

RESEARCH

Open Access



# The convergence rate of the proximal alternating direction method of multipliers with indefinite proximal regularization

Min Sun<sup>1,2\*</sup> and Jing Liu<sup>3</sup>

\*Correspondence:

ziyouxiaodou@163.com

<sup>1</sup>School of Management, Qufu Normal University, Shandong, 276826, P.R. China

<sup>2</sup>School of Mathematics and Statistics, Zaozhuang University, Shandong, 277160, P.R. China  
Full list of author information is available at the end of the article

## Abstract

The proximal alternating direction method of multipliers (P-ADMM) is an efficient first-order method for solving the separable convex minimization problems. Recently, He *et al.* have further studied the P-ADMM and relaxed the proximal regularization matrix of its second subproblem to be indefinite. This is especially significant in practical applications since the indefinite proximal matrix can result in a larger step size for the corresponding subproblem and thus can often accelerate the overall convergence speed of the P-ADMM. In this paper, without the assumptions that the feasible set of the studied problem is bounded or the objective function's component  $\theta_i(\cdot)$  of the studied problem is strongly convex, we prove the worst-case  $\mathcal{O}(1/t)$  convergence rate in an ergodic sense of the P-ADMM with a general Glowinski relaxation factor  $\gamma \in (0, \frac{1+\sqrt{5}}{2})$ , which is a supplement of the previously known results in this area. Furthermore, some numerical results on compressive sensing are reported to illustrate the effectiveness of the P-ADMM with indefinite proximal regularization.

**MSC:** 90C25; 90C30

**Keywords:** proximal alternating direction method of multipliers; two-block separable convex minimization problem; compressive sensing

## 1 Introduction

Let  $\theta_i : \mathcal{R}^{n_i} \rightarrow (-\infty, +\infty]$  ( $i = 1, 2$ ) be two lower semicontinuous proper (not necessarily smooth) functions. This work aims to solve the following two-block separable convex minimization problem:

$$\min\{\theta_1(x_1) + \theta_2(x_2) | A_1x_1 + A_2x_2 = b\}, \quad (1)$$

where  $A_i \in \mathcal{R}^{l \times n_i}$  ( $i = 1, 2$ ),  $b \in \mathcal{R}^l$ . If there are convex set constraints  $x_i \in \mathcal{X}_i$  ( $i = 1, 2$ ), where  $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}$  ( $i = 1, 2$ ) are some simple convex set, such as the nonnegative cones or positive semi-definite cones, etc. Then, we can define the indicator function as  $I_{\mathcal{X}_i}(\cdot)$  ( $I_{\mathcal{X}_i}(x_i) = 0$  if  $x_i \in \mathcal{X}_i$ ; otherwise,  $I_{\mathcal{X}_i}(x_i) = +\infty$ ), by which we can incorporate the constraints  $x_i \in \mathcal{X}_i$  ( $i = 1, 2$ ) into the objective function of (1), and get the following equivalent form:

$$\min\{\theta_1(x_1) + I_{\mathcal{X}_1}(x_1) + \theta_2(x_2) + I_{\mathcal{X}_2}(x_2) | A_1x_1 + A_2x_2 = b\}.$$

---

**Algorithm 1** The P-ADMM for (1)

---

Input  $\beta > 0, \gamma \in (0, \frac{1+\sqrt{5}}{2})$ , two symmetric matrices  $G_1 \in \mathcal{R}^{n_1 \times n_1}$  and  $G_2 \in \mathcal{R}^{n_2 \times n_2}$ . Initialize  $(x_1, x_2, \lambda) := (x_1^0, x_2^0, \lambda^0)$ ,  $k := 0$ .

**while** ‘not converged,’ **do**

(1) Compute  $(x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$  in the alternating order by the following P-ADMM procedure.

$$\begin{cases} x_1^{k+1} \in \operatorname{argmin}_{x_1 \in \mathcal{R}^{n_1}} \{ \theta_1(x_1) + \frac{\beta}{2} \|A_1 x_1 + A_2 x_2^k - b - \frac{\lambda^k}{\beta}\|^2 \\ \quad + \frac{1}{2} \|x_1 - x_1^k\|_{G_1}^2 \}, \\ x_2^{k+1} \in \operatorname{argmin}_{x_2 \in \mathcal{R}^{n_2}} \{ \theta_2(x_2) + \frac{\beta}{2} \|A_1 x_1^{k+1} + A_2 x_2 - b - \frac{\lambda^k}{\beta}\|^2 \\ \quad + \frac{1}{2} \|x_2 - x_2^k\|_{G_2}^2 \}, \\ \lambda^{k+1} = \lambda^k - \gamma \beta (A_1 x_1^{k+1} + A_2 x_2^{k+1} - b). \end{cases} \tag{2}$$

(2) Set  $k := k + 1$ .

**end while**

Output  $x_1^{k+1}, x_2^{k+1}$ .

---

Then, we can further introduce some auxiliary variables and functions to rewrite the above problem as problem (1) (Please refer to [1] for more details). Therefore, problem (1) is quite general, and in fact problems like (1) come from diverse applications, such as the latent variable graphical model selection [2], the sparse inverse covariance selection [3], stable principal component pursuit with nonnegative constraint [4], and robust alignment for linearly correlated images [5], etc.

As one of the first-order methods, the following Algorithm 1, that is proximal alternating direction method of multipliers (P-ADMM) [6–8] is quite efficient for solving (1) or related problems, especially for large scale case.

The parameter  $\gamma$  in the P-ADMM is called the Glowinski relaxation factor in the literature, and  $\gamma > 1$  can often accelerate the P-ADMM [9]. Due to its high efficiency, the P-ADMM has been intensively studied during the past few decades, and many scholars presented a lot of customized variants of the P-ADMM for some concrete separable minimization problems [10–12].

In this paper, we only focus our attention on the P-ADMM. In fact, the theory developed in this work can easily be extended to its various variants. Now, let us briefly analyze the structure advantages of the P-ADMM. Obviously, the P-ADMM fully utilizes the separable structure inherent to the original problem (1), which decouples the primal variable  $(x_1, x_2)$  and get two subproblems with lower-dimension. Then, at each iteration, the computation of P-ADMM is dominated by solving its two subproblems. Fortunately, the two subproblems in (2) often admit closed-form solutions provided that  $\theta_i(\cdot)$  ( $i = 1, 2$ ) are some the functions (such as  $\theta_i(\cdot) = \|\cdot\|_1, \|\cdot\|_2$  or  $\|\cdot\|_*$ ) and the matrices  $A_i$  ( $i = 1, 2$ ) are unitary (i.e.  $A_i^T A_i$  ( $i = 1, 2$ ) are the identity matrices). Even if  $A_i$  ( $i = 1, 2$ ) are not unitary, we can judiciously set  $G_i = rI_{n_i} - \beta A_i^T A_i$  with  $r > \beta \|A_i^T A_i\|$  ( $i = 1, 2$ ), and then the two subproblems in the P-ADMM also have closed-form solutions in many practical applications. The global convergence of the P-ADMM with  $\gamma = 1$  has been proved in [10, 11] for some concrete models of (1), and in [13], Xu and Wu presented an elegant analysis of the global convergence of the P-ADMM with  $\gamma \in (0, \frac{1+\sqrt{5}}{2})$  for the general model (1). Quite recently,

He *et al.* [14] have further studied the P-ADMM and get some substantial advances by relaxing the matrix  $G_2$  in the proximal regularization term of its second subproblem to be indefinite. This is quite preferred in practical applications since the indefinite proximal matrix can result in a larger step size for the subproblem and thus maybe accelerate the overall convergence speed of the P-ADMM.

