

# Sparsity-Regularized Robust Non-Negative Matrix Factorization for Hyperspectral Unmixing

Wei He, *Student Member, IEEE*, Hongyan Zhang, *Member, IEEE*, and Liangpei Zhang, *Senior Member, IEEE*

**Abstract**—Hyperspectral unmixing (HU) is one of the crucial steps for many hyperspectral applications, including material classification and recognition. In the last decade, non-negative matrix factorization (NMF) and its extensions have been widely studied and have achieved advanced performances in HU. Unfortunately, most of the existing NMF-based methods make the assumption that the hyperspectral data are only corrupted by Gaussian noise. In real applications, the hyperspectral data are inevitably corrupted by sparse noise, which includes impulse noise, stripes, deadlines, and others types of noise. By separately modeling the sparse noise and Gaussian noise, a robust NMF (RNMF) model is subsequently introduced to unmix the hyperspectral data. The proposed RNMF model is able to simultaneously handle Gaussian noise and sparse noise, and can be efficiently learned with elegant update rules. In addition, sparsity regularizers are added to restrict the abundance maps in the RNMF, with the consideration of the sparse property of the material types within the hyperspectral scene. The experimental results with simulated and real data confirm the superiority of the proposed sparsity-regularized RNMF methods compared to the traditional NMF methods.

**Index Terms**—Hyperspectral unmixing (HU), robust non-negative matrix factorization (RNMF), sparse noise, sparsity regularizer.

## I. INTRODUCTION

IN recent years, with the wealth of available spectral information, HSIs have been widely used in practical applications such as food safety, pharmaceutical process monitoring and quality control, biomedical applications, and so on [1], [2]. However, due to the low spatial resolution and the complexity of the terrain, a single pixel in an HSI sometimes contains different materials, and is called a “mixed pixel.” The existence of mixed pixels introduces a challenge to HSI analysis because the underlying assumption governing the clustering and classification tasks is that each pixel vector comprises the response of a single underlying material [3], [4]. Thus, in order to make full use of the data, hyperspectral unmixing (HU) is an important step for HSI analysis.

The task of HU is to decompose a mixed pixel into a collection of constituent materials (also called endmembers) and

Manuscript received July 11, 2015; revised October 11, 2015 and December 21, 2015; accepted January 14, 2016. Date of publication February 10, 2016; date of current version September 30, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 41571362, Grant 61201342, and Grant 41431175, and in part by the Fundamental Research Funds for Central Universities under Grant 2015904020202. (*Corresponding author: Hongyan Zhang.*)

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, and the Collaborative Innovation Center for Geospatial Technology, Wuhan University, Wuhan 430072, China (e-mail: zhanghongyan@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2016.2519498

their relative proportions (also called abundances) [1], [3]. HU algorithms mainly rely on the expected type of mixing, which can be characterized as either a linear or nonlinear model [1]. The nonlinear unmixing model describes the mixed spectrum by assuming that the observed pixel is generated from a nonlinear function of the abundance associated with the endmembers [5]. The linear mixing model (LMM) assumes that the different endmembers do not interfere with each other [6] and lies at the center of interest of many important developments in HU [1], [4], [7]–[11]. Despite the fact that the LMM is not always true, especially under certain scenarios that exhibit strong nonlinearity, it is generally recognized as an acceptable model for many real-world scenarios [12]. Based on the LMM, many different HU strategies have been proposed, e.g., endmember determination plus inversion [1], [13]–[16], dictionary-based sparse regression [4], [17], [18], and statistical approaches [10]. In this paper, we focus on the non-negative matrix factorization (NMF) [19] based approaches to HU.

NMF attempts to decompose a high-dimensional dataset into two non-negative matrices: one consisting of “basis vectors” and the other containing “coefficient vectors” [10], [20]. From the data analysis point of view, NMF is very attractive because it usually provides a part-based representation of the data, making the decomposition matrices more intuitive and interpretable. NMF has drawn a lot attention in HU, as it does not require the pure pixel assumption and can simultaneously determine endmember spectra and the corresponding abundances. Unfortunately, the solution space of NMF is very large if no further constraints are taken into consideration. In addition, due to the nonconvexity of the objective function of NMF, the algorithm may fall into local minima. To alleviate this situation, the abundance sum-to-one constraint (ASC), which is the basis of the LMM, was first added to constrain the solution space. To further shrink the solution space, additional constraints have also been imposed upon the abundances [8], [9], [21], [22], as well as upon the endmembers [6], [10]. We give a brief review of these methods below.

By employing the variance of the spectral matrix to constrain the recovered spectra to be flat and preserve the possible spectral singularities, Huck *et al.* [23] proposed minimum dispersion constrained NMF (MiniDisCo) to unmix HSIs. By combining the **simplex volume minimization** and NMF models, Miao and Qi [6] proposed the **minimum volume constrained NMF** (MVCNMF) approach for HU. Wang *et al.* [10] proposed the endmember dissimilarity constrained NMF (EDCNMF) method, which assumes that the endmember signal should itself be smooth. Jia and Qian [7] proposed a piecewise smoothness NMF approach with sparseness constraint (PSnsNMF), which

imposes both piecewise smoothness and sparseness constraints on the NMF. Another approach named abundance separation and abundance smoothness constrained NMF (ASSNMF) [22] has been proposed by adding an abundance separation constraint and an abundance smoothness constraint into the basic NMF. In recent years, sparsity-based methods have also been adopted for NMF HU. They assume that, in HSIs, most pixels are mixed by a subset of endmembers, rather than all the endmembers, and thus employ various kinds of sparse constraints on the abundances [4], [24]. The  $L_1$  regularizer was first adopted since it yields sparse solutions. However, for spectral unmixing, the  $L_1$  regularizer does not enforce the full additivity constraint. The S-measure constraint was used to measure the sparsity of the abundance in [21], and achieved better results than the  $L_1$  regularizer. The authors in [8] also explored the use of the  $L_{1/2}$  regularizer, which is an alternative to its  $L_1$  counterpart. The  $L_{1/2}$  regularizer has been theoretically found to be a sparsity-promoting function. Furthermore, the  $L_{1/2}$  regularizer can not only provide sparse solutions which are closer to the truth than those yielded when  $L_1$  is used, but is also computationally efficient.

Although the NMF-related methods have achieved commendable performances, they are far from the optimal solution. Most of the existing LMM-based methods are on the basis that the hyperspectral data are only corrupted by Gaussian noise. However, real HSI data are inevitably corrupted by sparse noise [25]. Sparse noise, which includes impulse noise, deadlines, and stripes, is defined as noise of arbitrary magnitude that contaminates certain specific bands or pixels [26]. The existence of sparse noise seriously reduces the effectiveness of the NMF-related HU methods since the objective functions of these methods are easily dominated by the sparse noise, leading to a failure of the material decomposition. From another aspect, in most real HSI unmixing experiments, the bands with high noise levels are often discarded by manual selection, and only the high signal-to-noise ratio (SNR) bands are utilized for the material decomposition in the HSI unmixing procedure. This processing strategy requires prior information about the noise distribution in the bands of the HSI to be unmixed. In addition, it raises a question: can these high-noise bands actually provide helpful information for the unmixing of the HSI data?

To alleviate the negative influence of the sparse noise existing in HSI data, we propose a new method named sparsity-regularized robust NMF (RNMF) for the HU task. In the proposed method, the corruptions in the data can take arbitrary values, but are assumed to be sparse. Specifically, we introduce a sparse matrix to explicitly capture the sparse corruption. This strategy can significantly improve the robustness of NMF with respect to sparse noise. In addition, sparsity regularizers are adopted to enforce the sparsity of the material abundances. The main contributions of this paper can be summarized as follows.

- 1) The RNMF model is introduced to unmix hyperspectral data considering sparse corruptions. A sparse matrix is used to model the sparse noise in the NMF objective function. As the support of the sparse matrix usually exists in some of the pixels of some of the bands, we use the  $L_{1,2}$  norm to restrict the sparse noise.
- 2) We propose a new coupled blind HU and mixed noised removal method, and attempt to simultaneously unmix

the low-noise and noisy bands of an HSI in real data experiments. We investigate the assumption that with the appropriate modeling of the noise, the noisy bands can also provide useful information for the HU task.

This paper is organized as follows. After introducing the related work in Section II, the proposed sparsity-regularized RNMF is described in Section III. In Section IV, both simulated and real data experiments are described and analyzed, and the conclusions are drawn in Section V.

## II. BACKGROUND

### A. Linear Mixing Model

The classical LMM assumes that a pixel in a hyperspectral dataset is a linear mixture of  $K$  known material signatures, called endmembers:  $\mathbf{A} := [a_1, a_2, \dots, a_K]$ , in which  $a_i \in \mathbf{R}^{L \times 1}$  is the spectral signature of the  $i$ th endmember. The corresponding proportion is called the abundance and is denoted as  $s \in \mathbf{R}^{K \times 1}$ . Based on the LMM, each pixel in an HSI dataset can be expressed as

$$x = \mathbf{A}s + n \quad (1)$$

where  $x \in \mathbf{R}^{L \times 1}$  is a signature vector corresponding to one pixel in the HSI and  $n$  stands for the noise, which is usually assumed to obey a Gaussian distribution. Adopting a matrix notation, an HSI with  $L$  bands can be described by a matrix  $\mathbf{X} \in \mathbf{R}^{L \times N}$ , where  $N$  is the total number of pixels. The LMM can then be described as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (2)$$

where  $\mathbf{S} := [s_1, s_2, \dots, s_N] \in \mathbf{R}^{K \times N}$  denotes the relative end-member abundances and  $\mathbf{N}$  is the noise. In general, two constraints—the *abundance non-negative constraint* (ANC) and the *abundance sum-to-one constraint* (ASC)—are added to restrict the LMM model, and can be explicitly given by

$$\mathbf{S} \geq 0 \quad (3)$$

$$\mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T \quad (4)$$

in which  $\mathbf{1}_K^T$  and  $\mathbf{1}_N^T$  represent all-one vectors with size  $K$  and size  $N$ , respectively.

