# Nonlocal Tensor-Based Sparse Hyperspectral Unmixing

Jie Huang, *Member, IEEE*, Ting-Zhu Huang, Xi-Le Zhao, and Liang-Jian Deng, *Member, IEEE*

***Abstract*— Sparse unmixing is an important technique for analyzing and processing hyperspectral images (HSIs). Simultaneously exploiting spatial correlation and sparsity improves substantially abundance estimation accuracy. In this article, we propose to exploit nonlocal spatial information in the HSI for the sparse unmixing problem. Specifically, we first group similar patches in the HSI, and then unmix each group by imposing simultaneous a low-rank constraint and joint sparsity in the corresponding third-order abundance tensor. To this end, we build an unmixing model with a mixed regularization term consisting of the sum of the weighted tensor trace norm and the weighted tensor $\ell_{2,1}$-norm of the abundance tensor. The proposed model is solved under the alternating direction method of multipliers framework. We term the developed algorithm as the nonlocal tensor-based sparse unmixing algorithm. The effectiveness of the proposed algorithm is illustrated in experiments with both simulated and real hyperspectral data sets.**

***Index Terms*— Hyperspectral unmixing, joint sparsity, low-rank, nonlocal similarity, tensor.**

## I. Introduction

SPECTRAL unmixing of hyperspectral images (HSIs) has received much attention in recent years [1], [2]. Its basic task is to identify the spectral signatures of the materials (*endmembers*) and also the fractions (*abundances*) of these endmembers in each mixed HSI pixel. In the literature, a plethora of geometrical and statistical methods have been put forward to identifying the spectral signatures of endmembers, known as endmembers' extraction; see [3]–[7] and references therein. On the other hand, nonnegative matrix factorization-based blind hyperspectral algorithms have been well studied for the task [8]–[10]. In addition, another attractive research topic is to assume the spectral signatures of endmembers available and belonging to a predefined set usually called *endmembers' dictionary*, then diverse abundance estimation algorithms have been proposed in a semisupervised fashion [11]–[14].

Manuscript received May 14, 2020; revised September 6, 2020; accepted October 8, 2020. This work was supported in part by the NSFC under Grant 61772003, Grant 61876203, and Grant 61702083; in part by the Key Projects of Applied Basic Research in Sichuan Province under Grant 2020YJ0216; in part by the Science Strength Promotion Program of UESTC; and in part by the Fundamental Research Funds for the Central Universities under Grant ZYGX2019J093. *(Corresponding authors: Jie Huang; Ting-Zhu Huang.)*

The authors are with the School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: happyjie07mo@163.com; tingzhuhuang@126.com).

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

Digital Object Identifier 10.1109/TGRS.2020.3030233

The mixing process in numerous state-of-the-art unmixing algorithms is either characterized as linear [2] or nonlinear [15]. Linear mixture model (LMM) assumes that the spectral signature of each pixel in HSI is a linear combination of the spectral signatures of the endmembers, weighted by their corresponding abundances. Due to its simplicity and tractability, LMM has been widely adopted by diverse spectral unmixing algorithms (see [1], [2] and reference therein). For the remainder of this article, we will focus on the linear model in a semisupervised way.

Incorporating *a prior* information improves the abundance estimation accuracy. The *sparsity* assumption plays a key role in many semisupervised algorithms. Its adoption is justified by the fact that only a few of spectral signatures participate in the linear mixture process of each pixel, particularly compared with large-scale available dictionaries [11]. Practically, sparsity is imposed on abundances via the standard $\ell_1$-norm regularization [11] and its variants [8], [16]–[19]. Collaborative (or joint) sparsity scheme provides a refinement by simultaneously imposing the sparsity on all pixel in the scene [12], [20]–[22].

