]Equipe MAISI, Ibn Zohr University, ENSA, BP 1136, Agadir, Morocco
[2]General Education Center, National Yunlin University of Science and Technology, Douliou 64002, Taiwan

Abstract. In this paper, based on the square quadratic proximal (SQP) method and the inertial proximal Peaceman-Rachford splitting method (PRSM), we propose an inertial PRSM with the SQP regularization for solving a separable convex minimization model with positive orthant constraints. The new algorithm can be viewed as an interior version of the inertial PRSM with the SQP regularization. Under standard assumptions, the global convergence of the proposed method is proved. We show that the proposed method can find an approximate solution of the mixed variational inequalities with an accuracy of $o(1/\sqrt{k})$.
Keywords. Variational inequalities; Square-quadratic proximal method; Peaceman-Rachford splitting method; Alternating direction method.

1. Introduction

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

$$\min\left\{\theta_1(x)+\theta_2(y) | Ax+By=b, x\in\mathscr{X}, y\in\mathscr{Y}\right\}, \tag{1.1}$$

where $\mathscr{X}\subset\mathbb{R}^{n_1}$ and $\mathscr{Y}\subset\mathbb{R}^{n_2}$ are closed convex sets, $\theta_1:\mathscr{X}\to\mathbb{R}$ and $\theta_2:\mathscr{Y}\to\mathbb{R}$ 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 multiplier method (ADMM) is one of the most successful and influential approaches for solving (1.1) and has been studied extensively both in theory and in practice. One now has a variety of techniques and methods to investigate various iterative ADMMs in the literature. Some of these variants include proximal terms in the subproblems of the ADMM in order to make them easier to solve. Since the choice of the step size selection strategies is important for the algorithms efficiency. Others added a step size parameter in the Lagrangian multiplier updating to improve the performance of the method; see, for example,

*Corresponding author.
E-mail addresses: babedallah@yahoo.com (A. Bnouhachem), qinxl@yuntech.edu.tw (X. Qin).
Received October 12, 2020; Accepted November 26, 2020.

[1, 2, 3, 4, 6, 7]. Various inexact, relaxed and accelerated variants of the ADMM with different error conditions were also very well studied in the literature, see, e.g., [8, 9, 10, 11].

The augmented Lagrangian function of (1.1) is defined by

$$L_\beta(x,y,\lambda) = \theta_1(x) + \theta_2(y) - \lambda^\top(Ax+By-b) + \frac{\beta}{2}\|Ax+By-b\|^2,$$

where $\lambda \in \mathbb{R}^l$ is the Lagrange multiplier and $\beta > 0$ is a penalty parameter. Some operator splitting methods [12, 13, 14] have been developed for solving the dual problem of (1.1). The Peaceman-Rachford operator splitting method (PRSM) [15] is also a splitting method. Different from the ADMM, the PRSM updates the Lagrange multiplier twice at each iteration. Based on PRSM, He *et al.* [5] proposed a strictly contractive PRSM by introducing a parameter $\alpha \in (0,1)$ to the update scheme of the dual variable, yielding the following procedure:

$$\begin{cases} x^{k+1} = \arg\min\left\{L_\beta(x,y^k,\lambda^k)|x\in\mathscr{X}\right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1}+By^k-b), \\ y^{k+1} = \arg\min\left\{L_\beta(x^{k+1},y,\lambda^{k+\frac{1}{2}})|y\in\mathscr{Y}\right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(Ax^{k+1}+By^{k+1}-b). \end{cases} \tag{1.2}$$

Inspired by [5], Gu, Jiang and Han [16] proposed a modification of the Peaceman-Rachford splitting method by introducing two different parameters in updating the dual variable, and by introducing semi-proximal terms to the subproblems in updating the primal variables. From a given $w^k = (x^k,y^k,\lambda^k) \in \mathscr{W}$, the new iterate $w^{k+1} = (x^{k+1},y^{k+1},\lambda^{k+1})$ is obtained via solving the following:

$$\begin{cases} x^{k+1} = \arg\min\left\{L_\beta(x,y^k,\lambda^k) + \frac{1}{2}\|x-x^k\|_L^2|x\in\mathscr{X}\right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1}+By^k-b), \\ y^{k+1} = \arg\min\left\{L_\beta(x^{k+1},y,\lambda^{k+\frac{1}{2}}) + \frac{1}{2}\|y-y^k\|_T^2|y\in\mathscr{Y}\right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1}+By^{k+1}-b), \end{cases} \tag{1.3}$$

where L and T are two positive semi-definite matrices. The global convergence is proved under $\alpha \in (0,1)$ and $\gamma \in \left(0, \frac{1-\alpha+\sqrt{(1-\alpha)^2+4(1-\alpha^2)}}{2}\right)$.

Recently, Li and Yuan [17] proposed a new method by combining the strictly contractive PRSM and the LQP regularization. More specifically, from a given $w^k = (x^k,y^k,\lambda^k) \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++} \times \mathbb{R}^l$, the new iterate $w^{k+1} = (x^{k+1},y^{k+1},\lambda^{k+1})$ is obtained by solving the following:

$$\begin{cases} x^{k+1} = \arg\min\left\{L_\beta(x,y^k,\lambda^k) + rd(x,x^k)|x\in\mathbb{R}^{n_1}_{++}\right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(Ax^{k+1}+By^k-b), \\ y^{k+1} = \arg\min\left\{L_\beta(x^{k+1},y,\lambda^{k+\frac{1}{2}}) + sd(x,x^k)|y\in\mathbb{R}^{n_2}_{++}\right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1}+By^{k+1}-b), \end{cases} \tag{1.4}$$

where r,s are positive scalars, $\alpha,\gamma \in (0,1)$ are two relaxation factors,

$$d(u,v) = \begin{cases} \sum_{j=1}^{N}\left[\frac{1}{2}(u_j-v_j)^2 + \mu(v_j^2 log\frac{v_j}{u_j} + u_jv_j - v_j^2)\right], & \text{if} \quad u\in\mathbb{R}^N_{++} \\ +\infty, & \text{otherwise,} \end{cases}$$

and $\mu \in (0,1)$.

Based on the work of Chen *et al.* [18], Dou, Li and Liu [19] proposed an inertial proximal Peaceman-Rachford splitting method. The iterative scheme reads as

$$\begin{cases} (\bar{x}^k,\bar{y}^k,\bar{\lambda}^k) = (x^k,y^k,\lambda^k)+\alpha_k(x^k-x^{k-1},y^k-y^{k-1},\lambda^k-\lambda^{k-1}), \\ x^{k+1} = \arg\min\left\{L_\beta(x,\bar{y}^k,\bar{\lambda}^k)+\frac{1}{2}\|x-\bar{x}^k\|_L^2 | x\in\mathscr{X}\right\}, \\ \lambda^{k+\frac{1}{2}} = \bar{\lambda}^k-\alpha\beta(Ax^{k+1}+B\bar{y}^k-b), \\ y^{k+1} = \arg\min\left\{L_\beta(x^{k+1},y,\lambda^{k+\frac{1}{2}})+\frac{1}{2}\|y-\bar{y}^k\|_T^2 | y\in\mathscr{Y}\right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}}-\alpha\beta(Ax^{k+1}+By^{k+1}-b), \end{cases} \tag{1.5}$$

where $\alpha \in (0,1)$ and the sequence $\{\alpha_k\}$ is chosen such that, for all $k \geq 0$, $0 < \alpha_k < t$, where t is a real number in $(0,1)$.

In this paper, we focus on the case of (1.1) where the constraint sets $\mathscr{X} = \mathbb{R}^{n_1}_+, \mathscr{Y} = \mathbb{R}^{n_2}_+$; i.e.,

$$\min\left\{\theta_1(x)+\theta_2(y) | Ax+By=b, x\in\mathbb{R}^{n_1}_+, y\in\mathbb{R}^{n_2}_+\right\}. \tag{1.6}$$

The main contribution of this paper is to show that the recently proposed inertial PRSM [18] can also be integrated with the SQP regularization for solving problem (1.6). The new algorithm can be viewed as an interior version of the inertial PRSM with the SQP regularization. We establish the global convergence of the proposed algorithm under suitable assumptions. Our results can be viewed as significant extensions of associated previously known results.