### B. Non-Negative Matrix Factorization

NMF has received considerable attention in the field of HU due to its many advantages. The LMM assumes that the HSI consists of spectral signatures of endmembers with corresponding non-negative abundances. Therefore, the non-negativity of  $\mathbf{A}$  and  $\mathbf{S}$  mentioned earlier is a natural property of the measured quantities in the hyperspectral data. In most NMF-related works, the squared Euclidean distance is adopted as the cost function, which is built for independent identically distributed Gaussian noise. The objective function of the NMF method is presented as follows:

$$\min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2, \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0 \quad (5)$$

where  $\|\mathbf{g}\|_F$  is the Frobenius norm of the matrix.

The solution for the NMF model has been widely studied in the past few years [27]–[29]. Although the above minimization problem is separately convex in  $\mathbf{A}$  and  $\mathbf{S}$ , it is not simultaneously convex with respect to both of the matrices. In this paper, a multiplicative iterative algorithm is applied to solve the NMF model, which minimizes a multivariate objective function by dividing the parameters into two sets and adopting a dual-step process. In the first step, a subset of the parameters is updated while the others remain fixed. The second step proceeds conversely by fixing the newly updated parameters while estimating the solution of the second subset.

The update rule and the convergence were first investigated by Lee and Seung [27]. When applied to (5), this multiplicative update becomes

$$\mathbf{A} \leftarrow \mathbf{A} * \mathbf{X}\mathbf{S}^T ./ \mathbf{A}\mathbf{S}\mathbf{S}^T \quad (6a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T \mathbf{X} ./ \mathbf{A}^T \mathbf{A}\mathbf{S} \quad (6b)$$

in which  $(\cdot)^T$  denotes the transpose of the matrix, and  $*$  and  $./$  denote the element-wise multiplication and division, respectively.

### C. NMF With a Sparse Regularizer

Since the objective function is nonconvex, a large number of minima occur during unmixing, and it is clear that the solution of the function may not be unique, which leads to instability in the unmixing. Therefore, more constraints should be added to the traditional NMF. As shown in [8], sparsity of abundance is an intrinsic property of hyperspectral data. For this reason, NMF with a sparsity constraint is considered as the objective function for the minimization problem, which was introduced in [8] and [9]. The objective function can be shown as follows:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{S}} f(\mathbf{A}, \mathbf{S}) &= \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \gamma g(\mathbf{S}) \\ \text{s.t. } \mathbf{A} &\geq 0, \mathbf{S} \geq 0 \end{aligned} \quad (7)$$

where  $\gamma \geq 0$  is the parameter used to control the contribution of the sparsity measure function  $g(\cdot)$  of the matrix  $\mathbf{S}$ , which is regarded as the regularization term.

In this paper, we introduce two kinds of sparsity regularizers:  $L_1$  and  $L_{1/2}$  regularizers. The corresponding  $L_1$ -NMF is given as follows:

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{S}} f(\mathbf{A}, \mathbf{S}) &= \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \gamma \|\mathbf{S}\|_1 \\ \text{s.t. } \mathbf{A} &\geq 0, \mathbf{S} \geq 0. \end{aligned} \quad (8)$$

The  $L_{1/2}$ -NMF is then written as

$$\begin{aligned} \min_{\mathbf{A}, \mathbf{S}} f(\mathbf{A}, \mathbf{S}) &= \frac{1}{2} \|\mathbf{X} - \mathbf{A}\mathbf{S}\|_F^2 + \gamma \|\mathbf{S}\|_{1/2} \\ \text{s.t. } \mathbf{A} &\geq 0, \mathbf{S} \geq 0 \end{aligned} \quad (9)$$

in which  $\|\mathbf{S}\|_{1/2} := \sum_{i,j}^{L,N} (\mathbf{S}_{i,j})^{1/2}$  and  $\mathbf{S}_{i,j}$  is the abundance fraction for the  $i$ th endmember at the  $j$ th pixel in the HSI.

A multiplicative iterative algorithm can also be used in sparsity-regularized NMF-based HU [8]. For the  $L_1$ -NMF model (8), the update rule developed by Hoyer [30] is presented as follows:

$$\mathbf{A} \leftarrow \mathbf{A} * \mathbf{X}\mathbf{S}^T ./ \mathbf{A}\mathbf{S}\mathbf{S}^T \quad (10a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T \mathbf{X} ./ (\mathbf{A}^T \mathbf{A}\mathbf{S} + \gamma). \quad (10b)$$

As for the  $L_{1/2}$ -NMF optimization (9), the update rule is [8]

$$\mathbf{A} \leftarrow \mathbf{A} * \mathbf{X}\mathbf{S}^T ./ \mathbf{A}\mathbf{S}\mathbf{S}^T \quad (11a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T \mathbf{X} ./ \left( \mathbf{A}^T \mathbf{A}\mathbf{S} + \frac{\gamma}{2} \mathbf{S}^{-1/2} \right). \quad (11b)$$

Likewise,  $\mathbf{S}^{-1/2}$  is given by the inverse of the element-wise square root of each entry in the matrix  $\mathbf{S}$ . A very small value is also added to  $\mathbf{S}$  to avoid a trivial solution.

As we can see, since the optimization of sparsity-regularized NMF is not globally convex, the solution of (9) [and (8)] is sensitive to noise. That is to say, the existence of sparse noise of arbitrary magnitude can significantly degrade the performance of sparse NMF in HU. Hence, further efforts should be made to improve the resistance of the HU method to various kinds of noise.

## III. SPARSITY-REGULARIZED RNMF

### A. Extended LMM Model

The classical LMM is based on the assumption that the model is only corrupted by Gaussian noise. Unfortunately, real HSI data are also inevitably corrupted by sparse noise [26], which is defined as noise of arbitrary magnitude that contaminates certain specific bands or pixels. Sparse noise occurs for the following reasons. First, due to the high spectral resolution of hyperspectral sensors, the energy used to produce each hyperspectral band, which is partitioned from a narrow wavelength, is limited. As a result, the imaging information can be easily overwhelmed by various kinds of noise. In addition, the poor imaging conditions are also responsible for the degraded bands, due to water vapor and atmospheric effects [31]. Fig. 1 shows some bands of the Hyperspectral Digital Imagery Collection Experiment (HYDICE) ‘‘Urban’’ dataset, which include low-noise bands, noisy bands, and water-absorption bands. Second, even in the low-noise bands of the HSI, there can often be sparse and large digital number (DN) variations in some pixels of the bands with high SNR, due to the sudden and unexpected change of the spectral range of the illumination source or corruption of the electronic charge of the sensor [32]. For example, we present the spectral curves of three pixels which belong to the same material in Fig. 2. From Fig. 2(b), it can be observed that the signatures of pixels 1 and 3 have rapid fluctuations in the low-noise bands. To some extent, these two pixels can be regarded as outliers. In addition, as these rapid fluctuations only exist in some specific bands, we can model these rapid fluctuations as sparse noise.

Based on these assumptions, the LMM can be extended to consider sparse noise [25]

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{E} + \mathbf{N} \quad (12)$$

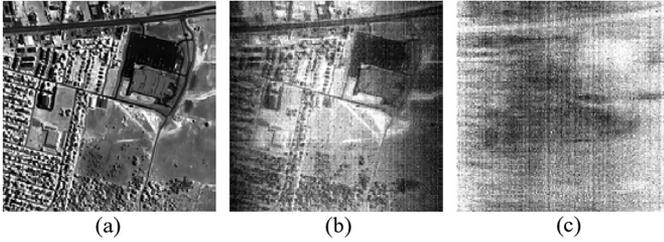


Fig. 1. Some bands of the HYDICE Urban dataset. (a) Low-noise band. (b) Noisy band. (c) Water-absorption band.

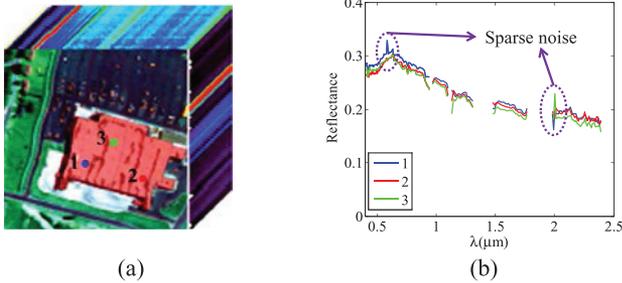


Fig. 2. Representation of sparse noise in HSI data. (a) Subset of the HYDICE Urban data, with noisy and water-absorption bands removed. The three pixels pointed out in the image belong to the same material. (b) Spectral curves of the three pixels. The rapid fluctuations in the signatures are assumed to be sparse noise.

where  $\mathbf{E}$  represents the sparse noise, which is assumed to corrupt certain specific bands. The ANC and ASC constraints are also taken into consideration.

Compared to the LMM, the extended LMM splits the noise into two parts: sparse noise and Gaussian noise. In the extended LMM, we respectively model the sparse noise and Gaussian noise, which can significantly improve the results in a real HSI unmixing procedure. In the next part, we introduce the RNMF model to unmix HSIs, taking this extended LMM into consideration.

### B. Sparsity-Regularized RNMF Model

The two widely adopted cost functions of NMF (the squared Euclidean distance and the generalized Kullback–Leibler divergence) are optimal for Gaussian noise and Poisson noise, respectively [32]. However, in the extended LMM, the sparse errors in the data may be arbitrarily large, as shown in Figs. 1 and 2. The traditional NMF may break down under this case, since the error assumptions are apparently violated.