Compared with the study of the global convergence of the P-ADMM, the research of its convergence rate is quite insubstantial in the literature. In [14, 15], under the assumption that the feasible set of (1) is bounded, He *et al.* have proved the worst-case  $\mathcal{O}(1/t)$  convergence rate of the P-ADMM with  $\gamma = 1$ , where  $t$  denotes the iteration counter. In [1], Lin *et al.* have presented a parallel version of the P-ADMM with the adaptive penalty  $\beta$ , and proved that the convergence rate of their new method is also  $\mathcal{O}(1/t)$ . In addition, Goldstein *et al.* [16] proved a better convergence rate than  $\mathcal{O}(1/t)$  for the P-ADMM scheme with  $\gamma = 1$  and  $G_1 = 0, G_2 = 0$  under the assumption that  $\theta_i(\cdot)$  ( $i = 1, 2$ ) are both strongly convex, which is usually violated in practice, and thus excludes many practical applications of the P-ADMM. Then, by introducing some free parameters  $\alpha_k$  and  $\gamma_k$ , Xu [17] developed a new variant of the P-ADMM for (1), which refined the results in [16]. In fact, only under the assumption that the function  $\theta_2(\cdot)$  is strongly convex, Xu [17] proved that the new method has  $\mathcal{O}(1/t)$  convergence rate with constant parameters and enjoys  $\mathcal{O}(1/t^2)$  convergence rate with adaptive parameters.

In this paper, we aim to further improve the above results by removing the assumptions of the strong convexity of  $\theta_2(\cdot)$  and the boundedness of the feasible set of (1), and prove that the P-ADMM for the convex minimization problem (1) has a worst-case  $\mathcal{O}(1/t)$  convergence rate in an ergodic sense, which partially improves the results in [8, 13–15, 17].

The remaining of the paper is organized as follows. Section 2 gives some useful preliminaries. In Section 3, we prove the convergence rate of the P-ADMM in detail. In Section 4, a simple experiment on compressive sensing is conducted to demonstrate the effectiveness of the P-ADMM.

## 2 Preliminaries

In this section, we summarize some basic concepts and preliminaries that will be used in the later discussion.

First, we list some notation to be used in this paper.  $\langle \cdot, \cdot \rangle$  denotes the inner product of  $\mathcal{R}^n$ ;  $G \succ 0$  (or  $G \geq 0$ ) denotes that the symmetric matrix  $G$  is positive definite (or positive semi-definite); If  $G$  is symmetric, we set  $\|x\|_G^2 = x^\top Gx$  though  $G$  maybe not positive definite. The effective domain of a function  $f : \mathcal{X} \rightarrow (-\infty, +\infty]$  is defined as  $\text{dom}(f) := \{x \in \mathcal{X} | f(x) < +\infty\}$ . The set of all relative interior points of a given nonempty convex set  $\mathcal{C}$  is denoted by  $\text{ri}(\mathcal{C})$ .

A function  $f : \mathcal{R}^n \rightarrow \mathcal{R}$  is convex iff

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \forall x, y \in \mathcal{R}^n, \alpha \in [0, 1].$$

Then, if  $f : \mathcal{R}^n \rightarrow \mathcal{R}$  is convex, we have the following first-order necessary condition:

$$f(x) \geq f(y) + \langle \xi, x - y \rangle, \quad \forall x, y \in \mathcal{R}^n, \xi \in \partial f(y), \tag{3}$$

where  $\partial f(y) = \{\xi \in \mathcal{R}^n : f(\bar{y}) \geq f(y) + \langle \xi, \bar{y} - y \rangle, \text{ for all } \bar{y} \in \mathcal{R}^n\}$  denotes the subdifferential of  $f(\cdot)$  at the point  $y$ .

The following equality is used frequently in the paper:

$$\langle x - y, x - z \rangle = \frac{1}{2} (\|x - y\|^2 + \|x - z\|^2 - \|y - z\|^2), \quad \forall x, y, z \in \mathcal{R}^n. \tag{4}$$

From now on, we denote

$$x = (x_1, x_2), \quad \theta(x) = \theta_1(x_1) + \theta_2(x_2), \quad A = (A_1, A_2).$$

Throughout this paper, we make the following assumptions.

**Assumption 2.1** The functions  $\theta_i(\cdot)$  ( $i = 1, 2$ ) are both convex.

**Assumption 2.2** There is a point  $(\hat{x}_1, \hat{x}_2) \in \text{ri}(\text{dom}(\theta_1) \times \text{dom}(\theta_2))$  such that  $A_1\hat{x}_1 + A_2\hat{x}_2 = b$ .

Then, under Assumption 2.2, it follows from Corollaries 28.2.2 and 28.3.1 in [18] that  $(x_1^*, x_2^*) \in \text{ri}(\text{dom}(\theta_1) \times \text{dom}(\theta_2))$  is an optimal solution to problem (1) iff there exists a Lagrangian multiplier  $\lambda^* \in \mathcal{R}^l$  such that  $(x_1^*, x_2^*, \lambda^*)$  is a solution of the following KKT systems:

$$\begin{cases} 0 \in \partial\theta_1(x_1^*) - A_1^\top \lambda^*, \\ 0 \in \partial\theta_2(x_2^*) - A_2^\top \lambda^*, \\ A_1 x_1^* + A_2 x_2^* = b. \end{cases} \tag{5}$$

The set of the solutions of (5) is denoted by  $\mathcal{W}^*$ . By Assumption 2.1, (3), and (5), for any  $(x^*, \lambda^*) = (x_1^*, x_2^*, \lambda^*) \in \mathcal{W}^*$ , we have the following useful inequality:

$$\theta(x) - \theta(x^*) - \langle \lambda^*, Ax - b \rangle \geq 0, \quad \forall x = (x_1, x_2) \in \mathcal{R}^{m_1+m_2}. \tag{6}$$

**Assumption 2.3** The solution set  $\mathcal{W}^*$  of the KKT systems (5) is nonempty, and at least one  $(x_1^*, x_2^*, \lambda^*) \in \mathcal{W}^*$  with  $\lambda^* \neq 0$ .

### 3 Convergence rate of the P-ADMM

In this section, we aim to prove the convergence rate of the P-ADMM, and to accomplish this, we need to make some restrictions of the matrices  $A_i, G_i$  ( $i = 1, 2$ ) included in the P-ADMM as follows.

**Assumption 3.1** (1) The matrix  $G_1 \succeq 0$ , and  $A_1$  is full-column rank if  $G_1 = 0$ .

(2) The matrix  $G_2$  is set as  $G_2 = \alpha\tau I_{n_2} - \beta A_2^\top A_2$  with  $\tau > \beta \|A_2^\top A_2\|$ ,  $\alpha \in (0, 1]$ , and  $\alpha \geq (5 - \min\{\gamma, 1 + \gamma - \gamma^2\})/5$ .

