Available online at www.tjnsa.com
J. Nonlinear Sci. Appl. 9 (2016), 392–403
Research Article

Journal of Nonlinear Science and Applications

Print: ISSN 2008-1898 Online: ISSN 2008-1901

A proximal splitting method for separable convex programming and its application to compressive sensing

Hongchun Sun[a], Min Sun[b,*], Houchun Zhou[a]

[a]School of Sciences, Linyi University, Shandong, 276005, P. R. China.

[b]School of Mathematics and Statistics, Zaozhuang University, Shandong, 277160, P. R. China.

Communicated by C. Park

Abstract

Recently, by taking full exploitation to the special structure of the separable convex programming, some splitting methods have been developed. However, in some practical applications, these methods need to compute the inverse of a matrix, which maybe slow down their convergence rate, especially when the dimension of the matrix is large. To solve this issue, in this paper we shall study the Peaceman-Rachford splitting method (PRSM) by adding a proximal term to its first subproblem and get a new method named proximal Peaceman-Rachford splitting method (PPRSM). Under mild conditions, the global convergence of the PPRSM is established. Finally, the efficiency of the PPRSM is illustrated by testing some applications arising in compressive sensing.

Keywords: Proximal Peaceman-Rachford splitting method, separable convex programming, compressive sensing.
2010 MSC: 90C25, 90C30.

1. Introduction

In this paper, we are interested in the following convex minimization model with linear constraints and separable objective function:

$$\min\{\theta_1(x_1) + \theta_2(x_2) | A_1x_1 + A_2x_2 = b, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}, \tag{1.1}$$

*Corresponding author
Email addresses: sunhc68@126.com (Hongchun Sun), ziyouxiaodou@163.com (Min Sun), zhouhouchun@163.com (Houchun Zhou)

Received 2015-04-28

where $A_i \in \mathcal{R}^{l\times n_i}(i=1,2)$, $b \in \mathcal{R}^l$ and $\mathcal{X}_i \subset \mathcal{R}^{n_i}(i=1,2)$ are nonempty closed convex sets, $\theta_i : \mathcal{R}^{n_i} \to \mathcal{R}(i=1,2)$ are convex but not necessarily smooth functions, such as in compressive sensing, θ_1 refers to a data-fidelity term and θ_2 denotes a regularization term. Throughout this paper, we assume that the solution set of (1.1) is nonempty and $A_i(i=1,2)$ are full column-rank matrices.

A fundamental method for solving (1.1) is the Peaceman-Rachford splitting method (PRSM)[6, 7], which was presented originally in [1]. The standard PRSM iterative scheme is:

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1\in\mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \beta(A_1x_1^{k+1} + A_2x_2^k - b), \\ x_2^{k+1} = \operatorname{argmin}_{x_2\in\mathcal{X}_2}\{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top(A_1x_1^{k+1} + A_2x_2 - b) + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b). \end{cases} \tag{1.2}$$

The PRSM has been well studied in the literature[3, 4, 5, 8]. The PRSM scheme is always efficient when it is convergent. However, according to [1, 4], the sequence generated by PRSM maybe does not satisfy the strictly contractive property, which results in divergence of the PRSM. To deal with this issue, He et al. [4] developed a strictly contractive Peaceman-Rachford splitting method (SCPRSM), and its iterative scheme is:

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1\in\mathcal{X}_1}\{\theta_1(x_1) - (\lambda^k)^\top(A_1x_1 + A_2x_2^k - b) + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(A_1x_1^{k+1} + A_2x_2^k - b), \\ x_2^{k+1} = \operatorname{argmin}_{x_2\in\mathcal{X}_2}\{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top(A_1x_1^{k+1} + A_2x_2 - b) + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(A_1x_1^{k+1} + A_2x_2^{k+1} - b), \end{cases} \tag{1.3}$$

where the parameter $\alpha \in (0,1)$. Obviously, the iterative scheme (1.3) reduces to the (1.2) if $\alpha = 1$. However, to ensure the global convergence of (1.3), the parameter α must be restricted in the interval $(0,1)$. The most important property of the SCPRSM is that its generated sequence satisfies the strictly contractive property.

However, similar to other splitting methods, the SCPRSM also has to compute the inverse of a matrix in some practical applications, such as the numerical examples in [4]: the statistical learning problems and the image reconstruction models (see iterative schemes (6.5)-(6.11) or (6.20)-(6.21) in [4]). In fact, we need to compute the matrix $(D^\top D + \beta I)^{-1}$, where $D \in \mathcal{R}^{n\times d}$ is the design matrix and I is the identity matrix, which is quite time consuming if the dimension d is large. In order to solve this issue, in this paper, we propose a proximal Peaceman-Rachford splitting method (PPRSM), which regularizes the first subproblem in (1.3) by the proximal regularization $\frac{1}{2}\|A_1(x_1 - x_1^k)\|_R^2$, where $R \in \mathcal{R}^{l\times l}$ is a positive definite matrix. The relationship between SCPRSM and PPRSM can be summarized as follows: in fact, the matrix $(D^\top D+\beta I)^{-1}$ in SCPRSM is resulted from the quadratic term in the objective function. Similar to [2], we can linearize the quadratic term and add a proximal term, and get an implementable iterative scheme, which is just SCPRSM with a special matrix R.

The rest of this paper is organized as follows. In Section 2, we describe the proximal Peaceman-Rachford splitting method and prove its global convergence in detail. The application of the PPRSM to compressive sensing are discussed in Section 3. In Section 4, we compare our algorithm with SCPRSM to illustrate the efficiency by performing numerical experiments. Finally, some conclusions are drawn in Section 5.

