J. Appl. Math. Comput. (2016) 50:349–363
DOI 10.1007/s12190-015-0874-x

ORIGINAL RESEARCH

A proximal Peaceman–Rachford splitting method for compressive sensing

Min Sun · **Jing Liu**

Received: 14 October 2014 / Published online: 15 February 2015

Abstract Recently, He et al. proposed a modified Peaceman–Rachford splitting method (MPRSM) for separable convex programming, which includes compressive sensing (CS) as a special case. In this paper, we further study MPRSM for CS, and regularize its first subproblem by the proximal regularization. Thus the computational load of the subproblem is substantially alleviated. That is, it is easy enough to have a closed-form solution for CS. Convergence of the new method can be guaranteed under the same assumptions as MPRSM. Finally, numerical results, including comparisons with MPPSM are reported to demonstrate the efficiency of the new method.

Keywords Proximal Peaceman–Rachford splitting method · Compressive sensing · Global convergence

Mathematics Subject Classfication 90C25 · 90C30

1 Introduction

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$_\mu$) problem, which can be depicted as

M. Sun (✉)
School of Mathematics and Statistics, Zaozhuang University, Zaozhuang 277160, Shandong, China
e-mail: ziyouxiaodou@163.com

J. Liu
School of Mathematics and Statistics, Zhejiang University of Finance and Economics,
Hangzhou 310018, China

$$\min_{x\in\mathcal{R}^n} \frac{1}{2}\|Ax-y\|_2^2+\mu\|x\|_1, \tag{1}$$

where $\mu > 0$ is the regularization parameter and $\|x\|_1$ is the l_1-norm of the vector x defined as $\|x\|_1 = \sum_{i=1}^n |x_i|$. Throughout this paper, we assume that the solution set of (1) is nonempty.

It is obviously that QP_μ is a special case of the famous separable convex programming, which is studied intensively by many researchers [4,10–12]. In fact, by setting $x_1 = x, x_2 = x$, we can reformulate (1) as

$$\begin{aligned} &\min\ \frac{1}{2}\|Ax_1-y\|_2^2+\mu\|x_2\|_1\\ &\text{s.t. } x_1 - x_2 = 0,\\ &\qquad x_1\in\mathcal{R}^n, x_2\in\mathcal{R}^n. \end{aligned} \tag{2}$$

Thus, all the numerical methods which can solve the separable convex programming are applicable to the above QP_μ, including the inexact/linearized alternating directions method[6,7], the Peaceman–Rachford splitting method (PRSM) of multipliers[3,4], etc. In this paper, we are going to study the PRSM for CS.

Applying the PRSM [1,2] to (2), Bertsekas [3] obtained an iterative scheme of PRSM for (1), which is always efficient when it is convergent. However, according to [3], it "is less 'robust' in that it converges under more restrictive assumptions than alternating direction method of multipliers (ADMM)". Here, ADMM is another efficient method for (1) [6–8]. To guarantee the convergence of PPSM in [3] under mild conditions, He et al. [4] developed a modified Peaceman–Rachford splitting method (MPRSM) by attaching an underdetermined relaxation factor α to the penalty parameter β in the steps of Lagrange multiplier updating, and yield the following iterative scheme

$$\begin{cases} x_1^{k+1} = (A^\top A+\beta I)^{-1}(A^\top y+\beta x_2^k+\lambda^k),\\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(x_1^{k+1}-x_2^k),\\ x_2^{k+1} = \text{shrink}_{\frac{\mu}{\beta}}(x_1^{k+1}-\frac{1}{\beta}\lambda^{k+\frac{1}{2}}),\\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(x_1^{k+1}-x_2^{k+1}), \end{cases} \tag{3}$$

where the parameter $\alpha \in (0, 1)$, $\beta > 0$ is a penalty parameter, and for any $c > 0$, $\text{shrink}_c(\cdot)$ is the soft-thresholding operator defined as

$$\text{shrink}_c(g) := g - \min\{c, |g|\}\frac{g}{|g|}, \forall g\in\mathcal{R}^n,$$

and $(g/|g|)_i$ should be taken 0 if $|g|_i = 0$. The global convergence of MPRSM can be guaranteed under some standard assumptions and its efficiency was verified numerically in [4].