**Remark 3.1** In [14], the parameter  $\alpha$  can take any value of the interval  $[0.8, 1)$ . Obviously, the parameter  $\alpha$  in this paper can also obtain the lower bound 0.8 if  $\gamma = 1$ .

Let us introduce some matrices to simplify our notation in the subsequent analysis. More specifically, we set

$$G = \begin{pmatrix} G_1 & 0 \\ 0 & G_2 \end{pmatrix}$$

and

$$\begin{aligned} \bar{G}_2 &= \tau I_{n_2} - \beta A_2^\top A_2, & M &= G_1 + \beta A_1^\top A_1, \\ N &= G_2 + \beta A_2^\top A_2, & H &= G_2 + \beta \min\{\gamma, 1 + \gamma - \gamma^2\} A_2^\top A_2. \end{aligned} \tag{7}$$

**Remark 3.2** From Assumption 3.1 and  $\gamma \in (0, \frac{1+\sqrt{5}}{2})$ , we see that the matrices  $\bar{G}_2, M, N, H$  defined by (7) are all positive definite. However, the matrix  $G_2$  defined in Assumption 3.1 may be indefinite. For example, when  $\gamma = 1, \alpha = 0.8$ , and  $\tau = 1.01\beta\|A_2^\top A_2\|$ , then  $G_2 = -0.198\beta A_2^\top A_2$ , which is obviously indefinite if the matrix  $A_2$  is full-column rank.

**Remark 3.3** From the definitions of  $G_2$  and  $\bar{G}_2$ , we have

$$G_2 = \alpha \bar{G}_2 - (1 - \alpha)\beta A_2^\top A_2. \tag{8}$$

Now, we start proving the convergence rate of the P-ADMM under Assumptions 2.1-2.3 and Assumption 3.1. Firstly, we prove three lemmas step by step.

**Lemma 3.1** Let  $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$  be the sequence generated by the P-ADMM. Under Assumptions 2.1-2.3, for any  $(x_1, x_2, \lambda) \in \mathcal{R}^{m+n_2+l}$  such that  $A_1 x_1 + A_2 x_2 = b$ , we have

$$\begin{aligned} &\theta(x) - \theta(x^{k+1}) \\ &\geq \frac{1}{2} (\|x^{k+1} - x\|_G^2 - \|x^k - x\|_G^2 + \|x^{k+1} - x^k\|_G^2) + \langle Ax^{k+1} - b, -\lambda \rangle + \frac{2-\gamma}{2\beta\gamma^2} \|\lambda^{k+1} - \lambda^k\|^2 \\ &\quad + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) + \beta \langle Ax^{k+1} - b, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &\quad + \frac{\beta}{2} (\|A_2 x_2^{k+1} - A_2 x_2\|^2 - \|A_2 x_2^k - A_2 x_2\|^2 + \|A_2 x_2^{k+1} - A_2 x_2^k\|^2). \end{aligned} \tag{9}$$

*Proof* Note that the optimality condition for the first subproblem (i.e., the subproblem with respect to  $x_1$ ) in (2) is

$$\begin{aligned} 0 &= \nabla\theta_1(x_1^{k+1}) + \beta A_1^\top \left( A_1 x_1^{k+1} + A_2 x_2^k - b - \frac{1}{\beta} \lambda^k \right) + G_1(x_1^{k+1} - x_1^k) \\ &= \nabla\theta_1(x_1^{k+1}) - A_1^\top \tilde{\lambda}^k + \beta A_1^\top A_2(x_2^k - x_2^{k+1}) + G_1(x_1^{k+1} - x_1^k), \end{aligned} \tag{10}$$

where  $\nabla\theta_1(x_1^{k+1})$  is a subgradient of  $\theta_1(\cdot)$  at  $x_1^{k+1}$ ,  $\tilde{\lambda}^k = \lambda^k - \beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b)$ , and the second equality uses the updating formula for  $\lambda$  in (2). Then (10) can be rewritten as

$$\begin{aligned} 0 &= \langle x_1^{k+1} - x_1, \nabla\theta_1(x_1^{k+1}) \rangle + \langle x_1^{k+1} - x_1, G_1(x_1^{k+1} - x_1^k) + \beta A_1^\top A_2(x_2^k - x_2^{k+1}) \rangle \\ &\quad + \langle A_1(x_1^{k+1} - x_1), -\tilde{\lambda}^k \rangle \\ &\geq \theta_1(x_1^{k+1}) - \theta_1(x_1) + \langle x_1^{k+1} - x_1, G_1(x_1^{k+1} - x_1^k) + \beta A_1^\top A_2(x_2^k - x_2^{k+1}) \rangle \\ &\quad + \langle A_1(x_1^{k+1} - x_1), -\tilde{\lambda}^k \rangle, \end{aligned} \tag{11}$$

where the inequality comes from the convexity of  $\theta_1(\cdot)$  and (3). Similarly, the optimality condition for the second subproblem (*i.e.*, the subproblem with respect to  $x_2$ ) in (2) gives

$$\begin{aligned} 0 &= \nabla\theta_2(x_2^{k+1}) + \beta A_2^\top \left( A_1 x_1^{k+1} + A_2 x_2^{k+1} - b - \frac{1}{\beta} \lambda^k \right) + G_2(x_2^{k+1} - x_2^k) \\ &= \nabla\theta_1(x_1^{k+1}) - A_2^\top \tilde{\lambda}^k + G_2(x_2^{k+1} - x_2^k), \end{aligned}$$

*i.e.*,

$$\begin{aligned} 0 &= \langle x_2^{k+1} - x_2, \nabla\theta_2(x_2^{k+1}) \rangle + \langle A_2(x_2^{k+1} - x_2), -\tilde{\lambda}^k \rangle + \langle x_2^{k+1} - x_2, G_2(x_2^{k+1} - x_2^k) \rangle \\ &\geq \theta_2(x_2^{k+1}) - \theta_2(x_2) + \langle x_2^{k+1} - x_2, G_2(x_2^{k+1} - x_2^k) \rangle + \langle A_2(x_2^{k+1} - x_2), -\tilde{\lambda}^k \rangle, \end{aligned} \tag{12}$$

where the inequality follows from the convexity of  $\theta_2(\cdot)$  and (3). Then, adding (11) and (12), we obtain

$$\begin{aligned} &\theta(x) - \theta(x^{k+1}) \\ &\geq \sum_{i=1}^2 \langle x_i^{k+1} - x_i, G_i(x_i^{k+1} - x_i^k) \rangle + \sum_{i=1}^2 \langle A_i(x_i^{k+1} - x_i), -\tilde{\lambda}^k \rangle \\ &\quad + \beta \langle A_1 x_1^{k+1} - A_1 x_1, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &= \sum_{i=1}^2 \langle x_i^{k+1} - x_i, G_i(x_i^{k+1} - x_i^k) \rangle + \langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle + \beta \langle A_1 x_1^{k+1} - A_1 x_1, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &= \sum_{i=1}^2 \langle x_i^{k+1} - x_i, G_i(x_i^{k+1} - x_i^k) \rangle + \langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle \\ &\quad + \beta \langle A_1 x_1^{k+1} + A_2 x_2^{k+1} - A_1 x_1 - A_2 x_2, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &\quad + \beta \langle A_2 x_2 - A_2 x_2^{k+1}, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &= \sum_{i=1}^2 \langle x_i^{k+1} - x_i, G_i(x_i^{k+1} - x_i^k) \rangle + \langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle \\ &\quad + \beta \langle Ax^{k+1} - b, A_2 x_2^k - A_2 x_2^{k+1} \rangle + \beta \langle A_2 x_2 - A_2 x_2^{k+1}, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &= \frac{1}{2} \sum_{i=1}^2 (\|x_i^{k+1} - x_i\|_{G_i}^2 - \|x_i^k - x_i\|_{G_i}^2 + \|x_i^{k+1} - x_i^k\|_{G_i}^2) + \langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle \\ &\quad + \beta \langle Ax^{k+1} - b, A_2 x_2^k - A_2 x_2^{k+1} \rangle \\ &\quad + \frac{\beta}{2} (\|A_2 x_2^{k+1} - A_2 x_2\|^2 - \|A_2 x_2^k - A_2 x_2\|^2 + \|A_2 x_2^{k+1} - A_2 x_2^k\|^2), \end{aligned} \tag{13}$$

