Research Article

A Proximal Fully Parallel Splitting Method for Stable Principal Component Pursuit

Hongchun Sun, Jing Liu, and Min Sun

1School of Mathematics and Statistics, Linyi University, Shandong 276005, China
2School of Data Sciences, Zhejiang University of Finance and Economics, Zhejiang 310018, China
3School of Mathematics and Statistics, Zaozhuang University, Shandong 277160, China
4School 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

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 \( \eta \in (\sqrt{\frac{3}{2}}, \frac{2}{\sqrt{3}}) \) 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:

\[
\min_{L,S} \text{rank}(L) + \tau \, \|S\|_0,
\]

s.t. \( L + S = D \),

where \( D \in \mathbb{R}^{p \times q} \) is a given matrix, \( \text{rank}(L) \) is the rank of the matrix \( L \in \mathbb{R}^{p \times q} \), \( \|S\|_0 \) 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:

\[
\min_{L,S} \|L\|_* + \tau \, \|S\|_1,
\]

s.t. \( L + S = D \),

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:
\[
\min_{L,S} \|L\|_* + \tau \|S\|_1, \\
\text{s.t. } \|P_\Omega (D - L - S)\| \leq \delta,
\]
where \(\delta > 0\) is related to the impulsive or Gaussian noise level, \(\|\cdot\|\) denotes the Frobenius-norm, \(\Omega \subseteq \{1, 2, \ldots, p\} \times \{1, 2, \ldots, q\}\) is the index set of the observable entries of \(D\), and \(P_\Omega : \mathbb{R}^{p\times q} \rightarrow \mathbb{R}^{p\times q}\) is the projection operator defined by

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

\(1 \leq i \leq p, \ 1 \leq j \leq q.\)

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

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

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 \(U \in \mathbb{R}^{p\times q}\), the model (5) can be rewritten in the following equivalent form:

\[
\min_{L,S,U} \left\{ \|L\|_* + \tau \|S\|_1 + \frac{1}{2\mu} \|P_\Omega (U)\|^2 \right\},
\]

\(s.t. \quad L + S + U = D,\)

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 (FFDLM) 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 \(A = (a_{ij}) \in \mathbb{R}^{p \times q}\), let \(\|A\| = \sqrt{\sum_{i,j} a_{ij}^2}\) denote the Frobenius-norm of \(A\). For any two matrices \(A, B \in \mathbb{R}^{p \times q}\), \((A, B) = \text{Trace}(A^T B)\) denotes their Frobenius inner product. Then we have \((A, B) \leq \|A\|\|B\|\). If \(G \in \mathbb{R}^{p \times p}\) is symmetric positive definite matrix, we denote by \(\|A\|_G = \sqrt{(A, GA)} = (\text{Trace}(A^T GA))^{1/2}\) the \(G\)-norm of the matrix \(A\).

The effective domain of a function \(f : X \rightarrow (-\infty, +\infty]\) is defined as \(\text{dom}(f) = \{x \in X \mid f(x) < +\infty\}\). The set of all relative interior points of a given nonempty convex set \(\mathcal{E}\) is denoted by \(\text{ri}(\mathcal{E})\). \(\text{diag}(A_1, A_2, \ldots, A_m)\) has \(A_i\) as its \(i\)th block on the diagonal.

A function \(f : \mathbb{R}^n \rightarrow \mathbb{R}\) is convex if and only if

\[
f(ax + (1 - a)y) \leq af(x) + (1 - a)f(y),
\]

\(\forall x, y \in \mathbb{R}^n, \ a \in [0, 1].\)
Then, for a convex function \( f : \mathbb{R}^n \to \mathbb{R} \), we have the following basic inequality:

\[
 f(x) \geq f(y) + \langle \xi, x - y \rangle,
\]

\[\forall x, y \in \mathbb{R}^n, \xi \in \partial f(y), \quad (8)\]

where \( \partial f(y) = \{ \xi \in \mathbb{R}^n : f(y) + f(\xi) - \langle \xi, y \rangle \} \), for all \( y \in \mathbb{R}^n \) 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 \( \theta_j(X_i) = \| I \|, \theta_2(X_2) = r\|S\|, \theta_3(X_3) = (1/2\mu)\|P_1(U)\|^2 \). Then \( \theta_j(\cdot) (i = 1, 2, 3) \) are all convex functions.

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 \( (X_1^0, X_2^0, X_3^0) \in \mathbb{R}^{p \times q} \times \mathbb{R}^{p \times q} \times \mathbb{R}^{p \times q} \mid X_1 + X_2 + X_3 = D \).