Obviously, MPRSM has a shortcoming. That is we need to compute the matrix $(A^\top A+\beta I)^{-1}$, where $A\in\mathcal{R}^{m\times n}$ is the sensing matrix and I is the identity matrix, which is quite time consuming when the dimension n is large. In order to solve this issue, motivated the linearized ADMM in [5–7,9], we propose a proximal Peaceman–Rachford splitting method (PPRSM), whose main idea is to regularize the equivalent minimization problem of x_1's iterative scheme in (3) by the proximal regularization $\frac{1}{2}\|x_1 - x_1^k\|_R^2$. Here $R \in \mathcal{R}^{n\times n}$ is defined as $R = \frac{1}{\tau} I_n - A^\top A$, where the parameter τ is restricted to $\tau \in (0, 1/\lambda_{\max}(A^\top A))$ to ensure that R is a positive definite matrix. By doing so, the proposed PPRSM does not need to compute the matrix $(A^\top A + \beta I)^{-1}$.

The paper is organized as follows. In Sect. 2, we characterize problem (2) by a mixed variational inequality problem and summarize some useful preliminaries. In Sect. 3, we describe the PPRSM for CS and prove its global convergence in detail. In Sect. 4, some numerical experiments and comparisons with MPRSM in CS are given to illustrate the efficiency of the proposed method. Finally, some concluding remarks are drawn in Sect. 5.

2 Preliminaries

In this section, we characterize problem (2) by a mixed variational inequality problem and summarize some preliminaries which are useful for further discussions.

First, we define some auxiliary variables and functions: $x = (x_1, x_2)$, $w = (x, \lambda)$, $\theta_1(x) = \frac{1}{2}\|Ax_1 - y\|_2^2$, $\theta_2(x) = \mu\|x_2\|_1$ and $\theta(x) = \theta_1(x) + \theta_2(x)$. Then, by invoking the first-order optimality condition for convex programming, we can reformulate problem (2) as the following mixed variational inequality problem (denoted by MVI($\mathcal{W}$, F, θ)): Finding a vector $w^* \in \mathcal{W}$ such that

$$\theta(x) - \theta(x^*) + (w - w^*)^\top F(w^*) \geq 0, \quad \forall w \in \mathcal{W}, \tag{4}$$

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

$$w = \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix} \quad \text{and} \quad F(w) = \begin{pmatrix} -\lambda \\ \lambda \\ x_1 - x_2 \end{pmatrix}. \tag{5}$$

Since the mapping $F(w)$ defined in (5) is affine with a skew-symmetric matrix, it is monotone. We denote by $\mathcal{W}^*$ the set of such w^* that satisfies (4). Then, $\mathcal{W}^*$ is nonempty under nonempty assumption onto the solution set of problem (1).

Now, let us define some matrices in order to present our analysis in a compact way. Let

$$M = \begin{pmatrix} I_n & 0 & 0 \\ 0 & I_n & 0 \\ 0 & \alpha\beta I_n & 2\alpha I_n \end{pmatrix}, \quad \text{and} \quad Q = \begin{pmatrix} R & 0 & 0 \\ 0 & \beta I_n & \alpha I_n \\ 0 & I_n & \frac{1}{\beta} I_n \end{pmatrix}, \tag{6}$$

where $R \in \mathcal{R}^{n\times n}$ is the positive definite matrix defined in the introduction. Then, let

$$H = \begin{pmatrix} R & 0 & 0 \\ 0 & \frac{2-\alpha}{2}\beta I_n & \frac{1}{2} I_n \\ 0 & \frac{1}{2} I_n & \frac{1}{2\alpha\beta} I_n \end{pmatrix}. \tag{7}$$

The matrices M, Q, H just defined satisfy the following properties.

Lemma 2.1 *(1) The matrices M, Q, H defined, respectively, in (6), (7) have the following relationship:*

$$HM = Q. \tag{8}$$

(2) The matrix H defined in (7) is positive definite.

Proof (1) By (6) and (7), we have

$$HM = \begin{pmatrix} R & 0 & 0 \\ 0 & \frac{2-\alpha}{2}\beta I_n & \frac{1}{2}I_n \\ 0 & \frac{1}{2}I_n & \frac{1}{2\alpha\beta}I_n \end{pmatrix} \begin{pmatrix} I_n & 0 & 0 \\ 0 & I_n & 0 \\ 0 & \alpha\beta I_n & 2\alpha I_n \end{pmatrix}$$

$$= \begin{pmatrix} R & 0 & 0 \\ 0 & \beta I_n & \alpha I_n \\ 0 & I_n & \frac{1}{\beta}I_n \end{pmatrix} = Q.$$

Then the first assertion is proved.