where the last equality comes from the identity (4). Now, let us deal with the term  $\langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle$  on the right side of (13). Specifically, from the updating formula for  $\lambda$  in (2) again, we can get

$$\begin{aligned} &\langle Ax^{k+1} - b, -\tilde{\lambda}^k \rangle \\ &= \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{\beta} \langle \lambda^k - \tilde{\lambda}^k, \lambda - \tilde{\lambda}^k \rangle \end{aligned}$$

$$\begin{aligned}
 &= \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{\beta} \left\langle \lambda^k - \frac{1}{\gamma} \lambda^{k+1} + \frac{1-\gamma}{\gamma} \lambda^k, \lambda - \frac{1}{\gamma} \lambda^{k+1} + \frac{1-\gamma}{\gamma} \lambda^k \right\rangle \\
 &= \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{\beta} \left\langle \frac{1}{\gamma} \lambda^k - \frac{1}{\gamma} \lambda^{k+1}, \lambda - \lambda^k + \frac{1}{\gamma} \lambda^k - \frac{1}{\gamma} \lambda^{k+1} \right\rangle \\
 &= \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{\beta\gamma} \langle \lambda^k - \lambda^{k+1}, \lambda - \lambda^k \rangle + \frac{1}{\beta\gamma^2} \|\lambda^{k+1} - \lambda^k\|^2 \\
 &= \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) + \frac{2-\gamma}{2\beta\gamma^2} \|\lambda^{k+1} - \lambda^k\|^2, \tag{14}
 \end{aligned}$$

where the second equality comes from  $\lambda^{k+1} = \lambda^k - \gamma(\lambda^k - \tilde{\lambda}^k)$ , and the last equality uses the identity (4). Then, substituting (14) into (13) yields (9). This completes the proof.  $\square$

The following lemma aims to further refine the crossing term  $\beta \langle Ax^{k+1} - b, A_2x_2^k - A_2x_2^{k+1} \rangle$  on the right side of (9).

**Lemma 3.2** *Let  $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$  be the sequence generated by the P-ADMM. Under Assumptions 2.1-2.3, for any  $(x_1, x_2, \lambda) \in \mathcal{R}^{m+n_2+l}$  such that  $A_1x_1 + A_2x_2 = b$ , we have*

$$\begin{aligned}
 &\theta(x) - \theta(x^{k+1}) \\
 &\geq \frac{1}{2} (\|x^{k+1} - x\|_G^2 - \|x^k - x\|_G^2 + \|x^{k+1} - x^k\|_G^2) + \langle Ax^{k+1} - b, -\lambda \rangle \\
 &\quad + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) + \frac{2-\gamma}{2\beta\gamma^2} \|\lambda^{k+1} - \lambda^k\|^2 \\
 &\quad + (1-\gamma)\beta \langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle + \|x_2^{k+1} - x_2^k\|_{G_2}^2 \\
 &\quad + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k-1}) \rangle \\
 &\quad + \frac{\beta}{2} (\|A_2x_2^{k+1} - A_2x_2\|^2 - \|A_2x_2^k - A_2x_2\|^2 + \|A_2x_2^{k+1} - A_2x_2^k\|^2). \tag{15}
 \end{aligned}$$

*Proof* Setting  $x_2 = x_2^k$  in (12), we get

$$0 \geq \theta_2(x_2^{k+1}) - \theta_2(x_2^k) + \langle x_2^{k+1} - x_2^k, G_2(x_2^{k+1} - x_2^k) \rangle + \langle A_2(x_2^{k+1} - x_2^k), -\tilde{\lambda}^k \rangle.$$

That is,

$$\theta_2(x_2^k) - \theta_2(x_2^{k+1}) - \langle A_2(x_2^{k+1} - x_2^k), -\tilde{\lambda}^k \rangle \geq \|x_2^{k+1} - x_2^k\|_{G_2}^2. \tag{16}$$

Similarly, taking  $x_2 = x_2^{k+1}$  in (12) for  $k := k - 1$ , and thus we have

$$0 \geq \theta_2(x_2^k) - \theta_2(x_2^{k+1}) + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k-1}) \rangle + \langle A_2(x_2^k - x_2^{k+1}), -\tilde{\lambda}^{k-1} \rangle.$$

That is,

$$\theta_2(x_2^{k+1}) - \theta_2(x_2^k) - \langle A_2(x_2^k - x_2^{k+1}), -\tilde{\lambda}^{k-1} \rangle \geq \|x_2^k - x_2^{k+1}\|_{G_2}^2. \tag{17}$$

Adding (16) and (17), we obtain

$$\langle A_2(x_2^k - x_2^{k+1}), \tilde{\lambda}^{k-1} - \tilde{\lambda}^k \rangle \geq \|x_2^{k+1} - x_2^k\|_{G_2}^2 + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k-1}) \rangle. \tag{18}$$

We have

$$\begin{aligned} & \tilde{\lambda}^{k-1} - \tilde{\lambda}^k \\ &= \tilde{\lambda}^{k-1} - [\lambda^k - \beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b)] \\ &= \tilde{\lambda}^{k-1} - [\lambda^{k-1} - \gamma\beta(A_1x_1^k + A_2x_2^k - b) - \beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b)] \\ &= -(1 - \gamma)\beta(A_1x_1^k + A_2x_2^k - b) + \beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b). \end{aligned}$$

Substituting the above equality into (18), we obtain

$$\begin{aligned} & \beta\langle A_2(x_2^k - x_2^{k+1}), A_1x_1^{k+1} + A_2x_2^{k+1} - b \rangle \\ & \geq (1 - \gamma)\beta\langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle + \|x_2^{k+1} - x_2^k\|_{G_2}^2 \\ & \quad + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k+1}) \rangle. \end{aligned} \tag{19}$$

Then, substituting (19) into (9) yields (15). The proof is completed.

