

RESEARCH

Open Access



# A hybrid splitting method for smoothing Tikhonov regularization problem

Yu-Hua Zeng<sup>1,2</sup>, Zheng Peng<sup>3</sup> and Yu-Fei Yang<sup>1\*</sup>

\*Correspondence:

yfyang@hnu.edu.cn

<sup>1</sup>College of Mathematics and  
Econometrics, Hunan University,  
Changsha, 410082, China

Full list of author information is  
available at the end of the article

## Abstract

In this paper, a hybrid splitting method is proposed for solving a smoothing Tikhonov regularization problem. At each iteration, the proposed method solves three subproblems. First of all, two subproblems are solved in a parallel fashion, and the multiplier associated to these two block variables is updated in a rapid sequence. Then the third subproblem is solved in the sense of an alternative fashion with the former two subproblems. Finally, the multiplier associated to the last two block variables is updated. Global convergence of the proposed method is proven under some suitable conditions. Some numerical experiments on the discrete ill-posed problems (DIPPs) show the validity and efficiency of the proposed hybrid splitting method.

**Keywords:** Tikhonov regularization; augmented Lagrangian method; parallel splitting method; alternative direction method of multipliers

## 1 Introduction

In this paper, we consider a smoothing Tikhonov regularization problem, which is an unconstrained minimization of the form [1]

$$\min \|Ax - b\|^2 + \delta \|\Delta x\|^2 + \eta \|x\|_p^p, \quad (1.1)$$

where  $A \in R^{m \times n}$ ,  $b \in R^m$ , and  $x \in \mathcal{X} \subset R^n$ , and  $\|\cdot\|$  denotes the Euclid norm. The parameters  $\delta \geq 0$  and  $\eta \geq 0$  are used to control the smoothness and size of the approximate solution. Matrix  $\Delta$  is a (tridiagonal, Toeplitz) matrix,  $\Delta x$  represents a measure of the variation or smoothness of  $x$ , where

$$\Delta = n^2 \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -2 & 1 \end{pmatrix} \in R^{(n-2) \times n}. \quad (1.2)$$

The smoothing regularization problem (1.1) has numerous applications in many fields, including mathematical programs with vanishing constraints [2], maximum-likelihood estimation problem [3], language modeling [4], and so on.

The last term of the problem (1.1),  $\eta \|x\|_p^p$ , is a regularization term. As a common regularization method,  $\ell_1$  regularization ( $p = 1$ ) problem has many good properties since it is a convex programming problem. In recent years, there has been an increasing interest in the  $\ell_1$  regularizer. The  $\ell_1$  regularization model can reconstruct the original signal with less observed signals, when the original signal is sparse or the observed signal contains noise. Especially, the  $\ell_1$  formulation suits significantly better for denoising data containing so-called outliers, *i.e.*, observations containing large measurement errors [5]. Therefore, mathematical models and large-scale fast algorithms associated with  $\ell_1$  regularization can be seen everywhere in compressed sensing, signal/image processing, machine learning, statistics, and many other fields [6–9].

By  $\ell_1$  regularization, the problem (1.1) reduces to

$$\min \|Ax - b\|_2^2 + \delta \|\Delta x\|_2^2 + \eta \|x\|_1. \quad (1.3)$$

It is identical to a separable convex minimization of the form

$$\begin{cases} \min \frac{1}{2} \|Ax - b\|^2 + \frac{1}{2} \delta \|\Delta y\|^2 + \eta \|z\|_1, \\ \text{s.t. } x - y = 0, \\ y - z = 0. \end{cases} \quad (1.4)$$

The augmented Lagrangian function associated to the problem (1.4) is

$$\begin{aligned} L(x, y, \lambda, z, \mu) = & \frac{1}{2} \|Ax - b\|^2 + \frac{1}{2} \delta \|\Delta y\|^2 - \lambda^T (x - y) + \frac{\beta}{2} \|x - y\|^2 \\ & - \mu^T (y - z) + \eta \|z\|_1 + \frac{\rho}{2} \|y - z\|^2. \end{aligned} \quad (1.5)$$

Indeed, there are many methods for solving the problem (1.4) in the literature. Among these methods, the parallel splitting augmented Lagrangian method and the alternating direction method of multipliers are two power tools. The recent research indicates that, due to the separable convex optimization with three block variables, the direct extension of the alternating direction method of multipliers is not necessarily convergent [10]. Thus, some hybrid splitting methods can be found in the literature. For example, see He [11], Peng and Wu [12], and Glowinski and Le Tallec [13].

The saddle point of a Lagrange function associated to the convex optimization problem (1.4),  $w^* = (x^*, y^*, \lambda^*, z^*, \mu^*) \in \mathcal{W}$ , satisfies the following variational inequalities:

$$\begin{cases} (x' - x^*)^T [A^T (Ax^* - b) - \lambda^*] \geq 0, \\ (y' - y^*)^T [\delta \Delta^T \Delta y^* + \lambda^* - \mu^*] \geq 0, \\ (\lambda' - \lambda^*)^T (x^* - y^*) \geq 0, \\ \eta \|z'\|_1 - \eta \|z^*\|_1 + (z' - z^*)^T [\mu^*] \geq 0, \\ (\mu' - \mu^*)^T (y^* - z^*) \geq 0, \end{cases} \quad \forall w' = (x', y', \lambda', z', \mu') \in \mathcal{W}, \quad (1.6)$$

where

$$\mathcal{W} = \mathcal{X} \times \mathcal{X} \times R^n \times \mathcal{X} \times R^n \subset R^{5 \times n}.$$

In this paper, we will propose a new splitting method to solve the separable convex programming problem (1.4) as well as the structured variational inequalities (1.6). The pro-

posed method, referred to as the hybrid splitting method (HSM), will combine a parallel splitting (augmented Lagrangian) method and an alternating directions method of multipliers. In the HSM, the predictor of the new iterate,  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$ , is got in the following way: find  $(\tilde{x}^k, \tilde{y}^k)$  in a parallel manner, and update  $\tilde{\lambda}^k$ . Then compute  $\tilde{z}^k$  alternately with  $(\tilde{x}^k, \tilde{y}^k)$ , and update  $\tilde{\mu}^k$  at the last. The new iterate is produced by a correct operator. The global convergence of the HSM is proven under some wild assumptions.

