J. Appl. Math. Comput. (2016) 51:605–622
DOI 10.1007/s12190-015-0922-6

ORIGINAL RESEARCH

Generalized Peaceman-Rachford splitting method for separable convex programming with applications to image processing

Min Sun[1] · **Jing Liu**[2]

Received: 4 July 2015 / Published online: 15 August 2015

Abstract Recently, a globally convergent variant of Peaceman–Rachford splitting method (PRSM) has been proposed by He et al. In this paper, motivated by the idea of the generalized alternating direction method of multipliers, we propose, analyze and test a generalized PRSM for separable convex programming, which removes some restrictive assumptions of He's PRSM. Furthermore, both subproblems are approximated by the linearization technique, and the resulting subproblems thus may have closed-form solution, especially in some practical applications. We prove the global convergence of the proposed method and report some numerical results about the image deblurring problems, which demonstrate that the new method is efficient and promising.

Keywords Peaceman–Rachford splitting method · Convex programming · Image deblurring problems

Mathematics subject classification 90C25 · 90C30

1 Introduction

Many problems arising from compressed sensing, image deblurring, and statistical learning can be casted as the following canonical convex programming with linear constraints and a separable objective function:

✉ Min Sun
ziyouxiaodou@163.com

[1] School of Mathematics and Statistics, Zaozhuang University, Shandong 277160, China

[2] School of Mathematics and Statistics, Zhejiang University of Finance and Economics, Hangzhou 310018, China

$$\min \{\theta_1(x_1) + \theta_2(x_2) | A_1x_1 + A_2x_2 = b, x_1 \in \mathcal{X}_1, x_2 \in \mathcal{X}_2\}, \tag{1}$$

where $\theta_i : \mathcal{R}^{n_i} \to \mathcal{R}(i = 1, 2)$ are closed proper convex but not necessarily smooth functions, $A_i \in \mathcal{R}^{l\times n_i}(i = 1, 2)$, $b \in \mathcal{R}^l$ and $\mathcal{X}_i \subseteq \mathcal{R}^{n_i}(i = 1, 2)$ are nonempty closed convex sets. Throughout our discussion, the solution set of (1) is assumed to be nonempty and the matrices A_1, A_2 are assumed to have full column rank.

To solve (1), there are many efficient augmented Lagrangian method (ALM) based solvers, such as the alternating direction method of multipliers (ADMM) [1,2], the generalized ADMM [3,4], and the Peaceman–Rachford splitting method (PRSM) [5–9], and so on. Recently, it was discovered that the ADMM, the PRSM and their variants are quite efficient for solving some practical problems. For example, in the compressed sensing, the subproblems in the iterative schemes of these methods often have closed-form or easily computable solutions. Due to their high efficiency, these methods are well studied by many researchers.

In this paper, let us focus on the PRSM, which was original proposed in [5]. Applying it to problem (1), we get the following iterative scheme [6]:

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

where $\lambda \in \mathcal{R}^l$ is the Lagrange multiplier associated with the linear constraints in (1) and $\beta > 0$ is a penalty parameter. Compared with the ADMM, the PRSM updates the Lagrange multiplier twice at each iteration. The PRSM is always efficient when it is convergent. However, according to [6], it "is less 'robust' in that it converges under more restrictive assumptions than alternating direction method of multipliers". Thus, compared with the ADMM, the PRSM has received much less attention. To ensure the global convergence of the PRSM, He et al. [7] developed a strictly contractive PRSM (SCPRSM) by attaching an underdetermined relaxation factor r to the penalty parameter β in the steps of Lagrange multiplier updating, and yield the following iterative scheme:

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1\in\mathcal{X}_1} \left\{\theta_1(x_1) - (\lambda^k)^\top A_1x_1 + \frac{\beta}{2}\|A_1x_1 + A_2x_2^k - b\|^2\right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta\left(A_1x_1^{k+1} + A_2x_2^k - b\right), \\ x_2^{k+1} = \operatorname{argmin}_{x_2\in\mathcal{X}_2} \left\{\theta_2(x_2) - \left(\lambda^{k+\frac{1}{2}}\right)^\top A_2x_2 + \frac{\beta}{2}\|A_1x_1^{k+1} + A_2x_2 - b\|^2\right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - r\beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right), \end{cases} \tag{3}$$

where the parameter $r \in (0, 1)$. Obviously, when $r = 1$, the iterative scheme (3) reduces to (2). However, to ensure the global convergence of (3), the parameter r must be restricted in the interval $(0, 1)$. The global convergence of SCPRSM was proved via the analytic framework of contractive type methods and its efficiency was verified numerically by some applications in statistical learning and image processing. Furthermore, the experiments in [7] also indicate that some aggressive values of r close to 1 (e.g., [0.8, 0.9]) are preferred.

It is well known that the generalized ADMM [3,4] includes the original ADMM as a special case, and can numerically accelerate the original ADMM with some values of the relaxation factor. Therefore, inspired by the relationship between the ADMM and the generalized ADMM, we propose the following generalized version of the SCPRSM:

$$\begin{cases} x_1^{k+1} = \operatorname{argmin}_{x_1 \in \mathcal{X}_1} \left\{ \theta_1(x_1) - (\lambda^k)^\top A_1 x_1 + \frac{\beta}{2} \|A_1 x_1 + A_2 x_2^k - b\|^2 + \frac{1}{2} \|x_1 - x_1^k\|_{G_1}^2 \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k - r\beta \left(A_1 x_1^{k+1} + A_2 x_2^k - b \right), \\ x_2^{k+1} = \operatorname{argmin}_{x_2 \in \mathcal{X}_2} \left\{ \theta_2(x_2) - (\lambda^{k+\frac{1}{2}})^\top A_2 x_2 + \frac{\beta}{2} \|\alpha A_1 x_1^{k+1} - (1-\alpha)(A_2 x_2^k - b) \right. \\ \qquad \left. + A_2 x_2 - b\|^2 + \frac{1}{2} \|x_2 - x_2^k\|_{G_2}^2 \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} - \beta \left[\alpha A_1 x_1^{k+1} - (1-\alpha)(A_2 x_2^k - b) + A_2 x_2^{k+1} - b \right], \end{cases} \tag{4}$$

