The symmetric ADMM with positive-indefinite proximal regularization and its application

Min Sun[1,2], Maoying Tian[3], Hongchun Sun[4]

1. School of Mathematics and Statistics, Zaozhuang University, Shandong, 277160, P. R. China.

2. School of Management, Qufu Normal University, Shandong, 276826, P. R. China

3. Department of Physiology, Shandong Coal Mining Health School, Taozhuang, Shandong, 277011, P. R. China

4. School of Sciences, Linyi University, Shandong, 276005, P. R. China.

Abstract. Due to update the Lagrangian multiplier twice at each iteration, the symmetric alternating direction method of multipliers (S-ADMM) often performs better than other ADMM-type methods. In practice, some proximal terms with positive definite proximal matrices are often added to its subproblems, and it is commonly known that large proximal parameter of the proximal term often results in "too-small-step-size" phenomenon. In this paper, we generalize the proximal matrix from positive definite to positive-indefinite, and give a new S-ADMM with positive-indefinite proximal regularization (IPS-ADMM) for the two-block separable convex programming with linear constraints. Without any additional assumptions, we prove the global convergence of the IPS-ADMM and analyze its convergence rate under the ergodic sense by the iteration complexity. Finally, numerical results reveal that the IPS-ADMM is more efficient than some ADMM-type methods with positive definite proximal regularization.

Keywords: Symmetric alternating direction method of multipliers; positive-indefinite proximal regularization; global convergence.

AMS subject classifications. 90C25, 94A08

1 Introduction

Let $\mathcal{R}^{n_i}$ stand for an n_i-dimensional Euclidean space, and let $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$ be nonempty, closed and convex set, where $i = 1, 2$. For two continuous closed convex functions $\theta_i(x_i) : \mathcal{R}^{n_i} \to \mathcal{R}$ $(i = 1, 2)$, the canonical two-block separable convex programming with linear equality constraints is

$$\min \left\{ \theta_1(x_1) + \theta_2(x_2) | A_1 x_1 + A_2 x_2 = b \right\}, \tag{1}$$

where $A_i \in \mathcal{R}^{m\times n_i}(i=1,2)$, $b \in \mathcal{R}^m$. Throughout, the solution set of (1) is assumed to be nonempty. Convex programming (1) has promising applicability in modeling concrete problems arising frequently in a wide range of disciplines, such as statistical learning, inverse problems and image processing, see, e.g. [1, 2, 3] for more details.

Convex programming (1) is studied extensively in the literature, and to solve it researchers have developed a group of splitting methods based on the well-known Douglas-Rachford splitting method [4, 5] and the Peachmen-Rachford splitting method [5, 6] in partial differential equation literature. Concretely, applying the Douglas-Rachford splitting method to the dual of (1) [7, 8], we get the alternating direction of multipliers (ADMM) [9, 10], whose iterative schemes reads as

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1\in\mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\}, \\ x_2^{k+1} \in \operatorname{argmin}_{x_2\in\mathcal{X}_2}\{\theta_2(x_2) - (\lambda^k)^\top(A_1x_k^{k+1} + A_2x_2 - b) + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2\}, \\ \lambda^{k+1} = \lambda^k - s\beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b), \end{cases} \tag{2}$$

where $\lambda \in \mathcal{R}^m$ is the Lagrangian multiplier; $\beta > 0$ is a penalty parameter, and $s \in (0, \frac{1+\sqrt{5}}{2})$ is a relaxation factor. Analogously, applying the Peachmen-Rachford splitting method to the dual of (1), we get the symmetric ADMM [11, 12, 13], which generates its sequence via the scheme

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1\in\mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta(A_1x_1^{k+1} + A_2x_2^k - b), \\ x_2^{k+1} \in \operatorname{argmin}_{x_2\in\mathcal{X}_2}\{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top(A_1x_k^{k+1} + A_2x_2 - b) + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b), \end{cases} \tag{3}$$

where the feasible region of r, s is

$$\mathcal{D} = \{(r,s) | r \in (-1,1), s \in (0, \frac{1+\sqrt{5}}{2})\ \&\ r+s>0, |r| < 1+s-s^2\}. \tag{4}$$

Both methods make full use of the separable structure of (1), and minimize the primal variables x_1 and x_2 individually in the Gauss-Seidel way. As elaborated in [13], the S-ADMM updates the Lagrangian multiplier twice at each iteration and thus the variables x_1, x_2 are treated in a symmetric manner. The S-ADMM includes some known ADMM-based schemes as special cases. In particular, it reduces to the original ADMM (2) when $r = 0$, and reduces to the generalized ADMM [14] when $r \in (-1, 1), s = 1$. Therefore, the S-ADMM provides a unified framework to study the ADMM-type methods. The convergence results of the S-ADMM with any $(r, s) \in \mathcal{D}$, including global convergence, the worst-case $\mathcal{O}(1/t)$ convergence rate in the ergodic sense, have been established in [13]. However, compared to the ADMM (2), the worst-case $\mathcal{O}(1/t)$ convergence rate in the non-ergodic sense of the S-ADMM hasn't been proved.

In the application, the two essential subproblems related to x_1 and x_2 dominate the computation of the S-ADMM, which are often time consuming, and the difficulty comes from the quadratic term $\frac{\beta}{2}\|A_ix_i\|^2(i = 1, 2)$. To reduce computational load, some proximal terms are often added to these two subproblems, which can linearize the quadratic term of the subproblems, and get the following proximal S-ADMM (PS-ADMM) [15, 16, 17]

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1 \in \mathcal{X}_1} \{\theta_1(x_1) - (\lambda^k)^\top (A_1 x_1 + A_2 x_2^k - b) + \frac{\beta}{2}\|A_1 x_1 + A_2 x_2^k - b\|^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta(A_1 x_1^{k+1} + A_2 x_2^k - b), \\ x_2^{k+1} \in \operatorname{argmin}_{x_2 \in \mathcal{X}_2} \{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top (A_1 x_k^{k+1} + A_2 x_2 - b) \\ \qquad\qquad + \frac{\beta}{2}\|A_1 x_1^{k+1} + A_2 x_2 - b\|^2 + \frac{1}{2}\|x_2 - x_2^k\|_G^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b), \end{cases} \tag{5}$$

where $G \in \mathcal{R}^{n_2 \times n_2}$ is a positive definite matrix. When $G = \tau I_{n_2} - \beta A_2^\top A_2$ with $\tau > \beta\|A_2^\top A_2\|$, the quadratic term $\frac{\beta}{2}\|A_2 x_2\|^2$ in the third subproblem of the PS-ADMM disappears and thus the quadratic term $\frac{\beta}{2}\|A_1 x_1^{k+1} + A_2 x_2 - b\|^2$ is linearized. Then, if $\mathcal{X}_2 = \mathcal{R}^{n_2}$, the PS-ADMM only needs to compute the proximal mapping of the involved convex function $\theta_2(\cdot)$ at each iteration, which is often simple enough to have a closed-form solution in many practical applications, such as $\theta_2(x_2) = \|x_2\|_1$ in the compressive sensing problems [3], $\theta_2(x_2) = \|x_2\|_*$ (Here x_2 is a square matrix) in the robust principal component analysis models [18].

The curse accompanying the above improvement in solvability is that the proximal (linearization) parameter τ is not easy to determine for some problems in practice. Large τ prompts the weight of the quadratic term $\frac{1}{2}\|x_2 - x_2^k\|_G^2$ in the objective function of the x_2-subproblem and inevitably results in the "too-small-step-size" phenomenon. Then, the advance of x_2 is tiny at the $k-$th iteration, which often result in the slow convergence. Since the feasible set of τ is $\{\tau | \tau > \beta\|A_2^\top A_2\|\}$, we have to compute the norm $\|A_2^\top A_2\|$, which determines the infimum of τ. Therefore, it is meaningful to reduce the infimum of τ. In fact, if we further reduce the infimum of τ, the proximal matrix G will become indefinite, and we cannot ensure the global convergence of the corresponding method. Quite recently there are some advances in this subject. For the ADMM (2) with $\gamma = 1$, He et al. [19] have proved the infimum of τ is $0.8\beta\|A_2^\top A_2\|$, and for the ADMM (2) with $\gamma \in (0, \frac{1+\sqrt{5}}{2})$, Sun et al. [20] have proved the infimum of τ is $(5 - \min\{\gamma, 1 + \gamma - \gamma^2\})\beta\|A_2^\top A_2\|/5$. Then, for the S-ADMM with $r \in (-1, 1), s = 1$, Gao et al. [21] have proved the infimum of τ is $(r^2 - r + 4)\beta\|A_2^\top A_2\|/(r^2 - 2r + 5)$. In this paper, we continue to study along this direction, and show that for any $(r, s) \in \mathcal{D}$, the global convergence of the S-ADMM with positive-indefinite proximal regularization can be guaranteed. The infimum of τ derived from this paper generalizes that in [19], and is often smaller than those in [20, 21].

The rest of the paper is organized as follows. In Sect. 2, we summarize some preliminaries which are useful for further discussion. Then, in Sect. 3, we list the iterative scheme of the IPS-ADMM and prove some important properties of the generated sequence. Its convergence results, including the global convergence and the convergence rate are proved in Sect. 4. Some preliminary numerical results are reported in Sect. 5. Finally, some concluding remarks are drawn in Sect. 6.

2 Preliminaries

In this section, we first list some notation used in this paper, and then characterize problem (1) by a mixed variational inequality problem. Some matrices and variables to simplify the notation of our later analysis are also defined.