To end this section, some notations used in this paper are list. We use $\mathcal{R}_+^n$ to denote the nonnegative quadrant in $\mathcal{R}^n$; the vector x_+ denotes the orthogonal projection of vector $x \in \mathcal{R}^n$ onto $\mathcal{R}_+^n$, that is, $(x_+)_i := \max\{x_i, 0\},\ 1 \le i \le n$; the norm $\|\cdot\|_1$, $\|\cdot\|$ and $\|\cdot\|_M$ denote the Euclidean 1-norm, 2-norm and M-norm, respectively. For $x, y \in R^n$, we use $(x;y)$ to denote the column vector $(x^\top, y^\top)^\top$, and use I_m to denote an identity matrix of order m. The transpose of a matrix M is denoted by $M^\top$.

2. Algorithm and Global Convergence

In this section, we first develop an equivalent reformulation of the problem (1.1) by a mixed variational inequality problem (denoted by $\mathrm{VI}(\mathcal{W}, F, \theta)$). Then we describe a proximal Peaceman-Rachford splitting method (PPRSM) for the $\mathrm{VI}(\mathcal{W}, F, \theta)$, and establish its global convergence.

First, we define some auxiliary variables: $x = (x_1, x_2)$, $w = (x, \lambda)$ and $\theta(x) = \theta_1(x_1) + \theta_2(x_2)$. Then, by invoking the first-order optimality condition for convex programming, we can reformulate problem (1.1) as the following variational inequality problem (denoted by VI$(\mathcal{W}, F, \theta)$): Finding a vector $w^* \in \mathcal{W}$ such that

$$\theta(x) - \theta(x^*) + (w - w^*)^\top F(w^*) \geq 0, \quad \forall w \in \mathcal{W}, \tag{2.1}$$

where $\mathcal{W} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{R}^l$, $w = (x, \lambda) = (x_1, x_2, \lambda)$ and

$$F(w) = \begin{pmatrix} -A_1^\top \lambda \\ -A_2^\top \lambda \\ A_1 x_1 + A_2 x_2 - b \end{pmatrix} = \bar{M} w + \bar{p}, \tag{2.2}$$

where

$$\bar{M} = \begin{pmatrix} 0 & 0 & -A_1^\top \\ 0 & 0 & -A_2^\top \\ A_1 & A_2 & 0 \end{pmatrix}, \bar{p} = \begin{pmatrix} 0 \\ 0 \\ b \end{pmatrix}.$$

We denote the set of (2.1) by $\mathcal{W}^*$. Then, $\mathcal{W}^*$ is nonempty under nonempty assumption onto the solution set of problem (1.1).

In this following, a proximal Peaceman-Rachford splitting method (PPRSM) for solving the VI$(\mathcal{W}, F, \theta)$ is outlined.

Algorithm 2.1. PPRSM

Step 0. Choose the parameters $\alpha \in (0, 1), \beta > 0$, a positive definite matrix $R \in \mathcal{R}^{l \times l}$, the tolerance $\varepsilon > 0$ and the initial iterate $w^0 = (x_1^0, x_2^0, \lambda^0) \in \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{R}^l$. Set $k := 0$.

Step 1. The new iterate $w^{k+1} = (x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ by solving the following problem

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1 \in \mathcal{X}_1} \{\theta_1(x_1) - (\lambda^k)^\top A_1 x_1 + \frac{\beta}{2}\|A_1 x_1 + A_2 x_2^k - b\|^2 + \frac{1}{2}\|A_1(x_1 - x_1^k)\|_R^2\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b), \\ x_2^{k+1} = \operatorname{argmin}_{x_2 \in \mathcal{X}_2} \{\theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top A_2 x_2 + \frac{\beta}{2}\|A_1 x_1^{k+1} + A_2 x_2 - b\|^2\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b). \end{cases} \tag{2.3}$$

Step 2. If

$$\max\{\|A_1(x_1^k - x_1^{k+1})\|, \|A_2(x_2^k - x_2^{k+1})\|, \|\lambda^k - \lambda^{k+1}\|\} < \epsilon, \tag{2.4}$$

stop, then $(x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ is a solution of VI$(\mathcal{W}, F, \theta)$; otherwise, set $k := k + 1$, go to Step 1.

For the ease of description,we denote the sequence $\{\hat{w}^k\}$ as

$$\hat{w}^k = \begin{pmatrix} \hat{x}_1^k \\ \hat{x}_2^k \\ \hat{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^k - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) \end{pmatrix}. \tag{2.5}$$

By the second equality of (2.3), one has

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha[\lambda^k - (\lambda^k - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b))] = \lambda^k - \alpha(\lambda^k - \hat{\lambda}^k). \tag{2.6}$$

Combining this with the fourth equality of (2.3), one has

$$\begin{aligned} \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \alpha\beta[(A_1 x_1^{k+1} + A_2 x_2^k - b) - (A_2 x_2^k - A_2 x_2^{k+1})] \\ &= \lambda^{k+\frac{1}{2}} - \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b) + \alpha\beta A_2(x_2^k - x_2^{k+1}) \\ &= \lambda^{k+\frac{1}{2}} - \alpha[\lambda^k - (\lambda^k - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b))] + \alpha\beta A_2(x_2^k - x_2^{k+1}) \\ &= \lambda^{k+\frac{1}{2}} - \alpha(\lambda^k - \hat{\lambda}^k) + \alpha\beta A_2(x_2^k - x_2^{k+1}) \\ &= \lambda^k - [2\alpha(\lambda^k - \hat{\lambda}^k) - \alpha\beta A_2(x_2^k - \hat{x}_2^k)]. \end{aligned} \tag{2.7}$$