Inspired by the recent work in robust principal component analysis (RPCA) [33], a novel algorithm named RNMF [34] was proposed to handle a case with gross errors, and has also been introduced to treat the nonlinear effect in HSI data [35]–[37]. On the basis of the extended LMM, we assume that some entries of the data matrix may be arbitrarily corrupted, but the corruption is sparse, and is mainly concentrated on some specific bands of the data. To obtain  $\mathbf{A}$  and  $\mathbf{S}$ , RNMF can be performed by minimizing the difference between  $(\mathbf{X} - \mathbf{E})$  and  $\mathbf{AS}$ , and enforcing non-negativity on  $\mathbf{A}$  and  $\mathbf{S}$ . We also assume that the noise matrix  $\mathbf{N}$  is Gaussian distributed, and use the

Euclidean distance [8] to measure the difference between the matrices. The loss function for RNMF is given by

$$f(\mathbf{A}, \mathbf{S}, \mathbf{E}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2. \quad (13)$$

Due to the presence of  $\mathbf{E}$ ,  $\mathbf{A}$ , and  $\mathbf{S}$  are protected from the sparse noise corruption, contributing to more robustness than the traditional NMF. Let  $\|g\|_{0,2}$  be the matrix  $L_{0,2}$  norm, which counts the number of nonzero rows in its arguments. As the sparse noise  $\mathbf{E}$  only exists in a few of the specific bands, we add the  $L_{0,2}$  norm constraint to  $\mathbf{E}$  to enforce the row sparsity of the sparse noise. Finally, the optimization problem of the RNMF model is

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0, \|\mathbf{E}\|_{0,2} \leq r \quad (14)$$

in which  $r$  is the parameter that specifies the maximum number of nonzero rows in  $\mathbf{E}$ . Since the  $L_{0,2}$  norm is often difficult to solve, we relax the  $L_{0,2}$  norm with the  $L_{1,2}$  norm. The optimization problem (14) can then be reformulated as

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{E}\|_{1,2} \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0 \quad (15)$$

where  $\|\mathbf{E}\|_{1,2} := \sum_{i=1}^L \|\mathbf{E}_{i,:}\|_2$  and  $\lambda \geq 0$  is the regularization parameter, which controls the row sparsity of  $\mathbf{S}$ .

As with NMF, we also consider RNMF with a sparsity constraint as the objective function for our minimization problem, which is the same strategy as that used in [8] and [9]. The objective function can be shown as follows:

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} f(\mathbf{A}, \mathbf{S}, \mathbf{E}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{E}\|_{1,2} + \gamma g(\mathbf{S}) \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0 \quad (16)$$

where  $\gamma \geq 0$  is the parameter used to control the contribution of the sparsity measure function  $g(\cdot)$  of the matrix  $\mathbf{S}$ , which is regarded as the regularization term.

In this paper, we adopt RNMF combined with  $L_1$  and  $L_{1/2}$  regularizers, respectively, to fully demonstrate the performance of the RNMF model for sparse noise processing. The corresponding  $L_1$ -RNMF is given as follows:

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} f(\mathbf{A}, \mathbf{S}, \mathbf{E}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{E}\|_{1,2} + \gamma \|\mathbf{S}\|_1 \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0. \quad (17)$$

The  $L_{1/2}$ -RNMF is then written as

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} f(\mathbf{A}, \mathbf{S}, \mathbf{E}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{E}\|_{1,2} + \gamma \|\mathbf{S}\|_{1/2} \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0. \quad (18)$$

In the next part, we demonstrate the expandability of the NMF model to the RNMF model from the aspect of optimization, convergence, and implementation issues, for the cases of  $L_1$ -RNMF and  $L_{1/2}$ -RNMF.

### C. Update Rules for Sparsity-Regularized RNMF-Based HU

In this part, we extend the multiplicative iterative algorithm for sparsity-regularized RNMF.

If we regard  $\mathbf{X} - \mathbf{E}$  as one variable, then the update rule of variables  $\mathbf{A}$  and  $\mathbf{S}$  for *sparsity-regularized* RNMF (16) is the same as that for *sparsity-regularized* NMF (7). In addition, the update of  $\mathbf{E}$  can be recast as an  $L_{1,2}$  norm based optimization problem, and we can adopt a *soft* threshold [38] to solve it. We describe this *soft* threshold as Lemma 1 below.

*Lemma 1:* [38]: Let  $\mathbf{Q}$  be a given matrix. If the optimal solution to

$$\min_{\mathbf{W}} \lambda \|\mathbf{W}\|_{1,2} + \frac{1}{2} \|\mathbf{W} - \mathbf{Q}\|_F^2$$

is  $\mathbf{W}^* = \text{soft}_\lambda(\mathbf{Q})$ , then the  $i$ th row of  $\mathbf{W}^*$  is

$$[\mathbf{W}^*]_{i,:} = \begin{cases} \frac{\|\mathbf{Q}\|_{i,:} - \lambda}{\|\mathbf{Q}\|_{i,:}} [\mathbf{Q}]_{i,:}, & \text{if } \|\mathbf{Q}\|_{i,:} \geq \lambda \\ 0, & \text{otherwise.} \end{cases}$$

By combining (10) and Lemma 1, the update rule for  $L_1$ -RNMF becomes

$$\mathbf{A} \leftarrow \mathbf{A} * (\mathbf{X} - \mathbf{E}) \mathbf{S}^T ./ \mathbf{A} \mathbf{S} \mathbf{S}^T \quad (19a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T (\mathbf{X} - \mathbf{E}) ./ (\mathbf{A}^T \mathbf{A} \mathbf{S} + \gamma) \quad (19b)$$

$$\mathbf{E} \leftarrow \text{soft}_\lambda(\mathbf{X} - \mathbf{A} \mathbf{S}). \quad (19c)$$

Analogously, we can also easily extend the update rule of  $L_{1/2}$ -NMF to  $L_{1/2}$ -RNMF by introducing the soft threshold. We present it as

$$\mathbf{A} \leftarrow \mathbf{A} * (\mathbf{X} - \mathbf{E}) \mathbf{S}^T ./ \mathbf{A} \mathbf{S} \mathbf{S}^T \quad (20a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T (\mathbf{X} - \mathbf{E}) ./ \left( \mathbf{A}^T \mathbf{A} \mathbf{S} + \frac{\gamma}{2} \mathbf{S}^{-1/2} \right) \quad (20b)$$

$$\mathbf{E} \leftarrow \text{soft}_\lambda(\mathbf{X} - \mathbf{A} \mathbf{S}). \quad (20c)$$

There is no doubt that the update rules of the RNMF-based methods have a lot of similar properties to those of the NMF-based methods, due to the high correlation of the two approaches. Next, we take  $L_{1/2}$ -RNMF as an example to investigate the convergence of the proposed update rule.

### D. Convergence of $L_{1/2}$ -RNMF

To ensure the reliability of RNMF-based HU, it is necessary to analyze the convergence property of the update rule. The objective function of  $L_{1/2}$ -RNMF is denoted as  $f(\mathbf{A}, \mathbf{S}, \mathbf{E})$ , as shown in (18). Our purpose is to prove that this objective function is nonincreasing in each update step shown in (20). That is to say, if we set  $\mathbf{A}^k, \mathbf{S}^k, \mathbf{E}^k$  as the values of the  $k$ th iteration, and  $\mathbf{A}^{k+1}, \mathbf{S}^{k+1}, \mathbf{E}^{k+1}$  are the values obtained by update rule (20), then we should prove

$$f(\mathbf{A}^{k+1}, \mathbf{S}^k, \mathbf{E}^k) \leq f(\mathbf{A}^k, \mathbf{S}^k, \mathbf{E}^k) \quad (21a)$$

$$f(\mathbf{A}^{k+1}, \mathbf{S}^{k+1}, \mathbf{E}^k) \leq f(\mathbf{A}^{k+1}, \mathbf{S}^k, \mathbf{E}^k) \quad (21b)$$

$$f(\mathbf{A}^{k+1}, \mathbf{S}^{k+1}, \mathbf{E}^{k+1}) \leq f(\mathbf{A}^{k+1}, \mathbf{S}^{k+1}, \mathbf{E}^k). \quad (21c)$$

In each update step of (20a)–(20b), if we assume  $\mathbf{M} := \mathbf{X} - \mathbf{E}^k$ , in which  $\mathbf{M}$  is a constant non-negative matrix, then

the update rule for  $\mathbf{A}$  and  $\mathbf{S}$  is the same as in (11). Thus, the inequality of (21a)–(21b) is the same as  $L_{1/2}$ -NMF, which has been proven in [8] and [27]. As a result, we only need to present the non-negativity of  $\mathbf{X} - \mathbf{E}$  and the inequality of (21c). The inequality of (21c) can be rewritten as

$$\begin{aligned} & \|\mathbf{X} - \mathbf{E}^{k+1} - \mathbf{A}^{k+1} \mathbf{S}^{k+1}\|_F^2 + \lambda \|\mathbf{E}^{k+1}\|_{1,2} \\ & \leq \|\mathbf{X} - \mathbf{E}^k - \mathbf{A}^{k+1} \mathbf{S}^{k+1}\|_F^2 + \lambda \|\mathbf{E}^k\|_{1,2} \end{aligned} \quad (22)$$

which can be easily deduced by Lemma 1 as  $\mathbf{E}^{k+1} = \text{soft}_\lambda(\mathbf{X} - \mathbf{A}^{k+1} \mathbf{S}^{k+1})$ .

The non-negativity of  $\mathbf{X} - \mathbf{E}^k$  can also be achieved as long as the initial values of  $\mathbf{A}$  and  $\mathbf{S}$  are set to be strictly positive matrices, and  $\mathbf{E}$  is a zero matrix. If  $\mathbf{X} - \mathbf{E}$  is non-negative, the update rule of (20a)–(20b) guarantees that the elements of the two matrices  $\mathbf{A}$  and  $\mathbf{S}$  remain non-negative. In addition, the non-negativity of  $\mathbf{X} - \mathbf{E}$  can also be retained after the *soft* thresholding (20c) if  $\mathbf{X}$ ,  $\mathbf{A}$ , and  $\mathbf{S}$  are all non-negative. All the above make the objective function  $f(\mathbf{A}, \mathbf{S}, \mathbf{E})$  decrease monotonically at each iteration until convergence has been reached.

