

## Research Article

# A Proximal Fully Parallel Splitting Method for Stable Principal Component Pursuit

Hongchun Sun,<sup>1</sup> Jing Liu,<sup>2</sup> and Min Sun<sup>3,4</sup>

<sup>1</sup>School of Mathematics and Statistics, Linyi University, Shandong 276005, China

<sup>2</sup>School of Data Sciences, Zhejiang University of Finance and Economics, Zhejiang 310018, China

<sup>3</sup>School of Mathematics and Statistics, Zaozhuang University, Shandong 277160, China

<sup>4</sup>School of Management, Qufu Normal University, Shandong 276826, China

Correspondence should be addressed to Min Sun; ziyouxiaodou@163.com

Received 30 March 2017; Revised 8 June 2017; Accepted 29 August 2017; Published 25 October 2017

Academic Editor: Laurent Bako

Copyright © 2017 Hongchun Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a special three-block separable convex programming, the stable principal component pursuit (SPCP) arises in many different disciplines, such as statistical learning, signal processing, and web data ranking. In this paper, we propose a proximal fully parallel splitting method (PFPSM) for solving SPCP, in which the resulting subproblems all admit closed-form solutions and can be solved in distributed manners. Compared with other similar algorithms in the literature, PFPSM attaches a Glowinski relaxation factor  $\eta \in (\sqrt{3}/2, 2/\sqrt{3})$  to the updating formula for its Lagrange multiplier, which can be used to accelerate the convergence of the generated sequence. Under mild conditions, the global convergence of PFPSM is proved. Preliminary computational results show that the proposed algorithm works very well in practice.

## 1. Introduction

Recovering the low-rank and sparse components of a given matrix is a fundamental task in many scientific fields, which can be modeled as the following two-block separable nonconvex and nonsmooth programming:

$$\begin{aligned} \min_{L,S} \quad & \text{rank}(L) + \tau \|S\|_0, \\ \text{s.t.} \quad & L + S = D, \end{aligned} \quad (1)$$

where  $D \in \mathcal{R}^{p \times q}$  is a given matrix,  $\text{rank}(L)$  is the rank of the matrix  $L \in \mathcal{R}^{p \times q}$ ,  $\|S\|_0$  denotes the  $\ell_0$ -norm of  $S$ , which is defined by the number of the nonzeros of its entries, and  $\tau > 0$  is a trade-off parameter balancing the low rank and sparsity. Recent studies [1–3] indicate that the NP-hard task (1) can be accurately accomplished by solving one of its convex relaxation problems: the principal component pursuit (PCP) problem, in which the rank and the  $\ell_0$ -norm are replaced by their tightest convex relaxations, that is, the nuclear norm and the  $\ell_1$ -norm, respectively. The PCP

problem can be mathematically modeled as the following optimization problem:

$$\begin{aligned} \min_{L,S} \quad & \|L\|_* + \tau \|S\|_1, \\ \text{s.t.} \quad & L + S = D, \end{aligned} \quad (2)$$

where the nuclear norm  $\|L\|_*$  is the sum of all singular values of  $L$ , which is to induce the low-rank component of the given matrix  $D$ , and the  $\ell_1$ -norm  $\|S\|_1$  is the sum of the absolute values of all entries of  $S$ , which is to induce the sparse component of the given matrix  $D$ .

To capture even more applications, Zhou et al. [4] considered that the observation matrix  $D$  is corrupted by both small entry-wise noise and gross sparse errors. Then more general case was studied by Tao and Yuan [5] wherein only a fraction of entries of the given matrix  $D$  can be observed, and they suggested recovering the low-rank and sparse components of  $D$  by solving the following stable principal component pursuit (SPCP) problem:

$$\begin{aligned} \min_{L,S} \quad & \|L\|_* + \tau \|S\|_1, \\ \text{s.t.} \quad & \|P_\Omega(D - L - S)\| \leq \delta, \end{aligned} \quad (3)$$

where  $\delta > 0$  is related to the impulsive or Gaussian noise level,  $\|\cdot\|$  denotes the Frobenius-norm,  $\Omega \subseteq \{1, 2, \dots, p\} \times \{1, 2, \dots, q\}$  is the index set of the observable entries of  $D$ , and  $P_\Omega : \mathcal{R}^{p \times q} \rightarrow \mathcal{R}^{p \times q}$  is the projection operator defined by

$$[P_\Omega(X)]_{ij} = \begin{cases} X_{ij}, & \text{if } (i, j) \in \Omega \\ 0, & \text{otherwise,} \end{cases} \quad (4)$$

$$1 \leq i \leq p, 1 \leq j \leq q.$$

In [6], Hou et al. transformed (3) into the following unconstrained problem:

$$\min_{L,S} \left\{ \|L\|_* + \tau \|S\|_1 + \frac{1}{2\mu} \|P_\Omega(D - L - S)\|^2 \right\}, \quad (5)$$

where  $\mu > 0$  is a penalty parameter. Problems (3) and (5) are both convex optimization programming, which can be equivalently reformulated as some semidefinite programming (SDP) problems and then solved by using some state-of-the-art SDP solvers in polynomial time, such as SeDuMi [7] and SDPT3 [8]. However, these solvers are not applicable to high-dimensional (3) and (5), because a linear system needs to be solved (in)exactly at each iteration, which may be more costly and require very large amount of memory in actual implementations [9].

By introducing an auxiliary variable  $U \in \mathcal{R}^{p \times q}$ , the model (5) can be rewritten as the following equivalent form:

$$\begin{aligned} \min_{L,S,U} \quad & \left\{ \|L\|_* + \tau \|S\|_1 + \frac{1}{2\mu} \|P_\Omega(U)\|^2 \right\}, \\ \text{s.t.} \quad & L + S + U = D, \end{aligned} \quad (6)$$

which decouples the term  $\|P_\Omega(D - L - S)\|^2$  of (5). Recently, there has been an intensive study on the iteration methods for solving (6) [5, 6, 10–12]. Interestingly, these methods usually only exploit the first-order information of the objective function of (6), and most of them are based on the well-known *alternating direction method with multipliers* (ADMM), such as the sequential ADMM [11], the partially parallel ADMMs [5, 6, 10, 12], and the fully parallel ADMMs [13, 14]. As an influential first-order iteration method in optimization, ADMM was independently proposed by Glowinski and Marrocco [15] and Gabay and Mercier [16] about forty years ago, and due to its high efficiency in the numerical simulation, ADMM has been receiving considerable attention in recent years. ADMM can be viewed as an application of the Douglas-Rachford splitting method to the dual of the two-block separable convex programming (2) [17] or a special case of the proximal point method for the general convex programming [18], or a generalization of the classical Uzawa method for solving the saddle-point problems [19]. Chen et al. [20] showed that the sequential ADMM cannot directly extend to three-block separable convex programming by a simple

counterexample, and similarly He et al. [13] showed that the full parallel ADMM may be divergent even for two-block separable convex programming. There are three strategies to deal with the issue: (1) the substitution procedure [13, 21]; (2) the technique of regularization [14, 22]; (3) the technique of small step size [23]. For other recent developments of ADMM, including the (sub)linear convergence rate, various acceleration techniques, indefinite regularized matrices, larger step size, and its generalization to solve nonconvex and nonsmooth multiblock separable programming, we refer to [24–28].

In this paper, we focus the attention on the fully parallel ADMM, which is based on the substitution procedure and is more suitable for distributed computation. Following the full Jacobian decomposition of the augmented Lagrangian method (FJDALM) in [13], we shall propose a new fully parallel ADMM, termed as the proximal fully parallel splitting method (PFPSM) in the following analysis. Compared to other similar methods in the literature, the new iteration method has three properties. The first one is that all of its subproblems are regularized by some proximal terms, which can enhance its numerical efficiency in practice. The second one is that a Glowinski relaxation factor  $\eta \in (\sqrt{3}/2, 2/\sqrt{3})$  is attached to the updating formula for its Lagrange multiplier, and larger values of  $\eta$  often can accelerate the convergence of the proposed method (as to be reported in Section 4). To the best of our knowledge, this is the first fully parallel ADMM-type method with the Glowinski relaxation factor for solving problem (6). The third one is that it dynamically updates the step size  $\alpha_k$ .

The remainder of this paper is organized as follows. In Section 2, we present some necessary notations and useful lemmas. In Section 3, we propose the proximal fully parallel splitting method for solving problem (6) and prove its global convergence step by step. Some numerical results about the proposed method are shown and discussed in Section 4. Finally, in Section 5, some concluding remarks are drawn and several future research directions are discussed also.

## 2. Preliminaries

In this section, we briefly introduce some necessary notations, definitions, and lemmas.

For any matrix  $A = (a_{ij}) \in \mathcal{R}^{p \times q}$ , let  $\|A\| = \sqrt{\sum_{i=1}^p \sum_{j=1}^q a_{ij}^2}$  denote the Frobenius-norm of  $A$ . For any two matrices  $A, B \in \mathcal{R}^{p \times q}$ ,  $\langle A, B \rangle := \text{Trace}(A^T B)$  denotes their Frobenius inner product. Then, we have  $\langle A, B \rangle \leq \|A\| \|B\|$ . If  $G \in \mathcal{R}^{p \times p}$  is symmetric positive definite matrix, we denote by  $\|A\|_G = \sqrt{\langle A, GA \rangle} = (\text{Trace}(A^T GA))^{1/2}$  the  $G$ -norm of the matrix  $A$ . The effective domain of a function  $f : \mathcal{X} \rightarrow (-\infty, +\infty]$  is defined as  $\text{dom}(f) := \{x \in \mathcal{X} \mid f(x) < +\infty\}$ . The set of all relative interior points of a given nonempty convex set  $\mathcal{C}$  is denoted by  $\text{ri}(\mathcal{C})$ .  $\text{diag}\{A_1, A_2, \dots, A_m\}$  has  $A_i$  as its  $i$ th block on the diagonal.