Combining this with (2.5), we obtian

$$
\begin{aligned}
w^{k+1} &= \begin{pmatrix} x_1^k \\ x_2^k \\ \lambda^k \end{pmatrix} - \begin{pmatrix} x_1^k - x_1^{k+1} \\ x_2^k - x_2^{k+1} \\ \lambda^k - \lambda^{k+1} \end{pmatrix} \\
&= \begin{pmatrix} x_1^k \\ x_2^k \\ \lambda^k \end{pmatrix} - \begin{pmatrix} x_1^k - \hat{x}_1^k \\ x_2^k - \hat{x}_2^k \\ -\alpha\beta A_2(x_2^k - \hat{x}_2^k) + 2\alpha(\lambda^k - \hat{\lambda}^k) \end{pmatrix} \\
&= w^k - M(w^k - \hat{w}^k),
\end{aligned}
\tag{2.8}
$$

where

$$
M = \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\alpha\beta A_2 & 2\alpha I_l \end{pmatrix}. \tag{2.9}
$$

Based on above analysis, we would show that the algorithm is globally convergent. To this end, we first give the following needed lemma.

Lemma 2.1. *If $A_i x_i^k = A_i x_i^{k+1} (i = 1, 2)$ and $\lambda^k = \lambda^{k+1}$, then $w^{k+1} = (x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ produced by PPRSM is a solution of VI($\mathcal{W}, F, \theta$).*

Proof. By deriving the first-order optimality condition of x_1-subproblem in (2.3), for any $x_1 \in \mathcal{X}_1$, we have

$$
\theta_1(x_1) - \theta_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top \left\{ A_1^\top[-\lambda^k + \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) + RA_1(x_1^{k+1} - x_1^k)] \right\} \geq 0. \tag{2.10}
$$

From the definition of $\hat{\lambda}^k$ in (2.5), (2.10) can be written as

$$
\theta_1(x_1) - \theta_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top \left\{ -A_1^\top \hat{\lambda}^k + A_1^\top R A_1(x_1^{k+1} - x_1^k) \right\} \geq 0, \forall x_1 \in \mathcal{X}_1. \tag{2.11}
$$

Similarly, from the x_2-subproblem in (2.3), we have

$$
\theta_2(x_2) - \theta_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top \left\{ -A_2^\top \lambda^{k+\frac{1}{2}} + \beta A_2^\top (A_1 x_1^{k+1} + A_2 x_2^{k+1} - b) \right\} \geq 0, \forall x_2 \in \mathcal{X}_2. \tag{2.12}
$$

By the definition of $\lambda^{k+\frac{1}{2}}$ in (2.3) and $\hat{\lambda}^k$ in (2.5), one has

$$
\begin{aligned}
& - A_2^\top \lambda^{k+\frac{1}{2}} + \beta A_2^\top (A_1 x_1^{k+1} + A_2 x_2^{k+1} - b) \\
&\quad = A_2^\top[-\lambda^k + \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b) + \beta(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b)] \\
&\quad = A_2^\top[-\lambda^k + \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) + \alpha\beta(A_1 x_1^{k+1} + A_2 x_2^k - b) - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) \\
&\qquad + \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) + \beta(A_2 x_2^{k+1} - A_2 x_2^k)] \\
&\quad = -A_2^\top \hat{\lambda}^k + \beta A_2^\top A_2(x_2^{k+1} - x_2^k) - \alpha A_2^\top (\hat{\lambda}^k - \lambda^k)].
\end{aligned}
\tag{2.13}
$$

Combining (2.12) with (2.13), one has

$$
\theta_2(x_2) - \theta_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top \left\{ -A_2^\top \hat{\lambda}^k + \beta A_2^\top A_2(x_2^{k+1} - x_2^k) - \alpha A_2^\top (\hat{\lambda}^k - \lambda^k)] \right\} \geq 0, \forall x_2 \in \mathcal{X}_2. \tag{2.14}
$$

In addition, from (2.5) again, we have

$$
(A_1 x_1^{k+1} + A_2 x_2^{k+1} - b) - A_2(x_2^{k+1} - x_2^k) + \frac{1}{\beta}(\hat{\lambda}^k - \lambda^k) = 0. \tag{2.15}
$$

Using (2.11), (2.14), (2.15) and $x_i^{k+1} = \hat{x}_i^k (i = 1, 2)$, for any $w = (x_1, x_2, \lambda) \in \mathcal{W}$, it holds that

$$
\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top \left\{ \begin{pmatrix} -A_1^\top \hat{\lambda}^k \\ -A_2^\top \hat{\lambda}^k \\ A_1 \hat{x}_1^k + A_2 \hat{x}_2^k - b \end{pmatrix} + \begin{pmatrix} A_1^\top R A_1(\hat{x}_1^k - x_1^k) \\ \beta A_2^\top A_2(\hat{x}_2^k - x_2^k) - \alpha A_2^\top (\hat{\lambda}^k - \lambda^k) \\ -A_2(\hat{x}_2^k - x_2^k) + (\hat{\lambda}^k - \lambda^k)/\beta \end{pmatrix} \right\} \geq 0.
$$

For any $w \in \mathcal{W}$, the above inequality can be written as

$$\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k) \geq (w - \hat{w}^k)^\top Q(w^k - \hat{w}^k), \tag{2.16}$$

where

$$Q = \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & \beta A_2^\top A_2 & -\alpha A_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_l \end{pmatrix}. \tag{2.17}$$