Exploiting *spatial correlation* in sparse unmixing framework has offered stimulating results. Typically, the total variation (TV) is included in sparse unmixing to promote piecewise smooth in each abundance map [13], [23]–[26]. The low-rank property is another structural characteristic that has been recently adopted for abundances [22], [27], [28]. It transfers the spatial correlation among pixels' spectral signatures into linear dependence among their corresponding abundance vectors. The resulting abundance matrix therefore admits a low-rank behavior and its rank is usually approximated by the weighted nuclear norm in practice. Nonlocal sparse unmixing presents an alternative way of modeling the spatial correlation. Nonlocal means method is incorporated into the TV regularization to exploit the similar patterns and structures in abundance maps [29], [30]. As an improvement of nonlocal means TV spatial consideration, an adaptive nonlocal Euclidean medians approach is proposed to better utilize the spatial information by means of filtering [31]. In addition, the nonlocal means method has been used in a collaborative sparse unmixing framework to suppress the abundance estimation error [32]. Recently, the nonlocal similarity of abundance maps is employed in unmixing problem by first grouping nonlocal similar abundance patches and then imposing a low-rank constraint [33], [34]. These nonlocal methods in [29], [31]–[34] mainly focus on exploring the nonlocal spatial information in abundance maps. Note that the abundances are

estimated iteratively, so the nonlocal spatial information is correspondingly updated as soon as the latest estimation of abundances is acquired.

In this article, we introduce a novel idea for exploring the nonlocal spatial information and spectral correlation. Departing from the usual paradigm, we explore the nonlocal spatial similarity in the HSI for hyperspectral unmixing. That said, we first group similar patches in the HSI and each group is stacked to be a third-order tensor. The technique of exploiting the spatial nonlocal similarity in HSIs has been well studied for HSI denoising [35]–[39]. It helps to improve the denoising performance, in which repetitive patches can provide more complementary information especially under heavy noise. Following this technique, we incorporate nonlocal spatial similarity and spectral correlation of HSIs into the unmixing process in an attempt to achieve better abundance estimation results. To the best of our knowledge, it is new for hyperspectral unmixing. Particularly, there is no need to recalculate the nonlocal spatial information in abundance maps at each iteration round. Then, we unmix each group by transferring the spatial and spectral correlation in each group into the low-rank property and the joint sparsity in the corresponding third-order abundance tensor. The low-rank and joint-sparsity constraints give rise to a mixed penalty term that regularizes a least squares data fitting function, through the *weighted tensor trace norm* and the proposed *weighted tensor $\ell_{2,1}$-norm.* To minimize the cost function, we propose an algorithm under the classic *alternating direction method of multipliers* (ADMM) framework. After unmixing each group of similar HSI patches, we reconstruct the abundance tensor for the HSI by aggregating estimated small-size abundance tensors. Thus, we name our algorithm *nonlocal tensor-based sparse unmixing* algorithm (NL-TSUn). The proposed algorithm is compared with state-of-the-art unmixing techniques, and the efficacy is demonstrated via simulated and real-data experiments.

*Notation:* We use nonbold letters for scalars, e.g., $x$, boldface lowercase letters for vectors, e.g., $\boldsymbol{x}$, and boldface capital letters for matrices, e.g., $\boldsymbol{X}$. Tensors are denoted by boldface calligraphic letters, e.g., $\boldsymbol{\mathcal{X}}$. Consider a third-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. The Frobenius norm of $\boldsymbol{\mathcal{X}}$ is defined as

$$\|\boldsymbol{\mathcal{X}}\|_F := \sqrt{\sum_{i_1,i_2,i_3} |x_{i_1 i_2 i_3}|^2}$$

where $x_{i_1 i_2 i_3}$s are elements of $\boldsymbol{\mathcal{X}}$. The inverse operator of unfolding is denoted by "fold", i.e., $\boldsymbol{\mathcal{X}} = \text{fold}_n(\boldsymbol{X}_{(n)})$. We note that different articles use different orderings of the columns for the mode-$n$ unfolding; see [40]–[43]. Here, we follow the order in [40] throughout this article. Recall from [40] that the *mode-n multiplication*, denoted by the symbol $\times_n$, defines the multiplication of an array $\boldsymbol{\mathcal{G}}$ by a matrix $\boldsymbol{S}$ along mode $n$ of the array and the outcome $\tilde{\boldsymbol{\mathcal{G}}}$ is equivalent to the matrix multiplication $\tilde{\boldsymbol{G}}_{(n)} = \boldsymbol{S}\boldsymbol{G}_{(n)}$. In addition, Table I lists the notations of the variables that we usually use in this article.