Now, let us deal with the term  $\frac{2-\gamma}{2\beta\gamma^2}\|\lambda^{k+1} - \lambda^k\|^2 + (1 - \gamma)\beta\langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle + \frac{\beta}{2}\|A_2x_2^{k+1} - A_2x_2^k\|^2$  on the right side of (15). □

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

$$\begin{aligned} & \frac{2 - \gamma}{2\beta\gamma^2}\|\lambda^{k+1} - \lambda^k\|^2 + (1 - \gamma)\beta\langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle + \frac{\beta}{2}\|A_2x_2^{k+1} - A_2x_2^k\|^2 \\ & \geq \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\}(\|\lambda^{k+1} - \lambda^k\|^2 - \|\lambda^k - \lambda^{k-1}\|^2) \\ & \quad + \frac{\beta\varrho}{2} \left( \frac{1}{\gamma^3\beta^2}\|\lambda^{k+1} - \lambda^k\|^2 + \|A_2(x_2^k - x_2^{k+1})\|^2 \right), \end{aligned} \tag{20}$$

where  $\varrho = \min\{\gamma, 1 + \gamma - \gamma^2\}$ .

*Proof* Obviously, by the updating formula for  $\lambda$  in (2), we have

$$\langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle = \frac{1}{\beta\gamma}\langle A_2(x_2^k - x_2^{k+1}), \lambda^{k-1} - \lambda^k \rangle. \tag{21}$$

Then, applying the Cauchy-Schwartz inequality, we can get

$$\begin{cases} (1 - \gamma)\beta\langle A_2(x_2^k - x_2^{k+1}), (\lambda^{k-1} - \lambda^k)/(\beta\gamma) \rangle \\ \geq -\frac{(1-\gamma)\beta}{2}(\|A_2(x_2^k - x_2^{k+1})\|^2 + \frac{1}{\gamma^2\beta^2}\|\lambda^{k-1} - \lambda^k\|^2) \\ \text{if } \gamma \in (0, 1], \\ (1 - \gamma)\beta\langle A_2(x_2^k - x_2^{k+1}), (\lambda^{k-1} - \lambda^k)/(\beta\gamma) \rangle \\ \geq -\frac{(\gamma-1)\beta}{2}(\gamma\|A_2(x_2^k - x_2^{k+1})\|^2 + \frac{1}{\gamma^3\beta^2}\|\lambda^{k-1} - \lambda^k\|^2) \\ \text{if } \gamma \in (1, +\infty). \end{cases}$$



Then, substituting the above two inequalities into (21), and by some simple manipulations, we obtain

$$\begin{aligned} & \frac{2-\gamma}{2\beta\gamma^2} \|\lambda^{k+1} - \lambda^k\|^2 + (1-\gamma)\beta \langle A_2(x_2^k - x_2^{k+1}), A_1x_1^k + A_2x_2^k - b \rangle + \frac{\beta}{2} \|A_2x_2^{k+1} - A_2x_2^k\|^2 \\ & \geq \frac{1}{2\beta\gamma^2} \max\{1-\gamma, 1-\gamma^{-1}\} (\|\lambda^{k+1} - \lambda^k\|^2 - \|\lambda^k - \lambda^{k-1}\|^2) \\ & \quad + \frac{\beta}{2} \min\{\gamma, 1+\gamma-\gamma^2\} \left( \frac{1}{\gamma^3\beta^2} \|\lambda^{k+1} - \lambda^k\|^2 + \|A_2(x_2^k - x_2^{k+1})\|^2 \right), \end{aligned}$$

which is the same as the assertion (20), and the lemma is thus proved.

Substituting (20) into (15), we get the following important inequality:

$$\begin{aligned} & \theta(x) - \theta(x^{k+1}) \\ & \geq \frac{1}{2} (\|x^{k+1} - x\|_G^2 - \|x^k - x\|_G^2 + \|x^{k+1} - x^k\|_G^2) + \langle Ax^{k+1} - b, -\lambda \rangle \\ & \quad + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) \\ & \quad + \frac{1}{2\beta\gamma^2} \max\{1-\gamma, 1-\gamma^{-1}\} (\|\lambda^{k+1} - \lambda^k\|^2 - \|\lambda^k - \lambda^{k-1}\|^2) \\ & \quad + \|x_2^{k+1} - x_2^k\|_{G_2}^2 + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k-1}) \rangle \\ & \quad + \frac{\beta}{2} (\|A_2x_2^{k+1} - A_2x_2\|^2 - \|A_2x_2^k - A_2x_2\|^2) \\ & \quad + \frac{\beta}{2} \min\{\gamma, 1+\gamma-\gamma^2\} \left( \frac{1}{\gamma^3\beta^2} \|\lambda^{k+1} - \lambda^k\|^2 + \|A_2(x_2^k - x_2^{k+1})\|^2 \right). \end{aligned} \tag{22}$$

Now, let us deal with all the terms related with the variable  $x_2$  on the right side of (22). From the definition of the matrices  $G_2, N$  and (8), we have

$$\begin{aligned} & \frac{1}{2} (\|x_2^{k+1} - x_2\|_{G_2+\beta A_2^T A_2}^2 - \|x_2^k - x_2\|_{G_2+\beta A_2^T A_2}^2) \\ & \quad + \frac{1}{2} \|x_2^{k+1} - x_2^k\|_{G_2+\beta \varrho A_2^T A_2}^2 \\ & \quad + \|x_2^{k+1} - x_2^k\|_{G_2}^2 + \langle x_2^k - x_2^{k+1}, G_2(x_2^k - x_2^{k-1}) \rangle \\ & = \frac{1}{2} (\|x_2^{k+1} - x_2\|_N^2 - \|x_2^k - x_2\|_N^2) + \frac{1}{2} \|x_2^{k+1} - x_2^k\|_{3\alpha \bar{G}_2 - (3-3\alpha-\varrho)\beta A_2^T A_2}^2 \\ & \quad + \alpha (x_2^k - x_2^{k+1})^T \bar{G}_2 (x_2^k - x_2^{k-1}) - (1-\alpha)\beta (x_2^k - x_2^{k+1})^T (A_2^T A_2) (x_2^k - x_2^{k-1}) \\ & \geq \frac{1}{2} (\|x_2^{k+1} - x_2\|_N^2 - \|x_2^k - x_2\|_N^2) + \frac{1}{2} \|x_2^{k+1} - x_2^k\|_{3\alpha \bar{G}_2 - (3-3\alpha-\varrho)\beta A_2^T A_2}^2 \\ & \quad - \frac{\alpha}{2} (\|x_2^k - x_2^{k+1}\|_{\bar{G}_2}^2 + \|x_2^k - x_2^{k-1}\|_{\bar{G}_2}^2) \\ & \quad - \frac{(1-\alpha)\beta}{2} (\|A_2(x_2^k - x_2^{k+1})\|^2 + \|A_2(x_2^k - x_2^{k-1})\|^2) \\ & = \frac{1}{2} (\|x_2^{k+1} - x_2\|_N^2 - \|x_2^k - x_2\|_N^2) + \frac{\alpha}{2} \|x_2^{k+1} - x_2^k\|_{\bar{G}_2}^2 \\ & \quad + \frac{\alpha}{2} (\|x_2^{k+1} - x_2^k\|_{\bar{G}_2}^2 - \|x_2^k - x_2^{k-1}\|_{\bar{G}_2}^2) \end{aligned}$$

$$\begin{aligned}
 & + \frac{(1-\alpha)\beta}{2} (\|A_2(x_2^k - x_2^{k+1})\|^2 - \|A_2(x_2^k - x_2^{k-1})\|^2) \\
 & + \frac{\beta(\varrho - 5 + 5\alpha)}{2} \|A_2(x_2^k - x_2^{k+1})\|^2.
 \end{aligned}$$