A function  $f : \mathcal{R}^n \rightarrow \mathcal{R}$  is convex if and only if

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad (7)$$

$$\forall x, y \in \mathcal{R}^n, \alpha \in [0, 1].$$

Then, for a convex function  $f : \mathcal{R}^n \rightarrow \mathcal{R}$ , we have the following basic inequality:

$$f(x) \geq f(y) + \langle \xi, x - y \rangle, \quad (8)$$

$$\forall x, y \in \mathcal{R}^n, \xi \in \partial f(y),$$

where  $\partial f(y) = \{\xi \in \mathcal{R}^n : f(\bar{y}) \geq f(y) + \langle \xi, \bar{y} - y \rangle\}$ , for all  $\bar{y} \in \mathcal{R}^n$  denotes the subdifferential of  $f(\cdot)$  at the point  $y$ . The above definition and property (8) can be easily extended to the matrix-valued function.

For the convenience of our analysis, let us relabel  $L$  as  $X_1$ ,  $S$  as  $X_2$ , and  $U$  as  $X_3$  and denote  $\theta_1(X_1) = \|L\|_*$ ,  $\theta_2(X_2) = \tau \|S\|_1$ ,  $\theta_3(X_3) = (1/2\mu) \|P_\Omega(U)\|^2$ . Then  $\theta_i(\cdot)$  ( $i = 1, 2, 3$ ) are all convex function.

We make the following standard assumption about problem (6) in this paper.

*Assumption 1.* The generalized Slater's condition holds; that is, there is a point  $(\widehat{X}_1, \widehat{X}_2, \widehat{X}_3) \in \text{ri}(\text{dom}(\theta_1) \times \text{dom}(\theta_2) \times \text{dom}(\theta_3)) \cap P$ , where the set  $P$  is defined by

$$P = \{(X_1, X_2, X_3) \mid X_1 + X_2 + X_3 = D\}. \quad (9)$$

Under Assumption 1, it follows from Theorems 3.22 and 3.23 of [29] that  $(X_1^*, X_2^*, X_3^*)$  is an optimal solution of problem (6) if and only if there exists a matrix  $\Lambda^* \in \mathcal{R}^{p \times q}$  such that  $(X_1^*, X_2^*, X_3^*, \Lambda^*)$  is a solution of the following KKT systems:

$$0 \in \partial \theta_i(X_i^*) - \Lambda^*, \quad i = 1, 2, 3, \quad (10)$$

$$X_1^* + X_2^* + X_3^* = D.$$

The following lemma is the basis to reformulate (10) as a mixed variational inequality problem.

**Lemma 2.** For any two matrices  $X_i^* \in \mathcal{R}^{p \times q}$ ,  $\Lambda^* \in \mathcal{R}^{p \times q}$ , the relationship  $0 \in \partial \theta_i(X_i^*) - \Lambda^*$  is equivalent to the following mixed variational inequality problem:

$$\theta_i(X_i) - \theta_i(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \quad (11)$$

$$\forall X_i \in \mathcal{R}^{p \times q}.$$

*Proof.* From  $0 \in \partial \theta_i(X_i^*) - \Lambda^*$ , one has

$$\Lambda^* \in \partial \theta_i(X_i^*), \quad (12)$$

which together with the convexity of  $\theta_i(\cdot)$  and the subgradient inequality (8) implies

$$\theta_i(X_i) - \theta_i(X_i^*) \geq \langle X_i - X_i^*, \Lambda^* \rangle, \quad \forall X_i \in \mathcal{R}^{p \times q}; \quad (13)$$

that is,

$$\theta_i(X_i) - \theta_i(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \quad (14)$$

$$\forall X_i \in \mathcal{R}^{p \times q}.$$

Conversely, from  $\theta_i(X_i) - \theta_i(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0$ ,  $\forall X_i \in \mathcal{R}^{p \times q}$ , one has

$$\theta_i(X_i) + \langle X_i, -\Lambda^* \rangle \geq \theta_i(X_i^*) + \langle X_i^*, -\Lambda^* \rangle, \quad (15)$$

$$\forall X_i \in \mathcal{R}^{p \times q},$$

which implies

$$X_i^* = \operatorname{argmin}_{X_i \in \mathcal{R}^{p \times q}} \{\theta_i(X_i) + \langle X_i, -\Lambda^* \rangle\}. \quad (16)$$

From this and Theorem 3.22 of [29], we have  $0 \in \partial \theta_i(X_i^*) - \Lambda^*$ . The proof is completed.  $\square$

*Remark 3.* Based on (10) and Lemma 2, the matrix  $X^* = (X_1^*, X_2^*, X_3^*)$  is an optimal solution to problem (6) if and only if there exists a matrix  $\Lambda^* \in \mathcal{R}^{p \times q}$  such that

$$\theta_i(X_i) - \theta_i(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \quad (17)$$

$$\forall X_i \in \mathcal{R}^{p \times q}, \quad i = 1, 2, 3;$$

$$X_1^* + X_2^* + X_3^* = D.$$

Moreover, any  $\Lambda^* \in \mathcal{R}^{p \times q}$  satisfying (17) is an optimal solution to the dual of problem (6).

Furthermore, we set

$$X = (X_1, X_2, X_3),$$

$$\theta(X) = \theta_1(X_1) + \theta_2(X_2) + \theta_3(X_3), \quad (18)$$

$$W = (X, \Lambda).$$

Obviously, (17) can be written as the following mixed variational inequality problem, abbreviated as  $\text{VI}(\mathcal{W}, F, \theta)$ : find a matrix  $W^* = (X_1^*, X_2^*, X_3^*, \Lambda^*) \in \mathcal{W}$  such that

$$\theta(X) - \theta(X^*) + \langle W - W^*, F(W^*) \rangle \geq 0, \quad (19)$$

$$\forall W \in \mathcal{W},$$

where  $\mathcal{W} = \mathcal{R}^{p \times q} \times \mathcal{R}^{p \times q} \times \mathcal{R}^{p \times q} \times \mathcal{R}^{p \times q}$ , and

$$F(W) = \begin{pmatrix} -\Lambda \\ -\Lambda \\ -\Lambda \\ X_1 + X_2 + X_3 - D \end{pmatrix} \quad (20)$$

$$= \begin{pmatrix} 0 & 0 & 0 & -I \\ 0 & 0 & 0 & -I \\ 0 & 0 & 0 & -I \\ I & I & I & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ \Lambda \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 0 \\ D \end{pmatrix}.$$

The set of the solutions of  $\text{VI}(\mathcal{W}, F, \theta)$ , denoted by  $\mathcal{W}^*$ , is nonempty by Assumption 1 and Remark 3. Because  $F(W)$  is a linear mapping and its coefficient matrix is skew-symmetric, it satisfies the following nice property:

$$\langle W' - W, F(W') - F(W) \rangle = 0, \quad \forall W', W \in \mathcal{W}. \quad (21)$$

### 3. The Proximal Fully Parallel Splitting Method

In this section, we shall present the proximal fully parallel splitting method for solving (6) and discuss its convergence results.

Firstly, the augmented Lagrangian function of the convex programming (6) can be written as

$$\begin{aligned} \mathcal{L}_\beta(X_1, X_2, X_3, \Lambda) &= \sum_{i=1}^3 \theta_i(X_i) \\ &\quad - \langle \Lambda, X_1 + X_2 + X_3 - D \rangle \\ &\quad + \frac{\beta}{2} \|X_1 + X_2 + X_3 - D\|^2, \end{aligned} \quad (22)$$

where  $\beta > 0$  is the penalty parameter and  $\Lambda \in \mathcal{R}^{p \times q}$  is the Lagrange multiplier for the linear constraint of (6). To simplify our notation in later analysis, we define a block diagonal matrix as follows:

$$G = \text{diag} \left\{ \beta I_p + R_1, \beta I_p + R_2, \beta I_p + R_3, \frac{I_p}{(\eta\beta)} \right\}, \quad (23)$$

where  $R_i \in \mathcal{R}^{p \times p}$  ( $i = 1, 2, 3$ ) are some symmetric positive semidefinite matrices and  $\eta \in (\sqrt{3}/2, 2/\sqrt{3})$  is a predefined parameter. Note that the matrix  $G$  is symmetric positive definite.

*Algorithm 4* (proximal fully parallel splitting method (termed as PFPSM)).

*Step 0.* Given the tolerance  $\varepsilon > 0$ , two constants  $\eta \in (\sqrt{3}/2, 2/\sqrt{3})$ ,  $\gamma \in (0, 2)$ , and the initial iterate  $W^0 = (X_1^0, X_2^0, X_3^0, \Lambda^0) \in \mathcal{W}$ . Set  $k := 0$ .

*Step 1* (prediction step). Compute the temporary iterate  $\widetilde{W}^k = (\widetilde{X}_1^k, \widetilde{X}_2^k, \widetilde{X}_3^k, \widetilde{\Lambda}^k)$  by

$$\begin{aligned} &\widetilde{X}_1^k \\ &:= \arg \min \left\{ \mathcal{L}_\beta(X_1, X_2^k, X_3^k, \Lambda^k) + \frac{1}{2} \|X_1 - X_1^k\|_{R_1}^2 \right\}, \\ &\widetilde{X}_2^k \\ &:= \arg \min \left\{ \mathcal{L}_\beta(X_1^k, X_2, X_3^k, \Lambda^k) + \frac{1}{2} \|X_2 - X_2^k\|_{R_2}^2 \right\}, \quad (24) \\ &\widetilde{X}_3^k \\ &:= \arg \min \left\{ \mathcal{L}_\beta(X_1^k, X_2^k, X_3, \Lambda^k) + \frac{1}{2} \|X_3 - X_3^k\|_{R_3}^2 \right\}, \\ &\widetilde{\Lambda}^k := \Lambda^k - \eta\beta \left( \sum_{i=1}^3 \widetilde{X}_i^k - D \right). \end{aligned}$$

*Step 2* (convergence verification). If the stopping criterion

$$\max \{ \|X_1^k - \widetilde{X}_1^k\|, \|X_2^k - \widetilde{X}_2^k\|, \|X_3^k - \widetilde{X}_3^k\|, \|\Lambda^k - \widetilde{\Lambda}^k\| \} < \varepsilon \quad (25)$$

is satisfied, then stop.