### E. Implementation Issues

As mentioned before, the sparsity-regularized RNMF-based methods are simultaneously nonconvex with respect to  $\mathbf{A}$  and  $\mathbf{S}$ . As a result, the initializations of matrices  $\mathbf{A}$ ,  $\mathbf{S}$ , and  $\mathbf{E}$  are important, and different initial values will result in different results [39]. The sparse noise matrix  $\mathbf{E}$  is initialized to be a zero matrix. For the endmember matrix  $\mathbf{A}$ , there are generally two strategies: random initialization and selection from the original data. The former strategy randomly allocates values between 0 and 1 as entries of  $\mathbf{A}$ . The latter strategy consists of two methods. The first method is spectral information divergence (SID) based selection to determine  $\mathbf{A}$ , which was used in [10]. The second method is to utilize unsupervised endmember extraction methods to identify the endmembers as the input of  $\mathbf{A}$  [8], [9]. After determining the endmember matrix  $\mathbf{A}$  in both methods, a fully constrained least squares (FCLS) solution [40] is adopted to generate the abundance matrix  $\mathbf{S}$ . All of these initialization methods can guarantee the non-negativity of the matrices  $\mathbf{A}$  and  $\mathbf{S}$ .

The second issue we are concerned about is the ANC and ASC constraints, which can reduce the solution space of the optimization. Actually, the non-negativity is guaranteed under the update rule (20) [(19) for  $L_1$ -RNMF] if the initial matrices  $\mathbf{A}$  and  $\mathbf{S}$  are non-negative, and the initial sparse matrix  $\mathbf{E}$  is a zero matrix. The ASC constraint can be implemented by the effective and widely used method proposed in [40], where the data matrix  $\mathbf{X} - \mathbf{E}$  and the endmember matrix  $\mathbf{A}$  are augmented by a row of constants defined by

$$(\mathbf{X} - \mathbf{E})_f = \begin{bmatrix} \mathbf{X} - \mathbf{E} \\ \delta \mathbf{1}_N^T \end{bmatrix}, \quad \mathbf{A}_f = \begin{bmatrix} \mathbf{A} \\ \delta \mathbf{1}_K^T \end{bmatrix} \quad (23)$$

where  $\delta$  is used to adjust the effect of the ASC. A larger value of  $\delta$  can lead to a more accurate result, but with a much lower convergence rate. In order to achieve the desired tradeoff, a

relatively small value  $\delta = 15$  was selected in our experiments, as introduced in [9].

The third issue is about the parameters in the model. Parameter  $\lambda$  controls the degree of intensity of the sparse noise. When  $\lambda$  is set to 0, the optimal solution of model (16) is  $\mathbf{A} = \mathbf{S} = \mathbf{0}$  and  $\mathbf{E} = \mathbf{X}$ . On the other hand, the RNMF model reduces to a standard NMF model if  $\lambda$  is set to a large enough value. The analysis of parameter  $\lambda$  is given in the experimental section. The value of parameter  $\gamma$  is dependent on the sparsity of the material abundances. Since these abundances cannot be obtained *a priori*, we use an estimator for  $\gamma$  based on the sparseness criteria in [8] and [30], which is defined as

$$\gamma = \frac{1}{\sqrt{L}} \sum_l \frac{\sqrt{N} - \|x_l\|_1 / \|x_l\|_2}{\sqrt{N} - 1}$$

where  $x_l$  denotes the  $l$ th band in the HSI.

With the proposed methods, the selection of parameters  $\delta$  and  $\gamma$  may not be optimal for all cases. However, in the experiments, we fixed parameters  $\delta$  and  $\gamma$  using the same strategy as [8] and [9], to allow a fair comparison between the different methods, and, in addition, to reduce the complexity of the proposed sparsity-regularized RNMF.

We also adopt two stopping criteria for the optimization. The first criterion is to set an error tolerance that is predefined. Once the error is successively within the limits of the tolerance 10 times, the iteration is stopped. The other criterion is to set a maximum iteration number, which was adopted in our experiments, using a maximum iteration number of 3000. Once either of these criteria is met, the optimization ends. In general, they are sufficient to guarantee the convergence.

In addition, before the HU, we should point out that we assume the number of endmembers to be known. Although a proper estimation is important to the result, it is another independent topic, which can be solved by resorting to virtual dimensionality [41] or HySime [42].

The last issue is specifically for  $L_{1/2}$ -RNMF. As shown in [8], in order to improve the robustness of the algorithm, not all the elements in  $\mathbf{S}$  are updated following the application of (20b). For those elements less than a predefined threshold, we omit the additional term corresponding to the  $L_{1/2}$  sparsity operator. In our experiments, the threshold was set to  $10^{-4}$ .

The proposed sparsity-regularized RNMF approach is summarized in Algorithm 1.

#### F. Computational Complexity Analysis

The computational complexity is important for algorithms. Based on the update rules (10), (11), (19), and (20), it is clear that the additional calculation of the RNMF-based methods, compared to the NMF-based methods, is centralized on the update step of  $\mathbf{E}$ . We first analyze the additional computational cost of RNMF compared to the standard NMF. The update rules of RNMF are very similar to (19) and (20), with a slight adjustment in the update of  $\mathbf{S}$ , as follows:

$$\mathbf{A} \leftarrow \mathbf{A} * (\mathbf{X} - \mathbf{E})\mathbf{S}^T ./ \mathbf{A}\mathbf{S}\mathbf{S}^T \quad (24a)$$

$$\mathbf{S} \leftarrow \mathbf{S} * \mathbf{A}^T (\mathbf{X} - \mathbf{E}) ./ (\mathbf{A}^T \mathbf{A} \mathbf{S}) \quad (24b)$$

$$\mathbf{E} \leftarrow \text{soft}_\lambda(\mathbf{X} - \mathbf{A}\mathbf{S}). \quad (24c)$$

---

#### Algorithm 1. Sparsity-Regularized RNMF for Hyperspectral Unmixing

---

- 1 **Input:** The observed mixture data  $\mathbf{X} \in \mathbb{R}^{L \times N}$ , the number of endmembers  $K$  the parameters  $\lambda$  and  $\gamma$ .
  - 2 **Output:** Endmember signature matrix  $\mathbf{A}$  and abundance matrix  $\mathbf{S}$ .
  - 3 **Initialize**  $\mathbf{A}$ ,  $\mathbf{S}$ , and  $\mathbf{E}$  by the initialization rules mentioned before.
  - 4 **Repeat** until convergence:
  - 5 Update  $\mathbf{A}$  by (20a);
  - 6 Augment  $\mathbf{X} - \mathbf{E}$  and  $\mathbf{A}$  to obtain  $(\mathbf{X} - \mathbf{E})_f$  and  $\mathbf{A}_f$ , respectively;
  - 7 Update  $\mathbf{S}$  by (20b) (the update rule is (19b) for  $L_1$ -RNMF);
  - 8 Update  $\mathbf{E}$  by (20c).
- 

The floating-point calculation of each step in (6) and (24) is presented in Table I. From Table I, we can see that the total additional floating-point calculation of RNMF compared to NMF is  $2LNK + 6LN + 2L$  in each step. Without any obstacles, we can also deduce that the additional calculation of  $L_{1/2}$ -RNMF compared to  $L_{1/2}$ -NMF (the same as  $L_1$ -RNMF compared to  $L_1$ -NMF) is also  $2LNK + 6LN + 2L$  in each step. That is to say, the computational complexities of the RNMF-based methods and the NMF-based methods are of the same magnitude.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

Both synthetic and real image data experiments were undertaken to demonstrate the effectiveness of the proposed methods for spectral unmixing. In the synthetic image data experiments, we compared  $L_1$ -RNMF and  $L_{1/2}$ -RNMF with  $L_1$ -NMF and  $L_{1/2}$ -NMF, respectively, to demonstrate the robustness of the proposed sparsity-regularized RNMF methods. In the real data experiments, we compared the performance of  $L_1$ -RNMF,  $L_{1/2}$ -RNMF,  $L_1$ -NMF,  $L_{1/2}$ -NMF, and  $L_{1/2}$ -NMF with a data whitening step (denoted as  $L_{1/2}$ -WNMF), robust NMF (rNMF) [35], EDCNMF [10], and simplex identification via split augmented Lagrangian (SISAL) [16]. For the case of  $L_{1/2}$ -WNMF, the strategy in [43] was adopted for the data whitening step, in which multiple regression theory [42], [44], [45] is used to estimate the noise variance of the HSI. Subsequently,  $L_{1/2}$ -NMF was adopted to unmix the whitened HSI data. The results were evaluated using the spectral angle distance (SAD) and root-mean-square error (RMSE). The SAD was used to compare the similarity of the endmember signature  $\mathbf{A}_k$  and its estimate  $\hat{\mathbf{A}}_k$ , and is defined as

$$\text{SAD}_k = \arccos \left( \frac{\mathbf{A}_k^T \hat{\mathbf{A}}_k}{\|\mathbf{A}_k\| \|\hat{\mathbf{A}}_k\|} \right). \quad (25)$$

The RMSE is defined as

$$\text{RMSE}_k = \left( \frac{1}{N} \|\mathbf{S}_k - \hat{\mathbf{S}}_k\|^2 \right)^{1/2} \quad (26)$$

where  $\hat{\mathbf{S}}_k$  is the ground-truth abundance matrix for the  $k$ th endmember.

TABLE I  
FLOATING-POINT CALCULATION AT EACH ITERATION IN NMF AND RNMF

Methods	Update A	Update S	Update E	Total
NMF	Addition	$LNK + (L+N)K^2$	$LNK + (L+N)K^2$	$2LNK + 2(L+N)K^2$
	Multiplication	$LNK + (L+N)K^2 + LK$	$LNK + (L+N)K^2 + NK$	$2LNK + 2(L+N)K^2 + LK + NK$
	Division	$LK$	$NK$	$LK + NK$
RNMF	Addition	$LNK + (L+N)K^2 + LN$	$LNK + (L+N)K^2 + LN$	$3LNK + 2(L+N)K^2 + 4LN + L$
	Multiplication	$LNK + (L+N)K^2 + LK$	$LNL + (L+N)K^2 + NK$	$3LNK + 2(L+N)K^2 + LK + NK + 2LN$
	Division	$LK$	$NK$	$LK + NK + L$
Additional calculation	$LN$	$LN$	$2LNK + 4LN + 2L$	$2LNK + 6LN + 2L$

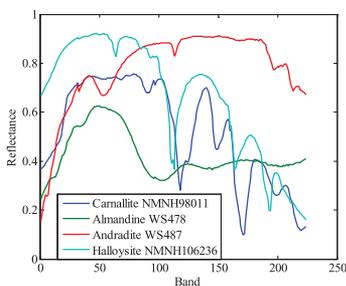


Fig. 3. Example endmember spectra for the synthetic data.