2. Preliminaries

We state some preliminaries that are useful in later analysis.
Let $\mathbb{R}^n_+ = \{x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n. : x_i \geq 0, \forall i=1,2,\ldots,n\}$ and $\mathbb{R}^n_{++} = \{x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n : x_i > 0, \forall i=1,2,\ldots,n\}$. For any vector $u\in\mathbb{R}^n$, $\|u\|^2 = u^\top u$. Let $D\in\mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. We denote the D-norm of u by $\|u\|_D^2 = u^\top Du$.

Lemma 2.1. *Assume that $\{a_k\}$ is a sequence of nonnegative real numbers such that*

(1) a_k is monotonically non-increasing,

(2) $\sum_{k=1}^\infty a_k < \infty$.

Then, $a_k = o(1/k)$.

Proof. Since a_k is monotonically non-increasing, one has

$$ka_{2k} \leq a_{k+1}+\cdots+a_{2k} \longrightarrow 0 \quad \text{as} \quad k\longrightarrow\infty.$$

Then, $a_k = o(1/k)$. □

2.1. The square quadratic proximal regularization.

Auslender, Teboulle and Ben-Tiba [20] proposed a new type of proximal interior algorithms by replacing the quadratic function $\frac{1}{2}\|x-x^k\|^2$ by $d_\phi(x,x^k)$, which could be defined as

$$d_\phi(x,y) := \sum_{j=1}^n y_j^2\phi(x_j/y_j).$$

We consider the function defined by

$$\phi(t) = \begin{cases} \frac{1}{4}(t-1)^2+\mu(\sqrt{t}-1)^2, & \text{if} \quad t>0, \\ +\infty, & \text{otherwise.} \end{cases}$$

Then, $d_\phi(x,y)$ can be written as

$$d_\phi(x,y) = \begin{cases} \frac{1}{4}\|x-y\|^2 + \mu \sum_{j=1}^{n} \left(y_j x_j - 2(y_j)^2 \sqrt{\frac{x_j}{y_j}} + (y_j)^2 \right), & \text{if} \quad x \in \mathbb{R}^n_{++}, \\ +\infty, & \text{otherwise.} \end{cases}$$

It is easy to see that

$$\begin{aligned} \nabla_x d_\phi(x,y) &= \frac{1}{2}(x-y) + \mu \sum_{j=1}^{n} \left(y_j - \frac{(y_j)^2}{\sqrt{y_j}} \frac{1}{\sqrt{x_j}} \right) \\ &= \frac{1}{2}(x-y) + \mu (y - Y(\sqrt{x})^{-1}), \end{aligned} \tag{2.1}$$

where $Y = \text{diag}(\sqrt{y_1}^3, \ldots, \sqrt{y_n}^3)$ and $\sqrt{x} = (\sqrt{x_1}, \ldots, \sqrt{x_n})$.

The following lemma provides useful facts, which facilitate the analysis of the convergence of the proposed method.

Lemma 2.2. *Let $P := diag(p_1, p_2, \ldots, p_n) \in \mathbb{R}^{n\times n}, q(z) \in \mathbb{R}^n$ be a monotone mapping of z with respect to $\mathbb{R}^n_+$ and $f : \mathbb{R}^n \longrightarrow \mathbb{R}$. Let $\mu \in (0,1)$ be a constant. For given $z, \tilde{z} \in \mathbb{R}^n_{++}$, we define $Z = diag(\sqrt{\tilde{z}_1}^3, \ldots, \sqrt{\tilde{z}_n}^3)$ and $\sqrt{z} = (\sqrt{z_1}, \ldots, \sqrt{z_n})$,*

$$\Psi(\tilde{z}, z) := \nabla_z d_\phi(z, \tilde{z}) = \frac{1}{2}(z - \tilde{z}) + \mu(\tilde{z} - Z(\sqrt{z})^{-1}). \tag{2.2}$$

Thus, the variational inequality

$$f(z') - f(z) + (z'-z)^\top [q(z) + P\Psi(\tilde{z}, z)] \geq 0, \quad \forall z' \in \mathbb{R}^n_+ \tag{2.3}$$

has the unique positive solution z. In addition, for this positive solution $z \in \mathbb{R}^n_{++}$ and $z' \in \mathbb{R}^n_+$, one has

$$f(z) - f(z') + (z-z')^\top [q(z) + \frac{(1+\mu)}{2} P(z-\tilde{z})] \leq \frac{\mu}{2}\|z-\tilde{z}\|^2_P. \tag{2.4}$$

Proof. For each $t > 0$, one has $\frac{1}{2}\left(1-\frac{1}{t}\right) \leq 1 - \frac{1}{\sqrt{t}} \leq \frac{1}{2}(t-1)$. We obtain after multiplication by $z'_j \tilde{z}_j \geq 0$, for each $j = 1, \ldots, n$,

$$z'_j \tilde{z}_j \left(1 - \frac{\sqrt{\tilde{z}_j}}{\sqrt{z_j}}\right) \leq z'_j \tilde{z}_j \frac{1}{2}\left(\frac{z_j}{\tilde{z}_j} - 1\right) = \frac{1}{2} z'_j (z_j - \tilde{z}_j)$$

and, after multiplication by $z_j \tilde{z}_j \geq 0$, for each $j = 1, \ldots, n$,

$$-z_j \tilde{z}_j \left(1 - \frac{\sqrt{\tilde{z}_j}}{\sqrt{z_j}}\right) \leq z_j \tilde{z}_j \frac{1}{2}\left(\frac{\tilde{z}_j}{z_j} - 1\right) = \frac{1}{2}\tilde{z}_j(\tilde{z}_j - z_j).$$

Adding the two inequalities, we get

$$\begin{aligned} & p_j\left(z'_j - z_j\right)\left(\frac{1}{2}(z_j - \tilde{z}_j) + \mu\left(\tilde{z}_j - \left(\sqrt{\tilde{z}_j}\right)^3 (\sqrt{z_j})^{-1}\right)\right) \\ \leq\ & \frac{1}{2}\mu p_j (z'_j - \tilde{z}_j)(z_j - \tilde{z}_j) + \frac{1}{2} p_j (z_j - \tilde{z}_j)(z'_j - z_j). \end{aligned} \tag{2.5}$$

Using the identities

$$\frac{1}{2}(z'_j - \tilde{z}_j)(z_j - \tilde{z}_j) = \frac{1}{4}\left((z_j - \tilde{z}_j)^2 - (z_j - z'_j)^2 + (z'_j - \tilde{z}_j)^2\right)$$

$$\frac{1}{2}(z_j-\tilde{z}_j)(z'_j-z_j)=\frac{1}{4}\left((z'_j-\tilde{z}_j)^2-(z'_j-z_j)^2-(z_j-\tilde{z}_j)^2\right)$$

and rearranging the terms appropriately, one obtains

$$\begin{aligned} & p_j\left(z_j-z'_j\right)\left(\frac{1}{2}(z_j-\tilde{z}_j)+\mu\left(\tilde{z}_j-\left(\sqrt{\tilde{z}_j}\right)^3\left(\sqrt{z_j}\right)^{-1}\right)\right) \\ \geq\ & \frac{(1+\mu)p_j}{4}\left((z_j-z'_j)^2-(\tilde{z}_j-z'_j)^2\right)+\frac{(1-\mu)p_j}{4}(\tilde{z}_j-z_j)^2. \end{aligned} \tag{2.6}$$

Summing over $j=1,\ldots,n$, and using the definition of Ψ in (2.2), one arrives at

$$(z-z')^\top P\Psi(\tilde{z},z)\geq\frac{1+\mu}{4}\left(\|z-z'\|_P^2-\|\tilde{z}-z'\|_P^2\right)+\frac{1-\mu}{4}\|\tilde{z}-z\|_P^2. \tag{2.7}$$

It follows from (2.3) and (2.7) that

$$f(z')-f(z)+(z'-z)^\top q(z)\geq\frac{1+\mu}{4}\left(\|z-z'\|_P^2-\|\tilde{z}-z'\|_P^2\right)+\frac{1-\mu}{4}\|\tilde{z}-z\|_P^2. \tag{2.8}$$