The rest of this article is organized as follows. Section II briefly reviews sparse unmixing models. Section III derives the proposed algorithm. Then, the effectiveness of the proposed algorithm is demonstrated by both simulated experiments in Section IV and a real-data experiment in Section V. Finally, Section VI gives some concluding remarks.

TABLE I
SOME NOTATIONS USED IN THIS ARTICLE

| Notation | Description |
|---|---|
| $n_1$ | spatial width of HSI |
| $n_2$ | spatial height of HSI |
| $J$ | spectral bands of HSI |
| $K$ | number of similar patches |
| $m$ | number of endmembers |
| $N$ | number of pixels in HSI |
| $I$ | number of pixels in a patch |
| $\boldsymbol{Y} \in \mathbb{R}^{J \times N}$ | HSI matrix |
| $\boldsymbol{A} \in \mathbb{R}^{J \times m}$ | spectral library |
| $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ | abundance matrix |
| $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{I \times J \times K}$ | tensor of similar HSI patches |
| $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{I \times m \times K}$ | abundance tensor of $\boldsymbol{\mathcal{Y}}$ |
| $\boldsymbol{X}_{(n)}$ | mode-$n$ unfolding of $\boldsymbol{\mathcal{X}}$ |
| $\boldsymbol{X}_{i::}$ | $i$-th horizontal slice of $\boldsymbol{\mathcal{X}}$ |
| $\boldsymbol{x}^{[i]}$ | $i$-th row of $\boldsymbol{X}$ |
| $\|\boldsymbol{X}\|_* = \sum_i \sigma_i(\boldsymbol{X})$ | nuclear norm |
| $\|\boldsymbol{X}\|_{\boldsymbol{w},*} = \sum_i w_i \sigma_i(\boldsymbol{X})$ | weighted nuclear norm |
| $\|\boldsymbol{\mathcal{X}}\|_{\boldsymbol{W},*} = \sum_n \|\boldsymbol{X}_{(n)}\|_{\boldsymbol{w}_n,*}$ | weighted tensor trace norm |
| $\|\boldsymbol{X}\|_{2,1} = \sum_i \|\boldsymbol{x}^{[i]}\|_2$ | $\ell_{2,1}$-norm |
| $\|\boldsymbol{X}\|_{\boldsymbol{z},2,1} = \sum_i z_i \|\boldsymbol{x}^{[i]}\|_2$ | weighted $\ell_{2,1}$-norm |
| $\|\boldsymbol{\mathcal{X}}\|_{\boldsymbol{Z},2,1} = \sum_i \|\boldsymbol{X}_{i::}\|_{\boldsymbol{z}_i,2,1}$ | weighted tensor $\ell_{2,1}$-norm |

## II. SPARSE UNMIXING MODEL

Let $\boldsymbol{\mathcal{Y}}_o \in \mathbb{R}^{n_1 \times n_2 \times J}$ denote the observed HSI with $n_1 \times n_2$ pixels and $J$ spectral bands. Then, the third-order tensor $\boldsymbol{\mathcal{Y}}_o$ is rearranged to a matrix $\boldsymbol{Y} \in \mathbb{R}^{J \times N}$ with $N = n_1 \times n_2$ [34], [40]. Let $\boldsymbol{A} \in \mathbb{R}^{J \times m}$ denote the dictionary with $m$ spectral signatures. The classic LMM can be described as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{E} \tag{1}$$

where $\boldsymbol{X} \in \mathbb{R}^{m \times N}$ is the fractional abundance matrix whose columns correspond with the abundance vectors of $N$ pixels in $\boldsymbol{Y}$, and $\boldsymbol{E} \in \mathbb{R}^{J \times N}$ is an independent and identically distributed (i.i.d.) zero-mean Gaussian noise matrix. With large and available $\boldsymbol{A}$, sparse unmixing aims to find a relatively small number of spectral signatures to linearly mix every pixel in $\boldsymbol{Y}$. Exploiting a prior information helps to improve the estimation accuracy. The so-called *abundance nonnegativity constraint* (ANC) and the *abundance sum-to-one constraint* (ASC)

$$\boldsymbol{X} \geq \boldsymbol{0}, \quad \boldsymbol{1}^T \boldsymbol{X} = \boldsymbol{1}^T$$

respectively, are widely considered for physical meaning [44]. Here, $\boldsymbol{X} \geq \boldsymbol{0}$ denote that each element is nonnegative and $\boldsymbol{1}$ is a column vector of 1's. Nevertheless, we relax the ASC to focus

on the exploitation of structural characters of $\boldsymbol{X}$, similarly as in [13], [12], [28], and [22]; see more details in [11].