If $A_i x_i^k = A_i x_i^{k+1} (i = 1, 2)$ and $\lambda^k = \lambda^{k+1}$, combining (2.5) with (2.7), we have $A_i x_i^k = A_i \hat{x}_i^k (i = 1, 2)$ and $\lambda^k = \hat{\lambda}^k$. Thus,

$$Q(w^k - \hat{w}^k) = 0. \tag{2.18}$$

Combining this with (2.16), one has

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

which implies that $\hat{w}^k = (\hat{x}_1^k, \hat{x}_2^k, \hat{\lambda}^k)$ is a solution of $\mathrm{VI}(\mathcal{W}, F, \theta)$. Since $\hat{w}^k = w^{k+1}$, so w^{k+1} is a solution of $\mathrm{VI}(\mathcal{W}, F, \theta)$. □

Lemma 2.2. *The matrices M, Q defined in* (2.9) *and* (2.17), *respectively. Then, we have*

$$HM = Q, \tag{2.19}$$

and the matrix H is positive definite, where

$$H = \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & \frac{2-\alpha}{2} \beta A_2^\top A_2 & -\frac{1}{2} A_2^\top \\ 0 & -\frac{1}{2} A_2 & \frac{1}{2\alpha\beta} I_l \end{pmatrix}. \tag{2.20}$$

Proof. Using (2.9) and (2.17), one has

$$\begin{aligned} HM &= \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & \frac{2-\alpha}{2} \beta A_2^\top A_2 & -\frac{1}{2} A_2^\top \\ 0 & -\frac{1}{2} A_2 & \frac{1}{2\alpha\beta} I_l \end{pmatrix} \begin{pmatrix} I_{n_2} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\alpha\beta A_2 & 2\alpha I_l \end{pmatrix} \\ &= \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & \beta A_2^\top A_2 & -\alpha A_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_l \end{pmatrix} = Q. \end{aligned}$$

Then the first assertion is proved.

Since R is a positive definite matrix, there exists positive definite matrix $R_1 \in \mathcal{R}^{l \times l}$, such that $R = R_1^\top R_1$. By a simple manipulation, we obtain

$$H = \begin{pmatrix} A_1^\top R_1^\top & 0 & 0 \\ 0 & \sqrt{\beta} A_2^\top & 0 \\ 0 & 0 & \frac{1}{\sqrt{\beta}} I_l \end{pmatrix} \begin{pmatrix} I_l & 0 & 0 \\ 0 & \frac{2-\alpha}{2} I_l & \frac{1}{2} I_l \\ 0 & \frac{1}{2} I_l & \frac{1}{2\alpha} I_l \end{pmatrix} \begin{pmatrix} R_1 A_1 & 0 & 0 \\ 0 & \sqrt{\beta} A_2 & 0 \\ 0 & 0 & \frac{1}{\sqrt{\beta}} I_l \end{pmatrix}.$$

Since the matrix

$$\begin{pmatrix} I_l & 0 & 0 \\ 0 & \frac{2-\alpha}{2} I_l & \frac{1}{2} I_l \\ 0 & \frac{1}{2} I_l & \frac{1}{2\alpha} I_l \end{pmatrix}$$

is positive definite if $\alpha \in (0, 1)$, and $A_i (i = 1, 2)$ are full column-rank matrices. Thus, H is positive definite if $\alpha \in (0, 1)$. □

Lemma 2.3. *Let the sequence $\{w^k\}$ be generated by PPRSM. Then, for any $w \in \mathcal{W}$, one has*

$$(w-\hat{w}^k)^\top Q(w^k-\hat{w}^k) \geq \frac{1}{2}(\|w-w^{k+1}\|_H^2 - \|w-w^k\|_H^2) + \frac{1}{2}\|w^k-\hat{w}^k\|_N^2, \tag{2.21}$$

where the matrices Q, H defined in (2.17) *and* (2.20), *respectively, and*

$$N = \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & \frac{\beta(1-\alpha)}{4}A_2^\top A_2 & 0 \\ 0 & 0 & \frac{2(1-\alpha)}{3\beta}I_l \end{pmatrix}.$$

Proof. From the fact that

$$(a-b)^\top H(c-d) = \frac{1}{2}(\|a-d\|_H^2 - \|a-c\|_H^2) + \frac{1}{2}(\|c-b\|_H^2 - \|d-b\|_H^2), \forall a,b,c,d \in R^n.$$

Setting $a=w, b=\hat{w}^k, c=w^k, d=w^{k+1}$, one has

$$(w-\hat{w}^k)^\top H(w^k-w^{k+1}) = \frac{1}{2}(\|w-w^{k+1}\|_H^2 - \|w-w^k\|_H^2) + \frac{1}{2}(\|w^k-\hat{w}^k\|_H^2 - \|w^{k+1}-\hat{w}^k\|_H^2).$$

Using the above equality, combining (2.19) with (2.8), we obtain

$$\begin{aligned}(w-\hat{w}^k)^\top Q(w^k-\hat{w}^k) &= (w-\hat{w}^k)^\top HM(w^k-\hat{w}^k) \\ &= (w-\hat{w}^k)^\top H(w^k-w^{k+1}) \\ &= \frac{1}{2}(\|w-w^{k+1}\|_H^2 - \|w-w^k\|_H^2) + \frac{1}{2}(\|w^k-\hat{w}^k\|_H^2 - \|w^{k+1}-\hat{w}^k\|_H^2).\end{aligned} \tag{2.22}$$

For the last term of (2.22), using (2.8), one has

$$\begin{aligned}&\|w^k-\hat{w}^k\|_H^2 - \|w^{k+1}-\hat{w}^k\|_H^2 \\ &= \|w^k-\hat{w}^k\|_H^2 - \|(w^k-\hat{w}^k)-(w^k-w^{k+1})\|_H^2 \\ &= \|w^k-\hat{w}^k\|_H^2 - \|(w^k-\hat{w}^k)-M(w^k-\hat{w}^k)\|_H^2 \\ &= 2(w^k-\hat{w}^k)^\top HM(w^k-\hat{w}^k) - (w^k-\hat{w}^k)^\top M^\top HM(w^k-\hat{w}^k) \\ &= (w^k-\hat{w}^k)^\top (Q^\top + Q - M^\top HM)(w^k-\hat{w}^k).\end{aligned} \tag{2.23}$$

