

Inertial proximal strictly contractive Peaceman–Rachford splitting method with an indefinite term for convex optimization

Zhao Deng*, Sanyang Liu

School of Mathematics and Statistics, Xidian University, Xi'an 710071, PR China



ARTICLE INFO

Article history:

Received 17 May 2019

Received in revised form 19 January 2020

Keywords:

Convex programming

Peaceman–Rachford splitting method

Inertial proximal point

Indefinite

Variational inequality

Global convergence

ABSTRACT

We consider an inertial proximal strictly contractive Peaceman–Rachford splitting method (abbreviated as IPSCPRSM) with an indefinite proximal term for solving convex optimization problems with linear constraints. With the aid of variational inequality, proximal point method and fundamental inequality, we prove global convergence of the proposed method and analyze iteration complexity in the best function value and feasibility residues. The experimental results demonstrate the efficiency of the inertial extrapolation step and the indefinite proximal term even compared with the state-of-the-art methods.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

We consider the following separable convex programming model

$$\min\{\theta_1(x_1) + \theta_2(x_2) \mid A_1x_1 + A_2x_2 = b, x_i \in \mathcal{X}_i, i = 1, 2\}, \quad (1.1)$$

where $\theta_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$ ($i = 1, 2$) are closed convex (possibly nonsmooth) functions, $\mathcal{X}_i \subset \mathbb{R}^{n_i}$ ($i = 1, 2$) are nonempty closed convex sets, $A_i \in \mathbb{R}^{l \times n_i}$ are linear operators, and $b \in \mathbb{R}^l$ is a given vector. Problem (1.1) arises from statistical learning, image processing, matrix decomposition [1–6] and many other applications.

The augmented Lagrangian function of (1.1) is

$$\mathcal{L}_\beta(x_1, x_2, \lambda) = \theta_1(x_1) + \theta_2(x_2) - \lambda^T(A_1x_1 + A_2x_2 - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2 - b\|^2, \quad (1.2)$$

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

Some operator splitting methods have been developed for solving the dual problem of (1.1) such as the Douglas–Rachford splitting method (DRSM) [7,8], the Peaceman–Rachford splitting method (PRSM) [9], the forward–backward splitting method [10], and so forth. As we all know, a benchmark solver for problem (1.1) is the alternating direction method of multipliers (ADMM) [11,12], which is equivalent to the DRSM [13].

Unfortunately, a counterexample is given in [14] that the PRSM is not necessarily convergent without additional assumptions. To obtain convergence, He et al. [15] suggested to attach a small relaxation factor α to the iteration of λ so

* Corresponding author.

E-mail address: dengzhao2016@163.com (Z. Deng).

that the strictly contractive property is guaranteed. The iterative scheme of the strictly contractive Peaceman–Rachford splitting method (SCPRSM) is

$$\begin{cases} x_1^{k+1} = \arg \min_{x_1 \in \mathcal{X}_1} \mathcal{L}_\beta(x_1, x_2^k, \lambda^k), \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b), \\ x_2^{k+1} = \arg \min_{x_2 \in \mathcal{X}_2} \mathcal{L}_\beta(x_1^{k+1}, x_2, \lambda^{k+\frac{1}{2}}), \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b), \end{cases} \quad (1.3)$$

where $\alpha \in (0, 1)$ is a step size. Its global convergence and convergence rate were derived in both ergodic and nonergodic senses.

Many practical problems show that often one of the subproblems is expensive to solve. Accordingly, an effective acceleration method is to linearize the quadratic term in the subproblem without calculating $(A_2^T A_2)^{-1}$, especially when A_2 has a high-dimensional column number, see [16–18]. This viewpoint is in fact the motivation to derive the optimal lower bound of the linearization parameter. Li et al. [19] presented a majorized ADMM with indefinite proximal terms and step size in $(0, (1 + \sqrt{5})/2)$. He et al. [20] first proved that the optimal lower bound of the linearized ADMM is $3/4$ by a counterexample. Recently, Gao and Ma [21] considered a linearized symmetric ADMM with indefinite proximal regularization as below

$$\begin{cases} x_1^{k+1} = \arg \min_{x_1 \in \mathcal{X}_1} \mathcal{L}_\beta(x_1, x_2^k, \lambda^k), \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b), \\ x_2^{k+1} = \arg \min_{x_2 \in \mathcal{X}_2} \{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^T A_2 x_2 + \frac{\beta}{2} \|A_1 x_1^{k+1} + A_2 x_2 - b\|^2 \\ + \frac{1}{2} \|x_2 - x_2^k\|_D^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b), \end{cases} \quad (1.4)$$

where $D = \tau r_2 I_{n_2} - \beta A_2^T A_2$, $r_2 > \beta \|A_2^T A_2\|$ and $\tau \in [(\alpha^2 - \alpha + 4)/(\alpha^2 - 2\alpha + 5), 1)$. The solution of x_2 -subproblem can be obtained via

$$x_2^{k+1} = \arg \min_{x_2 \in \mathcal{X}_2} \{\theta_2(x_2) + \frac{\tau r_2}{2} \|x_2 - (x_2^k + q_k)\|^2\},$$

where $q_k = \frac{\tau r_2}{2} A_2^T [\lambda^{k+\frac{1}{2}} - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b)]$.

Gao and Ma [21] also pointed out that the proximal term can be extended to indefinite and it is beneficial to get an improved numerical performance. Immediately, Jiang et al. [22] reduced the lower bound of parameter τ to $(3 + \alpha)/4$. For more relevant variants of the improved ADMM with indefinite proximal regularization, we refer the readers to [23–25].

The inertial technique comes from dynamical system and can be traced back to the heavy ball method in [26]. The inertial technique is a linear combination and an extrapolation at the current point along the direction of the previous step. It is designed to speed up the convergence properties in [27–30] and efficiently accelerate the ADMM-based algorithm in [31–34].

Based on the work of Dou et al. [33], we propose an inertial proximal strictly contractive Peaceman–Rachford splitting method with an indefinite proximal term (denoted by IPSCPRSM). The iterative scheme reads as

$$\begin{cases} (\bar{x}_1^k, \bar{x}_2^k, \bar{\lambda}^k) = (x_1^k, x_2^k, \lambda^k) + \rho_k(x_1^k - x_1^{k-1}, x_2^k - x_2^{k-1}, \lambda^k - \lambda^{k-1}), \\ x_1^{k+1} = \arg \min_{x_1 \in \mathcal{X}_1} \{\theta_1(x_1) - (\bar{\lambda}^k)^T A_1 x_1 + \frac{\beta}{2} \|A_1 x_1 + A_2 \bar{x}_2^k - b\|^2 \\ + \frac{1}{2} \|x_1 - \bar{x}_1^k\|_C^2\}, \\ \lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha\beta(A_1 x_1^{k+1} + A_2 \bar{x}_2^k - b), \\ x_2^{k+1} = \arg \min_{x_2 \in \mathcal{X}_2} \{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^T A_2 x_2 + \frac{\beta}{2} \|A_1 x_1^{k+1} + A_2 x_2 - b\|^2 \\ + \frac{1}{2} \|x_2 - \bar{x}_2^k\|_D^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b), \end{cases} \quad (1.5)$$

where $\rho_k \in [0, 1/3)$ is an inertial parameter, $\alpha \in (0, 1)$ is a step size, $\beta > 0$ is a penalty parameter and

$$C = r_1 I_{n_1} - \beta A_1^T A_1, \quad D = \tau r_2 I_{n_2} - \beta A_2^T A_2, \quad \text{with } r_1 \geq \beta \|A_1^T A_1\|, \quad r_2 > \beta \|A_2^T A_2\|. \quad (1.6)$$

In the above, $r_1 \geq \beta \|A_1^T A_1\|$ implies the positive semi-definiteness of matrix C and $\tau < 1$ indicates the indefiniteness of matrix D .

Remark 1.1. In IPSCPRSM (1.5), the parameter ρ_k can be an adjusted adaptively number in the interval $[0, 1/3)$, for instance $\rho_k = 1/(k + 2)$. We set it as a constant $\rho_k = 0.2$ in this paper. What we find interesting is that when $\tau \in ((1 + \alpha)/2, 1)$, the matrix D is indefinite, but matrices H (2.4) and G (2.5) still satisfy positive semi-definiteness.

Remark 1.2. Notice that the proposed IPSCPRSM (1.5) reduces to the SCPRSM presented in (1.3) when $C = 0$, $D = 0$ and $\rho_k = 0$. The work of [33] is a relaxed version of IPSCPRSM when $\tau \geq 1$. Moreover, the classic PRSM is a special case of IPSCPRSM if $\rho_k = 0$, $C = 0$, $D = 0$ and $\alpha = 1$.

In a nutshell, we prove convergent result of the introduced IPSCPRSM (1.5) and analyze iteration complexity in the best function value and feasibility residues. We present a counterexample to explain that IPSCPRSM is not necessarily convergent when $\tau \in (0, (1+\alpha)/2)$. Some numerical experiments are carried out to show the feasibility and effectiveness of the proposed method.

The rest of the paper is organized as follows. In Section 2, we summarize some useful preliminary results. Then, in Section 3, we give a prediction–correction reformulation of IPSCPRSM (1.5). The global convergence and iteration complexity in the best function value and feasibility residues are established. The optimal proximal parameter is discussed in Section 4. In Section 5, we perform some numerical experiments to illustrate the efficiency of the proposed method. Concluding remarks and future work are drawn in Section 6.

2. Preliminaries

In this section, we list some preliminaries and simple results for further analysis. First, we use \mathbb{R} to represent the usual Euclidean space, $\|\cdot\|$ and $\langle \cdot, \cdot \rangle$ denote the Euclidean norm and inner product in \mathbb{R} . If x is a vector in \mathbb{R}^n and D is a $n \times n$ symmetric matrix, we use the notation $\|x\|_D^2 := x^T D x$ to denote D-norm.

Then, we introduce several auxiliary sequences and matrices

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad w = \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix}, \quad F(w) = \begin{pmatrix} -A_1^T \lambda \\ -A_2^T \lambda \\ r \end{pmatrix}, \quad r = A_1 x_1 + A_2 x_2 - b, \quad (2.1)$$

and

$$\tilde{w}^k = \begin{pmatrix} \tilde{x}_1^k \\ \tilde{x}_2^k \\ \tilde{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \tilde{\lambda}^k - \beta(A_1 x_1^{k+1} + A_2 \tilde{x}_2^k - b) \end{pmatrix}. \quad (2.2)$$

Let

$$Q = \begin{pmatrix} C & 0 & 0 \\ 0 & \tau r_2 I_{n_2} & -\alpha A_2^T \\ 0 & -A_2 & \frac{1}{\beta} I_l \end{pmatrix}, \quad M = \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\alpha \beta A_2 & 2\alpha I_l \end{pmatrix}, \quad (2.3)$$

$$H = \begin{pmatrix} C & 0 & 0 \\ 0 & D + \frac{2-\alpha}{2} \beta A_2^T A_2 & -\frac{1}{2} A_2^T \\ 0 & -\frac{1}{2} A_2 & \frac{1}{2\alpha\beta} I_l \end{pmatrix}, \quad (2.4)$$

$$G = \begin{pmatrix} C & 0 & 0 \\ 0 & \tau r_2 I_{n_2} - \alpha \beta A_2^T A_2 & (\alpha - 1) A_2^T \\ 0 & (\alpha - 1) A_2 & \frac{2(1-\alpha)}{\beta} I_l \end{pmatrix}. \quad (2.5)$$