#### A. Simulated Data Experiments

In the synthetic experiments, eight spectral signatures were chosen from the USGS digital spectral library. Fig. 3 shows some of the endmember signatures and their names. The generation of abundances was similar to the method used in [8] and [6] and can be described as follows: 1) an image of size  $64 \times 64$  is divided into  $8 \times 8$  patches; 2) each patch is filled up by only one type of signature, which is randomly selected from the eight; 3) a  $7 \times 7$  low-pass filter is utilized to generate mixed pixels; 4) for the pixels whose abundance is larger than 0.8, a mixture composed of all the endmembers with the abundances of  $1/K$  takes their place; and 5) the noise is simulated.

Two kinds of noise were simultaneously added to the synthetic data. First, zero-mean Gaussian noise was added. Here, the SNR is defined as follows:

$$\text{SNR} = 10 \log_{10} \frac{E[x^T x]}{E[n^T n]} \quad (27)$$

in which  $x$  and  $n$  represent the observation and noise of a pixel, respectively, and  $E[\cdot]$  denotes the expectation operator. Second, sparse noise was added to certain bands of the synthetic image. We added impulse noise to the image, since the intensity of the impulse noise can be easily measured. We use *ratio* to denote the percentage of the bands corrupted by the impulse noise, and *sp* to denote the intensity of the impulse noise added to each band. For example, if *ratio* = 0.1 and *sp* = 0.1, this means that 10% of the synthetic image bands are corrupted by impulse noise, and for each corrupted band image, 10% of the pixels are corrupted by impulse noise.

By following these procedures, a synthetic image corrupted by Gaussian noise and sparse noise was produced. In the simulated experiments,  $\mathbf{A}$  was initialized by setting its entries to random values in the interval  $[0, 1]$ , then FCLS [40] was used to acquire the initial  $\mathbf{S}$ . Each experiment was repeated 10 times to ensure a reliable comparison.

1) *Experiment 1 (Parameter Analysis)*: In this experiment, the influence of parameter  $\lambda$  in the RNMF-based methods was considered when SNR = 30dB,  $K = 4$ , *ratio* = 0.2, and *sp* = 0.2. Here, the results of  $L_1$ -NMF and  $L_{1/2}$ -NMF are taken as the references. Fig. 4 shows how the performance of  $L_1$ -RNMF varies with parameter  $\lambda$ , and Fig. 5 shows the case of  $L_{1/2}$ -RNMF. As shown in Figs. 4 and 5, both  $L_1$ -RNMF and  $L_{1/2}$ -RNMF perform worse than the baseline methods when the value of parameter  $\lambda$  is relatively low. As the value of  $\lambda$  increases, the performances of the RNMF-based methods gradually improve. In particular,  $L_1$ -RNMF attains a better performance than  $L_1$ -NMF in both SAD and RMSE when  $\lambda$  is larger than 0.8, and  $L_{1/2}$ -RNMF obtains a better performance when  $\lambda$  is larger than 1.8. This verifies the effectiveness of incorporating the robust estimation in the HU by adopting a sparse matrix to model the sparse noise. On the other hand, when  $\lambda$  increases to a certain level, the performance of the RNMF-based methods decreases. In particular, the performance of the RNMF-based methods will reduce to the baseline methods if  $\lambda$  is set to infinitely large. In fact, the optimal value of parameter  $\lambda$  in the RNMF-based methods is related to the size of the image and the intensity of the sparse noise. The adaptive selection of parameter  $\lambda$  still remains a key problem. In the simulated experiments, we set  $\lambda$  by searching the range  $[0.1, 5]$  at 20 equally spaced values.

2) *Experiment 2 (Robustness Analysis to Gaussian Noise)*: In this experiment, four different HU algorithms were performed with different Gaussian noise levels, with  $K = 4$ , *ratio* = 0.2, and *sp* = 0.2. The SNR was assigned as 10, 20, 30, and 40 dB, respectively. Fig. 6 shows the SAD values of the estimated endmember signatures, and the RMSE of the estimated abundance maps as functions of the SNR. From the figure, we can see that the performances of the RNMF-based methods have the same trend as those of the NMF-based methods. This indicates that the RNMF-based methods are robust to Gaussian noise, as are the NMF-based methods. In addition, the

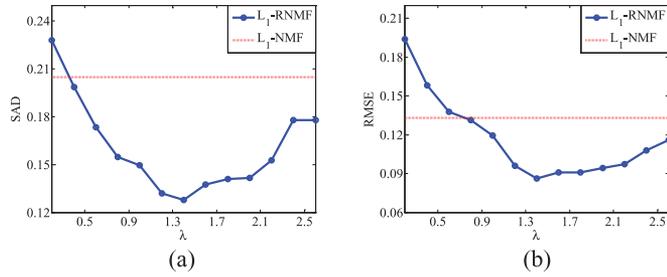


Fig. 4. Performance of  $L_1$ -RNMF with respect to parameter  $\lambda$  in terms of (a) SAD and (b) RMSE.

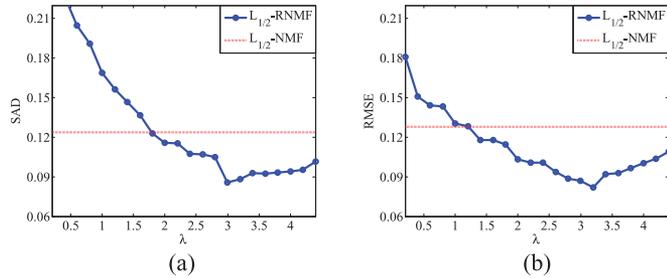


Fig. 5. Performance of  $L_{1/2}$ -RNMF with respect to parameter  $\lambda$  in terms of (a) SAD and (b) RMSE.

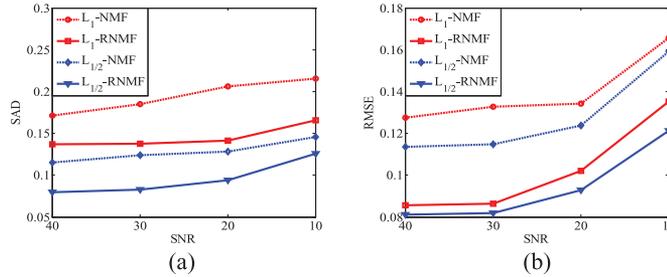


Fig. 6. Comparison of the algorithms at different Gaussian noise levels in terms of (a) SAD and (b) RMSE.

performance of  $L_1$ -RNMF is better than that of  $L_1$ -NMF, and  $L_{1/2}$ -RNMF outperforms  $L_{1/2}$ -NMF. This is mainly because the RNMF-based methods are robust to sparse noise. On the other hand, the NMF-based methods perform worse than the RNMF-based methods under the corruption of sparse noise.

3) *Experiment 3 (Robustness Analysis to Sparse Noise)*: We also tested the robustness of the proposed RNMF-based methods with respect to sparse noise. The number of endmembers  $K$  was fixed as 4, and the Gaussian noise level SNR was set to 30 dB. Fig. 7 presents the SAD and RMSE values of the four HU results in the cases of four different sparse noise intensities, namely,  $ratio = sp = 0, 0.1, 0.2, 0.3$ , respectively. Here,  $ratio = sp = 0$  means that no sparse noise was added to the simulated data. In this case, as shown in Fig. 7, the performances of the RNMF-based methods are the same as those of the NMF-based methods. However, as the intensity of the sparse noise increases, the SAD and RMSE values of the NMF-based methods increase rapidly. This indicates that the performances of  $L_1$ -NMF and  $L_{1/2}$ -NMF are seriously impacted by the sparse noise. On the other hand, the RNMF-based methods are robust to the sparse noise. The SAD and RMSE values of  $L_1$ -RNMF and  $L_{1/2}$ -RNMF are stable

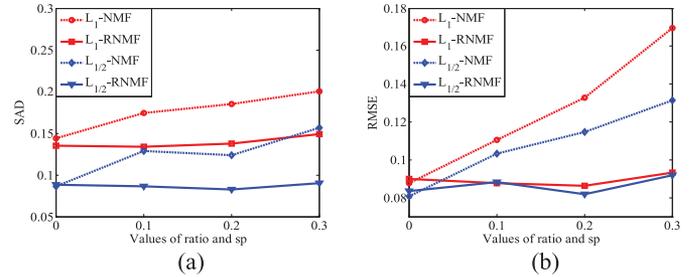


Fig. 7. Comparison of the algorithms at different sparse noise levels in terms of (a) SAD and (b) RMSE.

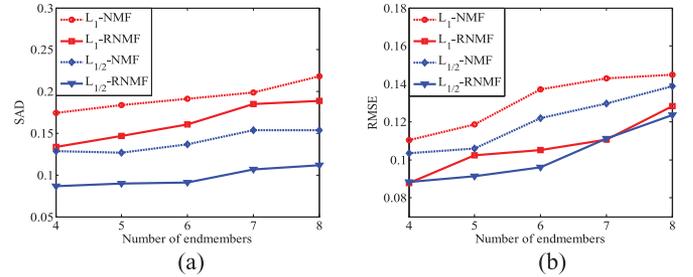


Fig. 8. Comparison of the algorithms at different numbers of endmembers in terms of (a) SAD and (b) RMSE.

with the increase in the sparse noise intensity. This confirms the great advantage of the RNMF-based methods in solving the HU problem in the presence of sparse noise.

4) *Experiment 4 (Effect of Different Numbers of Endmembers)*: This experiment evaluated the performances of the four unmixing methods when the data were made up of different numbers of endmembers. The SNR was set as 30 dB,  $ratio = sp = 0.1$ , and  $K$  was varied from 4 to 8. The results are shown in Fig. 8. From the figure, we can see that the performances of the four methods decrease when the number of endmembers increases. On the whole, the results of the RNMF-based methods are superior to those of the NMF-based methods for all the cases of  $K$ . This is mainly due to the sparse noise, which greatly affects the performance of the NMF-based methods. Meanwhile, the RNMF-based methods can overcome the impact of the sparse noise. This further indicates the robustness of RNMF to sparse noise.