Then, substituting the above inequality into (22), we can obtain

$$\begin{aligned}
 & \theta(x) - \theta(x^{k+1}) \\
 & \geq \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{2} (\|x_1^{k+1} - x_1\|_{G_1}^2 - \|x_1^k - x_1\|_{G_1}^2 + \|x_1^{k+1} - x_1^k\|_{G_1}^2) \\
 & \quad + \frac{1}{2} (\|x_2^{k+1} - x_2\|_N^2 - \|x_2^k - x_2\|_N^2) \\
 & \quad + \frac{1}{2} (\|x_2^k - x_2^{k+1}\|_{\alpha\bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 - \|x_2^{k-1} - x_2^k\|_{\alpha\bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2) \\
 & \quad + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) \\
 & \quad + \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\} (\|\lambda^{k+1} - \lambda^k\|^2 - \|\lambda^k - \lambda^{k-1}\|^2) \\
 & \quad + \frac{\alpha}{2} \|x_2^{k+1} - x_2^k\|_{G_2}^2 + \frac{\beta(\varrho - 5 + 5\alpha)}{2} \|A_2(x_2^k - x_2^{k+1})\|^2 + \frac{\varrho}{2\gamma^3\beta} \|\lambda^{k+1} - \lambda^k\|^2 \\
 & \geq \langle Ax^{k+1} - b, -\lambda \rangle + \frac{1}{2} (\|x_1^{k+1} - x_1\|_{G_1}^2 - \|x_1^k - x_1\|_{G_1}^2) + \frac{1}{2} (\|x_2^{k+1} - x_2\|_N^2 - \|x_2^k - x_2\|_N^2) \\
 & \quad + \frac{1}{2} (\|x_2^k - x_2^{k+1}\|_{\alpha\bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 - \|x_2^{k-1} - x_2^k\|_{\alpha\bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2) \\
 & \quad + \frac{1}{2\beta\gamma} (\|\lambda^{k+1} - \lambda\|^2 - \|\lambda^k - \lambda\|^2) \\
 & \quad + \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\} (\|\lambda^{k+1} - \lambda^k\|^2 - \|\lambda^k - \lambda^{k-1}\|^2), \tag{23}
 \end{aligned}$$

where the inequality comes from  $\alpha \in (0, 1]$  and  $\alpha \geq \frac{5-\varrho}{5}$ . Based on (23), we can prove the worst-case  $\mathcal{O}(1/t)$  convergence rate in an ergodic sense of the P-ADMM.  $\square$

**Theorem 3.1** *Suppose that Assumptions 2.1-2.3 and Assumption 3.1 hold. Let  $\{(x^k, \lambda^k)\} = \{(x_1^k, x_2^k, \lambda^k)\}$  be the sequence generated by the P-ADMM and let  $\bar{x}^t = \frac{1}{t} \sum_{k=1}^t x^{k+1}$ , where  $t$  is a positive integer. Then,*

$$\begin{cases} |\theta(\bar{x}^t) - \theta(x^*)| \leq \frac{D}{t}, \\ \|A\bar{x}^t - b\| \leq \frac{D}{t\|\lambda^*\|}, \end{cases} \tag{24}$$

where  $(x^*, \lambda^*) = (x_1^*, x_2^*, \lambda^*)$  with  $\lambda^* \neq 0$  is a point satisfying the KKT conditions in (5), and  $D$  is a constant defined by

$$\begin{aligned}
 D & = \frac{1}{2} \|x^1 - x^*\|_{G_1}^2 + \frac{1}{2} \|x_2^1 - x_2^*\|_N^2 + \frac{1}{2} \|x_2^0 - x_2^1\|_{\alpha\bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 \\
 & \quad + \frac{1}{\beta\gamma} (\|\lambda^1\|^2 + \|\lambda\|^2) + \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\} \|\lambda^1 - \lambda^0\|^2. \tag{25}
 \end{aligned}$$

*Proof* Setting  $x = x^*$  in the inequality (22) and summing it over  $k = 1, 2, \dots, t$ , we obtain

$$\begin{aligned} & \sum_{k=1}^t [\theta(x^{k+1}) - \theta(x^*) - \langle Ax^{k+1} - b, \lambda \rangle] \\ & \leq \frac{1}{2} \|x^1 - x^*\|_{G_1}^2 + \frac{1}{2} \|x_2^1 - x_2^*\|_N^2 + \frac{1}{2} \|x_2^0 - x_2^1\|_{\alpha \bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 \\ & \quad + \frac{1}{2\beta\gamma} \|\lambda^1 - \lambda\|^2 \\ & \quad + \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\} \|\lambda^1 - \lambda^0\|^2 \\ & \leq \frac{1}{2} \|x^1 - x^*\|_{G_1}^2 + \frac{1}{2} \|x_2^1 - x_2^*\|_N^2 + \frac{1}{2} \|x_2^0 - x_2^1\|_{\alpha \bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 \\ & \quad + \frac{1}{\beta\gamma} (\|\lambda^1\|^2 + \|\lambda\|^2) \\ & \quad + \frac{1}{2\beta\gamma^2} \max\{1 - \gamma, 1 - \gamma^{-1}\} \|\lambda^1 - \lambda^0\|^2, \end{aligned}$$

which together with the convexity of the function  $\theta(\cdot)$  implies

$$\begin{aligned} & \theta(\bar{x}^t) - \theta(x^*) - \langle A\bar{x}^t - b, \lambda \rangle \\ & \leq \frac{1}{2t} \|x^1 - x^*\|_{G_1}^2 + \frac{1}{2t} \|x_2^1 - x_2^*\|_N^2 + \frac{1}{2t} \|x_2^0 - x_2^1\|_{\alpha \bar{G}_2 + \beta(1-\alpha)A_2^\top A_2}^2 \\ & \quad + \frac{1}{\beta\gamma t} (\|\lambda^1\|^2 + \|\lambda\|^2) \\ & \quad + \frac{1}{2\beta\gamma^2 t} \max\{1 - \gamma, 1 - \gamma^{-1}\} \|\lambda^1 - \lambda^0\|^2. \end{aligned}$$

Using the Lemmas 2.2 and 2.3 of [17] with  $\rho = 2\|\lambda^*\|$  ( $\rho$  is a parameter defined in Lemmas 2.2 and 2.3 of [17]), we can get (24). This completes the proof.  $\square$

**Remark 3.4** From (24) and (25), we can conclude that larger values of  $\gamma$  is more beneficial for accelerating the convergence of the P-ADMM, as the larger  $\gamma$ , the smaller  $D$ , which controls the upper bounds of  $|\theta(\bar{x}^t) - \theta(x^*)|$  and  $\|A\bar{x}^t - b\|$ .

#### 4 Numerical experiments

In this section, we apply the P-ADMM to solve the compressive sensing, a concrete problem of the general model (1). The codes were written by Matlab R2010a and conducted on a ThinkPad notebook with Pentium(R) Dual-Core CPU T4400@2.2 GHz, 2 GB of RAM using Windows 7.