For any two vectors $x, y \in \mathcal{R}^n$, $\langle x, y\rangle$ or $x^\top y$ denote their inner product. For any two matrices $A \in \mathcal{R}^{s\times m}, B \in \mathcal{R}^{n\times s}$, the Kronecker product of A and B is defined as $A \otimes B = (a_{ij}B)$. We let $\|\cdot\|_1$ and $\|\cdot\|$ be the ℓ_1-norm and ℓ_2-norm for vector variables, respectively. I_n denotes the $n-$dimensional identity matrix. If the matrix $G \in \mathcal{R}^{n\times n}$ is symmetric, we use the symbol $\|x\|_G^2$ to denote $x^\top Gx$ even if G is indefinite; $G \succ 0$ (resp., $G \succeq 0$) denotes that the matrix G is positive definite (resp., semi-definite).

Let us split the feasible set $\mathcal{D}$ of the parameters (r, s) into the following five subsets:

$$\begin{cases} \mathcal{D}_1 = \{(r,s)|r \in (-1,1), s \in (0,1), r+s > 0\}, \\ \mathcal{D}_2 = \{(r,s)|r \in (-1,1), s = 1\}, \\ \mathcal{D}_3 = \{(r,s)|r = 0, s \in (1, \frac{1+\sqrt{5}}{2})\}, \\ \mathcal{D}_4 = \{(r,s)|r \in (0,1), s \in (1, \frac{1+\sqrt{5}}{2}) \ \& \ r < 1+s-s^2\}, \\ \mathcal{D}_5 = \{(r,s)|r \in (-1,0), s \in (1, \frac{1+\sqrt{5}}{2}) \ \& \ -r < 1+s-s^2\}. \end{cases} \tag{6}$$

The set $\{\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3, \mathcal{D}_4, \mathcal{D}_5\}$ is a simplicial partition of $\mathcal{D}$.

Throughout, the proximal matrix G is defined by

$$G = \alpha\tau I_{n_2} - \beta A_2^\top A_2, \tag{7}$$

where $\tau > \beta\|A_2^\top A_2\|, \alpha \in (c(r,s), 1)$, and $c(r, s)$ is defined by

$$c(r,s) = \begin{cases} s + \frac{(1-s)^2}{2-r-s}, & \text{if } (r,s) \in \mathcal{D}_1 \\ \frac{4-r-r^2}{5-3r}, & \text{if } (r,s) \in \mathcal{D}_2 \\ \frac{7s^2-22s+23}{5s^2-20s+25}, & \text{if } (r,s) \in \mathcal{D}_3 \\ \frac{r^3+r^2-r-5}{3r^2-2r-5}, & \text{if } (r,s) \in \mathcal{D}_4 \\ \frac{(r^2+r-4)s^2-(r^2+4r-9)s-(r-1)^2}{s(2-s)(5-3r)}, & \text{if } (r,s) \in \mathcal{D}_5. \end{cases} \tag{8}$$

Furthermore, we define the matrix

$$G_0 = \alpha(\tau I_{n_2} - \beta A_2^\top A_2). \tag{9}$$

Obviously, $G_0 \succ 0$ by $\tau > \beta\|A_2^\top A_2\|$.

Invoking the first-order optimality condition for convex programming, we can reformulate problem (1) as the following mixed variational inequality problem (denoted by MVI$(\theta, F, \mathcal{W})$): Finding a vector $w^* \in \mathcal{W}$ such that

$$\theta(x) - \theta(x^*) + (w - w^*)^\top F(w^*) \geq 0, \quad \forall w \in \mathcal{W}, \tag{10}$$

where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad w = \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix}, \quad \theta(x) = \theta_1(x_1) + \theta_2(x_2),$$

$$F(w) = \begin{pmatrix} -A_1^\top \lambda \\ -A_2^\top \lambda \\ A_1 x_1 + A_2 x_2 - b \end{pmatrix}, \quad \mathcal{W} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{R}^m. \tag{11}$$

The mapping $F(w)$ is not only monotone, but also satisfies the property

$$w^\top (F(w) - F(\tilde{w})) = \tilde{w}^\top (F(w) - F(\tilde{w})), \quad \forall w, \tilde{w} \in \mathcal{R}^{m+n_1+n_2}. \tag{12}$$

Furthermore, the soultion set of MVI$(\theta, F, \mathcal{W})$, denote by $\mathcal{W}^*$, is nonempty under nonempty assumption onto the solution set of problem (1).

Now, let us define three matrices in order to make our proof more succinctness. Set

$$M = \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -s\beta A_2 & (r+s) I_m \end{pmatrix}, \quad Q = \begin{pmatrix} P & 0 & 0 \\ 0 & G + \beta A_2^\top A_2 & -r A_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_m \end{pmatrix}. \tag{13}$$

$$H = \begin{pmatrix} P & 0 & 0 \\ 0 & G + (1 - \frac{rs}{r+s})\beta A_2^\top A_2 & -\frac{r}{r+s} A_2^\top \\ 0 & -\frac{r}{r+s} A_2 & \frac{1}{(r+s)\beta} I_m \end{pmatrix}. \tag{14}$$

Lemma 2.1. Suppose the matrix A_2 is full column rank and the parameter α in (7) satisfies

$$\alpha > \alpha_0 \doteq \frac{rs + r^2}{r+s}. \tag{15}$$

Then, the matrices M, Q, H defined, respectively, in (6), (7) satisfies

$$HM = Q. \tag{16}$$

$$H \succ 0. \tag{17}$$

Proof. The proof of (16) is trivial, and we only need to prove (17). By the positive definiteness of P, we only need to prove $H(2:3, 2:3)$ is positive definite. Here $H(2:3, 2:3)$ denotes the corresponding sub-matrix formed from the rows and columns with the indices $(2:3)$ and $(2:3)$ as in Matlab. Substituting (7) into the right-hand side of (17), we get

$$
\begin{aligned}
H(2:3,2:3) &= \begin{pmatrix} \alpha\tau I_{n_2} - \frac{rs}{r+s}\beta A_2^\top A_2 & -\frac{r}{r+s}A_2^\top \\ -\frac{r}{r+s}A_2 & \frac{1}{(r+s)\beta}I_m \end{pmatrix} \\
&= \begin{pmatrix} \alpha(\tau I_{n_2} - \beta A_2^\top A_2) & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} (\alpha - \frac{rs}{r+s})\beta A_2^\top A_2 & -\frac{r}{r+s}A_2^\top \\ -\frac{r}{r+s}A_2 & \frac{1}{(r+s)\beta}I_m \end{pmatrix} \\
&\succeq \begin{pmatrix} (\alpha - \frac{rs}{r+s})\beta A_2^\top A_2 & -\frac{r}{r+s}A_2^\top \\ -\frac{r}{r+s}A_2 & \frac{1}{(r+s)\beta}I_m \end{pmatrix} \\
&= \frac{1}{r+s}\begin{pmatrix} A_2^\top & 0 \\ 0 & I_m \end{pmatrix}\begin{pmatrix} \beta((r+s)\alpha - rs)I_m & -rI_m \\ -rI_m & \frac{1}{\beta}I_m \end{pmatrix}\begin{pmatrix} A_2 & 0 \\ 0 & I_m \end{pmatrix},
\end{aligned}
$$

where the relationship $\succeq$ comes from $\alpha > 0$ and $\tau > \beta\|A_2^\top A_2\|$. Since the matrix A_2 is full column rank, we only need to prove the positive definiteness of the matrix

$$\begin{pmatrix} \beta((r+s)\alpha - rs)I_m & -rI_m \\ -rI_m & \frac{1}{\beta}I_m \end{pmatrix},$$

which can be further written as

$$\begin{pmatrix} \beta((r+s)\alpha - rs) & -r \\ -r & \frac{1}{\beta} \end{pmatrix} \otimes I_m,$$

where $\otimes$ denotes the matrix Kronecker product. Then, we only need to show the 2-by-2 matrix

$$\begin{pmatrix} \beta((r+s)\alpha - rs) & -r \\ -r & \frac{1}{\beta} \end{pmatrix}$$

is positive definite. In fact, by (15), we have

$$\beta((r+s)\alpha - rs) \times \frac{1}{\beta} - r^2 = (r+s)\alpha - rs - r^2 > 0.$$

Therefore, the matrix H is positive definite. The proof is completes.

At the end of this section, let us summarize two criteria to measure the worst-case $\mathcal{O}(1/t)$ convergence rate of the ADMM-type methods in an ergodic sense.

(1) For a given compact set $\bar{\mathcal{D}} \subset \mathcal{R}^{m+n}$, let $d = \sup\{\|w - w^0\| | w \in \bar{\mathcal{D}}\}$, where w^0 is the initial iterate. He et al. [22] established the following criterion

$$\sup_{w \in \bar{\mathcal{D}}} \{\theta(x_t) - \theta(x) + (w_t - w)^\top F(w)\} \le \frac{Cd^2}{t+1}, \tag{18}$$

where $w_t = \frac{1}{t+1}\sum_{k=0}^{t} w^k$, $C > 0$, and t is the iteration counter. This criterion is used in [19, 21]. Obviously, we can only ensure that any $w \in \bar{\mathcal{D}}$ satisfies (18). Therefore, the criterion (18) is not very reasonable.

(2) In [23], Lin et al. proposed the following criterion

$$\theta(x_t) - \theta(x^*) + (x_t - x^*)^\top(-A^\top\lambda^*) + \frac{c}{2}\|Ax_t - b\|^2 \le \frac{C}{t+1}, \tag{19}$$

where $c > 0$. Proposition 1 in [23] indicates that the vector $x_t \in \mathcal{X}_1 \times \mathcal{X}_2$ is an optimal solution to (1) if and only if the left-hand side of (19) equals to zero. Compared with (18), the criterion (19) is more reasonable. Therefore, we shall use a criterion similar to (19) to measure the $\mathcal{O}(1/t)$ convergence rate of our new method.

