



# Sparse unmixing of hyperspectral data with bandwise model

Chang Li<sup>a</sup>, Yu Liu<sup>a</sup>, Juan Cheng<sup>a</sup>, Rencheng Song<sup>a</sup>, Jiayi Ma<sup>b</sup>, Chenhong Sui<sup>c</sup>,  
Xun Chen<sup>d,\*</sup>

<sup>a</sup> Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China

<sup>b</sup> Electronic Information School, Wuhan University, Wuhan 430072, China

<sup>c</sup> School of Opto-electronic Information Science and Technology, YanTai University, YanTai 264005, China

<sup>d</sup> Department of Electronic Engineering & Information Science, University of Science and Technology of China, Hefei 230026, China

## ARTICLE INFO

### Article history:

Received 26 July 2019

Revised 10 October 2019

Accepted 21 October 2019

Available online 1 November 2019

### Keywords:

Sparse unmixing

Hyperspectral image (HSI)

Bandwise model

Alternative direction method of multipliers

(ADMM)

## ABSTRACT

Sparse unmixing has long been a hot research topic in the area of hyperspectral image (HSI) analysis. Most of the traditional sparse unmixing methods usually assume to only take the Gaussian noise into consideration. However, there are also other types of noise in real HSI, *i.e.*, impulse noise, stripes, dead lines and so on. In addition, the intensity of Gaussian noise is usually different for each band of real HSI. To this end, we propose a novel sparse unmixing method with the bandwise model (SUBM) to address the above mentioned problems simultaneously. Besides, the alternative direction method of multipliers (ADMM) is adopted for solving the proposed SUBM. Moreover, we conduct extensive experiments on synthetic and real datasets to demonstrate effectiveness of the proposed sparse unmixing method under the bandwise model.

© 2019 Elsevier Inc. All rights reserved.

## 1. Introduction

Hyperspectral imaging has long been an effective tool along with the continuous progress of remote sensing technology, and it has been widely used in different kinds of applications, such as geoinformation science, space research, material mapping, and geological exploration [1–7]. However, the phenomenon of mixed pixels is widely existed due to insufficient spatial resolution of HSI, multiple scattering and homogeneous mixture [8,9]. Thus, hyperspectral unmixing is a common way to recover a series of constituent pure spectra (endmembers), and obtain the corresponding proportions (abundances) [10,11].

Unmixing methods can be generally divided into three categories according to the prior knowledge of endmembers, *i.e.*, supervised, unsupervised, and semi-supervised unmixing methods [12]. Supervised unmixing methods can obtain desirable performance when the endmembers are known. However, it is hard to know the spectra of endmembers before hand in real HSI. On the contrary, unsupervised unmixing methods do not need the prior knowledge of endmembers, which can obtain the endmembers and abundances of the given HSI simultaneously. However, the unsupervised unmixing methods are usually derived based on the pure pixel assumption [13], which are usually very hard to satisfy for real HSI. The semi-supervised methods can overcome the above mentioned problems by adopting a set of pure spectra (spectral library) existing in

\* Corresponding author.

E-mail addresses: [changli@hfut.edu.cn](mailto:changli@hfut.edu.cn) (C. Li), [yuliu@hfut.edu.cn](mailto:yuliu@hfut.edu.cn) (Y. Liu), [chengjuan@hfut.edu.cn](mailto:chengjuan@hfut.edu.cn) (J. Cheng), [rcsong@hfut.edu.cn](mailto:rcsong@hfut.edu.cn) (R. Song), [jiayima@whu.edu.cn](mailto:jiayima@whu.edu.cn) (J. Ma), [sui6662008@163.com](mailto:sui6662008@163.com) (C. Sui), [xunchen@ustc.edu.cn](mailto:xunchen@ustc.edu.cn) (X. Chen).

nature, and they use the optimal subset to represent the mixed pixel in HSI, so the abundances obtained by semi-supervised unmixing methods lead to sparse solution, which are essentially sparse unmixing.

Sparse unmixing methods were generally divided into five categories, *i.e.*, Bayesian methods [14–16], greedy methods [17–20], convex relaxation methods [21–28], nonconvex optimization methods [29–33], and multi-objective optimization methods [34–37]. The Bayesian methods were usually derived from the Bayesian analysis and probability statistics. Themelis et al. proposed a hierarchical Bayesian model for HSI unmixing [14], Seyyedsalehi et al. proposed a probabilistic joint sparse method for HSI unmixing [15], and the computational complexity of these methods were usually very high. The greedy methods could obtain the locally optimal solution at each stage, and approximated to the global optimal solution gradually. Shi et al. proposed a novel subspace matching pursuit (SMP) method for HSI unmixing, which aimed at reconstructing the whole HSI by finding each subspace iteratively [18]. Tang et al. proposed a novel regularized simultaneous forward-backward greedy algorithm (RSFoBa) to exploit the joint sparsity among pixels, and the spatial-contextual coherence was enforced within the whole HSI to make it more efficient [19]. Although the greedy methods had the advantage of low computational complexity, they were easily trapped into the local optimum owing to high correlation among endmembers. The convex relaxation methods usually relaxed the original  $\ell_0$  norm to the convex one, such as  $\ell_1$  norm and  $\ell_{2,1}$  norm. Bioucas et al. adopted the  $\ell_1$  norm to characterize the sparsity of abundances, and proposed a variable splitting and augmented Lagrangian (SUnSAL) method for HSI unmixing [21]. Iordache et al. adopted the  $\ell_{2,1}$  norm to characterize sparsity among all pixels, and proposed a collaborative SUnSAL (CLSUnSAL) for HSI unmixing [23]. Li et al. used the weighting  $\ell_1$  regularization to further enforce the sparsity of abundances [28]. The nonconvex optimization methods relaxed the original  $\ell_0$  norm to the nonconvex one. Chen et al. adopted the  $\ell_p$  ( $0 < p < 1$ ) norm to represent the sparsity of abundances, which had numerical advantages over the  $\ell_1$  norm [29]. Shi et al. proposed a collaborative sparse HSI unmixing method with  $\ell_{2,0}$  norm to impose sparsity to all pixels simultaneously, which could obtain a more sparse solution with higher accuracy [32]. Uezato et al. adopted the double sparsity for the selection of both endmember classes and endmember spectra [33]. However, the nonconvex optimization methods usually had higher computational complexity than these of the convex relaxation methods. The multi-objective optimization methods formulated the problem of sparse unmixing as multi-objective optimization problem. Jiang et al. proposed a two phase multi-objective sparse unmixing method [35]. Firstly, it minimized the residuals and estimated number of endmembers simultaneously. Secondly, it minimized the unmixing residuals and the total variation simultaneously. Xu et al. proposed a classification-based model for individual generating under the framework of multi-objective optimization based sparse unmixing [37].

To sum up, the above mentioned sparse unmixing methods usually only take the Gaussian noise into consideration, which usually occurs due to sensor noise or dark current variations [38]. In most circumstances, the intensities of Gaussian noise are assumed to be the same for all bands of HSI. However, the intensities of Gaussian noise usually vary for different bands of HSI. Besides, there are also other types of noise in real HSI, *i.e.*, impulse noise, stripes, dead lines and so on [39]. The impulse noise usually occurs when some of the sensors do not work or become saturated [40]. Dead pixels occur when the detectors are abnormal, they are usually missing or present as zero values [41]. Stripes usually occur due to temporal variations of the detector characteristics [38].

In this paper, we propose a new sparse unmixing method with the bandwise unmixing model (SUBM) to overcome the above mentioned problems. Different from the most traditional unmixing methods, the proposed SUBM can take not only different intensities of Gaussian noise for different bands of HSI, but also some other kinds of noise into consideration. In summary, the main contributions of this paper are summarized as follows.

1. We propose a bandwise unmixing model, which can formulate each band of HSI as the linear combination of spectral library and the abundances in the presence of various kinds of noise. Besides, the Gaussian noise in each band is assumed to be independent and identically distributed, and the standard deviation of Gaussian noise in different bands can be different from each other.
2. We propose a new sparse unmixing method based on the bandwise unmixing model, which can take different kinds of noise into consideration under the maximum a posteriori framework. Besides, the alternative direction method of multipliers (ADMM) is adopted for solving the proposed SUBM.
3. We conduct extensive experiments on both synthetic and real datasets to demonstrate advantages of the proposed unmixing model and the corresponding unmixing method SUBM.

The rest of this paper is structured as follows. In Section 2, a new sparse unmixing method SUBM is proposed, and the ADMM is adopted for solving the proposed SUBM. Section 3 gives the sparse unmixing experimental results, and the conclusion is drawn in Section 4. For convenience of reference, Table 1 shows the list of notations used in this paper.

## 2. Bandwise unmixing model and algorithm

### 2.1. Formulation of the bandwise unmixing model and the corresponding algorithm

The linear mixing model (LMM) has been widely used in sparse unmixing of HSI due to its simplicity and flexibility, and a mixed pixel can be represented by a linear combination of endmembers weighted by the corresponding abundances. Let  $\mathbf{y} \in \mathbf{R}^{D \times 1}$  denote the spectral vector of a given HSI with  $D$  bands, and  $\mathbf{E} \in \mathbf{R}^{D \times M}$  denotes the spectral library with  $M$  pure

**Table 1**

List of notations.

Notation	Description	Notation	Description
$\mathbf{R}^{D \times M}, \mathbf{R}^{D \times 1}$	$D \times M$ and $D \times 1$ Euclidean space	$\mathbf{y}, \mathbf{Y}$	Spectral vector and matrix
$\mathbf{a}, \mathbf{A}$	Abundance vector and matrix	$\mathbf{n}, \mathbf{N}$	Noise vector and matrix
$\mathbf{E}$	Spectral library	$D, M, P$	Number of bands, endmembers and pixels
$\mathbf{Y}_i, (\mathbf{EA})_i, \mathbf{S}_i, \mathbf{N}_i$	$i$ th band of $\mathbf{Y}, \mathbf{EA}, \mathbf{S}$ and $\mathbf{N}$	$\sigma_i^2$	Variance of Gaussian noise in $i$ th band
$p(\mathbf{X})$	Probability function of $\mathbf{X}$	$\mathbf{W}$	Diagonal matrix
$\mathbf{S}$	Sparse noise	$\ \cdot\ _F, \ \cdot\ _0, \ \cdot\ _1, \ \cdot\ _{2,1}$	Frobenius, $\ell_0$ , $\ell_1$ and $\ell_{2,1}$ norm
$\hat{\mathbf{A}}, \hat{\mathbf{S}}$	Estimation of $\mathbf{A}$ and $\mathbf{S}$	$\lambda, \alpha$	Regularization parameters
$\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3$	Auxiliary variables	$\mathbf{G}, \mathbf{Q}, \mathbf{H}, \mathbf{V}, \mathbf{Z}$	Auxiliary matrices
$\mu, \Lambda/\mu$	Penalty parameter and Lagrange multipliers	$r^{k+1}, d^{k+1}$	Primal and dual residuals
$\varepsilon, N$	Error threshold and maximum number of iterations	$\mathbf{A}^{k+1}, \mathbf{S}^{k+1}, \mathbf{V}_1^{k+1}, \mathbf{V}_2^{k+1}, \mathbf{V}_3^{k+1}$	$k$ th iteration of $\mathbf{A}, \mathbf{S}, \mathbf{V}_1, \mathbf{V}_2$ and $\mathbf{V}_3$

endmembers. Mathematically, the LMM can be formulated as follows:

$$\mathbf{y} = \mathbf{E}\mathbf{a} + \mathbf{n}, \quad (1)$$

where  $\mathbf{a} \in \mathbf{R}^{M \times 1}$  represents the abundance, and  $\mathbf{n}$  denotes the noise. The LMM is usually assumed to satisfy the abundance non-negativity constraint (ANC) and abundance sum to one constraint (ASC) as follows:

$$\begin{aligned} \mathbf{a} &\geq \mathbf{0}, \\ \mathbf{1}^T \mathbf{a} &= 1. \end{aligned} \quad (2)$$

However, the spectral vector of real HSI is usually variable [17], and the abundances cannot always meet the ASC, so the ASC is not considered in this paper. Thus, the LMM for HSI with total  $P$  pixels can be written as follows:

$$\mathbf{Y} = \mathbf{E}\mathbf{A} + \mathbf{N}, \quad (3)$$

where  $\mathbf{Y} \in \mathbf{R}^{D \times P}$  is the reshaped HSI,  $\mathbf{A} \in \mathbf{R}^{M \times P}$  denotes the abundance matrix for all  $P$  pixels, and  $\mathbf{N} \in \mathbf{R}^{D \times P}$  represents the noise matrix.

