Sun et al. *Journal of Inequalities and Applications* (2017) 2017:172
DOI 10.1186/s13660-017-1447-3

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH **Open Access**

The symmetric ADMM with indefinite proximal regularization and its application

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

*Correspondence:
ziyouxiaodou@163.com
[3]School of Mathematics and Statistics, Zaozhuang University, Zaozhuang, Shandong 277160, P.R. China
[4]School of Management, Qufu Normal University, Qufu, Shandong 276826, P.R. China
Full list of author information is available at the end of the article

Abstract
Due to updating 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 practical applications, 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 indefinite, and propose a new S-ADMM with indefinite proximal regularization (termed 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 worst-case $\mathcal{O}(1/t)$ convergence rate in an ergodic sense by the iteration complexity. Finally, some numerical results are included to illustrate the efficiency of the IPS-ADMM.

MSC: 90C25; 94A08

Keywords: symmetric alternating direction method of multipliers; indefinite proximal regularization; global convergence

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

$$\min\left\{\theta_1(x_1) + \theta_2(x_2) | A_1x_1 + A_2x_2 = b, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\right\}, \tag{1}$$

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 many concrete problems arising in a wide range of disciplines, such as statistical learning, inverse problems and image processing; see, *e.g.* [1–3] for more details.

Convex programming (1) has been studied extensively in the literature, researchers have developed many numerical methods to solve it during the last decades, which are mainly based on the well-known Douglas-Rachford splitting method [4, 5] and the Peachmen-Rachford splitting method [5, 6], which originate with the partial differential equation

(PDE) literature. Concretely, applying the Douglas-Rachford splitting method to the dual of (1) [7, 8], we get the well-known alternating direction of multipliers (ADMM) [9, 10], whose iterative schemes reads

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

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 [11–13], which generates its sequence via the scheme

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

where the feasible region of r, s is

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

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 well-known ADMM-based schemes as special cases. For example, 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 an ergodic sense, have been established in [13]. To the best of the authors' knowledge, the worst-case $\mathcal{O}(1/t)$ convergence rate in some non-ergodic sense of the S-ADMM is still missing.

In practical applications, the two essential subproblems related to x_1 and x_2 dominate the computation of the S-ADMM, which are often either linear or easily solvable, but nevertheless challenging. In order to solve the issue, some proximal terms are often added to these subproblems, which can linearize the quadratic term $\frac{\beta}{2}\|A_ix_i\|^2$ ($i = 1,2$) of these subproblems, and as a result we have the following proximal S-ADMM (termed PS-ADMM)

[15–17]:

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1\in\mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) \\ \qquad + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta(A_1x_1^{k+1} + A_2x_2^k - b), \\ x_2^{k+1} \in \operatorname{argmin}_{x_2\in\mathcal{X}_2}\{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top(A_1x_k^{k+1} + A_2x_2 - b) \\ \qquad + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2 + \frac{1}{2}\|x_2 - x_2^k\|_G^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - s\beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b), \end{cases} \tag{5}$$

where $G \in \mathcal{R}^{n_2\times n_2}$ is a positive definite matrix. When we set $G = \tau I_{n_2} - \beta A_2^\top A_2$ with $\tau > \beta\|A_2^\top A_2\|$, the quadratic term $\frac{\beta}{2}\|A_2x_2\|^2$ in the subproblem related to x_2 of the PS-ADMM is offset and thus the quadratic term $\frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2$ 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]. $\|x_2\|_*$ is defined by the sum of all singular values of x_2.