3 Algorithm

In this section, we first present the symmetric ADMM with positive-indefinite proximal regularization (denoted by IPS-ADMM), and prove some important properties of the sequence generated by the IPS-ADMM.

Algorithm 3.1 IPS-ADMM for problem (1)

Step 0 Input $\beta > 0, (r,s) \in \mathcal{D}$, the tolerance $\varepsilon > 0$, and the proximal matrices $P \in \mathcal{R}^{n_1 \times n_1}$ with $P \succ 0$ and $G \in \mathcal{R}^{n_2 \times n_2}$ defined by (7). Initialize $(x_1, x_2, \lambda) := (x_1^0, x_2^0, \lambda^0)$, and set $k := 0$.

Step 1 Compute the new iterate $w^{k+1} = (x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ by the following iterative scheme:

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1 \in \mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2 + \frac{1}{2}\|x_1 - x_1^k\|_P^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta(A_1x_1^{k+1} + A_2x_2^k - b), \\ x_2^{k+1} \in \operatorname{argmin}_{x_2 \in \mathcal{X}_2}\{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top(A_1x_k^{k+1} + A_2x_2 - b) \\ \qquad\qquad + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2 + \frac{1}{2}\|x_2 - x_2^k\|_G^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b), \end{cases} \tag{20}$$

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

To prove the convergence results of the IPS-ADMM, we first define and one block matrix two vectors.

$$A = (A_1, A_2), \quad \tilde{w}^k = \begin{pmatrix} \tilde{x}_1^k \\ \tilde{x}_2^k \\ \tilde{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^k - \beta(A_1x_1^{k+1} + A_2x_2^k - b) \end{pmatrix}.$$

Now, we give an important lemma.

Lemma 3.1 For the sequence $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ generated by the IPS-ADMM, we have

$$\theta(x) - \theta(x^{k+1}) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \ge (w - \tilde{w}^k)^\top Q(w^k - \tilde{w}^k), \quad \forall w \in \mathcal{W}, \tag{21}$$

and

$$\theta(x) - \theta(x^{k+1}) + (w - \tilde{w}^k)^\top F(w) \ge \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2) + \frac{1}{2}\|w^k - \tilde{w}^k\|_R^2, \quad \forall w \in \mathcal{W}, \tag{22}$$

where $R = Q^\top + Q - M^\top H M$.

Proof The proof of this lemma is similar to that of Lemma 3.1 and Theorem 4.2 in [13], which is omitted.

Remark 3.1 By the definition of $F(\cdot)$ in (11), and (12), for any $(x_1, x_2, \lambda) \in \mathcal{R}^{m+n_1+n_2}$ such that $A_1x_1 + A_2x_2 = b$, the left-hand side of (22) can be written as

$$\begin{aligned}
&\theta(x) - \theta(x^{k+1}) + (w - \tilde{w}^k)^\top F(w) \\
&= \theta(x) - \theta(x^{k+1}) + (w - \tilde{w}^k)^\top F(\tilde{w}^k) \\
&= \theta(x) - \theta(x^{k+1}) + (x_1 - \tilde{x}_1^k)^\top(-A_1\tilde{\lambda}^k) + (x_2 - \tilde{x}_2^k)^\top(-A_2\tilde{\lambda}^k) \\
&\quad + (\lambda - \tilde{\lambda}^k)^\top(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k - b) \\
&= \theta(x) - \theta(x^{k+1}) + (A_1\tilde{x}_1^k + A_2\tilde{x}_2^k - b)^\top\tilde{\lambda}^k + (\lambda - \tilde{\lambda}^k)^\top(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k - b) \\
&= \theta(x) - \theta(x^{k+1}) + \lambda^\top(A_1\tilde{x}_1^k + A_2\tilde{x}_2^k - b) \\
&= \theta(x) - \theta(x^{k+1}) + \lambda^\top(Ax^{k+1} - Ax) \\
&= \theta(x) - \theta(x^{k+1}) + (x - x^{k+1})^\top(-A^\top\lambda).
\end{aligned} \tag{23}$$

Then, substituting the above equality into the left-hand side of (22), we get

$$\theta(x^{k+1}) - \theta(x) + (x^{k+1} - x)^\top(-A^\top\lambda) \le \frac{1}{2}(\|w - w^k\|_H^2 - \|w - w^{k+1}\|_H^2) - \frac{1}{2}\|w^k - \tilde{w}^k\|_R^2, \tag{24}$$

where $(x_1, x_2, \lambda) \in \mathcal{R}^{m+n_1+n_2}$ satisfies $A_1x_1 + A_2x_2 = b$.

Comparing all the terms between (19) and (24), we find that the left-hand side of (24) doesn't have the term $\|Ax^{k+1} - b\|^2$ temporarily, and due to the indefinite of R, the term $\|v^k - \tilde{v}^k\|_R^2$ on the right-hand side of (24) maybe negative. Let us deal with the term $\|v^k - \tilde{v}^k\|_R^2$, and by doing so, the term $\|Ax^{k+1} - b\|^2$ will also appear. By some simple computations, we get the concrete expression of the matrix R, which is as follows.

$$R = \begin{pmatrix} P & 0 & 0 \\ 0 & G + (1-s)\beta A_2^\top A_2 & -(1-s)A_2^\top \\ 0 & -(1-s)A_2 & \frac{2-(r+s)}{\beta} I_m \end{pmatrix}. \tag{25}$$

Lemma 3.2 Let $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. Then we have

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&= \|x_1^k - x_1^{k+1}\|_P^2 + \|x_2^k - x_2^{k+1}\|_G^2 + (1-r)\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + (2 - r - s)\beta\|Ax^{k+1} - b\|^2 + 2(1-r)\beta(Ax^{k+1} - b)^\top A_2(x_2^{k+1} - x_2^k).
\end{aligned} \tag{26}$$

Proof The proof of this lemma is similar to that of Lemma 5.1 in [13], which is omitted.

The following lemma deals with the crossing term $(Ax^{k+1} - b)^\top A_2(x_2^{k+1} - x_2^k)$ on the right-hand side of (26), whose proof is mainly motivated by those of Lemma 3.2 in [24] and Lemma 5.2 in [13].

Lemma 3.3 Let $\{(x^k,\lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. Then we have

$$\begin{aligned}
&(Ax^{k+1}-b)^\top A_2(x_2^{k+1}-x_2^k)\\
&\geq \tfrac{1-s}{1+r}(Ax^k-b)^\top A_2(x_2^k-x_2^{k+1}) - \tfrac{r}{1+r}\|A_2(x_2^k-x_2^{k+1})\|^2\\
&\quad+\tfrac{\alpha}{2(1+r)\beta}(\|x_2^k-x_2^{k+1}\|_{G_0}^2-\|x_2^{k-1}-x_2^k\|_{G_0}^2)-\tfrac{1-\alpha}{2(1+r)}(3\|A_2(x_2^k-x_2^{k+1})\|^2+\|A_2(x_2^{k-1}-x_2^k)\|^2).
\end{aligned} \tag{27}$$

Proof The first-order optimality condition of x_2-subproblem in (5) indicates that, for any $x_2 \in \mathcal{X}_2$, it holds that

$$\theta_2(x_2)-\theta_2(x_2^{k+1})+(x_2-x_2^{k+1})^\top\left\{-A_2^\top\lambda^{k+\frac{1}{2}}+\beta(Ax^{k+1}-b)+G(x_2^{k+1}-x_2^k)\right\}\geq 0. \tag{28}$$

Setting $x_2 = x_2^k$ in (28), we get

$$\theta_2(x_2^k)-\theta_2(x_2^{k+1})+(x_2^k-x_2^{k+1})^\top\left\{-A_2^\top\lambda^{k+\frac{1}{2}}+\beta(Ax^{k+1}-b)+G(x_2^{k+1}-x_2^k)\right\}\geq 0.$$

Similarly, taking $x_2 = x_2^{k+1}$ in (28) for $k := k-1$, we have

$$\theta_2(x_2^{k+1})-\theta_2(x_2^k)+(x_2^{k+1}-x_2^k)^\top\left\{-A_2^\top\lambda^{k-\frac{1}{2}}+\beta(Ax^k-b)+G(x_2^k-x_2^{k-1})\right\}\geq 0.$$

Then, adding the above two inequalities, we get

$$\begin{aligned}
&(x_2^k-x_2^{k+1})^\top A_2^\top\{(\lambda^{k-\frac{1}{2}}-\lambda^{k+\frac{1}{2}})-\beta(Ax^k-b)+\beta(Ax^{k+1}-b)\}\\
&\geq\|x_2^{k+1}-x_2^k\|_G^2+(x_2^k-x_2^{k+1})^\top G(x_2^k-x_2^{k-1}).
\end{aligned} \tag{29}$$

From the update formula for λ in (5), we have

$$\begin{aligned}
&\lambda^{k+\frac{1}{2}}\\
&=\lambda^k-r\beta(A_1x_1^{k+1}+A_2x_2^k-b)\\
&=\lambda^{k-\frac{1}{2}}-s\beta(A_1x_1^k+A_2x_2^k-b)-r\beta(A_1x_1^{k+1}+A_2x_2^k-b).
\end{aligned}$$

Substituting the above equality into the left-hand side of (30), we get

$$\begin{aligned}
&(x_2^k-x_2^{k+1})^\top A_2^\top\{(1+r)\beta(Ax^{k+1}-b)-(1-s)\beta(Ax^k-b)+r\beta A_2(x_2^k-x_2^{k+1})\}\\
&\geq\|x_2^{k+1}-x_2^k\|_G^2+(x_2^k-x_2^{k+1})^\top G(x_2^k-x_2^{k-1}).
\end{aligned} \tag{30}$$