*Step 3* (correction step). Generate the new iterate by

$$W^{k+1} = W^k - \gamma\alpha_k Q^{-1}G(W^k - \widetilde{W}^k), \quad (26)$$

where  $Q$  is any symmetric positive definite matrix and  $\alpha_k$  is the step size defined by

$$\alpha_k = \frac{\varphi(W^k, \widetilde{W}^k)}{\|Q^{-1}G(W^k - \widetilde{W}^k)\|_Q^2}, \quad (27)$$

with

$$\begin{aligned} \varphi(W^k, \widetilde{W}^k) &= \|W^k - \widetilde{W}^k\|_G^2 \\ &\quad + \frac{1}{\eta} \left\langle \Lambda^k - \widetilde{\Lambda}^k, \sum_{i=1}^3 (X_i - \widetilde{X}_i) \right\rangle \\ &\quad + \frac{1-\eta}{\eta^2\beta} \|\Lambda^k - \widetilde{\Lambda}^k\|^2. \end{aligned} \quad (28)$$

Set  $k := k + 1$  and go to Step 1.

*Remark 5.* In PFPSM, we apply the classical proximal point algorithm (PPA) to regularize its subproblems. Compared to the iteration method in [13], the new iteration method PFPSM extends the scope of  $\eta$  from 1 to the interval  $(\sqrt{3}/2, 2/\sqrt{3})$ , and larger values of  $\eta$  are often more preferred in practice; see Figure 2. Furthermore, compared to the iteration method in [14], the new iteration method PFPSM removes the restriction imposed on the regularized matrices  $R_i$  ( $i = 1, 2, 3$ ); see condition (2.10) in [14].

The following lemma provides a lower bound of the  $\varphi(W^k, \widetilde{W}^k)$ , that is, the numerator of the step size  $\alpha_k$ , and this lower bound plays an important role in the proof of the global convergence of PFPSM.

**Lemma 6.** *Let  $\widetilde{W}^k = (\widetilde{X}_1^k, \widetilde{X}_2^k, \widetilde{X}_3^k, \widetilde{\Lambda}^k)$  be the matrix generated by PFPSM from the given  $W^k$  and  $\varphi(W^k, \widetilde{W}^k)$  be defined by (28). Then we have*

$$\varphi(W^k, \widetilde{W}^k) \geq c_0 \|W^k - \widetilde{W}^k\|_G^2, \quad (29)$$

where  $c_0 = \min\{(2\eta - \sqrt{3})/2\eta, (2 - \sqrt{3}\eta)/2\eta\}$ .

*Proof.* By the matrix form of the inequality

$$ab \geq -\frac{1}{2} \left( va^2 + \frac{1}{v}b^2 \right) \quad (30)$$

with  $a = X_i^k - \widetilde{X}_i^k$ ,  $b = \Lambda^k - \widetilde{\Lambda}^k$  and  $v = \sqrt{3}\beta$ , one has

$$\begin{aligned} &\langle X_i^k - \widetilde{X}_i^k, \Lambda^k - \widetilde{\Lambda}^k \rangle \\ &\geq -\frac{1}{2} \left( \sqrt{3}\beta \|X_i^k - \widetilde{X}_i^k\|^2 + \frac{1}{\sqrt{3}\beta} \|\Lambda^k - \widetilde{\Lambda}^k\|^2 \right). \end{aligned} \quad (31)$$

It follows from (23) and (28) that

$$\begin{aligned} \varphi(W^k, \widetilde{W}^k) &= \sum_{i=1}^3 \|X_i^k - \widetilde{X}_i^k\|_{\beta I_p + R_i}^2 + \frac{1}{\eta^2\beta} \|\Lambda^k - \widetilde{\Lambda}^k\|^2 \\ &\quad + \frac{1}{\eta} \left\langle \Lambda^k - \widetilde{\Lambda}^k, \sum_{i=1}^3 (X_i - \widetilde{X}_i) \right\rangle \geq \sum_{i=1}^3 \|X_i^k - \widetilde{X}_i^k\|_{\beta I_p + R_i}^2 \\ &\quad + \frac{1}{\eta^2\beta} \|\Lambda^k - \widetilde{\Lambda}^k\|^2 \end{aligned}$$

$$\begin{aligned}
 & -\frac{1}{2\eta} \left( \sqrt{3}\beta \sum_{i=1}^3 \|X_i^k - \bar{X}_i^k\|^2 + \frac{3}{\sqrt{3}\beta} \|\Lambda^k - \bar{\Lambda}^k\|^2 \right) \\
 & = \frac{(2\eta - \sqrt{3})\beta}{2\eta} \sum_{i=1}^3 \|X_i^k - \bar{X}_i^k\|^2 + \sum_{i=1}^3 \|X_i^k - \bar{X}_i^k\|_{R_i}^2 \\
 & + \frac{2 - \sqrt{3}\eta}{2\eta^2\beta} \|\Lambda^k - \bar{\Lambda}^k\|^2 \geq \min \left\{ \frac{2\eta - \sqrt{3}}{2\eta}, \frac{2 - \sqrt{3}\eta}{2\eta} \right\} \\
 & \cdot \left( \beta \sum_{i=1}^3 \|X_i^k - \bar{X}_i^k\|^2 + \sum_{i=1}^3 \|X_i^k - \bar{X}_i^k\|_{R_i}^2 + \frac{1}{\eta\beta} \|\Lambda^k - \bar{\Lambda}^k\|^2 \right) \\
 & = c_0 \|W^k - \bar{W}^k\|_G^2, \tag{32}
 \end{aligned}$$

where the first inequality comes from (31) and the second inequality follows from  $\eta \in (\sqrt{3}/2, 2/\sqrt{3})$ . The proof is completed.  $\square$

*Remark 7.* It follows from (27) and (29) that

$$\alpha_k = \frac{\varphi(W^k, \bar{W}^k)}{\|Q^{-1}G(W^k - \bar{W}^k)\|_Q^2} \geq \frac{c_0 \lambda_{\min}(G)}{\|GQ^{-1}G\|}, \quad \forall k \geq 0, \tag{33}$$

where  $\lambda_{\min}(G)$  is the minimum eigenvalue of the positive definite matrix  $G$ . Therefore the sequence  $\{\alpha_k\}$  generated by PFPSM is bounded away from zero. Specially, if we set  $Q = G$ , then

$$\alpha_k \geq \frac{c_0 \|W^k - \bar{W}^k\|_G^2}{\|W^k - \bar{W}^k\|_G^2} = c_0, \quad \forall k \geq 0. \tag{34}$$

Now, we shall show that the stopping criterion (25) is reasonable.

**Lemma 8.** *If  $X_i^k = \bar{X}_i^k$  ( $i = 1, 2, 3$ ),  $\Lambda^k = \bar{\Lambda}^k$ , then the matrix  $(X_1^k, X_2^k, X_3^k, \Lambda^k)$  is a solution of VI( $\mathcal{W}, F, \theta$ ).*

*Proof.* Deriving the first-order optimal condition of  $X_1$  subproblem in (24), we have

$$\begin{aligned}
 & \theta_1(X_1) - \theta_1(\bar{X}_1^k) + \langle X_1 - \bar{X}_1^k, -\Lambda^k \\
 & + \beta(\bar{X}_1^k + X_2^k + X_3^k - D) + R_1(\bar{X}_1^k - X_1^k) \rangle \geq 0, \tag{35} \\
 & \forall X_1 \in \mathcal{R}^{p \times q}.
 \end{aligned}$$

From the updating formula for  $\Lambda$  in (24), one has

$$\Lambda^k = \bar{\Lambda}^k + \eta\beta \left( \sum_{i=1}^3 \bar{X}_i^k - D \right), \tag{36}$$

and, substituting it into the above inequality, we get

$$\begin{aligned}
 & \theta_1(X_1) - \theta_1(\bar{X}_1^k) + \langle X_1 - \bar{X}_1^k, -\bar{\Lambda}^k \\
 & + \beta(1 - \eta)(\bar{X}_1^k + \bar{X}_2^k + \bar{X}_3^k - D) + \beta(X_2^k - \bar{X}_2^k) \\
 & + \beta(X_3^k - \bar{X}_3^k) + R_1(\bar{X}_1^k - X_1^k) \rangle \geq 0, \tag{37} \\
 & \forall X_1 \in \mathcal{R}^{p \times q},
 \end{aligned}$$

which can be reduced to

$$\begin{aligned}
 & \theta_1(X_1) - \theta_1(\bar{X}_1^k) + \left\langle X_1 - \bar{X}_1^k, -\bar{\Lambda}^k \right. \\
 & \left. + \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle \geq \langle X_1 \\
 & - \bar{X}_1^k, (\beta I_p + R_1)(X_1^k - \bar{X}_1^k) \rangle, \quad \forall X_1 \in \mathcal{R}^{p \times q}.
 \end{aligned} \tag{38}$$