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

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

Since the matrix

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

is positive definite if $\alpha \in (0, 1)$, H is also positive definite. The proof is complete. □

3 Algorithm and global convergence

In this section, we further describe our motivation and then present the PPRSM for MVI$(\mathcal{W}, F, \theta)$. We also establish the new method's global convergence in a contraction perspective in this section.

The equivalent minimization problem of x_1's iterative scheme in MPRSM (3) is as follows

$$x_1^{k+1} = \text{argmin}_{x_1\in\mathcal{R}^n} \left\{ \frac{1}{2}\|Ax_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 \right\}, \tag{9}$$

and its closed-form solution is just the x_1's iterative scheme in MPRSM (3)

$$x_1^{k+1} = (A^\top A + \beta I_n)^{-1}(A^\top y + \lambda^k + \beta x_2^k).$$

However, as pointed in the introduction, the computation of $(A^\top A + \beta I_n)^{-1}$ is very time consuming when n is large. Then, we linearize $\frac{1}{2}\|Ax_1 - y\|_2^2$ at the current point x_1^k and add a proximal term, i.e.,

$$\begin{aligned}&\frac{1}{2}\|Ax_1 - y\|_2^2\\ &\quad\approx \frac{1}{2}\|Ax_1^k - y\|_2^2 + (g^k)^\top(x_1 - x_1^k) + \frac{1}{2\tau}\|x_1 - x_1^k\|^2,\end{aligned}$$

where $g^k = A^\top(Ax_1^k - y)$ denotes the gradient at x_1^k, and $\tau > 0$ is a parameter. Thus, (9) is approximated by the following problem

$$\begin{aligned}x_1^{k+1} = \operatorname{argmin}_{x_1\in\mathcal{R}^n} &\left\{\frac{1}{2}\|Ax_1^k - y\|_2^2 + (g^k)^\top(x_1 - x_1^k)\right.\\ &\left.+\frac{1}{2\tau}\|x_1 - x_1^k\|^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2\right\},\end{aligned}$$

which can be written as

$$x_1^{k+1} = \operatorname{argmin}_{x_1\in\mathcal{R}^n}\left\{(g^k)^\top x_1 + \frac{1}{2\tau}\|x_1 - x_1^k\|^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2\right\}. \tag{10}$$

Obviously, the above minimization problem has the following closed-form solution

$$x_1^{k+1} = \frac{\tau}{1+\beta\tau}\left(\lambda^k + \frac{1}{\tau}x_1^k + \beta x_2^k - g^k\right). \tag{11}$$

In the following, we show that (10) is exactly the x_1-subproblem of (9) regularized by the proximal regularization $\frac{1}{2}\|x_1 - x_1^k\|_R^2$ with $R = \frac{1}{\tau}I_n - A^\top A$. In fact, if we regularize the x_1-subproblem of (9) by $\frac{1}{2}\|x_1 - x_1^k\|_R^2$, then we get

$$\begin{aligned}&\operatorname{argmin}_{x_1\in\mathcal{R}^n}\left\{\frac{1}{2}\|Ax_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 + \frac{1}{2}\|x_1 - x_1^k\|^2_{\frac{1}{\tau}I_n - A^\top A}\right\}\\ &= \operatorname{argmin}_{x_1\in\mathcal{R}^n}\left\{\frac{1}{2}\|Ax_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2\right.\\ &\qquad\left.+\frac{1}{2}(x_1 - x_1^k)^\top\left(\frac{1}{\tau}I_n - A^\top A\right)(x_1 - x_1^k)\right\}\\ &= \operatorname{argmin}_{x_1\in\mathcal{R}^n}\left\{\frac{1}{2}\|Ax_1 - y\|_2^2 - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2\right.\\ &\qquad\left.+\frac{1}{2\tau}\|x_1 - x_1^k\|^2 - \frac{1}{2}\|Ax_1 - Ax_1^k\|^2\right\}\end{aligned}$$

$$= \operatorname{argmin}_{x_1 \in \mathcal{R}^n} \left\{ (x_1)^\top A^\top (Ax_1^k - y) - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 + \frac{1}{2\tau}\|x_1 - x_1^k\|^2 \right\}.$$

The last expression is just (10). Now, we describe our new method for CS in detail.

Algorithm 3.1 *PPRSM*

Step 0. Choose the parameters $\alpha \in (0, 1)$, $\beta > 0$, $0 < \tau < 1/\lambda_{\max}(A^\top A)$, the tolerance $\varepsilon > 0$ and the initial iterate $(x_1^0, x_2^0, \lambda^0) \in \mathcal{W}$. Set $R = \frac{1}{\tau} I_n - A^\top A$ and $k := 0$.