The rest of this paper is organized as follows. In Section 2, we describe the proposed hybrid splitting method. Section 3 is devoted to showing that the sequence  $\{w^k\}$  generated by the HSM is Fejér monotone with respect to the solution set. Then, the convergence of the HSM is proved. In Section 4, some preliminary numerical results are presented which indicate the feasibility and efficiency of the proposed method. Finally, some concluding remarks are made in Section 5.

## 2 The hybrid splitting method

In this section, we first propose a hybrid splitting method for the problem (1.6), and then we give some remarks on the described method.

**Algorithm 2.1** (The hybrid splitting method (HSM)) For a given  $w^k = (x^k, y^k, \lambda^k, z^k, u^k) \in \mathcal{W}$ ,  $\beta_k > 0$ , and  $\rho_k > 0$ , the HSM produces the new iterate  $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1}, z^{k+1}, u^{k+1}) \in \mathcal{W}$  by the following scheme:

S1. Produce  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  by s1.1 to s1.5.

s1.1. Find  $\tilde{x}^k \in \mathcal{X}$  (with fixed  $y^k, \lambda^k, z^k, \mu^k, \beta_k, \rho_k$ ) such that

$$(x' - \tilde{x}^k) \{A^T(A\tilde{x}^k - b) - \lambda^k + \beta_k(\tilde{x}^k - y^k)\} \geq 0, \quad \forall x' \in \mathcal{X}. \quad (2.1)$$

s1.2. Find  $\tilde{y}^k \in \mathcal{X}$  (with fixed  $x^k, \lambda^k, z^k, \mu^k, \beta_k, \rho_k$ ) such that

$$(y' - \tilde{y}^k) \{\delta \Delta^T \Delta \tilde{y}^k + \lambda^k - \mu^k - \beta_k(x^k - \tilde{y}^k) + \rho_k(\tilde{y}^k - z^k)\} \geq 0, \quad \forall y' \in \mathcal{X}. \quad (2.2)$$

s1.3. Update  $\tilde{\lambda}^k$  via

$$\tilde{\lambda}^k = \lambda^k - \beta_k(\tilde{x}^k - \tilde{y}^k). \quad (2.3)$$

s1.4. Find  $\tilde{z}^k \in \mathcal{X}$  (with fixed  $\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \mu^k, \beta_k, \rho_k$ ) such that

$$\eta \|z'\|_1 - \eta \|\tilde{z}^k\|_1 + (z' - \tilde{z}^k) \{\mu^k - \rho_k(\tilde{y}^k - \tilde{z}^k)\} \geq 0, \quad \forall z' \in \mathcal{X}. \quad (2.4)$$

s1.5. Update  $\tilde{\mu}^k$  via

$$\tilde{\mu}^k = \mu^k - \rho_k(\tilde{y}^k - \tilde{z}^k). \quad (2.5)$$

S2. Convergence verification: for a given small  $\varepsilon > 0$ , if  $\|w^k - \tilde{w}^k\|_\infty < \varepsilon$  then stop, and accept  $w^k$  to be the approximate solution. Else, go to S3.

S3. Produce the new iterate by

$$w^{k+1} = w^k - \alpha_k d(w^k, \tilde{w}^k), \quad (2.6)$$

where

$$\alpha_k = \gamma \alpha_k^*, \quad \gamma \in (0, 2), \quad (2.7)$$

$$\alpha_k^* = \frac{\varphi(w^k, \tilde{w}^k)}{\|d(w^k, \tilde{w}^k)\|_G^2}, \quad (2.8)$$

and

$$\varphi(w^k, \tilde{w}^k) = \|w^k - \tilde{w}^k\|_G^2, \quad d(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k), \quad (2.9)$$

$$G = \begin{pmatrix} \frac{\beta_k}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{2\beta_k + \rho_k}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\beta_k} & 0 & 0 \\ 0 & 0 & 0 & \frac{\rho_k}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\rho_k} \end{pmatrix}, \quad M = \begin{pmatrix} \beta_k & 0 & 0 & 0 & 0 \\ 0 & \beta_k + \rho_k & -\frac{\rho_k}{2} & 0 & 0 \\ 0 & -\frac{\rho_k}{2} & \frac{2}{\beta_k} & 0 & 0 \\ 0 & 0 & 0 & \rho_k & 0 \\ 0 & 0 & 0 & 0 & \frac{3}{\rho_k} \end{pmatrix}.$$

**Remark 2.1** The parameters  $\beta_k$  and  $\rho_k$  are updated in the same style as proposed in He *et al.* [14]. By Remark 5.1 in [14], the sequences  $\{\beta_k\}$  and  $\{\rho_k\}$  are bounded and finally constants. Thus, there is a  $\kappa > 0$  such that  $\|M\|_G^2 := \sum \|M(:, j)\|_G^2 \leq \kappa$ .