To explore the sparsity of each abundance vector, the classic sparse regression model for hyperspectral unmixing problem is as follows:

$$\min_{\boldsymbol{X}} \ \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}\|_F^2+\lambda\|\boldsymbol{X}\|_{1,1} \quad \text{s.t. } \boldsymbol{X}\geq\boldsymbol{0} \tag{2}$$

where $\|\boldsymbol{X}\|_{1,1}=\sum_{i=1}^{m}\sum_{j=1}^{N}|x_{i,j}|$ is the $\ell_1$-norm of $\boldsymbol{X}$, $x_{i,j}$ denotes the $(i,j)$th element of $\boldsymbol{X}$, and $\lambda\geq 0$ is the regularization parameter. A sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) algorithm is introduced to solve the above model in [11]. In addition, the $\ell_p$-norm-based regularization has been considered to promote sparseness in the abundance and also in other HSI processing problems, see [16], [17], [45] and reference therein.

With the assumption that all pixels in the data set share the same active set of endmembers, the collaborative sparse regression framework is

$$\min_{\boldsymbol{X}} \ \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}\|_F^2+\lambda\|\boldsymbol{X}\|_{2,p} \quad \text{s.t. } \boldsymbol{X}\geq\boldsymbol{0} \tag{3}$$

where $p\in\{0,1\}$ and $\lambda\geq 0$ is the regularization parameter. Denote $\boldsymbol{x}^{[i]}$ be the $i$th row of $\boldsymbol{X}$. Suppose $p=1$, then $\|\boldsymbol{X}\|_{2,1}=\sum_{i=1}^{m}\|\boldsymbol{x}^{[i]}\|_2$ is the so-called $\ell_{2,1}$-norm of $\boldsymbol{X}$. It follows a convex optimization problem and the corresponding algorithm is termed as collaborative SUnSAL (CLSUnSAL) in [12]. Consider instead $p=0$ and $\|\boldsymbol{X}\|_{2,0}=\sum_{i=1}^{m}\mathbf{1}(\|\boldsymbol{x}^{[i]}\|_2>0)$ is the $\ell_{2,0}$ quasi-norm of $\boldsymbol{X}$, it is shown in [21] that the $\ell_0$ quasi-norm regularization provides sparser and more accurate estimations. The corresponding algorithm is termed as the collaborative sparse hyperspectral unmixing algorithm using $\ell_0$ quasi-norm (CSUnL0). In addition, a more general $\ell_{2,p}$-norm ($0<p<1$) regularization has been considered in [20] to promote sparsity along rows of the abundance matrix.

To exploit local collaborative sparsity property, a joint-sparse-blocks regression model has been proposed to encourage that similar adjacent pixels share the same support set of endmembers [22]. To this end, the abundance matrix is viewed as a block matrix

$$\boldsymbol{X}=[\boldsymbol{X}_1,\ldots,\boldsymbol{X}_s] \tag{4}$$

where the column number of $\boldsymbol{X}_j\in\mathbb{R}^{m\times d_j}$ satisfies $\sum_{j=1}^{s}d_j=N$ and the block number $s$ is a positive integer for $1\leq s\leq N$. Then, each block $\boldsymbol{X}_j$ is assumed joint-sparse. The so-called $\ell_{2,1}$-blocks regularized hyperspectral unmixing model is

$$\min_{\boldsymbol{X}} \ \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}\|_F^2+\lambda\sum_{j=1}^{s}\|\boldsymbol{X}_j\|_{2,1} \quad \text{s.t. } \boldsymbol{X}\geq\boldsymbol{0} \tag{5}$$

where $\lambda\geq 0$ is the regularization parameter. Particularly, the model in (5) generalizes both the SUnSAL model by setting $s=N$ and the CLSUnSAL model by setting $s=1$.