By the definitions of G and G_0 (see (7) and (9)), we have

$$\begin{aligned}
&\|x_2^{k+1}-x_2^k\|_G^2+(x_2^k-x_2^{k+1})^\top G(x_2^k-x_2^{k-1})\\
&=\alpha\|x_2^{k+1}-x_2^k\|_{G_0}^2-(1-\alpha)\beta\|A_2(x_2^{k+1}-x_2^k\|^2+\alpha(x_2^k-x_2^{k+1})^\top G_0(x_2^k-x_2^{k-1})\\
&\quad-(1-\alpha)\beta(A_2x_2^k-A_2x_2^{k+1})^\top(A_2x_2^k-A_2x_2^{k-1})\\
&\geq\frac{\alpha}{2}(\|x_2^k-x_2^{k+1}\|_{G_0}^2-\|x_2^{k-1}-x_2^k\|_{G_0}^2)-\frac{(1-\alpha)\beta}{2}(3\|A_2(x_2^k-x_2^{k+1})\|^2+\|A_2(x_2^{k-1}-x_2^k)\|^2),
\end{aligned}$$

where the last inequality comes from the Cauchy-Schwartz inequality. Substituting the above inequality into the right-hand side of (30) and arranging terms, we get the assertion (28) immediately.

Then, substituting (28) into the right-hand side of (26), we get the following main theorem, which provides a lower bound of as $\|w^k - \tilde{w}^k\|_R^2$, and the lower-bound is composed of the the term $\|Ax^{k+1} - b\|^2$, some terms in the form $\|w - w^{k+1}\|^2 - \|w - w^k\|^2$, and some others.

Theorem 3.1 Let $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. Then we have

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&\geq \|x_1^k - x_1^{k+1}\|_P^2 + \alpha\|x_2^k - x_2^{k+1}\|_{G_0}^2 - (1-\alpha)\beta\|A_2(x_2^k - x_2^{k+1})\|^2 + \tfrac{(1-r)^2}{1+r}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + (2-r-s)\beta\|Ax^{k+1} - b\|^2 + \tfrac{2(1-r)(1-s)}{1+r}\beta(Ax^k - b)^\top A_2(x_2^k - x_2^{k+1}) \\
&\quad + \tfrac{(1-r)\alpha}{1+r}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad - \tfrac{(1-r)(1-\alpha)\beta}{1+r}(3\|A_2(x_2^k - x_2^{k+1})\|^2 + \|A_2(x_2^{k-1} - x_2^k)\|^2).
\end{aligned} \tag{31}$$

According to the simplicial partition $\mathcal{D}_i (i = 1, 2, \ldots, 5)$ of the set $\mathcal{D}$ in (6), the following analysis is divided into five cases.

3.1 For $(r, s) \in \mathcal{D}_1$

Theorem 3.2 For any fixed $(r, s) \in \mathcal{D}_1$, if $\alpha > \alpha_1 \doteq s + \frac{(1-s)^2}{2-r-s}$, then there are constants $C_1, C_2 > 0$ such that

$$\|w^k - \tilde{w}^k\|_R^2 \geq \|x_1^k - x_1^{k+1}\|_P^2 + C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 + C_2\beta\|Ax^{k+1} - b\|^2. \tag{32}$$

Furthermore, $\alpha_1 \in (\alpha_0, 1)$, for any $(r, s) \in \mathcal{D}_1$, where α_0 is defined in (15).

Proof Define a matrix R_0 as

$$R_0 = \begin{pmatrix} P & 0 & 0 \\ 0 & (\alpha - s)\beta A_2^\top A_2 & -(1-s)A_2^\top \\ 0 & -(1-s)A_2 & \frac{2-(r+s)}{\beta} I_m \end{pmatrix}.$$

By the expression of R in (25), we have

$$\begin{aligned}
R(2:3, 2:3) &= \begin{pmatrix} \alpha G_0 + (\alpha - s)\beta A_2^\top A_2 & -(1-s)A_2^\top \\ -(1-s)A_2 & \frac{2-(r+s)}{\beta} I_m \end{pmatrix} \\
&= \begin{pmatrix} \alpha G_0 & 0 \\ 0 & 0 \end{pmatrix} + R_0(2:3, 2:3) \\
&\succeq R_0(2:3, 2:3) \\
&= \begin{pmatrix} A_2^\top & 0 \\ 0 & I_m \end{pmatrix} \begin{pmatrix} \beta(\alpha - s)I_m & -(1-s)I_m \\ -(1-s)I_m & \frac{2-r-s}{\beta} I_m \end{pmatrix} \begin{pmatrix} A_2 & 0 \\ 0 & I_m \end{pmatrix}.
\end{aligned}$$

Now, let us verify the positive definiteness of the matrix

$$\begin{pmatrix} \beta(\alpha - s)I_m & -(1-s)I_m \\ -(1-s)I_m & \frac{2-r-s}{\beta} I_m \end{pmatrix},$$

which can be written as

$$\begin{pmatrix} \beta(\alpha - s) & -(1-s) \\ -(1-s) & \frac{2-r-s}{\beta} \end{pmatrix} \otimes I_m.$$

Obviously, when $\alpha > \alpha_1$, the above matrix is positive definite. Therefore, the matrices R and R_0 are both positive definite by the full column rank of A_2 and the positive definiteness of P. Then, we get

$$\|w^k - \tilde{w}^k\|_R^2 \geq \|w^k - \tilde{w}^k\|_{R_0}^2,$$

Following the proof of Lemma 5.4 in [13], we get the assertion (33). By the definitions of α_0 and α_1, we have

$$\alpha_1 - \alpha_0 = \frac{(1-r)^2}{2-r-s} > 0, \quad \forall (r,s) \in \mathcal{D}_1.$$

Therefore, $\alpha_1 > \alpha_0$, for any $(r,s) \in \mathcal{D}_1$. By some manipulations, we have

$$1 - \alpha_1 = \frac{(1-r)(1-s)}{2-r-s} > 0, \quad \forall (r,s) \in \mathcal{D}_1.$$

Therefore, $\alpha_1 \in (\alpha_0, 1)$, for any $(r,s) \in \mathcal{D}_1$.

3.2 For $(r,s) \in \mathcal{D}_2$

Theorem 3.3 For any $(r,s) \in \mathcal{D}_2$, if $\alpha > \alpha_2 \doteq \frac{4-r-r^2}{5-3r}$, then we have

$$\begin{aligned} &\|w^k - \tilde{w}^k\|_R^2 \\ &\geq \|x_1^k - x_1^{k+1}\|_P^2 + C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\ &\quad + C_2\beta\|Ax^{k+1} - b\|^2 + C_3(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\ &\quad + C_4\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2), \end{aligned} \tag{33}$$

where $C_i (i = 1,2,3,4)$ are four positive constants defined by

$$C_1 = \frac{(5-3r)\alpha + r^2 + r - 4}{1+r}, \quad C_2 = 1 - r, \quad C_3 = \frac{(1-r)\alpha}{1+r}, \quad C_4 = \frac{(1-r)(1-\alpha)}{1+r}.$$

Furthermore, $\alpha_2 \in (\alpha_0, 1)$, for any $(r,s) \in \mathcal{D}_2$.

Proof Setting $s = 1$ in (31), we have

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&\geq \|x_1^k - x_1^{k+1}\|_P^2 + \alpha\|x_2^k - x_2^{k+1}\|_{G_0}^2 - (1-\alpha)\beta\|A_2(x_2^k - x_2^{k+1})\|^2 + \frac{(1-r)^2}{1+r}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + (1-r)\beta\|Ax^{k+1} - b\|^2 + \frac{(1-r)\alpha}{1+r}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad - \frac{(1-r)(1-\alpha)\beta}{1+r}(3\|A_2(x_2^k - x_2^{k+1})\|^2 + \|A_2(x_2^{k-1} - x_2^k)\|^2) \\
&\geq \frac{(5-3r)\alpha + r^2 + r - 4}{1+r}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + (1-r)\beta\|Ax^{k+1} - b\|^2 + \frac{(1-r)\alpha}{1+r}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad + \frac{(1-r)(1-\alpha)\beta}{1+r}(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2),
\end{aligned}$$

which proves (33). From $\alpha > \alpha_2$, it is obvious that $C_1 > 0$, and from $r \in (-1, 1), \alpha \in (\alpha_2, 1)$, we have $C_2, C_3, C_4 > 0$. By the definition of α_2, we get

$$\alpha_2 - \alpha_0 = \frac{2(1-r)(2-r)}{5-3r} > 0, \quad \forall (r, s) \in \mathcal{D}_2.$$

Therefore, $\alpha_2 > \alpha_0$, for any $(r, s) \in \mathcal{D}_2$. By some manipulations, we have

$$1 - \alpha_2 = \frac{(r-1)^2}{5-3r} > 0, \quad \forall (r, s) \in \mathcal{D}_2.$$

Therefore, $\alpha_2 \in (\alpha_0, 1)$, for any $(r, s) \in \mathcal{D}_2$.

Remark 3.1 In [21], for any $(r, s) \in \mathcal{D}_2$, Gao et al. have prove that $\alpha_G \doteq \frac{r^2-r+4}{r^2-2r+5}$ is a lower bound of α. The curves of α_2 and α_G in $r \in [-1, 1]$ are drawn in Figure 1, which illustrates that $\alpha_2 < \alpha_G$ if $r \in (-1, 0)$, and $\alpha_2 > \alpha_G$ if $r \in (0, 1)$. When $r \in (-1, 1)$, the infimum and supremum of α_2 are 0.5 and 1, while the infimum and supremum of α_G are 0.75 and 1. Therefore, the infimum of α_2 is obviously less than that of α_G.