The curse accompanying the above improvement in solvability is that the proximal 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 kth iteration, which often slows down the convergence of the corresponding method. Therefore, it is meaningful to expand the feasible set of τ. Obviously, if we further reduce τ to $\tau \leq \beta\|A_2^\top A_2\|$, the proximal matrix G will become indefinite, and it is thus natural to ask whether or not the corresponding method with such G is still globally convergent? Quite recently the authors in [19–21] partially answered the question. More specifically, for the ADMM (2) with $s = 1$, He *et al.* [19] have proved that the feasible set of τ can be expanded to $\{\tau|\tau > 0.8\beta\|A_2^\top A_2\|\}$, and for the ADMM (2) with $s \in (0, \frac{1+\sqrt{5}}{2})$, Sun *et al.* [20] have proved that the feasible set of τ can be expanded to $\{\tau|\tau > (5 - \min\{s, 1 + s - s^2\})\beta\|A_2^\top A_2\|/5\}$. Then, for the S-ADMM with $r \in (-1, 1)$, $s = 1$, Gao *et al.* [21] have proved that the feasible set of τ can be expanded to $\{\tau|\tau > (r^2 - r + 4)\beta\|A_2^\top A_2\|/(r^2 - 2r + 5)\}$. Other relevant studies can be found in [22, 23]. In this paper, we continue to study along this direction, and present a new feasible set of τ, which generalizes those in [19–21] to any $(r, s) \in \mathcal{D}$. Furthermore, we show that for any $(r, s) \in \mathcal{D}$, the global convergence of the S-ADMM with some indefinite proximal regularization can be guaranteed.

The rest of the paper is organized as follows. In Section 2, we summarize some preliminaries which are useful for further discussion. Then, in Section 3, we list the iterative scheme of the IPS-ADMM and prove its convergence results, including the global convergence and the convergence rate. Some preliminary numerical results are reported in Section 4. Finally, some conclusions are drawn in Section 5.

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:

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

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

Throughout, the proximal matrix G is defined by

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

where we set $\tau = \alpha\tilde{\tau}$ with $\tilde{\tau} > \beta\|A_2^\top A_2\|$, $\alpha \in (c(r,s), +\infty)$, and $c(r,s)$ is defined by

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

Remark 2.1 Note that $c(r,s) \le 1$ if $(r,s) \in \mathcal{D}$; see Lemmas 3.4-3.8 in Section 3. Therefore, the feasible set of τ is expanded from $\{\tau | \tau > \beta\|A_2^\top A_2\|\}$ to $\{\tau | \tau > c(r,s)\beta\|A_2^\top A_2\|\}$, which provides more choices for researchers or practitioners.

Furthermore, we define an auxiliary matrix as follows:

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

which is positive definite by $\tilde{\tau} > \beta\|A_2^\top A_2\|$.

Invoking the first-order optimality condition for convex programming, we get the following equivalent form of problem (1): Finding a vector $w^* \in \mathcal{W}$ such that

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

where

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

$$F(w) = \begin{pmatrix} -A_1^\top \lambda \\ -A_2^\top \lambda \\ A_1x_1 + A_2x_2 - b \end{pmatrix}, \qquad \mathcal{W} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{R}^m. \tag{11}$$

Obviously, the problem (10) is a mixed variational inequality problem, which is denoted by $\mathrm{MVI}(\theta, F, \mathcal{W})$. The mapping $F(w)$ defined in (11) is not only monotone, but also satisfies the property

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

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

Now, let us define three matrices in order to make our following analysis more succinct. Set

$$M = \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -s\beta A_2 & (r+s)I_m \end{pmatrix}, \qquad Q = \begin{pmatrix} P & 0 & 0 \\ 0 & G + \beta A_2^\top A_2 & -rA_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_m \end{pmatrix}, \tag{13}$$

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

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 (14), we get

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

where the relationship $\succeq$ comes from $\alpha > 0$ and $\tilde{\tau} > \beta\|A_2^\top A_2\|$. 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

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

Therefore, the matrix H is positive definite. The proof is completed. □

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 $d = \sup\{\|w - w^0\| \mid w \in \bar{\mathcal{D}}\}$, where w^0 is the initial iterate. He *et al.* [24] established the following criterion:

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

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 reasonable.

(2) In [25], Lin *et al.* proposed the following criterion:

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

where $c > 0$. Proposition 1 in [25] 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 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 and convergence results

In this section, we first present the symmetric ADMM with indefinite proximal regularization (termed IPS-ADMM), and then prove the convergence results of the sequence generated by the IPS-ADMM.

Algorithm 3.1 (The IPS-ADMM for problem (1))

Step 0. Input four parameters $(r,s) \in \mathcal{D}$, $\alpha \in (c(r,s),+\infty)$, $\beta > 0$, 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:

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

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

Remark 3.1 Since the global convergence of IPS-ADMM with $\alpha \ge 1$ has been established in the literature [16, 26–28], in the following, we restrict $\alpha \in (c(r,s),1)$.