## B. Real Data Experiments

We now present the results of applying the RNMF-based methods to a real-world dataset. The Urban dataset was collected by the HYDICE sensor and can be downloaded online at <http://www.agc.army.mil/>. In the Urban dataset, there are 210 bands that cover the wavelength range of 400–2500 nm. Fig. 1 shows some bands of the original data, which include low-noise bands, noisy bands, and water-absorption bands. Here, we use “low-noise” to describe the high-SNR bands, which include bands 5–75, 77–86, 88–100, 112–135, and 154–197. In most of the recent studies [8], [21], [22], only these high-SNR bands were adopted to unmix the Urban data. However, it is worth mentioning that these low-noise bands are also corrupted by sparse noise, as presented in Fig. 2. The existence of sparse noise can degrade the performance of most of the existing unmixing methods. The noisy bands include bands 1–4, 76,

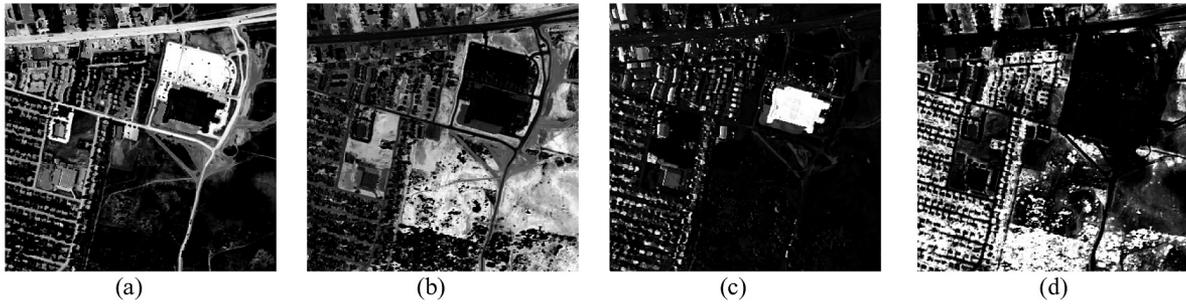


Fig. 9. Benchmark abundance maps for four targets in the Urban HYDICE data. (a) Asphalt. (b) Grass. (c) Roof. (d) Tree.

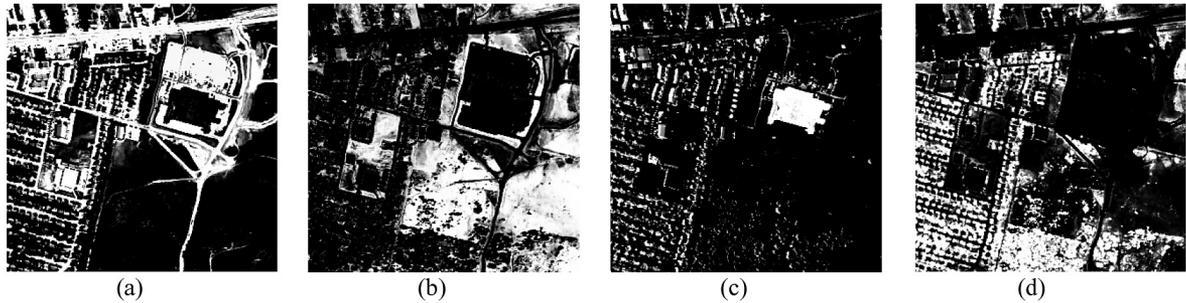


Fig. 10. Abundance maps of different endmembers using  $L_{1/2}$ -NMF with the *low-noise image*. (a) Asphalt. (b) Grass. (c) Roof. (d) Tree.

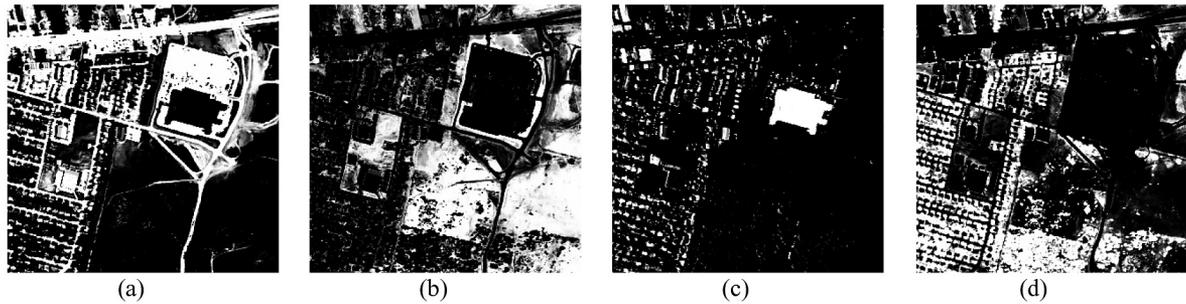


Fig. 11. Abundance maps of different endmembers using  $L_{1/2}$ -RNMF with the *low-noise image*. (a) Asphalt. (b) Grass. (c) Roof. (d) Tree.

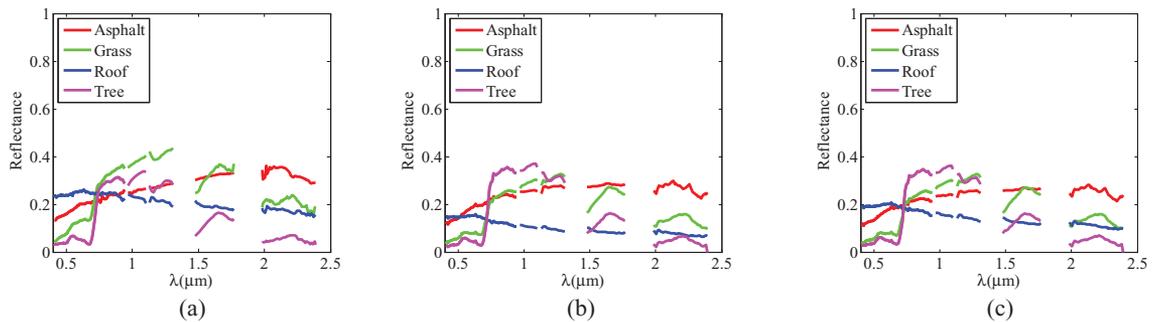


Fig. 12. Endmember signatures selected from the spectral library and extracted by  $L_{1/2}$ -NMF and  $L_{1/2}$ -RNMF with the *low-noise image*. (a) Collected from the spectral library. (b) Extracted by  $L_{1/2}$ -NMF. (c) Extracted by  $L_{1/2}$ -RNMF.

87, 101–104, 110–111, 136–138, 152–153, and 198–207. The rest of the bands are categorized as water-absorption bands, which cannot provide any useful information, only noise. As described above, the noisy and water-absorption bands are discarded in the traditional unmixing procedures. In this paper, we first present the HU results of the proposed methods with the *low-noise image*, which consists of 162 low-noise bands of

the Urban HYDICE data, and we then present the results of the four HU methods with the *noisy image*, which consists of 162 low-noise bands and 27 noisy bands of the Urban HYDICE data. This strategy can further prove the robustness of the proposed RNMF-based methods to sparse noise. In the real data experiments, parameter  $\lambda$  for all the RNMF-based methods was set to 2.

TABLE II  
SAD VALUES OF THE DIFFERENT METHODS WITH THE LOW-NOISE IMAGE

	$L_{1/2}$ -RNMF	$L_{1/2}$ -NMF	$L_{1/2}$ -WNMF	$L_1$ -RNMF	$L_1$ -NMF	rNMF	EDCNMF	SISAL
Asphalt	0.0871	<u>0.0865</u>	0.0982	0.0886	0.0871	<b>0.0654</b>	0.0976	0.2046
Grass	<u>0.0838</u>	0.0864	0.1151	0.0869	0.0873	0.1471	<b>0.0822</b>	0.2031
Roof	<b>0.0590</b>	0.1596	0.1103	0.2745	0.3380	0.3024	0.2232	<u>0.1045</u>
Tree	<b>0.0811</b>	0.0934	<u>0.0834</u>	0.1359	0.1441	0.1346	0.1416	0.1580
Mean	<b>0.0777</b>	0.1075	<u>0.1018</u>	0.1465	0.1641	0.1623	0.1361	0.1675

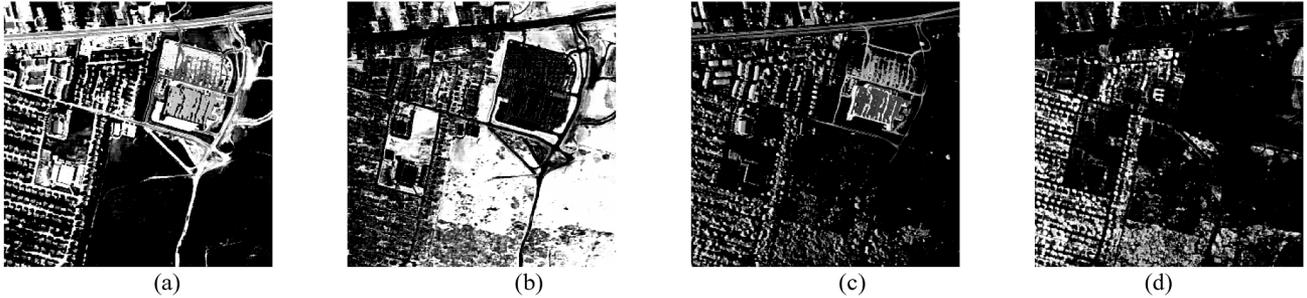


Fig. 13. Abundance maps of different endmembers using  $L_{1/2}$ -NMF with the noisy image. (a) Asphalt. (b) Grass. (c) Roof. (d) Tree.