Similarly, for the other two variables  $X_2, X_3$ , we have

$$\begin{aligned}
 & \theta_2(X_2) - \theta_2(\bar{X}_2^k) + \left\langle X_2 - \bar{X}_2^k, -\bar{\Lambda}^k \right. \\
 & \left. + \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle \geq \langle X_2 \\
 & - \bar{X}_2^k, (\beta I_p + R_2)(X_2^k - \bar{X}_2^k) \rangle, \quad \forall X_2 \in \mathcal{R}^{p \times q},
 \end{aligned} \tag{39}$$

$$\begin{aligned}
 & \theta_3(X_3) - \theta_3(\bar{X}_3^k) + \left\langle X_3 - \bar{X}_3^k, -\bar{\Lambda}^k \right. \\
 & \left. + \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle \geq \langle X_3 \\
 & - \bar{X}_3^k, (\beta I_p + R_3)(X_3^k - \bar{X}_3^k) \rangle, \quad \forall X_3 \in \mathcal{R}^{p \times q}.
 \end{aligned}$$

Furthermore, the updating formula for  $\Lambda$  in (24) can be rewritten as the following variational inequality problem:

$$\left\langle \Lambda - \bar{\Lambda}^k, \left( \sum_{i=1}^3 \bar{X}_i^k - D \right) + \frac{1}{\eta\beta} (\bar{\Lambda}^k - \Lambda^k) \right\rangle \geq 0, \tag{40}$$

$$\forall \Lambda \in \mathcal{R}^{p \times q}.$$

Then, adding the above four inequalities, and by manipulation, one has

$$\begin{aligned}
 & \theta(X) - \theta(\bar{X}^k) + \left\langle W - \bar{W}^k, F(\bar{W}^k) + (I_p, I_p, I_p)^\top \right. \\
 & \cdot \left[ \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right] \right\rangle \geq \langle W \\
 & - \bar{W}^k, G(W^k - \bar{W}^k) \rangle, \quad \forall W \in \mathcal{W}.
 \end{aligned} \tag{41}$$

Therefore, substituting  $X_i^k = \bar{X}_i^k$  ( $i = 1, 2, 3$ ),  $\Lambda^k = \bar{\Lambda}^k$  into (41), we get

$$\theta(X) - \theta(\bar{X}^k) + \langle W - \bar{W}^k, F(\bar{W}^k) \rangle \geq 0, \tag{42}$$

$$\forall W \in \mathcal{W};$$

that is,

$$\theta(X) - \theta(X^k) + \langle W - W^k, F(W^k) \rangle \geq 0, \tag{43}$$

$$\forall W \in \mathcal{W},$$

which indicates that  $(X_1^k, X_2^k, X_3^k, \Lambda^k)$  is a solution of  $\text{VI}(\mathcal{W}, F, \theta)$ . This completes the proof.  $\square$

If the iteration method PFPSM terminates at Step 2, then the iterate  $W^k$  is an approximate solution of  $\text{VI}(\mathcal{W}, F, \theta)$ . Thus, it is assumed, without loss of generality, that the iteration method PFPSM generates an infinite sequence  $\{W^k\}$ . Now, we intend to prove that  $-Q^{-1}G(W^k - \bar{W}^k)$  is a descent direction of the merit function  $(1/2)\|W - W^*\|_Q^2$  at the point  $W = W^k$ , where  $W^* \in \mathcal{W}^*$ .

**Lemma 9.** *Let  $\{W^k\}$  and  $\{\bar{W}^k\}$  be the two sequences generated by PFPSM. Then, one has*

$$\langle W^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \varphi(W^k, \bar{W}^k), \quad (44)$$

$$\forall W^* \in \mathcal{W}^*.$$

*Proof.* Since  $W^* \in \mathcal{W}^* \subseteq \mathcal{W}$ , it follows from (41) that

$$\begin{aligned} & \theta(X^*) - \theta(\bar{X}^k) + \left\langle W^* - \bar{W}^k, F(\bar{W}^k) \right. \\ & \quad \left. + (I_p, I_p, I_p)^\top \cdot \left[ \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right] \right\rangle \geq \langle W^* \\ & \quad - \bar{W}^k, G(W^k - \bar{W}^k) \rangle. \end{aligned} \quad (45)$$

Thus,

$$\begin{aligned} & \langle \bar{W}^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \theta(\bar{X}^k) - \theta(X^*) \\ & \quad + \left\langle \bar{W}^k - W^*, F(\bar{W}^k) + (I_p, I_p, I_p)^\top \right. \\ & \quad \cdot \left[ \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right] \right\rangle \\ & \geq \left\langle \bar{W}^k - W^*, (I_p, I_p, I_p)^\top \right. \\ & \quad \cdot \left[ \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right] \right\rangle \\ & = \left\langle \sum_{i=1}^3 \bar{X}_i^k - D, \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) \right. \\ & \quad \left. + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle = \frac{1}{\eta\beta} \left\langle \Lambda^k \right. \\ & \quad \left. - \bar{\Lambda}^k, \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle, \end{aligned} \quad (46)$$

where the second inequality follows from  $W^* \in \mathcal{W}^*$ , and the second equality comes from (24). Therefore, we have

$$\begin{aligned} & \langle W^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \|W^k - \bar{W}^k\|_G^2 \\ & \quad + \frac{1}{\eta\beta} \left\langle \Lambda^k - \bar{\Lambda}^k, \frac{1-\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) \right. \\ & \quad \left. + \beta \sum_{i=1}^3 (X_i^k - \bar{X}_i^k) \right\rangle = \varphi(W^k, \bar{W}^k), \end{aligned} \quad (47)$$

which completes the proof of the lemma.  $\square$

From inequalities (29) and (44), one has

$$\begin{aligned} & \langle Q(W^k - W^*), -Q^{-1}G(W^k - \bar{W}^k) \rangle \\ & \leq -c_0 \|W^k - \bar{W}^k\|_G^2, \end{aligned} \quad (48)$$

which indicates that  $-Q^{-1}G(W^k - \bar{W}^k)$  is a descent direction of the merit function  $(1/2)\|W - W^*\|_Q^2$  at  $W = W^k$ . Therefore, it is natural to generate the new iterate  $W^{k+1}$  along this direction as the iterative scheme (26). Now, let us investigate how to choose a suitable step size to achieve progress as much as possible. The new iterate with an undetermined step size  $\alpha$  is denoted by  $W^{k+1}(\alpha)$ , so one has

$$W^{k+1}(\alpha) = W^k - \alpha Q^{-1}G(W^k - \bar{W}^k). \quad (49)$$

Then, let

$$\Theta_k(\alpha) := \|W^k - W^*\|_Q^2 - \|W^{k+1}(\alpha) - W^*\|_Q^2 \quad (50)$$

be the profit function of the  $k$ th iteration. However, we cannot maximize  $\Theta_k(\alpha)$  directly because  $W^* \in \mathcal{W}^*$  is unknown. Motivated by the strategy in [13], we derive a tight lower bound of  $\Theta_k(\alpha)$  in the next lemma, which is independent of the unknown matrix  $W^*$ , to obtain a suboptimal step size.

**Lemma 10.** *For any given matrix  $W^k$ , let  $\bar{W}^k = (\bar{X}_1^k, \bar{X}_2^k, \bar{X}_3^k, \bar{\Lambda}^k)$  be the matrix generated by (24),  $\varphi(W^k, \bar{W}^k)$  be defined by (28), and  $\Theta_k(\alpha)$  be defined by (50). Then, one has*

$$\begin{aligned} & \Theta_k(\alpha) \geq \Phi_k(\alpha) \\ & := 2\alpha\varphi(W^k, \bar{W}^k) - \alpha^2 \|Q^{-1}G(W^k - \bar{W}^k)\|_Q^2. \end{aligned} \quad (51)$$

*Proof.* The proof is similar to that of Lemma 4.6 in [13] and is omitted here.  $\square$

It follows from Lemma 9 that  $\Phi_k(\alpha)$  is a lower bound of  $\Theta_k(\alpha)$  for any  $\alpha > 0$ . Therefore, we can maximize  $\Phi_k(\alpha)$  to accelerate the convergence of PFPSM. Note that  $\Phi_k(\alpha)$  is a quadratic function of  $\alpha$  and it reaches its maximum at  $\alpha_k$  defined by (27). The following theorem shows that the sequence  $\{W^k\}$  generated by PFPSM is Fejér monotone with respect to the solution set  $\mathcal{W}^*$  of  $\text{VI}(\mathcal{W}, F, \theta)$ .