Most of the traditional sparse unmixing methods are implicitly degraded by Gaussian noise, and the intensities of Gaussian noise are usually the same for all bands of HSI. However, the intensities of Gaussian noise usually vary for different bands of HSI. Besides, for the HSI acquired in the real world, it is usually contaminated not merely by Gaussian noise, and there are also other different kinds of noise [39,42,43]. Most of the HSI are usually contaminated by Gaussian noise, which can be categorized as dense noise. Only a small fraction of the HSI is degraded by other types of noise, which can be categorized as sparse noise, such as dead pixels or lines, stripes and impulse noise. Thus, the proposed bandwise model for sparse unmixing in the presence of mixed noise can be written as follows:

$$\mathbf{Y}_i = (\mathbf{EA})_i + \mathbf{S}_i + \mathbf{N}_i, \quad i = 1, \dots, D, \quad (4)$$

where  $\mathbf{Y}_i, (\mathbf{EA})_i, \mathbf{S}_i$  and  $\mathbf{N}_i$  denote the  $i$ th band of  $\mathbf{Y}, \mathbf{EA}, \mathbf{S}$  and  $\mathbf{N}$ , respectively, and  $\mathbf{S}$  denotes the sparse noise. It is assumed that the Gaussian noise in each band is identically distributed and independent from each other, it can be written as  $\mathbf{N}_i \sim \mathcal{N}(\mathbf{0}, \sigma_i^2 \mathbf{I}_P)$  ( $i = 1, \dots, D$ ),  $\sigma_i^2$  denotes the variance of Gaussian noise in  $i$ th band, and it is naturally to obtain  $p(\mathbf{Y}_i | [(\mathbf{EA})_i + \mathbf{S}_i]) \sim \mathcal{N}((\mathbf{EA})_i + \mathbf{S}_i, \sigma_i^2 \mathbf{I}_P)$  ( $i = 1, \dots, M$ ). Thus, we can obtain

$$\begin{aligned} p(\mathbf{Y} | (\mathbf{EA} + \mathbf{S})) &= \prod_{i=1}^D p(\mathbf{Y}_i | [(\mathbf{EA})_i + \mathbf{S}_i]) \\ &= c \exp \left( - \sum_{i=1}^D \frac{1}{2\sigma_i^2} \|\mathbf{Y}_i - (\mathbf{EA})_i - \mathbf{S}_i\|_2^2 \right) \\ &= c \exp \left( - \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{S})\|_F^2 \right), \end{aligned} \quad (5)$$

where  $c$  is a constant,  $\mathbf{W} \in \mathbb{R}^{D \times D}$  denotes the diagonal matrix,  $\mathbf{W}_{i,i} = \frac{1}{\sigma_i^2}$  ( $i = 1, \dots, D$ ), and  $\|\cdot\|_F$  denotes the Frobenius norm. Thus, we can estimate the abundance and sparse noise based on the maximum a posteriori criterion as follows:

$$\begin{aligned} [\hat{\mathbf{A}}, \hat{\mathbf{S}}] &= \arg \max_{\mathbf{A}, \mathbf{S}} p((\mathbf{EA} + \mathbf{S}) | \mathbf{Y}) \\ &= \arg \max_{\mathbf{A}, \mathbf{S}} p(\mathbf{Y} | (\mathbf{EA} + \mathbf{S})) p(\mathbf{EA} + \mathbf{S}) \\ &= \arg \min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{S})\|_F^2 - \ln(p(\mathbf{EA} + \mathbf{S})), \end{aligned} \quad (6)$$

where  $p(\mathbf{EA} + \mathbf{S})$  denotes the prior distribution,  $\mathbf{E}$  is the spectral library known in advance, and  $\ln(p(\mathbf{EA} + \mathbf{S}))$  is a function of  $\mathbf{A}$  and  $\mathbf{S}$ , so it can be seen as the prior knowledge or regularization constraint on  $\mathbf{A}$  and  $\mathbf{S}$ .

Since the pixels should have the same active set of endmembers in the spectral library, so the abundance should have only a small number of nonzero lines [23], which has the underlying collaborative sparse property, and it can be modeled by the  $\ell_{2,1}$  norm. Besides, the sparse noise has the underlying sparse property, and it can be modeled by the  $\ell_0$  norm. However, the  $\ell_0$  norm optimization problem is usually NP-hard to solve, and we adopt the  $\ell_1$  norm to surrogate the original  $\ell_0$  norm. Moreover, the abundance should satisfy the ANC constraint. Thus, (6) can be reformulated as follows:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{S})\|_F^2 + \lambda \|\mathbf{A}\|_{2,1} + \alpha \|\mathbf{S}\|_1, \\ \text{s.t. } \mathbf{A} \geq \mathbf{0}, \end{aligned} \quad (7)$$

where  $\lambda$  and  $\alpha$  denote the regularization parameters,  $\ell_{2,1}$  norm is defined as  $\|\mathbf{A}\|_{2,1} = \sum_{i=1}^M \sqrt{\sum_{j=1}^P \mathbf{A}_{ij}^2}$ , and  $\ell_1$  norm is defined as  $\|\mathbf{S}\|_1 = \sum_{i=1}^D \sum_{j=1}^P |\mathbf{S}_{ij}|$ .

## 2.2. Solving the proposed SUBM with ADMM

Since the ADMM has been widely used for solving HSI unmixing models and the other various fields, and they have achieved desirable performances, so we also adopt the ADMM to solve the proposed SUBM, which comes down to solving (7). For more details of ADMM, please refer to Boyd et al. [44].

Three auxiliary variables  $\mathbf{V}_1$ ,  $\mathbf{V}_2$  and  $\mathbf{V}_3$  are introduced to split variables, and (7) can be rewritten as follows:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{S}, \mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3} \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{V}_1)\|_F^2 + \lambda \|\mathbf{V}_2\|_{2,1} + l_{\mathcal{R}_+}(\mathbf{V}_3) + \alpha \|\mathbf{S}\|_1, \\ \text{s.t. } \mathbf{V}_1 = \mathbf{S}, \\ \mathbf{V}_2 = \mathbf{A}, \\ \mathbf{V}_3 = \mathbf{A}, \end{aligned} \quad (8)$$

where  $l_{\mathcal{R}_+}(\mathbf{X}) = \sum_{i,j} l_{\mathcal{R}_+}(\mathbf{X}_{i,j})$ ,  $\mathbf{X}_{i,j}$  represents the  $i$ ,  $j$ th element of  $\mathbf{X}$ , and  $l_{\mathcal{R}_+}(\mathbf{X}_{i,j})$  is zero when  $\mathbf{X}_{i,j}$  is in the nonnegative orthant, and  $+\infty$  otherwise.

For simplicity, (8) can be also written as follows:

$$\min_{\mathbf{V}, \mathbf{Q}} g(\mathbf{V}, \mathbf{Q}) \text{ s.t. } \mathbf{GQ} + \mathbf{HV} = \mathbf{Z}, \quad (9)$$

where  $g(\mathbf{V}, \mathbf{Q}) = \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{V}_1)\|_F^2 + \lambda \|\mathbf{V}_2\|_{2,1} + l_{\mathcal{R}_+}(\mathbf{V}_3) + \alpha \|\mathbf{S}\|_1$ ,  $\mathbf{G} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \end{bmatrix}$ ,  $\mathbf{Q} = \begin{bmatrix} \mathbf{S} \\ \mathbf{A} \\ \mathbf{A} \end{bmatrix}$ ,  $\mathbf{H} = \begin{bmatrix} -\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -\mathbf{I} \end{bmatrix}$ ,

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \mathbf{V}_3 \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}.$$

The augmented Lagrangian for (9) is

$$\mathcal{L}(\mathbf{V}, \mathbf{Q}, \Lambda) = g(\mathbf{V}, \mathbf{Q}) + \frac{\mu}{2} \|\mathbf{GQ} + \mathbf{HV} - \mathbf{Z} - \Lambda\|_F^2, \quad (10)$$

where  $\mu > 0$  denotes the penalty parameter,  $\Lambda/\mu$  represents the Lagrange multipliers. Then, the ADMM scheme minimizes the augmented Lagrangian  $\mathcal{L}(\mathbf{V}, \mathbf{Q}, \Lambda)$  sequentially, i.e., minimizes one single variable by fixing the remaining variables.

When updating  $\mathbf{A}$ , these terms independent of  $\mathbf{A}$  in (10) can be ignored, and the corresponding optimization problem becomes

$$\begin{aligned} \mathbf{A}^{k+1} &= \arg \min_{\mathbf{A}} \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA} - \mathbf{V}_1^k)\|_F^2 + \frac{\mu}{2} (\|\mathbf{V}_2^k - \mathbf{A} - \Lambda_2^k\|_F^2 + \|\mathbf{V}_3^k - \mathbf{A} - \Lambda_3^k\|_F^2) \\ &= [(\mathbf{WE})^T (\mathbf{WE}) + 2\mu \mathbf{I}]^{-1} [(\mathbf{WE})^T \mathbf{W}(\mathbf{Y} - \mathbf{V}_1^k) + \mu(\mathbf{V}_2^k - \Lambda_2^k + \mathbf{V}_3^k - \Lambda_3^k)]. \end{aligned} \quad (11)$$

When updating  $\mathbf{S}$ , the corresponding optimization problem becomes

$$\begin{aligned} \mathbf{S}^{k+1} &= \arg \min_{\mathbf{S}} \alpha \|\mathbf{S}\|_1 + \frac{\mu}{2} \|\mathbf{V}_1^k - \mathbf{S} - \Lambda_1^k\|_F^2 \\ &= S_{\alpha/\mu}(\mathbf{V}_1^k - \Lambda_1^k), \end{aligned} \quad (12)$$

where  $S_\tau[x] = \text{sgn}(x) \max(|x| - \tau, 0)$  represents the well-known soft shrinkage operator [45], and  $\tau$  represents the threshold.