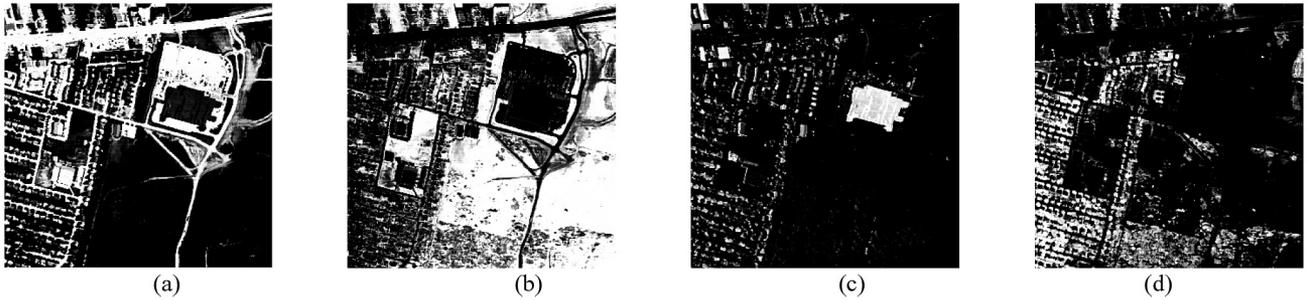


Fig. 14. Abundance maps of different endmembers using  $L_{1/2}$ -RNMF with the noisy image. (a) Asphalt. (b) Grass. (c) Roof. (d) Tree.

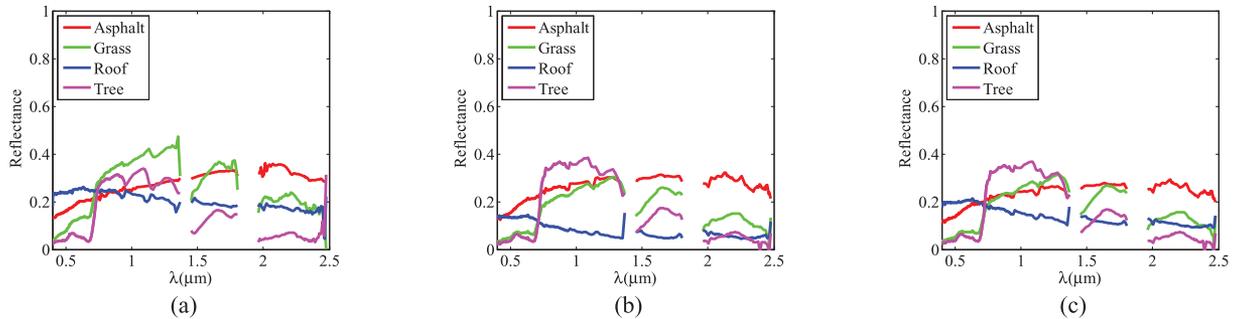


Fig. 15. Endmember signatures selected from the spectral library and extracted by  $L_{1/2}$ -NMF and  $L_{1/2}$ -RNMF with the noisy image. (a) Collected from the spectral library. (b) Extracted by  $L_{1/2}$ -NMF. (c) Extracted by  $L_{1/2}$ -RNMF.

1) *Unmixing Results With the Low-Noise Image:* In light of the previous analysis [8], [24], four types of signatures named “Asphalt,” “Grass,” “Roof,” and “Tree” were estimated in the image. The reference signatures were collected from the spectral library which can be downloaded from <http://www.agc.army.mil/>. The benchmark abundance maps of each signature shown in Fig. 9 were achieved via the method introduced in [7] and [24]. For all the methods, we used the SID-based method to select four pixels as the initial endmember matrix. In addition, the initial endmember matrix was the

same for all the unmixing methods, and the experiments were repeated 10 times, to ensure a reliable comparison.

To save space, we take the  $L_{1/2}$  sparse regularization as an example and only show the results of  $L_{1/2}$ -NMF and  $L_{1/2}$ -RNMF. We first present the abundance maps of each endmember obtained by the different unmixing methods. Fig. 10 illustrates the resulting abundances by  $L_{1/2}$ -NMF, and Fig. 11 shows the results of  $L_{1/2}$ -RNMF. Figs. 10 and 11 are both grayscale abundance maps, where a dark pixel denotes a low abundance of the corresponding endmember. Subsequently,

TABLE III  
SAD VALUES OF THE DIFFERENT METHODS WITH THE NOISY IMAGE

	$L_{1/2}$ -RNMF	$L_{1/2}$ -NMF	$L_{1/2}$ -WNMF	$L_1$ -RNMF	$L_1$ -NMF	rNMF	EDCNMF	SISAL
Asphalt	<u>0.0844</u>	0.0948	0.0982	0.0942	0.0914	0.0767	0.1001	0.2144
Grass	<b>0.0786</b>	<u>0.0918</u>	0.1329	0.1004	0.1066	0.1579	0.1055	0.2320
Roof	<b>0.0652</b>	0.1746	0.1403	0.2675	0.4014	0.3982	0.2418	0.1083
Tree	<b>0.0692</b>	<u>0.1028</u>	0.1234	0.1285	0.1444	0.1401	0.1413	0.1592
Mean	<b>0.0744</b>	<u>0.1160</u>	0.1237	0.1476	0.1859	0.1932	0.1472	0.1785

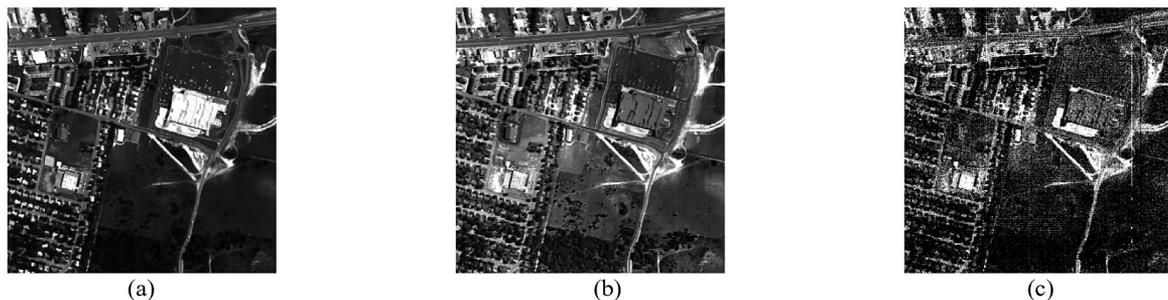


Fig. 16. Some bands of the HYDICE Urban dataset. (a) Band 28. (b) Band 112. (c) Band 206.

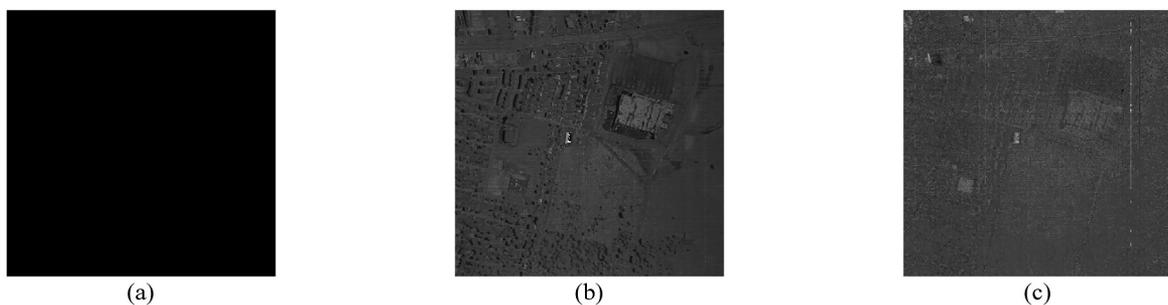


Fig. 17. Some bands of the sparse error matrix  $\mathbf{E}$  obtained by  $L_{1/2}$ -RNMF on the noisy image. (a) Band 28. (b) Band 112. (c) Band 206.

the endmember signatures selected from the spectral library and extracted by  $L_{1/2}$ -NMF and  $L_{1/2}$ -RNMF are presented in Fig. 12. Meanwhile, Table II gives the mean SAD values obtained by the four HU methods to present the quantitative evaluations. The best results for the quality index are labeled in bold. From the table, it can be clearly observed that the  $L_{1/2}$ -regularized methods achieve lower mean SAD values than the  $L_1$ -regularized methods. Furthermore,  $L_{1/2}$ -RNMF achieves a lower mean SAD value than  $L_{1/2}$ -NMF, because of the sparse modeling of the sparse noise. Similarly, the unmixing result of  $L_1$ -RNMF is better than that of  $L_1$ -NMF.  $L_{1/2}$ -WNMF achieves higher SAD values than  $L_{1/2}$ -RNMF, which suggests that it is more appropriate to model the Urban data as corrupted by mixed noise rather than noniid Gaussian noise. EDCNMF and rNMF perform better than the  $L_1$ -NMF method, but not as well as the  $L_{1/2}$ -RNMF method. SISAL has the largest SAD values. All in all, it can be concluded that the proposed RNMF-based methods are generally more robust than the corresponding NMF-based methods, with respect to the sparse noise.

2) *Unmixing Results With the Noisy Image*: To further demonstrate the robustness of the proposed RNMF-based methods to sparse noise, we also applied the unmixing methods

to the noisy image, which contains 189 bands in total. To ensure a reliable comparison between the unmixing results with the low-noise image and the noisy image, the initial endmember matrices were composed of the pixels selected from the same places in the Urban data. Fig. 13 presents the resulting abundances by  $L_{1/2}$ -NMF, and Fig. 14 shows the results of  $L_{1/2}$ -RNMF. Table III presents the mean SAD values. Note that the endmember signatures collected from the spectral library were also corrupted by noise, as shown in Fig. 15(a). Therefore, before computing the SAD values, we removed the spectra of the noisy bands and only calculated the spectra of the low-noise bands. For the unmixing results with the noisy image, similar to the low-noise image case, it was found that the proposed RNMF-based methods perform better than the corresponding NMF-based methods, which clearly shows the strong robustness to the sparse noise existing in the noisy image.