**Remark 2.2** It is easy to show,  $\|M(w - \tilde{w})\|_G^2 \leq \|M\|_G^2 \cdot \|w - \tilde{w}\|_G^2$ . Thus by (2.8) we have

$$\alpha_k^* \geq \frac{1}{\|M\|_G^2} \geq \frac{1}{\kappa}, \quad \forall k. \quad (2.10)$$

For convenience in the analysis, the following notations are useful:

$$F(w) = \begin{pmatrix} A^T(Ax - b) - \lambda \\ \sigma \Delta^T \Delta y + \lambda - \mu \\ x - y \\ \mu \\ y - z \end{pmatrix}, \quad H = \begin{pmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & 2I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 3I \end{pmatrix},$$

$$g(w, \tilde{w}) = \begin{pmatrix} \beta(x - \tilde{x}) - \beta(y - \tilde{y}) \\ -\beta(x - \tilde{x}) + \beta(y - \tilde{y}) + \rho(y - \tilde{y}) - \rho(z - \tilde{z}) \\ 0 \\ \rho(z - \tilde{z}) - \rho(y - \tilde{y}) \\ 0 \end{pmatrix}.$$

By these notations, the variational inequalities (1.6) can be rewritten in a compact form: find  $w^* \in \mathcal{W}$  such that

$$(w' - w^*)^T F(w^*) \geq 0, \quad \forall w' \in \mathcal{W}. \quad (2.11)$$

In the HSM, (2.1)-(2.5) can be written in the compact form: find  $\tilde{w}^k \in \mathcal{W}$  such that

$$\eta \|z'\|_1 - \eta \|\tilde{z}^k\|_1 + (w' - \tilde{w}^k)^T [F(\tilde{w}^k) + g(w^k, \tilde{w}^k) - M(w^k - \tilde{w}^k)] \geq 0, \quad \forall w' \in \mathcal{W}. \quad (2.12)$$

### 3 The convergence

To prove the convergence of the HSM, we will show first in this section that the sequence  $\{w^k\}$  generated by the HSM is Fejér monotone with respect to the solution set  $\mathcal{W}^*$  of the problem (2.12).

Due to (1.6), multiplying both sides of the third inequality by 2, and multiplying both sides of the last inequality by 3, respectively, we get

$$\begin{cases} (x' - x^*)^T [A^T (Ax^* - b) - \lambda^*] \geq 0, \\ (y' - y^*)^T [\delta \Delta^T \Delta y^* + \lambda^* - \mu^*] \geq 0, \\ 2 \times (\lambda' - \lambda^*)^T (x^* - y^*) \geq 0, \\ \eta \|z'\|_1 - \eta \|z^*\|_1 + (z' - z^*)^T [\mu^*] \geq 0, \\ 3 \times (\mu' - \mu^*)^T (y^* - z^*) \geq 0, \end{cases} \quad \forall w' = (x', y', \lambda', z', \mu') \in \mathcal{W}, \quad (3.1)$$

and (3.1) can be written as

$$(w' - w^*)^T HF(w^*) \geq 0, \quad \forall w' \in \mathcal{W}. \quad (3.2)$$

**Lemma 3.1** For a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , if  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5), then we have

$$(\tilde{w}^k - w^*)^T F(\tilde{w}^k) \geq 0 \quad (3.3)$$

and

$$(\tilde{w}^k - w^*)^T HF(\tilde{w}^k) \geq 0, \quad (3.4)$$

where  $w^* = (x^*, y^*, \lambda^*, z^*, \mu^*) \in \mathcal{W}^*$  is a solution.

*Proof* It is easy to show that  $F(w)$  is linear and consequently monotone. Indeed, let

$$Q = \begin{pmatrix} A^T A & 0 & -I & 0 & 0 \\ 0 & \delta \Delta^T \Delta & I & 0 & -I \\ I & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I \\ 0 & I & 0 & -I & 0 \end{pmatrix} \quad \text{and} \quad e = \begin{pmatrix} -A^T b \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

then  $F(w) = Qw + e$ . By the monotonicity of  $F$  and  $HF$  we have

$$(\tilde{w}^k - w^*)^T (F(\tilde{w}^k) - F(w^*)) \geq 0$$

and

$$(\tilde{w}^k - w^*)^T (HF(\tilde{w}^k) - HF(w^*)) \geq 0,$$

respectively, which results in (by (2.11))

$$(\tilde{w}^k - w^*)^T F(\tilde{w}^k) \geq (\tilde{w}^k - w^*)^T F(w^*) \geq 0,$$

and by (3.2)

$$(\tilde{w}^k - w^*)^T HF(\tilde{w}^k) \geq (\tilde{w}^k - w^*)^T HF(w^*) \geq 0. \quad \square$$

**Lemma 3.2** For a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , if  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5), then we have

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] \\ & \geq (\lambda^k - \tilde{\lambda}^k)^T (x^k - \tilde{x}^k) - (\lambda^k - \tilde{\lambda}^k)^T (y^k - \tilde{y}^k) \\ & \quad + (\mu^k - \tilde{\mu}^k)^T (y^k - \tilde{y}^k) - (\mu^k - \tilde{\mu}^k)^T (z^k - \tilde{z}^k). \end{aligned} \quad (3.5)$$

*Proof* By a direct computation, we get

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] \\ & = (\tilde{w}^k - w^*)^T HF(\tilde{w}^k) + (\tilde{w}^k - w^*)^T g(w^k, \tilde{w}^k) \\ & \geq (\tilde{w}^k - w^*)^T g(w^k, \tilde{w}^k) \\ & = (\tilde{x}^k - x^*)^T \beta_k (x^k - \tilde{x}^k) - (\tilde{x}^k - x^*)^T \beta_k (y^k - \tilde{y}^k) - (\tilde{y}^k - y^*) \beta_k (x^k - \tilde{x}^k) \\ & \quad + (\tilde{y}^k - y^*)^T \beta_k (y^k - \tilde{y}^k) + (\tilde{y}^k - y^*) \rho_k (y^k - \tilde{y}^k) - (\tilde{y}^k - y^*) \rho_k (z^k - \tilde{z}^k) \\ & \quad + (\tilde{z}^k - z^*) \rho_k (z^k - \tilde{z}^k) - (\tilde{z}^k - z^*) \rho_k (y^k - \tilde{y}^k) \\ & = [(\tilde{x}^k - \tilde{y}^k) - (x^* - y^*)]^T \beta_k (x^k - \tilde{x}^k) - [(\tilde{x}^k - \tilde{y}^k) - (x^* - y^*)]^T \beta_k (y^k - \tilde{y}^k) \\ & \quad + [(\tilde{y}^k - \tilde{z}^k) - (y^* - z^*)]^T \rho_k (y^k - \tilde{y}^k) - [(\tilde{y}^k - \tilde{z}^k) - (y^* - z^*)]^T \rho_k (z^k - \tilde{z}^k). \end{aligned}$$