When updating  $\mathbf{V}_1$ , the corresponding optimization problem becomes

$$\begin{aligned} \mathbf{V}_1^{k+1} &= \arg \min_{\mathbf{V}_1} \frac{1}{2} \|\mathbf{W}(\mathbf{Y} - \mathbf{EA}^{k+1} - \mathbf{V}_1)\|_F^2 + \frac{\mu}{2} \|\mathbf{V}_1 - \mathbf{S}^{k+1} - \Lambda_1^k\|_F^2 \\ &= [\mathbf{W}^T \mathbf{W} + \mu \mathbf{I}]^{-1} [\mathbf{W}^T \mathbf{W}(\mathbf{Y} - \mathbf{EA}^{k+1}) + \mu(\mathbf{S}^{k+1} + \Lambda_1^k)]. \end{aligned} \quad (13)$$

When updating  $\mathbf{V}_2$ , the corresponding optimization problem becomes

$$\mathbf{V}_2^{k+1} = \arg \min_{\mathbf{V}_2} \lambda \|\mathbf{V}_2\|_{2,1} + \frac{\mu}{2} \|\mathbf{V}_2 - \mathbf{A}^{k+1} - \Lambda_2^k\|_F^2, \quad (14)$$

it can be solved by applying the well-known vect-soft threshold operator [46] to each row of  $\zeta$  as follows:

$$\mathbf{V}_2^{k+1}(i, :) = \text{vect-soft} \left( \zeta(i, :), \frac{\lambda}{\mu} \right), \quad i = 1, \dots, M, \quad (15)$$

where  $\zeta = \mathbf{A}^{k+1} + \Lambda_2^k$ , and  $\text{vect-soft}(b, \tau) = b \frac{\max\{\|b\|_2 - \tau, 0\}}{\max\{\|b\|_2 - \tau, 0\} + \tau}$ .

When updating  $\mathbf{V}_3$ , the corresponding optimization problem becomes

$$\begin{aligned} \mathbf{V}_3^{k+1} &= \arg \min_{\mathbf{V}_3} l_{\mathcal{R}_+}(\mathbf{V}_3) + \frac{\mu}{2} \|\mathbf{V}_3 - \mathbf{A}^{k+1} - \Lambda_3^k\|_F^2 \\ &= \max(\mathbf{A}^{k+1} + \Lambda_3^k, \mathbf{0}). \end{aligned} \quad (16)$$

The primal residual  $r^{k+1}$  and dual residual  $d^{k+1}$  become

$$r^{k+1} = \mathbf{G}\mathbf{Q}^{k+1} + \mathbf{H}\mathbf{V}^{k+1}, \quad (17)$$

$$d^{k+1} = \mu \mathbf{G}^T \mathbf{H}(\mathbf{Q}^{k+1} - \mathbf{Q}^k). \quad (18)$$

We update  $\mu$  to keep the ratio between the primal residual norm and dual residual norm within a positive interval. The stopping criterion becomes

$$\begin{aligned} \|r^{k+1}\|_F / \sqrt{(3M+D)P} &\leq \varepsilon \text{ and} \\ \|d^{k+1}\|_F / \sqrt{(3M+D)P} &\leq \varepsilon. \end{aligned} \quad (19)$$

When using the ADMM to solve the proposed SUBM, the diagonal matrix  $\mathbf{W}$  should be known, so the standard deviation of Gaussian noise in each band should be estimated before hand. We adopt a popular noise estimation method named hyperspectral signal identification by minimum error (HySime) [47] to estimate the standard deviation of Gaussian noise in each band of HSI. The key idea of HySime is based on the fact that the neighboring bands of HSI usually have high correlation, and it is modeled by the multiple regression theory. To sum up, we summarize the procedure to solve the proposed SUBM in Algorithm 1.

---

**Algorithm 1:** Solving the proposed SUBM with ADMM.

---

**Input:**  $\mathbf{Y}, \mathbf{E};$

**Output:**  $\mathbf{A}, \mathbf{S};$

```

1 while  $k < N$  & (19) is not satisfied do
2   Update  $\mathbf{W}$  via HySime [47];
3    $\mathbf{A}^{k+1} = [(\mathbf{W}\mathbf{E})^T(\mathbf{W}\mathbf{E}) + 2\mu\mathbf{I}]^{-1}[(\mathbf{W}\mathbf{E})^T\mathbf{W}(\mathbf{Y} - \mathbf{V}_1^k) + \mu(\mathbf{V}_2^k - \Lambda_2^k + \mathbf{V}_3^k - \Lambda_3^k)];$ 
4    $\mathbf{S}^{k+1} = \mathcal{S}_{\alpha/\mu}(\mathbf{V}_1^k - \Lambda_1^k);$ 
5    $\mathbf{V}_1^{k+1} = [\mathbf{W}^T\mathbf{W} + \mu\mathbf{I}]^{-1}[\mathbf{W}^T\mathbf{W}(\mathbf{Y} - \mathbf{E}\mathbf{A}^{k+1}) + \mu(\mathbf{S}^{k+1} + \Lambda_1^k)];$ 
6   Update  $\mathbf{V}_2^{k+1}$  by (15);
7    $\mathbf{V}_3^{k+1} = \max(\mathbf{A}^{k+1} + \Lambda_3^k, \mathbf{0});$ 
8    $\Lambda_1^{k+1} = \Lambda_1^k - (\mathbf{V}_1^{k+1} - \mathbf{S}^{k+1});$ 
9    $\Lambda_2^{k+1} = \Lambda_2^k - (\mathbf{V}_2^{k+1} - \mathbf{A}^{k+1});$ 
10   $\Lambda_3^{k+1} = \Lambda_3^k - (\mathbf{V}_3^{k+1} - \mathbf{A}^{k+1});$ 
11   $k = k + 1;$ 
12 end
13 return  $\mathbf{A} = \mathbf{A}^{k+1}, \mathbf{S} = \mathbf{S}^{k+1}.$ 

```

---

### 2.3. Computational complexity of the proposed SUBM

Concerning the computational complexity of the proposed SUBM, updating  $\mathbf{W}$  has the computational complexity of  $4D^2P + 6D^3$ . Since the number of pixels  $P$  is usually bigger than the number of bands  $D$  in HSI, so updating  $\mathbf{W}$  has the complexity of  $\mathcal{O}(D^2P)$ . For updating of the other variables, the main time complexity lies in the computation of  $\mathbf{A}$ , which has the complexity of  $\mathcal{O}(D^2P)$ . Thus, the computational complexity of the proposed SUBM is  $\mathcal{O}(D^2P)$ .

**Table 2**

The parameter setting of convex relaxation methods SUnSAL, CLSUnSAL and the proposed SUBM.

Algorithm	$\lambda$	$\alpha$	$\mu$	$\varepsilon$	$N$
SUnSAL	$\{10^{-5}, 10^{-4}, \dots, 1, \dots, 10^4, 10^5\}$	Not applicable	$10^{-2}$	$10^{-6}$	1000
CLSUnSAL	$\{10^{-5}, 10^{-4}, \dots, 1, \dots, 10^4, 10^5\}$	Not applicable	$10^{-2}$	$10^{-6}$	1000
SUBM	$\{10^{-5}, 10^{-4}, \dots, 1, \dots, 10^4, 10^5\}$	$\{10^{-5}, 10^{-4}, \dots, 1, \dots, 10^4, 10^5\}$	$10^{-2}$	$10^{-6}$	1000

### 3. Experiments

In this section, we verify the unmixing performance of the proposed SUBM using both synthetic and real datasets experiments, and the compared methods include SMP [18], RSFoBa [19], SUnSAL [21] and CLSUnSAL [23]. The SMP and RSFoBa belong to greedy methods, SUnSAL and CLSUnSAL belong to convex relaxation methods. For the greedy method SMP, two parameters have large impact on the final unmixing performance, *i.e.*, the block size and the threshold  $t$ . We tune the block size  $\in \{5, 10, 15, 20, 25\}$ , and the threshold  $t$  is fixed to 0.96 according to Shi et al. [18]. Besides, the computational complexity of SMP is  $\mathcal{O}(DPM)$  according to Shi et al. [18]. For the greedy method RSFoBa, there are mainly two parameters to be tuned, *i.e.*, the block size and the regularization parameter. We tune the block size  $\in \{5, 10, 15, 20, 25\}$ , and tune the regularization parameter  $\in \{0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.3, 0.5, 1, 3, 5, 10\}$ . Besides, the computational complexity of RSFoBa is  $\mathcal{O}(\max(DPM, P^2/L))$  according to Tang et al. [19], where  $L$  is the block number in the HSI. For the convex relaxation methods SUnSAL, CLSUnSAL and the proposed SUBM, it is found that the regularization parameters  $\lambda$  and  $\alpha$  have larger impact on the final sparse unmixing performance than the Lagrange multiplier  $\mu$ , error tolerance  $\varepsilon$  and the maximum number of iterations  $N$ , so we set  $\mu$ ,  $\varepsilon$  and  $N$  to fixed value, and tune  $[\lambda, \alpha] \in \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10^1, 10^2, 10^3, 10^4, 10^5\}$ , which can be shown in Table 2. Moreover, the computational complexity of SUnSAL and CLSUnSAL are  $\mathcal{O}(D^2P)$  according to Bioucas-Dias and Figueiredo [21], lordache et al. [23].

We use the root mean square error (RMSE) to evaluate the performance of the proposed SUBM quantitatively [48]. The definition of RMSE is as follows:

$$\text{RMSE} = \frac{\|\mathbf{A} - \hat{\mathbf{A}}\|_F^2}{\sqrt{MP}}, \quad (20)$$

where  $\mathbf{A}$  and  $\hat{\mathbf{A}}$  denote the actual and estimated abundance, respectively. In general, the smaller RMSE indicates better sparse unmixing performance.

#### 3.1. Experimental results with synthetic data

In this subsection, we adopt subset of the United States Geological Survey (USGS)<sup>1</sup> as our spectral library, which has 240 endmembers, and each endmember has 224 bands ranging from 0.4  $\mu\text{m}$  to 2.5  $\mu\text{m}$ .