Note that

$$\frac{1}{2}(z'-z)^\top P(z-\tilde{z})=\frac{1}{4}\left(\|\tilde{z}-z'\|_P^2-\|z-z'\|_P^2\right)-\frac{1}{4}\|z-\tilde{z}\|_P^2. \tag{2.9}$$

Adding (2.8) and (2.9), one obtains (2.4). This completes the proof. □

2.2. **Variational inequality characterization.** Let the Lagrangian function of (1.6) be

$$L(x,y,\lambda)=\theta_1(x)+\theta_2(y)-\lambda^\top(Ax+By-b),$$

which is defined on $\mathscr{W}=\mathbb{R}_+^{n_1}\times\mathbb{R}_+^{n_2}\times\mathbb{R}^l$. We call $w^*=(x^*,y^*,\lambda^*)\in\mathscr{W}$ a saddle point of the Lagrangian function if it satisfies

$$L(x^*,y^*,\lambda)\leq L(x^*,y^*,\lambda^*)\leq L(x,y,\lambda^*),\quad\forall(x,y,\lambda)\in\mathscr{W}.$$

Accordingly, for any$(x,y,\lambda)\in\mathscr{W}$, we deduce the following inequalities

$$\begin{cases} \theta_1(x)+\theta_1(y)-(\theta_2(x^*)+\theta_2(y^*))+(x-x^*)^\top(-A^\top\lambda^*)+(y-y^*)^\top(-B^\top\lambda^*)\geq 0, \\ (\lambda-\lambda^*)^\top(Ax^*+By^*-b)\geq 0. \end{cases}$$

Then, problem (1.6) is equivalent to the mixed variational inequalities: find $w^*\in\mathscr{W}$ such that

$$\theta(u)-\theta(u^*)+(w-w^*)^\top F(w^*)\geq 0,\quad\forall\, w\in\mathscr{W}, \tag{2.10}$$

where

$$u=\begin{pmatrix} x \\ y \end{pmatrix},\quad \theta(u)=\theta_1(x)+\theta_2(y),\quad w=\begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}\quad\text{and}\quad F(w)=\begin{pmatrix} -A^\top\lambda \\ -B^\top\lambda \\ Ax+By-b. \end{pmatrix}.$$

Clearly, F is monotone, i.e., $(F(w_1)-F(w_2)^\top(w_1-w_2)\geq 0,\forall w_1,w_2\in\mathscr{W}$.

2.3. **Some notations.** Our analysis needs several matrices defined by

$$M=\begin{pmatrix} I_{n_2} & 0\\ \alpha\beta B & (\alpha+\gamma)\beta I_l \end{pmatrix},\quad Q=\begin{pmatrix} \frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma}\beta B^\top B & -\frac{\alpha}{\alpha+\gamma}B^\top\\ -\frac{\alpha}{\alpha+\gamma}B & \frac{1}{(\alpha+\gamma)\beta}I_l \end{pmatrix},$$

$$H=\begin{pmatrix} \frac{(1+\mu)}{2}R & 0 & 0\\ 0 & \frac{(1+\mu)}{2}S+\frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma}\beta B^\top B & -\frac{\alpha}{\alpha+\gamma}B^T\\ 0 & -\frac{\alpha}{\alpha+\gamma}B & \frac{1}{(\alpha+\gamma)\beta}I_l \end{pmatrix}$$

and

$$N=\begin{pmatrix} \frac{\mu}{2}R & 0 & 0\\ 0 & \frac{\mu}{2}S & 0\\ 0 & 0 & 0 \end{pmatrix},\quad K=\begin{pmatrix} (1-\alpha)\beta B^\top B & (1-\alpha)\beta B^\top\\ (1-\alpha)\beta B & (2-\alpha-\gamma)\beta I_l \end{pmatrix}$$

Lemma 2.3. *If $\mu\in(0,1),\gamma\in(0,1)$ and $\alpha\in(0,1)$, then*

(1) the matrices M and Q defined in Subsection (2.3) satisfy

$$M^\top QM=\begin{pmatrix} (1-\alpha)\beta B^\top B & 0\\ 0 & (\alpha+\gamma)\beta I_l \end{pmatrix}.$$

(2) the matrices Q,H and K are symmetric positive definite.

Proof. (1) is obvious. Next, we prove (2). It is obvious that Q,H and K are symmetric. Next, we prove that they are positive definite. Since $\gamma\in(0,1)$ and $\alpha\in(0,1)$, we have $\alpha+\gamma-\alpha\gamma\geq\alpha\geq\alpha^2$. Then, for any $v=(y,\lambda)\neq(0,0)$,

$$\begin{aligned} v^\top Qv &= \frac{1}{(\alpha+\gamma)\beta}\Big((\alpha+\gamma-\alpha\gamma)\beta^2\|By\|^2-2\alpha\beta y^\top B^\top\lambda+\|\lambda\|^2\Big)\\ &\geq \frac{1}{(\alpha+\gamma)\beta}\Big(\alpha^2\beta^2\|By\|^2-2\alpha\beta y^\top B^\top\lambda+\|\lambda\|^2\Big)\\ &\geq \frac{1}{(\alpha+\gamma)\beta}\|\alpha\beta By-\lambda\|^2\\ &> 0. \end{aligned}$$

Thus, Q is positive definite. For any $w=(x,y,\lambda)\neq(0,0,0)$, we have

$$w^\top Hw=\frac{(1+\mu)}{2}\|x\|_R^2+\frac{(1+\mu)}{2}\|y\|_S^2+v^\top Qv>0.$$

Then, H is positive definite. For any $v=(y,\lambda)\neq(0,0)$, we have

$$\begin{aligned} v^\top Kv &= (1-\alpha)\beta\|By\|^2+2(1-\alpha)\beta y^\top B^\top\lambda+(2-\alpha-\gamma)\beta\|\lambda\|^2\\ &\geq (1-\alpha)\beta(\|By\|^2+2y^\top B^\top\lambda+\|\lambda\|^2)\\ &= (1-\alpha)\beta\|By+\lambda\|^2\\ &> 0. \end{aligned}$$

Therefore K is positive definite. This completes the proof. □

3. THE PROPOSED METHOD

Let $\beta > 0$, $r > 0, s > 0$, $R \in \mathbb{R}^{n_1 \times n_1}$, $S \in \mathbb{R}^{n_2 \times n_2}$ be positive definite diagonal matrices, where $R = rI_{n_1 \times n_1}$ and $S = sI_{n_2 \times n_2}$. Let ρ_k be a sequence of positive real numbers such that $0 < \rho_k < \rho$, where $0 < \rho < 1$. We propose the following inertial PRSM with the SQP regularization for solving (1.6).

Algorithm 3.1.

Step 0. *The initial step:*
Give $\varepsilon > 0$, $\mu \in (0,1), \alpha \in (0,1), \gamma \in (0,1)$ and $w^0 = (x^0, y^0, \lambda^0) = (x^{-1}, y^{-1}, \lambda^{-1}) \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++} \times \mathbb{R}^l$.
Set $k = 0$.

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

$$(\bar{x}^k, \bar{y}^k, \bar{\lambda}^k) = (x^k, y^k, \lambda^k) + \rho_k(x^k - x^{k-1}, y^k - y^{k-1}, \lambda^k - \lambda^{k-1}), \tag{3.1a}$$

$$\begin{aligned} x^{k+1} = \arg\min\Big\{ &\theta_1(x) - (\bar{\lambda}^k)^\top (Ax + B\bar{y}^k - b) + \frac{\beta}{2}\|Ax + B\bar{y}^k - b\|^2 \\ &+ rd_\phi(x, \bar{x}^k), \forall x \in \mathbb{R}^{n_1}_{++} \Big\}, \end{aligned} \tag{3.1b}$$

$$\lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha\beta(Ax^{k+1} + B\bar{y}^k - b), \tag{3.1c}$$

$$\begin{aligned} y^{k+1} = \arg\min\Big\{ &\theta_2(y) - (\lambda^{k+\frac{1}{2}})^\top (Ax^{k+1} + By - b) + \frac{\beta}{2}\|Ax^{k+1} + By - b\|^2 \\ &+ rd_\phi(y, \bar{y}^k), \forall y \in \mathbb{R}^{n_2}_{++} \Big\}, \end{aligned} \tag{3.1d}$$

$$\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b). \tag{3.1e}$$

Step 2. if $\|w^{k+1} - w^k\| < \varepsilon$, then stop; otherwise set $k = k+1$ and go to Step 1.