**Theorem 11.** For any  $W^* = (X_1^*, X_2^*, X_3^*, \Lambda^*) \in \mathcal{W}^*$ , there exists a positive constant  $\eta$  such that the sequence  $\{W^k\}$  generated by PFPSM satisfies

$$\|W^{k+1} - W^*\|_Q^2 \leq \|W^k - W^*\|_Q^2 - \tau \|W^k - \widetilde{W}^k\|_G^2, \quad (52)$$

where

$$\tau = \frac{c_0^2 \gamma (2 - \gamma) \lambda_{\min}(G)}{\|G^T Q^{-1} G\|}. \quad (53)$$

*Proof.* It follows from (50) and (51) that

$$\begin{aligned} \|W^{k+1} - W^*\|_G^2 &\leq \|W^k - W^*\|_G^2 - \Phi_k(\gamma \alpha_k) \\ &= \|W^k - W^*\|_G^2 - 2\gamma \alpha_k \varphi(W^k, \widetilde{W}^k) \\ &\quad + (\gamma \alpha_k)^2 \|Q^{-1} G (W^k - \widetilde{W}^k)\|_Q^2 \\ &= \|W^k - W^*\|_G^2 \\ &\quad - \gamma (2 - \gamma) \alpha_k \varphi(W^k, \widetilde{W}^k) \\ &\leq \|W^k - W^*\|_G^2 - \tau \|W^k - \widetilde{W}^k\|_G^2, \end{aligned} \quad (54)$$

where the second equality follows from the definition of  $\alpha_k$  in (27), and the second inequality comes from (29) and (33). Then, assertion (52) is thus proved.  $\square$

*Remark 12.* For some very large-scale (6), computing the step size  $\alpha_k$  defined by (27) is a fussy work needing much calculating. For these instances, set  $Q = G$  in PFPSM, and the step size  $\alpha_k$  can be fixed as the constant  $c_0$  defined in Lemma 6. The new iterate can be generated by

$$W^{k+1} = W^k - \gamma c_0 Q^{-1} G (W^k - \widetilde{W}^k). \quad (55)$$

Then, referring to the proof of the above theorem, we have

$$\begin{aligned} \|W^{k+1} - W^*\|_G^2 &\leq \|W^k - W^*\|_G^2 - \Phi_k(\gamma c_0) \\ &= \|W^k - W^*\|_G^2 - 2\gamma c_0 \varphi(W^k, \widetilde{W}^k) \\ &\quad + (\gamma c_0)^2 \|W^k - \widetilde{W}^k\|_G^2 \\ &= \|W^k - W^*\|_G^2 - 2\gamma c_0^2 \|W^k - \widetilde{W}^k\|_G^2 \\ &\quad + (\gamma c_0)^2 \|W^k - \widetilde{W}^k\|_G^2 \\ &\leq \|W^k - W^*\|_G^2 \\ &\quad - \gamma (2 - \gamma) c_0^2 \|W^k - \widetilde{W}^k\|_G^2. \end{aligned} \quad (56)$$

Based on the Fejér monotonicity of the sequence  $\{W^k\}$  generated by PFPSM, the global convergence of PFPSM can be proved by similar arguments to those in [13], which is omitted.

**Theorem 13.** The sequence  $\{W^k\}$  generated PFPSM converges to some  $W^\infty \in \mathcal{W}^*$ .

## 4. Numerical Results

In this section, we run a series of numerical experiments and apply the proposed PFPSM to solve some concrete examples of SPCP (6). We compare it with the splitting method for separable convex programming (denoted by HTY) in [30], the full Jacobian decomposition of the augmented Lagrangian method (denoted by FJDALM) in [13], and the fully parallel ADMM (denoted by PADMM) in [14]. Through these numerical experiments, we want to illustrate that (1) suitable proximal terms can enhance PFPSM's numerical efficiency in practice, (2) larger values of the Glowinski relaxation factor  $\eta$  can often accelerate PFPSM's convergence speed, and (3) compared with the other three ADMM-type methods the dynamically updated step size  $\alpha_k$  defined in (27) can accelerate the convergence of PFPSM. All codes were written by Matlab R2010a and conducted on a ThinkPad notebook with Pentium(R) Dual-Core CPU T4400@2.2 GHz, 4 GB of memory.

Now, let us first introduce two shrinkage operators, which are used to solve the subproblems of the four tested methods. Let  $c > 0$  be a constant and  $E \in \mathcal{R}^{p \times q}$  be a given matrix, and let  $E = U \Sigma V^T$  be the singular value decomposition (SVD) of  $E$ . Then, the solution of the optimization problem

$$\min_{X \in \mathcal{R}^{p \times q}} \left\{ c \|X\|_* + \frac{1}{2} \|X - E\|^2 \right\} \quad (57)$$

is given by the operator  $\mathcal{D}_c(E) \in \mathcal{R}^{p \times q}$ , which is defined by [31]

$$\mathcal{D}_c(E) := U \text{diag}(\mathcal{S}_c(\Sigma)) V^T. \quad (58)$$

The solution of the optimization problem

$$\min_{S \in \mathcal{R}^{p \times q}} \left\{ c \|S\|_1 + \frac{1}{2} \|S - E\|^2 \right\} \quad (59)$$

is given by the operator  $\mathcal{S}_c(E) \in \mathcal{R}^{p \times q}$ , which is defined componentwisely by [32]

$$\begin{aligned} (\mathcal{S}_c(E))_{ij} &:= \text{sign}(E_{ij}) \cdot \max\{|E_{ij}| - c, 0\}, \\ &1 \leq i \leq p, 1 \leq j \leq q, \end{aligned} \quad (60)$$

where  $\text{sign}(\cdot)$  is the sign function.

In the following, let us discuss how to solve the three subproblems of PFPSM. In fact, these three subproblems are simple enough to have closed-form solutions, which are listed as follows. For simplicity, we set  $R_i = \nu \beta I_p$  ( $i = 1, 2, 3$ ) with  $\nu \geq 0$  in PFPSM. Based on the two shrinkage operators (58) and (60), the closed-form solutions of the three subproblems in the prediction step of PFPSM are summarized as follows:

- (i) The explicit solution  $\widetilde{X}_1^k$  is

$$\begin{aligned}\bar{X}_1^k &= \operatorname{argmin}_{X_1 \in \mathcal{R}^{p \times q}} \left\{ \|X_1\|_* + \frac{\beta}{2} \left\| X_1 + X_2^k + X_3^k - D - \frac{1}{\beta} \Lambda^k \right\|^2 + \frac{\nu\beta}{2} \|X_1 - X_1^k\|^2 \right\} \\ &= \mathcal{D}_{1/\beta(1+\nu)} \left( \frac{1}{1+\nu} \left( D + \frac{1}{\beta} \Lambda^k - X_2^k - X_3^k \right) + \frac{\nu}{1+\nu} X_1^k \right).\end{aligned}\quad (61)$$

(ii) The explicit solution  $\bar{X}_2^k$  is

$$\begin{aligned}\bar{X}_2^k &= \operatorname{argmin}_{X_2 \in \mathcal{R}^{p \times q}} \left\{ \tau \|X_2\|_1 + \frac{\beta}{2} \left\| X_1^k + X_2 + X_3^k - D - \frac{1}{\beta} \Lambda^k \right\|^2 + \frac{\nu\beta}{2} \|X_2 - X_2^k\|^2 \right\} \\ &= \mathcal{S}_{\tau/\beta(1+\nu)} \left( \frac{1}{1+\nu} \left( D + \frac{1}{\beta} \Lambda^k - X_1^k - X_3^k \right) + \frac{\nu}{1+\nu} X_2^k \right).\end{aligned}\quad (62)$$

(iii) The subproblem related to  $\bar{X}_3$  is

$$\begin{aligned}\bar{X}_3^k &= \operatorname{argmin}_{X_3 \in \mathcal{R}^{p \times q}} \left\{ \frac{1}{2\mu} \|P_\Omega(X_3)\|^2 + \frac{\beta}{2} \left\| X_1^k + X_2^k + X_3 - D - \frac{1}{\beta} \Lambda^k \right\|^2 + \frac{\nu\beta}{2} \|X_3 - X_3^k\|^2 \right\} \\ &= \operatorname{argmin}_{X_3 \in \mathcal{R}^{p \times q}} \left\{ \frac{1}{2\mu} \|P_\Omega(X_3)\|^2 + \frac{\beta(1+\nu)}{2} \left\| X_3 - \frac{1}{1+\nu} \left( D + \frac{1}{\beta} \Lambda^k - X_1^k - X_2^k \right) - \frac{\nu}{1+\nu} X_3^k \right\|^2 \right\}.\end{aligned}\quad (63)$$

Then, the explicit solution  $\bar{X}_3^k$  is given as follows:

$$[\bar{X}_3^k]_{ij} = \begin{cases} \frac{\mu\beta(1+\nu)}{1+\mu\beta(1+\nu)} \widehat{Z}_{ij}, & \text{if } (i, j) \in \Omega, \\ \widehat{Z}_{ij}, & \text{otherwise,} \end{cases}\quad (64)$$

$$\text{where } \widehat{Z} = \frac{1}{1+\nu} \left( D + \frac{1}{\beta} \Lambda^k - X_1^k - X_2^k \right) + \frac{\nu}{1+\nu} X_3^k.$$

**4.1. Synthetic Data.** In the experiment, following [5], we used the following synthetic data: the true low-rank matrix  $X_1^* = L^* = V_1 V_2^T$  is generated using normally distributed  $V_1 \in \mathcal{R}^{p \times r}$  and  $V_2 \in \mathcal{R}^{q \times r}$ . Hence, the rank of  $L^*$  is  $r$ , and the low-rank ratio  $\text{rr}$  of  $L^*$  is defined by  $\text{rr} = r/p$ . The nonzero entries of sparse matrix  $X_2^*$  or  $S^*$  are generated uniformly in  $[-500, 500]$ , whose sparsity ratio is denoted by  $\text{spr}$ , defined by  $\|S^*\|_0/pq$ , where  $\|S^*\|_0$  represents the number of nonzero entries of  $S^*$ . Then, we get the matrix  $D = L^* + S^*$ . Here, we consider that only a fraction of entries of  $D$  can be observed, and indices of the observed entries are collected in the index set  $\Omega$ , which is determined randomly with a certain sample ratio (the ratio is denoted by  $\text{sr} = (100|\Omega|/pq)\%$ , where  $|\Omega|$  is the cardinality of  $\Omega$ ). Then, we get the matrix  $P_\Omega(D)$ , which is further corrupted by adding Gaussian noise with standard deviation  $\sigma = 10^{-3}$ . Finally, we get the true observed matrix  $P_\Omega(D) + \sigma \cdot \text{rand}(p, q)$ . Here, we test the following scenarios:

$p = q \in \{100, 200\}$ ,  $\text{sr} \in \{0.9, 0.8, 0.7\}$ ,  $\text{rr} \in \{0.05, 0.1\}$ , and  $\text{spr} \in \{0.05, 0.1\}$ . The parameters  $\tau$  and  $\mu$  in (6) are set as  $\tau = 1/\sqrt{p}$ ,  $\mu = \sqrt{p + \sqrt{8p\sigma}}/10$ , respectively.