Let us briefly review the compressive sensing. Compressive sensing (CS) is to recover a sparse signal  $\bar{x} \in \mathcal{R}^n$  from an undetermined linear system  $y = A\bar{x}$ , where  $A \in \mathcal{R}^{m \times n}$  ( $m \ll n$ ) is a linear mapping and  $y \in \mathcal{R}^m$  is an observation. An important decoding model of CS is

$$\min_{x \in \mathcal{R}^n} \frac{1}{2} \|Ax - y\|^2 + \nu \|x\|_1, \tag{26}$$

where the parameter  $\nu > 0$  is used to trade off both terms for minimization. This is a special case of the general two-block separable convex minimization model (1). In fact, setting

$x_1 = x, x_2 = x$ , (26) can be recast as

$$\begin{aligned} & \min \frac{1}{2} \|Ax_1 - y\|^2 + v\|x_2\|_1 \\ & \text{s.t. } x_1 - x_2 = 0, \\ & \quad x_1 \in \mathcal{R}^n, x_2 \in \mathcal{R}^n, \end{aligned}$$

which is a special case of (1) with

$$\begin{aligned} \theta_1(x_1) &= \frac{1}{2} \|Ax_1 - y\|^2, & \theta_2(x_2) &= v\|x_2\|_1, \\ A_1 &= I_n, & A_2 &= -I_n, & b &= 0, & \mathcal{X}_1 &= \mathcal{X}_2 = \mathcal{R}^n, \end{aligned}$$

and thus, the P-ADMM can be used to solve CS.

In our experiment, the stopping criterion of the P-ADMM is set as

$$\frac{\|f_k - f_{k-1}\|}{\|f_{k-1}\|} < 10^{-5},$$

where  $f_k$  denotes the function value of (26) at the iterate  $x_1^k$ . The initial points of  $x_1, x_2, \lambda$  are all set as  $A^\top y$ , and due to the limit of EMS memory of our computer, we only test a medium scale of (26) with  $n = 1,000, m = 300, k = 60$ , where  $k$  is the number of random nonzero elements contained in the original signal. In addition, we set

$$\bar{A} = \text{randn}(m, n), \quad [Q, R] = \text{qr}(\bar{A}', 0), \quad A = Q',$$

and  $v = 0.01, G_1 = \tau I_n - \beta A^\top A$  with  $\tau = 2, \beta = \text{mean}(\text{abs}(b))$ . In the literature, the relative error (RelErr) is usually used to measure the quality of recovered signal and is defined by

$$\text{RelErr} = \frac{\|\tilde{x} - \bar{x}\|}{\|\bar{x}\|},$$

where  $\tilde{x}$  and  $\bar{x}$  denote the recovered signal and the original signal, respectively.

First, let us illustrate the sensitivity of  $\gamma$  for the P-ADMM. We choose different values of  $\gamma$  in the interval  $[0.1, 1.6]$  (More specifically, we take  $\gamma = 0.10, 0.15, \dots, 1.60$ ). The numerical results of the objective value of (26) and the CPU time in seconds required by the P-ADMM are depicted in Figure 1, and the numerical results of the numbers of iteration and the RelErr required by P-ADMM are depicted in Figure 2.

According to the curves in Figures 1-2, we can see that the relaxation factor  $\gamma$  works well for a wide range of values and, based on this experiment, the values greater than 0.5 are more preferred.

Now let us test the effectiveness of the P-ADMM with the indefinite proximal matrix  $G_2 = (\alpha\tau - \beta)I_n$ . Here we set  $\alpha = 0.8, \tau = 1.1, \beta = 1$ , and  $\gamma = 1$ . The numerical results of one experiment are as follows: the objective value is 0.4291; the CPU time is 1.0920; the numbers of iteration is 378 and the RelErr is 5.75%. The original signal, the measurement and the signal recovered by the P-ADMM for this test scenario are given in Figure 3. The recovered results are marked by a red circle in the third subplot of Figure 3, which shows

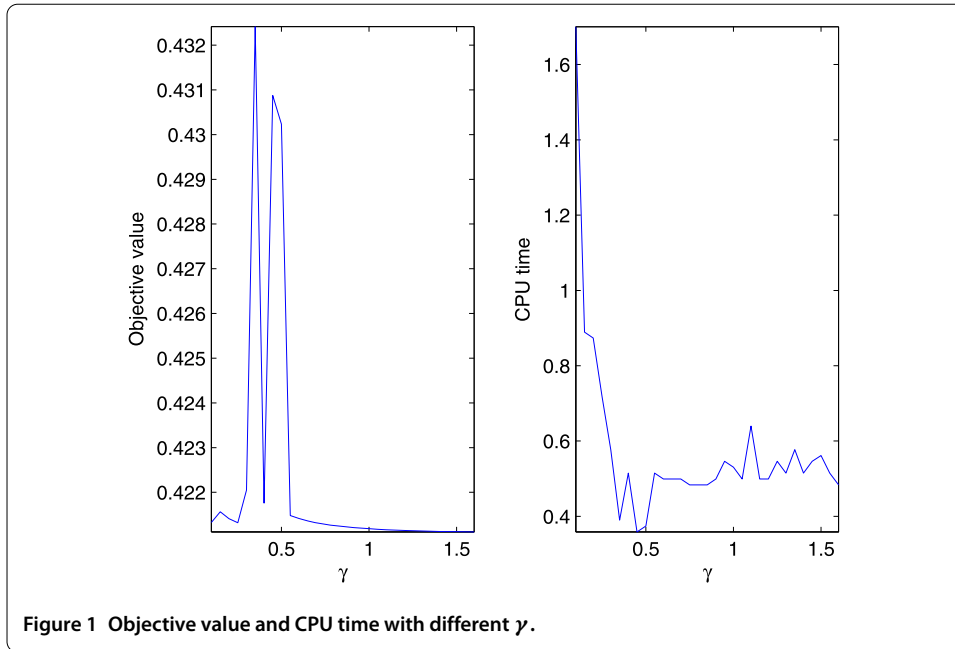


Figure 1 Objective value and CPU time with different  $\gamma$ .

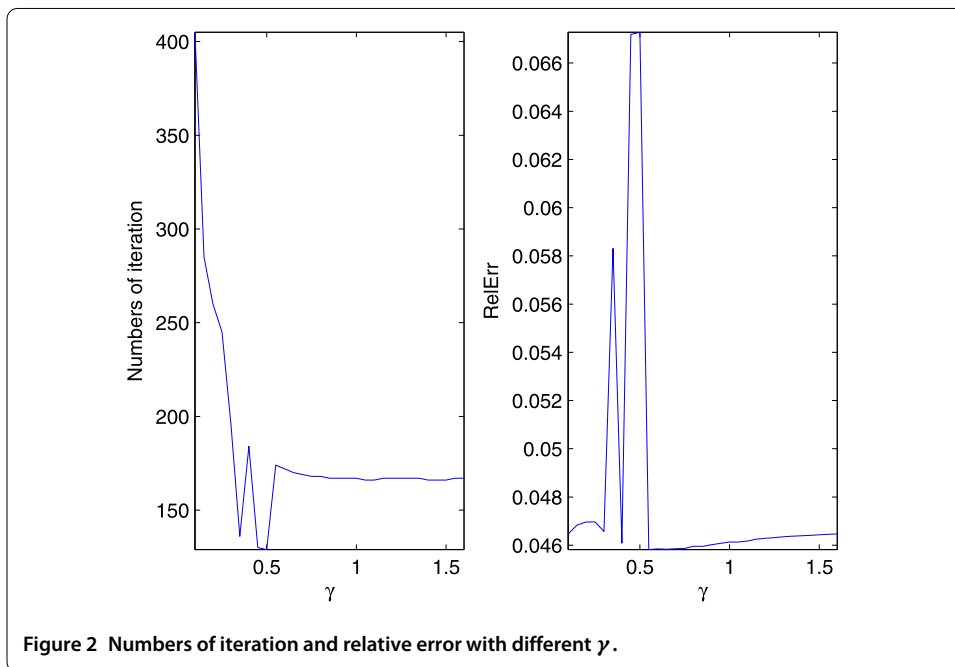


Figure 2 Numbers of iteration and relative error with different  $\gamma$ .