To prove the convergence results of the IPS-ADMM, we first define a block matrix and an auxiliary variable.

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

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

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

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

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

Remark 3.2 By the definition of $F(\cdot)$ in (11), (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

$$\begin{aligned} &(w - \tilde{w}^k)^\top F(w) \\ &\quad = (w - \tilde{w}^k)^\top F(\tilde{w}^k) \end{aligned}$$

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

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

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

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

Comparing all the terms appeared in (19) and (24), we find that the left-hand side of (24) does not 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. Now 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 a manipulation, 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*

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

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 [26] 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*

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

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

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

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

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

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

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

Then, adding the above two inequalities, we get

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

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

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

Substituting the above equality into the left-hand side of (29), we get

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

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

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

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 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*

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

Now, let us rewrite all the terms on the right-hand side of (31) by some quadratic terms, and mainly deal with the term $\|x_2^k-x_2^{k+1}\|_{G_0}^2$ and the crossing term $(Ax^k-b)^\top A_2(x_2^k-x_2^{k+1})$. 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, which are discussed in the following five subsections.

3.1 Case 1: $(r,s)\in\mathcal{D}_1$

Lemma 3.4 *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* $C_{11},C_{12}>0$ *such that*

$$\|w^k-\tilde{w}^k\|_R^2\geq\|x_1^k-x_1^{k+1}\|_P^2+C_{11}\beta\|A_2(x_2^k-x_2^{k+1})\|^2+C_{12}\beta\|Ax^{k+1}-b\|^2.\tag{32}$$

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

Proof We prove the assertion (32) from the definition of the matrix R directly. Define an auxiliary 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

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

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 matrix S is positive definite, and then the matrices R and R_0 are both positive definite by the full column rank of A_2 and the positive definiteness of P. By a manipulation, we get

$$\begin{aligned} &\|w^k - \tilde{w}^k\|_R^2 \\ &\quad \geq \|w^k - \tilde{w}^k\|_{R_0}^2 \\ &\quad = \|x_1^k - x_1^{k+1}\|_P^2 + \left\| \begin{matrix} A_2(x_2^k - x_2^{k+1}) \\ Ax^{k+1} - b \end{matrix} \right\|_{\tilde{S}}^2, \end{aligned}$$

where

$$\tilde{S} = L^\top SL, \quad L = \begin{pmatrix} I_m & 0 \\ \beta I_m & \beta I_m \end{pmatrix}.$$

By the positive definiteness of the matrix S, we get the assertion (32). 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 Case 2: $(r,s) \in \mathcal{D}_2$

Lemma 3.5 *For any $(r,s) \in \mathcal{D}_2$, if $\alpha > \alpha_2 \doteq \frac{4-r-r^2}{5-3r}$, then we have*

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

where C_{2i} $(i = 1,2,3,4)$ *are four positive constants defined by*

$$C_{21} = \frac{(5-3r)\alpha + r^2 + r - 4}{1+r}, \qquad C_{22} = 1 - r,$$
$$C_{23} = \frac{(1-r)\alpha}{1+r}, \qquad C_{24} = \frac{(1-r)(1-\alpha)}{1+r}.$$

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

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

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

which proves (33). From $\alpha > \alpha_2$, it is obvious that $C_{21} > 0$, and from $r \in (-1,1)$, $\alpha \in (\alpha_2, 1)$, we have $C_{22}, C_{23}, C_{24} > 0$. 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 3.3 For any $(r,s) \in \mathcal{D}_2$, Gao *et al.* [21] have proved 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 with $r \in (-1,1)$ are drawn in Figure 1, from which we have $\alpha_2 < \alpha_G$ if $r \in (-1,0)$, and $\alpha_2 > \alpha_G$ if $r \in (0,1)$. Therefore, compared with that in [21], the feasible set of τ in this paper is expanded if $r \in (-1,0)$, and is shrunk if $r \in (0,1)$. However, Gao *et al.* only established the worst-case convergence rate of the IPS-ADMM using the criterion (18), and we shall prove the worst-case convergence rate of the IPS-ADMM using the more reasonable criterion (19); see the following Theorem 3.4.