Throughout, we use the following stopping criterion:

$$\max \left\{ \frac{\|\bar{X}_1^k - X_1^k\|}{1 + \|X_1^k\|}, \frac{\|\bar{X}_2^k - X_2^k\|}{1 + \|X_2^k\|} \right\} < 10^{-5}, \quad (65)$$

and the initial point is set as  $(X_1^0, X_2^0, X_3^0, \Lambda^0) = (0, 0, 0, 0)$ . For the penalty parameter  $\beta$ , we take  $\beta = 0.06|\Omega|/\|P_\Omega(D)\|_1$  for all tested methods. The parameters in the tested methods are listed as follows:

HTY:  $\gamma = 2.01$ .

FJDALM:  $\gamma = 1.8$ .

PADMM:  $P_i = 2.4\beta I_p$ ,  $\gamma = 1.1$ .

PFPSM:  $Q = G$ ,  $\gamma = 1.5$ .

Let us first test the sensitivity of PFPSM with respect to the proximal parameter  $\nu$  and the Glowinski relaxation factor  $\eta$ , and the numerical advantage of attaching the parameters  $\nu$  and  $\eta$  to the iterative scheme of PFPSM is thus clear. We choose  $\nu = 0:0.1:1$  and fix  $\gamma = 1.5$ ,  $\eta = 1.15$  and  $p = q = 200$ ,  $\text{rr} = \text{spr} = 0.05$ ,  $\text{sr} = 0.8$ . The maximum number of iterations is set to 500. The numerical results

are plotted in Figure 1, in which “RLR” denotes the relative error of the recovered low-rank matrix ( $\|X_1^k - X_1^*\|/\|X_1^*\|$ ) and “RSP” denotes the relative error of the recovered sparse matrix ( $\|X_2^k - X_2^*\|/\|X_2^*\|$ ) when the stopping criterion (65) is satisfied. Figure 1 shows that small values of  $\nu > 0$  can improve the accuracy of the solutions, while large values of  $\nu > 0$  can accelerate the convergence speed of PFPSM. Then taking into consideration the numerical performance of other comparing methods, in the following we set  $\nu = 0.9$ . Now let us make the sensitivity analysis of the parameter  $\eta$ . We choose  $\eta = 0.86:0.02:1.15$  and fix  $\gamma = 1.5$ ,  $\nu = 0.9$  and  $p = q = 100$ ,  $\text{rr} = \text{spr} = 0.05$ ,  $\text{sr} = 0.8$ . The maximum number of iterations is also set to 500. The numerical results are plotted in Figure 2. The two subplots in Figure 1 indicate that the number of iterations is descent and the elapsed CPU time has a descent tendency as the parameter  $\eta$  increases, which verify the numerical advantage of PFPSM with large values of  $\eta$ . Then, based on this observation, we set  $\eta = 1.15$  in the following experiments.

Now, we compare PFPSM with PPSA, FJDALM, and PADMM to examine its numerical advantage from aspects of the number of iterations (denoted by “Iter.”), the elapsed CPU time in seconds (denoted by “Time”), the relative error of the recovered low-rank matrix, and the relative error of the recovered sparse matrix when the stopping criterion (65) is satisfied. Some numerical results are listed in Tables 1 and 2, which are the average results over 20 trials for each setting of parameters.

The numerical results in Tables 1 and 2 indicate the following: (1) From the “Iter.” criterion, the tested methods all successfully recover the given matrix  $D^*$ ; (2) from the “Iter.” and “Time” criteria, our new method PFPSM performs better than the other three tested methods. It is almost the fastest for all the cases, and the reason may be that the suitable choice of the parameter  $\eta$  can accelerate the convergence of PFPSM; (3) compared to HTY, PFPSM is a fully parallel splitting method, which only recorded the maximum time taken by its three subproblems. Thus, we observe in Tables 1 and 2 that PFPSM takes less CPU time than HTY, though it needs to calculate a complicated correction step at each iteration or/and takes more iterations to terminate than HTY. In fact, considering the case  $p = q = 100$ ,  $\text{rr} = \text{spr} = 0.05$ ,  $\text{sr} = 0.9$  as an example, we find that the correction step of PFPSM takes about 7.26% of its whole CPU time, while the computation of the recovered sparse matrix and the simple correction step of HTY takes about 13.10% of its whole CPU time. In conclusion, the proposed PFPSM can be used as a surrogate of HTY, FJDALM, and PADMM.

In Figures 3 and 4, we display the evolution of  $\|X_1^k - X_1^*\|/\|X_1^*\|$  and  $\|X_2^k - X_2^*\|/\|X_2^*\|$  with respect to the iteration counter  $k$ , respectively, from which we can see that PFPSM performs better than FJDALM and PADMM and slightly worse than HTY in terms of the accuracy. The reason may be that PFPSM, FJDALM, and PADMM are fully parallel methods, while HTY is a partially parallel method. Therefore, HTY can fully utilize the latest iteration information to speed up its convergence. The relationship is just like that of the Jacobi iteration method and the Gauss-Seidel iteration

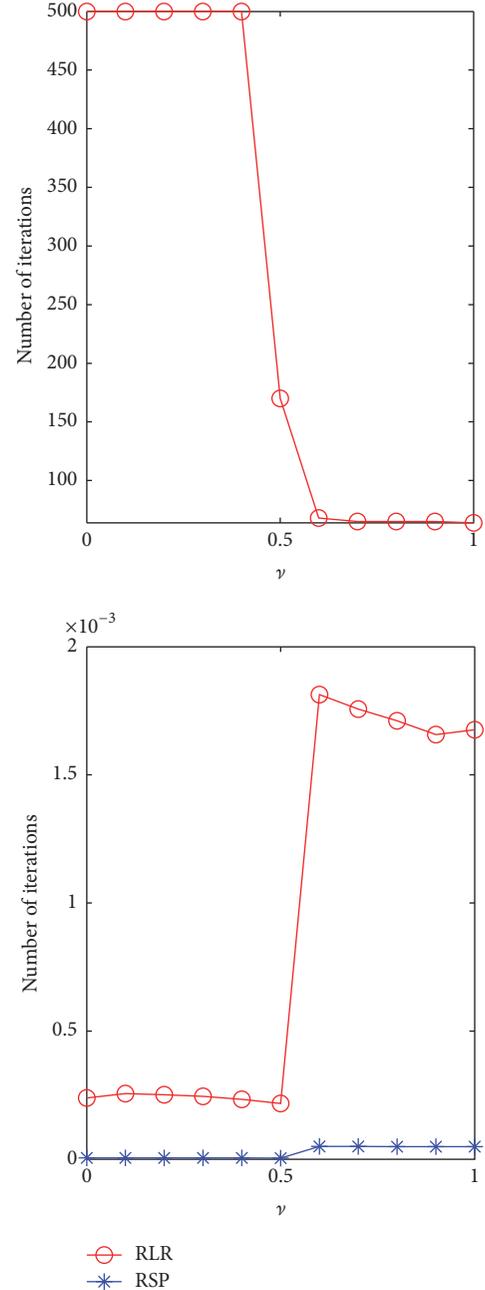


FIGURE 1: Sensitivity test on the parameter  $\nu$  with  $p = q = 100$ ,  $\text{rr} = \text{spr} = 0.05$ ,  $\text{sr} = 0.8$ .

method for linear system of equations. However, as illustrated in Table 1, the CPU time of PFPSM is almost always smaller than that of HTY, which is the main advantage of the parallel methods over the partially parallel or sequential methods.

**4.2. Background Extraction from Surveillance Video with Missing and Noisy Data.** In this subsection, we are going to show the efficiency of the proposed method by computing the real-world problem: background extraction from surveillance video with missing and noisy data. In the experiment, following [6], we focus on a well-tested video taken in an

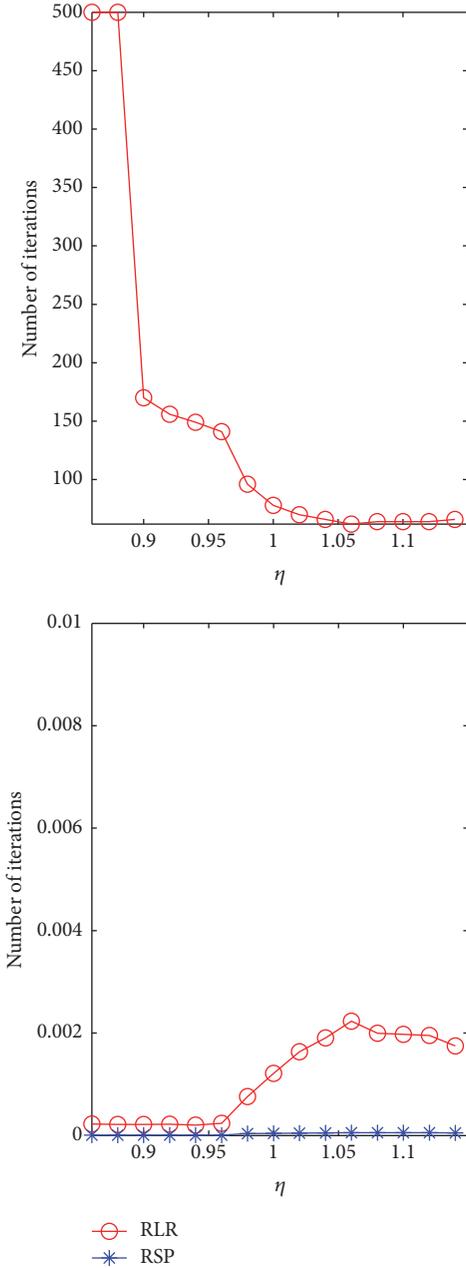


FIGURE 2: Sensitivity test on the parameter  $\eta$  with  $p = q = 100$ ,  $rr = spr = 0.05$ ,  $sr = 0.8$ .