Remark 3.1. Our method can be viewed as an extension for some known results, for example, the following:

- Choosing $\rho_k = 1, \forall k \geq 0$, we obtain the method in [17] with SQP regularization.
- Setting $\rho_k = 1, \forall k \geq 0, \mu = 0$ and the matrices $R = 2L, S = 2T$, we get the method in [16].
- Setting $\mu = 0, \gamma = \alpha \in (0,1)$ and the matrices $R = 2L, S = 2T$, we obtain the method in [19].
- If $\mu = 0, \gamma = \alpha \in (0,1)$ and the matrices $R = 0, S = 0$, the proposed method collapses to the strictly contractive PRSM proposed in [5].
- Setting $\mu = 0, \alpha \in (0,1), \gamma \in (0, \frac{1+\sqrt{5}}{2})$ and the matrices $R = 2L, S = 2T$, the proposed method reduces to the semi-proximal ADMM considered in [4, 21, 22].

Therefore, the new algorithm is expected to be widely applicable.

Remark 3.2. According to their first-order optimality conditions, the minimization problems in (3.1) can be be characterized as: find $(x^{k+1}, y^{k+1}) \in \mathbb{R}^{n_1}_{++} \times \mathbb{R}^{n_2}_{++}$ such that, for any $(x, y) \in \mathbb{R}^{n_1}_{+} \times \mathbb{R}^{n_2}_{+}$,

$$\theta_1(x) - \theta_1(x^{k+1}) + (x - x^{k+1})^\top \left(-A^\top [\bar{\lambda}^k - \beta(Ax^{k+1} + B\bar{y}^k - b)] + r\Psi(\bar{x}^k, x^{k+1}) \right) \geq 0, \tag{3.2}$$

$$\theta_2(y) - \theta_2(y^{k+1}) + (y - y^{k+1})^\top \left(-B^\top [\lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} + By^{k+1} - b)] + s\Psi(\bar{y}^k, y^{k+1}) \right) \geq 0, \tag{3.3}$$

Remark 3.3. Note that, if $\theta_1(x)$ and $\theta_2(y)$ are differentiable, then (3.1) reduces to

$$\begin{cases} (\bar{x}^k,\bar{y}^k,\bar{\lambda}^k) = (x^k,y^k,\lambda^k) + \rho_k(x^k - x^{k-1}, y^k - y^{k-1}, \lambda^k - \lambda^{k-1}), \\ \nabla\theta_1(x^{k+1}) - A^\top[\bar{\lambda}^k - \beta(Ax^{k+1} + B\bar{y}^k - b)] + r\Psi(\bar{x}^k, x^{k+1}) = 0, \\ \lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha\beta(Ax^{k+1} + B\bar{y}^k - b), \\ \nabla\theta_2(y^{k+1}) - B^\top[\lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} + By^{k+1} - b)] + s\Psi(\bar{y}^k, y^{k+1}) = 0, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \gamma\beta(Ax^{k+1} + By^{k+1} - b). \end{cases} \tag{3.4}$$

Compared to [4, 5, 16, 19, 21, 22] (instead of dealing with two sub-variational inequalities), the new iterate of the proposed method is obtained by solving two easier systems of nonlinear equations.

Throughout this paper, we make the following standard assumptions:

Assumption A. The sequence w^k generated by Algorithm 3.1 satisfies

$$\sum_{k=0}^{\infty} \rho_k \|\bar{w}^k - w^{k-1}\|_H^2 < \infty.$$

Assumption B. The solution set of (1.6), denoted by $\mathscr{W}^*$, is nonempty.

In the following, we prove some properties, which are useful for establishing the main result. The first lemma presents a lower bound of $\theta(u) - \theta(u^{k+1}) + (w - w^{k+1})^\top(F(w^{k+1}) + H(w^{k+1} - \bar{w}^k))$.

Lemma 3.4. *Let w^k be generated by Algorithm 3.1. Then, for any $w = (x, y, \lambda) \in \mathscr{W}^*$,*

$$\begin{aligned} & \theta(u) - \theta(u^{k+1}) + (w - w^{k+1})^\top(F(w^{k+1}) + H(w^{k+1} - \bar{w}^k)) \\ \geq\ & (1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top(B\bar{y}^k - By^{k+1}) \\ & + (1-\alpha-\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2 - \|\bar{w}^k - w^{k+1}\|_N^2. \end{aligned} \tag{3.5}$$

Proof. Applying Lemma 2.2 to (3.2) by setting $P = R, \tilde{z} = \bar{x}^k$, $z = x^{k+1}$, $f(\cdot) = \theta_1(\cdot), z' = x$ and $q(z) = -A^\top[\bar{\lambda}^k - \beta(Ax^{k+1} + B\bar{y}^k - b)]$ in (2.3), we get

$$\begin{aligned} & \theta_1(x^{k+1}) - \theta_1(x) + (x^{k+1} - x)^\top \Big\{ -A^\top\bar{\lambda}^k + \beta A^\top(Ax^{k+1} + By^{k+1} - b) \\ & + \beta A^\top(B\bar{y}^k - By^{k+1}) - \frac{(1+\mu)}{2} R(\bar{x}^k - x^{k+1}) \Big\} \leq \frac{\mu}{2}\|\bar{x}^k - x^{k+1}\|_R^2. \end{aligned} \tag{3.6}$$

It follows from (3.1c) and (3.1e) that

$$\bar{\lambda}^k = \lambda^{k+1} + (\alpha+\gamma)\beta(Ax^{k+1} + By^{k+1} - b) + \alpha\beta(B\bar{y}^k - By^{k+1}). \tag{3.7}$$

Substituting (3.7) into (3.6), we obtain

$$\begin{aligned} & \theta_1(x^{k+1}) - \theta_1(x) + (x^{k+1} - x)^\top \Big\{ -A^\top\lambda^{k+1} + \beta(1-\alpha-\gamma)A^\top(Ax^{k+1} + By^{k+1} - b) \\ & + \beta(1-\alpha)A^\top(B\bar{y}^k - By^{k+1}) - \frac{(1+\mu)}{2} R(\bar{x}^k - x^{k+1}) \Big\} \leq \frac{\mu}{2}\|\bar{x}^k - x^{k+1}\|_R^2. \end{aligned} \tag{3.8}$$

Similarly, by setting $P = S, \tilde{z} = \bar{y}^k$, $z = y^{k+1}$, $f(\cdot) = \theta_2(\cdot), z' = y$ and $q(z) = -B^\top[\lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} + By^{k+1} - b)]$ in (2.3), we conclude from Lemma 2.2 that

$$\begin{aligned} & \theta_2(y^{k+1}) - \theta_2(y) + (y^{k+1} - y)^\top \Big\{ -B^\top[\lambda^{k+\frac{1}{2}} - \beta(Ax^{k+1} + By^{k+1} - b)] \\ & - \tfrac{(1+\mu)}{2} S(\bar{y}^k - y^{k+1}) \Big\} \leq \tfrac{\mu}{2}\|\bar{y}^k - y^{k+1}\|_S^2. \end{aligned} \tag{3.9}$$

Using (3.1e), we have

$$
\begin{aligned}
& \theta_2(y^{k+1}) - \theta_2(y) + (y^{k+1}-y)^\top \Big\{ -B^\top \lambda^{k+1} - \frac{(1+\mu)}{2} S(\bar{y}^k - y^{k+1}) \\
& -\frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma} \beta B^\top B(\bar{y}^k - y^{k+1}) + \frac{\alpha}{\alpha+\gamma} B^\top (\bar{\lambda}^k - \lambda^{k+1}) \Big\} \\
\leq \;& (y^{k+1}-y)^\top \Big\{ \frac{\alpha}{\alpha+\gamma} B^\top (\bar{\lambda}^k - \lambda^{k+1}) - \frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma} \beta B^\top B(\bar{y}^k - y^{k+1}) \\
& -\beta(1-\gamma) B^\top (Ax^{k+1} + By^{k+1} - b) \Big\} + \frac{\mu}{2} \|\bar{y}^k - y^{k+1}\|_S^2.
\end{aligned}
$$