In the first synthetic data experiment, we generate a synthetic HSI containing 224 bands and  $75 \times 75$  pixels. The abundances are generated according to lordache et al. [22]<sup>2</sup>, which aim at generating both pure and mixed regions. Each mixed region is a mixture of two to five endmembers, and the generated abundance maps can be observed in Fig. 1. The fractional abundances of five endmembers in the background pixels are randomly set as 0.1149, 0.0741, 0.2003, 0.2055, and 0.4051, respectively. According to the bandwise unmixing model in (4), we add noise to the HSI as follows:

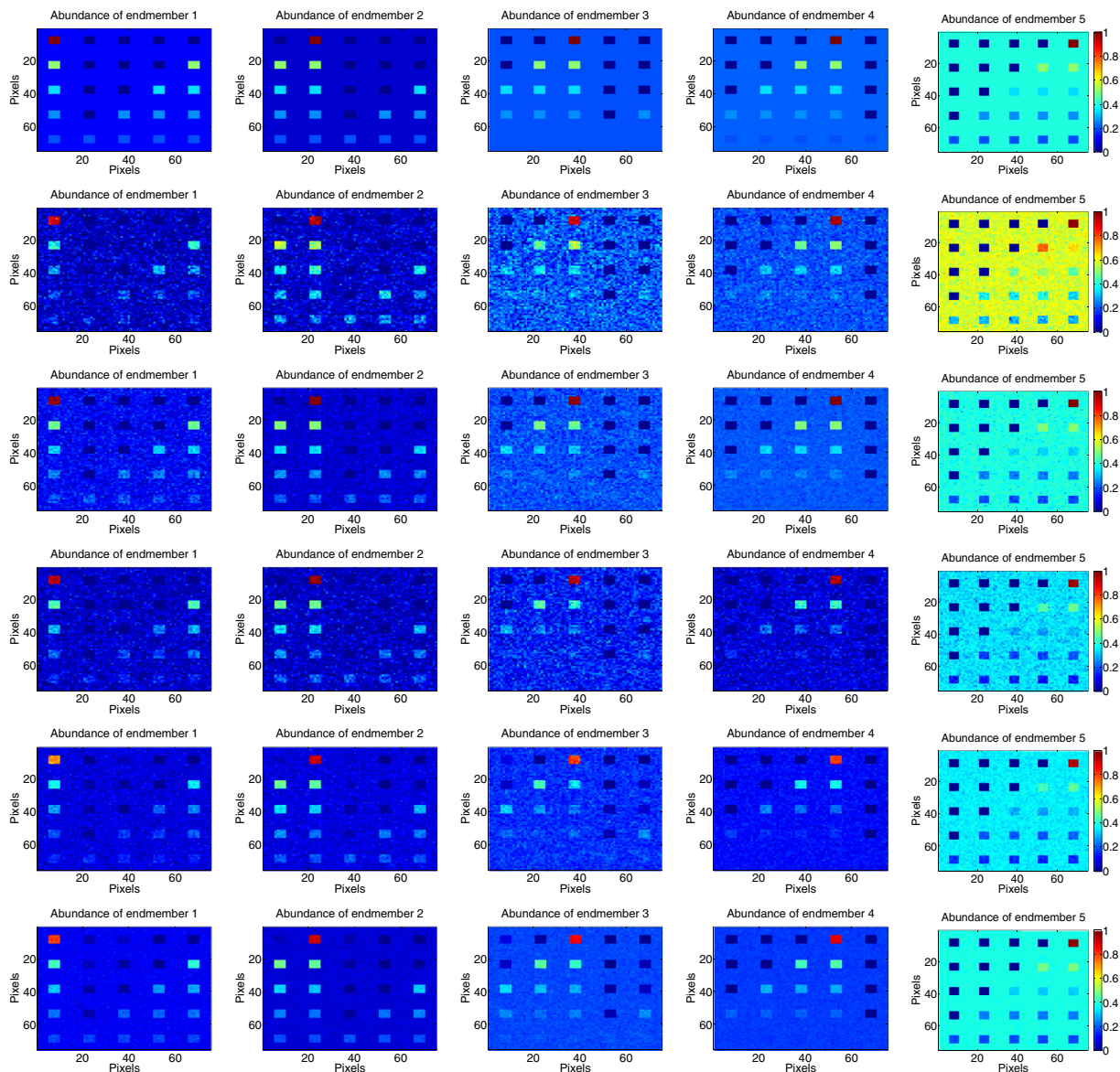
1. Gaussian noise: we add Gaussian noise to each band of HSI with the signal-to-noise ratio (SNR) ranging from 20dB to 40 dB, and the mean SNR is 29.52 dB.
2. Impulse noise: we only degrade 11 bands (60–70) of HSI by 30% impulse noise.
3. Dead lines: we only degrade 11 bands (120–130) of HSI by dead lines.

First, we conduct experiments to evaluate unmixing performance of different methods in the presence of different types of noise. Table 3 shows sparse unmixing performance and time consumption of the proposed SUBM and the compared methods. We can observe from Table 3 that the proposed SUBM can obtain smaller RMSE than CLSUnSAL when only contaminated by Gaussian noise, which is due to that the SUBM considers the variation of Gaussian noise in different bands of HSI. Fig. 1 presents the abundance maps of different methods when only contaminated by Gaussian noise. It can be observed that the abundances obtained by the SUBM are more close to the ground truth than the compared methods, which also demonstrate the effectiveness of considering the different noise level of Gaussian noise in different bands. Besides, the proposed SUBM can obtain much smaller RMSE than the compared methods in the presence of impulse noise or dead lines, the underlying reason is that the compared methods do not take the impulse noise or dead lines into consideration, and they all adopt the  $\ell_2$  norm based loss function, which is sensitive to non-Gauss noise. Only the proposed SUBM takes the underlying sparse property of impulse noise or dead lines. Moreover, our method can get much smaller RMSE than the

<sup>1</sup> Available at: <http://speclab.cr.usgs.gov/spectral-lib.html>.

<sup>2</sup> Available at: <http://www.lx.it.pt/~bioucas/publications.html>.





**Fig. 1.** Abundance maps obtained by different methods when only contaminated by Gaussian noise for the first synthetic data experiment. From top to bottom: ground truth, SMP, RSFoBa, SUnSAL, CLSUnSAL and SUBM ( $\lambda = 1$ ,  $\alpha = 10$ ).

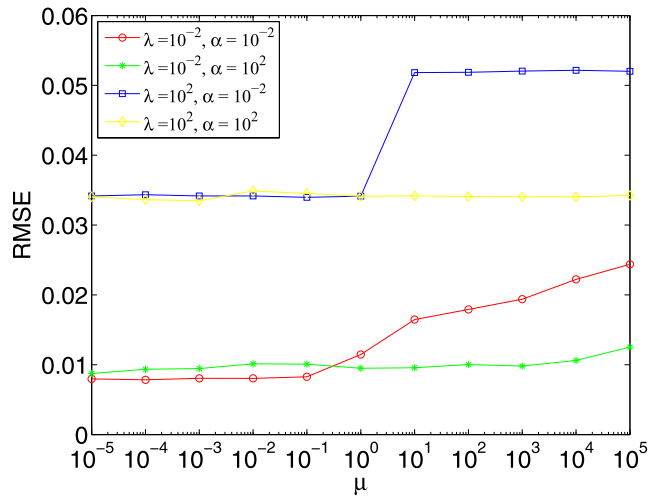
compared methods in the presence of mixed noise, which is due to that the SUBM can take different level of Gaussian noise of different bands and the underlying sparse property of sparse noise into account simultaneously. Furthermore, the SUBM and CLSUnSAL can generally obtain better unmixing performance than SUnSAL, this is due to that the SUnSAL adopts the  $\ell_1$  norm to represent the sparsity of abundances, while the proposed SUBM and CLSUnSAL employ the  $\ell_{2,1}$  norm to represent the collaborative sparse property of abundances. Finally, the proposed SUBM takes more time than the compared methods, which is due to that the proposed SUBM needs to estimate the noise level of different bands of HSI, and it may converge slower than SUnSAL and CLSUnSAL in practice.

Second, we assess the influence of the regularization parameters  $\lambda$ ,  $\alpha$  and Lagrange multiplier  $\mu$  on the final unmixing performance in the first synthetic data experiment. Fig. 2 shows the RMSE of SUBM with the varying Lagrange multiplier  $\mu$  for different  $\alpha$  and  $\lambda$  in the first data experiment. We can see that the RMSE of SUBM remains stable when  $\mu$  is less than  $10^0$ , and the main reason lies in that in the iteration process, we update  $\mu$  to keep the ratio between the primal residual norm and dual residual norm within a positive interval. This means that the  $\mu$  is changing during the whole iteration process to keep the ratio between the primal residual norm and dual residual norm within a positive interval, which makes the RMSE of SUBM remains stable when  $\mu$  is less than  $10^0$ . Thus, we adopt  $\mu = 10^{-2}$  for SUBM. Fig. 3 shows RMSE of SUBM with the varying regularization parameter  $\lambda$  for different  $\alpha$  in the presence of Gaussian noise, impulse noise and

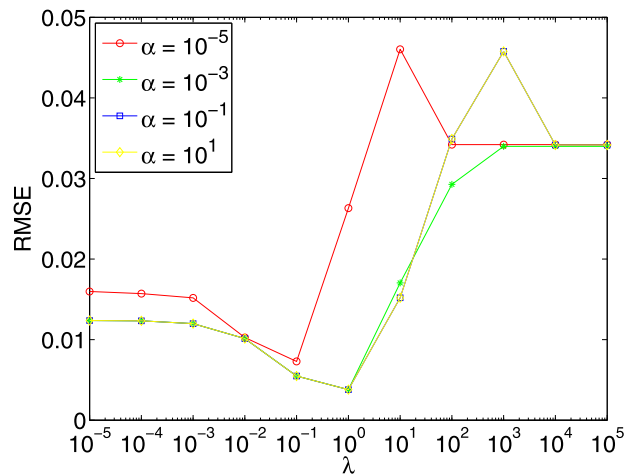
**Table 3**

Comparison of RMSEs ( $\times 10^{-2}$ ) and Time (s) of different methods under different noise condition for the first synthetic data experiment.

Type of noise		SMP	RSFoBa	SUNsAL	CLSunSAL	SUBM
Gaussian noise	RMSE	1.621	0.538	1.486	0.861	<b>0.354</b>
	Time	3.82	<b>1.75</b>	5.01	5.14	76.72
Impulse noise	RMSE	1.775	1.016	2.220	1.875	<b>0.120</b>
	Time	2.63	<b>2.28</b>	6.02	5.96	76.53
Dead lines	RMSE	1.897	0.994	2.233	1.963	<b>0.104</b>
	Time	2.63	<b>2.45</b>	6.04	5.90	76.58
Gaussian noise & Impulse noise	RMSE	1.808	1.140	2.301	1.999	<b>0.361</b>
	Time	2.48	<b>2.36</b>	6.02	5.95	75.42
Gaussian noise & Dead lines	RMSE	2.290	1.143	2.321	2.073	<b>0.370</b>
	Time	3.68	<b>2.08</b>	6.38	6.15	75.43
Impulse noise & Dead lines	RMSE	2.069	1.371	2.575	2.264	<b>0.107</b>
	Time	2.53	<b>1.91</b>	5.08	5.15	76.16
Gaussian noise & Impulse noise & Dead lines	RMSE	2.135	1.491	2.636	2.268	<b>0.379</b>
	Time	2.43	<b>2.01</b>	5.41	5.33	75.25



**Fig. 2.** RMSE of SUBM with the varying regularization parameter  $\mu$  for different  $\alpha$  and  $\lambda$  in the first synthetic data experiment.



**Fig. 3.** RMSE of SUBM with the varying regularization parameter  $\lambda$  for different  $\alpha$  in the first synthetic data experiment.



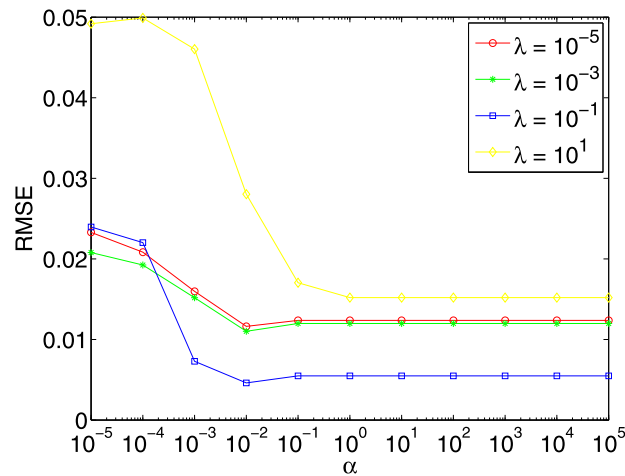


Fig. 4. RMSE of SUBM with the varying regularization parameter  $\alpha$  for different  $\lambda$  in the first synthetic data experiment.