Recently, imposing the low-rank property on the abundance matrix provides a new perspective for exploiting spatial information [22], [27], [28]. Simultaneously imposing the sparsity and the low-rank representation on the abundance leads to an unmixing model as follows:

$$\min_{\boldsymbol{X}} \ \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}\|_F^2+\lambda\|\boldsymbol{X}\|_{\boldsymbol{Z},1}+\tau\|\boldsymbol{X}\|_{\boldsymbol{w},*} \quad \text{s.t. } \boldsymbol{X}\geq\boldsymbol{0} \tag{6}$$

where $\lambda$ and $\tau$ are nonnegative regularization parameters

$$\|\boldsymbol{X}\|_{\boldsymbol{Z},1}=\sum_{i=1}^{m}\sum_{j=1}^{N}z_{i,j}|x_{i,j}|, \quad \|\boldsymbol{X}\|_{\boldsymbol{w},*}=\sum_{i=1}^{r}w_i\sigma_i$$

are the weighted $\ell_1$-norm and the weighted nuclear norm of $\boldsymbol{X}$, respectively, $\boldsymbol{Z}=[z_{i,j}]\in\mathbb{R}^{m\times N}$ and $\boldsymbol{w}=[w_1,\ldots,w_r]\in\mathbb{R}^r$ are nonnegative weighting arrays, $\sigma_i$ is the $i$th singular value of $\boldsymbol{X}$, and $r$ is the rank of $\boldsymbol{X}$. The model in (6) can be solved by an algorithm termed as the alternating direction sparse and low-rank unmixing algorithm (ADSpLRU) [28]. Instead simultaneously imposing the joint-sparse-blocks structure and the low-rank representation on the fractional abundance, the unmixing model becomes

$$\min_{\boldsymbol{X}} \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{X}\|_F^2+\lambda\sum_{j=1}^{s}\|\boldsymbol{X}_j\|_{\boldsymbol{z}_j,2,1}+\tau\|\boldsymbol{X}\|_{\boldsymbol{w},*} \quad \text{s.t. } \boldsymbol{X}\geq\boldsymbol{0} \tag{7}$$

where $\boldsymbol{X}_j$ is the $j$th column block of $\boldsymbol{X}$ in (4), $\|\boldsymbol{X}_j\|_{\boldsymbol{z}_j,2,1}$ is the weighted $\ell_{2,1}$-norm of $\boldsymbol{X}_j$ defined as

$$\|\boldsymbol{X}_j\|_{\boldsymbol{z}_j,2,1}=\sum_{i=1}^{m}z_{i,j}\left\|\boldsymbol{x}_j^{[i]}\right\|_2$$

$\boldsymbol{z}_j=[z_{1,j},\ldots,z_{m,j}]^T\in\mathbb{R}^m$ is a nonnegative weighting vector for $\boldsymbol{X}_j$, and $\boldsymbol{x}_j^{[i]}$ is the $i$th row of $\boldsymbol{X}_j$, for $i=1,\ldots,m$ and $j=1,\ldots,s$. The model in (7) is solved by the so-called joint-sparse-blocks and low-rank unmixing algorithm (JSpBLRU) and a reweighting strategy has been introduced to adaptively update $z_{i,j}$ at each iteration [22].

Finally, it is worth mentioning that tensor decomposition-based spectral unmixing has attracted considerable attention in recent years. A matrix-vector nonnegative tensor factorization (MV-NTF) has proposed for hyperspectral unmixing [46]. It approximates a third-order HSI tensor by sum of the outer products of an endmember and its corresponding abundance map, providing another explicit physical explanation of linear spectral mixture. Under the MV-NTF framework, more spatial and spectral prior information has been incorporated to improve the unmixing performance in [10], [47], [48]. Note that it can be strict to impose a low-rank characteristic on the abundance maps and the endmembers in real-world scenarios, so Imbiriba *et al.* [49] proposed a low-rank regularization by introducing a low-dimensional spatial-spectral structure into the abundance and the endmember tensors. In addition, Sun *et al.* [34] imposed a low-rank property on nonlocal third-order abundance tensors and develop a weighted nonlocal low-rank tensor decomposition method for HSI sparse unmixing. It can be observed that compared with matrix factorization-based unmixing algorithms, tensor decomposition-based unmixing algorithms offer stimulating