It follows from (3.7) that

$$
\begin{aligned}
& \theta_2(y^{k+1}) - \theta_2(y) + (y^{k+1}-y)^\top \Big\{ -B^\top \lambda^{k+1} - \frac{(1+\mu)}{2} S(\bar{y}^k - y^{k+1}) \\
& -\frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma} \beta B^\top B(\bar{y}^k - y^{k+1}) + \frac{\alpha}{\alpha+\gamma} B^\top (\bar{\lambda}^k - \lambda^{k+1}) \Big\} \\
\leq \;& (y^{k+1}-y)^\top \Big\{ (\alpha-1)\beta B^\top B(\bar{y}^k - y^{k+1}) + \beta(\alpha+\gamma-1) B^\top (Ax^{k+1} + By^{k+1} - b) \Big\} \\
& + \frac{\mu}{2} \|\bar{y}^k - y^{k+1}\|_S^2
\end{aligned} \tag{3.10}
$$

and

$$
Ax^{k+1} + By^{k+1} - b + \frac{\alpha}{(\alpha+\gamma)} (B\bar{y}^k - By^{k+1}) - \frac{1}{(\alpha+\gamma)\beta} (\bar{\lambda}^k - \lambda^{k+1}) = 0. \tag{3.11}
$$

Combining (3.8), (3.10) and (3.11) and using $Ax + By - b = 0$, we can get the assertion of this lemma immediately. □

Lemma 3.5. *Let w^k be generated by Algorithm 3.1. Then, for any $w^{k+1} \in \mathscr{W}$,*

$$
\begin{aligned}
& (1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top (B\bar{y}^k - By^{k+1}) + (1-\alpha-\gamma)\beta \|Ax^{k+1} + By^{k+1} - b\|^2 \\
\geq \;& -\frac{(\alpha+\gamma-\alpha\gamma)}{2(\alpha+\gamma)} \beta \|B\bar{y}^k - By^{k+1}\|^2 + \frac{\alpha}{(\alpha+\gamma)} (B\bar{y}^k - By^{k+1})^\top (\bar{\lambda}^k - \lambda^{k+1}) \\
& -\frac{1}{2(\alpha+\gamma)\beta} \|\bar{\lambda}^k - \lambda^{k+1}\|^2.
\end{aligned} \tag{3.12}
$$

Proof. Using $2a^\top b \geq -\|a\|^2 - \|b\|^2$, we have

$$
\begin{aligned}
& (1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top (B\bar{y}^k - By^{k+1}) + (1-\alpha-\gamma)\beta \|Ax^{k+1} + By^{k+1} - b\|^2 \\
\geq \;& \frac{(1-\alpha-2\gamma)}{2} \beta \|Ax^{k+1} + By^{k+1} - b\|^2 - \frac{(1-\alpha)}{2} \beta \|B\bar{y}^k - By^{k+1}\|^2.
\end{aligned} \tag{3.13}
$$

From $\bar{\lambda}^k - \lambda^{k+1} = (\alpha+\gamma)\beta(Ax^{k+1} + By^{k+1} - b) + \alpha\beta(B\bar{y}^k - By^{k+1})$, we get

$$
\begin{aligned}
& -\frac{(\alpha+\gamma-\alpha\gamma)}{2(\alpha+\gamma)} \beta \|B\bar{y}^k - By^{k+1}\|^2 + \frac{\alpha}{(\alpha+\gamma)} (B\bar{y}^k - By^{k+1})^\top (\bar{\lambda}^k - \lambda^{k+1}) \\
& -\frac{1}{2(\alpha+\gamma)\beta} \|\bar{\lambda}^k - \lambda^{k+1}\|^2 \\
= \;& -\frac{(\alpha+\gamma)\beta}{2} \|Ax^{k+1} + By^{k+1} - b\|^2 - \frac{(1-\alpha)}{2} \beta |B\bar{y}^k - By^{k+1}\|^2.
\end{aligned} \tag{3.14}
$$

Since $\gamma \in (0,1)$, it follows from (3.13) and (3.14) that

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

And the assertion of this lemma is proved □

4. The convergence of the proposed method

In this section, we prove the global convergence for the proposed method. Before proceeding, we need the following lemmas. The following result can be proved by using the technique of Theorem 1 in [23].

Lemma 4.1. *Let w^k be generated by Algorithm 3.1. Then, for any $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$, we have*

$$\sum_{k=0}^{\infty} \|w^k - w^*\|_H^2 < \infty.$$

Proof. Using Lemmas 3.4 and 3.5, we have

$$\begin{aligned}
& \theta(u^{k+1}) - \theta(u^*) + (w^{k+1}-w^*)^\top(F(w^{k+1}) + H(w^{k+1}-\bar{w}^k)) - \|\bar{w}^k - w^{k+1}\|_N^2 \\
\leq\ & \frac{(\alpha+\gamma-\alpha\gamma)}{2(\alpha+\gamma)}\beta\|B\bar{y}^k - By^{k+1}\|^2 - \frac{\alpha}{(\alpha+\gamma)}(B\bar{y}^k - By^{k+1})^\top(\bar{\lambda}^k - \lambda^{k+1}) \\
& + \frac{1}{2(\alpha+\gamma)\beta}\|\bar{\lambda}^k - \lambda^{k+1}\|^2.
\end{aligned} \tag{4.1}$$

Setting $w = w^{k+1}$ in (2.10), we obtain

$$\theta(u^{k+1}) - \theta(u^*) + (w^{k+1}-w^*)^\top F(w^*) \geq 0. \tag{4.2}$$

Combining (4.1) and (4.2), and using the monotonicity of F, we get

$$\begin{aligned}
& (w^{k+1}-w^*)^\top H(w^{k+1}-\bar{w}^k) - \|\bar{w}^k - w^{k+1}\|_N^2 \\
\leq\ & \frac{(\alpha+\gamma-\alpha\gamma)}{2(\alpha+\gamma)}\beta\|B\bar{y}^k - By^{k+1}\|^2 - \frac{\alpha}{(\alpha+\gamma)}(B\bar{y}^k - By^{k+1})^\top(\bar{\lambda}^k - \lambda^{k+1}) \\
& + \frac{1}{2(\alpha+\gamma)\beta}\|\bar{\lambda}^k - \lambda^{k+1}\|^2.
\end{aligned} \tag{4.3}$$

Since $\bar{w}^k = w^k + \rho_k(w^k - w^{k-1})$, it follows that

$$\begin{aligned}
& (w^{k+1}-w^*)^\top H(w^{k+1}-\bar{w}^k) \\
=\ & (w^{k+1}-w^*)^\top H(w^{k+1}-w^k) - \rho_k(w^{k+1}-w^*)^\top H(w^k - w^{k-1}).
\end{aligned} \tag{4.4}$$

Setting $\varphi_k := \frac{1}{2}\|w^k - w^*\|_H^2$, using the following identity

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

we obtain

$$(w^{k+1}-w^*)^\top H(w^{k+1}-w^k) = \varphi_{k+1} - \varphi_k + \frac{1}{2}\|w^{k+1}-w^k\|_H^2, \tag{4.5}$$

and

$$(w^{k+1}-w^*)^\top H(w^k - w^{k-1}) = \varphi_k - \varphi_{k-1} + \frac{1}{2}\|w^k - w^{k-1}\|_H^2 + (w^{k+1}-w^k)^\top H(w^k - w^{k-1}).$$

Substituting the above equalities into (4.4), we get

$$
\begin{aligned}
(w^{k+1}-w^*)^\top H(w^{k+1}-\bar{w}^k) &= \varphi_{k+1}-\varphi_k-\rho_k(\varphi_k-\varphi_{k-1})+\frac{1}{2}\|w^{k+1}-w^k\|_H^2 \\
&\quad -\frac{\rho_k}{2}\|w^k-w^{k-1}\|_H^2-\rho_k(w^{k+1}-w^k)^\top H(w^k-w^{k-1}). \\
&= \varphi_{k+1}-\varphi_k-\rho_k(\varphi_k-\varphi_{k-1})+\frac{1}{2}\|w^{k+1}-w^k-\rho_k(w^k-w^{k-1})\|_H^2 \\
&\quad -\frac{(\rho_k+\rho_k^2)}{2}\|w^k-w^{k-1}\|_H^2 \\
&= \varphi_{k+1}-\varphi_k-\rho_k(\varphi_k-\varphi_{k-1})+\frac{1}{2}\|w^{k+1}-\bar{w}^k\|_H^2 \\
&\quad -\frac{(\rho_k+\rho_k^2)}{2}\|w^k-w^{k-1}\|_H^2. \qquad (4.6)
\end{aligned}
$$