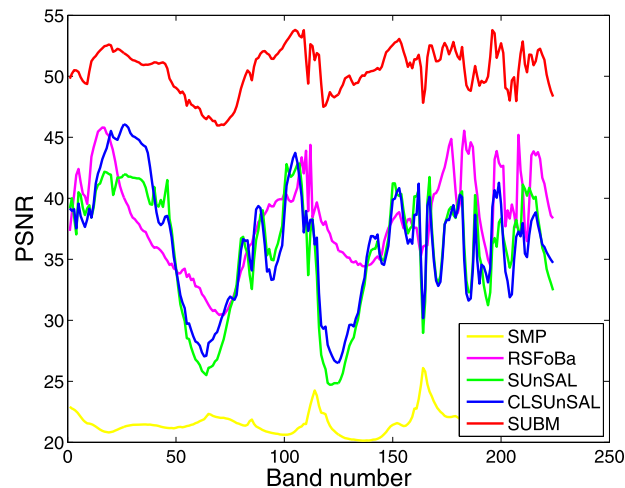
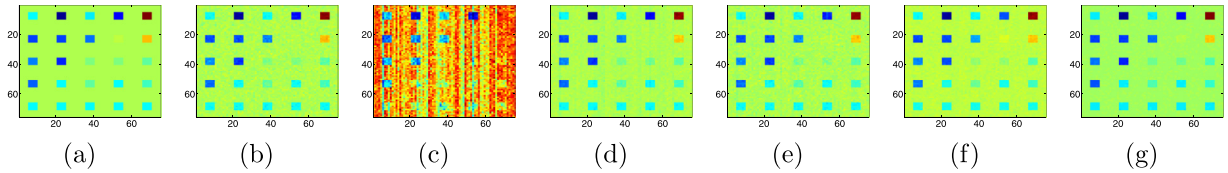


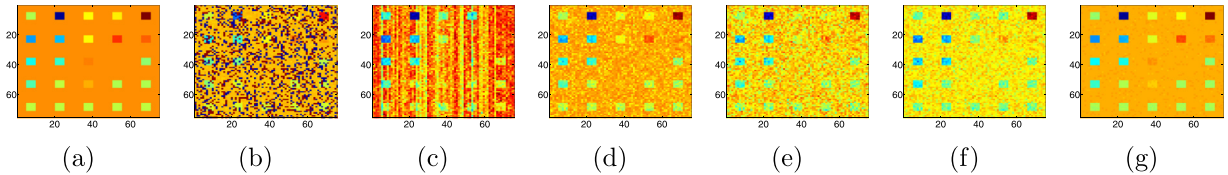
Fig. 5. PSNR of each band using the SUBM and the compared methods in the first synthetic data experiment.

dead lines. It can be clearly seen from Fig. 3 that the RMSE decreases as  $\lambda$  increases at first, and they obtain the best RMSE for different  $\alpha$ . However, they begin to increase when  $\lambda$  increases to a certain level. Fig. 4 shows RMSE of the proposed SUBM as a function of varying regularization parameter  $\alpha$  for different  $\lambda$  when contaminated by the three mixed noise. We see that the RMSE decreases when  $\alpha$  increases at first, and then they get the best RMSE. However, they gradually increase and remain stable when  $\alpha$  gets to a certain extent. From Figs. 3 and 4, we can be see that different  $\lambda$  and  $\alpha$  lead to different unmixing performances, the SUBM can obtain better unmixing performances when the two regularization terms make a proper contribution to the whole objective function, and it is beneficial to improve the unmixing performance when considering the underlying collaborative sparsity of abundance and the sparsity of sparse noise simultaneously.

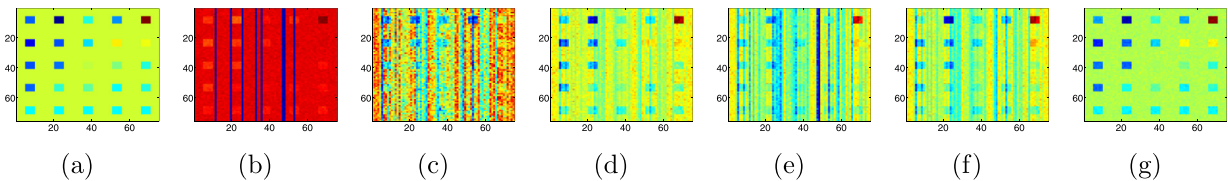
Third, we evaluate the denoising performance by different unmixing methods in the first synthetic data experiment. The denoised HSI can be obtained by multiplying the spectral library by the estimated abundances with different unmixing methods. Fig. 5 presents the peak-signal-to-noise ratio (PSNR) of each band with the proposed SUBM and the compared methods when contaminated by three mixed noise, and larger PSNR generally indicates better denoising performance. We can see from Fig. 5 that the SUBM can get the best PSNR for all bands of the denoised HSI reconstructed by different unmixing methods, which indicates that the SUBM can get the best denoising performance quantitatively. Besides, we also select three representative bands for qualitative comparison. Fig. 6 shows the denoising result of band 20 by different unmixing methods when it is only contaminated by Gaussian noise. It can be observed from Fig. 6 that the RSFoBa, SUnSAL, CLSUnSAL and SUBM can obtain obvious better denoising performance than SMP, and there still exist some Gaussian noise for RSFoBa, SUnSAL and CLSUnSAL. Fig. 7 shows the denoising result of band 64 by different unmixing methods when it is degraded by Gaussian and impulse noise. We can see from Fig. 7 that only the proposed SUBM can effectively remove the Gaussian and impulse noise simultaneously, and the other compared methods can just remove part of Gaussian noise and impulse noise. Fig. 8 shows the denoising result of band 122 by different unmixing methods when it is degraded by Gaussian noise and



**Fig. 6.** Restoration results by different unmixing methods in the first synthetic data experiment: (a) Original band 20, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUnSAL, (f) CLSUnSAL and (g) SUBM.



**Fig. 7.** Restoration results by different unmixing methods in the first synthetic data experiment: (a) Original band 64, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUnSAL, (f) CLSUnSAL and (g) SUBM.



**Fig. 8.** Restoration results by different unmixing methods in the first synthetic data experiment: (a) Original band 122, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUnSAL, (f) CLSUnSAL and (g) SUBM.

dead lines. We can see from Fig. 8 that there still exist some Gaussian noise and dead lines in the denoised HSI obtained by the compared methods, and only our method can remove the Gaussian noise and dead lines simultaneously.

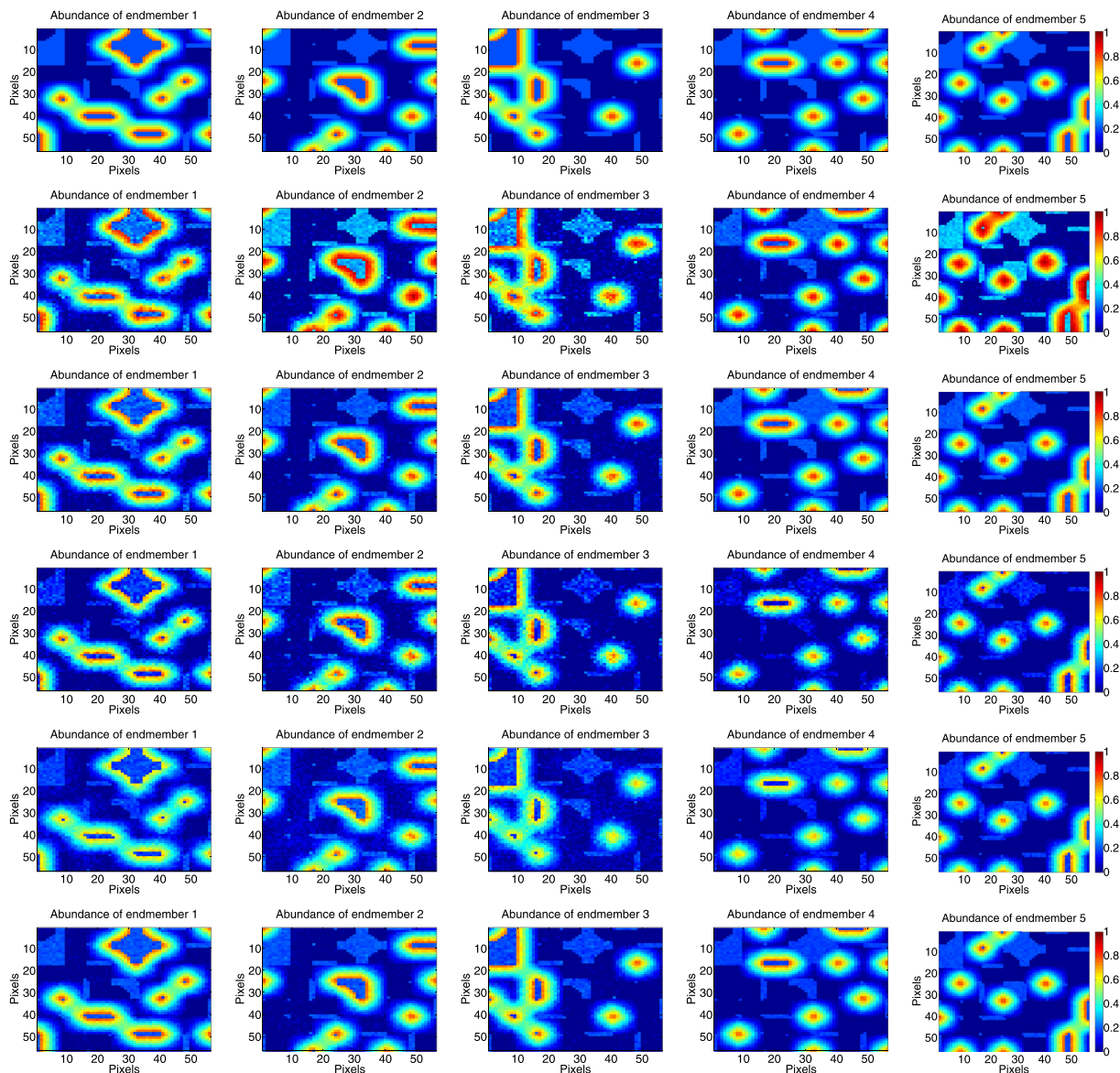
Fourth, we evaluate the influence the number of spectral signatures  $M$  on the final unmixing performance. Figs. 10 shows the RMSE of SUBM with the varying number of spectral signatures  $M$  in the first synthetic data experiment. It can be observed that the RMSE of SUBM generally decreases when the number of spectral signatures  $M$  increases, which is due to that the endmembers in the USGS spectral library are usually highly correlated.

In the second synthetic data experiment, we generate a synthetic HSI containing  $64 \times 64$  pixels and 224 bands. The abundances are generated according to Miao and Qi [49]<sup>3</sup>, which aim at generating synthetic HSI without pure regions. First, the whole HSI is segmented into  $8 \times 8$  sub-regions, each region is of size  $8 \times 8$ , and is randomly initialized by one endmember. Second, we generate the mixed pixels by using the  $9 \times 9$  low-pass filter. Third, we replace the abundances larger than 80% by an equal mixture of all endmembers, which guarantee that there are no pure pixels in the generated synthetic HSI. According to the bandwise unmixing model in (4), we also add noise to the HSI as follows:

1. Gaussian noise: we add Gaussian noise to each band of HSI with the signal-to-noise ratio (SNR) ranging from 20dB to 40 dB, and the mean SNR is 29.82 dB.
2. Impulse noise: we only degrade 11 bands (60–70) of HSI by 30% impulse noise.
3. Dead lines: we only degrade 11 bands (120–130) of HSI by dead lines.