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

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1 \in \mathcal{R}^n} \left\{ \theta_1(x_1) - (\lambda^k)^\top x_1 + \frac{\beta}{2}\|x_1 - x_2^k\|^2 + \frac{1}{2}\|x_1 - x_1^k\|_R^2 \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta \left(x_1^{k+1} - x_2^k \right), \\ x_2^{k+1} = \operatorname{argmin}_{x_2 \in \mathcal{R}^n} \left\{ \theta_2(x_2) + \left(\lambda^{k+\frac{1}{2}} \right)^\top x_2 + \frac{\beta}{2}\|x_1^{k+1} - x_2\|^2 \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta \left(x_1^{k+1} - x_2^{k+1} \right), \end{cases} \tag{12}$$

Step 2. If

$$\max \left\{ \|x_1^k - x_1^{k+1}\|, \|x_2^k - x_2^{k+1}\|, \|\lambda^k - \lambda^{k+1}\| \right\} < \epsilon, \tag{13}$$

then stop and return an approximate solution $(x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ of MVI($\mathcal{W}, F, \theta$); else set $k := k + 1$, and goto Step 1.

For further analysis, similar to [4], we also define an auxiliary sequence $\{\hat{w}^k\}$ as

$$\hat{w}^k = \begin{pmatrix} \hat{x}_1^k \\ \hat{x}_2^k \\ \hat{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^k - \beta(x_1^{k+1} - x_2^k) \end{pmatrix}. \tag{14}$$

Thus, from [4], we also get

$$\lambda^{k+\frac{1}{2}} = \lambda^k - \alpha(\lambda^k - \hat{\lambda}^k), \quad \text{and} \quad \lambda^{k+1} = \lambda^k - \left[2\alpha(\lambda^k - \hat{\lambda}^k) + \alpha\beta(x_2^k - \hat{x}_2^k) \right],$$

which together with (6) and (14) shows that

$$w^{k+1} = w^k - M(w^k - \hat{w}^k). \tag{15}$$

Now, we start to prove the global convergence of PPRSM, and firstly we show the stopping criterion (13) is reasonable.

Lemma 3.1 *If $x_i^k = x_i^{k+1} (i = 1, 2)$ 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 MVI($\mathcal{W}, F, \theta$).*

Proof By deriving the first-order optimality condition of x_1-subproblem in (12), for any $x_1 \in \mathcal{R}^n$, we have

$$\theta_1(x_1) - \theta_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top \left\{-\lambda^k + \beta(x_1^{k+1} - x_2^k) + R(x_1^{k+1} - x_1^k)\right\} \geq 0.$$

By the definition of $\hat{\lambda}^k$ in (14), the above inequality can be written as

$$\theta_1(x_1) - \theta_1(x_1^{k+1}) + (x_1 - x_1^{k+1})^\top \left\{-\hat{\lambda}^k + R(x_1^{k+1} - x_1^k)\right\} \geq 0, \forall x_1 \in \mathcal{R}^n. \quad (16)$$

Similarly, from the x_2-subproblem in (12), we have

$$\theta_2(x_2) - \theta_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top \left\{\hat{\lambda}^k + \beta(x_2^{k+1} - x_2^k) + \alpha(\hat{\lambda}^k - \lambda^k)\right\} \geq 0, \forall x_2 \in \mathcal{R}^n. \quad (17)$$

In addition, follows from (14) again, we have

$$(x_1^{k+1} - x_2^{k+1}) + (x_2^{k+1} - x_2^k) + \frac{1}{\beta}(\hat{\lambda}^k - \lambda^k) = 0. \quad (18)$$

Then, combining (16)–(18) 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

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

Then, recall the definition of Q in (6), the above inequality can be written as

$$\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k) \geq (w - \hat{w}^k)^\top Q(w^k - \hat{w}^k), \quad (19)$$

for any $w \in \mathcal{W}$. In addition, if $x_i^k = x_i^{k+1} (i = 1, 2)$ and $\lambda^k = \lambda^{k+1}$, then we have $x_i^k = \hat{x}_i^k (i = 1, 2)$ and $\lambda^k = \hat{\lambda}^k$. Thus,

$$Q(w^k - \hat{w}^k) = 0,$$

which together with (19) shows that

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

This implies that $\hat{w}^k = (\hat{x}_1^k, \hat{x}_2^k, \hat{\lambda}^k)$ is a solution of MVI($\mathcal{W}, F, \theta$). Since $\hat{w}^k = w^{k+1}$, therefore w^{k+1} is also a solution of MVI($\mathcal{W}, F, \theta$). This completes the proof. □

Now, we deal with the right-hand side of (19), and we want to find a lower bound in terms of the discrepancy between $\|w - w^{k+1}\|_H^2$ and $\|w - w^k\|_H^2$ for any $w \in \mathcal{W}$.