where $\alpha \in (0, 2)$, $r \in (0, 2 - \alpha)$ are two relaxation factors, and $G_1 \succ 0$, $G_2 \succeq 0$ or $G_1 \succeq 0$, $G_2 \succ 0$. Here for a matrix G, $G \succ 0$ (resp., $\succeq 0$) means that G is positive definite (resp., semidefinite). Note that: (I) We don't attach any relaxation factor before β in the second update of the Lagrangian multiplier. The reasons are: (i) If we attach a factor, then the matrix H (defined in (8)) will be nonsymmetric; (ii) As pointed out in [7], aggressive values of r are preferred in (3). (II) At each iteration, the iterative scheme (3) require solving a constrained subproblem with respect to x_1, which is in the form

$$\min \left\{ \theta_1(x_1) + \frac{\beta}{2} \|A_1 x_1 - a_1\|^2 | x_1 \in \mathcal{X}_1 \right\}$$

with a given vector $a_1 \in \mathcal{R}^l$. When $A_1 \neq I_{n_1}$, this subproblem is not easy to solve even when $\mathcal{X}_1 = \mathcal{R}^{n_1}$, and the resolvent of $\partial\theta_1$, that is $(I_{n_1} + \frac{1}{\beta}\partial\theta_1)^{-1}(\cdot)$, has a closed-form representation. Thus some iterative solver is needed to get an approximate solution of this subproblem. However, in this case, we can set $G_1 = \tau I_{n_1} - \beta A_1^\top A_1$ in the x_1 subproblem of (4) with the requirement $\tau > \beta \|A_1^\top A_1\|$, where $\|\cdot\|$ denotes the spectral norm of a matrix. With this choice, the quadratic term $\frac{\beta}{2}\|A_1 x_1 + A_2 x_2^k - b\|^2$ is linearized, and the x_1-subproblem in (4) reduces to estimating the resolvent of $\partial\theta_1$:

$$x_1^{k+1} = \operatorname*{argmin}_{x_1 \in \mathcal{R}^{n_1}} \left\{ \theta_1(x_1) + \frac{\tau}{2} \|x_1 - u^k\|^2 \right\},$$

where $u^k = \frac{1}{\tau}(Gx_1^k - \beta A_1^\top A_2 x_2^k + A^\top \lambda^k + \beta A_1^\top b)$, and thus it has closed-form solution. For example, when $\theta_1(x_1) = \|x_1\|_1$ (here $\|x_1\|_1 = \sum_{i=1}^{n_1} |(x_1)_i|$) or $\theta_1(x_1) = \|x_1\|$ (here $\|\cdot\|$ denotes the Euclidean 2-norm), then

$$x_1^{k+1} = \text{shrink}_{1,2}(u^k, 1/\tau) \doteq \text{sign}(u^k) \cdot \max\{0, |u^k| - 1/\tau\},$$

or

$$x_1^{k+1} = \text{shrink}_{2,2}(u^k, 1/\tau) \doteq \frac{u^k}{\|u^k\|} \cdot \max\{0, \|u^k\| - 1/\tau\},$$

where sign$(\cdot)$ is the sign function.

The rest of this paper is organized as follows. In Sect. 2, we give some preliminaries and characterize (1) by the variational inequality problem. In Sect. 3, we describe the generalized PRSM and prove its global convergence in detail. In Sect. 4, we apply the proposed method to solve the image deblurring problem and verify its numerical efficiency. Finally, we give some concluding remarks in Sect. 5.

2 Preliminaries

In this section, we summarize some preliminaries which are useful for further discussions. In particular, we characterize problem (1) by the variational inequality problem.

First, we define some auxiliary variables which will help us alleviate the notation in the following analysis. Set $x = (x_1, x_2)$, which is a column vector by stacking vectors x_1, x_2. Similarly, set $w = (x, \lambda)$ and $\theta(x) = \theta_1(x_1) + \theta_2(x_2)$. By invoking the first-order optimality condition for convex programming, we reformulate problem (1) as the following variational inequality problem (denoted by VI($\mathcal{W}, F, \theta$)): Finding a vector $w^* \in \mathcal{W}$ such that

$$\theta(x) - \theta(x^*) + (w - w^*)^\top F(w^*) \geq 0, \quad \forall w \in \mathcal{W}, \tag{5}$$

where $\mathcal{W} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{R}^l$, and

$$w = \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \lambda \end{pmatrix} \text{ and } F(w) = \begin{pmatrix} -A_1^\top \lambda \\ -A_2^\top \lambda \\ A_1 x_1 + A_2 x_2 - b \end{pmatrix}. \tag{6}$$

Obviously, the mapping $F(w)$ defined in (6) is affine with a skew-symmetric matrix; it is thus monotone. We denote by $\mathcal{W}^*$ the solution set of (5). Then, $\mathcal{W}^*$ is nonempty under assumption of the solution set of problem (1), and if $w^* = (x_1^*, x_2^*, \lambda^*) \in \mathcal{W}^*$, then (x_1^*, x_2^*) is a solution of (1).

Now, we define some matrices in order to present our analysis in a compact way. Set

$$M = \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\beta A_2 & (r+\alpha) I_l \end{pmatrix}, \text{ and } Q = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & \beta A_2^\top A_2 + G_2 & (1-r-\alpha)A_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_l \end{pmatrix}, \tag{7}$$

and

$$H = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & \frac{1}{r+\alpha}\beta A_2^\top A_2 + G_2 & \frac{1-r-\alpha}{r+\alpha} A_2^\top \\ 0 & \frac{1-r-\alpha}{r+\alpha} A_2 & \frac{1}{\beta(r+\alpha)} I_l \end{pmatrix}. \tag{8}$$

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

Lemma 2.1 *If $\alpha \in (0, 2)$, $r + \alpha < 2$ and $G_1 \succ 0$, $G_2 \succeq 0$ or $G_1 \succeq 0$, $G_2 \succ 0$, then we have*

(1). *The matrices M, Q, H defined, respectively, in (7), (8) have the following relationship:*

$$HM = Q. \tag{9}$$

(2). *If $G_1 \succ 0$, $G_2 \succeq 0$, then $H \succ 0$; If $G_1 \succeq 0$, $G_2 \succ 0$, then*

$$H_1 = \begin{pmatrix} \frac{1}{r+\alpha}\beta A_2^\top A_2 + G_2 & \frac{1-r-\alpha}{r+\alpha} A_2^\top \\ \frac{1-r-\alpha}{r+\alpha} A_2 & \frac{1}{\beta(r+\alpha)} I_l \end{pmatrix} \succ 0.$$

(3). $Q^\top + Q - M^\top H M \succeq 0$.