Using (2.9),(2.17) and (2.19), a direct computation yields that

$$Q^\top + Q - M^\top HM = Q^\top + Q - M^\top Q = \begin{pmatrix} A_1^\top R A_1 & 0 & 0 \\ 0 & (1-\alpha)\beta A_2^\top A_2 & -(1-\alpha)A_2^\top \\ 0 & -(1-\alpha)A_2 & \frac{2(1-\alpha)}{\beta}I_l \end{pmatrix}.$$

Combining this with (2.23), using the Cauchy-Schwartz Inequality, one has

$$\begin{aligned}&(w^k-\hat{w}^k)^\top (Q^\top + Q - M^\top HM)(w^k-\hat{w}^k) \\ = \ & \|A_1(x_1^k-\hat{x}_1^k)\|_R^2 + (1-\alpha)\{\beta\|A_2(x_2^k-\hat{x}_2^k)\|^2 - 2[A_2(x_2^k-\hat{x}_2^k)]^\top(\lambda^k-\hat{\lambda}^k) \\ &+\frac{2}{\beta}\|\lambda^k-\hat{\lambda}^k\|^2\} \\ = \ & \|A_1(x_1^k-\hat{x}_1^k)\|_R^2 + (1-\alpha)\{\frac{\beta}{4}\|A_2(x_2^k-\hat{x}_2^k)\|^2 + \frac{2}{3\beta}\|\lambda^k-\hat{\lambda}^k\|^2 \\ &+\frac{3\beta}{4}\|A_2(x_2^k-\hat{x}_2^k)\|^2 - 2[A_2(x_2^k-\hat{x}_2^k)]^\top(\lambda^k-\hat{\lambda}^k) + \frac{4}{3\beta}\|\lambda^k-\hat{\lambda}^k\|^2\} \\ \geq \ & \|A_1(x_1^k-\hat{x}_1^k)\|_R^2 + (1-\alpha)\{\frac{\beta}{4}\|A_2(x_2^k-\hat{x}_2^k)\|^2 + \frac{2}{3\beta}\|\lambda^k-\hat{\lambda}^k\|^2\} \\ = \ & \|w^k-\hat{w}^k\|_N^2.\end{aligned}$$

Combining this with (2.23), one has

$$\|w^k - \hat{w}^k\|_H^2 - \|w^{k+1} - \hat{w}^k\|_H^2 \geq \|w^k - \hat{w}^k\|_N^2.$$

Combining this with (2.22), we have that (2.21) holds. □

Theorem 2.4. *Let $\{w^k\}$ be the sequence generated by PPRSM. Then, for any $w \in \mathcal{W}$, we have*

$$\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(w) \geq \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2) + \frac{1}{2}\|w^k - \hat{w}^k\|_N^2. \tag{2.24}$$

Proof. From (2.2), a direct computation yields that

$$(w - \hat{w}^k)^\top (F(w) - F(\hat{w}^k)) = 0.$$

Thus, one has $(w - \hat{w}^k)^\top F(w) = (w - \hat{w}^k)^\top F(\hat{w}^k)$. Combining this, using (2.16) and (2.21), we obtain that (2.24) holds. □

Using the above theorem, we first prove that the stopping criterion (2.4) is reasonable.

Theorem 2.5. *Let $\{w^k\}$ be the sequence generated by PPRSM. Then, we have*

$$\lim_{k\to\infty} \|A_1(x_1^k - \hat{x}_1^k)\| = 0, \lim_{k\to\infty} \|A_2(x_2^k - \hat{x}_2^k)\| = 0, \quad \text{and} \quad \lim_{k\to\infty} \|\lambda^k - \hat{\lambda}^k\| = 0. \tag{2.25}$$

Proof. Setting $w = w^* \in \mathcal{W}^*$ in (2.24), we obtain

$$\begin{aligned} \|w^k - w^*\|_H^2 - \|w^k - \hat{w}^k\|_N^2 &\geq 2\{\theta(\hat{x}^k) - \theta(x^*) + (\hat{w}^k - w^*)^\top F(w^*)\} + \|w^{k+1} - w^*\|_H^2 \\ &\geq \|w^{k+1} - w^*\|_H^2, \end{aligned}$$

where the second inequality follows from $w^* \in \mathcal{W}^*$. Thus, one has

$$\|w^{k+1} - w^*\|_H^2 \leq \|w^k - w^*\|_H^2 - \|w^k - \hat{w}^k\|_N^2. \tag{2.26}$$

Using (2.26), a direct computation yields that

$$\sum_{k=0}^{\infty} \|w^k - \hat{w}^k\|_N^2 \leq \|w^0 - w^*\|_H^2,$$

which implies that

$$\lim_{k\to\infty} \|w^k - \hat{w}^k\|_N = 0. \tag{2.27}$$

Combining (2.27) with the matrix N defined in Lemma 2.3, we have that (2.25) holds. □

Now, we are ready to establish the global convergence of PPRSM for solving VI$(\mathcal{W}, F, \theta)$.