3.3 For $(r, s) \in \mathcal{D}_3$

Theorem 3.4 For any $(r, s) \in \mathcal{D}_3$, if $\alpha > \alpha_3 \doteq \frac{7s^2-22s+23}{5s^2-20s+25}$, then we have

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&\geq \|x_1^k - x_1^{k+1}\|_P^2 + C_0\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) + C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + C_2\beta\|Ax^{k+1} - b\|^2 + C_3(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad + C_4\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2)
\end{aligned} \tag{34}$$

where $C_i (i = 1, 2, 3, 4)$ are four positive constants defined by

$$C_0 = T_1 - s, \;\; C_1 = 1 - \frac{(1-s)^2}{T_1 - s} - 5(1-\alpha), \;\; C_2 = 2 - T_1, \;\; C_3 = \alpha, \;\; C_4 = 1 - \alpha, \;\; T_1 = \frac{1}{3}(s^2 - s + 5).$$

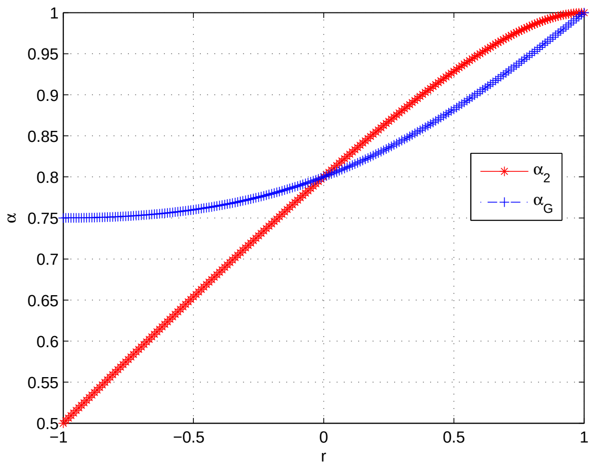

Figure 1: The curves of τ_2 and τ_G in $r \in [-1, 1]$

Furthermore, $\alpha_3 \in (\alpha_0, 1)$, for any $(r, s) \in \mathcal{D}_3$.

Proof By the inequality (5.26) in [13], we have

$$2(1-s)\beta(Ax^k - b)^\top A_2(x_2^k - x_2^{k+1}) \geq -(T_1 - s)\beta\|Ax^k - b\|^2 - \frac{(1-s)^2}{T_1 - s}\|A_2(x_2^k - x_2^{k+1})\|^2.$$

Then, substituting the above inequality into the right-hand side of (31), we get

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&\geq \|x_1^k - x_1^{k+1}\|_P^2 + (T_1 - s)\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) \\
&\quad + (1 - \frac{(1-s)^2}{T_1 - s} - 5(1-\alpha))\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\quad + (2 - T_1)\beta\|Ax^{k+1} - b\|^2 + \alpha(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad + (1-\alpha)\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2),
\end{aligned}$$

which proves (34). From (5.28) and (5.29) in [13], we have $C_0, C_2 > 0$, and from $\alpha \in (\alpha_3, 1)$, we have $C_3, C_4 > 0$. As for $C_1 > 0$, we have

$$C_1 = \frac{(5s^2 - 20s + 25)\alpha - 7s^2 + 22s - 23}{s^2 - 4s + 5} > 0, \quad \forall \alpha > \alpha_3.$$

By the definition of α_3, we get

$$\alpha_3 - \alpha_0 = \alpha_3 = \frac{7s^2 - 22s + 23}{5(s^2 - 4s + 5)} > 0, \quad \forall (r, s) \in \mathcal{D}_3.$$

By some manipulations, we have

$$1 - \alpha_3 = \frac{-2(s - \frac{1-\sqrt{5}}{2})(s - \frac{1+\sqrt{5}}{2})}{5(s^2 - 4s + 5)} > 0, \quad \forall (r, s) \in \mathcal{D}_3.$$

Therefore, $\alpha_3 \in (\alpha_0, 1)$, for any $(r, s) \in \mathcal{D}_3$.

3.4 For $(r,s) \in \mathcal{D}_4$

Theorem 3.5 For any $(r,s) \in \mathcal{D}_4$, if $\alpha > \alpha_4 \doteq \frac{r^3+r^2-r-5}{3r^2-2r-5}$, then we have

$$\begin{aligned} &\|w^k - \tilde{w}^k\|_R^2 \\ &\geq \|x_1^k - x_1^{k+1}\|_P^2 + C_0\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) + C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\ &\quad + C_2\beta\|Ax^{k+1} - b\|^2 + C_3(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\ &\quad + C_4\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2) \end{aligned} \tag{35}$$

where $C_i (i = 1,2,3,4)$ are four positive constants defined by

$$C_0 = T_2 - (r+s), \quad C_1 = r\frac{(1-r)^2}{(1+r)^2} - 4(1-\alpha)\frac{1-r}{1+r} + (\alpha - 1),$$

$$C_2 = 2 - T_2, \quad C_3 = \alpha\frac{1-r}{1+r}, \quad C_4 = (1-\alpha)\frac{1-r}{1+r}, \quad T_2 = r + s + (1-s)^2.$$

Furthermore, $\alpha_4 \in (\alpha_0, 1)$, for any $(r,s) \in \mathcal{D}_4$.

Proof By the inequality (5.31) in [13], we have

$$\begin{aligned} &2\frac{(1-r)(1-s)}{1+r}\beta(Ax^k - b)^\top A_2(x_2^k - x_2^{k+1}) \\ &\geq -[T_2 - (r+s)]\beta\|Ax^k - b\|^2 - \frac{(1-r)^2(1-s)^2}{(1+r)^2[T_2 - (r+s)]}\|A_2(x_2^k - x_2^{k+1})\|^2. \end{aligned}$$

Then, substituting the above inequality into the right-hand side of (31), we get

$$\begin{aligned} &\|w^k - \tilde{w}^k\|_R^2 \\ &\geq \|x_1^k - x_1^{k+1}\|_P^2 + [T_2 - (r+s)]\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) \\ &\quad + (r\frac{(1-r)^2}{(1+r)^2} - 4(1-\alpha)\frac{1-r}{1+r} + (\alpha - 1))\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\ &\quad + (2 - T_2)\beta\|Ax^{k+1} - b\|^2 + \alpha\frac{1-r}{1+r}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\ &\quad + (1-\alpha)\frac{1-r}{1+r}\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2), \end{aligned}$$

which proves (35). From the definition of T_2, $\alpha \in (\alpha_4, 1)$, $(r,s) \in \mathcal{D}_4$, it is easy to verify that $C_0, C_2, C_3, C_4 > 0$. From the definition of C_1, we get

$$C_1 = \frac{(r+1)(5-3r)\alpha + r^3 + r^2 - r - 5}{(r+1)^2} > 0, \quad \forall \alpha > \alpha_4.$$

By the definition of α_4, we get

$$\alpha_4 - \alpha_0 = \frac{(1-r)[2(1-r^2) + r + 3]}{(1+r)(5-3r)} > 0, \quad \forall (r,s) \in \mathcal{D}_4.$$

By some manipulations, we have

$$1 - \alpha_4 = \frac{r(r-1)^2}{(1+r)(5-3r)} > 0, \quad \forall (r,s) \in \mathcal{D}_4.$$

Therefore, $\alpha_4 \in (\alpha_0, 1)$, for any $(r,s) \in \mathcal{D}_4$.

3.5 For $(r,s)\in\mathcal{D}_5$

Theorem 3.6 For any $(r,s)\in\mathcal{D}_5$, if $\alpha>\alpha_5\doteq\frac{(r^2+r-4)s^2-(r^2+4r-9)s-(r-1)^2}{s(2-s)(5-3r)}$, then we have

$$\begin{aligned}&\|w^k-\tilde{w}^k\|_R^2\\&\geq\|x_1^k-x_1^{k+1}\|_P^2+C_0\beta(\|Ax^{k+1}-b\|^2-\|Ax^k-b\|^2)+C_1\beta\|A_2(x_2^k-x_2^{k+1})\|^2\\&\quad+C_2\beta\|Ax^{k+1}-b\|^2+C_3(\|x_2^k-x_2^{k+1}\|_{G_0}^2-\|x_2^{k-1}-x_2^k\|_{G_0}^2)\\&\quad+C_4\beta(\|A_2(x_2^k-x_2^{k+1})\|^2-\|A_2(x_2^{k-1}-x_2^k)\|^2)\end{aligned}\tag{36}$$

where $C_i(i=1,2,3,4)$ are four positive constants defined by

$$C_0=\frac{(s^2-s)(2-s)}{1+r},\quad C_1=\frac{(1-r)^2(1+s-s^2)}{s(1+r)(2-s)}-4(1-\alpha)\frac{1-r}{1+r}+(\alpha-1),$$

$$C_2=2-T_3,\quad C_3=\alpha\frac{1-r}{1+r},\quad C_4=(1-\alpha)\frac{1-r}{1+r},\quad T_3=r+s+\frac{(s^2-s)(2-s)}{1+r}.$$

Furthermore, $\alpha_5\in(\alpha_0,1)$, for any $(r,s)\in\mathcal{D}_5$.