Proof (1) By (7) and (8), we have

$$HM = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & \frac{1}{r+\alpha}\beta A_2^\top A_2 + G_2 & \frac{1-r-\alpha}{r+\alpha} A_2^\top \\ 0 & \frac{1-r-\alpha}{r+\alpha} A_2 & \frac{1}{\beta(r+\alpha)} I_l \end{pmatrix} \begin{pmatrix} I_{n_1} & 0 & 0 \\ 0 & I_{n_2} & 0 \\ 0 & -\beta A_2 & (r+\alpha) I_l \end{pmatrix}$$
$$= \begin{pmatrix} G_1 & 0 & 0 \\ 0 & \beta A_2^\top A_2 + G_2 & (1-r-\alpha)A_2^\top \\ 0 & -A_2 & \frac{1}{\beta} I_l \end{pmatrix} = Q.$$

Then the first assertion is proved.

(2) The proof is divided into two cases.

Case (I) If $G_1 \succ 0$, $G_2 \succeq 0$. For any $w = (x_1, x_2, \lambda) \neq 0$, since $r > 0$, $\alpha \in (0, 2)$ and $r + \alpha < 2$, we get

$$\begin{aligned} & w^\top H w \\ & = \|x_1\|_{G_1}^2 + \|x_2\|_{G_2}^2 + \frac{1}{r+\alpha} \left(\beta\|A_2 x_2\|^2 + 2(1-r-\alpha)\lambda^\top A_2 x_2 + \frac{1}{\beta}\|\lambda\|^2\right) \\ & \geq \|x_1\|_{G_1}^2 + \|x_2\|_{G_2}^2 + \frac{2}{r+\alpha} \min\{2-r-\alpha, r+\alpha\} \|A_2 x_2\| \cdot \|\lambda\|, \end{aligned} \tag{10}$$

where the inequality follows from the Cauchy-Schwartz inequality. If $x_1 \neq 0$, then from (10), we have $w^\top H w > 0$. Otherwise $(x_2, \lambda) \neq 0$: if $x_2 = 0$ or $\lambda = 0$, then from the equality in (10) and the full column rank of A_2, we also have $w^\top H w > 0$; if $x_2 \neq 0$ and $\lambda \neq 0$, then from (10) again, we have $w^\top H w > 0$. Thus, $H \succ 0$.

Case (II) If $G_1 \succeq 0$, $G_2 \succ 0$. For any $v = (x_2, \lambda) \neq 0$, as the above discussion, we have $v^\top H_1 v > 0$. Thus, $H_1 \succ 0$.

(3) Using (7) and (9), we have

$$
\begin{aligned}
&Q^\top + Q - M^\top H M \\
&\quad = Q^\top + Q - M^\top Q \\
&\quad = \begin{pmatrix} G_1 & 0 & 0 \\ 0 & G_2 & 0 \\ 0 & 0 & \frac{2-(r+\alpha)}{\beta} I_l \end{pmatrix} \succeq 0,
\end{aligned}
$$

where the inequality follows from $r + \alpha < 2$ and $G_1 \succ 0$, $G_2 \succeq 0$ or $G_1 \succeq 0$, $G_2 \succ 0$. The proof is complete. □

3 Algorithm and global convergence

In this section, we first describe the generalized PRSM (denoted by GPRSM) for VI($\mathcal{W}, F, \theta$) formally. Then, we establish its global convergence in a contraction perspective.

Algorithm 1 A generalized PRSM for VI($\mathcal{W}, F, \theta$)

Input $\alpha \in (0, 1)$, $r \in (0, 2 - \alpha)$, and $G_1 \succ 0$, $G_2 \succeq 0$ or $G_1 \succeq 0$, $G_2 \succ 0$. Initialize $(x_1, x_2, \lambda) = (x_1^0, x_2^0, \lambda^0)$, $k = 0$.
while "not converged", **do**
  (1) Compute $w^{k+1} = (x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ according to (4).
  (2) $k = k + 1$.
end while
Output x_1^{k+1}, x_2^{k+1}.

For further analysis, we define an auxiliary sequence $\{\hat{w}^k\}$ as

$$
\hat{w}^k = \begin{pmatrix} \hat{x}_1^k \\ \hat{x}_2^k \\ \hat{\lambda}^k \end{pmatrix} = \begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^k - \beta(A_1 x_1^{k+1} + A_2 x_2^k - b) \end{pmatrix}. \tag{11}
$$

Thus, from (4) and (11), we get

$$
\lambda^{k+\frac{1}{2}} = \lambda^k - r(\lambda^k - \hat{\lambda}^k),
$$

and

$$\begin{aligned}\lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \beta\left[\alpha A_1\hat{x}_1^k - (1-\alpha)\left(A_2x_2^k - b\right) + A_2\hat{x}_2^k - b\right]\\ &= \lambda^k - r\left(\lambda - \hat{\lambda}^k\right) - \left[\alpha\beta\left(A_1\hat{x}_1^k + A_2x_2^k - b\right) + \beta A_2(\hat{x}_2^k - x_2^k)\right]\\ &= \lambda^k - r\left(\lambda - \hat{\lambda}^k\right) - \left[\alpha\left(\lambda^k - \hat{\lambda}^k\right) - \beta A_2\left(\hat{x}_2^k - x_2^k\right)\right]\\ &= \lambda^k - \left[(r+\alpha)\left(\lambda^k - \hat{\lambda}^k\right) - \beta A_2\left(x_2^k - \hat{x}_2^k\right)\right],\end{aligned}$$

which together with (7) and (11) shows that

$$w^{k+1} = w^k - M\left(w^k - \hat{w}^k\right). \tag{12}$$

Now, we start to prove the global convergence of GPRSM.