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 \( \Lambda^* \in \mathbb{R}^{p \times q} \) such that \( (X_1^*, X_2^*, X_3^*, \Lambda^*) \) is a solution of the following KKT systems:

\[
 \begin{align*}
 0 & \in \partial \theta_j(X_i^*) - \Lambda^*, & i = 1, 2, 3, \\
 X_1^* + X_2^* + X_3^* = D. & (10)
\end{align*}
\]

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

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

\[
 \theta_j(X_i) - \theta_j(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \\
 \forall X_i \in \mathbb{R}^{p \times q}. & (11)
\]

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

\[
 \Lambda^* \in \partial \theta_j(X_i^*), & (12)
\]

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

\[
 \theta_j(X_i) - \theta_j(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \quad \forall X_i \in \mathbb{R}^{p \times q}. & (13)
\]

that is,

\[
 \theta_j(X_i) - \theta_j(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \quad \forall X_i \in \mathbb{R}^{p \times q}. & (14)
\]

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

\[
 \theta_j(X_i) + \langle X_i, -\Lambda^* \rangle \geq \theta_j(X_i^*) + \langle X_i^*, -\Lambda^* \rangle, \\
 \forall X_i \in \mathbb{R}^{p \times q}, & (15)
\]

which implies

\[
 X_i^* = \arg \min_{X_i \in \mathbb{R}^{p \times q}} \{ \theta_j(X_i) + \langle X_i, -\Lambda^* \rangle \}. & (16)
\]

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

**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 \( \Lambda^* \in \mathbb{R}^{p \times q} \) such that

\[
 \theta_j(X_i) - \theta_j(X_i^*) + \langle X_i - X_i^*, -\Lambda^* \rangle \geq 0, \\
 \forall X_i \in \mathbb{R}^{p \times q}, i = 1, 2, 3; & (17)
\]

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

Furthermore, we set

\[
 X = (X_1, X_2, X_3), \\
 \theta(X) = \theta_1(X_1) + \theta_2(X_2) + \theta_3(X_3), \\
 W = (X, \Lambda). & (18)
\]

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

\[
 \theta(X) - \theta(X^*) + \langle W - W^*, F(X^*) \rangle \geq 0, \\
 \forall W \in \mathcal{W}. & (19)
\]

where \( \mathcal{W} = \mathbb{R}^{p \times q} \times \mathbb{R}^{p \times q} \times \mathbb{R}^{p \times q} \times \mathbb{R}^{p \times q} \), and

\[
 F(W) = \begin{pmatrix}
 -\Lambda \\
 -\Lambda \\
 -\Lambda \\
 X_1 + X_2 + X_3 - D
\end{pmatrix}. & (20)
\]

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:

\[
 \langle W^* - W, F(W^*) - F(W) \rangle = 0, \quad \forall W^*, W \in \mathcal{W}. & (21)
\]

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
\[
\mathcal{L}_\beta(X_1, X_2, X_3, \Lambda) = \sum_{i=1}^{3} \theta_i(X_i) - \langle \Lambda, X_1 + X_2 + X_3 - D \rangle + \frac{\beta}{2} \|X_1 + X_2 + X_3 - D\|^2,
\]
where \(\beta > 0\) is the penalty parameter and \(\Lambda \in \mathbb{R}^{p \times d}\) 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:
\[
G = \text{diag} \left\{ \beta I_p + R_1, \beta I_p + R_2, \beta I_p + R_3, \frac{I_p}{\eta^2} \right\},
\]
where \(R_i \in \mathbb{R}^{p \times p} \ (i = 1, 2, 3)\) 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 \(\epsilon > 0\), 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 \(\bar{W}^k = (\bar{X}_1^k, \bar{X}_2^k, \bar{X}_3^k, \bar{\Lambda}^k)\) by
\[
\bar{X}_1^k = \arg \min \left\{ \mathcal{L}_\beta(X_1, X_2^k, X_3^k, \Lambda^k) + \frac{1}{2} \|X_1 - X_1^k\|^2_{R_1} \right\},
\]
\[
\bar{X}_2^k = \arg \min \left\{ \mathcal{L}_\beta(X_2, X_1^k, X_3^k, \Lambda^k) + \frac{1}{2} \|X_2 - X_2^k\|^2_{R_2} \right\},
\]
\[
\bar{X}_3^k = \arg \min \left\{ \mathcal{L}_\beta(X_3, X_1^k, X_2^k, \Lambda^k) + \frac{1}{2} \|X_3 - X_3^k\|^2_{R_3} \right\},
\]
\[
\bar{\Lambda}^k = \Lambda^k - \eta \beta \left( \sum_{i=1}^{3} \Lambda_i - D \right).
\]

**Step 2** (convergence verification). If the stopping criterion
\[
\max \{ \|X_1^k - \bar{X}_1^k\|, \|X_2^k - \bar{X}_2^k\|, \|X_3^k - \bar{X}_3^k\|, \|\Lambda^k - \bar{\Lambda}^k\| < \epsilon \}
\]

is satisfied, then stop.