Defining

$$D_0 = r_2 I_{n_2} - \beta A_2^T A_2, \quad (2.6)$$

invoking (1.6), we can deduce that matrix D_0 is positive definite and

$$D = \tau D_0 - (1 - \tau) \beta A_2^T A_2. \quad (2.7)$$

2.1. Variational inequality characterization

Let the Lagrangian function of (1.1) be

$$L(x_1, x_2, \lambda) = \theta_1(x_1) + \theta_2(x_2) - \lambda^T (A_1 x_1 + A_2 x_2 - b),$$

which is defined on $\Omega := \mathcal{X}_1 \times \mathcal{X}_2 \times \mathbb{R}^l$. We call $w^* = (x_1^*, x_2^*, \lambda^*) \in \Omega$ a saddle point of the Lagrangian function if it satisfies

$$L(x_1^*, x_2^*, \lambda) \leq L(x_1^*, x_2^*, \lambda^*) \leq L_{x_i \in \mathcal{X}_i, (i=1,2)}(x_1, x_2, \lambda^*).$$

Accordingly, we deduce the following inequalities

$$\begin{cases} \theta_1(x_1) - \theta_1(x_1^*) + (x_1 - x_1^*)^T(-A_1^T\lambda^*) \geq 0, & \forall x_1 \in \mathcal{X}_1, \\ \theta_2(x_2) - \theta_2(x_2^*) + (x_2 - x_2^*)^T(-A_2^T\lambda^*) \geq 0, & \forall x_2 \in \mathcal{X}_2, \\ (\lambda - \lambda^*)^T(A_1x_1^* + A_2x_2^* - b) \geq 0, & \forall \lambda \in \mathbb{R}^l. \end{cases} \quad (2.8)$$

Invoking (2.1), the variational inequalities (2.8) can be denoted into a compact form as below

$$\text{VI}(\Omega, F, \theta) : \quad \theta(x) - \theta(x^*) + (w - w^*)^T F(w^*) \geq 0, \quad \forall w \in \Omega, \quad (2.9)$$

where

$$\theta(x) = \theta(x_1) + \theta(x_2).$$

From the definition of mapping F in (2.1), we have

$$(w_1 - w_2)^T F(w_1) \geq (w_1 - w_2)^T F(w_2), \quad \forall w \in \Omega. \quad (2.10)$$

Since the mapping F defined in (2.1) is affine with a skew-symmetric matrix, it is monotone. We denote by Ω^* the solution set of $\text{VI}(\Omega, F, \theta)$, and it is nonempty.

We summarize some properties of the matrices defined above in the lemma below.

Lemma 2.1. Suppose that $\alpha \in (0, 1)$ and $\tau \in ((1 + \alpha)/2, 1)$, then the matrices Q , M and H defined respectively in (2.3) and (2.4) satisfy

$$H \succeq 0, \quad HM = Q, \quad (2.11)$$

and

$$G \succeq 0, \quad \text{where } G = Q^T + Q - M^T H M. \quad (2.12)$$

Proof. By using $C \succeq 0$ and $r_2 > \beta \|A_2^T A_2\|$, we know that $H \succeq 0$ is guaranteed if the following matrix

$$\tilde{H} = \begin{pmatrix} C & 0 & 0 \\ 0 & (\tau - \frac{\alpha}{2})\beta A_2^T A_2 & -\frac{1}{2}A_2^T \\ 0 & -\frac{1}{2}A_2 & \frac{1}{2\alpha\beta}I_l \end{pmatrix}$$

is positive semi-definite. Note that \tilde{H} can be expressed as

$$\tilde{H} = \frac{1}{2} \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & A_2^T & 0 \\ 0 & 0 & I \end{pmatrix} \begin{pmatrix} 2C & 0 & 0 \\ 0 & (2\tau - \alpha)\beta & -1 \\ 0 & -1 & \frac{1}{\alpha\beta} \end{pmatrix} \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & I \end{pmatrix}.$$

Since $\alpha \in (0, 1)$, we only need to check

$$\begin{pmatrix} 2C & 0 & 0 \\ 0 & (2\tau - \alpha)\beta & -1 \\ 0 & -1 & \frac{1}{\alpha\beta} \end{pmatrix} \succeq \begin{pmatrix} 2C & 0 & 0 \\ 0 & (2\tau - \alpha)\beta & -1 \\ 0 & -1 & \frac{1}{\beta} \end{pmatrix} \succeq 0.$$

Recalling $\tau \in ((1 + \alpha)/2, 1)$, we obtain the positive semi-definiteness of H explicitly.

In sequel, we show that the matrix G defined in (2.5) is positive semi-definite. Similarly, by using $\tau > (1 + \alpha)/2$ and $r_2 > \beta \|A_2^T A_2\|$, we know that $G \succeq 0$ is guaranteed if the following matrix

$$\tilde{G} = \begin{pmatrix} C & 0 & 0 \\ 0 & \frac{1+\alpha}{2}\beta A_2^T A_2 - \alpha\beta A_2^T A_2 & (\alpha - 1)A_2^T \\ 0 & (\alpha - 1)A_2 & \frac{2(1-\alpha)}{\beta}I_l \end{pmatrix}$$

is positive semi-definite, and it amounts to proving

$$\tilde{G} = \begin{pmatrix} C & 0 & 0 \\ 0 & \frac{1-\alpha}{2}\beta A_2^T A_2 & (\alpha - 1)A_2^T \\ 0 & (\alpha - 1)A_2 & \frac{2(1-\alpha)}{\beta}I_l \end{pmatrix} \succeq 0.$$

Since $\alpha \in (0, 1)$, the positive semi-definiteness of matrix G is naturally established. \square

3. Convergence analysis

In this section, we revisit IPSCPRSM (1.5) from the variational inequality perspective and show that it can be rewritten as a prediction–correction framework proposed by He et al. [35–37]. Then, we prove Lemma 3.8 by using the proof method in Dou et al. [33]. Finally, we show the global convergence of IPSCPRSM (1.5).

3.1. Prediction–correction reformulation

We make the following assumption on parameter sequence $\{\rho_k\}_{k=0}^\infty$ as in [28], it is elemental for guaranteeing the convergence.

Assumption 3.1 ([28]). Let $\{\rho_k\}_{k=0}^\infty$ be chosen such that (i) for all $k \geq 0$, $0 \leq \rho_k \leq \rho$ for some $\rho \in [0, 1/3]$, (ii) the sequence $\{w^k\}$ be generated by IPSCPRSM (1.5) satisfy

$$\sum_{k=0}^{\infty} \rho_k \|w^k - w^{k-1}\|_H^2 < \infty. \quad (3.1)$$

One approach to ensure Assumption 3.1 is to choose $\rho_k = \min\{\rho, 1/(k\|w^k - w^{k-1}\|_H)^2\}$.

Assumption 3.2. Let the solution set of (1.1) be nonempty.

The auxiliary variable $\{\tilde{w}^k\}$ denotes the predicted variable and $\{w^{k+1}\}$ represents the corrected variable, respectively. Here we give the result of the prediction step of IPSCPRSM (1.5).

Lemma 3.1. Let $\{w^k\}$ be generated by IPSCPRSM (1.5), and $\{\tilde{w}^k\}$ be defined in (2.2). Then we have $\tilde{w}^k \in \Omega$ and

$$\theta(x) - \theta(\tilde{x}^k) + (w - \tilde{w}^k)^T F(\tilde{w}^k) \geq (w - \tilde{w}^k)^T Q(\tilde{w}^k - \tilde{w}^k), \quad \forall w \in \Omega, \quad (3.2)$$

where Q is defined in (2.3).

Proof. From the first-order optimality condition of the x_1 -subproblem in (1.5), we conclude that $\tilde{x}_1^k \in \mathcal{X}_1$ and

$$\theta_1(x_1) - \theta_1(\tilde{x}_1^k) + (x_1 - \tilde{x}_1^k)^T [-A_1^T \tilde{\lambda}^k + \beta A_1^T (A_1 \tilde{x}_1^k + A_2 \tilde{x}_2^k - b) + C(\tilde{x}_1^k - \tilde{x}_1^k)] \geq 0, \quad \forall x_1 \in \mathcal{X}_1.$$

By the definition of $\tilde{\lambda}^k = \bar{\lambda}^k - \beta(A_1 x_1^{k+1} + A_2 \tilde{x}_2^k - b)$ in (2.2), we obtain

$$\theta_1(x_1) - \theta_1(\tilde{x}_1^k) + (x_1 - \tilde{x}_1^k)^T [-A_1^T \tilde{\lambda}^k + C(\tilde{x}_1^k - \tilde{x}_1^k)] \geq 0, \quad \forall x_1 \in \mathcal{X}_1. \quad (3.3)$$

Analogously, combining the optimality condition and (2.1), we have $\tilde{x}_2^k \in \mathcal{X}_2$ and

$$\theta_2(x_2) - \theta_2(\tilde{x}_2^k) + (x_2 - \tilde{x}_2^k)^T [-A_2^T \tilde{\lambda}^{k+\frac{1}{2}} + \beta A_2^T \tilde{r}^k + D(\tilde{x}_2^k - \tilde{x}_2^k)] \geq 0, \quad \forall x_2 \in \mathcal{X}_2.$$

With the notation of \tilde{w}^k and scheme (1.5) at hand, we deduce that

$$\lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha(\bar{\lambda}^k - \tilde{\lambda}^k) = \tilde{\lambda}^k + (\alpha - 1)(\tilde{\lambda}^k - \bar{\lambda}^k). \quad (3.4)$$

Substituting Eq. (3.4) and $\tilde{\lambda}^k = \bar{\lambda}^k - \beta(A_1 x_1^{k+1} + A_2 \tilde{x}_2^k - b)$ into the above inequality and recalling the definition of D (see (1.6)), we obtain

$$\theta_2(x_2) - \theta_2(\tilde{x}_2^k) + (x_2 - \tilde{x}_2^k)^T [\tau r_2(\tilde{x}_2^k - \bar{x}_2^k) - A_2^T \tilde{\lambda}^k - \alpha A_2^T (\tilde{\lambda}^k - \bar{\lambda}^k)] \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \quad (3.5)$$

Finally, we can rewrite the relation $\tilde{\lambda}^k = \bar{\lambda}^k - \beta(A_1 x_1^{k+1} + A_2 \tilde{x}_2^k - b)$ in (2.2) as

$$(\lambda - \tilde{\lambda}^k)^T [\tilde{r}^k - A_2(\tilde{x}_2^k - \bar{x}_2^k) + \frac{1}{\beta}(\tilde{\lambda}^k - \bar{\lambda}^k)] \geq 0, \quad \forall \lambda \in \mathbb{R}^l. \quad (3.6)$$

Combining (3.3), (3.5) and (3.6), we can assert that

$$\theta(x) - \theta(\tilde{x}^k) + (w - \tilde{w}^k)^T \begin{pmatrix} -A_1^T \tilde{\lambda}^k \\ -A_2^T \tilde{\lambda}^k \\ \tilde{r}^k \end{pmatrix} \geq (w - \tilde{w}^k)^T Q(\tilde{w}^k - \tilde{w}^k).$$

By using the notations in (2.1), (2.2) and (2.3), we get the assertion (3.2) immediately. \square

Now, we show how to generate the corrected point w^{k+1} by using the predicted variable \tilde{w}^k and the previous point \bar{w}^k .

Lemma 3.2. Let $\{w^k\}$ be generated by IPSCPRSM (1.5), and $\{\tilde{w}^k\}$ be defined in (2.2). Then we have

$$\bar{w}^k - w^{k+1} = M(\bar{w}^k - \tilde{w}^k), \quad (3.7)$$

where M is defined in (2.3).