Substituting  $x^* - y^* = 0$ ,  $y^* - z^* = 0$  into the last equation, we get

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] \\ & \geq (\tilde{x}^k - \tilde{y}^k) \beta_k (x^k - \tilde{x}^k) - (\tilde{x}^k - \tilde{y}^k) \beta_k (y^k - \tilde{y}^k) \\ & \quad + (\tilde{y}^k - \tilde{z}^k) \rho_k (y^k - \tilde{y}^k) - (\tilde{y}^k - \tilde{z}^k) \rho_k (z^k - \tilde{z}^k). \end{aligned} \quad (3.6)$$

By (2.3) and (2.5), we get

$$(\tilde{x}^k - \tilde{y}^k) = \frac{1}{\beta_k} (\lambda^k - \tilde{\lambda}^k), \quad (3.7)$$

$$(\tilde{y}^k - \tilde{z}^k) = \frac{1}{\rho_k} (\mu^k - \tilde{\mu}^k). \quad (3.8)$$

Substituting (3.7) and (3.8) into the right-hand side of (3.6), we obtain

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] \\ & \geq (\lambda^k - \tilde{\lambda}^k) (x^k - \tilde{x}^k) - (\lambda^k - \tilde{\lambda}^k) (y^k - \tilde{y}^k) \\ & \quad + (\mu^k - \tilde{\mu}^k) (y^k - \tilde{y}^k) - (\mu^k - \tilde{\mu}^k) (z^k - \tilde{z}^k). \end{aligned} \quad \square$$

**Lemma 3.3** For a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , if  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5), then we have

$$\begin{aligned} & \eta \|z'\|_1 - \eta \|\tilde{z}^k\|_1 + (w' - \tilde{w}^k)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k) - d(w^k, \tilde{w}^k)] \\ & \geq 0, \quad \forall z' \in \mathcal{X}, w' \in \mathcal{W}. \end{aligned} \quad (3.9)$$

*Proof* Combining (2.1)-(2.5) together, we have

$$\begin{aligned} & \eta \|z'\|_1 - \eta \|\tilde{z}^k\|_1 + (w' - \tilde{w}^k)^T \begin{pmatrix} A^T(A\tilde{x}^k - b) - \lambda^k + \beta_k(\tilde{x}^k - y^k) \\ \delta \Delta^T \Delta \tilde{y}^k + \lambda^k - \mu^k - \beta_k(x^k - \tilde{y}^k) + \rho_k(\tilde{y}^k - z^k) \\ 2(\tilde{x}^k - \tilde{y}^k) - \frac{1}{\beta_k}(\lambda^k - \tilde{\lambda}^k) \\ \mu^k - \rho_k(\tilde{y}^k - \tilde{z}^k) \\ 3(\tilde{y}^k - \tilde{z}^k) - \frac{1}{\rho_k}(\mu^k - \tilde{\mu}^k) \end{pmatrix} \\ & \geq 0. \end{aligned}$$

By a manipulation, we get

$$\begin{aligned} & \eta \|z'\|_1 - \eta \|\tilde{z}^k\|_1 + (w' - \tilde{w}^k)^T \left\{ \begin{pmatrix} A^T(A\tilde{x}^k - b) - \tilde{\lambda}^k \\ \delta \Delta^T \Delta \tilde{y}^k + \tilde{\lambda}^k - \tilde{\mu}^k \\ 2(\tilde{x}^k - \tilde{y}^k) \\ \tilde{\mu}^k \\ 3(\tilde{y}^k - \tilde{z}^k) \end{pmatrix} + g(w^k, \tilde{w}^k) - M(w^k - \tilde{w}^k) \right\} \\ & \geq 0. \end{aligned} \quad (3.10)$$

The assertion (3.9) is just a compact form of (3.10).  $\square$

**Theorem 3.1** For a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , if  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5), then for any  $w^* = (x^*, y^*, \lambda^*, z^*, \mu^*) \in \mathcal{W}^*$  we have

$$(\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] \geq \varphi(w^k, \tilde{w}^k) - (w^k - \tilde{w}^k)^T d(w^k, \tilde{w}^k). \quad (3.11)$$

*Proof* Note  $d(w^k, \tilde{w}^k) = M(w^k - \tilde{w}^k)$ , by (3.5) we get

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] + (w^k - \tilde{w}^k)^T d(w^k, \tilde{w}^k) \\ & \geq (\lambda^k - \tilde{\lambda}^k)^T (x^k - \tilde{x}^k) - (\lambda^k - \tilde{\lambda}^k)^T (y^k - \tilde{y}^k) \\ & \quad + (\mu^k - \tilde{\mu}^k)^T (y^k - \tilde{y}^k) - (\mu^k - \tilde{\mu}^k)^T (z^k - \tilde{z}^k) \\ & \quad + \beta_k \|x^k - \tilde{x}^k\| + \beta_k \|y^k - \tilde{y}^k\| + \rho_k \|y^k - \tilde{y}^k\| \\ & \quad - \frac{1}{2} \rho_k (y^k - \tilde{y}^k)^T (z^k - \tilde{z}^k) + \frac{2}{\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 \\ & \quad - \frac{\rho_k}{2} (z^k - \tilde{z}^k)^T (y^k - \tilde{y}^k) + \rho_k \|z^k - \tilde{z}^k\|^2 + \frac{3}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \\ & \geq (\lambda^k - \tilde{\lambda}^k)^T (x^k - \tilde{x}^k) - (\lambda^k - \tilde{\lambda}^k)^T (y^k - \tilde{y}^k) \end{aligned}$$