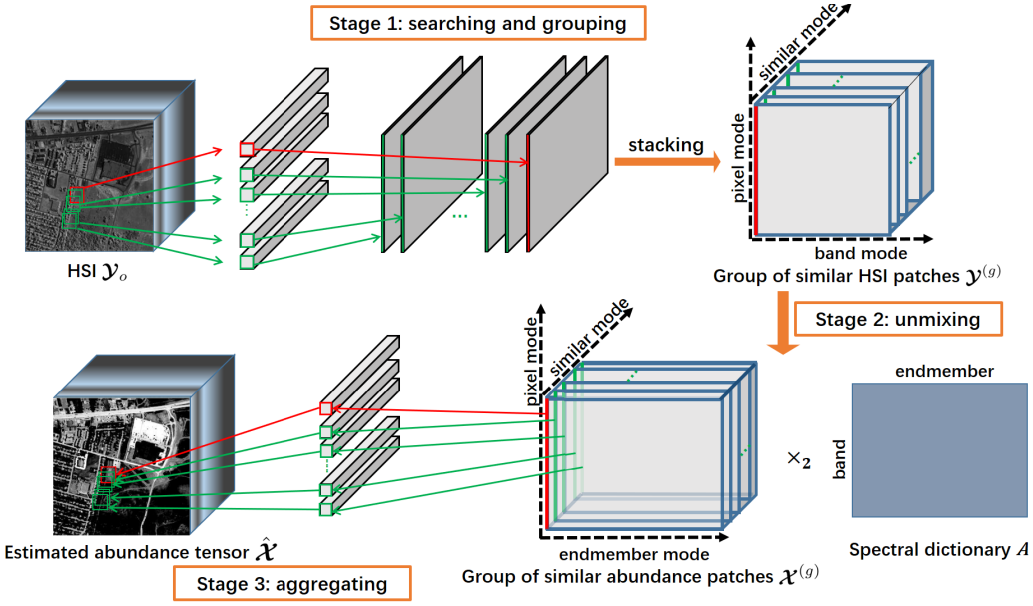



Fig. 1. Flowchart of the NL-TSUn algorithm. It consists of three stages. 1) Search and group similar HSI patches. 2) Unmix each third-order group. 3) Aggregate to the abundance tensor.

results. In the following, we propose a tensor-based unmixing algorithm that promotes the utilization of nonlocal spatial information and spectral correlation in HSIs in a different way in Section III. The advantages will be demonstrated later in Sections IV and V.

## III. NL-TSUn Algorithm

In this section, we will first propose a nonlocal tensor-based unmixing optimization model and then derive an algorithm under the ADMM framework.

### A. Proposed Unmixing Model

Unlike the authors in [29], [31]–[34] utilize the nonlocal spatial information in abundance maps, we consider directly exploiting the nonlocal spatial similarity in HSIs to further improve the abundance estimation performance. The process mainly consists of three stages: 1) group similar patches inside an HSI; 2) unmix each group by a novel tensor-based algorithm; and 3) aggregate to the abundance tensor of the HSI. The overall flowchart of the proposed algorithm is shown in Fig. 1. Specifically, at the first stage, we consider each key fullband patch (FBP), which is stacked by patches at the same location of the HSI over all bands, and there are many FBPs similar to the key FBP. We build a third-order tensor by stacking these similar FBPs, named *group of similar HSI patches* in Fig. 1. It is worth mentioning that utilizing the nonlocal spatial similarity has achieved state-of-the-art performance for the hyperspectral denoising problem [35]–[39]. Then, at the second stage, we unmix each group of similar HSI patches by encoding the nonlocal spatial similarity and spectral correlation in the corresponding abundance tensor. Finally, at the last stage, we aggregate the estimated *groups of similar abundance patches* to reconstruct the abundance tensor for the HSI. With the flowchart in mind, we study the second stage, that is, to unmix each group of similar HSI patches in the following.