Proof. It follows from (2.2) that

$$A_1 \tilde{x}_1^k + A_2 \tilde{x}_2^k - b = \bar{\lambda}^k - \tilde{\lambda}^k.$$

With the notations of \tilde{w}^k , (1.5), (2.1) and (3.4) at hand, we conclude that

$$\begin{aligned} \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \alpha \beta \tilde{r}^k \\ &= \bar{\lambda}^k - 2\alpha \beta (A_1 \tilde{x}_1^k + A_2 \tilde{x}_2^k - b) - \alpha \beta A_2 (\tilde{x}_2^k - \bar{x}_2^k) \\ &= \bar{\lambda}^k - [2\alpha (\bar{\lambda}^k - \tilde{\lambda}^k) - \alpha \beta A_2 (\tilde{x}_2^k - \bar{x}_2^k)]. \end{aligned}$$

Therefore, we have the following relationship

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \begin{pmatrix} \bar{x}_1^k \\ \bar{x}_2^k \\ \bar{\lambda}^k \end{pmatrix} - \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\alpha \beta A_2 & 2\alpha I_l \end{pmatrix} \begin{pmatrix} \bar{x}_1^k - \tilde{x}_1^k \\ \bar{x}_2^k - \tilde{x}_2^k \\ \bar{\lambda}^k - \tilde{\lambda}^k \end{pmatrix},$$

which can be rewritten into a concise form

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

where M is defined in (2.3). \square

3.2. Global convergence

In this subsection, we prove the global convergence and analyze the iteration complexity of IPSCPRSM (1.5).

Lemma 3.3. Let $\{w^k\}$ be generated by IPSCPRSM (1.5), $\{\tilde{w}^k\}$ be defined in (2.2). Then for any $\tilde{w}^k \in \Omega$, we have

$$(w - \tilde{w}^k)^T Q(\bar{w}^k - \tilde{w}^k) = \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - \bar{w}^k\|_H^2) + \frac{1}{2}\|\bar{w}^k - \tilde{w}^k\|_G^2. \quad (3.8)$$

Proof. By using $Q = HM$ in (2.11) and $M(\bar{w}^k - \tilde{w}^k) = (\bar{w}^k - w^{k+1})$ (see (3.7)), it holds that

$$(w - \tilde{w}^k)^T Q(\bar{w}^k - \tilde{w}^k) = (w - \tilde{w}^k)^T HM(\bar{w}^k - \tilde{w}^k) = (w - \tilde{w}^k)^T H(\bar{w}^k - w^{k+1}).$$

For any vectors $a, b, c, d \in \mathbb{R}^t$ and H is symmetrical, it follows that

$$(a - b)^T H(c - d) = \frac{1}{2}(\|a - d\|_H^2 - \|a - c\|_H^2) + \frac{1}{2}(\|c - b\|_H^2 - \|d - b\|_H^2).$$

Applying the above identity with $a = w$, $b = \tilde{w}^k$, $c = \bar{w}^k$ and $d = w^{k+1}$ gives

$$\begin{aligned} &(w - \tilde{w}^k)^T H(\bar{w}^k - w^{k+1}) \\ &= \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - \bar{w}^k\|_H^2) + \frac{1}{2}(\|\bar{w}^k - \tilde{w}^k\|_H^2 - \|w^{k+1} - \tilde{w}^k\|_H^2). \end{aligned} \quad (3.9)$$

For the last term of (3.9), we have

$$\begin{aligned} &\|\bar{w}^k - \tilde{w}^k\|_H^2 - \|w^{k+1} - \tilde{w}^k\|_H^2 \\ &= \|\bar{w}^k - \tilde{w}^k\|_H^2 - \|(\bar{w}^k - \tilde{w}^k) - (\bar{w}^k - w^{k+1})\|_H^2 \\ &= \|\bar{w}^k - \tilde{w}^k\|_H^2 - \|(\bar{w}^k - \tilde{w}^k) - M(\bar{w}^k - \tilde{w}^k)\|_H^2 \\ &= 2(\bar{w}^k - \tilde{w}^k)^T HM(\bar{w}^k - \tilde{w}^k) - (\bar{w}^k - \tilde{w}^k)^T M^T HM(\bar{w}^k - \tilde{w}^k) \\ &= (\bar{w}^k - \tilde{w}^k)^T (Q^T + Q - M^T HM)(\bar{w}^k - \tilde{w}^k) \\ &= \|\bar{w}^k - \tilde{w}^k\|_G^2. \end{aligned} \quad (3.10)$$

Now, combining (2.5), (3.9) and (3.10), we obtain the assertion (3.8). \square

With [Lemmas 3.1](#) and [3.3](#) at hand, we can get the result below.

Lemma 3.4. Let $\{w^k\}$ be generated by IPSCPRSM (1.5) and $\{\tilde{w}^k\}$ be defined in (2.2). Then we have

$$\begin{aligned} & \theta(x) - \theta(\tilde{x}^k) + (w - \tilde{w}^k)^T F(\tilde{w}^k) \\ & \geq \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - \tilde{w}^k\|_H^2) + \frac{1}{2}\|\tilde{w}^k - \tilde{w}^k\|_G^2, \quad \forall w \in \Omega. \end{aligned} \quad (3.11)$$

By using the proof method in Gu et al. [38] (see [Lemma 3.1](#)) and Dou et al. [33] (see [Lemma 3.3](#)), we first give the following lemma.

Lemma 3.5 ([33]). Let $\{w^k\}$ be generated by IPSCPRSM (1.5), then for any $w^{k+1} \in \Omega$, we have

$$\begin{aligned} & \theta(x) - \theta(x^{k+1}) + (w - w^{k+1})^T [F(w^{k+1}) + H(w^{k+1} - \tilde{w}^k)] \\ & \geq (1 - \alpha)\beta(r^{k+1})^T A_2(\tilde{x}_2^k - x_2^{k+1}) + (1 - 2\alpha)\beta\|r^{k+1}\|^2, \end{aligned} \quad (3.12)$$

where r is defined in (2.1).

With [Lemma 3.5](#) at hand, we use the fundamental inequality to derive the next lemma.

Lemma 3.6. Let $\{w^k\}$ be generated by IPSCPRSM (1.5), then for any $w^{k+1} \in \Omega$, we have

$$-(1 - \alpha)\beta(r^{k+1})^T A_2(\tilde{x}_2^k - x_2^{k+1}) - (1 - 2\alpha)\beta\|r^{k+1}\|^2 \leq \frac{1}{2}\|w^{k+1} - \tilde{w}^k\|_{M_0}^2, \quad (3.13)$$

where

$$M_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & D + \frac{2-\alpha}{2}\beta A_2^T A_2 & -\frac{1}{2}A_2^T \\ 0 & -\frac{1}{2}A_2 & \frac{1}{2\alpha\beta}I_l \end{pmatrix}, \quad (3.14)$$

and

$$H_0 = \begin{pmatrix} \frac{2-\alpha}{2}\beta A_2^T A_2 & -\frac{1}{2}A_2^T \\ -\frac{1}{2}A_2 & \frac{1}{2\alpha\beta}I_l \end{pmatrix}. \quad (3.15)$$

Proof. From the definition of M_0 in (3.14), we have

$$\frac{1}{2}\|w^{k+1} - \tilde{w}^k\|_{M_0}^2 = \frac{1}{2}\|v^{k+1} - \tilde{v}^k\|_{H_0}^2 + \frac{1}{2}\|x_2^{k+1} - \tilde{x}_2^k\|_D^2.$$

Similarly, by the relation of H_0 in (3.15), we obtain

$$\frac{1}{2}\|v^{k+1} - \tilde{v}^k\|_{H_0}^2 = \frac{1-\alpha}{2}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 + \alpha\beta\|r^{k+1}\|^2.$$

By invoking the definition of D in (1.6) and $\tau > (1 + \alpha)/2$, we deduce that

$$\begin{aligned} & \frac{1}{2}\|w^{k+1} - \tilde{w}^k\|_{M_0}^2 \\ & = \frac{1-\alpha}{2}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 + \alpha\beta\|r^{k+1}\|^2 + \frac{1}{2}\tau r_2\|x_2^{k+1} - \tilde{x}_2^k\|^2 - \frac{1}{2}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 \\ & \geq \frac{1-\alpha}{2}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 + \alpha\beta\|r^{k+1}\|^2 + \frac{1+\alpha}{4}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 - \frac{1}{2}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2 \\ & = \alpha\beta\|r^{k+1}\|^2 + \frac{1-\alpha}{4}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2. \end{aligned} \quad (3.16)$$

Since $\alpha \in (0, 1)$, applying the inequality $a^T b \leq a^2 + \frac{1}{4}b^2$ with $a = r^{k+1}$ and $b = A_2(x_2^{k+1} - \tilde{x}_2^k)$, we conclude that

$$-(1 - \alpha)\beta(r^{k+1})^T A_2(\tilde{x}_2^k - x_2^{k+1}) \leq (1 - \alpha)\beta\|r^{k+1}\|^2 + \frac{1-\alpha}{4}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2.$$

It follows that

$$\begin{aligned} & -(1 - \alpha)\beta(r^{k+1})^T A_2(\tilde{x}_2^k - x_2^{k+1}) - (1 - 2\alpha)\beta\|r^{k+1}\|^2 \\ & \leq \alpha\beta\|r^{k+1}\|^2 + \frac{1-\alpha}{4}\beta\|A_2(x_2^{k+1} - \tilde{x}_2^k)\|^2. \end{aligned} \quad (3.17)$$

Combining (3.16) and (3.17), we get the assertion (3.13) directly. \square

By recalling the definitions of H in (2.4), M_0 in (3.14) and the positive semi-definiteness of matrix C in (1.6), we obtain $M_0 \leq H$. The next lemma follows essentially from [33] ([Lemma 3.3](#)), we write proof details similarly.

Lemma 3.7. Let $\{w^k\}$ be generated by IPSCPRSM (1.5). Then for any $w^* \in \Omega^*$, we have $\sum_{k=0}^{\infty}\|w^k - w^*\|_H^2 < \infty$.

Proof. By setting $w = w^{k+1}$ in (2.9) and $w = w^*$ in (3.12) respectively, we obtain

$$\theta(x^{k+1}) - \theta(x^*) + (w^{k+1} - w^*)^T F(w^*) \geq 0,$$

and

$$\begin{aligned} & \theta(x^*) - \theta(x^{k+1}) + (w^* - w^{k+1})^T [F(w^{k+1}) + H(w^{k+1} - \bar{w}^k)] \\ & \geq (1 - \alpha)\beta(r^{k+1})^T A_2(\bar{x}_2^k - x_2^{k+1}) + (1 - 2\alpha)\beta\|r^{k+1}\|^2. \end{aligned}$$

Summing them up and combining (2.10), we conclude that

$$\begin{aligned} & (w^{k+1} - w^*)^T H(w^{k+1} - \bar{w}^k) + (1 - \alpha)\beta(r^{k+1})^T A_2(\bar{x}_2^k - x_2^{k+1}) \\ & + (1 - 2\alpha)\beta\|r^{k+1}\|^2 \leq 0. \end{aligned} \quad (3.18)$$

By recalling (1.5), we have

$$\bar{w}^k = w^k + \rho_k(w^k - w^{k-1}), \quad (3.19)$$

it follows that

$$\begin{aligned} & (w^{k+1} - w^*)^T H(w^{k+1} - \bar{w}^k) \\ & = (w^{k+1} - w^*)^T H(w^{k+1} - w^k) - \rho_k(w^{k+1} - w^*)^T H(w^k - w^{k-1}). \end{aligned} \quad (3.20)$$

Let $\varphi_k = \frac{1}{2}\|w^k - w^*\|_H^2$, we conclude that

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

and

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

Substituting them into (3.20), it reveals that

$$\begin{aligned} & (w^{k+1} - w^*)^T H(w^{k+1} - \bar{w}^k) = \varphi_{k+1} - \varphi_k - \rho_k(\varphi_k - \varphi_{k-1}) + \frac{1}{2}\|w^{k+1} - w^k\|_H^2 \\ & - \frac{\rho_k}{2}\|w^k - w^{k-1}\|_H^2 - \rho_k(w^{k+1} - w^k)^T H(w^k - w^{k-1}). \end{aligned} \quad (3.21)$$