First, we also conduct experiments to evaluate the sparse unmixing performance of the proposed SUBM and the compared methods in the presence of different types of noise. Fig. 9 presents the abundance maps estimated by different methods when only contaminated Gaussian noise. It can be clearly seen that our SUBM is more close to the ground truth than the compared methods. Table 4 shows sparse unmixing performance of the SUBM and compared methods when contaminated by three mixed noise. The proposed SUBM can get the best unmixing performance when only contaminated by Gaussian noise, which demonstrate the advantage of considering different level of Gaussian noise in different bands of HSI. When the HSI is contaminated by sparse noise, our SUBM can get much smaller RMSE than the compared methods, which demonstrate the advantage of considering the sparse noise. Besides, our SUBM and the CLSUnSAL can get smaller RMSE than SUnSAL, which demonstrate the advantage of taking the collaborative sparse property of abundances into consideration. Moreover, our SUBM can get much smaller RMSE than the compared methods when contaminated by three mixed noise, which also demonstrate the advantage of considering different level of Gaussian noise in different bands and the sparsity of sparse noise. Furthermore, the SUBM takes more time than the other compared methods.

<sup>3</sup> Available at: <https://bitbucket.org/aicp/mvcnmf>.



**Fig. 9.** Abundance maps obtained by different methods only in the presence of Gaussian noise for the second synthetic data experiment. From top to bottom: ground truth, SMP, RSFoBa, SUnSAL, CLSUnSAL and SUBM ( $\lambda = 10^{-1}$ ,  $\alpha = 10^{-1}$ ).

Second, we also assess the influence of the regularization parameters  $\lambda$ ,  $\alpha$  and Lagrange multiplier  $\mu$  in the second synthetic data experiment. Fig. 11 shows the RMSE of SUBM with the varying regularization parameter  $\mu$  for different  $\alpha$  and  $\lambda$  in the second data experiment. We can also see that the RMSE of SUBM remains stable when  $\mu$  is less than  $10^0$ , and the main reason lies in that in the iteration process, we update  $\mu$  to keep the ratio between the primal residual norm and dual residual norm within a positive interval. Thus, we also adopt  $\mu = 10^{-2}$  for SUBM. Fig. 12 shows RMSE of the proposed SUBM with the varying regularization parameter  $\lambda$  when contaminated by three mixed noise. We can see from Fig. 12 that the RMSE decreases when  $\lambda$  increases at first, and they get the best RMSE, which demonstrates the importance of considering the underlying collaborative sparsity of abundance. However, they begin to increase when  $\lambda$  gets to a certain extent. Fig. 13 presents RMSE of SUBM with the varying regularization parameter  $\alpha$  when contaminated by three mixed noise. We can see from Fig. 13 that the RMSE decreases when  $\alpha$  increases at first, and they obtain the best RMSE, which demonstrates the importance of considering the sparsity of sparse noise. However, they gradually increase and remain stable when  $\alpha$  gets to a certain extent, which is owing to that the best  $\alpha$  is highly related with the level of sparse noise in HSI.

Third, we also evaluate the denoising performance by different unmixing methods in the second synthetic data experiment. Fig. 14 presents the PSNR of each band using the SUBM and the compared methods in the presence of Gaussian noise,

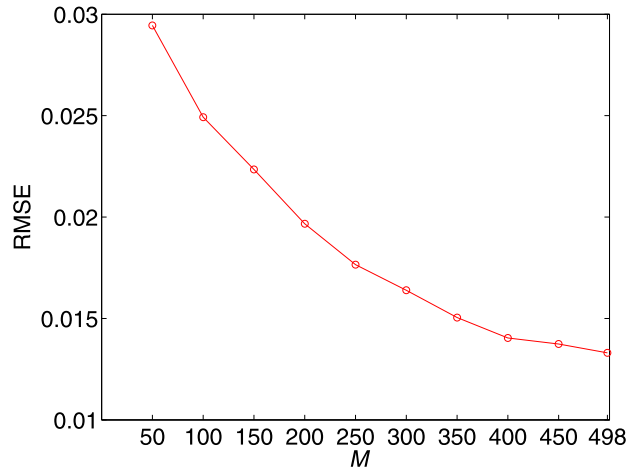


Fig. 10. RMSE of SUBM with the varying number of spectral signatures  $M$  in the first synthetic data experiment.

Table 4

Comparison of RMSEs ( $\times 10^{-2}$ ) and Time (s) of different methods under different noise condition for the second synthetic data experiment.

Type of noise		SMP	RSFoBa	SUNsAL	CLSUNsAL	SUBM
Gaussian noise	RMSE	1.021	0.408	1.003	0.781	<b>0.307</b>
	Time	<b>0.82</b>	1.77	4.00	3.98	54.09
Impulse noise	RMSE	1.204	0.792	1.922	1.741	<b>0.125</b>
	Time	<b>0.99</b>	1.82	3.98	3.96	53.96
Dead lines	RMSE	1.685	0.688	2.149	2.154	<b>0.118</b>
	Time	<b>1.33</b>	3.18	3.98	3.94	54.12
Gaussian noise & Impulse noise	RMSE	1.233	0.869	2.006	1.853	<b>0.322</b>
	Time	<b>1.23</b>	3.76	3.95	3.91	54.10
Gaussian noise & Dead lines	RMSE	1.222	0.736	2.214	2.204	<b>0.317</b>
	Time	<b>1.25</b>	3.23	3.95	3.92	54.08
Impulse noise & Dead lines	RMSE	1.343	0.985	2.536	2.491	<b>0.119</b>
	Time	<b>1.16</b>	3.31	3.94	3.85	53.74
Gaussian noise & Impulse noise & Dead lines	RMSE	1.274	1.014	2.593	2.503	<b>0.332</b>
	Time	<b>0.99</b>	2.79	3.96	3.88	54.02

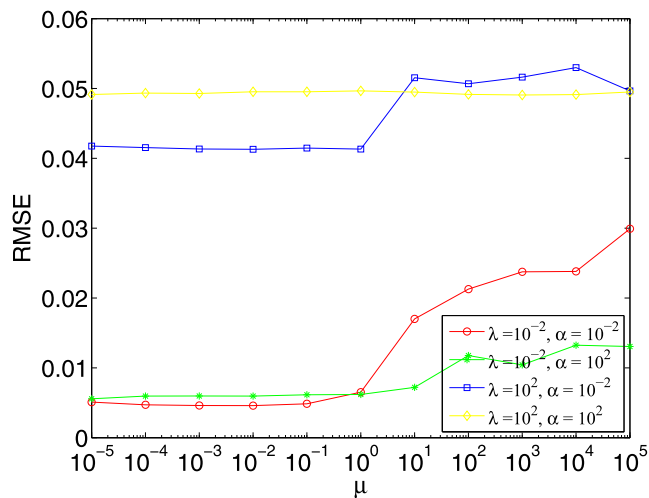


Fig. 11. RMSE of SUBM with the varying regularization parameter  $\mu$  for different  $\alpha$  and  $\lambda$  in the second synthetic data experiment.

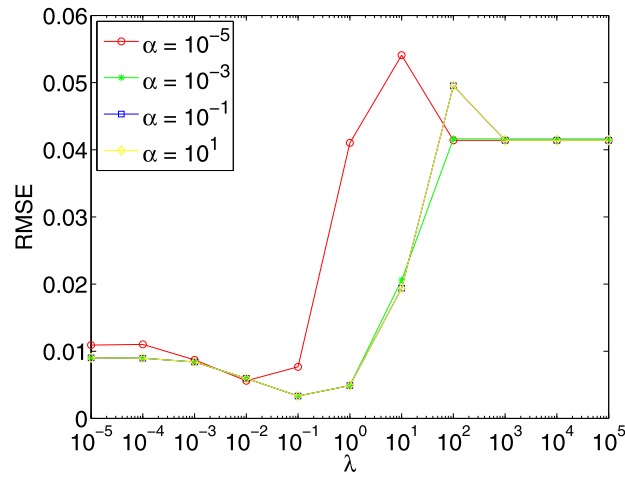


Fig. 12. RMSE of SUBM with the varying regularization parameter  $\lambda$  for different  $\alpha$  in the second synthetic data experiment.

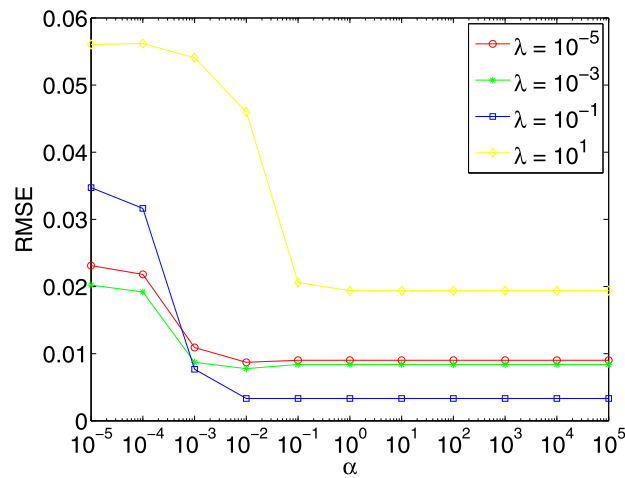


Fig. 13. RMSE of SUBM with the varying regularization parameter  $\alpha$  for different  $\lambda$  in the second synthetic data experiment.

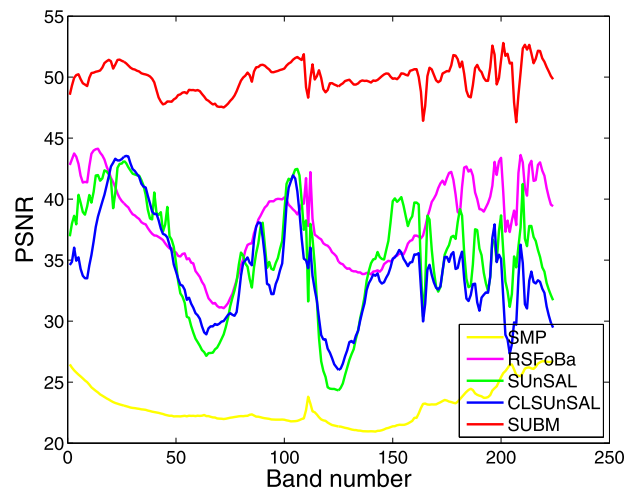


Fig. 14. PSNR of each band using the proposed SUBM and the compared methods in the second synthetic data experiment.



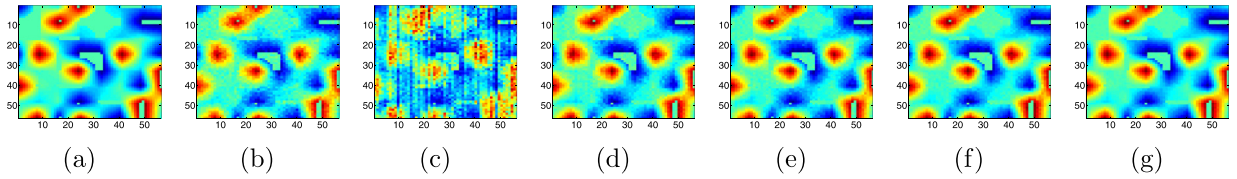


Fig. 15. Restoration results by different unmixing methods in the second synthetic data experiment: (a) Original band 30, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUNSAL, (f) CLSUnSAL and (g) SUBM.