**Step 3** (correction step). Generate the new iterate by
\[
W^{k+1} = W^k - \gamma \eta \alpha_k Q^{-1} G \left( W^k - \bar{W}^k \right),
\]
where \(Q\) is any symmetric positive definite matrix and \(\alpha_k\) is the step size defined by
\[
\alpha_k = \frac{\rho(W^k, \bar{W}^k)}{\|Q^{-1} G (W^k - \bar{W}^k)\|_Q},
\]
with
\[
\rho(W^k, \bar{W}^k) = \left\| W^k - \bar{W}^k \right\|_G^2 + \frac{1}{\eta} \left( \Lambda^k - \bar{\Lambda}^k, \sum_{i=1}^{3} (X_i - \bar{X}_i) \right) + \frac{1}{\eta^2} \|\Lambda^k - \bar{\Lambda}^k\|^2.
\]
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 \(\varphi(W^k, \bar{W}^k)\), 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 \(\bar{W}^k = (\bar{X}_1^k, \bar{X}_2^k, \bar{X}_3^k, \bar{\Lambda}^k)\) be the matrix generated by PFPSM from the given \(W^k\) and \(\varphi(W^k, \bar{W}^k)\) be defined by (28). Then we have
\[
\varphi(W^k, \bar{W}^k) \geq c_0 \|W^k - \bar{W}^k\|_G^2,
\]
where \(c_0 = \min\{(2\eta - \sqrt{3})/2\eta, (2 - \sqrt{3}\eta)/2\eta\}\).

**Proof.** By the matrix form of the inequality
\[
ab \geq -\frac{1}{2} \left( \frac{1}{\gamma^2} + \frac{1}{\eta^2} b^2 \right)
\]
with \(a = X_i^k - \bar{X}_i^k\), \(b = \Lambda^k - \bar{\Lambda}^k\) and \(\gamma = \sqrt{3}\beta\), one has
\[
\langle X_i^k - \bar{X}_i^k, \Lambda^k - \bar{\Lambda}^k \rangle \geq -\frac{1}{2} \left( \sqrt{3}\beta \|X_i^k - \bar{X}_i^k\|^2 + \frac{1}{\sqrt{3}\beta} \|\Lambda^k - \bar{\Lambda}^k\|^2 \right).
\]
It follows from (23) and (28) that
\[
\varphi(W^k, \bar{W}^k) = \sum_{i=1}^{3} \|X_i^k - \bar{X}_i^k\|_{R_i}^2 + \frac{1}{\eta^2} \|\Lambda^k - \bar{\Lambda}^k\|^2
\]
\[
+ \frac{1}{\eta} \left( \Lambda^k - \bar{\Lambda}^k, \sum_{i=1}^{3} (X_i - \bar{X}_i) \right) \geq \sum_{i=1}^{3} \|X_i^k - \bar{X}_i^k\|_{R_i}^2 + \frac{1}{\eta^2} \|\Lambda^k - \bar{\Lambda}^k\|^2.
\]
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.

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

$$\alpha_k \equiv \frac{\varphi(W^k, \overline{W}^k)}{\|Q^{-1}G(W^k - \overline{W}^k)\|_G} \geq \frac{\lambda_{\min}(G)}{\|Q^{-1}G\|_G}, \quad \forall k \geq 0,$$

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

$$\alpha_k \geq \frac{\varphi(W^k, \overline{W}^k)}{\|W^k - \overline{W}^k\|^2_G} = \alpha_0, \quad \forall k \geq 0.$$

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

**Lemma 8.** If $X_i^k = \overline{X}_i^k$ $(i = 1, 2, 3)$, $\Lambda^k = \overline{\Lambda}^k$, then the matrix $(\overline{X}_1^k, \overline{X}_2^k, \overline{X}_3^k, \Lambda^k)$ is a solution of VI$(\mathcal{W}, F, \theta)$.

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

$$\theta_1(x_1) - \theta_1(\overline{X}_1^k) + \langle X_1 - \overline{X}_1^k, -\Lambda^k \rangle + \beta(\overline{X}_1^k + X_2^k + X_3^k - D) + R_1(\overline{X}_1^k - X_1^k) \geq 0,$$

$$\forall X_1 \in \mathcal{R}^{PQ^q}.$$ (35)