Combining (3.13), (3.18) and (3.21), we obtain

$$\begin{aligned} & \varphi_{k+1} - \varphi_k - \rho_k(\varphi_k - \varphi_{k-1}) \leq -\frac{1}{2}\|w^{k+1} - w^k - \rho_k(w^k - w^{k-1})\|_H^2 \\ & + \frac{1}{2}(\rho_k + \rho_k^2)\|w^k - w^{k-1}\|_H^2 + \frac{1}{2}\|w^{k+1} - \bar{w}^k\|_{M_0}^2. \end{aligned}$$

Together with the relation (3.19), we get

$$\begin{aligned} & \varphi_{k+1} - \varphi_k - \rho_k(\varphi_k - \varphi_{k-1}) \leq -\frac{1}{2}\|w^{k+1} - \bar{w}^k\|_H^2 \\ & + \frac{1}{2}(\rho_k + \rho_k^2)\|w^k - w^{k-1}\|_H^2 + \frac{1}{2}\|w^{k+1} - \bar{w}^k\|_{M_0}^2. \end{aligned}$$

Notice that $M_0 \leq H$, we conclude that

$$\varphi_{k+1} - \varphi_k - \rho_k(\varphi_k - \varphi_{k-1}) \leq \frac{\rho_k + \rho_k^2}{2}\|w^k - w^{k-1}\|_H^2,$$

since $\rho_k \in [0, 1/3]$, it implies that $(\rho_k + \rho_k^2)/2 \leq \rho_k$, then we have

$$\varphi_{k+1} - \varphi_k - \rho_k(\varphi_k - \varphi_{k-1}) \leq \rho_k\|w^k - w^{k-1}\|_H^2. \quad (3.22)$$

Let $\theta_k = \varphi_k - \varphi_{k-1}$, $\delta_k = \rho_k\|w^k - w^{k-1}\|_H^2$. The relation (3.22) indicates that $\theta_{k+1} \leq \rho_k\theta_k + \delta_k \leq \rho[\theta_k]_+ + \delta_k$, where $[\rho]_+ = \max\{\rho, 0\}$. Moreover, 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 $\theta_0 = [\theta_0]_+ = 0$, $\delta_0 = 0$. By invoking Assumption 3.1, we obtain

$$\sum_{k=0}^{\infty} [\theta_k]_+ \leq \frac{1}{1-\rho} \sum_{k=0}^{\infty} \delta_k = \frac{1}{1-\rho} \sum_{k=1}^{\infty} \delta_k < \infty. \quad (3.23)$$

Let $m_k = \varphi_k - \sum_{j=1}^k [\theta_j]_+$, combining (3.23) with $\varphi_k \geq 0$, we get a lower bound of sequence $\{m_k\}$. On the other hand

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

it means that $\{m_k\}$ is a monotonous non-increasing sequence, so $\{m_k\}$ converges and

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

thus $\{\varphi_k\}$ converges, and hence $\sum_{k=0}^{\infty} \|w^k - w^*\|_H^2 < \infty$ follows accordingly. \square

The next lemma gathers several useful facts which facilitate the convergence analysis of the proposed IPSCPRSM (1.5).

Lemma 3.8. Let $\{w^k\}$ be generated by IPSCPRSM (1.5). Then, for any $w^{k+1} \in \Omega$, we have

$$\sum_{k=0}^{\infty} \|w^{k+1} - \tilde{w}^k\|_G^2 < \infty, \quad \forall w \in \Omega. \quad (3.24)$$

Proof. Using the monotonicity of F , it follows that

$$(w - \tilde{w}^k)^T F(w) \geq (w - \tilde{w}^k)^T F(\tilde{w}^k).$$

Substituting the above inequality into (3.2), we have

$$\theta(x) - \theta(\tilde{x}^k) + (w - \tilde{w}^k)^T F(w) \geq (w - \tilde{w}^k)^T Q(\tilde{w}^k - \tilde{w}^k), \quad \forall w \in \Omega.$$

By setting $w = w^*$ and recalling (2.9), we get

$$(\tilde{w}^k - w^*)^T Q(\tilde{w}^k - \tilde{w}^k) \geq 0, \quad \forall w \in \Omega,$$

which together with (3.8) implies that

$$\|\tilde{w}^k - \tilde{w}^k\|_G^2 \leq \|\tilde{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2.$$

By the definition of \tilde{w}^k and the relation $\|a + b\|_H^2 \leq 2\|a\|_H^2 + 2\|b\|_H^2$, we have

$$\begin{aligned} \|\tilde{w}^k - \tilde{w}^k\|_G^2 &\leq \|\tilde{w}^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 \\ &\leq \|w^k - w^* + \rho_k(w^k - w^{k-1})\|_H^2 - \|w^{k+1} - w^*\|_H^2 \\ &\leq \|w^k - w^*\|_H^2 - \|w^{k+1} - w^*\|_H^2 + \|w^k - w^*\|_H^2 + 2\|\rho_k(w^k - w^{k-1})\|_H^2. \end{aligned}$$

Recalling Assumption 3.1, Lemma 3.7 and summing the inequality over $k = 0, \dots, \infty$, we deduce that

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

Note that M defined in (2.3) is nonsingular, together with (3.7), the assertion (3.24) follows directly. \square

With Lemmas 3.7 and 3.8 at hand, we are able to establish the following theorem about the feasibility of IPSCPRSM (1.5) and the convergence of objective function.

Theorem 3.1. Let $\{w^k\}$ be generated by IPSCPRSM (1.5). There hold as $k \rightarrow \infty$

- (i) $\sum_{k=1}^{\infty} \|r^k\|^2 < \infty$, and hence $\lim_{k \rightarrow \infty} \|r^k\| = 0$;
- (ii) The objective function value $\theta_1(x_1^k) + \theta_2(x_2^k)$ converges to the optimal value of (1.1);
- (iii) The sequence $\{w^k\}$ converges to a point w^∞ in Ω^* .

Proof. (i) First, by invoking (1.5) and (2.1), we have

$$\|r^{k+1}\|^2 = \|\frac{A_2}{2}(x_2^{k+1} - \tilde{x}_2^k) - \frac{1}{2\alpha\beta}(\lambda^{k+1} - \tilde{\lambda}^k)\|^2.$$

Since $G \geq 0$, $\alpha \in (0, 1)$ and $\beta > 0$, there must exist a $\kappa > 0$ such that

$$\|\frac{A_2}{2}(x_2^{k+1} - \tilde{x}_2^k) - \frac{1}{2\alpha\beta}(\lambda^{k+1} - \tilde{\lambda}^k)\|^2 \leq \kappa \|w^{k+1} - \tilde{w}^k\|_G^2.$$

Together with Lemma 3.8, we get the claim $\sum_{k=1}^{\infty} \|r^k\|^2 < \infty$ naturally.

(ii) Let $w^* = (x_1^*, x_2^*, \lambda^*) \in \Omega$. By setting $w = (x_1^k, x_2^k, \lambda^*)$ in (2.9), it holds that

$$\theta_1(x_1^k) + \theta_2(x_2^k) - \theta_1(x_1^*) - \theta_2(x_2^*) \geq (x_1^k - x_1^*)^T A_1^T \lambda^* + (x_2^k - x_2^*)^T A_2^T \lambda^* = (\lambda^*)^T r^k. \quad (3.25)$$

Together with $\lim_{k \rightarrow \infty} \|r^k\| = 0$ yields

$$\liminf_{k \rightarrow \infty} [\theta_1(x_1^k) + \theta_2(x_2^k)] \geq \theta_1(x_1^*) + \theta_2(x_2^*). \quad (3.26)$$

On the other hand, applying $w = w^*$ in (3.12), we conclude that

$$\begin{aligned} \theta_1(x_1^*) + \theta_2(x_2^*) - \theta_1(x_1^{k+1}) - \theta_2(x_2^{k+1}) &\geq (w^{k+1} - w^*)^T H(w^{k+1} - \bar{w}^k) \\ &\quad - (\lambda^*)^T r^{k+1} + (1 - \alpha)\beta(r^{k+1})^T A_2(\bar{x}_2^k - x_2^{k+1}) + (1 - 2\alpha)\beta\|r^{k+1}\|^2. \end{aligned} \quad (3.27)$$

It follows from Lemma 3.7 that $\{\|w^k - w^*\|_H\}$ is bounded. Combining Lemma 3.8 and the result of (i), it implies that

$$\lim_{k \rightarrow \infty} (w^{k+1} - w^*)^T H(w^{k+1} - \bar{w}^k) - (\lambda^*)^T r^{k+1} = 0.$$

Note that M_0 defined in (3.14) and G defined in (2.5) are positive semi-definite, with Lemma 3.8 at hand, we have $\lim_{k \rightarrow \infty} \|w^{k+1} - \bar{w}^k\|_{M_0}^2 = 0$. Recalling (3.13) and (3.27), we obtain

$$\limsup_{k \rightarrow \infty} [\theta_1(x_1^k) + \theta_2(x_2^k)] \leq \theta_1(x_1^*) + \theta_2(x_2^*). \quad (3.28)$$

Associating (3.26) with (3.28), we confirm the assertion (ii) directly.

(iii) Lemma 3.7 reveals that $\lim_{k \rightarrow \infty} \|w^k - w^*\|_H$ exists, then the sequence $\{w^k\}$ must be bounded. Combining (3.24) with $G \geq 0$, we know that

$$\lim_{k \rightarrow \infty} \|w^{k+1} - \bar{w}^k\| = 0. \quad (3.29)$$

Thus, the sequence $\{\bar{w}^k\}$ must be bounded. By the relation (3.7), we deduce that $\{\tilde{w}^k\}$ is also bounded. There exists a subsequence $\{\tilde{w}^{k_j}\}$ that converges to a cluster point w^∞ . Let the triple $(x_1^\infty, x_2^\infty, \lambda^\infty) \in \Omega$. Then, it follows from (3.2) that

$$\tilde{w}^{k_j} \in \Omega, \quad \theta(x) - \theta(\tilde{x}^{k_j}) + (w - \tilde{w}^{k_j})^T F(\tilde{w}^{k_j}) \geq (w - \tilde{w}^{k_j})^T Q(\bar{w}^{k_j} - \tilde{w}^{k_j}), \quad \forall w \in \Omega.$$

Since the matrix Q is positive semi-definite, it follows from the continuity of $\theta(x)$ and $F(w)$ that

$$w^\infty \in \Omega, \quad \theta(x) - \theta(x^\infty) + (w - w^\infty)^T F(w^\infty) \geq 0, \quad \forall w \in \Omega,$$

which means that w^∞ is a solution of $VI(\Omega, F, \theta)$ in Ω^* and the whole sequence $\{w^k\}$ converges to w^∞ . Moreover, the sequence $\{w^k\}$ has a unique cluster point [39]. \square

Inspired by the Opial's Theorem in [40] and the convergence rate analysis of proximal method of multipliers in [41], the following theorem establishes certain asymptotic complexity results of the proposed IPSCPRSM (1.5). Specifically, parts (ii) and (iii) of the theorem present asymptotic $o(1/\sqrt{k})$ iteration complexity measured by residues in primal feasibility and the best function value, respectively. We show that the IPSCPRSM (1.5) can find an approximate solution of $VI(\Omega, F, \theta)$ with an accuracy of $o(1/\sqrt{k})$.

Theorem 3.2. Let $\{w^k\}$ be generated by IPSCPRSM (1.5). There hold as $k \rightarrow \infty$

- (i) $\min_{1 \leq i \leq k} \|w^i - \bar{w}^{i-1}\|_G = o(1/\sqrt{k})$;
- (ii) $\min_{1 \leq i \leq k} \|r^i\| = o(1/\sqrt{k})$;
- (iii) $\min_{1 \leq i \leq k} |\theta_1(x_1^i) + \theta_2(x_2^i) - \theta_1(x_1^*) - \theta_2(x_2^*)| = o(1/\sqrt{k})$.