Combining (4.3) and (4.6), we obtain

$$
\begin{aligned}
\varphi_{k+1}-\varphi_k-\rho_k(\varphi_k-\varphi_{k-1}) &\leq \|\bar{w}^k-w^{k+1}\|_N^2+\frac{(\alpha+\gamma-\alpha\gamma)}{2(\alpha+\gamma)}\beta\|B\bar{y}^k-By^{k+1}\|^2 \\
&\quad -\frac{\alpha}{(\alpha+\gamma)}(B\bar{y}^k-By^{k+1})^\top(\bar{\lambda}^k-\lambda^{k+1}) \\
&\quad +\frac{1}{2(\alpha+\gamma)\beta}\|\bar{\lambda}^k-\lambda^{k+1}\|^2-\frac{1}{2}\|w^{k+1}-\bar{w}^k\|_H^2 \\
&\quad +\frac{(\rho_k+\rho_k^2)}{2}\|w^k-w^{k-1}\|_H^2.
\end{aligned} \qquad (4.7)
$$

Since $\frac{1}{2}(\rho_k+\rho_k^2)\leq\rho_k$, using the definition of the matrix H in Subsection 2.3, we have that (4.7) is reduced to

$$
\varphi_{k+1}-\varphi_k-\rho_k(\varphi_k-\varphi_{k-1})\leq\rho_k\|w^k-w^{k-1}\|_H^2. \qquad (4.8)
$$

Let $\theta_k:=\varphi_k-\varphi_{k-1}$ and $\delta_k:=\rho_k\|w^k-w^{k-1}\|_H^2$. It follows from (4.8) that $\theta_{k+1}\leq\rho_k\theta_k+\delta_k\leq\rho[\theta_k]_++\delta_k$, where $[t]_+=\max\{0,t\}$. Furthermore, we deduce that

$$
[\theta_{k+1}]_+\leq\rho[\theta_k]_++\delta_k\leq\rho^{k+1}[\theta_0]_++\sum_{j=0}^{k}\rho^j\delta_{k-j}.
$$

Note that $w^0=w^{-1}$. Then $\theta_0=[\theta_0]_+=0,\delta_0=0$. Therefore, from Assumption A, we obtain

$$
\sum_{k=0}^{\infty}[\theta_k]_+\leq\frac{1}{1-\rho}\sum_{k=0}^{\infty}\delta_k=\frac{1}{1-\rho}\sum_{k=1}^{\infty}\delta_k<\infty. \qquad (4.9)
$$

Let $m_k:=\varphi_k-\sum_{j=1}^{k}[\theta_j]_+$. From (4.9) and $\varphi_k\geq 0$, we get a lower bound of sequence $\{m_k\}$. On the other hand,

$$
m_{k+1}=\varphi_{k+1}-[\theta_{k+1}]_+-\sum_{j=1}^{k}[\theta_j]_+\leq\varphi_{k+1}-\theta_{k+1}-\sum_{j=1}^{k}[\theta_j]_+=\varphi_k-\sum_{j=1}^{k}[\theta_j]_+=m_k,
$$

we have that $\{m_k\}$ is a monotone non-increasing sequence. So $\{m_k\}$ converges and

$$
\lim_{k\to\infty}\varphi_k=\sum_{j=1}^{\infty}[\theta_j]_++\lim_{k\to\infty}m_k.
$$

Therefore, $\{\varphi_k\}$ converges and $\sum_{k=0}^{\infty}\|w^k-w^*\|_H^2<\infty$. The proof is completed. □

Lemma 4.2. *Let w^k be generated by Algorithm 3.1. Then, for any $w^{k+1} \in \mathscr{W}$, we have*

$$\sum_{k=0}^{\infty} \|w^{k+1} - \bar{w}^k\|_H^2 < \infty.$$

Proof. From (3.5) and (4.2), we have

$$\begin{array}{ll} & (w^* - w^{k+1})^\top H(w^{k+1} - \bar{w}^k) \\ \geq & (1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top(B\bar{y}^k - By^{k+1}) \\ & +(1-\alpha-\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2 - \|\bar{w}^k - w^{k+1}\|_N^2. \end{array} \tag{4.10}$$

Using the following identity

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

for $a = \bar{w}^k - w^*, b = w^{k+1} - w^*$, we obtain

$$\|\bar{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 = \|\bar{w}^k - w^{k+1}\|_H^2 + 2(w^* - w^{k+1})^\top H(w^{k+1} - \bar{w}^k).$$

It follows from (4.10) that

$$\begin{array}{rcl} \|\bar{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 & \geq & \|\bar{w}^k - w^{k+1}\|_H^2 + 2(1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top(B\bar{y}^k - B\bar{y}^k) \\ & & +2(1-\alpha-\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2 - 2\|\bar{w}^k - w^{k+1}\|_N^2. \end{array} \tag{4.11}$$

On the other hand, since $\bar{\lambda}^k - \lambda^{k+1} = (\alpha+\gamma)\beta(Ax^{k+1} + By^{k+1} - b) + \alpha\beta(B\bar{y}^k - By^{k+1})$, it is easy to show that

$$\begin{array}{rcl} \|\bar{w}^k - w^{k+1}\|_H^2 & = & \dfrac{(1+\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + (1-\alpha)\beta\|B\bar{y}^k - By^{k+1}\|^2 \\ & & +(\alpha+\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2 \end{array}$$

and consequently

$$\begin{array}{ll} & \|\bar{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 \\ \geq & \dfrac{(1-\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + (1-\alpha)\beta\|B\bar{y}^k - By^{k+1}\|^2 \\ & +2(1-\alpha)\beta(Ax^{k+1} + By^{k+1} - b)^\top(B\bar{y}^k - B\bar{y}^k) \\ & +(2-\alpha-\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2. \end{array} \tag{4.12}$$

Denote $v = \begin{pmatrix} y \\ \lambda \end{pmatrix}$. Then

$$\bar{v}^k - v^{k+1} = \begin{pmatrix} \bar{y}^k - y^{k+1} \\ \bar{\lambda}^k - \lambda^{k+1} \end{pmatrix} = M\begin{pmatrix} \bar{y}^k - y^{k+1} \\ Ax^{k+1} + By^{k+1} - b \end{pmatrix}.$$

Using the definition of the matrix K in Subsection 2.3, we have that (4.12) becomes

$$\begin{array}{rcl} \|\bar{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 & \geq & \dfrac{(1-\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) \\ & & +(\bar{v}^k - v^{k+1})^\top M^{-T}KM^{-1}(\bar{v}^k - v^{k+1}). \end{array} \tag{4.13}$$

It follows from Lemma 2.3 (1) that

$$\begin{aligned} & K-\tau M^\top QM \\ = & \begin{pmatrix} (1-\tau)(1-\alpha)\beta B^\top B & (1-\alpha)\beta B^\top \\ (1-\alpha)\beta B & (2-(1+\tau)(\alpha+\gamma))\beta I_l \end{pmatrix} \\ = & \begin{pmatrix} \sqrt{\beta}B^\top & 0 \\ 0 & \sqrt{\beta}I_l \end{pmatrix}\begin{pmatrix} (1-\tau)(1-\alpha)I_l & (1-\alpha)I_l \\ (1-\alpha)I_l & (2-(1+\tau)(\alpha+\gamma))I_l \end{pmatrix}\begin{pmatrix} \sqrt{\beta}B & 0 \\ 0 & \sqrt{\beta}I_l \end{pmatrix}. \end{aligned}$$