Lemma 3.1 *If $A_ix_i^k = A_ix_i^{k+1}(i = 1, 2)$ and $\lambda^k = \lambda^{k+1}$, then $w^{k+1} = (x_1^{k+1}, x_2^{k+1}, \lambda^{k+1})$ produced by the GPRSM is a solution of VI($\mathcal{W}, F, \theta$).*

Proof Deriving the first-order optimality condition of x_1-subproblem in (4), for any $x_1 \in \mathcal{X}_1$, we have

$$\begin{aligned}&\theta_1(x_1) - \theta_1\left(x_1^{k+1}\right) + \left(x_1 - x_1^{k+1}\right)^\top\left\{-A_1^\top\left[\lambda^k - \beta\left(A_1x_1^{k+1} + A_2x_2^k - b\right)\right]\right.\\ &\quad\left. + G_1\left(x_1^{k+1} - x_1^k\right)\right\} \geq 0.\end{aligned}$$

By the definition of $\hat{\lambda}^k$ in (11) and using $\hat{x}_1^k = x_1^{k+1}$, the above inequality can be written as

$$\theta_1(x_1) - \theta_1\left(\hat{x}_1^k\right) + \left(x_1 - \hat{x}_1^k\right)^\top\left\{-A_1^\top\hat{\lambda}^k + G_1\left(\hat{x}_1^k - x_1^k\right)\right\} \geq 0, \qquad \forall x_1 \in \mathcal{X}_1. \tag{13}$$

Similarly, from the x_2-subproblem in (4), we have

$$\begin{aligned}&\theta_2(x_2) - \theta_2(x_2^{k+1}) + \left(x_2 - x_2^{k+1}\right)^\top\left\{-A_2^\top[\lambda^{k+\frac{1}{2}} - \alpha\beta A_1x_1^{k+1} - (1-\alpha)\beta(b - A_2x_2^k)\right.\\ &\quad\left. - \beta A_2x_2^{k+1} + \beta b] + G_2\left(x_2^{k+1} - x_2^k\right)\right\} \geq 0, \quad \forall x_2 \in \mathcal{X}_2.\end{aligned} \tag{14}$$

By (11), we have

$$\begin{aligned}&\lambda^{k+\frac{1}{2}} - \alpha\beta A_1x_1^{k+1} - (1-\alpha)\beta\left(b - A_2x_2^k\right) - \beta A_2x_2^{k+1} + \beta b\\ &\quad= \lambda^k - r\left(\lambda^k - \hat{\lambda}^k\right) - \alpha\left(\lambda^k - \hat{\lambda}^k\right) - \beta A_2\left(x_2^{k+1} - x_2^k\right)\\ &\quad= \hat{\lambda}^k + (1 - r - \alpha)\left(\lambda^k - \hat{\lambda}^k\right) - \beta A_2\left(x_2^{k+1} - x_2^k\right).\end{aligned}$$

Substituting this into (14) and using $\hat{x}_2^k = x_2^{k+1}$, we have

$$\begin{aligned}&\theta_2(x_2) - \theta_2\left(\hat{x}_2^k\right) + \left(x_2 - \hat{x}_2^k\right)^\top \left\{ -A_2^\top \hat{\lambda}^k - (1-r-\alpha)A_2^\top \left(\lambda^k - \hat{\lambda}^k\right) \right. \\ &\quad \left. + \beta A_2^\top A_2 \left(\hat{x}_2^k - x_2^k\right) + G_2\left(\hat{x}_2^k - x_2^k\right)\right\} \geq 0, \forall \quad x_2 \in \mathcal{X}_2.\end{aligned} \tag{15}$$

In addition, follows from (11) again, we have

$$\left(A_1\hat{x}_1^k + A_2\hat{x}_2^k - b\right) - A_2\left(\hat{x}_2^k - x_2^k\right) + \frac{1}{\beta}\left(\hat{\lambda}^k - \lambda^k\right) = 0. \tag{16}$$

Then, combining (13), (15), (16), for any $w = (x_1, x_2, \lambda) \in \mathcal{W}$, it holds that

$$\begin{aligned}&\theta(x) - \theta\left(\hat{x}^k\right) + (w - \hat{w}^k)^\top \left\{ \begin{pmatrix} -A_1^\top \hat{\lambda}^k \\ -A_2^\top \hat{\lambda}^k \\ A_1\hat{x}_1^k + A_2\hat{x}_2^k - b \end{pmatrix} \right. \\ &\quad \left. + \begin{pmatrix} G_1(\hat{x}_1^k - x_1^k) \\ (1-r-\alpha)A_2^\top\left(\hat{\lambda}^k - \lambda^k\right) + \beta A_2^\top A_2\left(\hat{x}_2^k - x_2^k\right) + G_2\left(\hat{x}_2^k - x_2^k\right) \\ -A_2\left(\hat{x}_2^k - x_2^k\right) + \left(\hat{\lambda}^k - \lambda^k\right)/\beta \end{pmatrix} \right\} \geq 0.\end{aligned}$$

Then, recalling the definition of Q in (7), the above inequality can be written as

$$\theta(x) - \theta\left(\hat{x}^k\right) + \left(w - \hat{w}^k\right)^\top F\left(\hat{w}^k\right) \geq \left(w - \hat{w}^k\right)^\top Q\left(w^k - \hat{w}^k\right), \quad \forall w \in \mathcal{W}. \tag{17}$$

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

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

which together with (17) indicates that

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

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

According to (17), the accuracy of $\hat{w}^k$ to a solution of VI($\mathcal{W}, F, \theta$) is measured by the quantity $(w - \hat{w}^k)^\top Q(w^k - \hat{w}^k)$. In the next lemma, we further refine this term and express it in terms of some quadratic terms.

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

$$\left(w - \hat{w}^k\right)^\top Q\left(w^k - \hat{w}^k\right) \geq \frac{1}{2}\left(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2\right) + \frac{1}{2}\|w^k - \hat{w}^k\|_N^2, \tag{18}$$

where $N = Q + Q^\top - M^\top HM$.