Denote $\mathcal{Y}^{(g)} \in \mathbb{R}^{I\times J\times K}$ as the third-order tensor grouped by similar nonlocal HSI patches, where $I$ is the number of pixels, $J$ is the number of spectral bands, and $K$ is the number of similar patches. In the following, we use $\mathcal{Y}$ to denote $\mathcal{Y}^{(g)}$ for notation simplicity. Notice that the mode-2 of the third-order tensor $\mathcal{Y}$ is the spectral mode. Then, the LMM for $\mathcal{Y}$ is written as

$$\boldsymbol{Y}_{(2)} = \boldsymbol{A}\boldsymbol{X}_{(2)} + \boldsymbol{E}_{(2)}$$

which is equivalent to, under the definition of the tensor mode-$n$ multiplication

$$\mathcal{Y} = \mathcal{X} \times_2 \boldsymbol{A} + \mathcal{E} \tag{8}$$

where $\mathcal{X}$ is the $I \times m \times K$ fractional abundance tensor with $m$ being the number of endmembers, $\boldsymbol{A} \in \mathbb{R}^{J\times m}$ is the spectral library, and $\mathcal{E} \in \mathbb{R}^{I\times J\times K}$ is the noise tensor.

In order to exploit the structure information of the abundance tensor $\mathcal{X}$ stacking by similar abundance patches, we show the mode-$n$ unfolding of $\mathcal{X}$ in Fig. 2. From Fig. 2, we see that each mode-$n$ unfolding of $\mathcal{X}$ is low-rank. We remark that the low-rank property of the mode-$n$ unfolding stems from the spatial correlation of pixels within a small size HSI patch for $n = 1$, from the high mutual coherence of the spectral libraries for $n = 2$, and from the nonlocal similarity of the HSI for $n = 3$. Therefore, to exploit the low-rankness characteristic of each mode-$n$ unfolding, we define the following *weighted tensor trace norm* [50]:

$$\|\mathcal{X}\|_{\boldsymbol{w},*} = \sum_{n=1}^{3} \|\boldsymbol{X}_{(n)}\|_{\boldsymbol{w}_n,*} = \sum_{n=1}^{3} \sum_{l=1}^{\operatorname{rank}(\boldsymbol{X}_{(n)})} w_{l,n}\sigma_{l,n}(\boldsymbol{X}_{(n)}) \tag{9}$$