Theorem 2.6. *Let $\{w^k\}$ be the sequence generated by PPRSM. Then, the sequence $\{w^k\}$ converges to some $\bar{w} \in \mathcal{W}^*$.*

Proof. From (2.26), we have

$$\|w^{k+1} - w^*\|_H^2 \leq \|w^k - w^*\|_H^2. \tag{2.28}$$

Using (2.28), a direct computation yields that

$$\|w^{k+1} - w^*\|_H^2 \leq \|w^0 - w^*\|_H^2,$$

which indicates that the sequence $\{w^k\}$ is bounded. Combining this with (2.27), the sequence $\{\hat{w}^k\}$ is also bounded. Therefore, it has at least one cluster point. Let $\bar{w}$ be a cluster point of $\{\hat{w}^k\}$ and the subsequence $\{\hat{w}^{k_j}\}$ converges to $\bar{w}$.

On the other hand, combining (2.25) with the definition of Q in (2.17), we have

$$\lim_{k\to\infty} Q(w^k - \hat{w}^k) = 0.$$

Combining this with (2.16), we get

$$\lim_{k\to\infty} \{\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k)\} \geq 0, \quad \forall w \in \mathcal{W}. \tag{2.29}$$

Substituting $\hat{x}^k$ in (2.29) with $\hat{w}^{k_j}$, and letting $k_j \to \infty$, we have

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

which implies that $\bar{w} \in \mathcal{W}^*$.

From $\lim_{k\to\infty} \|w^k - \hat{w}^k\|_N = 0$, we can deduce $\lim_{k\to\infty} \|w^k - \hat{w}^k\|_H = 0$, combining this with $\{\hat{w}^{k_j}\} \to \bar{w}$, for any given $\epsilon > 0$, there exists an integer l such that

$$\|w^{k_l} - \hat{w}^{k_l}\|_H < \frac{\epsilon}{2}, \quad \text{and} \quad \|\hat{w}^{k_l} - \bar{w}\|_H < \frac{\epsilon}{2}.$$

Thus, for any $k \geq k_l$, using the above two equalities and (2.28) that

$$\|w^k - w^\infty\|_H \leq \|w^{k_l} - \bar{w}\|_H \leq \|w^{k_l} - \hat{w}^{k_l}\|_H + \|\hat{w}^{k_l} - \bar{w}\|_H < \epsilon.$$

Thus, the sequence $\{w^k\}$ converges to $\bar{w} \in \mathcal{W}^*$. □

3. Application to Compressive Sensing

Compressive sensing (CS) is to recover a sparse signal $\bar{x} \in \mathcal{R}^n$ from an undetermined linear system $y = A\bar{x}$, where $A \in \mathcal{R}^{m\times n}$ ($m \ll n$) is the sensing matrix, and a fundamental decoding model in CS is the so-called unconstrained basis pursuit denoising (QP_ρ) problem, which can be depicted as

$$\min_{x\in\mathcal{R}^n} f(x) = \frac{1}{2}\|Ax - y\|_2^2 + \rho\|x\|_1, \tag{3.1}$$

where $\rho > 0$ is the regularization parameter and $\|x\|_1$ is the l_1-norm of the vector x defined as $\|x\|_1 = \sum_{i=1}^n |x_i|$.

Now, by introducing auxiliary variables μ_i and ν_i, $i = 1, 2, \cdots, n$, and letting

$$\mu_i + \nu_i = |x_i|, \mu_i - \nu_i = x_i, i = 1, 2, \cdots, n.$$

Thus, the problem QP_ρ is written as

$$\begin{aligned} \min_{(\mu;\nu)\in\mathcal{R}^{2n}} \quad & \frac{1}{2}\|(A, -A)(\mu;\nu) - y\|_2^2 + \rho(e^\top, e^\top)(\mu;\nu), \\ \text{s.t.} \quad & (\mu;\nu) \geq 0, \end{aligned} \tag{3.2}$$

where $e \in R^n$ denote the vector composed by elements 1.

Now, we can transform the problem (3.2) into the following problem. Letting $x_1 = (\mu;\nu), x_2 = (\mu;\nu)$, one has

$$\begin{aligned} \min \quad & \frac{1}{2}\|(A, -A)x_1 - y\|_2^2 + \rho(e^\top, e^\top)x_2, \\ \text{s.t.} \quad & x_1 - x_2 = 0, x_1 \geq 0, x_2 \geq 0, \end{aligned} \tag{3.3}$$

which is a special case of (1.1) with

$$\theta_1(x_1) = \frac{1}{2}\|(A,-A)x_1 - y\|_2^2, \theta_2(x_2) = \rho(e^\top, e^\top)x_2, A_1 = I_n, A_2 = -I_n, b = 0, \mathcal{X}_1 = \mathcal{X}_2 = \mathcal{R}_+^{2n}.$$

Combining PPRSM with (3.3), we first consider the following problem

$$\begin{aligned} \min\ & \frac{1}{2}\|(A,-A)x_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|_2^2, \\ \text{s.t.}\ & x_1 \in \mathcal{R}_+^{2n}. \end{aligned} \tag{3.4}$$

By a direct computation, we can establish the following equivalent formulation of (3.4)

$$\begin{aligned} \min\ & \frac{1}{2}\|x_1 - ([(A,-A)^\top(A,-A) + \beta I_{2n}]^{-1}[(A,-A)^\top y + \lambda^k + \beta x_2^k])\|_{\hat{M}}^2, \\ \text{s.t.}\ & x_1 \in \mathcal{R}_+^{2n}, \end{aligned} \tag{3.5}$$

where $\hat{M} = (A,-A)^\top(A,-A) + \beta I_{2n}$. Thus, the solution of (3.5) is given by

$$x_1^{k+1} = ([(A,-A)^\top(A,-A) + \beta I_{2n}]^{-1}[(A,-A)^\top y + \lambda^k + \beta x_2^k])_+.$$