Proof Setting $a = w, b = \hat{w}^k, c = w^k, d = w^{k+1}$ in the identity

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

we obtain

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

Combining the above inequality, (9) and (12), we have

$$\begin{aligned}(w - \hat{w}^k)^\top Q(w^k - \hat{w}^k) = &\frac{1}{2}\left(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2\right)\\ &+\frac{1}{2}\left(\|w^k - \hat{w}^k\|_H^2 - \|w^{k+1} - \hat{w}^k\|_H^2\right).\end{aligned} \tag{19}$$

For the last term of (19), we have

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

Substituting the above equality into (19), we obtain (18). The proof is complete. □

The following theorem indicates the the sequence generated by the GPRSM is Fejèr monotone with respect to $\mathcal{W}^*$.

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

$$\|w^{k+1} - w^*\|_H^2 \le \|w^k - w^*\|_H^2 - \Big(\|x_1^k - \hat{x}_1^k\|_{G_1}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \frac{2-r-\alpha}{\beta}\|\lambda^k - \hat{\lambda}^k\|^2\Big). \tag{20}$$

Proof Setting $w = w^*$ in (17) and (18), we get

$$\begin{aligned}&\theta(x^*) - \theta(\hat{x}^k) + (w^* - \hat{w}^k)^\top F(\hat{w}^k)\\ &\ge (w^* - \hat{w}^k)^\top Q(w^k - \hat{w}^k)\\ &\ge \frac{1}{2}\Big(\|w^* - w^{k+1}\|_H^2 - \|w^* - w^k\|_H^2\Big) + \frac{1}{2}\|w^k - \hat{w}^k\|_N^2.\end{aligned}$$

On the other hand, by $w^* \in \mathcal{W}^*$ and the monotonicity of F, we have

$$\begin{aligned}&\theta(\hat{x}^k) - \theta(x^*) + (\hat{w}^k - w^*)^\top F(\hat{w}^k)\\ &\ge \theta(\hat{x}^k) - \theta(x^*) + (\hat{w}^k - w^*)^\top F(\hat{w}^*)\\ &\ge 0.\end{aligned}$$

Adding the above two inequalities yields (20). The theorem is proved.

We are now ready to prove the global convergence of the proposed GPRSM. □

Theorem 3.2 *Let $\{w^k\}$ be the sequence generated by the GPRSM. Then, if $\alpha \in (0,2), r \in (0, 2-\alpha)$ and $G_1 \succeq 0, G_2 \succ 0$ or $G_1 \succ 0, G_2 \succeq 0$, the sequence $\{w^k\}$ converges to some w^∞, which belongs to $\mathcal{W}^*$.*

Proof Since $\alpha \in (0,2), r \in (0, 2-\alpha)$ and $G_1 \succ 0, G_2 \succeq 0$ or $G_1 \succeq 0, G_2 \succ 0$, by (20), we have

$$\|w^{k+1} - w^*\|_H^2 \le \|w^k - w^*\|_H^2 \le \cdots \le \|w^0 - w^*\|_H^2. \tag{21}$$

The following proof is divided into three cases.

Case (I) If $G_1 \succ 0, G_2 \succ 0$, by Lemma 2.1, we have $H \succ 0$. Then, by (21), the sequence $\{w^k\}$ is bounded. Furthermore, summing (20) over $k = 0, 1, \ldots, \infty$, it yields

$$\begin{aligned}&\sum_{k=0}^{\infty}\Big(\|x_1^k - \hat{x}_1^k\|_{G_1}^2 + \|x_2^k - \hat{x}_2^k\|_{G_2}^2 + \frac{2-r-\alpha}{\beta}\|\lambda^k - \hat{\lambda}^k\|^2\Big)\\ &\quad \le \|w^0 - w^*\|_H^2 < +\infty,\end{aligned}$$

which implies that

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

Then, by (17), we can get

$$\lim_{k\to\infty}\Big\{\theta(x) - \theta(\hat{x}^k) + (w - \hat{w}^k)^\top F(\hat{w}^k)\Big\} \ge 0, \quad \forall w \in \mathcal{W}. \tag{22}$$

Since the sequence $\{\hat{w}^k\}$ is bounded, 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 (22) 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 $\lim_{k\to\infty} \|w^k - \hat{w}^k\|_H = 0$ and $\{\hat{w}^{k_j}\} \to w^\infty$, 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 (21) 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 $w^\infty \in \mathcal{W}^*$.

Case (II) If $G_1 \succ 0$, $G_2 = 0$, by Lemma 2.1, we also have $H \succ 0$. Then, by (21), the sequence $\{w^k\}$ is also bounded. Furthermore, summing (20) over $k = 0, 1, \ldots, \infty$, it yields

$$\sum_{k=0}^{\infty} \left(\|x_1^k - \hat{x}_1^k\|_{G_1}^2 + \frac{2 - r - \alpha}{\beta} \|\lambda^k - \hat{\lambda}^k\|^2 \right) \leq \|w^0 - w^*\|_H^2 < +\infty,$$

which implies that

$$\lim_{k\to\infty} \|x_1^k - \hat{x}_1^k\| = 0, \quad \text{and} \quad \lim_{k\to\infty} \|\lambda^k - \hat{\lambda}^k\| = 0. \tag{23}$$

Note that the optimality condition for the x_2 subproblem in (4) can be written:

$$\theta_2(x_2) - \theta_2(x_2^{k+1}) + (x_2 - x_2^{k+1})^\top (-A_2^\top \lambda^{k+1}) \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \tag{24}$$

Since (24) is true for any integer $k \geq 0$, thus we have

$$\theta_2(x_2) - \theta_2(x_2^k) + (x_2 - x_2^k)^\top (-A_2^\top \lambda^k) \geq 0, \quad \forall x_2 \in \mathcal{X}_2. \tag{25}$$

Setting $x_2 = x_2^k$ and $x_2 = x_2^{k+1}$ in (24) and (25), respectively, we get

$$\theta_2(x_2^k) - \theta_2(x_2^{k+1}) + (x_2^k - x_2^{k+1})^\top (-A_2^\top \lambda^{k+1}) \geq 0,$$

and

$$\theta_2(x_2^{k+1}) - \theta_2(x_2^k) + (x_2^{k+1} - x_2^k)^\top (-A_2^\top \lambda^k) \geq 0.$$