Lemma 3.2 *Let the sequence $\{w^k\}$ be generated by PPRSM. Then, for any $w \in \mathcal{W}$, we have*

$$(w-\hat{w}^k)^\top Q(w^k-\hat{w}^k) \geq \frac{1}{2}\Big(\|w-w^{k+1}\|_H^2 - \|w-w^k\|_H^2\Big) + \frac{1}{2}\|w^k-\hat{w}^k\|_N^2, \quad (20)$$

where

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

Proof Applying the identity

$$(a-b)^\top H(c-d) = \frac{1}{2}\Big(\|a-d\|_H^2 - \|a-c\|_H^2\Big) + \frac{1}{2}\Big(\|c-b\|_H^2 - \|d-b\|_H^2\Big),$$

with

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

we obtain

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

Combining the above equality, (8) and (15), we have

$$\begin{aligned}(w-\hat{w}^k)^\top Q(w^k-\hat{w}^k) &= \frac{1}{2}\Big(\|w-w^{k+1}\|_H^2 - \|w-w^k\|_H^2\Big) \\ &+ \frac{1}{2}\Big(\|w^k-\hat{w}^k\|_H^2 - \|w^{k+1}-\hat{w}^k\|_H^2\Big). \quad (21)\end{aligned}$$

For the last term of (21), we have

$$\begin{aligned}&\|w^k-\hat{w}^k\|_H^2 - \|w^{k+1}-\hat{w}^k\|_H^2 \\ &= \|w^k-\hat{w}^k\|_H^2 - \|(w^k-\hat{w}^k)-(w^k-w^{k+1})\|_H^2 \\ &= \|w^k-\hat{w}^k\|_H^2 - \|(w^k-\hat{w}^k)-M(w^k-\hat{w}^k)\|_H^2 \text{(Using(15))} \\ &= 2(w^k-\hat{w}^k)^\top HM(w^k-\hat{w}^k) - (w^k-\hat{w}^k)^\top M^\top HM(w^k-\hat{w}^k) \\ &= (w^k-\hat{w}^k)(Q^\top + Q - M^\top HM)(w^k-\hat{w}^k). \quad (22)\end{aligned}$$

Then, by (6)–(8) and a simple manipulation, we can get

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

Thus, it follows from the Cauchy–Schwartz Inequality that

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

Then, the above inequality and (22) indicate that

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

Substituting this inequality into (21), we can get (20). The proof is complete. □

Theorem 3.1 *Let $\{w^k\}$ be the sequence generated by PPRSM. Then, for any $w \in \mathcal{W}$, we have*

$$\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(w) \geq \frac{1}{2}(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2) + \frac{1}{2}\|w^k - \hat{w}^k\|_N^2. \tag{23}$$

Proof First, combining (19) and (20), we get

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

From the monotonicity of $F(\cdot)$, we have

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

Adding the above two inequalities, we obtain the assertion (23). Hence, the theorem is proved. □

With the above theorem, we are now ready to establish the global convergence of PPRSM for solving MVI($\mathcal{W}$, F, θ).

Theorem 3.2 *Let $\{w^k\}$ be the sequence generated by PPRSM. Then, the sequence $\{w^k\}$ converges to some w^∞, which belongs to $\mathcal{W}^*$.*

Proof Setting $w = w^*$ in (23), we have

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

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

$$\|w^{k+1} - w^*\|_H^2 \leq \|w^k - w^*\|_H^2 - \|w^k - \hat{w}^k\|_N^2. \tag{24}$$

Summing over $k = 0, 1, \ldots, \infty$, it yields

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

which implies that

$$\lim_{k\to\infty} \|w^k - \hat{w}^k\|_N = 0. \tag{25}$$

This indicates that

$$\lim_{k\to\infty} \|x_1^k - \hat{x}_1^k\| = 0, \lim_{k\to\infty} \|x_2^k - \hat{x}_2^k\| = 0, \text{ and } \lim_{k\to\infty} \|\lambda^k - \hat{\lambda}^k\| = 0.$$