However, the computation of $[(A,-A)^\top(A,-A) + \beta I_{2n}]^{-1}$ is very time consuming if n is large. Then, we linearize $\frac{1}{2}\|(A,-A)x_1 - y\|_2^2$ at the current point x_1^k and add a proximal term, i.e.,

$$\frac{1}{2}\|(A,-A)x_1 - y\|_2^2 \approx \frac{1}{2}\|(A,-A)x_1^k - y\|_2^2 + (g^k)^\top(x_1 - x_1^k) + \frac{1}{2\tau}\|x_1 - x_1^k\|^2,$$

where $g^k = (A,-A)^\top((A,-A)x_1^k - y)$ denotes the gradient at x_1^k, and $\tau > 0$ is a parameter. Thus, (3.4) is approximated by the following problem

$$\begin{aligned} \min\ & \frac{1}{2}\|(A,-A)x_1^k - y\|_2^2 + (g^k)^\top(x_1 - x_1^k) + \frac{1}{2\tau}\|x_1 - x_1^k\|^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 \\ \text{s.t.}\ & x_1 \in \mathcal{R}_+^{2n}, \end{aligned}$$

which can be written as

$$\begin{aligned} \min\ & (g^k)^\top x_1 + \frac{1}{2\tau}\|x_1 - x_1^k\|^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 \\ \text{s.t.}\ & x_1 \in \mathcal{R}_+^{2n}. \end{aligned} \tag{3.6}$$

Obviously, the above problem has the solution

$$x_1^{k+1} = \frac{\tau}{1+\beta\tau}(\lambda^k + \frac{1}{\tau}x_1^k + \beta x_2^k - g^k)_+.$$

In the following, we show that (3.6) is the x_1-subproblem of (2.3) with $R = \frac{1}{\tau}I_{2n} - (A,-A)^\top(A,-A)$. In fact, setting $R = \frac{1}{\tau}I_{2n} - (A,-A)^\top(A,-A)$ in (2.3), we have

$$\begin{aligned} & \operatorname{argmin}_{x_1 \in \mathcal{R}_+^{2n}}\{\frac{1}{2}\|(A,-A)x_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 + \frac{1}{2}\|x_1 - x_1^k\|_R^2\} \\ = \ & \operatorname{argmin}_{x_1 \in \mathcal{R}_+^{2n}}\left\{ \begin{array}{l} \frac{1}{2}\|(A,-A)x_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 \\ +\frac{1}{2}(x_1 - x_1^k)^\top(\frac{1}{\tau}I_n - (A,-A)^\top(A,-A))(x_1 - x_1^k) \end{array} \right\} \\ = \ & \operatorname{argmin}_{x_1 \in \mathcal{R}_+^{2n}}\left\{ \begin{array}{l} \frac{1}{2}\|(A,-A)x_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 \\ +\frac{1}{2\tau}\|x_1 - x_1^k\|^2 - \frac{1}{2}\|(A,-A)x_1 - (A,-A)x_1^k\|^2 \end{array} \right\} \\ = \ & \operatorname{argmin}_{x_1 \in \mathcal{R}_+^{2n}}\{(x_1)^\top(A,-A)^\top((A,-A)x_1^k - y) - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 + \frac{1}{2\tau}\|x_1 - x_1^k\|^2\}. \end{aligned}$$

In addition, if $0 < \tau < 1/\lambda_{\max}((A, -A)^\top (A, -A))$, then R is a positive definite matrix.

Using the newly generated x_1^{k+1}, the Lagrange multiplier λ is updated via

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(x_1^{k+1} - x_2^k). \tag{3.7}$$

Using the updated x_1^{k+1}, $\lambda^{k+\frac{1}{2}}$, the x_2-subproblem in (2.3) is given

$$\begin{aligned} x_2^{k+1} &= \operatorname{argmin}_{x_2 \in \mathcal{R}_+^{2n}} \{\rho(e^\top, e^\top)x_2 + (\lambda^{k+\frac{1}{2}})^\top x_2 + \frac{\beta}{2}\|x_2 - x_1^{k+1}\|^2\} \\ &= \operatorname{argmin}_{x_2 \in \mathcal{R}_+^{2n}} \{\frac{\beta}{2}\|x_2 - (x_1^{k+1} - \frac{1}{\beta}\lambda^{k+\frac{1}{2}} - \frac{\rho}{\beta}(e; e))\|^2\} \end{aligned}$$

and its solution is given by

$$x_2^{k+1} = (x_1^{k+1} - \frac{1}{\beta}\lambda^{k+\frac{1}{2}} - \frac{\rho}{\beta}(e; e))_+. \tag{3.8}$$

Using the newly generated x_1^{k+1}, $\lambda^{k+\frac{1}{2}}$ and x_2^{k+1}, the Lagrange multiplier λ is second updated via

$$\lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(x_1^{k+1} - x_2^{k+1}). \tag{3.9}$$

4. Numerical Experiments

In this section, we conduct some numerical experiments about compressive sensing to verify the efficiency of the proposed PPRSM, and compared it with the strictly contractive Peaceman-Rachford splitting method (SCPRSM) in [4]. All the code were written by Matlab 7.0 and were performed on a ThinkPad computer equipped with Windows XP, 997MHz and 4.00 GB of memory.

For two methods, the stop criterion is

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

where f_k denotes the function value at iteration x_k. All the initial points are set as $[A^\top y; 0_{n\times 1}]$.