In addition, comparing the unmixing results of the low-noise image with those of the noisy image, we can further observe that the results of the RNMF-based methods  $L_{1/2}$ -RNMF and  $L_1$ -RNMF are relatively stable; meanwhile, the results of the NMF-based methods  $L_{1/2}$ -NMF,  $L_1$ -NMF, EDCNMF, and SISAL show a significant decline. This further indicates the robustness of the RNMF-based methods to sparse noise.

Furthermore, by comparing the results of  $L_{1/2}$ -RNMF with the low-noise image and the noisy image, we can see that the average SAD value of the low-noise image shown in Table II is slightly higher than that of the noisy image presented in Table III. This suggests that the noisy bands of the Urban dataset can also provide useful information for the unmixing if appropriate and rational constraints for the noise are enforced.

To further investigate the effectiveness of the proposed RNMF method, we also present some bands of the sparse error matrix  $\mathbf{E}$  obtained by  $L_{1/2}$ -RNMF with the noisy image. Fig. 16 presents some bands of the Urban data, and Fig. 17 shows the corresponding bands of sparse matrix  $\mathbf{E}$  obtained by  $L_{1/2}$ -RNMF on the noisy image. From the figure, it can be observed that most of the deadlines, stripes, and impulse noise present in the original band 206 are included in the sparse error matrix  $\mathbf{E}$ . In addition, some edge information is also included in the sparse error matrix  $\mathbf{E}$ , as shown in Fig. 17(b) and (c). It can therefore be concluded that after excluding the sparse noise, the endmembers and abundances estimated by the proposed RNMF method are more accurate.

## V. CONCLUSION

In this paper, we first proposed an extended LMM which separately models the sparse noise and Gaussian noise. Based on the extended LMM model, sparsity-regularized RNMF methods were then proposed to unmix hyperspectral data. Compared to the existing sparsity-regularized NMF methods, the proposed sparsity-regularized RNMF methods can lead to more desirable results when handling sparse noise. In addition, we studied the real HYDICE Urban dataset in depth and partitioned it into three subsets, i.e., low-noise bands, noisy bands, and water-absorption bands. We then compared the results of the proposed sparsity-regularized RNMF methods with a low-noise image consisting of low-noise bands and a noisy image consisting of both low-noise and noisy bands, and concluded that the noisy bands can also provide appropriate and useful information for urban data unmixing.

Nevertheless, the proposed methods still have room for improvement, and the adaptive selection of parameter  $\lambda$  remains a key problem. Moreover, the proposed sparsity-regularized RNMF methods are special cases of the following RNMF-based unmixing model:

$$\min_{\mathbf{A}, \mathbf{S}, \mathbf{E}} f(\mathbf{A}, \mathbf{S}, \mathbf{E}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E} - \mathbf{AS}\|_F^2 + \lambda \|\mathbf{E}\|_{1,2} + \beta h(\mathbf{A}) + \gamma g(\mathbf{S}) \quad \text{s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0 \quad (28)$$

where  $h$  and  $g$  are regularizers on the endmember matrix  $\mathbf{A}$  and abundance matrix  $\mathbf{S}$ , respectively. The robustness of the RNMF model with various other regularizers besides the sparsity regularizer will be explored and implemented in our future work.

## REFERENCES

- [1] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [2] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2056–2065, Jun. 2013.
- [3] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 9, pp. 44–57, Jan. 2002.
- [4] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.
- [5] N. Dobigeon, J. Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.
- [6] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.
- [7] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan. 2009.
- [8] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via  $L_{1/2}$  sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.
- [9] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, May 2013.
- [10] N. Wang, B. Du, and L. Zhang, "An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 554–569, Apr. 2013.
- [11] D. C. Zanotta, V. Haertel, Y. E. Shimabukuro, and C. D. Renno, "Linear spectral mixing model for identifying potential missing endmembers in spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 3005–3012, May 2014.
- [12] W. K. Ma *et al.*, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, Jan. 2014.
- [13] J. M. P. Nascimento and J. M. Bioucas Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [14] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE's Int. Symp. Opt. Sci. Eng. Instrum.*, 1999, pp. 266–275.
- [15] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.
- [16] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. 1st Workshop Hyperspectr. Image Signal Process.: Evol. Remote Sens. (WHISPERS'09)*, 2009, pp. 1–4.
- [17] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.
- [18] M. D. Iordache, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "MUSIC-CSR: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4364–4382, Jul. 2014.
- [19] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
- [20] D. Donoho and V. Stodden, "When does non-negative matrix factorization give a correct decomposition into parts?" in *Proc. Adv. Neural Inf. Process. Syst. Conf.*, 2004, pp. 1141–1148.
- [21] Z. Yang, G. Zhou, S. Xie, S. Ding, J.-M. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1112–1125, Apr. 2011.
- [22] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.
- [23] A. Huck, M. Guillaume, and J. Blanc-Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2590–2602, Jun. 2010.
- [24] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng, and C. Pan, "Spectral unmixing via data-guided sparsity," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5412–5427, Dec. 2014.

- [25] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 178–188, Jan. 2016.
- [26] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 4729–4743, Aug. 2014.
- [27] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.
- [28] C.-J. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Comput.*, vol. 19, pp. 2756–2779, 2007.
- [29] N. Guan, D. Tao, Z. Luo, and B. Yuan, "NeNMF: An optimal gradient method for nonnegative matrix factorization," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 2882–2898, Jun. 2012.
- [30] P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints," *J. Mach. Learn. Res.*, vol. 5, pp. 1457–1469, 2004.
- [31] G. A. Shaw and H.-H. K. Burke, "Spectral imaging for remote sensing," *Lincoln Lab. J.*, vol. 14, pp. 3–28, 2003.
- [32] A. Cichocki, R. Zdunek, and S.-I. Amari, "Csiszar's divergences for non-negative matrix factorization: Family of new algorithms," in *Independent Component Analysis and Blind Signal Separation*. New York, NY, USA: Springer, 2006, pp. 32–39.
- [33] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, p. 11, 2011.
- [34] L. Zhang, Z. Chen, M. Zheng, and X. He, "Robust non-negative matrix factorization," *Front. Electr. Electron. Eng. China*, vol. 6, pp. 192–200, 2011.
- [35] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, Dec. 2015.
- [36] N. Dobigeon and C. Févotte, "Robust nonnegative matrix factorization for nonlinear unmixing of hyperspectral images," in *Proc. IEEE Workshop Hyperspectr. Image Signal Process.: Evol. Remote Sens. (WHISPERS'13)*, 2013, pp. 1–4.
- [37] Y. Altmann, S. McLaughlin, and A. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEE Trans. Comput. Imag.*, vol. 1, no. 2, pp. 74–85, Jun. 2015.
- [38] G. Liu, Z. Lin, S. Yan, J. Sun, Y. Yu, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
- [39] A. Plaza and C.-I. Chang, "Impact of initialization on design of endmember extraction algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3397–3407, Nov. 2006.
- [40] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
- [41] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
- [42] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [43] Y.-Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 296–308, Jan. 2015.
- [44] W. He, H. Zhang, L. Zhang, and H. Shen, "Hyperspectral image denoising via noise-adjusted iterative low-rank matrix approximation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 3050–3061, Jun. 2015.
- [45] W. He, H. Zhang, L. Zhang, and H. Shen, "A noise-adjusted iterative randomized singular value decomposition method for hyperspectral image denoising," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGRASS'14)*, Jul. 13–18, 2014.



**Wei He** (S'14) received the B.S. degree in mathematics from the School of Mathematics and Statistics, Wuhan University, Wuhan, China, in 2012. Currently, he is pursuing the Ph.D. degree in photogrammetry and remote sensing at the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing (LIESMARS), Wuhan University.

His research interests include image quality improvement, remote sensing image processing, and low rank representation.



**Hongyan Zhang** (M'13) received the B.S. degree in geographic information system and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2005 and 2010, respectively.

He has been an Associate Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He has authored/coauthored more than 50 research papers. His research interests include image reconstruction, hyperspectral image processing, sparse representation, and low rank methods for sensing image imagery.

Dr. Zhang is a Reviewer of more than 10 international academic journals, including *IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING*, *IEEE TRANSACTION ON IMAGE PROCESSING*, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *IEEE Geoscience and remote sensing Letters* and so on.



**Liangpei Zhang** (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, ChangSha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics of Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He has been the Head with the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University. He is also a Chang-Jiang Scholar Chair Professor appointed by the Ministry of Education of China. He has been a Principal Scientist with the China State Key Basic Research Project (2011–2016) appointed by the Ministry of National Science and Technology of China to lead the remote sensing program in China. He has authored more than 410 research papers. He is the holder of 15 patents. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is the Founding Chair of the IEEE Geoscience and Remote Sensing Society (GRSS) Wuhan Chapter, a Fellow of the Institution of Engineering and Technology, an Executive Member (Board of Governors) of the China National Committee of the International Geosphere–Biosphere Programme, and an Executive Member of the China Society of Image and Graphics. He regularly serves as a Co-Chair of the series International Society for Optics and Photonics (SPIE) Conferences on Multispectral Image Processing and Pattern Recognition, the Conference on Asia Remote Sensing, and many other conferences. He edits several conference proceedings, issues, and geoinformatics symposiums. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the *International Journal of Digital Multimedia Broadcasting*, the *Journal of Geo-Spatial Information Science*, the *Journal of Remote Sensing*, and the *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*. He is a Guest Editor of the *Journal of Applied Remote Sensing* and the *Journal of Sensors*. He was the General Chair for the Fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) and a Guest Editor of the *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (JSTARS). He was the recipient of the 2010 Best Paper Boeing Award and the 2013 Best Paper ERDAS Award of American Society of Photogrammetry and Remote Sensing (ASPRS), respectively. He was the recipient of the Best Reviewer Awards from the IEEE GRSS for his service to the IEEE JSTARS in 2012 and the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2014, the 2010 Best Paper Boeing Award and the 2013 Best Paper ERDAS Award from the *American Society of Photogrammetry and Remote Sensing*. His research teams were the recipient of the top three prizes of the IEEE GRSS 2014 Data Fusion Contest, and his students have been selected as the winners or finalists of the IEEE International Geoscience and Remote Sensing Symposium Student Paper Contest in recent years.