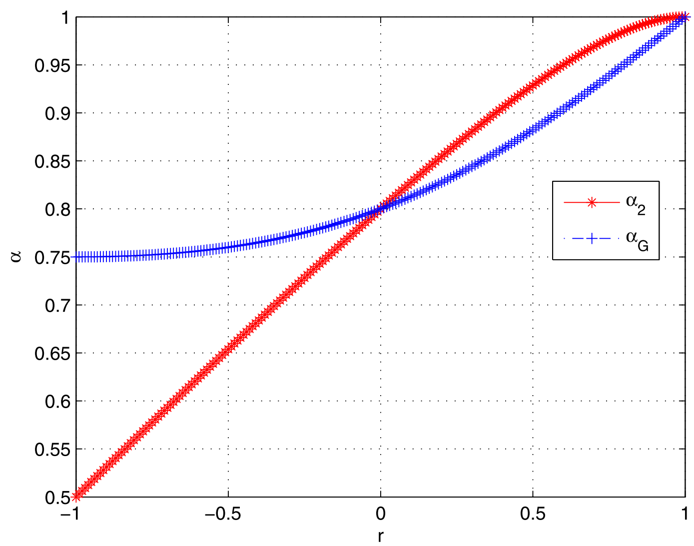

Figure 1 **The curves of α_2 and α_G in $r \in (-1, 1)$.**

3.3 Case 3: $(r, s) \in \mathcal{D}_3$

Lemma 3.6 *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*

$$\begin{aligned}
&\|w^k - \tilde{w}^k\|_R^2 \\
&\quad \geq \|x_1^k - x_1^{k+1}\|_P^2 + C_{30}\beta\big(\|Ax^{k+1} - b\|^2 - \|Ax^k - b\|^2\big) + C_{31}\beta\|A_2(x_2^k - x_2^{k+1})\|^2 \\
&\qquad + C_{32}\beta\|Ax^{k+1} - b\|^2 + C_{33}\big(\|x_2^k - x_2^{k+1}\|_{G_0}^2 - \|x_2^{k-1} - x_2^k\|_{G_0}^2\big) \\
&\qquad + C_{34}\beta\big(\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^{k-1} - x_2^k)\|^2\big),
\end{aligned} \tag{34}$$

where C_{3i} $(i = 0, 1, 2, 3, 4)$ are five positive constants defined by

$$\begin{aligned}
&C_{30} = T_1 - s, \qquad C_{31} = 1 - \frac{(1-s)^2}{T_1 - s} - 5(1-\alpha), \qquad C_{32} = 2 - T_1, \\
&C_{33} = \alpha, \qquad C_{34} = 1 - \alpha, \quad T_1 = \frac{1}{3}\big(s^2 - s + 5\big).
\end{aligned}$$

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

Proof By the Cauchy-Schwartz inequality, we have

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

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

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

$$
\begin{aligned}
&+ (2 - T_1)\beta \left\| Ax^{k+1} - b \right\|^2 + \alpha \left(\left\| x_2^k - x_2^{k+1} \right\|_{G_0}^2 - \left\| x_2^{k-1} - x_2^k \right\|_{G_0}^2 \right) \\
&+ (1 - \alpha)\beta \left(\left\| A_2 \left(x_2^k - x_2^{k+1} \right) \right\|^2 - \left\| A_2 \left(x_2^{k-1} - x_2^k \right) \right\|^2 \right),
\end{aligned}
$$

which proves (34). From $\alpha \in (\alpha_3, 1)$, it is obvious that

$$
\begin{aligned}
&C_{30} = \frac{1}{3}\left[(s-2)^2 + 1\right] > \frac{1}{3}, \\
&C_{31} = \frac{(5s^2 - 20s + 25)\alpha - 7s^2 + 22s - 23}{s^2 - 4s + 5} > 0, \\
&C_{32} = \frac{1}{3}\left(1 + s - s^2\right) > 0, \\
&C_{33} = \alpha > 0, \qquad C_{34} = 1 - \alpha > 0.
\end{aligned}
$$