Proof By the inequality (5.34) in [13], we have

$$\begin{aligned}&2\frac{(1-r)(1-s)}{1+r}\beta(Ax^k-b)^\top A_2(x_2^k-x_2^{k+1})\\&\geq-[T_3-(r+s)]\beta\|Ax^k-b\|^2-\frac{(1-r)^2(1-s)^2}{(1+r)^2[T_3-(r+s)]}\|A_2(x_2^k-x_2^{k+1})\|^2.\end{aligned}$$

Then, substituting the above inequality into the right-hand side of (31), we get

$$\begin{aligned}&\|w^k-\tilde{w}^k\|_R^2\\&\geq\|x_1^k-x_1^{k+1}\|_P^2+[T_3-(r+s)]\beta(\|Ax^{k+1}-b\|^2-\|Ax^k-b\|^2)\\&\quad+(\frac{(1-r)^2(1+s-s^2)}{s(1+r)(2-s)}-4(1-\alpha)\frac{1-r}{1+r}+(\alpha-1))\beta\|A_2(x_2^k-x_2^{k+1})\|^2\\&\quad+(2-T_3)\beta\|Ax^{k+1}-b\|^2+\alpha\frac{1-r}{1+r}(\|x_2^k-x_2^{k+1}\|_{G_0}^2-\|x_2^{k-1}-x_2^k\|_{G_0}^2)\\&\quad+(1-\alpha)\frac{1-r}{1+r}\beta(\|A_2(x_2^k-x_2^{k+1})\|^2-\|A_2(x_2^{k-1}-x_2^k)\|^2),\end{aligned}$$

which proves (36). From the definition of T_3, $\alpha\in(\alpha_5,1)$, $(r,s)\in\mathcal{D}_5$, it is easy to verify that $C_0,C_2,C_3,C_4>0$. From the definition of C_1, for any $(r,s)\in\mathcal{D}_5$, we get

$$C_1=\frac{s(2-s)(5-3r)\alpha-(r^2+r-4)s^2+(r^2+4r-9)s+(r-1)^2}{s(2-s)(r+1)}>0,\quad\forall\alpha>\alpha_5.$$

By the definition of α_5, for any $(r,s)\in\mathcal{D}_5$, we have

$$\begin{aligned}&\alpha_5-\alpha_0\\&=\frac{(1-r)[2(r-2)s^2+(9-5r)s+r-1]}{s(5-3r)(2-s)}\\&\geq\frac{(1-r)[2(r-2)(1+r+s)+(9-5r)s+r-1]}{s(5-3r)(2-s)}\\&=\frac{(1-r)[5(s-1)+r(2r-3s-1)]}{s(5-3r)(2-s)}\\&>0,\end{aligned}$$

where the first inequality follows from $s^2 < 1 + r + s$, and the second inequality comes from $r < 0, s \in (1, \frac{1+\sqrt{5}}{2}), r < \frac{3s+1}{2}$. By some manipulations, we obtain

$$1 - \alpha_5 = \frac{-(r-1)^2(s - \frac{1-\sqrt{5}}{2})(s - \frac{1+\sqrt{5}}{2})}{s(5-3r)(2-s)} > 0, \quad \forall (r,s) \in \mathcal{D}_5.$$

Therefore, $\alpha_5 \in (\alpha_0, 1)$, for any $(r,s) \in \mathcal{D}_5$.

4 Convergence results

In this section, we shall establish the convergence results of the sequence generated by the IPS-ADMM. First, based on (24) and Theorems 3.2-3.6, we can get the following theorem.

Theorem 4.1 Let $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. Then, for any $(r,s) \in \mathcal{D}$, $\alpha \in (c(r,s), 1)$, where $c(r,s)$ is defined in (8), we have

$$\begin{aligned}
&\theta(x^{k+1}) - \theta(x) + (x^{k+1} - x)^\top(-A^\top \lambda) \\
&\leq \tfrac{1}{2}(\|v - v^k\|_H^2 - \|v - v^{k+1}\|_H^2) - \tfrac{1}{2}\|x_1^k - x_1^{k+1}\|_P^2 - \tfrac{C_0}{2}\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) \\
&\quad - \tfrac{C_1}{2}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 - \tfrac{C_2}{2}\beta\|Ax^{k+1} - b\|^2 - \tfrac{C_3}{2}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad - \tfrac{C_4}{2}\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2),
\end{aligned} \tag{37}$$

where $(x_1, x_2, \lambda) \in \mathcal{R}^{m+n_1+n_2}$ satisfies $A_1x_1 + A_2x_2 = b$, and $C_0, C_3, C_4 \geq 0, C_1, C_2 > 0$.

With the above theorem in hand, now we are ready to prove the global convergence of the IPS-ADMM.

Theorem 4.2 Let $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. Then, if A_1, A_2 are both full column rank, we have $\{(x^k, \lambda^k)\}$ is bounded and converges to a point $(x^\infty, \lambda^\infty) \in \mathcal{W}^*$.

Proof. Choose an arbitrary $(x_1^*, x_2^*, \lambda^*) \in \mathcal{W}^*$ and setting $x_1 = x_1^*, x_2 = x_2^*, \lambda = \lambda^*$ in (37), we get

$$\begin{aligned}
&\theta(x^{k+1}) - \theta(x^*) + (x^{k+1} - x^*)^\top(-A^\top \lambda^*) \\
&\leq \frac{1}{2}(\|w^* - w^k\|_H^2 - \|w^* - w^{k+1}\|_H^2) - \frac{1}{2}\|x_1^k - x_1^{k+1}\|_P^2 - \frac{C_0}{2}\beta(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2) \\
&\quad - \frac{C_1}{2}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 - \frac{C_2}{2}\beta\|Ax^{k+1} - b\|^2 - \frac{C_3}{2}(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2) \\
&\quad - \frac{C_4}{2}\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2).
\end{aligned}$$

Then, from $x \in \mathcal{X}_1 \times \mathcal{X}_2$, $(x_1^*, x_2^*, \lambda^*) \in \mathcal{W}^*$ and (23), we have

$$\begin{aligned}
&\|w^{k+1} - w^*\|_H^2 + C_0\beta\|Ax^{k+1} - b\|^2 + C_3\|x_2^k - x_2^{k+1}\|_{G_0}^2 + C_4\beta(\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\leq \|w^k - w^*\|_H^2 + C_0\beta\|Ax^k - b\|^2 + C_3\|x_2^{k-1} - x_2^k\|_{G_0}^2 + C_4\beta\|A_2(x_2^{k-1} - x_2^k)\|^2 \\
&\quad - \|x_1^k - x_1^{k+1}\|_P^2 - C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 - C_2\beta\|Ax^{k+1} - b\|^2,
\end{aligned} \tag{38}$$

which together with $C_0, C_3, C_4 \geq 0, C_1, C_2 > 0, H, G_0 \succ 0$ implies that

$$\begin{aligned}&\sum_{k=1}^{\infty}(\|x_1^k - x_1^{k+1}\|_P^2 + C_1\beta\|A_2(x_2^k - x_2^{k+1})\|^2 + C_2\beta\|Ax^{k+1} - b\|^2)\\ &\leq \|w^1 - w^*\|_H^2 + C_0\beta\|Ax^1 - b\|^2 + C_3\|x_2^0 - x_2^1\|_{G_0}^2 + C_4\beta\|A_2(x_2^0 - x_2^1)\|^2 < +\infty.\end{aligned}$$

This and the full column rank of A_2, the positive definiteness of P indicate that

$$\lim_{k\to\infty}\|x^k - x^{k+1}\| = \lim_{k\to\infty}\|Ax^{k+1} - b\| = 0. \tag{39}$$

Furthermore, it follows from (38) that the sequences $\{w^k\}$ and $\{Ax^k - b\}$ are both bounded. Therefore, $\{w^k\}$ has at least one cluster point, saying w^∞, and suppose that the subsequence $\{w^{k_i}\}$ converges to w^∞. Then, taking limits on both sides of (21) along the subsequence $\{w^{k_i}\}$ and using (39), we have

$$\theta(x) - \theta(x^\infty) + (w - w^\infty)^\top F(w^\infty) \geq 0, \quad \forall w \in \mathcal{W}.$$

Therefore, $w^\infty \in \mathcal{W}^*$.

Hence, replacing w^* by w^∞ in (38), we get

$$\begin{aligned}&\|w^{k+1} - w^\infty\|_H^2 + C_0\beta\|Ax^{k+1} - b\|^2 + C_3\|x_2^k - x_2^{k+1}\|_{G_0}^2 + C_4\beta(\|A_2(x_2^k - x_2^{k+1})\|^2\\ &\leq \|w^k - w^\infty\|_H^2 + C_0\beta\|Ax^k - b\|^2 + C_3\|x_2^{k-1} - x_2^k\|_{G_0}^2 + C_4\beta\|A_2(x_2^{k-1} - x_2^k)\|^2.\end{aligned}$$

From (39), we have that for any given $\varepsilon > 0$, there exists $l_0 > 0$, such that

$$C_0\beta\|Ax^k - b\|^2 + C_3\|x_2^{k-1} - x_2^k\|_{G_0}^2 + C_4\beta\|A_2(x_2^{k-1} - x_2^k)\|^2 < \frac{\varepsilon}{2}, \quad \forall k \geq l_0.$$

Since $w^{k_i} \to w^\infty$ for $i \to \infty$, there exists $k_l > l_0$, such that

$$\|w^{k_l} - w^\infty\|_H^2 < \frac{\varepsilon}{2}.$$

Then, the above three inequalities lead to that, for any $k > k_l$, we have

$$\begin{aligned}&\|w^k - w^\infty\|_H^2\\ &\leq \|w^{k_l} - w^\infty\|_H^2 + C_0\beta\|Ax^{k_l} - b\|^2 + C_3\|x_2^{k_l-1} - x_2^{k_l}\|_{G_0}^2 + C_4\beta\|A_2(x_2^{k_l-1} - x_2^{k_l})\|^2\\ &< \varepsilon.\end{aligned}$$

Therefore the whole sequence $\{w^k\}$ converges to the w^∞. The proof is completed.