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

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

and, substituting it into the above inequality, we get

$$\theta_1(x_1) - \theta_1(\overline{X}_1^k) + \langle X_1 - \overline{X}_1^k, -\Lambda^k \rangle + \beta(\overline{X}_1^k + X_2^k + X_3^k - D) + \beta(X_2^k - \overline{X}_2^k)$$

$$+ \beta(X_3^k - \overline{X}_3^k) + R_1(\overline{X}_1^k - X_1^k) \geq 0,$$

$$\forall X_1 \in \mathcal{R}^{PQ^q},$$

which can be reduced to

$$\theta_1(x_1) - \theta_1(\overline{X}_1^k) + \langle X_1 - \overline{X}_1^k, -\Lambda^k \rangle$$

$$+ \frac{1 - \eta}{\eta} \left(\Lambda^k - \overline{\Lambda}^k + \beta \sum_{i=1}^{3} (X_i^k - \overline{X}_i^k)\right) \geq \langle X_1 \rangle$$

$$- X_1^k \left(\beta I_p + R_1\right)(X_i^k - \overline{X}_i^k), \quad \forall X_1 \in \mathcal{R}^{PQ^q}. $$

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

$$\theta_2(X_2) - \theta_2(\overline{X}_2^k) + \langle X_2 - \overline{X}_2^k, -\Lambda^k \rangle$$

$$+ \frac{1 - \eta}{\eta} \left(\Lambda^k - \overline{\Lambda}^k + \beta \sum_{i=1}^{3} (X_i^k - \overline{X}_i^k)\right) \geq \langle X_2 \rangle$$

$$- X_2^k \left(\beta I_p + R_2\right)(X_i^k - \overline{X}_i^k), \quad \forall X_2 \in \mathcal{R}^{PQ^q};$$

$$\theta_3(X_3) - \theta_3(\overline{X}_3^k) + \langle X_3 - \overline{X}_3^k, -\Lambda^k \rangle$$

$$+ \frac{1 - \eta}{\eta} \left(\Lambda^k - \overline{\Lambda}^k + \beta \sum_{i=1}^{3} (X_i^k - \overline{X}_i^k)\right) \geq \langle X_3 \rangle$$

$$- X_3^k \left(\beta I_p + R_3\right)(X_i^k - \overline{X}_i^k), \quad \forall X_3 \in \mathcal{R}^{PQ^q}. $$

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

$$\langle \Lambda - \overline{\Lambda}^k, \left(3 \sum_{i=1}^{3} \overline{X}_i^k - D\right) \rangle + \frac{1}{\eta \beta} \left(\Lambda^k - \overline{\Lambda}^k\right) \geq 0, \quad \forall \Lambda \in \mathcal{R}^{PQ^q}. $$

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

$$\theta(X) - \theta(\overline{X}^k) + \langle W - \overline{W}^k, F(\overline{W}^k) + (I_p, I_p, I_p)^\top \rangle$$

$$\geq \frac{1}{\eta} \left(\Lambda^k - \overline{\Lambda}^k + \beta \sum_{i=1}^{3} (X_i^k - \overline{X}_i^k)\right) \geq \langle W \rangle$$

$$- \overline{W}^k G(W^k - \overline{W}^k), \quad \forall W \in \mathcal{W}. $$

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

$$\theta(X) - \theta(\overline{X}^k) + \langle W - \overline{W}^k, F(\overline{W}^k) \rangle \geq 0, \quad \forall W \in \mathcal{W};$$

that is,

$$\theta(X) - \theta(\overline{X}^k) + \langle W - \overline{W}^k, F(\overline{W}^k) \rangle \geq 0, \quad \forall W \in \mathcal{W}. $$

Finally, it follows from (43) that

$$\theta(X) - \theta(\overline{X}^k) + \langle W - \overline{W}^k, F(\overline{W}^k) \rangle \geq 0, \quad \forall W \in \mathcal{W}. $$
which indicates that \((X^k_1, X^k_2, X^k_3, \Lambda^k)\) is a solution of \(VI(H, F, \theta)\). This completes the proof.

If the iteration method PFPSM terminates at Step 2, then the iterate \(W^k\) is an approximate solution of \(VI(H, F, \theta)\). 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 \((1/2)||W - W^*||_2^2\) at the point \(W = W^k\), where \(W^* \in H^*\).

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

\[
\langle W^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \varphi(W^k, \bar{W}^k),
\]

\[\forall W^* \in H^*.\]

**Proof.** Since \(W^* \in H^* \subseteq H\), it follows from (41) that

\[
\theta(X^*) - \theta(\bar{X}^k) + \left\langle W^* - \bar{W}^k, F(\bar{W}^k) + (I_p, I_p, I_p)^T \right\rangle + \left\langle \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X^k_i - \bar{X}^k_i) \right\rangle \geq \langle W^* - \bar{W}^k, G(W^k - \bar{W}^k) \rangle.
\]