This together with the definition of Q implies that

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

Then, by (19), we can get

$$\lim_{k\to\infty} \left\{\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k)\right\} \geq 0, \quad \forall w \in \mathcal{W}. \tag{26}$$

On the other hand, from (24) again, we have

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

which indicates that the sequence $\{w^k\}$ is bounded. Then, from (25), the sequence $\{\hat{w}^k\}$ is also bounded. Therefore, it has at least one cluster point. Let w^∞ be a cluster point of $\{\hat{w}^k\}$ and the subsequence $\{\hat{w}^{k_j}\}$ converges to w^∞. It follows from (26) that

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

which implies that $w^\infty \in \mathcal{W}^*$. From $\lim_{k\to\infty} \|w^k - \hat{w}^k\|_N = 0$, we can deduce $\lim_{k\to\infty} \|w^k - \hat{w}^k\|_H = 0$, which together with $\{\hat{w}^{k_j}\} \to w^\infty$ implies that, for any given $\epsilon > 0$, there exists an integer l, such that

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

Therefore, for any $k \geq k_l$, it follows from the above two equalities and (24) that

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

This shows that the sequence $\{w^k\}$ converges to $w^\infty \in \mathcal{W}^*$. This completes the proof.□

4 Numerical experiments

In this section, we conduct some numerical experiments about CS to verify the efficiency of the proposed PPRSM, and compared it with the MPRSM in [4]. All the code were written by Matlab 7.0 and were performed on a ThinkPad computer equipped with Windows XP, 2.60GHz and 1.96 GB of memory.

Obviously, for CS, (12) in PPRSM can be written in the following compact form

$$\begin{cases} x_1^{k+1} = \frac{\tau}{1+\beta\tau}(\lambda^k + \frac{1}{\tau}x_1^k + \beta x_2^k - g^k), \\ \lambda^{k+\frac{1}{2}} = \lambda^k - \alpha\beta(x_1^{k+1} - x_2^k), \\ x_2^{k+1} = \text{shrink}_{\frac{\mu}{\beta}}(x_1^{k+1} - \frac{1}{\beta}\lambda^{k+\frac{1}{2}}), \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \alpha\beta(x_1^{k+1} - x_2^{k+1}), \end{cases} \tag{27}$$

where $g^k = A^\top(Ax_1^k - y)$.

For two methods, the stop criterion is

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

where f_k denotes the objective function value of (1) at iteration x_k. All the initial points are set as $A^\top y$, and we set $n = 1000$, $m = \text{floor}(\gamma \times n)$, $k = \text{floor}(\sigma \times m)$, where k is the number of random nonzero elements contained in the original signal. In addition, we set $\mu = 0.01$, $\alpha = 0.9$, and the sensing matrix A is generated by:

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

4.1 Sensitivity to τ

In this subsection, we are going to test the sensitivity of τ for the PPRSM (27). We set $\gamma = 0.3$, $\sigma = 0.2$, $y = A\bar{x} + s_w$, where s_w is the additive Gaussian white noise of zero mean and standard derivation 0.01, and $\beta = \text{mean}(|y|)$. Define