Proof. Let $1 \leq i \leq k$ be an arbitrarily given integer. By invoking (3.24) and the Cauchy principle, we conclude the assertion (i) directly. Part (ii) follows immediately by Theorem 3.1(i). Combining (3.26) with (3.27), we deduce that

$$\begin{aligned} &|\theta_1(x_1^i) + \theta_2(x_2^i) - \theta_1(x_1^*) - \theta_2(x_2^*)| \\ &\leq |(\lambda^*)^T r^i| + |(w^i - w^*)^T H(w^i - \bar{w}^{i-1})| + |(1 - \alpha)\beta(r^i)^T A_2(\bar{x}_2^{i-1} - x_2^i)| + |(1 - 2\alpha)\beta\|r^i\|^2|, \end{aligned}$$

which, together with (i), (ii) and Lemma 3.8, completes the proof of (iii). \square

4. A counterexample when $\tau < (1 + \alpha)/2$

In Section 2, we show that $\tau \in ((1 + \alpha)/2, 1)$ is sufficient to ensure the convergence of IPSCPRSM (1.5). In this section, inspired by the work of He et al. [20] and Jiang et al. [22], we show by an extremely simple example that the convergence of (1.5) is not necessary guaranteed for any $\tau \in (0, (1 + \alpha)/2)$. Hence, $(1 + \alpha)/2$ is the optimal lower bound for the linearization parameter τ .

Consider a special case of the problem (1.1) as

$$\min\{0 \cdot x_1 + 0 \cdot x_2 \mid 0 \cdot x_1 + x_2 = 0, x_1 \in \{0\}, x_2 \in \mathbb{R}\}. \quad (4.1)$$

We take $\beta = 1$ and the augmented Lagrangian function of model (4.1) yields that

$$\mathcal{L}_1(x_1, x_2, \lambda) = -\lambda^T x_2 + \frac{1}{2}x_2^2. \quad (4.2)$$

The iterative scheme of IPSCPRSM (1.5) for solving (4.1) is

$$\begin{cases} (\bar{x}_1^k, \bar{x}_2^k, \bar{\lambda}^k) = (x_1^k, x_2^k, \lambda^k) + \rho_k(x_1^k - x_1^{k-1}, x_2^k - x_2^{k-1}, \lambda^k - \lambda^{k-1}), \\ x_1^{k+1} = \arg \min \{ \mathcal{L}_1(x_1, \bar{x}_2^k, \bar{\lambda}^k) + \frac{1}{2} \|x_1 - \bar{x}_1^k\|_C^2 \mid x_1 \in \{0\} \}, \\ \lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha \bar{x}_2^k, \\ x_2^{k+1} = \arg \min \{ -x_2^T(\lambda^{k+\frac{1}{2}}) + \frac{1}{2} x_2^2 + \frac{1}{2} \|x_2 - \bar{x}_2^k\|_D^2 \mid x_2 \in \mathbb{R} \}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha x_2^{k+1}. \end{cases} \quad (4.3)$$

Since $\beta = 1$ and $A_2^T A_2 = 1$, it follows from (1.6) and (2.7) that

$$D = \tau r_2 - 1 \quad \text{and} \quad D_0 = r_2 - 1, \quad \forall r_2 > 1,$$

and the recursion (4.3) yields that

$$\begin{cases} x_1^{k+1} = 0, \\ \lambda^{k+\frac{1}{2}} = \bar{\lambda}^k - \alpha \bar{x}_2^k, \\ -\lambda^{k+\frac{1}{2}} + x_2^{k+1} + (\tau r_2 - 1)(x_2^{k+1} - \bar{x}_2^k) = 0, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha x_2^{k+1}. \end{cases} \quad (4.4)$$

For any $k > 0$, we have $x_1^k = 0$. We thus just need to study the iterative sequence $\{v^k = (x_2^k, \lambda^k)\}$. For any given $\tau < (1 + \alpha)/2$, there is an $r_2 > 1$ that makes $\tau r_2 < (1 + \alpha)/2$ hold. By setting $\xi = \tau r_2$, the iterative scheme (4.4) can be rewritten as

$$\begin{cases} x_2^{k+1} = \frac{1}{\xi} [(\xi - \alpha - 1)\bar{x}_2^k + \bar{\lambda}^k], \\ \lambda^{k+1} = \frac{1}{\xi} [(\alpha^2 + \alpha - 2\alpha\xi)\bar{x}_2^k + (\xi - \alpha)\bar{\lambda}^k], \end{cases} \quad (4.5)$$

which can be denoted by

$$v^{k+1} = P(\xi)v^k \quad \text{with} \quad P(\xi) = \frac{1}{\xi} \begin{pmatrix} \xi - \alpha - 1 & 1 \\ \alpha^2 + \alpha - 2\alpha\xi & \xi - \alpha \end{pmatrix}. \quad (4.6)$$

Let $f_1(\xi)$ and $f_2(\xi)$ be the two eigenvalues of matrix $P(\xi)$. Then we have

$$f_1(\xi) = \frac{(2\xi - 2\alpha - 1) + \sqrt{(2\alpha + 1 - 2\xi)^2 - 4\xi(\xi - 1)}}{2\xi} \quad \text{and} \quad f_2(\xi) = \frac{(2\xi - 2\alpha - 1) - \sqrt{(2\alpha + 1 - 2\xi)^2 - 4\xi(\xi - 1)}}{2\xi}.$$

For function $f_2(\xi)$, we have $f_2(\frac{1+\alpha}{2}) = -1$ and

$$f_2'(\xi) = \frac{2\alpha + 1 + \sqrt{(2\alpha + 1 - 2\xi)^2 - 4\xi(\xi - 1)}}{2\xi^2} + \frac{2\alpha}{\xi \sqrt{(2\alpha + 1 - 2\xi)^2 - 4\xi(\xi - 1)}} > 0, \quad \forall \xi \in (0, \frac{1+\alpha}{2}).$$

Therefore, we have

$$f_2(\xi) = \frac{(2\xi - 2\alpha - 1) - \sqrt{(2\alpha + 1 - 2\xi)^2 - 4\xi(\xi - 1)}}{2\xi} < -1, \quad \forall \xi \in (0, \frac{1+\alpha}{2}).$$

That is, for any $\forall \xi \in (0, (1 + \alpha)/2)$, the matrix $P(\xi)$ in (4.6) has an eigenvalue less than -1 . The iterative scheme (4.4), i.e., the application of IPSCPRSM (1.5) to the problem (4.1), is not necessarily convergent for any $\tau \in (0, (1 + \alpha)/2)$.

5. Numerical experiments

In this section, we conduct the IPSCPRSM (1.5) for solving the least absolute shrinkage and selection operator (LASSO) model in [2,5], the total variation (TV) based denoising model in [1,42] and the computed tomography (CT) problem in [33,43]. To further verify its numerical efficiency, we also compare it with the state-of-the-art methods for solving these problems. We will verify the following three statements.

(1) The proposed IPSCPRSM (1.5) is effective and efficient for problem (1.1).

(2) The convergence results are related to the dual step size, the inertial parameter and the linearization parameter.

(3) Numerically, the introduced IPSCPRSM (1.5) outperforms the inertial ADMM (IADMM) in [31], the classic ADMM (CADMM) in [2], the linearized symmetric ADMM with indefinite regularization (IDSADMM) in [21], the generalized ADMM with optimal indefinite proximal term (IPGADMM) in [22] and the generalization of linearized ADMM (GLADMM) in [44] for solving LASSO and TV models. From the result of CT problem, we can directly see the improvement of numerical performances by inertial step and indefinite linearization term.

All experiments are performed on an Intel(R) Core(TM) i5-6500 CPU@ 3.20 GHz PC with 8 GB of RAM.

Table 5.1Numerical results for LASSO on $\alpha = 0.3$.

m	n	CADMM		IDSADMM		IPGADMM		GLADMM		IADMM		IPSCPRSM	
		Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time
900	3000	34	2.11	31	2.08	30	2.01	28	2.20	24	1.61	19	1.23
1050	3500	36	3.15	32	2.93	32	2.96	29	3.24	25	2.26	20	1.78
1200	4000	32	4.33	29	3.72	28	3.44	26	3.62	23	2.62	18	2.04
1350	4500	35	5.14	31	4.74	31	4.73	28	5.01	24	3.56	19	2.76
1500	5000	29	5.20	25	4.76	25	4.67	24	5.20	20	3.66	16	2.82

5.1. LASSO model

Considering that one of the subproblems is expensive to solve, we can rewrite the LASSO model as

$$\min\{\frac{1}{2}\|x - b\|^2 + \sigma\|y\|_1 \mid x - Ay = 0, x \in \mathbb{R}^m, y \in \mathbb{R}^n\}. \quad (5.1)$$

Then applying the proposed IPSCPRSM (1.5) to (5.1), the x -subproblem can be addressed by

$$x^{k+1} = \frac{1}{r_1 + \beta}(b + \bar{\lambda}^k + \beta A \bar{y}^k + (r_1 - 1)\bar{x}^k),$$

and the y -subproblem can be solved via

$$y^{k+1} = \text{shrink}\{\bar{y}^k + \frac{q_k}{\tau r_2}, \frac{\sigma}{\tau r_2}\},$$

where $q_k = -A^T(\lambda^{k+\frac{1}{2}} - \beta(x^{k+1} - A\bar{y}^k))$ and the shrink is the soft thresholding operator (see [45]) defined as

$$(S_\delta(t))_i = (1 - \delta/|t_i|)_+ \cdot t_i, \quad i = 1, 2, \dots, m.$$

The initial point is $(y^0, \lambda^0) = (0, 0)$, $(y^{-1}, \lambda^{-1}) = (0, 0)$. Let $r_1 = 1.001$, $r_2 = \beta\|A^T A\| + 0.001$. To ensure the fairness of comparison, we compare the results when the parameters of each algorithm are optimal. The first dual step size is recorded as α , and the second dual step size is unified to γ . The parameters in each algorithm are listed as below.

CADMM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = 1.001$, $\rho = 0$, $\gamma = 1.618$.

IDSADMM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = \frac{\alpha^2 - \alpha + 4}{\alpha^2 - 2\alpha + 5} + 0.001$, $\rho = 0$, $\gamma = 1$.

IPGADMM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = \frac{3+\alpha}{4} + 0.001$, $\rho = 0$, $\gamma = 1$.

GLADMM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = \frac{4(\alpha+\gamma)^2 - 5(\alpha+\gamma) + 10}{4(\alpha+\gamma)^2 - 8(\alpha+\gamma) + 16} + 0.001$, $\rho = 0$, $\gamma = 0.4$.

IADMM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = 1.001$, $\rho = 0.3$, $\gamma = 1$.

IPSCPRSM: $\beta = 1$, $\sigma = 0.1\|A^T b\|_\infty$, $\tau = \frac{1+\alpha}{2} + 0.001$, $\rho = 0.3$, $\alpha = \gamma$.

The stop criteria are

$$\|x^{k+1} - Ay^{k+1}\| \leq \epsilon^{\text{pri}} \quad \text{and} \quad \|\beta A(y^{k+1} - y^k)\| \leq \epsilon^{\text{dual}},$$

where $\epsilon^{\text{pri}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|x^{k+1}\|, \|Ay^{k+1}\|\}$ and $\epsilon^{\text{dual}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}}\|y^{k+1}\|$, with ϵ^{abs} and ϵ^{rel} set to be 10^{-4} and 10^{-2} .