Thus,

\[
\langle \bar{W}^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \theta(\bar{X}^k) - \theta(X^*) + \left\langle \bar{W}^k - W^*, F(\bar{W}^k) + (I_p, I_p, I_p)^T \right\rangle + \left\langle \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X^k_i - \bar{X}^k_i) \right\rangle \geq \langle \bar{W}^k - W^*, (I_p, I_p, I_p)^T \rangle + \left\langle \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X^k_i - \bar{X}^k_i) \right\rangle \geq 2\langle \bar{W}^k - W^*, (I_p, I_p, I_p)^T \rangle + \left\langle \frac{1 - \eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) + \beta \sum_{i=1}^3 (X^k_i - \bar{X}^k_i) \right\rangle \geq 0.
\]

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

\[
\langle W^k - W^*, G(W^k - \bar{W}^k) \rangle \geq \|W^k - \bar{W}^k\|^2_2
\]

\[
+ \frac{1}{\eta} \left\langle \Lambda^k - \bar{\Lambda}^k, 1 - \frac{\eta}{\eta} (\Lambda^k - \bar{\Lambda}^k) \right\rangle
\]

\[
+ \beta \sum_{i=1}^3 (X^k_i - \bar{X}^k_i),
\]

which completes the proof of the lemma.

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

\[
\langle Q(W^k - W^*), -Q^{-1}G(W^k - \bar{W}^k) \rangle \leq -c_0 \|W^k - \bar{W}^k\|^2_2,
\]

which indicates that \(-Q^{-1}G(W^k - \bar{W}^k)\) is a descent direction of the merit function \((1/2)||W - W^*||_2^2\) 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 undeterminate step size \(\alpha\) is denoted by \(W^{k+1}(\alpha)\), so one has

\[
W^{k+1}(\alpha) = W^k - \alpha Q^{-1}G(W^k - \bar{W}^k).
\]

Then, let

\[
\Theta_k(\alpha) = \|W^k - W^*\|^2_2 - \|W^{k+1}(\alpha) - W^*\|^2_2
\]

be the profit function of the \(k\)th iteration. However, we cannot maximize \(\Theta_k(\alpha)\) directly because \(W^* \in H^*\) 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 \(\bar{W}^k = (\bar{X}^k_1, \bar{X}^k_2, \bar{X}^k_3, \bar{\Lambda}^k)\) be the matrix generated by (24), \(q(W^k, \bar{W}^k)\) be defined by (28), and \(\Theta_k(\alpha)\) be defined by (50). Then, one has

\[
\Theta_k(\alpha) \geq \Phi_k(\alpha)
\]

\[
= 2\alpha \varphi(W^k, \bar{W}^k) - \alpha^2 \|Q^{-1}G(W^k - \bar{W}^k)\|^2_2.
\]

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

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 \(H^*\) of \(VI(H, F, \theta)\).
Theorem 11. For any $W^* = (X^*_{11}, X^*_{21}, X^*_{31}, \Lambda^*) \in \mathcal{W}^*$, there exists a positive constant $\eta$ such that the sequence $\{W^k\}$ generated by PFPSM satisfies

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

where

$$\tau = \frac{c_2^2 y (2 - y) \lambda_{\min}(G)}{\|G^T Q^{-1} G\|}.$$  \hspace{1cm} (53)

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

$$\|W^{k+1} - W^*\|_G^2 \leq \|W^k - W^*\|_G^2 - \Phi_k(y \alpha_k)$$

$$= \|W^k - W^*\|_G^2 - 2\gamma \alpha_k \varphi (W^k, \overline{W}^k)$$

$$+ (y \alpha_k)^2 \|Q^{-1} G (W^k - \overline{W}^k)\|_Q^2$$

$$= \|W^k - W^*\|_G^2$$

$$- \gamma (2 - \gamma) \alpha_k \varphi (W^k, \overline{W}^k)$$

$$\leq \|W^k - W^*\|_G^2 - \tau \|W^k - \overline{W}^k\|_G^2,$$  \hspace{1cm} (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. \qed

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

$$W^{k+1} = W^k - y c_0 Q^{-1} G (W^k - \overline{W}^k).$$  \hspace{1cm} (55)

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

$$\|W^{k+1} - W^*\|_G^2 \leq \|W^k - W^*\|_G^2 - \Phi_k(y c_0)$$

$$= \|W^k - W^*\|_G^2 - 2y c_0 \varphi (W^k, \overline{W}^k)$$

$$+ (y c_0)^2 \|W^k - \overline{W}^k\|_G^2$$

$$= \|W^k - W^*\|_G^2 - 2y c_0 \|W^k - \overline{W}^k\|_G^2$$

$$+ (y c_0)^2 \|W^k - \overline{W}^k\|_G^2$$

$$\leq \|W^k - W^*\|_G^2$$

$$- \gamma (2 - \gamma) y c_0 \|W^k - \overline{W}^k\|_G^2.$$  \hspace{1cm} (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@2.2GHz, 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 $E \in \mathbb{R}^{p \times q}$ 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

$$\min_{X \in \mathbb{R}^{p \times q}} \left\{c \|X\|_1 + \frac{1}{2} \|X - E\|_F^2\right\}$$  \hspace{1cm} (57)