$$\begin{aligned}
& + (\mu^k - \tilde{\mu}^k)^T (y^k - \tilde{y}^k) - (\mu^k - \tilde{\mu}^k)^T (z^k - \tilde{z}^k) \\
& - \rho_k (y^k - \tilde{y}^k)^T (z^k - \tilde{z}^k) + \beta_k \|x^k - \tilde{x}^k\|^2 \\
& + \beta_k \|y^k - \tilde{y}^k\|^2 + \rho_k \|y^k - \tilde{y}^k\|^2 + \rho_k \|z^k - \tilde{z}^k\|^2 \\
& + \frac{2}{\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 + \frac{3}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \\
& = \left[ \frac{1}{2} \beta_k \|x^k - \tilde{x}^k\|^2 + (\lambda^k - \tilde{\lambda}^k)^T (x^k - \tilde{x}^k) + \frac{1}{2\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 \right] \\
& + \left[ \frac{1}{2} \beta_k \|y^k - \tilde{y}^k\|^2 + (\lambda^k - \tilde{\lambda}^k)^T (y^k - \tilde{y}^k) + \frac{1}{2\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 \right] \\
& + \left[ \frac{1}{4} \rho_k \|y^k - \tilde{y}^k\|^2 + (\mu^k - \tilde{\mu}^k)^T (y^k - \tilde{y}^k) + \frac{1}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \right] \\
& + \left[ \frac{1}{4} \rho_k \|z^k - \tilde{z}^k\|^2 + (\mu^k - \tilde{\mu}^k)^T (z^k - \tilde{z}^k) + \frac{1}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \right] \\
& + \left[ \frac{1}{2} \rho_k \|y^k - \tilde{y}^k\|^2 - \rho_k (y^k - \tilde{y})^T (z^k - \tilde{z}^k) + \frac{\rho_k}{2} \|z^k - \tilde{z}^k\|^2 \right] \\
& + \frac{1}{2} \beta_k \|x^k - \tilde{x}^k\|^2 + \frac{1}{2} \beta_k \|y^k - \tilde{y}^k\|^2 + \frac{1}{4} \rho_k \|y^k - \tilde{y}^k\|^2 \\
& + \frac{1}{\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 + \frac{1}{4} \rho_k \|z^k - \tilde{z}^k\|^2 + \frac{1}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \\
& = \frac{1}{2} \left[ \sqrt{\beta_k} \|x^k - \tilde{x}^k\| + \frac{1}{\sqrt{\beta_k}} \|\lambda^k - \tilde{\lambda}^k\| \right]^2 + \frac{1}{2} \left[ \sqrt{\beta_k} \|y^k - \tilde{y}^k\| - \frac{1}{\sqrt{\beta_k}} \|\lambda^k - \tilde{\lambda}^k\| \right]^2 \\
& + \left[ \frac{\sqrt{\rho_k}}{2} \|y^k - \tilde{y}^k\| + \frac{1}{\sqrt{\rho_k}} \|\mu^k - \tilde{\mu}^k\| \right]^2 + \left[ \frac{\sqrt{\rho_k}}{2} \|z^k - \tilde{z}^k\| - \frac{1}{\sqrt{\rho_k}} \|\mu^k - \tilde{\mu}^k\| \right]^2 \\
& + \frac{1}{2} \left[ \sqrt{\rho_k} \|y^k - \tilde{y}^k\| - \sqrt{\rho_k} \|z^k - \tilde{z}^k\| \right]^2 \\
& + \frac{1}{2} \beta_k \|x^k - \tilde{x}^k\|^2 + \frac{1}{2} \beta_k \|y^k - \tilde{y}^k\|^2 + \frac{1}{4} \rho_k \|y^k - \tilde{y}^k\|^2 \\
& + \frac{1}{\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 + \frac{1}{4} \rho_k \|z^k - \tilde{z}^k\|^2 + \frac{1}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \\
& \geq \frac{1}{2} \beta_k \|x^k - \tilde{x}^k\|^2 + \frac{1}{2} \beta_k \|y^k - \tilde{y}^k\|^2 + \frac{1}{4} \rho_k \|y^k - \tilde{y}^k\|^2 + \frac{1}{\beta_k} \|\lambda^k - \tilde{\lambda}^k\|^2 \\
& + \frac{1}{4} \rho_k \|z^k - \tilde{z}^k\|^2 + \frac{1}{\rho_k} \|\mu^k - \tilde{\mu}^k\|^2 \\
& = (w^k - \tilde{w}^k)^T G (w^k - \tilde{w}^k),
\end{aligned}$$

where

$$G = \begin{pmatrix} \frac{1}{2} \beta_k & & & & \\ & \frac{2\beta_k + \rho_k}{4} & & & \\ & & \frac{1}{\beta_k} & & \\ & & & \frac{1}{4} \rho_k & \\ & & & & \frac{1}{\rho_k} \end{pmatrix}.$$



Thus, we have

$$\begin{aligned} & (\tilde{w}^k - w^*)^T [HF(\tilde{w}^k) + g(w^k, \tilde{w}^k)] + (w^k - \tilde{w}^k)^T d(w^k, \tilde{w}^k) \\ & \geq (w^k - \tilde{w}^k)^T G(w^k - \tilde{w}^k) = \|w^k - \tilde{w}^k\|_G^2 := \varphi(w^k, \tilde{w}^k). \end{aligned} \quad (3.12)$$

This theorem follows from (3.12) directly.  $\square$

**Lemma 3.4** *If  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5) from a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , then for any  $w^* = (x^*, y^*, \lambda^*, z^*, \mu^*) \in \mathcal{W}^*$ , we have*

$$(w^k - w^*)^T Gd(w^k, \tilde{w}^k) \geq \varphi(w^k, \tilde{w}^k). \quad (3.13)$$

*Proof* We omit the proof of Lemma 3.4 here. A similar proof can be found in [15].  $\square$

**Theorem 3.2** *For a given  $w^k = (x^k, y^k, \lambda^k, z^k, \mu^k)$ , if  $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k, \tilde{z}^k, \tilde{\mu}^k)$  is generated by (2.1)-(2.5), then for any  $w^* = (x^*, y^*, \lambda^*, z^*, \mu^*) \in \mathcal{W}^*$  we have*

$$\|w^{k+1} - w^*\|_G^2 \leq \|w^k - w^*\|_G^2 - \gamma(2 - \gamma) \frac{1}{\kappa} \|w^k - \tilde{w}^k\|_G^2, \quad (3.14)$$

where  $0 < \gamma < 2, \kappa > 0$ .