For a given dimension $m \times n$, the generation of data follows the way of [21].

```
p = 100/n;
x0 = sprandn(n, 1, p);
A = randn(m, n);
A = A*spdiags(1./sqrt(sum(A.^2)), 0, n, n);
b = A*x0 + sqrt(0.001)*randn(m, 1);
```

Table 5.1 shows the number of iterations and run time in seconds of CADMM, IDSADMM, IPGADMM, GLADMM, IADMM and IPSCPRSM (1.5) applied to problem (5.1) for matrix A of different dimensions. Bold characters represent the fewest number of iterations and time. Fig. 5.1 shows that the proposed IPSCPRSM (1.5) has great advantages over the other four algorithms for solving LASSO model (5.1).

5.2. TV denoising model

The TV model can be reshaped as

$$\min\{\frac{1}{2}\|y - b\|^2 + \eta\|x\|_1 \mid x - Dy = 0, x \in \mathbb{R}^n, y \in \mathbb{R}^n\}. \quad (5.2)$$

Applying IPSCPRSM (1.5) to (5.2), the x -subproblem indicates that

$$x^{k+1} = \text{shrink}\left\{\frac{(r_1-1)\bar{x}^k + \beta D \bar{y}^k + \bar{\lambda}^k}{\beta + r_1 - 1}, \frac{\eta}{\beta + r_1 - 1}\right\},$$

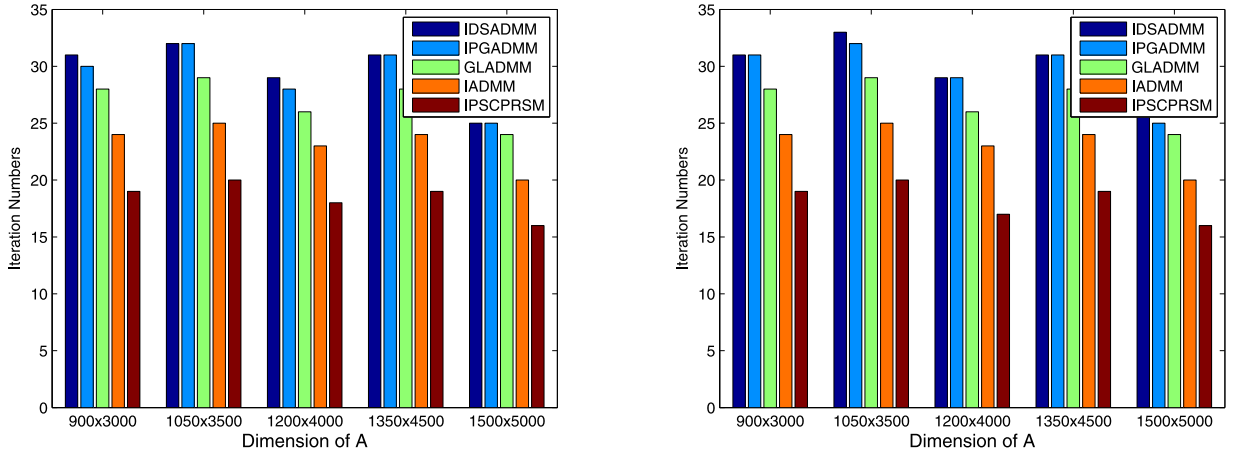


Fig. 5.1. Comparison results on $\alpha = 0.3$ and $\alpha = 0.4$ respectively for LASSO model.

Table 5.2

Numerical results for TV on $\alpha = 0.9$.

n	IADMM		CADMM		IDSADMM		IPGADMM		GLADMM		IPSCPRSM	
	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time	Iter.	Time
100	54	0.0049	50	0.0039	42	0.0033	44	0.0032	42	0.0039	37	0.0037
200	70	0.0088	64	0.0060	53	0.0040	55	0.0043	53	0.0056	46	0.0042
300	60	0.0063	55	0.0049	46	0.0042	48	0.0041	46	0.0050	40	0.0047
400	15	0.0020	21	0.0020	21	0.0022	21	0.0020	21	0.0027	19	0.0023
500	30	0.0036	29	0.0033	26	0.0028	26	0.0030	25	0.0032	22	0.0033

and the y -subproblem implies that

$$y^{k+1} = \frac{1}{1+\tau r_2}(b + \tau r_2 \tilde{y}^k + q_k).$$

where $q_k = -D^T(\lambda^{k+\frac{1}{2}} - \beta(x^{k+1} - D\tilde{y}^k))$. We set $(y^0, \lambda^0) = (0, 0)$, $(y^{-1}, \lambda^{-1}) = (0, 0)$, $r_1 = 1.001$, $r_2 = \beta\|D^T D\| + 0.001$. In addition, the parameters in each algorithm are listed as follows.

IADMM: $\beta = 1$, $\eta = 5$, $\tau = 1.001$, $\rho = 0.3$, $\gamma = 1$.

CADMM: $\beta = 1$, $\eta = 5$, $\tau = 1.001$, $\rho = 0$, $\gamma = 1.618$.

IDSADMM: $\beta = 1$, $\eta = 5$, $\tau = \frac{\alpha^2 - \alpha + 4}{\alpha^2 - 2\alpha + 5} + 0.001$, $\rho = 0$, $\gamma = 1$.

IPGADMM: $\beta = 1$, $\eta = 5$, $\tau = \frac{3+\alpha}{4} + 0.001$, $\rho = 0$, $\gamma = 1$.

GLADMM: $\beta = 1$, $\eta = 5$, $\tau = \frac{4(\alpha+\gamma)^2 - 5(\alpha+\gamma) + 10}{4(\alpha+\gamma)^2 - 8(\alpha+\gamma) + 16} + 0.001$, $\rho = 0$, $\gamma = 1.09$.

IPSCPRSM: $\beta = 1$, $\eta = 5$, $\tau = \frac{1+\alpha}{2} + 0.001$, $\rho = 0.2$, $\alpha = \gamma$.

We choose the termination criteria as

$$\|x^{k+1} - Dy^{k+1}\| \leq \epsilon^{\text{pri}} \quad \text{and} \quad \|\beta D(y^{k+1} - y^k)\| \leq \epsilon^{\text{dual}},$$

where $\epsilon^{\text{pri}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|x^{k+1}\|, \|Dy^{k+1}\|\}$ and $\epsilon^{\text{dual}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}}\|y^{k+1}\|$.

For a given dimension $n \times n$, we generate the data randomly as in [2].

```

for j = 1:3
    idx = randsample(n,1);
    k = randsample(1:10,1);
    y(ceil(idx/2):idx) = k*y(ceil(idx/2):idx);
end
b = y+randn(n,1);
e = ones(n,1);
D = spdiags([e -e],0:1,n,n);
    
```

Table 5.2 and Fig. 5.2 report the comparison between IADMM, CADMM, IDSADMM, IPGADMM, GLADMM and IPSCPRSM (1.5) for solving TV model (5.2). Calculating time is not very valuable, but the number of iterations of IPSCPRSM has obvious advantages.

To further show the high efficiency and robustness of the proposed IPSCPRSM (1.5), we demonstrate the numerical results of different α values for the LASSO and TV models, in which the α can be selected between intervals (0, 1).

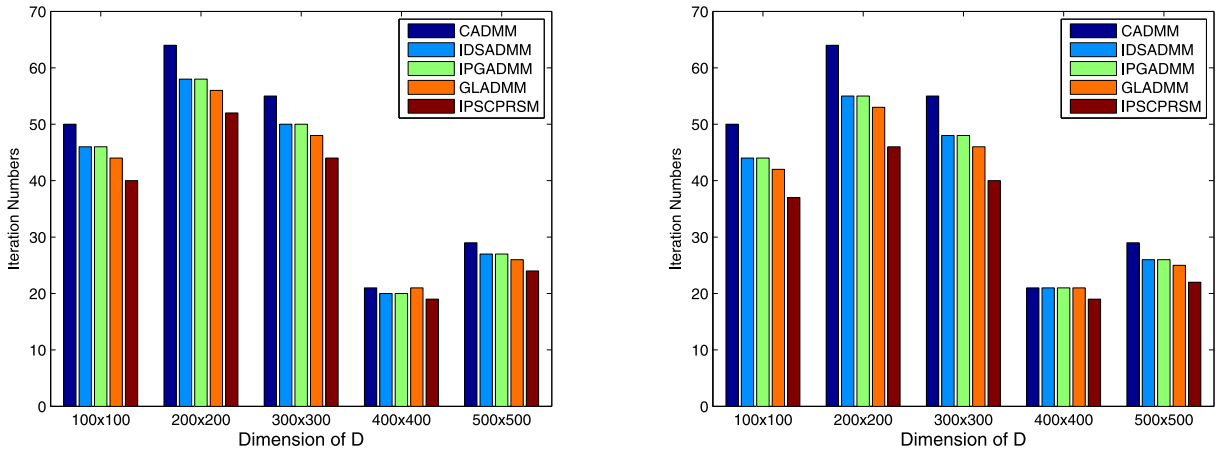


Fig. 5.2. Comparison results on $\alpha = 0.8$ and $\alpha = 0.9$ respectively for TV model.

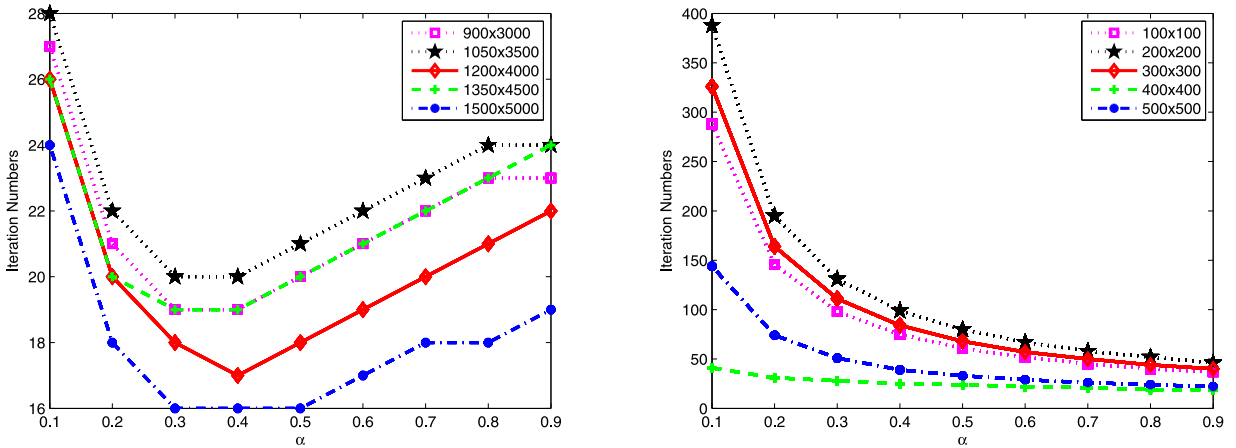


Fig. 5.3. Sensitivity tests on different α of IPSCPRSM for LASSO and TV models in different dimensions.

Fig. 5.3 shows the sensitivity tests of IPSCPRSM (1.5) for solving different dimensional LASSO and TV models with different α values. In the case of LASSO model, when α takes the value between intervals $[0.2, 0.6]$, the number of iterations under different dimensions is not more than 22, of which the minimum value is 16. In terms of TV model, the numerical results get better and better with the increase of α , especially when α takes the value in the interval $[0.7, 0.9]$.

5.3. CT problem

By introducing an auxiliary variable x , the CT model in [33,43] can be characterized as

$$\min\{\sigma\|x\|_1 + \frac{1}{2}\|Ry - y^0\|^2 \mid \nabla y - x = 0, x \in \mathbb{R}^{2n}, y \in \mathbb{R}^n\}, \quad (5.3)$$

where $\|\nabla \cdot\|$ is the total variation semi-norm, $\nabla^T = [\nabla_x^T, \nabla_y^T]^T$ is the discrete gradient operator, $\nabla_x : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\nabla_y : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are the finite difference operator in the horizontal and vertical directions respectively, $R \in \mathbb{R}^{m \times n}$ represents the Radon transform, $y \in \mathbb{R}^n$ denotes the cross-sectional image of the human body to be measured, ϵ is the noise and the X-ray image is $y^0 = Ry + \epsilon$.