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

$$\mathcal{D}_c(E) = U \text{ diag } \{\delta_c(\Sigma)\} V^T.$$  \hspace{1cm} (58)

The solution of the optimization problem

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

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

$$(\mathcal{D}'_c(E))_{ij} = \text{sign}(E_{ij}) \cdot \max\left\{|E_{ij}| - c, 0\right\},$$  \hspace{1cm} (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 $R_i = \nu I_p$ $(i = 1, 2, 3)$ 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 $X^3_{11}$ is
\[
X_k^i = \arg\min_{X_i \in \mathbb{R}^{p \times q}} \left\{ \|X_i\|_2 + \frac{\beta}{2} \|X_i + X_k^i - D - \frac{1}{\beta} \Lambda_k\|_2^2 + \frac{\nu \beta}{2} \|X_i - X_k^i\|_2^2 \right\}
= \mathcal{O}_{1/\beta(1 + \gamma)} \left( \frac{1}{1 + \nu} \left( D + \frac{1}{\beta} \Lambda_k - X_k^i - X_k^i \right) + \frac{\nu}{1 + \nu} X_k^i \right).
\]

(ii) The explicit solution \(X_2^i\) is

\[
X_2^i = \arg\min_{X_2 \in \mathbb{R}^{p \times q}} \left\{ \tau \|X_2\|_2 + \frac{\beta}{2} \|X_2 + X_3 - D - \frac{1}{\beta} \Lambda_3\|_2^2 + \frac{\nu \beta}{2} \|X_2 - X_3\|_2^2 \right\}
= \mathcal{O}_{\tau/\beta(1 + \gamma)} \left( \frac{1}{1 + \nu} \left( D + \frac{1}{\beta} \Lambda_3 - X_3\right) + \frac{\nu}{1 + \nu} X_3 \right).
\]

(iii) The subproblem related to \(X_3\) is

\[
X_3^k = \arg\min_{X_3 \in \mathbb{R}^{p \times q}} \left\{ \frac{\mu}{2} \|P_3 (X_3)\|^2 + \frac{\beta}{2} \|X_3 + X_3 - D - \frac{1}{\beta} \Lambda_3\|_2^2 + \frac{\nu \beta}{2} \|X_3 - X_3\|_2^2 \right\}
= \arg\min_{X_3 \in \mathbb{R}^{p \times q}} \left\{ \frac{\mu}{2} \|P_3 (X_3)\|^2 + \frac{\beta (1 + \gamma)}{2} \|X_3 - \frac{1}{1 + \gamma} \left( D + \frac{1}{\beta} \Lambda_3 - X_3\right) - \frac{\nu}{1 + \nu} X_3 \|_2^2 \right\}.
\]

Then, the explicit solution \(X_3^k\) is given as follows:

\[
[X_3^k]_{ij} = \begin{cases} 
\frac{\mu \beta (1 + \gamma)}{1 + \mu \beta (1 + \gamma)} Z_{ij}, & \text{if } (i, j) \in \Omega, \\
\frac{1}{1 + \nu} \left( D + \frac{1}{\beta} \Lambda_3 - X_3\right) + \frac{\nu}{1 + \nu} X_3, & \text{otherwise},
\end{cases}
\]

where \(Z = \frac{1}{1 + \nu} \left( D + \frac{1}{\beta} \Lambda_3 - X_3\right) + \frac{\nu}{1 + \nu} X_3.\)

4.1. Synthetic Data. In the experiment, following [5], we used the following synthetic data: the true low-rank matrix \(X_1^0 = L_1 = V_1 V_1^*\) is generated using normally distributed \(V_1 \in \mathbb{R}^{p \times r}\) and \(V_2 \in \mathbb{R}^{q \times r}\). Hence, the rank of \(L_1\) is \(r\), and the low-rank ratio \(\rho\) of \(L_1^*\) is defined by \(\rho = r / \rho\). The nonzero entries of sparse matrix \(X_1^0\) or \(S^*\) are generated uniformly in \([-500, 500]\), whose sparsity ratio is denoted by \(s\), defined by \(\|S^*\|_0 / pq\), where \(\|S^*\|_0\) represents the number of nonzero entries of \(S^*\). Then, we get the matrix \(D = L_1^* + S^\). 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 \(\mathcal{O}\), which is determined randomly with a certain sample ratio (the ratio is denoted by \(sr = (100 |\Omega| / pq) \%), where \(|\Omega|\) is the cardinality of \(\Omega\). Then, we get the matrix \(P_3 (D)\), which is further corrupted by adding Gaussian noise with standard deviation \(\sigma = 10^{-3}\). Finally, we get the true observed matrix \(P_3 (D) + \sigma \ast \text{rand}(p, q)\). Here, we test the following scenarios:

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

Throughout, we use the following stopping criterion:

\[
\max \left\{ \frac{\|X_1^k - X_2^k\|_2}{\|X_1^k\|_2}, \frac{\|X_2^k - X_3^k\|_2}{\|X_2^k\|_2} \right\} < 10^{-5},
\]

and the initial point is set as \((X_1^0, X_2^0, X_3^0, S^0) = (0, 0, 0, 0)\). For the penalty parameter \(\beta\), we take \(\beta = 0.06 |\Omega| / \|P_3 (D)\|_1\) for all tested methods. The parameters in the tested methods are listed as follows:

\[
\text{HTY: } \gamma = 2.01.
\text{FJDALM: } \gamma = 1.8.
\text{PADMM: } P_1 = 2.4 \beta l_p, \; \gamma = 1.1.
\text{PFPSM: } G = 0.1, \; \gamma = 1.5.
\]

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 \(\nu = 0.1:1\) and fix \(\gamma = 1.5, \eta = 1.15\) and \(p = q = 200, \; rr = spr = 0.05, \; sr = 0.8\). 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 \(\|X^k_1 - X^*_1\|/\|X^*_1\|\) and “RSP” denotes the relative error of the recovered sparse matrix \(\|X^k_2 - X^*_2\|/\|X^*_2\|\) 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 \(p = q = 100\), \(rr = spr = 0.05\), \(sr = 0.8\). 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 \(p = q = 100\), \(rr = spr = 0.05\), \(sr = 0.9\) 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 \(\|X^k_1 - X^*_1\|/\|X^*_1\|\) and \(\|X^k_2 - X^*_2\|/\|X^*_2\|\) 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 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
airport, which consists of 50 grayscale frames with each frame having 144 × 176 pixels. We vectorize all frames of the video and get a matrix \( D \in \mathbb{R}^{23344 \times 50} \) 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 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,
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 $rr, spr, sr$.

<table>
<thead>
<tr>
<th>$rr$</th>
<th>$spr$</th>
<th>$sr$</th>
<th>Method</th>
<th>Iter.</th>
<th>Time</th>
<th>$|X_1^k - X^<em>|_\infty / |X^</em>|_\infty$</th>
<th>$|X_1^k - X^<em><em>2|</em>\infty / |X^</em><em>2|</em>\infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>0.05</td>
<td>0.9</td>
<td>HTY</td>
<td>38</td>
<td>5.94</td>
<td>3.887e-004</td>
<td>4.561e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>104</td>
<td>18.69</td>
<td>1.531e-003</td>
<td>3.623e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>74</td>
<td>8.42</td>
<td>4.3589e-004</td>
<td>4.6012e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>44</td>
<td>4.02</td>
<td>3.8750e-004</td>
<td>4.5727e-005</td>
</tr>
<tr>
<td>0.05</td>
<td>0.05</td>
<td>0.8</td>
<td>HTY</td>
<td>48</td>
<td>7.53</td>
<td>4.4271e-004</td>
<td>4.576e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>121</td>
<td>20.45</td>
<td>2.5283e-003</td>
<td>4.2732e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>80</td>
<td>8.51</td>
<td>4.6868e-004</td>
<td>4.1213e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>48</td>
<td>3.98</td>
<td>4.3835e-004</td>
<td>4.582e-005</td>
</tr>
<tr>
<td>0.05</td>
<td>0.05</td>
<td>0.7</td>
<td>HTY</td>
<td>59</td>
<td>9.11</td>
<td>1.3515e-003</td>
<td>3.4954e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>160</td>
<td>27.73</td>
<td>4.9488e-003</td>
<td>5.8452e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>88</td>
<td>10.11</td>
<td>5.1563e-004</td>
<td>3.2281e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>56</td>
<td>5.23</td>
<td>6.7517e-004</td>
<td>5.5309e-005</td>
</tr>
<tr>
<td>0.05</td>
<td>0.1</td>
<td>0.9</td>
<td>HTY</td>
<td>54</td>
<td>8.24</td>
<td>1.2604e-003</td>
<td>5.1543e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>171</td>
<td>29.41</td>
<td>3.6625e-003</td>
<td>4.3575e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>78</td>
<td>8.94</td>
<td>4.7777e-004</td>
<td>3.9414e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>49</td>
<td>4.77</td>
<td>6.9882e-004</td>
<td>4.6485e-005</td>
</tr>
<tr>
<td>0.05</td>
<td>0.1</td>
<td>0.8</td>
<td>HTY</td>
<td>76</td>
<td>11.54</td>
<td>2.3014e-003</td>
<td>6.7335e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>219</td>
<td>38.30</td>
<td>6.6939e-003</td>
<td>1.529e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>91</td>
<td>9.46</td>
<td>1.5776e-003</td>
<td>5.3509e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>63</td>
<td>6.23</td>
<td>1.6119e-003</td>
<td>5.3545e-005</td>
</tr>
<tr>
<td>0.05</td>
<td>0.1</td>
<td>0.7</td>
<td>HTY</td>
<td>121</td>
<td>18.53</td>
<td>8.1084e-003</td>
<td>2.0286e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>299</td>
<td>47.70</td>
<td>1.5251e-002</td>
<td>4.7190e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>123</td>
<td>13.31</td>
<td>8.7675e-003</td>
<td>2.2203e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>99</td>
<td>8.51</td>
<td>7.3682e-003</td>
<td>1.9037e-004</td>
</tr>
<tr>
<td>0.1</td>
<td>0.05</td>
<td>0.9</td>
<td>HTY</td>
<td>37</td>
<td>5.58</td>
<td>4.7725e-003</td>
<td>4.7778e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>122</td>
<td>39.53</td>
<td>8.1084e-003</td>
<td>2.0286e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>73</td>
<td>10.73</td>
<td>3.5371e-004</td>
<td>3.209e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>44</td>
<td>4.44</td>
<td>2.8783e-004</td>
<td>3.1132e-005</td>
</tr>
<tr>
<td>0.1</td>
<td>0.05</td>
<td>0.8</td>
<td>HTY</td>
<td>61</td>
<td>8.88</td>
<td>2.9153e-003</td>
<td>1.6130e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>178</td>
<td>37.66</td>
<td>7.1372e-003</td>
<td>2.3228e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>78</td>
<td>7.45</td>
<td>1.1473e-003</td>
<td>6.8425e-005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>51</td>
<td>5.54</td>
<td>1.6505e-003</td>
<td>9.3970e-005</td>
</tr>
<tr>
<td>0.1</td>
<td>0.05</td>
<td>0.7</td>
<td>HTY</td>
<td>133</td>
<td>18.55</td>
<td>1.5750e-002</td>
<td>8.1458e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FJDALM</td>
<td>292</td>
<td>38.59</td>
<td>2.5434e-002</td>
<td>1.2862e-003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PADMM</td>
<td>140</td>
<td>11.54</td>
<td>1.5502e-002</td>
<td>8.0841e-004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PFPSM</td>
<td>116</td>
<td>9.91</td>
<td>1.4089e-002</td>
<td>7.3816e-004</td>
</tr>
</tbody>
</table>