Adding the above two inequalities, we get

$$(\lambda^k - \lambda^{k+1})^\top A_2(x_2^k - x_2^{k+1}) \geq 0. \tag{26}$$

On the other hand, from the equality above (12), we have

$$\begin{aligned}\|\lambda^k - \hat{\lambda}^k\|^2 &= \frac{1}{(r+\alpha)^2}\|(\lambda^k - \lambda^{k+1}) + \beta A_2(x_2^k - x_2^{k+1})\|^2 \\ &\geq \frac{1}{(r+\alpha)^2}(\|\lambda^k - \lambda^{k+1}\|^2 + \beta^2\|A_2(x_2^k - x_2^{k+1})\|^2),\end{aligned}$$

where the inequality follows from (26). The above inequality, (23) and $x_2^{k+1} = \hat{x}_2^k$ imply that

$$\lim_{k\to\infty} \|A_2(x_2^k - \hat{x}_2^k)\| = 0.$$

By the full column rank assumption of A_2, we get

$$\lim_{k\to\infty} \|x_2^k - \hat{x}_2^k\| = 0.$$

Therefore, by (17), we also have the inequality (22). The following proof is similar to that of Case (I).

Case (III) If $G_1 = 0$, $G_2 \succ 0$, by Lemma 2.1, we have $H_1 \succ 0$. Thus, from (21), we can deduce

$$\|v^{k+1} - v^*\|_{H_1}^2 \leq \|v^k - v^*\|_{H_1}^2 \leq \cdots \leq \|v^0 - v^*\|_{H_1}^2, \tag{27}$$

where $v = (x_2, \lambda)$. Thus the sequence $\{v^k\}$ is bounded, and

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

Thus, the sequence $\{\hat{v}^k\}$ is also bounded. On the other hand, by (11), we have

$$A_1 x_1^{k+1} = (\lambda^k - \hat{\lambda}^k)/\beta - A_2 x_2^k + b. \tag{28}$$

This and the full column rank of A_1 imply the sequences $\{x_1^k\}$ and $\{\hat{x}_1^k\}$ are also bounded. Overall, the sequences $\{w^k\}$ and $\{\hat{w}^k\}$ are bounded. In addition, by (17), we can get (22). Thus, a cluster point of the sequence $\{\hat{w}^k\}$ is a solution of VI($\mathcal{W}, F, \theta$). From $\lim_{k\to\infty} \|v^k - \hat{v}^k\|_{H_1} = 0$ and $\{\hat{v}^{k_j}\} \to v^\infty$, for any given $\epsilon > 0$, there exists an integer l, such that

$$\|v^{k_l} - \hat{v}^{k_l}\|_{H_1} < \frac{\epsilon}{2}, \text{ and } \|\hat{v}^{k_l} - v^\infty\|_{H_1} < \frac{\epsilon}{2}.$$

Therefore, for any $k \geq k_l$, it follows from the above two equalities and (27) that

$$\|v^k - v^\infty\|_{H_1} \leq \|v^{k_l} - v^\infty\|_{H_1} \leq \|v^{k_l} - \hat{v}^{k_l}\|_{H_1} + \|\hat{v}^{k_l} - v^\infty\|_{H_1} < \epsilon.$$

This shows that the sequence $\{v^k\}$ converges to v^∞. Then, by (28) again, the sequence $\{w^k\}$ converges to $w^\infty \in \mathcal{W}^*$. This completes the proof. □

4 Numerical experiments

In this section, we apply GPRSM to solving some constrained image deblurring problems, and compare its numerical performance with the linearized ADMM [10]. All the codes were written by Matlab 7.10 and were conducted on ThinkPad notebook with Pentium(R) Dual-Core CPU T4400@2.2GHz, 2GB of memory.

Given original image concatenated into an n-vector $\bar{x} \in \mathcal{R}^n$, and let $A \in \mathcal{R}^{n\times n}$ be a blurring operator (integral operator) acting on the image. Let $\omega \in \mathcal{R}^n$ be the Gaussian noise added onto the image. The observed image $c \in \mathcal{R}^n$ can be modeled by $c = A\bar{x} + \omega$. The constrained image deblurring problem (CIDP_λ) is to recover $\bar{x}$ from c, which can be depicted as

$$\min_{l\leq x\leq u} \frac{1}{2}\|Ax - c\|^2 + \frac{\mu^2}{2}\|Bx\|^2, \tag{29}$$

where $l, u \in \mathcal{R}^n_+$; $B \in \mathcal{R}^{n\times n}$ is a regularization operator (differential operator); $\mu^2 \in \mathcal{R}$ is the regularization parameter; $x \in \mathcal{R}^n$ is the restored image. The box constraints $l \leq x \leq u$ represent the dynamic range of the image (e.g., $l_i = 0$ and $u_i = 255$ for an 8-bit gray-scale image).

Obviously, introducing an auxiliary variable $y \in \mathcal{R}^n$, we can change CIDP_λ to the equivalent form

$$\begin{aligned} \min\ & \frac{1}{2}\|Ay - c\|^2 + \frac{\mu^2}{2}\|Bx\|^2 \\ \text{s.t. } & x - y = 0, \\ & x \in \Omega, y \in \mathcal{R}^n, \end{aligned} \tag{30}$$

where $\Omega = \{y | l \leq y \leq u\}$. In fact, (30) is a special case of (1) where $x_1 = y, x_2 = x$, $\theta_1(x_1) = \frac{1}{2}\|Ay - c\|^2$, $\theta_2(x_2) = \frac{\mu^2}{2}\|Bx\|^2$, $A_1 = -I_n, A_2 = I_n, d = 0$, and $\mathcal{X}_1 = \mathcal{R}^n, \mathcal{X}_2 = \Omega$. Then, the iterative scheme of the GPRSM (4) for (30) is

$$\begin{cases} y^{k+1} = \text{argmin}_{y\in\mathcal{R}^n} \left\{ \frac{1}{2}\|Ay-c\|^2 + \frac{\beta}{2}\|y-x^k+\frac{1}{\beta}\lambda^k\|^2 + \frac{1}{2}\|y-y^k\|^2_{G_1} \right\}, \\ \lambda^{k+\frac{1}{2}} = \lambda^k + r\beta(y^{k+1}-x^k), \\ x^{k+1} = \text{argmin}_{x\in\Omega} \left\{ \frac{\mu^2}{2}\|Bx\|^2 + \frac{\beta}{2}\|x-\alpha y^{k+1}-(1-\alpha)x^k - \frac{1}{\beta}\lambda^{k+\frac{1}{2}}\|^2 + \frac{1}{2}\|x-x^k\|^2_{G_2}\} \right\}, \\ \lambda^{k+1} = \lambda^{k+\frac{1}{2}} + \beta(y^{k+1}-x^{k+1}). \end{cases} \tag{31}$$

Now, let us elaborate on the two subproblems in (31).