*Proof* By the iterative formula (2.6), and Lemma 3.4, we have

$$\begin{aligned} \|w^{k+1} - w^*\|_G^2 &= \|w^k - w^* - \alpha_k d(w^k, \tilde{w}^k)\|_G^2 \\ &= \|w^k - w^*\|_G^2 - 2\alpha_k \langle w^k - w^*, Gd(w^k, \tilde{w}^k) \rangle + \alpha_k^2 \|d(w^k, \tilde{w}^k)\|_G^2 \\ &\leq \|w^k - w^*\|_G^2 - 2\alpha_k \varphi(w^k, \tilde{w}^k) + \alpha_k^2 \|d(w^k, \tilde{w}^k)\|_G^2. \end{aligned}$$

Following from (2.7) and (2.8), we have

$$\begin{aligned} \|w^{k+1} - w^*\|_G^2 &\leq \|w^k - w^*\|_G^2 - 2\gamma\alpha_k^{*2} \|d(w^k, \tilde{w}^k)\|_G^2 + \gamma^2\alpha_k^{*2} \|d(w^k, \tilde{w}^k)\|_G^2 \\ &= \|w^k - w^*\|_G^2 - \gamma(2 - \gamma)\alpha_k^{*2} \|d(w^k, \tilde{w}^k)\|_G^2 \\ &= \|w^k - w^*\|_G^2 - \gamma(2 - \gamma)\alpha_k^* \|w^k - \tilde{w}^k\|_G^2 \\ &\leq \|w^k - w^*\|_G^2 - \gamma(2 - \gamma) \frac{1}{\kappa} \|w^k - \tilde{w}^k\|_G^2. \end{aligned}$$

The last inequality follows from (2.10).  $\square$

Theorem 3.2 claims the Fejér monotonicity of the sequence  $\{w^k\}$  generated by the HSM. Adding (3.14) from 0 to  $\infty$  with respect to  $k$  yields

$$\gamma(2 - \gamma) \frac{1}{\kappa} \sum_{k=0}^{\infty} \|w^k - \tilde{w}^k\|_G^2 \leq \|w^0 - w^*\|_G^2 - \lim_{k \rightarrow \infty} \|w^{k+1} - w^*\|_G^2 \leq \|w^0 - w^*\|_G^2 < \infty.$$

Thus we have

$$\lim_{k \rightarrow \infty} \|w^k - \tilde{w}^k\|_G^2 = 0, \quad (3.15)$$

which results in both  $\{w^k\}$  and  $\{\tilde{w}^k\}$  being bounded sequences and having cluster points. Let  $w^\infty$  be a cluster point of  $\{\tilde{w}^k\}$  and  $\{\tilde{w}^{k_j}\}$  be a subsequence converging to  $w^\infty$ .

By the HSM,  $\tilde{w}^{k_j}$  is a solution of (2.12), thus

$$\begin{aligned} & \eta \|z'\|_1 - \eta \|z^{k_j}\|_1 + (w' - \tilde{w}^{k_j})^T [F(\tilde{w}^{k_j}) + g(w^{k_j}, \tilde{w}^{k_j}) - M(w^{k_j} - \tilde{w}^{k_j})] \\ & \geq 0, \quad \forall w' \in \mathcal{W}. \end{aligned} \quad (3.16)$$

The limit (3.15) also results in  $\lim_{k_j \rightarrow \infty} g(w^{k_j}, \tilde{w}^{k_j}) = 0$  and  $\lim_{k_j \rightarrow \infty} d(w^{k_j}, \tilde{w}^{k_j}) = M(w^{k_j} - \tilde{w}^{k_j}) = 0$  by positive-definiteness of  $G$  and  $M$ . Taking the limit on both sides of (3.16) we have

$$\eta \|z'\|_1 - \eta \|z^\infty\|_1 + (w' - w^\infty)^T F(w^\infty) \geq 0, \quad \forall w' \in \mathcal{W}, \quad (3.17)$$

which implies  $w^\infty$  is a solution of the problem (1.4) by the optimality condition (2.11).

Note that (3.14) holds for all solutions of (1.4), we get

$$\|w^{k+1} - w^\infty\|_G^2 \leq \|w^k - w^\infty\|_G^2, \quad \forall k. \quad (3.18)$$

Since  $\tilde{w}^{k_j} \rightarrow w^\infty$  and  $\|w^{k_j} - \tilde{w}^{k_j}\|_G^2 \rightarrow 0$  as  $k_j \rightarrow \infty$ , for  $\forall \varepsilon > 0$ , there exists an integer  $k_l > 0$  such that for all  $k$  and  $k_j$  larger than  $k_l$ , we have

$$\|\tilde{w}^{k_j} - w^\infty\|_G < \frac{\varepsilon}{2}, \quad \|w^k - \tilde{w}^{k_j}\|_G < \frac{\varepsilon}{2}. \quad (3.19)$$

It follows from (3.18) and (3.19) that

$$\|w^k - w^\infty\|_G = \|w^k - \tilde{w}^{k_j} + \tilde{w}^{k_j} - w^\infty\|_G \leq \|w^k - \tilde{w}^{k_j}\|_G + \|\tilde{w}^{k_j} - w^\infty\|_G < \varepsilon.$$

Thus, the sequence  $\{w^k\}$  converges to  $w^\infty$ , which is a solution of (1.4), or equivalently, of the problem (1.1).