airport, which consists of 50 grayscale frames with each frame having  $144 \times 176$  pixels. We vectorize all frames of the video and get a matrix  $D \in \mathcal{R}^{25344 \times 50}$  where each column represents a frame, which is further contaminated with additive Gaussian noise of mean zero and standard deviation  $\sigma = 10^{-3}$ . Similar to the last experiment, only a fraction of entries of  $D$  can be observed, and indices of the observed entries are also collected in the index set  $\Omega$ . In the experiment, we set  $\tau = 1/\sqrt{p}$ ,  $\delta = 0.01$  and the initial iteration is set to be zero. Other parameters remain the same as those used in the last experiment for all the tested algorithms. The numerical results are given in Table 3, where Obj. represents the value

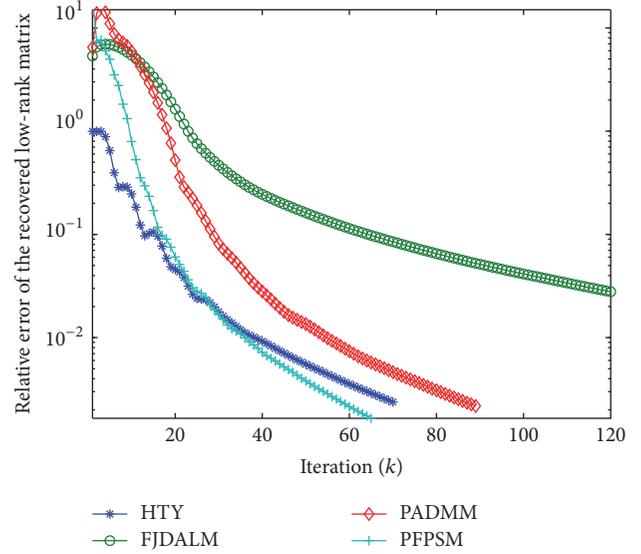


FIGURE 3: The relative error of the recovered low-rank matrix with  $p = q = 100$ ,  $rr = spr = 0.05$ ,  $sr = 0.8$ .

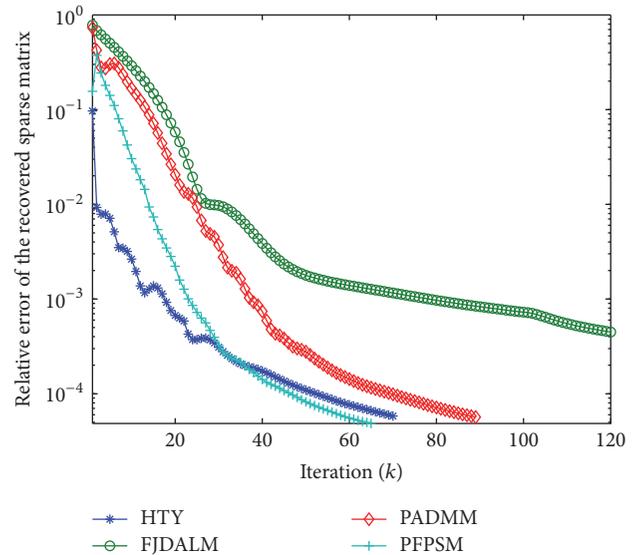


FIGURE 4: The relative error of the recovered sparse matrix with  $p = q = 100$ ,  $rr = spr = 0.05$ ,  $sr = 0.8$ .

of the objective function of (1) when the tested methods terminate.

Table 3 shows that although PFPSM implements more iterations than HTY for all test instances, the performance of PFPSM is much better than the other two methods in recovering low-rank and sparse components. In Figure 5, we exhibit a frame in the tested video and the separation results with 30% random missing data. Clearly, all the tested methods can extract the disturbed image quite well.

### 5. Conclusion

In this paper, we have proposed a new fully parallel splitting method for the stable principal component pursuit problem,

TABLE 1: Numerical results for SPCP with  $p = q = 100$  and different  $rr, spr, sr$ .

$rr$	$spr$	$sr$	Method	Iter.	Time	$\ X_1^k - X_1^*\  / \ X_1^*\ $	$\ X_2^k - X_2^*\  / \ X_2^*\ $
0.05	0.05	0.9	HTY	38	6.51	$3.8250e - 004$	$4.5595e - 005$
			FJDALM	104	15.50	$1.5313e - 003$	$5.5291e - 005$
			PADMM	74	6.46	$4.3589e - 004$	$4.6012e - 005$
			PFPSM	46	4.35	$3.8957e - 004$	$4.5757e - 005$
0.05	0.05	0.8	HTY	45	7.07	$4.7126e - 004$	$4.5788e - 005$
			FJDALM	121	16.41	$2.5283e - 003$	$7.3386e - 005$
			PADMM	80	6.37	$4.6868e - 004$	$4.1213e - 005$
			PFPSM	51	4.42	$4.4273e - 004$	$4.5894e - 005$
0.05	0.05	0.7	HTY	57	9.24	$8.2791e - 004$	$4.9784e - 005$
			FJDALM	160	21.09	$4.9488e - 003$	$1.4563e - 004$
			PADMM	88	7.43	$5.1563e - 004$	$3.2281e - 005$
			PFPSM	58	5.50	$5.2704e - 004$	$3.3050e - 005$
0.05	0.1	0.9	HTY	47	7.43	$1.0740e - 003$	$4.9652e - 005$
			FJDALM	171	25.36	$3.6625e - 003$	$8.5186e - 005$
			PADMM	78	7.85	$5.4060e - 004$	$3.9414e - 005$
			PFPSM	52	5.36	$5.1620e - 004$	$3.9418e - 005$
0.05	0.1	0.8	HTY	66	10.00	$2.0001e - 003$	$6.2475e - 005$
			FJDALM	219	32.29	$6.6939e - 003$	$1.5292e - 004$
			PADMM	91	7.82	$1.5776e - 003$	$5.3509e - 005$
			PFPSM	63	6.93	$1.4964e - 003$	$5.1892e - 005$
0.05	0.1	0.7	HTY	107	16.97	$7.5418e - 003$	$1.8982e - 004$
			FJDALM	299	39.63	$1.5251e - 002$	$3.6022e - 004$
			PADMM	123	10.49	$8.7675e - 003$	$2.2203e - 004$
			PFPSM	94	9.61	$7.7442e - 003$	$1.9931e - 004$
0.1	0.05	0.9	HTY	36	5.21	$3.2385e - 004$	$3.1248e - 005$
			FJDALM	122	17.56	$2.7233e - 003$	$1.2950e - 004$
			PADMM	73	5.91	$3.5371e - 004$	$3.2098e - 005$
			PFPSM	44	3.79	$2.8783e - 004$	$3.1123e - 005$
0.1	0.05	0.8	HTY	53	7.68	$2.4372e - 003$	$1.3577e - 004$
			FJDALM	178	22.78	$7.1372e - 003$	$3.6697e - 004$
			PADMM	78	5.91	$1.1473e - 003$	$6.8425e - 005$
			PFPSM	51	4.28	$1.6505e - 003$	$9.3970e - 005$
0.1	0.05	0.7	HTY	118	16.91	$1.5141e - 002$	$7.8531e - 004$
			FJDALM	292	37.57	$2.5434e - 002$	$1.2862e - 003$
			PADMM	140	11.63	$1.5502e - 002$	$8.0841e - 004$
			PFPSM	116	9.32	$1.4089e - 002$	$7.3816e - 004$

which inherits all merits of ADMM and parallel method. Compared to other similar methods in the literature, the new method attaches a Glowinski relaxation factor  $\eta$  to the updating formula for the Lagrange multiplier, and large values of  $\eta$  are often beneficial to its convergence in general. Preliminary numerical results indicate that the new method is quite efficient.

To ensure the convergence of the proposed iteration method, the scope of the involved parameter  $\eta$  is  $(\sqrt{3}/2, 2/\sqrt{3})$ , which is a proper subset of the interval  $(0, (1 + \sqrt{5})/2)$ , the scope of the Glowinski relaxation factor  $\eta$  of the sequential ADMM-type methods for two-block separable convex

programming [25, 33]. Therefore, it is worth extending the scope of the involved parameter  $\eta$  and studying its optimal choice.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Authors' Contributions

The first author has proved the convergence results; the second author and the third author have accomplished the

TABLE 2: Numerical results for SPCP with  $p = q = 200$  and different  $rr$ ,  $spr$ ,  $sr$ .