Since $\alpha\in(0,1)$ and $\gamma\in(0,1)$, we conclude that $K\geq\tau M^\top QM$, where $\tau=\frac{1-\sqrt{1-(\alpha+\gamma)(1-\gamma)}}{\alpha+\gamma}\in(0,1)$. It follows from (4.13) that

$$\begin{aligned} \|\bar{w}^k-w^*\|_H^2-\|w^{k+1}-w^*\|_H^2 \geq & \frac{(1-\mu)}{2}(\|\bar{x}^k-x^{k+1}\|_R^2+\|\bar{y}^k-y^{k+1}\|_S^2) \\ & +\tau(\bar{v}^k-v^{k+1})^\top Q(\bar{v}^k-v^{k+1}) \\ \geq & \frac{\tau(1-\mu)}{1+\mu}\|\bar{w}^k-w^{k+1}\|_H^2. \end{aligned} \tag{4.14}$$

From (4.5), we have

$$\begin{aligned} \|\bar{w}^k-w^*\|_H^2 &= \|w^k-w^*+\rho_k(w^k-w^{k-1})\|_H^2 \\ &= \|w^k-w^*\|_H^2+2\rho_k(w^k-w^*)^\top H(w^k-w^{k-1})+\rho_k^2\|w^k-w^{k-1}\|_H^2 \\ &= \|w^k-w^*\|_H^2+2\rho_k(\varphi_k-\varphi_{k-1})+(\rho_k+\rho_k^2)\|w^k-w^{k-1}\|_H^2 \\ &\leq \|w^k-w^*\|_H^2+2(\rho_k\theta_k+\delta_k) \\ &\leq \|w^k-w^*\|_H^2+2(\rho[\theta_k]_++\delta_k) \\ &\leq \|w^k-w^*\|_H^2+2(\rho^{k+1}[\theta_0]_++\sum_{j=0}^k\rho^j\delta_{k-j}). \end{aligned} \tag{4.15}$$

Since $\theta_0=[\theta_0]_+=0$, combining (4.14) and (4.15), we obtain

$$\|\bar{w}^k-w^{k+1}\|_H^2\leq\frac{1+\mu}{\tau(1-\mu)}\left(\|w^k-w^*\|_H^2+2\sum_{j=0}^k\rho^j\delta_{k-j}\right).$$

Using Lemma 4.1 and Assumption A, we get

$$\sum_{k=0}^\infty\|\bar{w}^k-w^{k+1}\|_H^2\leq\frac{1+\mu}{\tau(1-\mu)}\left(\sum_{k=0}^\infty\|w^k-w^*\|_H^2+\frac{2}{1-\rho}\sum_{k=1}^\infty\delta_k\right)<\infty.$$

The proof is completed. □

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

Theorem 4.3. *Let w^k be generated by Algorithm 3.1. Then*

(i)

$$\sum_{k=0}^\infty\|Ax^{k+1}+By^{k+1}-b\|^2<\infty \quad \textit{and} \quad \lim_{k\to\infty}\|Ax^{k+1}+By^{k+1}-b\|=0.$$

(ii) The objective function value $\theta_1(x^k)+\theta_2(y^k)$ converges to the optimal value of (1.6).

(iii) The sequence $\{w^k\}$ converges to a point $w^\infty\in\mathscr{W}^*$.

Proof. It follows from (3.1c) and (3.1e) that

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

Since $\alpha \in (0,1)$ and $\gamma \in (0,1)$, we have

$$\begin{aligned}
& \|w^{k+1} - \bar{w}^k\|_H^2 \qquad (4.16)\\
= \;& \frac{(1+\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + \frac{\alpha+\gamma-\alpha\gamma}{\alpha+\gamma}\beta\|By^{k+1} - B\bar{y}^k\|^2 \\
& + \frac{1}{(\alpha+\gamma)\beta}\|\lambda^{k+1} - \bar{\lambda}^k\|^2 - \frac{2\alpha}{\alpha+\gamma}(By^{k+1} - B\bar{y}^k)^\top(\lambda^{k+1} - \bar{\lambda}^k) \\
\geq \;& \frac{(1+\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + \frac{\alpha^2}{\alpha+\gamma}\beta\|By^{k+1} - B\bar{y}^k\|^2 \\
& + \frac{1}{(\alpha+\gamma)\beta}\|\lambda^{k+1} - \bar{\lambda}^k\|^2 - \frac{2\alpha}{\alpha+\gamma}(By^{k+1} - B\bar{y}^k)^\top(\lambda^{k+1} - \bar{\lambda}^k) \\
= \;& \frac{(1+\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + \\
& + (\alpha+\gamma)\beta\left\|\frac{1}{(\alpha+\gamma)\beta}(\bar{\lambda}^k - \lambda^{k+1}) - \frac{\alpha}{\alpha+\gamma}(B\bar{y}^k - By^{k+1})\right\|^2 \\
= \;& \frac{(1+\mu)}{2}(\|\bar{x}^k - x^{k+1}\|_R^2 + \|\bar{y}^k - y^{k+1}\|_S^2) + (\alpha+\gamma)\beta\|Ax^{k+1} + By^{k+1} - b\|^2.
\end{aligned} \tag{4.17}$$

From (4.16) and Lemma 4.2, we get

$$\sum_{k=0}^{\infty}\|Ax^{k+1} + By^{k+1} - b\|^2 \leq \sum_{k=0}^{\infty}\frac{1}{(\alpha+\gamma)\beta}\|w^{k+1} - \bar{w}^k\|_H^2 < \infty$$

and

$$\lim_{k\to\infty}\|Ax^{k+1} + By^{k+1} - b\| = 0,$$

we obtain the assertion (i). Let $w^* = (x^*, y^*, \lambda^*) \in \mathscr{W}^*$. Setting $w = (x^k, y^k, \lambda^*)$ in (2.10), we obtain

$$\theta_1(x^k) + \theta_2(y^k) - \theta_1(x^*) - \theta_2(y^*) - (\lambda^*)^\top(Ax^k + By^k - b) \geq 0.$$

It follows from (i) that

$$\liminf_{k\to\infty}(\theta_1(x^k) + \theta_2(y^k)) \geq \theta_1(x^*) + \theta_2(y^*). \tag{4.18}$$

On the other hand, setting $w = w^*$ and $w^k = (x^k, y^k, \lambda^*)$ in (3.5), we have

$$\begin{aligned}
\theta_1(x^*) + \theta_2(y^*) - \theta_1(x^k) - \theta_2(y^k) \;\geq\; & (\lambda^*)^\top(Ax^k + By^k - b) + (w^k - w^*)^\top H(w^k - \bar{w}^{k-1}) \\
& + (1-\alpha)\beta(Ax^k + By^k - b)^\top(B\bar{y}^{k-1} - By^k) \\
& + (1-\alpha-\gamma)\beta\|Ax^k + By^k - b\|^2 - \|\bar{w}^{k-1} - w^k\|_N^2.
\end{aligned} \tag{4.19}$$

From Lemma 3.5, it is easy to prove that

$$\begin{aligned}
& -(1-\alpha)\beta(Ax^k + By^k - b)^\top(B\bar{y}^{k-1} - By^k) - (1-\alpha-\gamma)\beta\|Ax^k + By^k - b\|^2 \\
& + \|\bar{w}^{k-1} + w^k\|_N^2 \leq \frac{1}{2}\|\bar{w}^{k-1} - w^k\|_H^2.
\end{aligned} \tag{4.20}$$

It follows from $\lim_{k\to\infty}\|Ax^k + By^k - b\| = 0, \lim_{k\to\infty}\|\bar{w}^{k-1} - w^k\|_H^2 = 0$, and $\sum_{k=0}^{\infty}\|w^k - w^*\|_H^2 < \infty$ that

$$\limsup_{k\to\infty}(\theta_1(x^k) + \theta_2(y^k)) \leq \theta_1(x^*) + \theta_2(y^*). \tag{4.21}$$

Combining (4.18) and (4.21), we obtain the assertion (ii).