#### 4 Numerical results

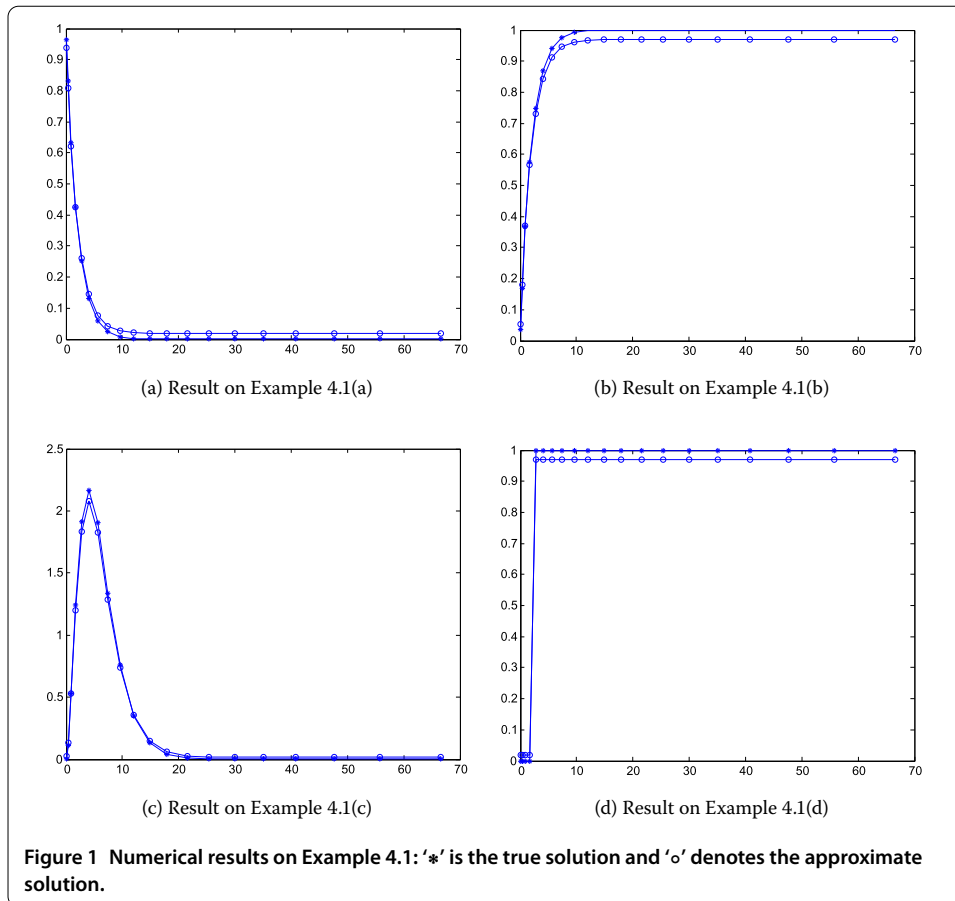
The main computational cost of the HSM is in solving the subproblems (2.1)-(2.2) and (2.4). For generality of the HSM, we ignore the special structure of the matrices  $A$  and  $\Delta$ , and employ the existing efficient method to solve those subproblems. In this paper, the subproblems (2.1) and (2.2) are solved by the projected Barzilai-Borwein method proposed by Dai and Fletcher [16], and the subproblem (2.4) is solved by a shrinkage-thresholding algorithm [17]. The test problems are discrete ill-posed problems selected from Hansen [18]. All tests are done on a laptop with Core(TM) i7 M620@D, 2.67 GHz, 4 GB RAM and Matlab R2009a.

A classical example of an ill-posed problem is a Fredholm integral equation of the first kind with a square integrable kernel, which has the form

$$\int_a^b K(s, t) f(t) dt = g(s), \quad s \in [c, d], \quad (4.1)$$

where the right-hand side  $g$  and the kernel  $K$  are given, and  $f$  is unknown.

The following examples are used to test our algorithm.



**Example 4.1** The kernel  $K$  is given by

$$K(s, t) = e^{-st},$$

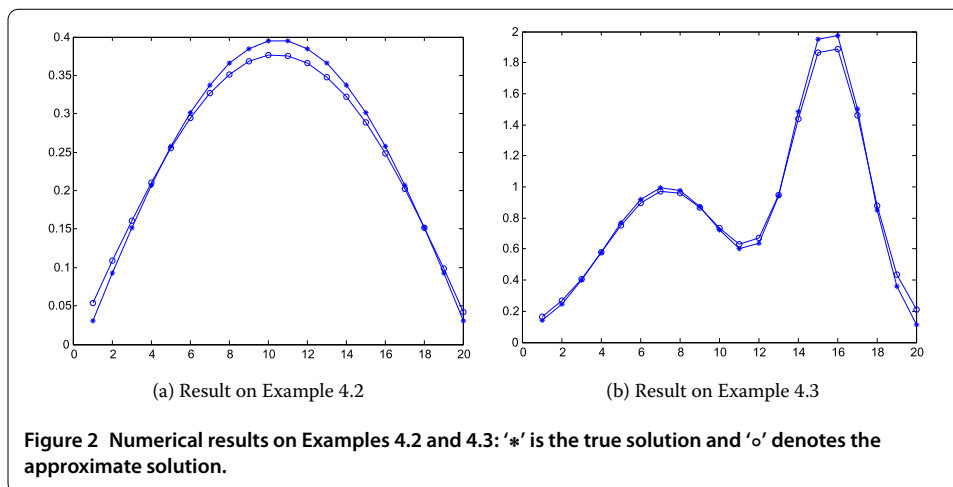
and the integration interval is  $[0, \infty)$ . The true solution  $f$  and the right-hand side  $g$  are given by

$$\begin{aligned} \text{(a)} \quad & f(t) = e^{-\frac{t}{2}}, \quad g(s) = \frac{1}{s+0.5}, \\ \text{(b)} \quad & f(t) = 1 - e^{-\frac{t}{2}}, \quad g(s) = \frac{1}{s} - \frac{1}{s+0.5}, \\ \text{(c)} \quad & f(t) = t^2 \times e^{-\frac{t}{2}}, \quad g(s) = \frac{2}{(s+0.5)^3}, \\ \text{(d)} \quad & f(t) = \begin{cases} 0, & t \leq 2, \\ 1, & t > 2, \end{cases} \quad g(s) = \frac{e^{-2s}}{s}. \end{aligned}$$

The numerical results are displayed in Figure 1.