Furthermore, by some manipulations, $\forall (r, s) \in \mathcal{D}_3$, we have

$$
\begin{aligned}
&\alpha_3 - \alpha_0 = \alpha_3 = \frac{7s^2 - 22s + 23}{5(s^2 - 4s + 5)} > 0, \\
&1 - \alpha_3 = \frac{-2(s - \frac{1-\sqrt{5}}{2})(s - \frac{1+\sqrt{5}}{2})}{5(s^2 - 4s + 5)} > 0.
\end{aligned}
$$

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

3.4 Case 4: $(r, s) \in \mathcal{D}_4$

Lemma 3.7 *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*

$$
\begin{aligned}
&\left\| w^k - \tilde{w}^k \right\|_R^2 \\
&\quad \geq \left\| x_1^k - x_1^{k+1} \right\|_P^2 + C_{40}\beta \left(\left\| Ax^{k+1} - b \right\|^2 - \left\| Ax^k - b \right\|^2 \right) + C_{41}\beta \left\| A_2 \left(x_2^k - x_2^{k+1} \right) \right\|^2 \\
&\qquad + C_{42}\beta \left\| Ax^{k+1} - b \right\|^2 + C_{43} \left(\left\| x_2^k - x_2^{k+1} \right\|_{G_0}^2 - \left\| x_2^{k-1} - x_2^k \right\|_{G_0}^2 \right) \\
&\qquad + C_{44}\beta \left(\left\| A_2 \left(x_2^k - x_2^{k+1} \right) \right\|^2 - \left\| A_2 \left(x_2^{k-1} - x_2^k \right) \right\|^2 \right),
\end{aligned} \tag{35}
$$

where C_{4i} $(i = 0, 1, 2, 3, 4)$ *are five positive constants defined by*

$$
\begin{aligned}
&C_{40} = T_2 - (r + s), \qquad C_{41} = r\frac{(1-r)^2}{(1+r)^2} - 4(1 - \alpha)\frac{1-r}{1+r} + (\alpha - 1), \\
&C_{42} = 2 - T_2, \qquad C_{43} = \alpha\frac{1-r}{1+r}, \qquad C_{44} = (1 - \alpha)\frac{1-r}{1+r}, \quad T_2 = r + s + (1 - s)^2.
\end{aligned}
$$

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

Proof By the Cauchy-Schwartz inequality, we have

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

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

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

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 $C_{40}, C_{42}, C_{43}, C_{44} > 0$. From the definition of C_{41}, we get

$$C_{41} = \frac{(r+1)(5-3r)\alpha + r^3 + r^2 - r - 5}{(r+1)^2} > 0, \quad \forall \alpha > \alpha_4.$$

Furthermore, by some manipulations, $\forall (r,s) \in \mathcal{D}_3$, we have

$$\begin{aligned}
\alpha_4 - \alpha_0 &= \frac{(1-r)[2(1-r^2) + r + 3]}{(1+r)(5-3r)} > 0, \\
1 - \alpha_4 &= \frac{r(r-1)^2}{(1+r)(5-3r)} > 0.
\end{aligned}$$

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

3.5 Case 5: $(r,s) \in \mathcal{D}_5$

Lemma 3.8 *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*

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

where C_{5i} $(i = 0,1,2,3,4)$ are five positive constants defined by

$$\begin{aligned}
&C_{50} = \frac{(s^2 - s)(2-s)}{1+r}, \qquad C_{51} = \frac{(1-r)^2(1+s-s^2)}{s(1+r)(2-s)} - 4(1-\alpha)\frac{1-r}{1+r} + (\alpha - 1), \\
&C_{52} = 2 - T_3, \qquad C_{53} = \alpha\frac{1-r}{1+r}, \\
&C_{54} = (1-\alpha)\frac{1-r}{1+r}, \quad T_3 = r + s + \frac{(s^2 - s)(2-s)}{1+r}.
\end{aligned}$$

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

Proof By the Cauchy-Schwartz inequality, we have

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

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

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

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 $C_{50},C_{52},C_{53},C_{54}>0$. From the definition of C_{51}, for any $(r,s)\in\mathcal{D}_5$, we get

$$C_{51}=\frac{s(2-s)(5-3r)\alpha-(r^2+r-4)s^2+(r^2+4r-9)s+(r-1)^2}{s(2-s)(r+1)}>0,\quad\forall\alpha>\alpha_5.$$

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

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

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 $(r,s)\in\mathcal{D}_5$. □

In the remainder of this section, we shall establish the convergence results of the sequence generated by the IPS-ADMM. First, based on (24) and Lemmas 3.4-3.8, we can get the following theorem.