- For the given y^k, x^k and λ^k, set $G_1 = 0$. With this choice, the y subproblem in (31) is amount to solving the normal equation

$$(A^\top A + \beta I_n)y = A^\top c + \beta x^k - \lambda^k.$$

  Under the periodic boundary conditions for y, $A^\top A$ will be a block circulant matrix with circulant blocks. Hence it can be diagonalized by the two-dimensional fast Fourier transform (FFT). Thus, the above equation can be solved exactly using two FFTs.
- For the given y^{k+1}, x^k and $\lambda^{k+\frac{1}{2}}$, set $G_2 = \mu^2(\tau I_n - B^\top B)$, where $\tau > \|B^\top B\|$, and the x subproblem in (31) reduces to

$$x^{k+1} = \text{argmin}_{x\in\Omega} \Big\{ \mu^2[(Bx^k)^\top B(x-x^k) + \frac{\tau}{2}\|x-x^k\|^2] + \frac{\beta}{2}\|x - \alpha y^{k+1} - (1-\alpha)x^k - \frac{1}{\beta}\lambda^{k+\frac{1}{2}}\|^2 \Big\}.$$

  The optimality condition of the above problem leads to the following variational inequality

$$(x' - x^{k+1})^\top \Big\{ \mu^2[B^\top Bx^k + \tau(x-x^k)] + \beta(x - \alpha y^{k+1} - (1-\alpha)x^k - \frac{1}{\beta}\lambda^{k+\frac{1}{2}}) \Big\} \geq 0, \quad \forall x' \in \Omega.$$

  Then, x^{k+1} can be given explicitly by

$$x^{k+1} = P_\Omega \left\{ \frac{1}{\mu^2\tau+\beta} \left[-\mu^2 B^\top Bx^k + \left(\beta - \alpha\beta + \mu^2\tau\right) x^k + \alpha\beta y^{k+1} + \lambda^{k+\frac{1}{2}} \right] \right\},$$

  where $P_\Omega(\cdot)$ denotes the projection operator onto Ω under the Euclidean norm.

The test images are 256-by-256 ($l_i = 0, u_i = 255$ for all $i = 1, 2, \ldots, n$) images as shown in Fig. 1: Fruits, Peppers, Lena and Cameraman. According, $n = 256^2$ in model (1) for these images. The blurring matrix A is chosen to be the out-of-focus blur and the matrix B is taken to be the gradient matrix. The observed image c is expressed as $c = A\bar{x} + \omega$, where ω is the Gaussian or impulse noise.

Fig. 1 The original test images: Fruits, Peppers, Lena and Cameraman

Here, we employ the Matlab scripts: $\texttt{A} = \texttt{fspecial}(\text{`}\texttt{average}\text{'},\texttt{alpha})$ and $\texttt{c} = \texttt{imfilter}(\texttt{x}, \texttt{A}, \text{`}\texttt{circular}\text{'},\text{`}\texttt{conv}\text{'}) + \omega$, in which $\texttt{alpha}$ is the size of the kernel. To assess the restoration performance qualitatively, we use the peak signal to noise ratio (PSNR) defined as

$$\text{PSNR} = 20\log_{10}\frac{x_{\max}}{\sqrt{\text{Var}(x,\bar{x})}} \quad \text{with} \quad \text{Var}(x,\bar{x}) = \frac{\sum_{j=1}^{n^2}[\bar{x}(j) - x(j)]^2}{n^2}.$$

Here $\bar{x}$ is the true image, and $\bar{x}_{\max}$ is the maximum possible pixel value of the image. Furthermore, the stopping criterion of all the tested methods is

$$\frac{\|x^{k-1} - x^k\|}{\|x^{k-1}\|} \le 10^{-4}.$$

In the experiment, we apply our proposed GPRSM and the linearized ADMM in [10] (denoted by LADMM) to solve model (30) with Gaussian noise and $\mu = 0.16$. Here, we set the Gaussian noise $\omega = \eta * \texttt{randn}(\texttt{n}, \texttt{n})$, and η is the level of noise. For the GPRSM, we set $\alpha = 1.1$, $\beta = 0.1$, $r = 0.89$, $\tau = 1.01 \cdot \rho(B^\top B)$. For the parameters of LADMM, we use the default values in [10]. All iterations start with the blurred images. For $\texttt{alpha} = 5$ and $\eta = 5$, the blurred images, the restored images by the two methods are showed in Fig. 2. From Fig. 2, we can see that both methods