By taking $C = 0$, $D = \tau r_2 I_n - R^T R - \beta \nabla^T \nabla$, $r_2 > \|R^T R + \beta \nabla^T \nabla\|$, and applying the introduced IPSCPRSM (1.5) to (5.3), we get the solution of the x -subproblem as follows

$$x^{k+1} = \text{shrink}\{\nabla \tilde{y}^k - \frac{\tilde{\lambda}^k}{\beta}, \frac{\sigma}{\beta}\}.$$

The y -subproblem can be solved by

$$y^{k+1} = \tilde{y}^k - \frac{1}{\tau r_2} R^T (R \tilde{y}^k - y^0) - \frac{1}{\tau r_2} \nabla^T (\beta \nabla \tilde{y}^k - \beta x^{k+1} - \lambda^{k+\frac{1}{2}}).$$

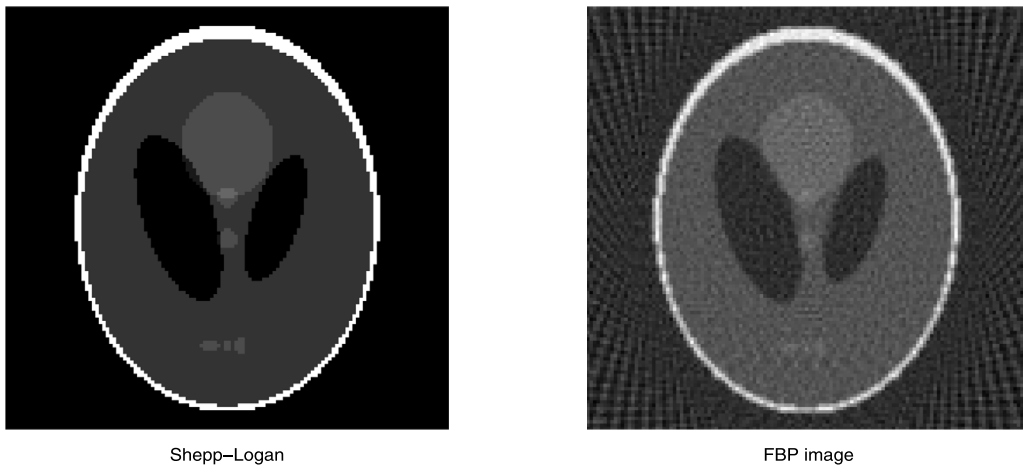


Fig. 5.4. Shepp-Logan image and its FBP image.

Table 5.3

Numerical results for FBP image restoration on $\alpha = 0.7$.

ρ	sPADMM		IsPADMM		sPSCPRSM		IsPPRSM		IPSCPRSM	
	ISNR	Time	ISNR	Time	ISNR	Time	ISNR	Time	ISNR	Time
0.1	19.92	13.19	20.09	12.79	20.20	12.27	20.31	12.81	20.45	13.67
0.2	20.15	12.92	20.49	12.66	20.44	11.99	20.66	12.70	20.76	12.93
0.3	20.19	13.56	20.65	13.53	20.48	12.97	20.75	13.66	20.81	12.86

Table 5.4

Numerical results for FBP image restoration on $\rho = 0.3$.

α	sPADMM		IsPADMM		sPSCPRSM		IsPPRSM		IPSCPRSM	
	ISNR	Time	ISNR	Time	ISNR	Time	ISNR	Time	ISNR	Time
0.6	20.25	13.47	20.70	13.33	20.50	12.92	20.79	13.44	20.87	13.34
0.7	20.24	13.06	20.67	12.54	20.51	12.08	20.76	12.69	20.82	12.71
0.8	20.17	12.69	20.63	12.80	20.49	12.13	20.75	12.48	20.79	12.75

We test the Shepp-Logan image (128×128) blurred by white Gaussian noise with standard deviation of 1. Fig. 5.4 shows the original image and the filtered back projection (FBP) image respectively.

The improved signal-to-noise ratio (ISNR) in decibel (dB) is defined by

$$\text{ISNR} = 10 * \log_{10} \frac{\|y^0 - y\|^2}{\|y - \hat{y}\|^2},$$

where y and y^0 represent the original image and the blurred image respectively, \hat{y} is the restored image.

To demonstrate the effectiveness of the introduced IPSCPRSM (1.5), we compare it with the semi-proximal ADMM (sPADMM), the inertial semi-proximal ADMM (IsPADMM), the semi-proximal PRSM (sPSCPRSM) and the inertial semi-proximal PRSM (IsPPRSM) in [33] for solving the FBP image restoration. We set $\beta = 0.1$ and $\sigma = 0.06$ for all the algorithms. Encouraged by the work of [33], we update the parameter sequence $\{\eta := \tau r_2\}$ as follows

$$\eta^{k+1} = \begin{cases} \eta^k, & \text{mod}(k, 20) \neq 0, \\ \tau \times \max\{\eta^k \times 0.6, 0.8\}, & \text{mod}(k, 20) = 0, \quad k \leq 100, \\ \min\{\eta^k \times 1.01, 5\}, & \text{mod}(k, 20) = 0, \quad k > 100, \end{cases}$$

where $\tau = (1 + \alpha)/2 + 0.001$ and $\eta^0 = 1.701$ can guarantee $r_2 > \|R^T R + \beta \nabla^T \nabla\|$.

In Table 5.3, we fix the step size $\alpha = 0.7$ to test three cases of $\rho = 0.1, 0.2$ and 0.3 . We display the ISNR values and the CPU time of sPADMM, IsPADMM, sPSCPRSM, IsPPRSM and IPSCPRSM (1.5) applied to the FBP image restoration. The results display that the inertial step is very helpful to improve the ISNR value, and the improvement by using the indefinite linearization is also significant. The ISNR level confirms that IPSCPRSM (1.5) is superior to the other four methods.

In sequel, we fix the inertial parameter $\rho = 0.3$ to test three cases of $\alpha = 0.6, 0.7$ and 0.8 . We test the ISNR history and the run time of sPADMM, IsPADMM, sPSCPRSM, IsPPRSM and IPSCPRSM (1.5) applied to the FBP image restoration. Numerical performances in Table 5.4 reveal that IPSCPRSM (1.5) is very robust and stable for solving FBP image restoration.

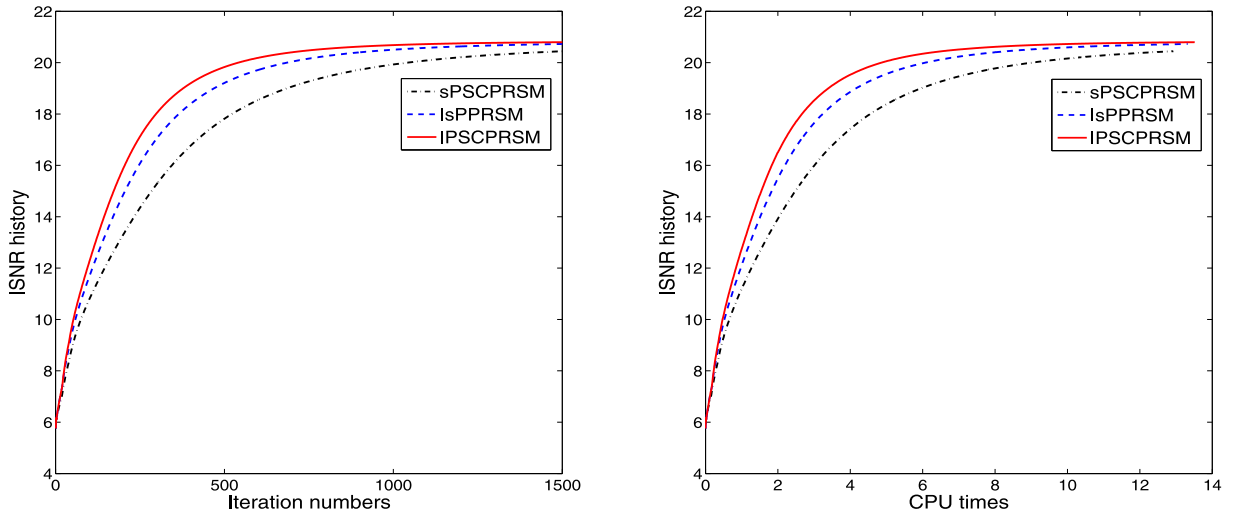


Fig. 5.5. ISNR values with respect to iteration numbers and CPU time.

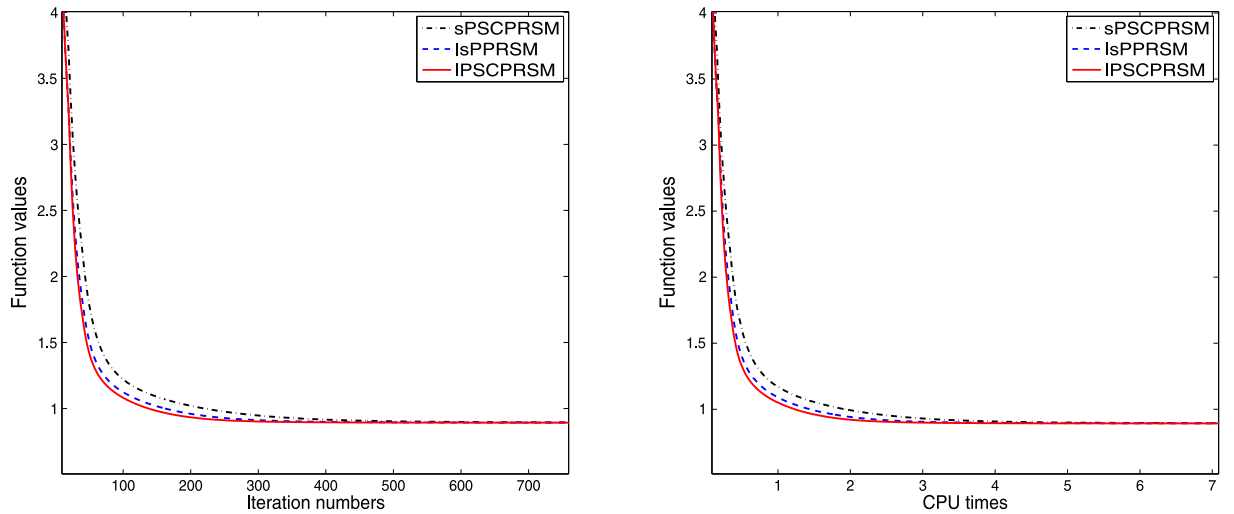


Fig. 5.6. Function values with respect to iteration numbers and CPU time.

In the following experiments, we set $\alpha = 0.7$ and $\rho = 0.28$. Fig. 5.5 reports the ISNR values with respect to iteration numbers and run time of sPSCPRSM, IsPPRSM and IPSCPRSM (1.5). Fig. 5.6 shows the function values corresponding to iteration numbers and CPU time. The results show that IPSCPRSM (1.5) has a faster increase of ISNR and a faster decrease of function value than sPSCPRSM and IsPPRSM.

The detailed restored images obtained by sPSCPRSM, IsPPRSM and IPSCPRSM are shown in Fig. 5.7. The ISNR values of restored images are 20.44 dB for sPSCPRSM, 20.73 dB for IsPPRSM and 20.80 dB for IPSCPRSM. It reveals that the ISNR value can be increased by 0.29 dB approximately by using inertial step and 0.07 dB by using indefinite linearization term.