### Table 3: Numerical results for background extraction problem.

<table>
<thead>
<tr>
<th>$sr$</th>
<th>Method</th>
<th>Iter.</th>
<th>Time</th>
<th>rank($X_1$)</th>
<th>$|X_1|_0$</th>
<th>Obj.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9</td>
<td>HTY</td>
<td>28</td>
<td>23.0</td>
<td>9</td>
<td>3.25e + 005</td>
<td>2.0481e + 003</td>
</tr>
<tr>
<td></td>
<td>PADMM</td>
<td>45</td>
<td>27.3</td>
<td>7</td>
<td>2.80e + 005</td>
<td>1.7666e + 003</td>
</tr>
<tr>
<td></td>
<td>PFPSM</td>
<td>38</td>
<td>25.5</td>
<td>5</td>
<td>1.22e + 005</td>
<td>770.4871</td>
</tr>
<tr>
<td>0.8</td>
<td>HTY</td>
<td>28</td>
<td>23.1</td>
<td>7</td>
<td>3.09e + 005</td>
<td>1.9492e + 003</td>
</tr>
<tr>
<td></td>
<td>PADMM</td>
<td>50</td>
<td>28.6</td>
<td>7</td>
<td>2.83e + 005</td>
<td>1.7853e + 003</td>
</tr>
<tr>
<td></td>
<td>PFPSM</td>
<td>41</td>
<td>28.0</td>
<td>4</td>
<td>1.14e + 005</td>
<td>723.1297</td>
</tr>
<tr>
<td>0.7</td>
<td>HTY</td>
<td>30</td>
<td>23.4</td>
<td>6</td>
<td>3.06e + 005</td>
<td>1.9291e + 003</td>
</tr>
<tr>
<td></td>
<td>PADMM</td>
<td>54</td>
<td>34.5</td>
<td>6</td>
<td>2.81e + 005</td>
<td>1.7692e + 003</td>
</tr>
<tr>
<td></td>
<td>PFPSM</td>
<td>46</td>
<td>34.4</td>
<td>3</td>
<td>1.09e + 005</td>
<td>687.3679</td>
</tr>
</tbody>
</table>
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 PFPSM (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


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