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

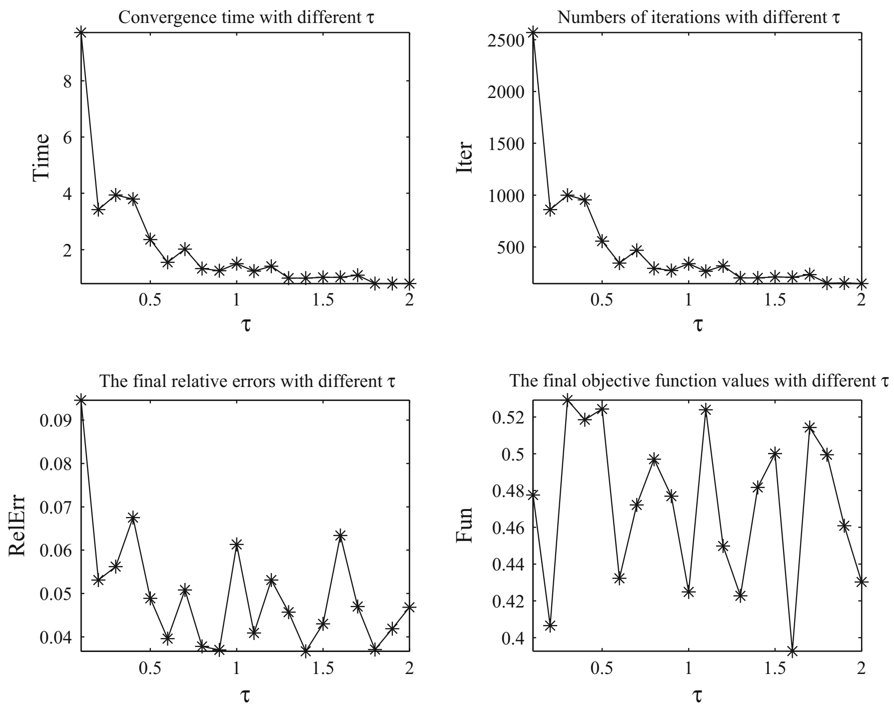

Fig. 1 Sensitivity test on the parameter τ

where $\tilde{x}$ denotes the reconstructive signal. We choose different values of τ in the interval $[0.1, 2]$. More specifically, we choose $\tau = \{0.1, 0.2, \ldots, 2\}$. The computing time (denoted by Time) in seconds, the numbers of iterations (denoted by Iter), the RelErr, and the objective function value (denoted by Fun) generated by the PPRSM (27) are recorded for each choice of τ. Then, we plot them in Fig. 1. According to the two top curves in Fig. 1, we see that the parameter τ works well when $\tau \geq 0.5$. In addition, Time and Iter have decreasing tendency when τ increases. However, other numerical results indicate that τ that bigger than 2 often leads to poor numerical performance. Therefore, based on our experiments, some values close to 2, such as $[1.8, 2]$, are preferred. In the following, we set $\tau = 2$.

4.2 Test on additive Gaussian white noise

In this subsection, we use PPRSM to recover a simulated sparse signal from the observation data corrupted by additive Gaussian white noise. Here, parameter values are set just the same to the previous. The original signal, the measurement and the reconstructed signal by PPRSM are given in Fig. 2. Compared the first and the last plots in Fig. 1, we clearly see that the original signal is recovered almost exactly. In addition, the RelErr=4.92%, the computing time is 0.7810, the number

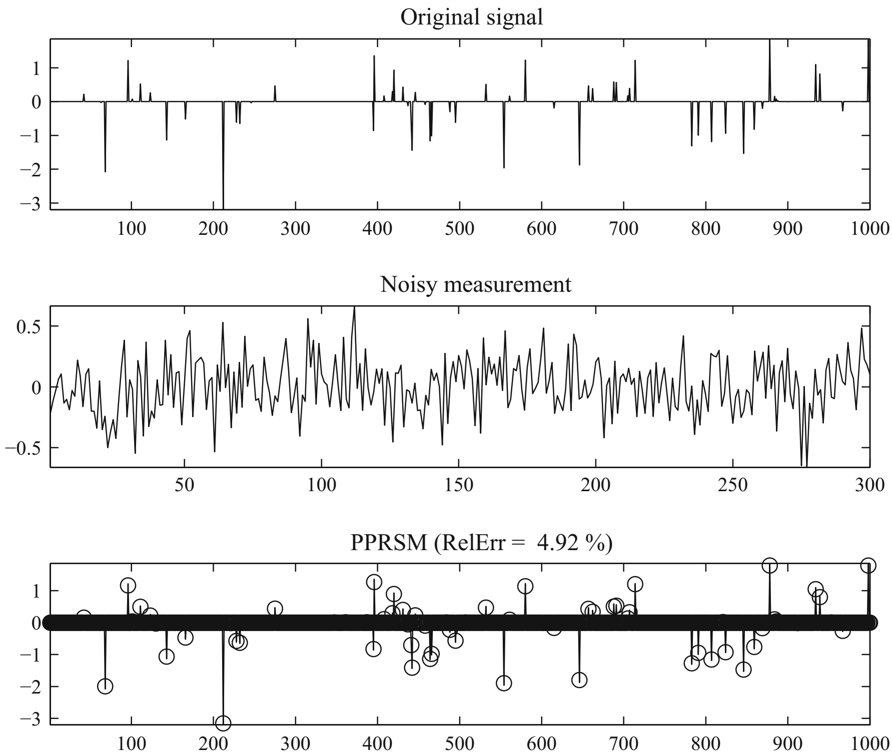

Fig. 2 The original signal, noisy measurement and reconstruction results

of iteration is 143. All together, this simple experiment shows that PPRSM works well.