Now, we are going to prove the worst-case $\mathcal{O}(1/t)$ convergence rate in an ergodic sense of IPS-ADMM.

Theorem 4.3 Let $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$ be the sequence generated by the IPS-ADMM. and let

$$x_t = \frac{1}{t}\sum_{k=1}^{t} x^{k+1}\ ,$$

where t is a positive integer. Then,

$$\theta(x_t) - \theta(x^*) + (x_t - x^*)^\top(-A^\top \lambda^*) + \frac{C_2}{2}\beta\|Ax_t - b\|^2 \le \frac{D}{t}, \tag{40}$$

where $(x^*, \lambda^*) \in \mathcal{W}^*$, and D is a constant defined by

$$D = \frac{1}{2}\|v^1 - v^*\|_H^2 + \frac{C_0}{2}\beta\|Ax^1 - b\|^2 + \frac{C_3}{2}\|x_2^0 - x_2^2\|_{G_0}^2 + \frac{C_4}{2}\beta\|A_2(x_2^0 - x_2^1)\|^2. \tag{41}$$

Proof. Setting $x = x^*, \lambda = \lambda^*$ in (37), and summing the resulted inequality over $k = 1, 2, \ldots, t$, we have

$$\begin{aligned} &\sum_{k=1}^{t}\left[\theta(x^{k+1}) - \theta(x^*) + (x^{k+1} - x^*)^\top(-A^\top\lambda^*) + \frac{C_2}{2}\beta\|Ax^{k+1} - b\|^2\right] \\ \le\ & \frac{1}{2}\|v^1 - v^*\|_H^2 + \frac{C_0}{2}\beta\|Ax^1 - b\|^2 + \frac{C_3}{2}\|x_2^0 - x_2^2\|_{G_0}^2 + \frac{C_4}{2}\beta\|A_2(x_2^0 - x_2^1)\|^2. \end{aligned} \tag{42}$$

Therefore, dividing (42) by t and using the convexity of $\theta(\cdot)$ lead to

$$\theta(x_t) - \theta(x^*) + (x_t - x^*)^\top(-A^\top\lambda^*) + \frac{C_2\beta}{2t}\sum_{k=1}^{t}\|A^{k+1} - b\|^2 \le \frac{D}{t}, \tag{43}$$

where the constant D is defined by (41).

Compared (43) with (19), we have to deal with the term $\frac{C_2\beta}{2t}\sum_{k=1}^{t}\|A^{k+1} - b\|^2$ of (43). In fact, from the convexity of $\|\cdot\|^2$, we get

$$\begin{aligned} &\frac{C_2\beta}{2t}\sum_{k=1}^{t}\|A^{k+1} - b\|^2 \\ =\ & \frac{C_2\beta}{2}\sum_{k=1}^{t}\frac{1}{t}\|Ax^{k+1} - b\|^2 \\ \ge\ & \frac{C_2\beta}{2}\left\|A\frac{\sum_{k=1}^{t}x^{k+1}}{t} - b\right\|^2 \\ =\ & \frac{C_2\beta}{2}\|Ax_t - b\|^2. \end{aligned}$$

Then, substituting the above inequality into (43), we get the desired result (40). This completes the proof.

5 Numerical results

We have established the convergence results of the IPS-ADMM in theory. In this section, by comparing the IPS-ADMM with the PS-ADMM [15], we are going to highlight its promising numerical behaviors in solving two image restoration problems: the image deblurring with wavelets and the total-variational denoising problem. All the codes were written by Matlab R2010a and all the numerical experiments were conducted on a THINKPAD notebook with Pentium(R) Dual-Core CPU@2.20 GHz and 4 GB RAM.

Problem 1. The image deblurring with wavelets

Let the matrix $W \in \mathcal{R}^{l\times n}$ be a wavelet dictionary with $l = l_1 \times l_2$, the vector $\mathbf{x} \in \mathcal{R}^l$ a digital image. Set $\mathbf{x} = Wx$, and x is a sparse vector. The image deblurring with wavelets is to recover the clean image $x \in \mathcal{R}^n$ based on some observation b, which can modelled as:

$$\min\{\|x\|_1 | BWx = b\}, \tag{44}$$

where $B \in \mathcal{R}^{m\times l}$ is a diagonal matrix. The matrix $B = SH$, where $S \in \mathcal{R}^{m\times l}$ is a downsampling matrix, and H is the blurry matrix generated by the Matlab script $H =$ `fspecial('disk',7)`. Obviously, (44) is a special case of (1) with $\theta_1(x_1) = 0, A_1 = 0, \theta_2(x_2) = \|x\|_1, A_2 = BW$, and the IPS-ADMM is applicable.

Now, let us elaborate on how to derive the closed-form solution for the subproblem resulted by the IPS-ADMM. Set $G = \alpha\tau I_n - \beta(BW)^\top(BW)$ with $\tau > \beta\|(BW)^\top(BW)\|$. For given (x^k, λ^k), the first subproblem is

$$x^{k+1} = \text{argmin}_{x\in\mathcal{R}^n}\left\{\|x\|_1 + \frac{\alpha\tau}{2}\left\|x - \frac{1}{\alpha\tau}(Gx^k + (BW)^\top\lambda^k + \beta(BW)^\top b)\right\|^2\right\},$$

which has a closed-from solution:

$$x^{k+1} = \text{shrink}_{1,2}(a^k, 1/(\alpha\tau)) \doteq \text{sign}(a^k)\cdot\max\{0, |a^k| - 1/(\alpha\tau)\},$$

in which $a^k = \frac{1}{\alpha\tau}(Gx^k + (BW)^\top\lambda^k + \beta(BW)^\top b)$. Note that all computations are component-wise.

In order to demonstrate its efficiency, we compare the IPS-ADMM with the PS-ADMM in [15]. In the numerical experiment, for both tested methods, we set $\alpha = 1.01c(r,s), \beta = 1, \tau = 5\beta\|(BW)^\top(BW)\|$. Furthermore, the maximize iteration number is set 500. The initial point is $(x^0, \lambda^0) = (W^\top(b), 0)$. We test the 256×256 image of Cameraman.tif for the image reconstruction problem, and use the signal to noise ratio (SNR) defined as

$$\text{SNR} = 20\text{log}_{10}\frac{\|\mathbf{x}\|}{\|\tilde{\mathbf{x}} - \mathbf{x}\|}.$$

to assess the restoration performance qualitatively, in which $\mathbf{x}$ is the true image, and $\tilde{\mathbf{x}}$ is the restored image.

The numerical results are listed in Table 1, in which, CPU, and SNR represent computing time in seconds, and the signal to noise ratio when both methods achieve the stopping criterion, which indicate that for all the tested cases the SNR of the IPS-ADMM is obviously bigger than that of the PS-ADMM. Of course, the advantage of the IPS-ADMM should be contributed to the indefinite proximal term. Figure 2 plots the images restored by the IPS-ADMM and the PS-ADMM when $s = 1.1$, which clearly illustrates that the IPS-ADMM recover the degraded image quite well.

Table 1: Comparison of PS-ADMM with IPS-ADMM for the image deblurring with wavelets.

s	PS-ADMM		IPS-ADMM	
	CPU	SNR	CPU	SNR
1.0	191.78	18.86	189.55	18.93
1.1	190.38	19.52	189.66	19.60
1.2	190.38	19.18	191.78	19.25
1.3	196.57	19.71	195.61	19.78
1.4	189.83	19.68	190.05	19.74
1.5	190.64	19.19	191.42	19.22

Problem 2 The total-variational (TV) denoising problem

Below, we consider the total-variational (TV) denoising problem [26]:

$$\min \frac{1}{2}\|y-b\|^2 + \frac{\eta}{2}\|Dy\|_1, \tag{45}$$

where $D^\top = [D_1^\top, D_2^\top]^\top$ is a discrete gradient operator with $D_1 : \mathcal{R}^n \to \mathcal{R}^n, D_2 : \mathcal{R}^n \to \mathcal{R}^n$ being the finite-difference operators in the horizontal and vertical directions, respectively; $\eta > 0$ is the regularization parameter. Here, we set $\eta = 5$.

Introducing an auxiliary variable $x \in \mathcal{R}^{2n}$, we can reformulate (45) as

$$\begin{array}{l} \min \ \eta \ \|x\|_1 + \ \dfrac{1}{2} \ \|y-b\|^2 \\ \text{s.t. } x - Dy = 0, \ x \in \mathcal{R}^{2n}, y \in \mathcal{R}^n, \end{array} \tag{46}$$

Obviously, (46) is a special case of (1), and therefore the IPS-ADMM is applicable. Now, let us elaborate on how to derive the closed-form solutions for the subproblems resulted by the IPS-ADMM.