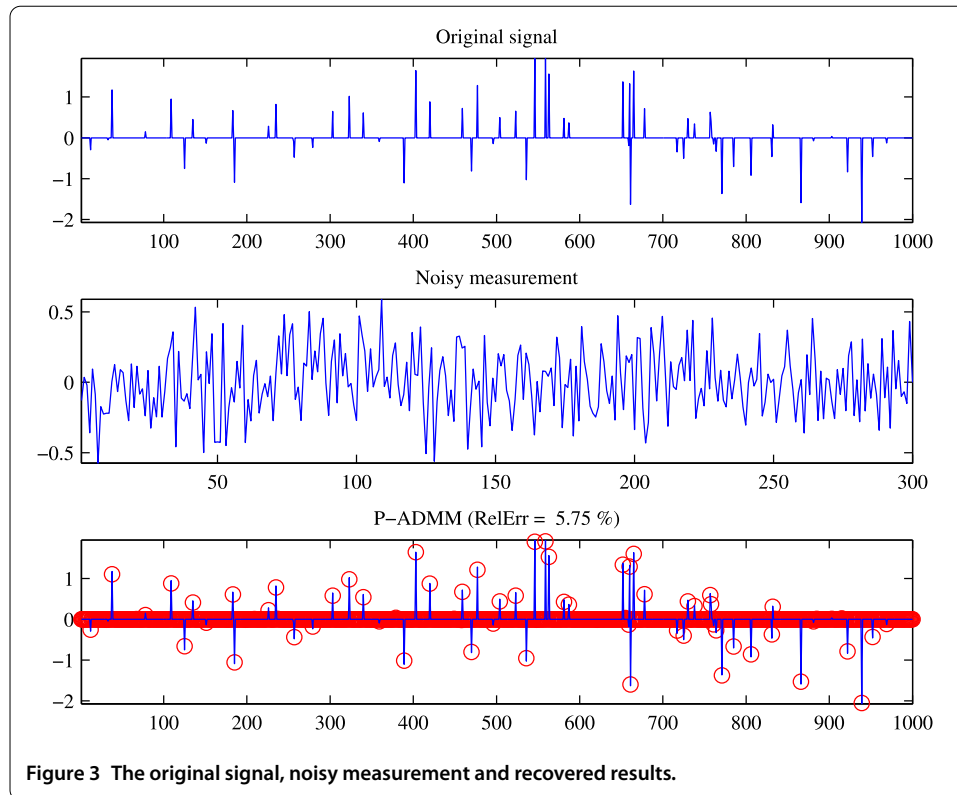
clearly that almost the original signal is recovered with high precision. This indicates that the P-ADMM is effective though the proximal matrix  $G_2$  is indefinite.

**Competing interests**

The authors declare that they have no competing interests.

**Authors' contributions**

The first author has proposed the motivations of the manuscript and the second author has proved the convergence rate. Both authors have equally contributed in the numerical results. All authors read and approved the final manuscript.



#### Author details

<sup>1</sup>School of Management, Qufu Normal University, Shandong, 276826, P.R. China. <sup>2</sup>School of Mathematics and Statistics, Zaozhuang University, Shandong, 277160, P.R. China. <sup>3</sup>School of Data Sciences, Zhejiang University of Finance and Economics, Hangzhou, 310018, China.

#### Acknowledgements

The authors gratefully acknowledge the helpful comments of the anonymous reviewers. This work is supported by the foundation of National Natural Science Foundation of China (No. 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 the Scientific Research Project of Shandong Universities (No. J15L111).

Received: 2 October 2016 Accepted: 4 January 2017 Published online: 14 January 2017

#### References

- Lin, Z, Liu, R, Li, H: Linearized alternating direction method with parallel splitting and adaptive penalty for separable convex programs in machine learning. *Mach. Learn.* **95**(2), 287-325 (2015)
- Chandrasekaran, V, Parrilo, P, Willsky, A: Latent variable graphical model selection via convex optimization. *Ann. Stat.* **40**(4), 1935-1967 (2012)
- Banerjee, O, El Ghaoui, L, d'Aspremont, A: Model selection through sparse maximum likelihood estimation for multivariate Gaussian for binary data. *J. Mach. Learn. Res.* **9**, 485-516 (2008)
- Zhou, Z, Li, X, Wright, J, Candès, E, Ma, Y: Stable principal component pursuit. In: *Proceedings of International Symposium on Information Theory* (2010)
- Peng, Y, Ganesh, A, Wright, J, Xu, W, Ma, Y: RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 2233-2246 (2012)
- Glowinski, R, Marrocco, A: Sur l'approximation, par éléments fins d'ordre  $n$ , et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet nonlinéaires. *Rev. Fr. Autom. Inform. Rech. Oper.* **9**, 41-76 (1975)
- Gabay, D, Mercier, B: A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. *Comput. Math. Appl.* **2**, 17-40 (1976)
- He, B, Liao, L, Han, D, Yang, H: A new inexact alternating directions method for monotone variational inequalities. *Math. Program., Ser. A* **92**, 103-118 (2002)
- He, B, Wang, S, Yang, H: A modified variable-penalty alternating directions method for monotone variational inequalities. *J. Comput. Math.* **21**, 495-504 (2003)
- Wang, X, Yuan, X: The linearized alternating direction method of multipliers for Dantzig selector. *SIAM J. Sci. Comput.* **34**, 2782-2811 (2012)
- Zhang, X, Burger, M, Osher, S: A unified primal-dual algorithm framework based on Bregman iteration. *J. Sci. Comput.* **6**, 20-46 (2010)

12. Sun, M, Liu, J: An inexact generalized PRSM with LQP regularization for structured variational inequalities and its applications to traffic equilibrium problems. *J. Inequal. Appl.* **2016** 150 (2016)
13. Xu, M, Wu, T: A class of linearized proximal alternating direction methods. *J. Optim. Theory Appl.* **155**, 321-337 (2011)
14. He, B, Ma, F, Yuan, X: Linearized alternating direction method of multipliers via positive-indefinite proximal regularization for convex programming. Manuscript (2016)
15. He, B, Yuan, X: On the  $\mathcal{O}(1/n)$  convergence rate of the Douglas-Rachford alternating direction method. *SIAM J. Numer. Anal.* **50**, 700-709 (2012)
16. Goldstein, T, O'Donoghue, B, Setzer, S, Baraniuk, R: Fast alternating direction optimization methods. *SIAM J. Imaging Sci.* **7**(3), 1588-1623 (2014)
17. Xu, Y: Accelerated first-order primal-dual proximal methods for linearly constrained composite convex programming. Manuscript (2016)
18. Rockafellar, R: *Convex Analysis*. Princeton University Press, Princeton (1970)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---