Theorem 3.2 *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*

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

where $(x_1,x_2,\lambda) \in \mathcal{R}^{m+n_1+n_2}$ satisfies $A_1x_1 + A_2x_2 = b$, $C_0, C_3, C_4 \ge 0$, $C_1, C_2 > 0$ with $C_j = C_{ij}$ if $(r,s) \in \mathcal{D}_i$, $i = 1,2,3,4,5$, $j = 0,1,2,3,4$.

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

Theorem 3.3 *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, the sequence $\{(x^k,\lambda^k)\}$ 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

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

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

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

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

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

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

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

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 the limits on both sides of (21) along the subsequence $\{w^{k_i}\}$ and using (39), we have

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

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

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

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

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

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

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

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

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

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

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 the IPS-ADMM.

Theorem 3.4 *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,*

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

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

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

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

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

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

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

where the constant D is defined by (41).

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

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

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

4 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 an image restoration problem: 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.

Below, we consider the total-variational (TV) denoising problem [29]:

$$\min \frac{1}{2}\|y-b\|^2 + \frac{\eta}{2}\|Dy\|_1, \tag{44}$$

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 (44) as

$$\begin{aligned} &\min \eta\|x\|_1 + \frac{1}{2}\|y-b\|^2 \\ &\quad \text{s.t. } x - Dy = 0, x \in \mathcal{R}^{2n}, y \in \mathcal{R}^n. \end{aligned} \tag{45}$$

Obviously, (45) 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

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

which has a closed-form 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

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

which has a closed-form solution:

$$y^{k+1} = \frac{1}{1+\alpha\tau_2}\left(b - D^\top \lambda^{k+\frac{1}{2}} + \beta D x^{k+1} + G y^k\right).$$

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 initialization is chosen as $x_0 = 0$, $y_0 = b$, $\lambda_0 = 0$. The stopping criterion is the same as that in [2]:

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

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

Table 1 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)

for (45) [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 1. Numerical results in Table 1 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.

5 Conclusions

In this paper, a symmetric ADMM with 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. Some 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 indefinite proximal regularization for the two-block separable convex problems. In the future, we shall study the ADMM-type method with indefinite proximal regularization for the multi-block case.

Acknowledgements
The authors gratefully acknowledge the valuable comments of the anonymous reviewers. This work is supported by the foundation of National Natural Science Foundation of China (Nos. 11601475, 11671228), the foundation of First Class Discipline of Zhejiang-A (Zhejiang University of Finance and Economics - Statistics), the foundation of National Natural Science Foundation of Shandong Province (No. ZR2016AL05) and Scientific Research Project of Shandong Universities (No. J15LI11).

Competing interests
The authors declare that they have no competing interests.

Authors' contributions

The first author has proposed the motivations of the manuscript; the second author has proved the convergence result, and the third author has accomplished the numerical results. All authors read and approved the final manuscript.

Author details

[1]School of Sciences, Linyi University, Linyi, Shandong 276005, P.R. China. [2]Department of Physiology, Shandong Coal Mining Health School, Taozhuang, Shandong 277011, P.R. China. [3]School of Mathematics and Statistics, Zaozhuang University, Zaozhuang, Shandong 277160, P.R. China. [4]School of Management, Qufu Normal University, Qufu, Shandong 276826, P.R. China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 10 May 2017 Accepted: 30 June 2017 Published online: 21 July 2017