Set $P = \tau_1 I_{2n}, G = \alpha\tau_2 I_n - \beta D^\top D$. For given (x^k, y^k, λ^k), the first subproblem is

$$x^{k+1} = \text{argmin}_{x\in\mathcal{R}^{2n}} \left\{ \eta\|x\|_1 - (\lambda^k)^\top(x - Dy^k) + \frac{\beta}{2}\|x - Dy^k\|^2 + \frac{1}{2}\|x - x^k\|_P^2 \right\},$$

which has a closed-from solution:

$$x^{k+1} = \text{shrink}_{1,2}\left(\frac{\tau_1 x^k + \beta Dy^k + \lambda^k}{\beta+\tau_1}, \frac{\eta}{\beta+\tau_1}\right).$$

For given $x^{k+1}, y^k, \lambda^{k+\frac{1}{2}}$, the third subproblem is

$$y^{k+1} = \text{argmin}_{y\in\mathcal{R}^n} \left\{ \frac{1}{2}\|y-b\|^2 - (\lambda^{k+\frac{1}{2}})^\top(x^{k+1} - Dy) + \frac{\beta}{2}\|x^{k+1} - Dy\|^2 + \frac{1}{2}\|y - y^k\|_G^2 \right\},$$

which has a closed-from solution:

$$y^{k+1} = \frac{1}{1+\alpha\tau_2}(b - D^\top\lambda^{k+\frac{1}{2}} + \beta Dx^{k+1} + Gy^k).$$

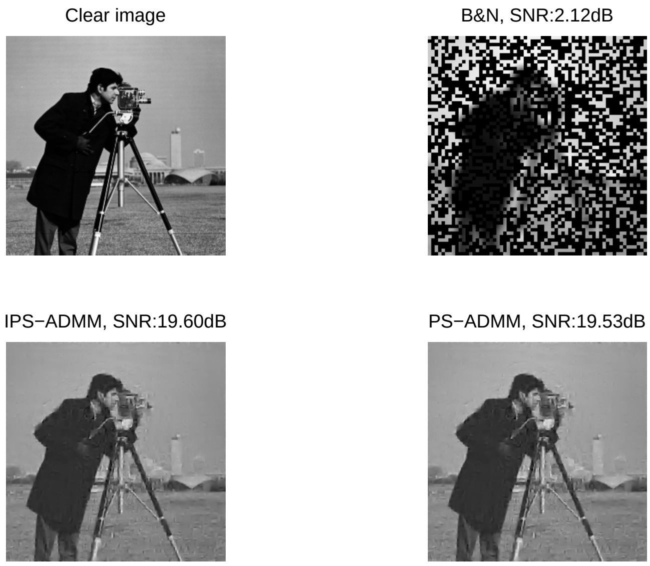

Figure 2: Clear image, blurring image and restored images

For the IPS-ADMM, we set $\beta = 1$, $\tau_1 = 0.001$, $\tau_2 = 1.01\beta\|D^\top D\|$, $\alpha = 1.01c(r, s)$. For the PS-ADMM, we set $G = \tau_2 I_n - \beta B^\top B$. The initial iterate is chosen as $x_0 = 0, y_0 = b, \lambda_0 = 0$. The stopping criterion is the same as that in [2]:

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

where $\epsilon^{\text{pri}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}} \max\{\|x^{k+1}\|, \|Dy^{k+1}\|\}$, and $\epsilon^{\text{dual}} = \sqrt{n}\epsilon^{\text{abs}} + \epsilon^{\text{rel}}\|y^{k+1}\|$ with $\epsilon^{\text{abs}} = 10^{-4}$ and $\epsilon^{\text{rel}} = 10^{-3}$. We use the following Matlab scripts to generate some synthetic data for (46) [21]:

```
for j = 1:3
   id = randsample(n,1);
   id = randsample(n,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);
```

We list some numerical results in Table 2. Numerical results in Table 2 illustrate that the IPS-ADMM often performs much better than the PS-ADMM, though the difference between them only lies in the proximal parameter. Then, the numerical advantage of smaller proximal parameter is verified.

Table 2: Comparison between the number of iterations (time in seconds) taken by PS-ADMM and IPS-ADMM for TV denoising problem.

n	PS-ADMM (r,s)=(-0.3,1.2)	IPS-ADMM (r,s)=(-0.3,1.2)	Ratio(%)	PS-ADMM (r,s)=(0.3,1.2)	IPS-ADMM (r,s)=(0.3,1.2)	Ratio(%)
100	176(0.04)	94(0.03)	0.53(0.60)	149(0.06)	97(0.02)	0.65(0.41)
200	213(0.05)	107(0.03)	0.50(0.49)	180(0.04)	117(0.03)	0.65(0.67)
300	189(0.06)	104(0.03)	0.55(0.45)	160(0.04)	105(0.03)	0.66(0.63)
400	47(0.02)	24(0.01)	0.51(0.43)	40(0.01)	27(0.01)	0.68(0.88)
500	99(0.03)	54(0.02)	0.55(0.56)	84(0.03)	56(0.02)	0.67(0.68)

6 Conclusions

In this paper, a symmetric ADMM with positive-indefinite proximal regularization for two-block linearly constrained convex programming is proposed. Under mild conditions, we have established the global convergence and the worst-case $\mathcal{O}(1/t)$ convergence rate in an ergodic sense of the new method. Two sets of numerical results are given, which illustrate that the new method often performs better than its counterpart with positive definite proximal regularization. Note that this paper only discusses the symmetric ADMM with positive-indefinite proximal regularization for the two-block separable convex problems. In the future, we shall study the ADMM-type method with positive-indefinite proximal regularization for the multi-block cases.

References

[1] Y.L. Wang, J.F. Yang, W.T. Yin, Y. Zhang, A new alternating minimization algorithm for total variation image reconstruction, SIAM Journal on Imaging Sciences, 1(3), 248-272, 2008.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learning, 3, 1-122, 2011.

[3] J.F. Yang, Y. Zhang, Alternating direction algorithms for L1-problems in compressive sensing, SIAM Journal on Scientific Computing, 33(1), 250-278, 2011.

[4] J. Douglas, H.H. Rachford, On the numerical solution of the heat conduction problem in 2 and 3 space variables, Transactions of the American Mathematical Society, 82(82), 421-439, 1956.

[5] P.L. Lions, B. Mercier, Splitting algorithms for the sum of two nonlinear operators, SIAM Journal on Numerical Analysis, 16(16), 964-979, 1979.

[6] D.H. Peaceman, H.H. Rachford, The numerical solution of parabolic elliptic differential equations, Journal of the Society for Industrial and Applied Mathematics, 3(1), 28-41, 1955.

[7] D. Gabay, Applications of the method of multipliers to variational inequalities, Augmented Lagrange Methods: Applications to the Solution of Boundary-valued Problems, M. Fortin and R. Glowinski, eds., North Holland, Amsterdam, The Netherlands, 299-331, 1983.

[8] R. Glowinski, P.L. Tallec, Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics, SIAM Studies in Applied Mathematics, Philadelphia, PA, 1989.

[9] R. Glowinski, A. Marrocco, Sur l′approximation, par éléments fins dordren, et la résolution, par pénalisation-dualité, dune classe de problèmes de dirichlet nonlinéares, Rev. Fr. Autom. Inform. Rech. Oper. 9, 41-76, 1975.

[10] D. Gabay, B. Mercier, A dual algorithm for the solution of nonlinear variational problems via finite-element approximations, Computers and Mathematics with Applications, 2(1), 17-40, 1976.

[11] K.C. Kiwiel, C.H. Rosa, A. Ruszczynski, Proximal decomposition via alternating linearization, SIAM Journal on Optimization, 9(3), 668-689, 1999.

[12] B.S. He, H. Liu, Z.R. Wang, X.M. Yuan, A strictly contractive Peaceman-Rachford splitting method for convex programming, SIAM Journal on Optimization, 24(3), 1011-1040, 2014.

[13] B.S. He, F. Ma, X.M. Yuan, Convergence study on the symmetric version of ADMM with larger step sizes, SIAM Journal on Imaging Sciences, 9(3), 1467-1501, 2016.

[14] J. Eckstein, D. Bertsekas, On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators, Mathematical Programming, 55(1), 293-318, 1992.

[15] X.X. Li, X.M. Li, A proximal strictly contractive Peaceman-Rachford splitting method for convex programming with applications to imaging, SIAM Journal on Imaging Sciences, 8(2), 1332-1365, 2015.

[16] M. Sun, J. Liu, A proximal Peaceman-Rachford splitting method for compressive sensing, Journal of Applied Mathematics and Computing, 50(1), 349-363, 2016.

[17] H.C. Sun, M. Sun, H.C. Zhou, A proximal splitting method for separable convex programming and its application to compressive sensing, Journal of Nonlinear Sciences and Applications, 9(2), 392-403, 2016.

[18] E.J. Candès, X.D. Li, Y. Ma, J. Wright, Robust principal component analysis? J. Assoc. Comput. Mach., 58, 1-37, 2011.

[19] B.S. He, F. Ma, X.M. Yuan, Linearized alternating direction method of multipliers via positive-indefinite proximal regularization for convex programming, manuscript, 2016.

[20] M. Sun, J. Liu, The convergence rate of the proximal alternating direction method of multipliers with indefinite proximal regularization, Journal of Inequalities and Applications, December 2017, 2017: 19.

[21] B. Gao, F. Ma, Symmetric ADMM with positive-indefinite proximal regularization for linearly ocnstrained ocnvex optimization, manuscript, 2017.

[22] B.S. He, X.M. Yuan, On the $\mathcal{O}(1/n)$ convergence rate of the Douglas-Rachford alternating direction method, SIAM Journal on Numerical Analysis, 50(2), 700-709, 2012.

[23] Z.C. Lin, R.S. Liu, H. Li, inearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning, Machine Learning, 99(2), 287-325, 2015.

[24] M.H. Xu, T. Wu, A class of linearized proximal alternating direction methods, Journal of Optimization Theory and Applications, 151(2), 321-337, 2011.

[25] R.H. Chan, M. Tao, X.M. Yuan, Linearized alternating direction method for constrained linear leastsquares problem, East Asian Journal on Applied Mathematics, 2(4), 326C341, 2012.

[26] A. Beck, M. Teboulle , Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, IEEE Trans. Image Process., 18, 2419-2434, 2009.