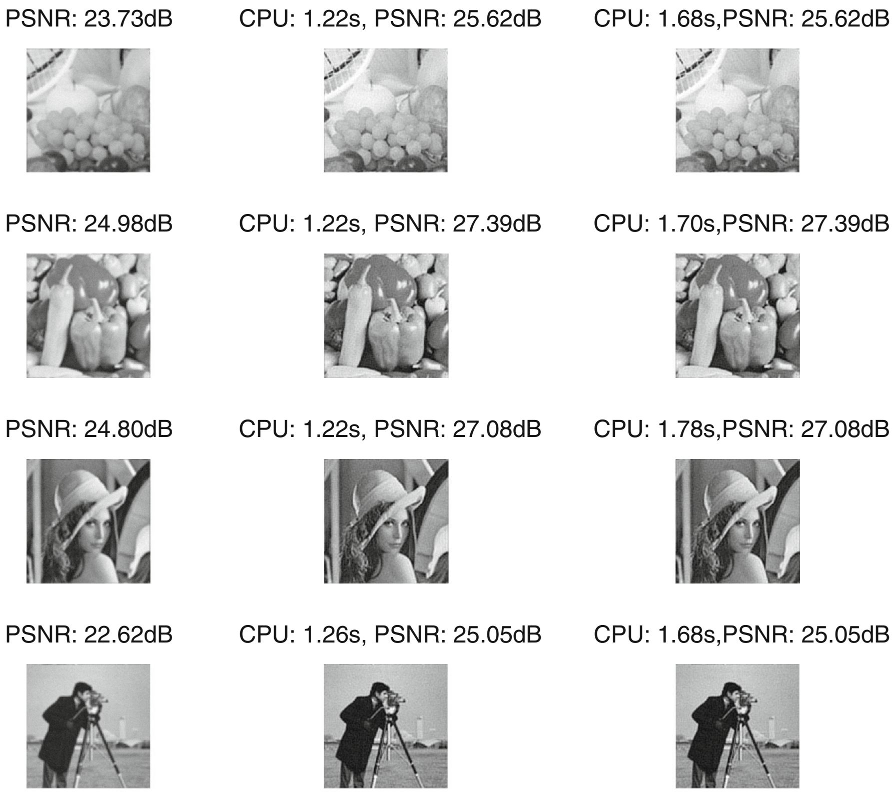

Fig. 2 From *left* to *right*: the blurred images, the restored images by GPRSM, LADMM

Table 1 Comparison of GPRSM with LADM

Image	alpha	η	GPRSM			LADMM		
			Time	Iter	PSNR	Time	Iter	PSNR
Fruits	3	3	0.81	11	28.41	1.00	12	28.41
	5	5	1.26	17	25.62	1.53	20	25.62
Peppers	3	3	0.94	11	30.72	1.25	16	30.72
	5	5	1.36	18	27.39	1.59	21	27.39
Lena	3	3	0.83	11	30.69	1.51	16	30.69
	5	5	1.11	18	27.08	1.97	22	27.08
Cameraman	3	3	0.97	11	28.36	1.22	16	28.36
	5	5	1.47	18	25.05	1.68	22	25.05

can recover the four blurred images very well, and our method is always faster than the LADMM. More numerical results are given in Table 1, and we report the CPU time (in seconds), the number of iterations (Iter) required for the whole deblurring process. The numerical results in Table 1 indicate that both methods reach almost the same

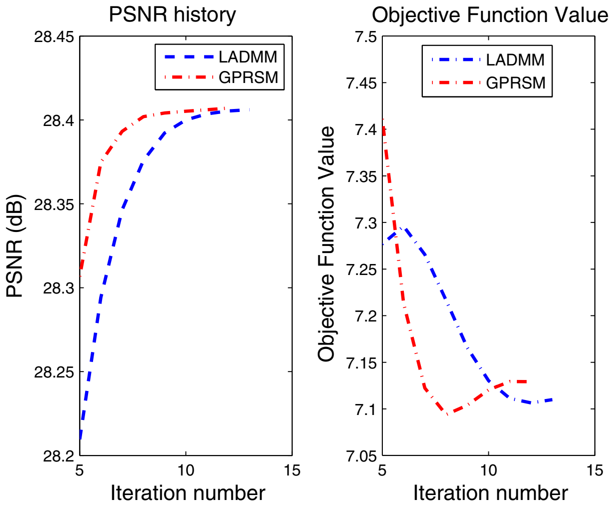

Fig. 3 Comparison results for Fruits

restored PSNR. In addition, GPRSM is always faster than LADMM, and the number of iterations of GPRSM is always smaller than that of LADMM. Thus, GPRSM is more efficient and robustness than the famous LADMM. Furthermore, in Fig. 3, we plot the evolution of the PRSN value and the objective function value with respect to the number of iterations for the Fruits with $\texttt{alpha} = 3$ and η=3. From Fig. 3, we see that, compared with the LADMM, the GPRSM is more efficient.

5 Conclusions

In this paper, we have proposed a generalized PRSM, which removes some restrictive assumptions of the strictly contractive PRSM proposed by He et al., and includes the original PRSM as a special case. Under mild conditions, we have proved the global convergence of the new method. Numerical results about the image deblurring problems indicate that the new method performs better than the linearized ADMM. In the future, we shall investigate the inexact version of our generalized PRSM and apply the generalized PRSM to solve other problems, such as the statistical learning problems, the compressive sensing etc.

Acknowledgments The authors gratefully acknowledge the helpful comments and suggestions of the anonymous reviewers. This work is supported by the National Natural Science Foundation of China (71371139, 11302188), the Shanghai Shuguang Talent Project (13SG24), the Shanghai Pujiang Talent Project (12PJC069), the foundation of Scientific Research Project of Shandong Universities (J15LI11), and the foundation of Zaozhuang University (2014YB03).

References

1. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. Comput. Math. Appl. **2**, 17–40 (1976)

2. Glowinski, R., Marrocco, A.: Approximation par elements finis dordre un, et la resolution, par penalisation-dualite, dune classe de problèmes de Dirichlet non lineaires. RAIRO Anal. Numer. **9**, 41–76 (1975)
3. Eckstein, J., Bertsekas, D.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**, 293–318 (1992)
4. Han, D., Yuan, X.M.: Local linear convergence of the alternating direction method of multipliers for quadratic programs. SIAM J. Numer. Anal. **51**(6), 3446–3457 (2013)
5. Peaceman, D.H., Rachford, H.H.: The numerical solution of parabolic elliptic differential equations. SIAM J. Appl. Math. **3**, 28–41 (1955)
6. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
7. 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)
8. Bertsekas, D.P.: Constrained Optimization and Lagrange Multiplier Methods. Academic Press, New York (1982)
9. Sun, M., Liu, J.: A proximal Peaceman–Rachford splitting method for compressive sensing, J. Appl. Math. Comput., 2015 (Accepted)
10. Chan, R.H., Tao, M., Yuan, X.M.: Linearized alternating direction method for constrained linear least-squares problem. East Asia J. Appl. Math. **2**, 326–341 (2012)