Firstly, we use PPRSM and SCPRSM to recover a simulated sparse signal from the observation data corrupted by additive Gaussian white noise, where $n = 1000, m = \text{floor}(\gamma \times n)$, $k = \text{floor}(\sigma \times m)$. Therefore k is the number of random nonzero elements contained in the original signal. In addition, we set $\gamma = 0.3, \sigma = 0.2$, $y = Ax + s_w$, where s_w is the additive Gaussian white noise of zero mean and standard derivation 0.01, $\beta = \text{mean}(|y|)$, $\tau = 1.1$, $\rho = 0.01$, and A is generated by:

$$B = \text{randn}(m, n), [Q, R] = \text{qr}(B^\top, 0), A = Q^\top.$$

Define

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

where $\tilde{x}$ denotes the reconstructive signal. The original signal, the measurement and the reconstructed signal by SCPRSM and PPRSM are given in Figure 1. Compared the first and the last two subplots in Figure 1, we clearly see that the original signal is recovered almost exactly by the test two methods. In addition, the RelErr of SCPRSM is a little smaller than that of our proposed method. In the following, we shall compare the two methods with respect to the computing time, the number of iterations and the RelErr. The parameters are set just same as the above discussion except γ and σ, and for SCPRSM, we use the same parameters as PPRSM. The codes of the two methods are repeatedly run 20 times, and the average numerical results are listed in Table 1.

From Table 1, we conclude that both methods are efficient in reconstructing the given sparse signals, and they attained the solutions successfully with comparable RelErr. However, the computing time of the PPRSM is a little less than that of the SCPRSM. Thus, we conclude that the PPRSM provides a valid approach for solving CS, and it is competitive with the SCPRSM.

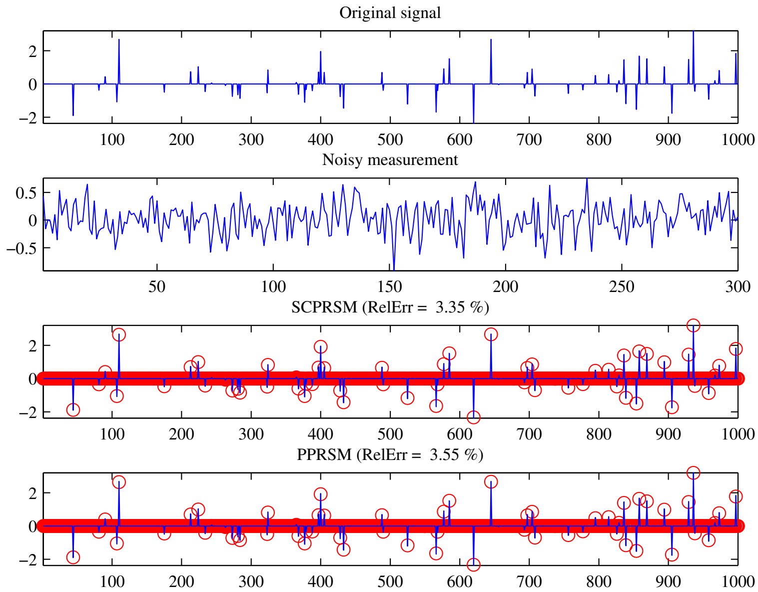

Figure 1: The original signal, noisy measurement and reconstruction results

γ	σ	PPRSM			SCPRSM		
		Time	Iter	RelErr	Time	Iter	RelErr
0.3	0.2	1.0772	255.8500	0.0492	1.0475	50.0000	0.0452
0.2	0.2	1.0608	370.8500	0.0837	1.0624	61.6500	0.0752
0.2	0.1	0.4017	145.8500	0.0562	1.0585	60.5500	0.0571
Average		0.8466	257.5167	0.0630	1.0561	57.4000	0.0592

Table 1: Comparison of PPRSM with SCPRSM

5. Conclusions

In this paper, we developed a new proximal Peaceman-Rachford splitting method (PPRSM), which does not need to compute the inverse of large matrix. Under mild conditions, we proved its global convergence. Numerical results of compressive sensing indicate that the new method is efficient for the compressive sensing.

Acknowledgements:

The first author was supported by the Natural Science Foundation of China (Nos. 11171180, 11101303, and 11271226), the Specialized Research Fund for the Doctoral Program of Chinese Higher Education (20113705110002), Shandong Province Science and Technology Development Projects (2013GGA13034), the Domestic Visiting Scholar Project for the Outstanding Young Teacher of Shandong Provincial Universities(2013,2014), and the Applied Mathematics Enhancement Program of Linyi University. The second author was supported by the foundation of Scientific Research Project of Shandong Universities (J15LI11), and the Foundation of Zaozhuang University Grants (2014YB03).

References

[1] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York-London, (1982). 1, 1

[2] S. Cao, Y. Xiao, H. Zhu, *Linearized alternating directions method for l_1-norm inequality constrained l_1-norm minimization*, Appl. Numer. Math., **85** (2014), 142–153. 1

[3] D. R. Han, X. M. Yuan, *Convergence analysis of the Peaceman-Rachford splitting method for nonsmooth convex optimization*, J. Optim. Theory Appl., (Under-revision). 1

[4] B. He, H. Liu, Z. Wang, X. Yuan, *A strictly contractive Peaceman-Rachford splitting method for convex programming*, SIAM J. Optim., **24** (2014), 1011–1040. 1, 1, 4

[5] M. Li, X. Yuan, *A strictly contractive Peaceman-Rachford splitting method with logarithmic-quadratic proximal regularization for convex programming*, Math. Oper. Res., **2015** (2015). 1

[6] P. L. Lions, B. Mercier, *Splitting algorithms for the sum of two nonlinear operators*, SIAM J. Num. Anal., **16** (1979), 964–979. 1

[7] D. H. Peaceman, H. H. Rachford, *The numerical solution of parabolic elliptic differential equations*, SIAM J. Appl. Math., **3** (1955), 28–41. 1

[8] M. Sun, J. Liu, *A proximal Peaceman-Rachford splitting method for compressive sensing*, J. Appl. Math. Comput., (2015), 1–15. 1