References

1. Wang, YL, Yang, JF, Yin, WT, Zhang, Y: A new alternating minimization algorithm for total variation image reconstruction. SIAM J. Imaging Sci. **1**(3), 248-272 (2008)
2. Boyd, S, Parikh, N, Chu, E, Peleato, B, Eckstein, J: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**, 1-122 (2011)
3. Yang, JF, Zhang, Y: Alternating direction algorithms for ℓ_1-problems in compressive sensing. SIAM J. Sci. Comput. **33**(1), 250-278 (2011)
4. Douglas, J, Rachford, HH: On the numerical solution of the heat conduction problem in 2 and 3 space variables. Trans. Am. Math. Soc. **82**(82), 421-439 (1956)
5. Lions, PL, Mercier, B: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**(16), 964-979 (1979)
6. Peaceman, DH, Rachford, HH: The numerical solution of parabolic elliptic differential equations. J. Soc. Ind. Appl. Math. **3**(1), 28-41 (1955)
7. Gabay, D: Applications of the method of multipliers to variational inequalities. In: Fortin, M, Glowinski, R (eds.) Augmented Lagrange Methods: Applications to the Solution of Boundary-Valued Problems, pp. 299-331. North-Holland, Amsterdam (1983)
8. Glowinski, R, Tallec, PL: Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. SIAM Studies in Applied Mathematics. SIAM, Philadelphia (1989)
9. Glowinski, R, Marrocco, A: Sur l'approximation, par éléments fins d'ordren, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet nonlinéares. Rev. Fr. Autom. Inform. Rech. Opér. **9**, 41-76 (1975)
10. Gabay, D, Mercier, B: A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. Comput. Math. Appl. **2**(1), 17-40 (1976)
11. Kiwiel, KC, Rosa, CH, Ruszczynski, A: Proximal decomposition via alternating linearization. SIAM J. Optim. **9**(3), 668-689 (1999)
12. He, BS, Liu, H, Wang, ZR, Yuan, XM: A strictly contractive Peaceman-Rachford splitting method for convex programming. SIAM J. Optim. **24**(3), 1011-1040 (2014)
13. He, BS, Ma, F, Yuan, XM: Convergence study on the symmetric version of ADMM with larger step sizes. SIAM J. Imaging Sci. **9**(3), 1467-1501 (2016)
14. Eckstein, J, Bertsekas, D: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**(1), 293-318 (1992)
15. Li, XX, Li, XM: A proximal strictly contractive Peaceman-Rachford splitting method for convex programming with applications to imaging. SIAM J. Imaging Sci. **8**(2), 1332-1365 (2015)
16. Sun, M, Liu, J: A proximal Peaceman-Rachford splitting method for compressive sensing. J. Appl. Math. Comput. **50**(1), 349-363 (2016)
17. Sun, HC, Sun, M, Zhou, HC: A proximal splitting method for separable convex programming and its application to compressive sensing. J. Nonlinear Sci. Appl. **9**(2), 392-403 (2016)
18. Candès, EJ, Li, XD, Ma, Y, Wright, J: Robust principal component analysis? J. Assoc. Comput. Mach. **58**, 1-37 (2011)
19. He, BS, Ma, F, Yuan, XM: Linearized alternating direction method of multipliers via indefinite proximal regularization for convex programming. Manuscript (2016)
20. Sun, M, Liu, J: The convergence rate of the proximal alternating direction method of multipliers with indefinite proximal regularization. J. Inequal. Appl. **2017**, 19 (2017)
21. Gao, B, Ma, F: Symmetric alternating direction method with indefinite proximal regularization for linearly constrained convex optimization. J. Optim. Theory Appl. (2017, submitted)
22. He, BS, Ma, F: Convergence study on the proximal alternating direction method with larger step size. Manuscript (2017)
23. Tao, M, Yuan, XM: On Glowinski's open question of alternating direction method of multipliers. Manuscript (2017)
24. He, BS, Yuan, XM: On the $\mathcal{O}(1/n)$ convergence rate of the Douglas-Rachford alternating direction method. SIAM J. Numer. Anal. **50**(2), 700-709 (2012)
25. Lin, ZC, Liu, RS, Li, H: Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning. Mach. Learn. **99**(2), 287-325 (2015)
26. Xu, MH, Wu, T: A class of linearized proximal alternating direction methods. J. Optim. Theory Appl. **151**(2), 321-337 (2011)
27. Fang, EX, He, BS, Liu, H, Yuan, XM: Generalized alternating direction method of multipliers: new theoretical insights and applications. Math. Program. Comput. **7**(2), 149-187 (2015)
28. Sun, M, Wang, YJ, Liu, J: Generalized Peaceman-Rachford splitting method for multiple-block separable convex programming with applications to robust PCA. Calcolo **54**(1), 77-94 (2017)
29. Beck, A, Teboulle, M: Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. IEEE Trans. Image Process. **18**, 2419-2434 (2009)