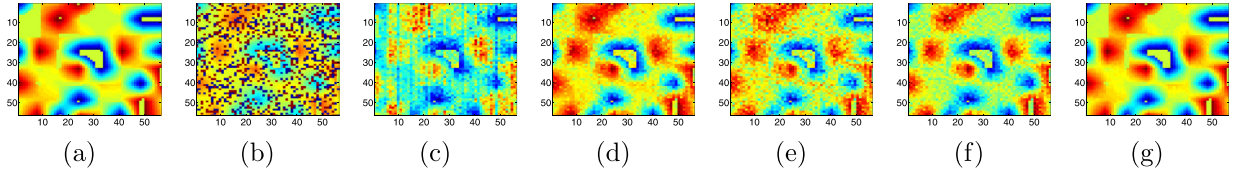


Fig. 16. Restoration results by different unmixing methods in the second synthetic data experiment: (a) Original band 65, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUNSAL, (f) CLSUnSAL and (g) SUBM.

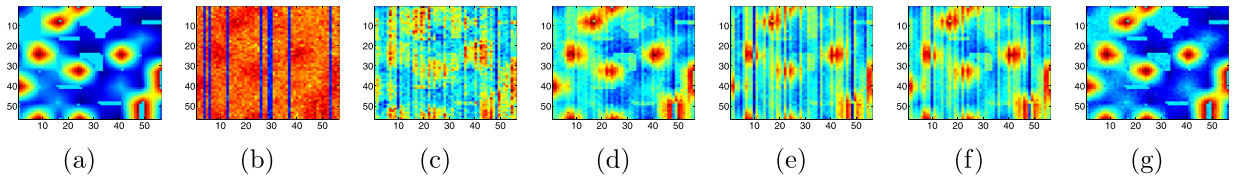


Fig. 17. Restoration results by different unmixing methods in the second synthetic data experiment: (a) Original band 125, (b) noisy band, (c) SMP, (d) RSFoBa, (e) SUNSAL, (f) CLSUnSAL and (g) SUBM.

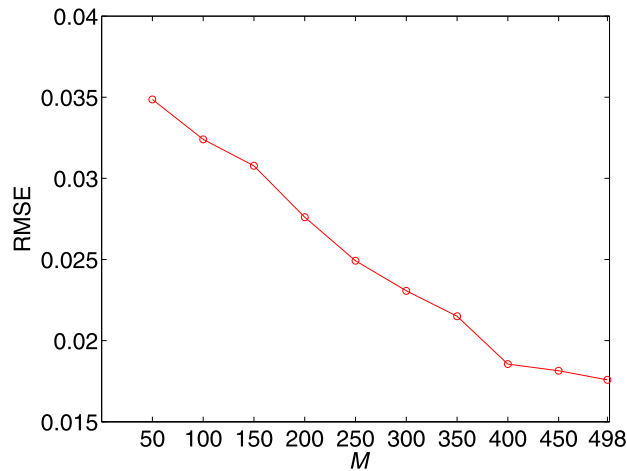


Fig. 18. RMSE of SUBM with the varying number of spectral signatures  $M$  in the first synthetic data experiment.

impulse noise and dead lines. We can see from Fig. 14 that our SUBM can get the best PSNR for all bands of the denoised HSI, which demonstrates that our SUBM can get the best denoising performance. Besides, Fig. 15 shows the denoising result of band 30 by different unmixing methods when it is only degraded by Gaussian noise. Fig. 16 presents the denoising result of band 65 by different unmixing methods when it is degraded by Gaussian and impulse noise. Fig. 17 shows the denoising result of band 125 by different unmixing methods. We can see from Figs. 15–17 that only the proposed SUBM can effectively remove the Gaussian noise, impulse noise and dead lines simultaneously.

Fourth, we also evaluate the influence the number of spectral signatures  $M$  on the final unmixing performance. Fig. 18 shows the RMSE of SUBM with the varying number of spectral signatures  $M$  in the second synthetic data experiment. We can also observe that the RMSE of SUBM generally decreases when the number of spectral signatures  $M$  increases, which is due to that the endmembers in the USGS spectral library are usually highly correlated.





Fig. 19. False-color image of Cuprite data set.

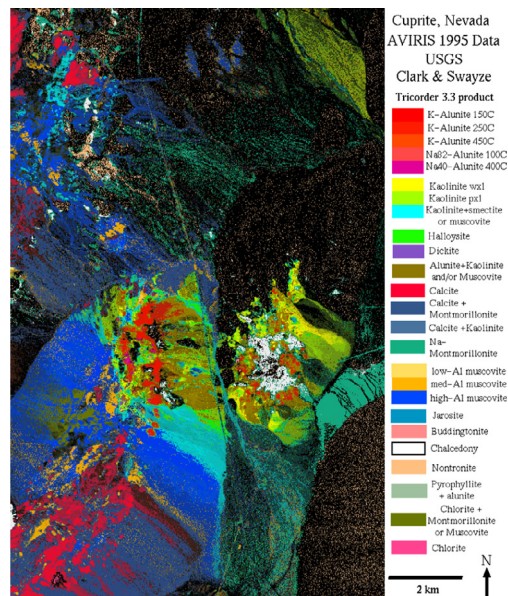


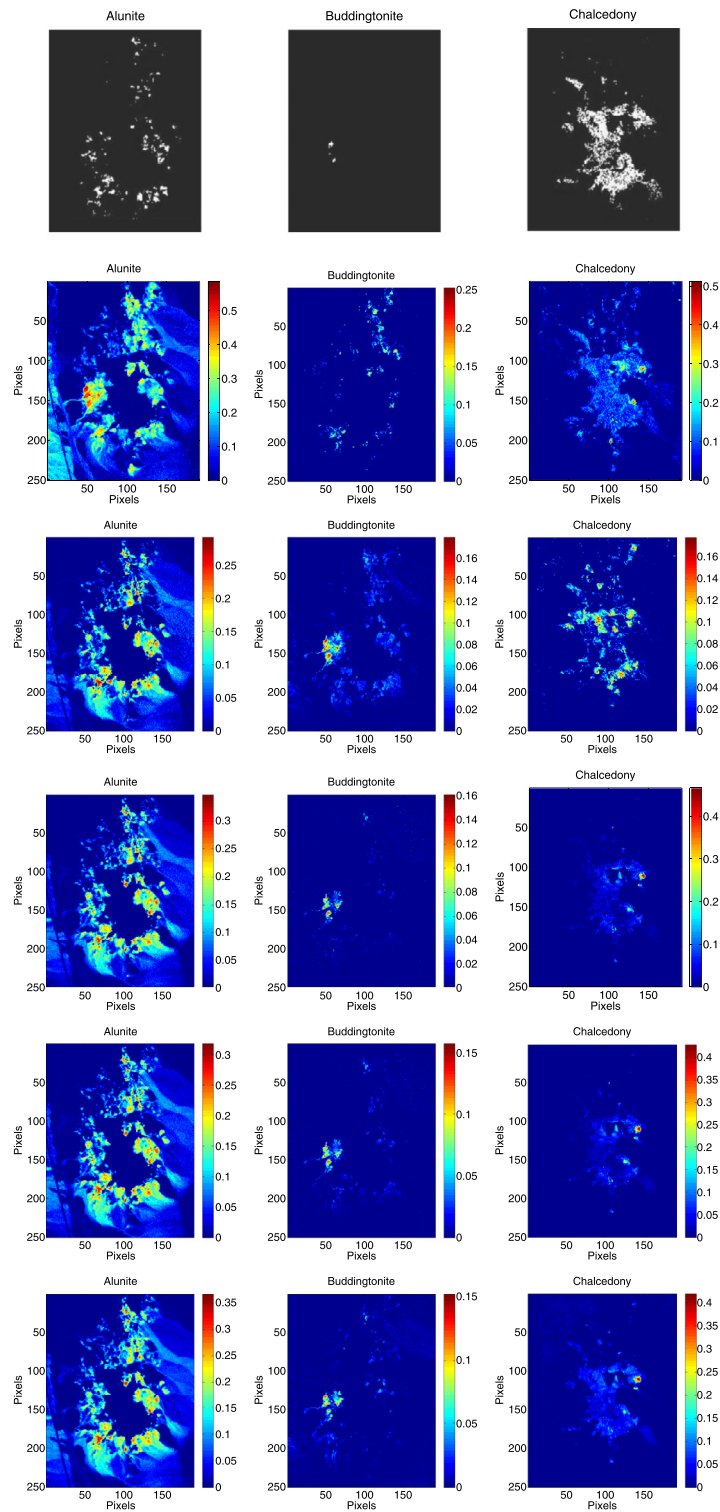
Fig. 20. USGS map of different minerals for the selected Cuprite data set.

To sum up, the proposed SUBM is superior to the compared methods, which can not only obtain the best unmixing performance but also remove the different level of Gaussian noise in different bands of HSI, impulse noise and dead lines simultaneously.

### 3.2. Experimental results with real data

In the real data experiment, we adopt the Cuprite data set, acquired by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), to test the sparse unmixing performance of all the methods. The region of interest used in this paper has  $250 \times 191$  pixels and 224 bands ranging from  $0.4\mu\text{m}$  to  $2.5\mu\text{m}$ , and the false color image can be observed in Fig. 19. We remove these bands (Bands 1–2, 104–113, 148–167 and 221–224) that have very low SNR and atmospheric absorption, and it remains 188 spectral bands.

Fig. 20 displays the mineral map obtained by the Tricorder 3.3 software product [50], which can be used for evaluation of the abundance obtained by our SUBM and the compared methods. The spectral library in the real data experiment is the same as that of synthetic data experiment. Fig. 21 presents the abundance maps of three typical endmembers obtained by our SUBM and the compared methods. We can see that the abundance maps obtained by our SUBM and the compared methods are different from the Tricorder maps to some extent, this is due to that sparse unmixing obtains the abundance maps denoting the fraction of the endmember in the pixel, while the Tricorder 3.3 software classifies each pixel to a specific



**Fig. 21.** Abundance maps of three representative endmembers obtained by the SUBM and compared methods. From top to bottom: USGS Tetracorder classification map, SMP, RSFoBa, SUNSAL, CLSUNSAL and SUBM.

**Table 5**  
Comparison of time (s) of different methods for the real CUPRITE data.

Method	SMP	RSFoBa	SUnSAL	CLSUnSAL	SUBM
Time	<b>20.77</b>	24.46	43.20	41.87	511.79

class. Besides, the average number of endmembers with abundances larger than 0.05 obtained by SMP, RSFoBa, SUnSAL, CLSUnSAL and SUBM, are 6.017, 4.870, 5.148, 5.038, 4.747 (per pixel), respectively, which indicates that the proposed SUBM can explain the real HSI with higher sparsity. Thus, our SUBM is valid for sparse unmixing of the real Cuprite HSI. Table 5 shows the time consumption of different methods for the real Cuprite data. We can see from Table 5 that our SUBM takes more time than the compared methods, and the reason is that the proposed SUBM needs to estimate the noise level of all bands of HSI, and it converges slower than SUnSAL and CLSUnSAL.

#### 4. Conclusion

In this paper, we propose a new bandwise unmixing model, and formulate each band of HSI as the linear combination of spectral library and the abundances in the presence of different types of noise. Besides, we propose the corresponding sparse unmixing method based on the bandwise unmixing model, which can take different level of Gaussian noise in different bands and the underlying sparsity of sparse noise into consideration under the maximum a posteriori framework. Moreover, the ADMM is adopted for solving our SUBM. Finally, experiments on both synthetic data set and real HSI demonstrate the advantages and effectiveness of the bandwise unmixing model and the corresponding unmixing method SUBM.