4.3 Test on PPRSM and MPRSM

In this subsection, we compare PPRSM with MPRSM with respect to the RelErr, the computing time etc. The parameters are set just the same as the above subsection except γ and σ, and for MPRSM, we use the same parameters as PPRSM. The codes of the two methods are repeatedly run 20 times with different combinations of n, γ and σ, and the numerical results are listed in Table 1.

As seen from Table 1, both methods are efficient in reconstructing the given sparse signals, and they attained the solutions successfully with comparable RelErr. However, the computing time of PPRSM is obviously less than that of MPRSM, which shows that PPRSM is faster. Meanwhile, the advantage of PPRSM becomes more clear as the dimension n increases. Taking everything all together, we conclude that PPRSM provides a valid approach for solving CS, and it performs better than MPRSM.

Table 1 Comparison of PPRSM with MPRSM

n	γ	σ	PPRSM				MPRSM			
			Time	Iter	RelErr	Fun	Time	Iter	RelErr	Fun
1000	0.3	0.2	0.8602	166.9500	0.4343	0.0495	1.4289	54.7000	0.4829	0.0453
	0.2	0.2	1.0524	265.8000	0.3032	0.0888	1.2908	56.8000	0.3097	0.0792
	0.2	0.1	0.8180	199.6500	0.1495	0.0605	1.2930	57.1000	0.1538	0.0556
Average			0.9102	210.8000	–	–	1.3376	56.2000	–	–
2000	0.3	0.2	4.2925	170.3500	0.9491	0.0430	9.8045	53.0000	0.9550	0.0485
	0.2	0.2	4.8570	274.6500	0.6035	0.0814	9.2580	65.0000	0.6812	0.0788
	0.2	0.1	3.5470	177.0000	0.3205	0.0598	8.9925	53.0000	0.3373	0.0559
Average			4.2322	207.3333	–	–	9.3517	57.0000	–	–

5 Conclusions

In this paper, based on the MPRSM recently proposed by He et al., we developed a new PPRSM, which is free from the computation of the inverse of large matrix. Under mild conditions, we proved its global convergence. Numerical results of CS indicate that the new method performs better than MPRSM. In the future, we will investigate the application of the PRSM to dual problems of CS, and design some more efficient solvers.

Acknowledgments The authors gratefully acknowledge the helpful comments and suggestions of the anonymous reviewers. This work was supported by the Foundation of Zaozhuang University Grants No. 2014YB03, Shandong Province Statistical Research Project No. 20143038, and the domestic visiting scholar project funding of Shandong Province outstanding young teachers in higher schools.

References

1. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
2. Peaceman, D.H., Rachford, H.H.: The numerical solution of parabolic elliptic differential equations. SIAM J. Appl. Math. **3**, 28–41 (1955)
3. Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Academic Press, New York (1982)
4. He, B.S., Liu, H., Wang, Z.R., Yuan, X.M.: A strictly contractive Peaceman–Rachford splitting method for convex programming. SIAM J. Optim **24**(3), 1011–1040 (2014)
5. Yang, J.F., Zhang, Y.: Alternating direction algorithms for l_1-problems in compressive sensing. SIAM J. Sci. Comput. **33**(1), 250–278 (2011)
6. Xiao, Y.H., Song, H.N.: An inexact alternating directions algorithm for constrained total variation regularized compressive sensing problems. J. Math. Imaging Vis. **44**(2), 114–127 (2012)
7. Cao, S.H., Xiao, Y.H., Zhu, H.: Linearized alternating directions method for l_1-norm inequality constrained l_1-norm minimization. Appl. Numer. Math. **85**, 142–153 (2014)
8. He, B.S., Yuan, X.M.: On the O(1/n) convergence rate of Douglas–Rachford alternating direction method. SIAM J. Numer. Anal. **50**, 700–709 (2012)
9. Yang, J.F., Yuan, X.M.: Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization. Math. Comput. **82**(281), 301–329 (2013)

10. He, B.S., Yuan, X.M.: On the direct extension of ADMM for multi-block separable convex programming and beyond: from variational inequality perspective. Manuscript, http://www.optimizationonline.org/DB_HTML/2014/03/4293.html. (2014)
11. Han, D.R., Yuan, X.M., Zhang, W.X., Cai, X.J.: An ADM-based splitting method for separable convex programming. Comput. Optim. Appl. **54**(2), 343–369 (2013)
12. Han, D.R., Yuan, X.M., Zhang, W.X.: An augmented Lagrangian based parallel splitting method for separable convex minimization with applications to image processing. Math. Comput. **83**(289), 2263–2291 (2014)