**Example 4.2** Shaw test problem: one-dimensional image restoration model. We have the kernel  $K$  and the solution  $f$ , which are given by

$$K(s, t) = (\cos(s) + \cos(t)) \times (\sin(u)/u)^2$$



and

$$f(t) = a_1 \times \exp(-c_1 \times (t - t_1)^2) + a_2 \times \exp(-c_2 \times (t - t_2)^2),$$

where  $u = \pi \times (\sin(s) + \sin(t))$ . Both  $K$  and  $f$  are discretized by simple quadrature to produce  $A$  and  $x$ . Then the discrete right-hand side is produced as  $b = Ax$ . In our test, the constants are assigned as follows:  $a_1 = 2$ ,  $c_1 = 6$ ,  $t_1 = 0.8$ ;  $a_2 = 1$ ,  $c_2 = 2$ ,  $t_2 = -0.5$ . The numerical result is displayed in Figure 2(a).

**Example 4.3** Baart test problem: the kernel  $K$  and right-hand side  $g$  of the discretization of a first-kind Fredholm integral equation are given by

$$K(s, t) = \exp(s \times \cos(t)), \quad g(s) = 2 \times \frac{\sinh(s)}{s},$$

where  $s \in [0, \pi/2]$ ,  $t \in [0, \pi]$ . The solution is given by  $f(t) = \sin(t)$ . The numerical result is displayed in Figure 2(b).

We can draw the conclusion from the above computational results: the proposed hybrid splitting method is valid and efficient for the smoothing Tikhonov regularization problem.

## 5 Conclusions

We have proposed a hybrid splitting method for solving the smoothing Tikhonov regularization problem, which is to minimize the sum of three convex functions over a simple closed convex set. The problem can be reformulated as a convex minimization problem with three separable block variables, and we get the variational inequalities formulation by the KKT conditions of the separable convex minimization problem. By the convexity of the problem, the solution of the resulting variational inequalities is the same as the solution of the convex minimization problem. At each iteration of the proposed method, three subproblems are solved and two multipliers are updated. The former two subproblems are first solved in a parallel fashion, and immediately, the multipliers associated to these are updated; then the third subproblem is solved in the sense of an alternative fashion with the former two subproblems, and the multipliers associated to the last two block variables are

updated. The proposed method is essentially to a hybrid splitting method since it combines the parallel splitting method and the alternating direction method, which are two power tools for the convex optimization problem with a separable structure. Under suitable assumptions, the global convergence of the hybrid splitting method is proved. The numerical results on the discrete ill-posed problems show that the proposed method has validity and efficiency.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

All authors contributed equally and significantly in this paper. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup>College of Mathematics and Econometrics, Hunan University, Changsha, 410082, China. <sup>2</sup>College of Mathematics and Computational Science, Hunan First Normal University, Changsha, 410205, China. <sup>3</sup>College of Mathematics and Computer Science, Fuzhou University, Fuzhou, 350108, China.

#### Acknowledgements

This work is supported by Natural Science Foundation of China (11571074), Natural Science Foundation of Fujian Province (2015J01010), Science and Technology Project of Hunan Province (2014SK3235), and Scientific Research Fund of Hunan Provincial Education Department (2015277).

Received: 3 December 2015 Accepted: 19 January 2016 Published online: 18 February 2016

#### References

1. Stephen, B, Lieven, V: Convex Optimization. Cambridge University Press, London (2004)
2. Achziger, W, Hoheisel, T, Kanzow, C: A smoothing-regularization approach to mathematical programs with vanishing constraints. *Comput. Optim. Appl.* **55**(3), 733-767 (2013)
3. Lusem, AN, Svaiter, BF: A new smoothing-regularization approach for a maximum-likelihood estimation problem. *Appl. Math. Optim.* **29**(3), 225-241 (1994)
4. Chen, SF, Goodman, J: An empirical study of smoothing techniques for language modeling. *Comput. Speech Lang.* **13**(4), 359-394 (1999)
5. Huber, PJ: Robust Statistics. Wiley-Interscience, New York (1981)
6. Tibshirani, R: Regression shrinkage and selection via the lasso. *J. R. Stat. Soc., Ser. B* **58**(1), 267-288 (1996)
7. Donoho, D: Compressed sensing. *IEEE Trans. Inf. Theory* **52**(4), 1289-1306 (2006)
8. Candes, EJ, Tao, T: Near optimal signal recovery from random projections: universal encoding strategies. *IEEE Trans. Inf. Theory* **52**(12), 5406-5425 (2006)
9. Zhang, Y: Theory of compressive sensing via  $\ell_1$ -minimization: a non-RIP analysis and extensions. *J. Oper. Res. Soc. China* **1**(1), 79-105 (2013)
10. Chen, CH, He, BS, Ye, YY, Yuan, XM: The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Math. Program.* (2014). doi:10.1007/s10107-014-0826-5
11. He, BS: Parallel splitting augmented Lagrangian methods for monotone structured variational inequalities. *Comput. Optim. Appl.* **42**(2), 195-212 (2009)
12. Peng, Z, Wu, DH: A partial parallel splitting augmented Lagrangian method for solving constrained matrix optimization problems. *Comput. Math. Appl.* **60**(6), 1515-1524 (2010)
13. Glowinski, R, Le Tallec, P: Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. SIAM Studies in Applied Mathematics. SIAM, Philadelphia (1989)
14. He, BS, Liao, LZ, Qian, MJ: Alternating projection based prediction-correction methods for structured variational inequalities. *J. Comput. Math.* **24**(6), 693-710 (2006)
15. He, BS, Xu, MH: A general framework of contraction methods for monotone variational inequalities. *Pac. J. Optim.* **4**(2), 195-212 (2008)
16. Dai, YH, Fletcher, R: Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming. *Numer. Math.* **100**(1), 21-47 (2005)
17. Amir, B, Teboulle, M: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sci.* **2**(1), 183-202 (2009)
18. Hansen, PC: Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Algorithms* **6**(1), 1-35 (1994)