#### Declaration of Competing Interest

None.

#### Acknowledgments

Financial support for this study was provided by the National Natural Science Foundation of China under Grants 41901350, 61922075, 61773295, 61601397, 61501164, 61701160, and 81571760, the Fundamental Research Funds for the Central Universities under Grants JZ2019HGTA0049, JZ2019HGGBZ0151 and JZ2018HGTB0228, the Young Elite Scientists Sponsorship Program by CAST under Grant 2017QNRC001, the Provincial Natural Science Foundation of Anhui under Grant 1808085QF186, and the SenseTime Research Fund.

#### References

- [1] L. Li, C. Sun, L. Lin, J. Li, S. Jiang, J. Yin, A dual-kernel spectral-spatial classification approach for hyperspectral images based on mahalanobis distance metric learning, *Inf. Sci.* 429 (2018) 260–283.
- [2] Z. Shao, L. Zhang, Sparse dimensionality reduction of hyperspectral image based on semi-supervised local fisher discriminant analysis, *Int. J. Appl. Earth Obs. Geoinf.* 31 (2014) 122–129.
- [3] Z. Shao, L. Zhang, X. Zhou, L. Ding, A novel hierarchical semisupervised SVM for classification of hyperspectral images, *IEEE Geosci. Remote Sens. Lett.* 11 (9) (2014) 1609–1613.
- [4] C. Hong, J. Yu, J. Wan, D. Tao, M. Wang, Multimodal deep autoencoder for human pose recovery, *IEEE Trans. Image Process.* 24 (12) (2015) 5659–5670.
- [5] C. Hong, J. Yu, J. Zhang, X. Jin, K.-H. Lee, Multi-modal face pose estimation with multi-task manifold deep learning, *IEEE Trans. Ind. Inf.* 15 (7) (2019) 3952–3961.
- [6] J. Zhang, J. Yu, D. Tao, Local deep-feature alignment for unsupervised dimension reduction, *IEEE Trans. Image Process.* 27 (5) (2018) 2420–2432.
- [7] J. Yu, C. Zhu, J. Zhang, Q. Huang, D. Tao, Spatial pyramid-enhanced NetVLAD with weighted triplet loss for place recognition, *IEEE Trans. Neural Netw. Learn. Syst.* (2019).
- [8] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, J. Chanussot, Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 5 (2) (2012) 354–379.
- [9] C. Li, Y. Ma, J. Huang, X. Mei, C. Liu, J. Ma, Gbm-based unmixing of hyperspectral data using bound projected optimal gradient method, *IEEE Geosci. Remote Sens. Lett.* 13 (7) (2016) 952–956.
- [10] X. Mei, Y. Ma, C. Li, F. Fan, J. Huang, J. Ma, Robust GBM hyperspectral image unmixing with superpixel segmentation based low rank and sparse representation, *Neurocomputing* 275 (2018) 2783–2797.
- [11] R. Feng, L. Wang, Y. Zhong, Joint local block grouping with noise-adjusted principal component analysis for hyperspectral remote-sensing imagery sparse unmixing, *Remote Sens.* 11 (10) (2019) 1223.
- [12] K. Themelis, A.A. Rontogiannis, K. Koutroumbas, Semi-supervised hyperspectral unmixing via the weighted lasso, in: *ICASSP*, 2010, pp. 1194–1197.
- [13] J.M. Nascimento, J.M. Dias, Vertex component analysis: a fast algorithm to unmix hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 43 (4) (2005) 898–910.
- [14] K.E. Themelis, A.A. Rontogiannis, K.D. Koutroumbas, A novel hierarchical Bayesian approach for sparse semisupervised hyperspectral unmixing, *IEEE Trans. Signal Process.* 60 (2) (2012) 585–599.
- [15] S.F. Seyyedsalehi, H.R. Rabiee, A. Soltani-Farani, A. Zarezade, A probabilistic joint sparse regression model for semisupervised hyperspectral unmixing, *IEEE Geosci. Remote Sens. Lett.* 14 (5) (2017) 592–596.
- [16] P. Chen, J.D. Nelson, J.-Y. Tourneret, Toward a sparse bayesian markov random field approach to hyperspectral unmixing and classification, *IEEE Trans. Image Process.* 26 (1) (2017) 426–438.
- [17] M.-D. Iordache, J.M. Bioucas-Dias, A. Plaza, Sparse unmixing of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 49 (6) (2011) 2014–2039.
- [18] Z. Shi, W. Tang, Z. Duren, Z. Jiang, Subspace matching pursuit for sparse unmixing of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 52 (6) (2014) 3256–3274.

- [19] W. Tang, Z. Shi, Y. Wu, Regularized simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 52 (9) (2014) 5271–5288.
- [20] X. Fu, W.-K. Ma, T.-H. Chan, J. Bioucas-Dias, Self-dictionary sparse regression for hyperspectral unmixing: greedy pursuit and pure pixel search are related, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9 (6) (2015) 1128–1141.
- [21] J.M. Bioucas-Dias, M.A. Figueiredo, Alternating direction algorithms for constrained sparse regression: application to hyperspectral unmixing, in: *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, IEEE, 2010, pp. 1–4.
- [22] M.-D. Iordache, J.M. Bioucas-Dias, A. Plaza, Total variation spatial regularization for sparse hyperspectral unmixing, *IEEE Trans. Geosci. Remote Sens.* 50 (11) (2012) 4484–4502.
- [23] M.-D. Iordache, J.M. Bioucas-Dias, A. Plaza, Collaborative sparse regression for hyperspectral unmixing, *IEEE Trans. Geosci. Remote Sens.* 52 (1) (2014) 341–354.
- [24] C. Li, Y. Ma, X. Mei, C. Liu, J. Ma, Hyperspectral unmixing with robust collaborative sparse regression, *Remote Sens.* 8 (7) (2016) 588.
- [25] R. Feng, Y. Zhong, L. Zhang, Adaptive spatial regularization sparse unmixing strategy based on joint map for hyperspectral remote sensing imagery, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9 (12) (2016) 5791–5805.
- [26] Y. Ma, C. Li, X. Mei, C. Liu, J. Ma, Robust sparse hyperspectral unmixing with  $\ell_{2,1}$  norm, *IEEE Trans. Geosci. Remote Sens.* 55 (3) (2017) 1227–1239.
- [27] C. Li, Y. Ma, X. Mei, F. Fan, J. Huang, J. Ma, Sparse unmixing of hyperspectral data with noise level estimation, *Remote Sens.* 9 (11) (2017) 1166.
- [28] X. Li, J. Huang, L.-J. Deng, T.-Z. Huang, Bilateral filter based total variation regularization for sparse hyperspectral image unmixing, *Inf. Sci.* 504 (2019) 334–353.
- [29] F. Chen, Y. Zhang, Sparse hyperspectral unmixing based on constrained  $l_p$ - $l_2$  optimization, *IEEE Geosci. Remote Sens. Lett.* 10 (5) (2013) 1142–1146.
- [30] Y. Xu, F. Fang, G. Zhang, Similarity-guided and  $l_p$ -regularized sparse unmixing of hyperspectral data, *IEEE Geosci. Remote Sens. Lett.* 12 (11) (2015) 2311–2315.
- [31] W. Tang, Z. Shi, Z. Duren, Sparse hyperspectral unmixing using an approximate  $l_0$  norm, *Optik* 125 (1) (2014) 31–38.
- [32] Z. Shi, T. Shi, M. Zhou, X. Xu, Collaborative sparse hyperspectral unmixing using  $l_0$  norm, *IEEE Trans. Geosci. Remote Sens.* (2018).
- [33] T. Uezato, M. Fauvel, N. Dobigeon, Hyperspectral unmixing with spectral variability using adaptive bundles and double sparsity, *IEEE Trans. Geosci. Remote Sens.* 57 (6) (2019) 3980–3992.
- [34] X. Xu, Z. Shi, Multi-objective based spectral unmixing for hyperspectral images, *ISPRS J. Photogramm. Remote Sens.* 124 (2017) 54–69.
- [35] X. Jiang, M. Gong, H. Li, M. Zhang, J. Li, A two-phase multiobjective sparse unmixing approach for hyperspectral data, *IEEE Trans. Geosci. Remote Sens.* 56 (1) (2018) 508–523.
- [36] X. Xu, Z. Shi, B. Pan,  $l_0$ -based sparse hyperspectral unmixing using spectral information and a multi-objectives formulation, *ISPRS J. Photogramm. Remote Sens.* 141 (2018) 46–58.
- [37] X. Xu, Z. Shi, B. Pan, X. Li, A classification-based model for multi-objective hyperspectral sparse unmixing, *IEEE Trans. Geosci. Remote Sens.* (2019).
- [38] C. Rogass, C. Mielke, D. Scheffler, N.K. Boesche, A. Lausch, C. Lubitz, M. Brell, D. Spengler, A. Eisele, K. Segl, et al., Reduction of uncorrelated striping noise applications for hyperspectral pushbroom acquisitions, *Remote Sens.* 6 (11) (2014) 11082–11106.
- [39] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, *IEEE Trans. Geosci. Remote Sens.* 52 (8) (2014) 4729–4743.
- [40] A. Majumdar, N. Ansari, H. Aggarwal, P. Biyani, Impulse denoising for hyper-spectral images: a blind compressed sensing approach, *Signal Process.* 119 (2016) 136–141.
- [41] M. Vidal, J.M. Amigo, Pre-processing of hyperspectral images. essential steps before image analysis, *Chemom. Intell. Lab. Syst.* 117 (2012) 138–148.
- [42] F. Fan, Y. Ma, C. Li, X. Mei, J. Huang, J. Ma, Hyperspectral image denoising with superpixel segmentation and low-rank representation, *Inf. Sci.* 397 (2017) 48–68.
- [43] B. Du, Z. Huang, N. Wang, A bandwise noise model combined with low-rank matrix factorization for hyperspectral image denoising, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 11 (4) (2018) 1070–1081.
- [44] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Found. Trends® Mach. Learn.* 3 (1) (2011) 1–122.
- [45] D.L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inf. Theory* 41 (3) (1995) 613–627.
- [46] S.J. Wright, R.D. Nowak, M.A. Figueiredo, Sparse reconstruction by separable approximation, *IEEE Trans. Signal Process.* 57 (7) (2009) 2479–2493.
- [47] J.M. Bioucas-Dias, J.M. Nascimento, Hyperspectral subspace identification, *IEEE Trans. Geosci. Remote Sens.* 46 (8) (2008) 2435–2445.
- [48] N. Yokoya, J. Chanussot, A. Iwasaki, Nonlinear unmixing of hyperspectral data using semi-nonnegative matrix factorization, *IEEE Trans. Geosci. Remote Sens.* 52 (2) (2014) 1430–1437.
- [49] L. Miao, H. Qi, Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization, *IEEE Trans. Geosci. Remote Sens.* 45 (3) (2007) 765–777.
- [50] R.N. Clark, G.A. Swayze, K.E. Livo, R.F. Kokaly, S.J. Sutley, J.B. Dalton, R.R. McDougal, C.A. Gent, Imaging spectroscopy: earth and planetary remote sensing with the usgs tetracorder and expert systems, *J. Geophys. Res.* 108 (E12) (2003).