6. Conclusion

In this paper, we propose an inertial proximal strictly contractive PRSM (short for IPSCPRSM), which unify the basic ideas of the inertial step and the indefinite linearization technique. By the aid of variational inequality, proximal point method and basic inequality, we prove global convergence of the introduced method and analyze iteration complexity in the best function value and feasibility residues. The optimal linearization parameter is proved while the convergence of the IPSCPRSM can still be ensured. Finally, some numerical results are included to illustrate the efficiency of the inertial step and the indefinite proximal term. Interesting topics for future research may include relaxing the conditions of $\{\rho^k\}_{k=0}^{\infty}$, considering two indefinite terms and better convergence rate.

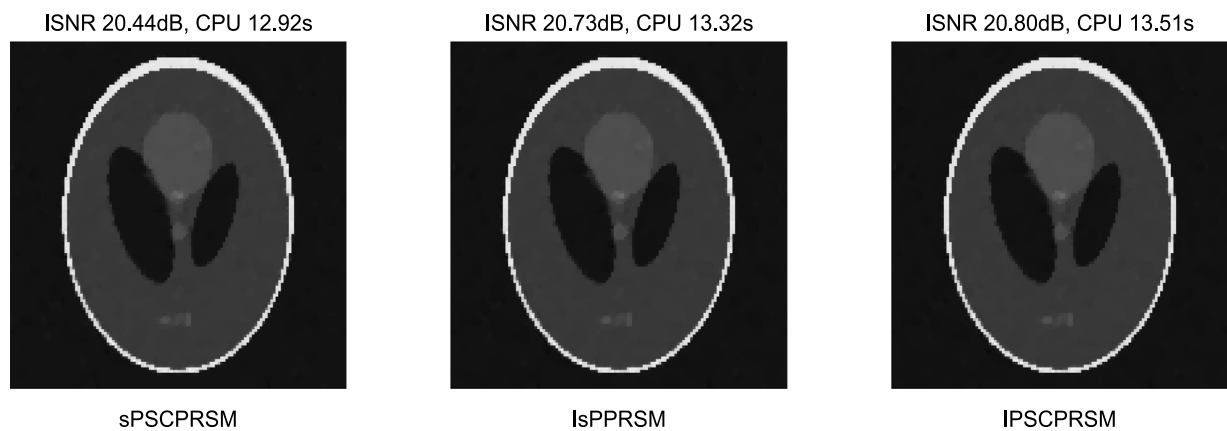


Fig. 5.7. The restored images obtained by sPSCPRSM, IsPPRSM and IPSCPRSM.

Acknowledgments

The authors are indebted to editors and anonymous referees for their valuable and constructive comments and suggestions. The authors wish to thank Prof. Xinwei Liu at Hebei University of Technology for his team's help in programming. This work was supported by the National Natural Science Foundation of China [61877046].

References

- [1] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM J. Imaging Sci.* 2 (1) (2009) 183–202, <https://academic.microsoft.com/paper/2100556411>.
- [2] S.P. 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 (1) (2011) 1–122, <https://academic.microsoft.com/paper/2164278908>.
- [3] P.L. Combettes, J.-C. Pesquet, *Proximal splitting methods in signal processing*, 2011, http://dx.doi.org/10.1007/978-1-4419-9569-8_10.
- [4] M. Fazel, T. Pong, D. Sun, P. Tseng, Hankel matrix rank minimization with applications to system identification and realization, *SIAM J. Matrix Anal. Appl.* 34 (3) (2013) 946–977, <https://academic.microsoft.com/paper/2056145308>.
- [5] R. Tibshirani, Regression shrinkage and selection via the LASSO: a retrospective, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 73 (3) (2011) 273–282, <https://academic.microsoft.com/paper/2119862467>.
- [6] A. Padcharoen, P. Kumam, J. Martínez-Moreno, Augmented Lagrangian method for TV- l_1 - l_2 based colour image restoration, *J. Comput. Appl. Math.* 354 (2019) 507–519, <https://academic.microsoft.com/paper/2904054353>.
- [7] 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, <https://academic.microsoft.com/paper/2003372231>.
- [8] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, *SIAM J. Numer. Anal.* 16 (6) (1979) 964–979, <https://academic.microsoft.com/paper/2019569173>.
- [9] D.W. Peaceman, H.H. Rachford, The numerical solution of parabolic and elliptic differential equations, *J. Soc. Ind. Appl. Math.* 3 (1) (1955) 28–41, <https://academic.microsoft.com/paper/2095036901>.
- [10] P.L. Combettes, V.R. Wajs, Signal recovery by proximal forward-backward splitting, *Multiscale Model. Simul.* 4 (4) (2006) 1168–1200, <https://academic.microsoft.com/paper/2006262045>.
- [11] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite element approximation, *Comput. Math. Appl.* 2 (1) (1976) 17–40, <https://academic.microsoft.com/paper/2045079045>.
- [12] 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, *Ann. Math. Stat.* 9 (1975) 41–76, <https://academic.microsoft.com/paper/2244987555>.
- [13] J. Eckstein, D.P. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, *Math. Program.* 55 (3) (1992) 293–318, <https://academic.microsoft.com/paper/2058532290>.
- [14] E. Corman, X. Yuan, A generalized proximal point algorithm and its convergence rate, *SIAM J. Optim.* 24 (4) (2014) 1614–1638, <https://academic.microsoft.com/paper/2027744538>.
- [15] B. He, H. Liu, Z. Wang, X. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, *SIAM J. Optim.* 24 (3) (2014) 1011–1040, <https://academic.microsoft.com/paper/2054489122>.
- [16] D. Han, X. Yuan, W. Zhang, X. Cai, An ADM-based splitting method for separable convex programming, *Comput. Optim. Appl.* 54 (2) (2013) 343–369, <https://academic.microsoft.com/paper/1982122564>.
- [17] X. Li, D. Sun, K. Toh, A Schur complement based semi-proximal ADMM for convex quadratic conic programming and extensions, *Math. Program.* 155 (1) (2016) 333–373, <https://academic.microsoft.com/paper/2093499918>.
- [18] J. Yang, X. Yuan, Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization, *Math. Comp.* 82 (281) (2012) 301–329, <https://academic.microsoft.com/paper/2079558799>.
- [19] M. Li, D. Sun, K.-C. Toh, A majorized ADMM with indefinite proximal terms for linearly constrained convex composite optimization, *SIAM J. Optim.* 26 (2) (2016) 922–950, <https://academic.microsoft.com/paper/2963647795>.
- [20] B. He, F. Ma, X. Yuan, Optimal linearized alternating direction method of multipliers for convex programming, 2017, <http://www.optimization-online.org/DBHTML/2016/07/5569.html>.
- [21] B. Gao, F. Ma, Symmetric alternating direction method with indefinite proximal regularization for linearly constrained convex optimization, *J. Optim. Theory Appl.* 176 (1) (2018) 178–204, <https://academic.microsoft.com/paper/2773075392>.

- [22] F. Jiang, Z. Wu, X. Cai, Generalized ADMM with optimal indefinite proximal term for linearly constrained convex optimization, *J. Ind. Manag. Optim.* 13 (5) (2017) 1–22, <https://academic.microsoft.com/paper/2907514291>.
- [23] J. Bai, J. Li, F. Xu, H. Zhang, Generalized symmetric ADMM for separable convex optimization, *Comput. Optim. Appl.* 70 (1) (2018) 129–170, <https://academic.microsoft.com/paper/2768736333>.
- [24] B. He, F. Ma, X. Yuan, Convergence study on the symmetric version of ADMM with larger step sizes, *SIAM J. Imaging Sci.* 9 (3) (2016) 1467–1501, <https://academic.microsoft.com/paper/2523748411>.
- [25] H. Sun, M. Tian, M. Sun, The symmetric ADMM with indefinite proximal regularization and its application, *J. Inequal. Appl.* 2017 (1) (2017) 172, <https://academic.microsoft.com/paper/2738296017>.
- [26] B. Polyak, Some methods of speeding up the convergence of iteration methods, *USSR Comput. Math. Math. Phys.* 4 (5) (1964) 1–17, <https://academic.microsoft.com/paper/1988720110>.
- [27] F. Alvarez, Weak convergence of a relaxed and inertial hybrid projection-proximal point algorithm for maximal monotone operators in Hilbert space, *SIAM J. Optim.* 14 (3) (2003) 773–782, <https://academic.microsoft.com/paper/2062184313>.
- [28] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, *Set-Valued Anal.* 9 (1–2) (2001) 3–11, <https://academic.microsoft.com/paper/199410564>.
- [29] D.A. Lorenz, T. Pock, An inertial forward-backward algorithm for monotone inclusions, *J. Math. Imaging Vision* 51 (2) (2015) 311–325, <https://academic.microsoft.com/paper/2095980290>.
- [30] A. Moudafi, M. Oliny, Convergence of a splitting inertial proximal method for monotone operators, *J. Comput. Appl. Math.* 155 (2) (2003) 447–454, <https://academic.microsoft.com/paper/1989464734>.
- [31] 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, <https://academic.microsoft.com/paper/1850691639>.
- [32] C. Chen, S. Ma, J. Yang, A general inertial proximal point algorithm for mixed variational inequality problem, *SIAM J. Optim.* 25 (2015) 2120–2142, <https://academic.microsoft.com/paper/1767939703>.
- [33] M. Dou, H. Li, X. Liu, An inertial proximal Peaceman-Rachford splitting method, *Sci. China Math.* 47 (2) (2017) 333–348, <https://academic.microsoft.com/paper/2564333494>.
- [34] Y. He, H. Li, X. Liu, Relaxed inertial proximal Peaceman-Rachford splitting method for separable convex programming, *Front. Math. China* 13 (3) (2018) 1–24, <https://academic.microsoft.com/paper/2805240903>.
- [35] B. He, M. Tao, X. Yuan, Alternating direction method with Gaussian back substitution for separable convex programming, *SIAM J. Optim.* 22 (2) (2012) 313–340, <https://academic.microsoft.com/paper/2007822845>.
- [36] B. He, X. Yuan, On the $O(1/n)$ convergence rate of the Douglas-Rachford alternating direction method, *SIAM J. Numer. Anal.* 50 (2) (2012) 700–709, <https://academic.microsoft.com/paper/1986080131>.
- [37] B. He, X. Yuan, Block-wise alternating direction method of multipliers for multiple-block convex programming and beyond, *J. Comput. Math.* 1 (2015) 145–174, <https://academic.microsoft.com/paper/2334805583>.
- [38] Y. Gu, B. Jiang, D. Han, A semi-proximal-based strictly contractive Peaceman-Rachford splitting method, 2015, arXiv preprint [arXiv:1506.02221](https://arxiv.org/abs/1506.02221), <https://academic.microsoft.com/paper/654292100>.
- [39] A. Chambolle, T. Pock, On the ergodic convergence rates of a first-order primal-dual algorithm, *Math. Program.* 159 (1–2) (2016) 253–287, <https://academic.microsoft.com/paper/2241435520>.
- [40] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.* 73 (4) (1967) 591–597, <https://academic.microsoft.com/paper/2053487511>.
- [41] R. Shefi, M. Teboulle, Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex minimization, *SIAM J. Optim.* 24 (1) (2014) 269–297, <https://academic.microsoft.com/paper/2042650576>.
- [42] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1–4) (1992) 259–268, <https://academic.microsoft.com/paper/2103559027>.
- [43] A.C. Kak, M. Slaney, G. Wang, Principles of computerized tomographic imaging, 1988, <https://academic.microsoft.com/paper/2037521346>.
- [44] X. Chang, S. Liu, P. Zhao, D. Song, A generalization of linearized alternating direction method of multipliers for solving two-block separable convex programming, *J. Comput. Appl. Math.* 357 (2019) 251–272, <https://academic.microsoft.com/paper/2919088647>.
- [45] D.L. Donoho, Y. Tsaig, Fast solution of l_1 -norm minimization problems when the solution may be sparse, *IEEE Trans. Inform. Theory* 54 (11) (2008) 4789–4812, <https://academic.microsoft.com/paper/2098012923>.