Clearly, from Lemma 4.1 we have that $\lim_{k\to\infty}\|w^k - w^*\|_H$ exists for any $w^* \in \mathscr{W}^*$. Since H is positive definite, it follows that $\{w^k\}$ is bounded sequence and must have at least one limit point. Let w^∞ be a cluster point of $\{w^k\}$ and the subsequence $\{w^{k_j}\}$ converge to w^∞. Since $\mathscr{W}$ is closed set, we have $w^\infty \in \mathscr{W}$. From Lemma 4.2, we have $\lim_{j\to\infty} \|w^{k_j+1} - \bar{w}^{k_j}\|_H = 0$. Thus, the positive definiteness of H implies that $\lim_{j\to\infty} H(w^{k_j+1} - \bar{w}^k) = 0$. On the other hand, from (3.5) and (4.20), we have

$$\theta(u^{k_j+1}) - \theta(u) - (w - w^{k_j+1})^\top (F(w^{k_j+1}) + H(w^{k_j+1} - \bar{w}^{k_j})) \le \frac{1}{2}\|\bar{w}^{k_j+1} - w^{k_j}\|_H^2.$$

Furthermore, by taking the limit $j \to \infty$, we obtain

$$\theta(u) - \theta(u^\infty) + (w - w^\infty)^\top F(w^\infty) \ge 0,$$

which means that $w^\infty \in \mathscr{W}^*$.

Finally, we show that the sequence $\{w^k\}$ has only one cluster point. Suppose that $\{w^k\}$ has two limit points w_1^* and w_2^*. By Lemma 4.1, $\lim_{k\to\infty} \|w^k - w_i^*\|_H$ exists for $i = 1,2$. Assume that $\lim_{k\to\infty} \|w^k - w_i^*\|_H = \eta_i$ exists for $i = 1,2$. Using the following identity

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

we obtain

$$2(w_1^* - w_2^*)^\top H(w^k - w_2^*) = \|w_1^* - w_2^*\|_H^2 - \|w^k - w_1^*\|_H^2 + \|w^k - w_2^*\|_H^2.$$

By taking the limit $k \to \infty$, we obtain $\eta_1 - \eta_2 = -\|w_1^* - w_2^*\|_H^2 = \|w_1^* - w_2^*\|_H^2$. Then, $\|w_1^* - w_2^*\|_H^2 = 0$. Since H is positive definite, this implies that $w_1^* = w_2^*$. Therefore, the sequence $\{w^k\}$ converges to $w^\infty \in \mathscr{W}^*$. □

In the following theorem, we prove that the proposed method can find an approximate solution of the mixed variational inequalities (2.10) with an accuracy of $o(1/\sqrt{k})$.

Theorem 4.4. *Let w^k be generated by Algorithm 3.1. Then,*

(a)

$$\min_{1\le i\le k} \|w^i - \bar{w}^{i-1}\|_G = o(1/\sqrt{k}),$$

(b)

$$\min_{1\le i\le k} \|Ax^i + By^i - b\| = o(1/\sqrt{k}),$$

(c)

$$\min_{1\le i\le k} |\theta_1(x^i) + \theta_2(y^i) - \theta_1(x^*) - \theta_2(y^*)| = o(1/\sqrt{k}),$$

Proof. Setting $a_k = \min_{1\le i\le k} \|w^i - \bar{w}^{i-1}\|_H^2$ in Lemma 2.1. It follows from Lemma 4.2 that $\sum_{k=0}^\infty a_k < \infty$. Thus, all the conditions of Lemma 2.1 are satisfied. Hence we deduce that $\min_{1\le i\le k} \|w^i - \bar{w}^{i-1}\|_G = o(1/\sqrt{k})$. In the same way, by using Lemma 2.1 and Theorem 4.3 (i), we get the assertion (b).

Next, we show the assertion (c). Observe from (4.18) and (4.19) that

$$\begin{aligned} |\theta_1(x^i) + \theta_2(y^i) - \theta_1(x^*) - \theta_2(y^*)| \quad \le \quad & |(\lambda^*)^\top (Ax^i + By^i - b)| + |(w^i - w^*)^\top H(w^i - \bar{w}^{i-1})| \\ & + |(1-\alpha)\beta(Ax^i + By^i - b)^\top (B\bar{y}^{i-1} - By^i)| \\ & + |(1-\alpha-\gamma)\beta\|Ax^i + By^i - b\|^2| + \|\bar{w}^{i-1} - w^i\|_N^2. \end{aligned}$$

Using the assertions (a) and (b), and Lemma 4.2, we obtain the desired result. □

5. The conclusion

In this paper, by combining the square quadratic proximal (SQP) method and the inertial proximal Peaceman-Rachford splitting method (PRSM), we proposed an inertial PRSM with the SQP regularization for solving a separable convex minimization model with positive orthant constraints. Instead of dealing with two sub-variational inequalities, the new iterate of the proposed method was obtained by solving two easier systems of nonlinear equations. Under standard assumptions, the global convergence of the proposed method was proved. Our results can be viewed as a extension of the previously known results.

Acknowledgments
The authors are grateful to the referees for their useful and insight comments which improved the presentation of this paper.

References

[1] Y. Cui, X. Li, D. Sun, K.C. Toh, On the convergence properties of a majorized ADMM for linearly constrained convex optimization problems with coupled objective functions, J. Optim. Theory Appl. 169 (2016), 1013-1041.
[2] W. Deng, W. Yin, On the global and linear convergence of the generalized alternating direction method of multipliers, J. Sci. Comput. 66 (2016), 889-916.
[3] B.S. He, L.Z. Liao, D.R. Han, H. Yang, A new inexact alternating directions method for monotone variational inequalities, Math. Program. 92 (2002), 103-118.
[4] B.S. He, X. Yuan, On the $o(1/n)$ convergence rate of the Douglas-Rachford alternating direction method, SIAM J. Numer. Anal. 50 (2012), 700-709.
[5] B.S. He, H. Liu, Z. Wang, X. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, SIAM J. Optim. 24 (2014), 1011-1040.
[6] B.S. He, X. Yuan, On non-ergodic convergence rate of Douglas-Rachford alternating direction method of multipliers, Numer. Math. 130 (2015), 567-577.
[7] M.H. Xu, Proximal alternating directions method for structured variational inequalities, J. Optim. Theory Appl. 134 (2007), 107-117.
[8] J. Eckstein, P.J.S. Silva, A practical relative error criterion for augmented Lagrangians, Math. Program. 141 (2013), 319-348.
[9] J. Eckstein, W. Yao, Approximate ADMM algorithms derived from Lagrangian splitting, Comput. Optim. Appl. 68 (2017), 363-405.
[10] J. Eckstein, W. Yao, Relative-error approximate versions of Douglas-Rachford splitting and special cases of the ADMM, Math. Program. 170 (2018), 417-444.
[11] J. Xie, A. Liao, X. Yang, An inexact alternating direction method of multipliers with relative error criteria, Optim. Lett. 11 (2017), 583-596.
[12] J. Douglas, H.H. Rachford, On the numerical solution of heat conduction problems in two and three space variables, Trans. Amer. Math. Soc. 82 (1956), 421-439.
[13] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Math. Program. 55 (1992), 293-318.
[14] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM J. Numer. Anal. 16 (1979), 964-979.
[15] D.W. Peaceman, H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, J. Soc. Ind. Appl. Math. 3 (1955), 28-41.
[16] Y. Gu, B. Jiang, D.R. Han, A semi-proximal-based strictly contractive Peaceman-Rachford splitting method, ArXiv: 1506.02221, 2015.
[17] M. Li, X. Yuan, A strictly contractive Peaceman-Rachford splitting method with logarithmic-quadratic proximal regularization for convex programming, Math. Oper. Res. 40 (2015), 842-858.

[18] C. Chen, R.H. Chan, S. Ma, J. Yang, Inertial proximal ADMM for linearly constrained separable convex optimization, SIAM J. Imaging Sci. 8 (2015), 2239-2267.
[19] M. Dou, H. Li, X. Liu, An inertial proximal Peaceman-Rachford splitting method, Sci. China Math. 47 (2017), 333-348.
[20] A. Auslender, M. Teboulle, S. Ben-Tiba, Interior proximal and multiplier methods based on second order homogenouskernels, Math. Oper. Res. 24 (1999), 646-668.
[21] M. Fazel, P.K. Pong, D. Sun, P. Tseng, Hankel matrix rank minimization with applications to system identification and realization, SIAM J. Matrix Anal. Appl. 34 (2013), 946-977.
[22] M. Xu, T. Wu, A class of linearized proximal alternating direction methods, J. Optim. Theory Appl. 151 (2011), 321-337.
[23] C. Chen, S. Ma, J. Yang, A general inertial proximal point algorithm for mixed variational inequality problem, SIAM J. Optim. 25 (2015), 2120-2142.