$rr$	$spr$	$sr$	Method	Iter.	Time	$\ X_1^k - X_1^*\  / \ X_1^*\ $	$\ X_2^k - X_2^*\  / \ X_2^*\ $
0.05	0.05	0.9	HTY	38	5.94	$3.8876e - 004$	$4.5610e - 005$
			FJDALM	104	18.69	$1.5313e - 003$	$3.6239e - 005$
			PADMM	74	8.42	$4.3589e - 004$	$4.6012e - 005$
			PFPSM	44	4.02	$3.8750e - 004$	$4.5727e - 005$
0.05	0.05	0.8	HTY	48	7.53	$4.4271e - 004$	$4.5768e - 005$
			FJDALM	121	20.45	$2.5283e - 003$	$4.2732e - 005$
			PADMM	80	8.51	$4.6868e - 004$	$4.1213e - 005$
			PFPSM	48	3.98	$4.3835e - 004$	$4.5828e - 005$
0.05	0.05	0.7	HTY	59	9.11	$1.3515e - 003$	$5.9535e - 005$
			FJDALM	160	27.73	$4.9488e - 003$	$5.8452e - 005$
			PADMM	88	10.11	$5.1563e - 004$	$3.2281e - 005$
			PFPSM	56	5.23	$6.7517e - 004$	$3.4954e - 005$
0.05	0.1	0.9	HTY	54	8.24	$1.2604e - 003$	$5.1543e - 005$
			FJDALM	171	29.41	$3.6625e - 003$	$4.3575e - 005$
			PADMM	78	8.94	$5.4060e - 004$	$3.9414e - 005$
			PFPSM	49	4.77	$6.9882e - 004$	$4.6485e - 005$
0.05	0.1	0.8	HTY	76	11.54	$2.3014e - 003$	$6.7335e - 005$
			FJDALM	219	38.30	$6.6939e - 003$	$1.5292e - 004$
			PADMM	91	9.46	$1.5776e - 003$	$5.3509e - 005$
			PFPSM	63	6.23	$1.6119e - 003$	$5.3545e - 005$
0.05	0.1	0.7	HTY	121	18.53	$8.1084e - 003$	$2.0286e - 004$
			FJDALM	299	47.70	$1.5251e - 002$	$4.7190e - 005$
			PADMM	123	13.31	$8.7675e - 003$	$2.2203e - 004$
			PFPSM	99	8.51	$7.3682e - 003$	$1.9037e - 004$
0.1	0.05	0.9	HTY	37	5.58	$4.7725e - 004$	$4.7778e - 005$
			FJDALM	122	39.53	$2.7233e - 003$	$3.2379e - 005$
			PADMM	73	10.73	$3.5371e - 004$	$3.2098e - 005$
			PFPSM	44	4.44	$2.8783e - 004$	$3.1123e - 005$
0.1	0.05	0.8	HTY	61	8.88	$2.9153e - 003$	$1.6130e - 004$
			FJDALM	178	37.66	$7.1372e - 003$	$2.3228e - 005$
			PADMM	78	7.45	$1.1473e - 003$	$6.8425e - 005$
			PFPSM	51	5.54	$1.6505e - 003$	$9.3970e - 005$
0.1	0.05	0.7	HTY	133	18.55	$1.5750e - 002$	$8.1458e - 004$
			FJDALM	292	38.59	$2.5434e - 002$	$1.2862e - 003$
			PADMM	140	11.54	$1.5502e - 002$	$8.0841e - 004$
			PFPSM	116	9.91	$1.4089e - 002$	$7.3816e - 004$

TABLE 3: Numerical results for background extraction problem.

$sr$	Method	Iter.	Time	$\text{rank}(X_1)$	$\ X_2\ _0$	Obj.
0.9	HTY	28	23.0	9	$3.25e + 005$	$2.0481e + 003$
	PADMM	45	27.3	7	$2.80e + 005$	$1.7668e + 003$
	PFPSM	38	25.5	5	$1.22e + 005$	770.4871
0.8	HTY	28	23.1	7	$3.09e + 005$	$1.9492e + 003$
	PADMM	50	28.6	7	$2.83e + 005$	$1.7853e + 003$
	PFPSM	41	28.0	4	$1.14e + 005$	723.1297
0.7	HTY	30	23.4	6	$3.06e + 005$	$1.9291e + 003$
	PADMM	54	34.5	6	$2.81e + 005$	$1.7692e + 003$
	PFPSM	46	34.4	3	$1.09e + 005$	687.3679



FIGURE 5: The 10th frame of the clean video and the corresponding corrupted frame with 30% random missing data (a); the extracted background and foreground by HTY (b), PADMM (c), and PFPSSM (d).

numerical experiment. All authors read and approved the final manuscript.

## Acknowledgments

This work is supported by the National Natural Science Foundation of China (nos. 11671228 and 11601475), 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).

## References

- [1] J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao, "Robust principal component analysis: exact recovery of corrupted low-rank matrices by convex optimization," in *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems (NIPS '09)*, pp. 2080–2088, Vancouver, Canada, December 2009.
- [2] E. J. Candes, X. Li, and Y. J. W. Ma, "Robust principle component analysis?" *Journal of the ACM*, vol. 58, pp. 1–37, 2011.
- [3] V. Chandrasekaran, S. Sanghavi, P. Parrilo, and A. S. Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.
- [4] Z. Zhou, X. Li, J. Wright, E. Candès, and Y. Ma, "Stable principal component pursuit," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '10)*, pp. 1518–1522, Austin, Tex, USA, June 2010.
- [5] M. Tao and X. Yuan, "Recovering low-rank and sparse components of matrices from incomplete and noisy observations," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 57–81, 2011.
- [6] L. Hou, H. He, and J. Yang, "A partially parallel splitting method for multiple-block separable convex programming with applications to robust PCA," *Computational Optimization and Applications*, vol. 63, no. 1, pp. 273–303, 2016.
- [7] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11/12, no. 1-4, pp. 625–653, 1999.
- [8] R. H. Tutuncu, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Math. Program*, vol. 95, pp. 189–217, 2003.
- [9] Y.-H. Xiao and Z.-F. Jin, "An alternating direction method for linear-constrained matrix nuclear norm minimization," *Numerical Linear Algebra with Applications*, vol. 19, no. 3, pp. 541–554, 2012.
- [10] K. Wang, J. Desai, and H. He, "A proximal partially parallel splitting method for separable convex programs," *Optimization Methods & Software*, vol. 32, no. 1, pp. 39–68, 2017.
- [11] D. Han, X. Yuan, W. Zhang, and X. Cai, "An ADM-based splitting method for separable convex programming," *Computational Optimization and Applications*, vol. 54, no. 2, pp. 343–369, 2013.
- [12] M. Sun, Y. Wang, and J. Liu, "Generalized Peaceman-Rachford splitting method for multi-block separable convex programming with applications to robust PCA," *Calcolo*, vol. 54, no. 1, pp. 77–94, 2017.
- [13] B. He, L. Hou, and X. Yuan, "On full Jacobian decomposition of the augmented Lagrangian method for separable convex programming," *SIAM Journal on Optimization*, vol. 25, no. 4, pp. 2274–2312, 2015.
- [14] W. Deng, M.-J. Lai, Z. Peng, and W. Yin, "Parallel multi-block ADMM with  $o(1/k)$  convergence," *Journal of Scientific Computing*, vol. 71, no. 2, pp. 712–736, 2017.
- [15] R. Glowinski and A. Marrocco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non-linéaires," *Revue Française D'automatique, Informatique, Recherche Opérationnelle. Analyse Numérique*, vol. 9, pp. 41–76, 1975.
- [16] D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," *Computers & Mathematics with Applications*, vol. 2, no. 1, pp. 17–40, 1976.
- [17] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, North-Holland, Amsterdam, The Netherlands, 1983.
- [18] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 3, Ser. A, pp. 293–318, 1992.
- [19] K. J. Arrow, L. Hurwicz, and H. Uzawa, *Studies in Linear And Non-Linear Programming*, Stanford University Press, Palo Alto, Calif, USA, 1958.
- [20] C. Chen, B. He, Y. Ye, and X. Yuan, "The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent," *Mathematical Programming*, vol. 155, no. 1-2, pp. 57–79, 2016.
- [21] B. He, M. Tao, and X. Yuan, "Alternating direction method with Gaussian back substitution for separable convex programming," *SIAM Journal on Optimization*, vol. 22, no. 2, pp. 313–340, 2012.
- [22] B. He, H.-K. Xu, and X. Yuan, "On the proximal Jacobian decomposition of ALM for multiple-block separable convex minimization problems and its relationship to ADMM," *Journal of Scientific Computing*, vol. 66, no. 3, pp. 1204–1217, 2016.
- [23] M. Hong and Z.-Q. Luo, "On the linear convergence of the alternating direction method of multipliers," *Mathematical Programming*, vol. 162, no. 1-2, Ser. A, pp. 165–199, 2017.
- [24] B. S. He and F. Ma, "Convergence study on the proximal alternating direction method with larger step size," *Optimization online*, 2017.
- [25] M. Sun and J. Liu, "The convergence rate of the proximal alternating direction method of multipliers with indefinite proximal regularization," *Journal of Inequalities and Applications*, Paper No. 19, 15 pages, 2017.
- [26] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *Journal of Scientific Computing*, vol. 66, no. 3, pp. 889–916, 2016.
- [27] K. Guo, D. Han, D. Z. Wang, and T. Wu, "Convergence of ADMM for multi-block nonconvex separable optimization models," *Frontiers of Mathematics in China*, vol. 12, no. 5, pp. 1139–1162, 2017.
- [28] W. Y. Tian and X. M. Yuan, "Faster alternating direction method of multipliers with  $O(1/n^2)$  convergence rate," *Mathematics of Computation*, 2016.
- [29] M. Fukushima, *Fundamentals of Nonlinear Optimization* (in Japanese), Asakura Shoten, Tokyo, 2001, Chinese edition: Translated by G.H. Lin, published by Science Press, Beijing, 2011.
- [30] B. He, M. Tao, and X. Yuan, "A splitting method for separable convex programming," *IMA Journal of Numerical Analysis (IMAJNA)*, vol. 35, no. 1, pp. 394–426, 2015.

- [31] J.-F. Cai, E. J. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.
- [32] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.
- [33] M. H. Xu, "Proximal alternating directions method for structured variational inequalities," *Journal of Optimization Theory and Applications*, vol. 134, no. 1, pp. 107–117, 2007.



# Hindawi

Submit your manuscripts at  
<https://www.hindawi.com>