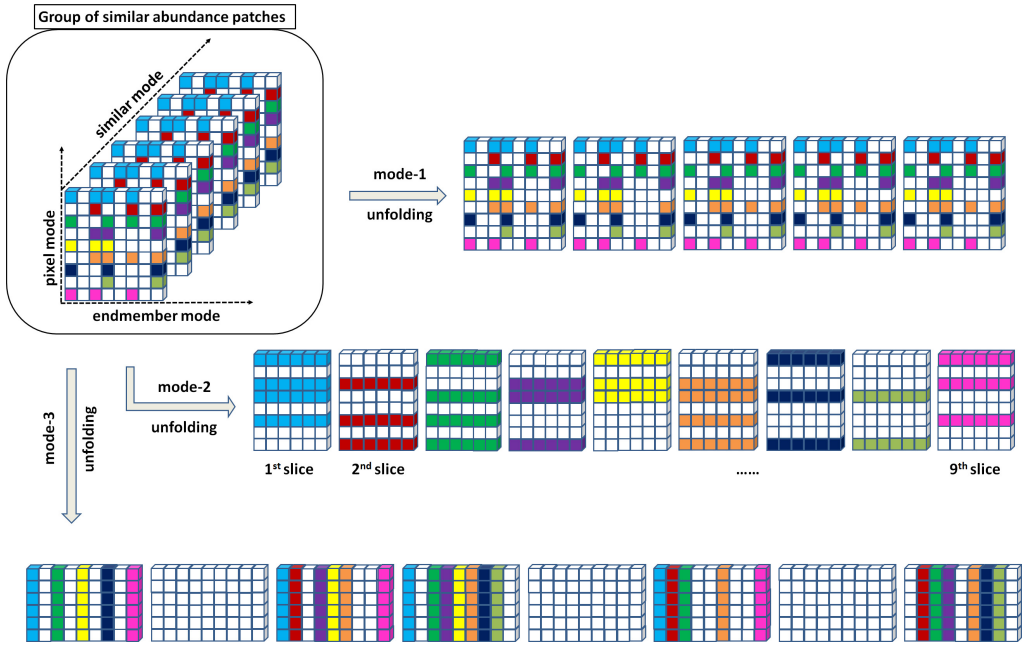



Fig. 2. Illustration of the low-rank characteristic and the joint sparsity in the mode-$n$ unfolding of a group of similar abundance patches.

where $\sigma_{l,n}(\boldsymbol{X}_{(n)})$ is the $l$th singular value of $\boldsymbol{X}_{(n)}$ and $\boldsymbol{w}_n = [w_{1,n}, \ldots, w_{\text{rank}(\boldsymbol{X}_{(n)}),n}]^T$ is a nonnegative weighting vector for $\boldsymbol{X}_{(n)}$. Recall that the value of the weight $w_{l,n}$ is set to 1/3 for any $l$ and $n$ in [50]. Here, we propose to enhance the sparsity of the singular values of $\boldsymbol{X}_{(n)}$ by assigning different weighting coefficients to different singular values, similarly as in [28], [36], and [22]. We will show the details in Section III-B. We note that the weighted nuclear norm has been used to characterize a global sparsity prior inside the tensor data in tensor completion problems [51].

We also see from Fig. 2 that each horizontal slice of the abundance tensor $\boldsymbol{\mathcal{X}}$ admits a joint-sparse property and it therefore leads to a joint-sparse-blocks structure in the mode-2 unfolding matrix. This mainly due to the sparse assumption and the spatial nonlocal similarity in the abundance tensor. It is assumed that in sparse hyperspectral unmixing, only a small portion of the $m$ endmembers in the spectral library $\boldsymbol{A}$ will be present in real HSIs. In other words, it is safe to assume that each abundance vector along the endmember mode in Fig. 2 has a sparse representation in terms of the potentially very large spectral library. Then, the spatial nonlocal similarity promotes that the corresponding abundance vectors of the similar pixels in the same horizontal slice share the same sparsity structure. It, therefore, follows a joint sparse structure in each horizontal slice of the abundance tensor $\boldsymbol{\mathcal{X}}$. Clearly, it is a generalization of the matrix case.

To utilize the joint sparsity of $\boldsymbol{\mathcal{X}}$, we define the following *weighted tensor $\ell_{2,1}$-norm*:

$$\|\boldsymbol{\mathcal{X}}\|_{\boldsymbol{Z},2,1} = \sum_{i=1}^{I} \|\boldsymbol{X}_{i::}\|_{\boldsymbol{z}_i,2,1} = \sum_{i=1}^{I}\sum_{k=1}^{m} z_{k,i} \left\|\boldsymbol{x}_{i::}^{[k]}\right\|_2 \quad (10)$$

where $\boldsymbol{x}_{i::}^{[k]}$ is the $k$th row of $\boldsymbol{X}_{i::}$ and $\boldsymbol{z}_i = [z_{1,i}, \ldots, z_{m,i}]^T$ is the $i$th column of the nonnegative weighting matrix $\boldsymbol{Z} \in \mathbb{R}^{m\times I}$. That said, we impose the joint sparsity on each horizontal slice of $\boldsymbol{\mathcal{X}}$, i.e., $\boldsymbol{X}_{i::} \in \mathbb{R}^{m\times K}$, via the weighted $\ell_{2,1}$-norm. Particularly, weighting coefficients $z_{k,i}$s are assigned differently to each row within each horizontal slice and they will be adaptively updated, similarly as in [22].

Simultaneously imposing the low-rank property via (9) and the joint sparsity via (10) on the abundance tensor $\boldsymbol{\mathcal{X}}$, we obtain the following tensor-based optimization problem:

$$\begin{aligned} &\min_{\boldsymbol{\mathcal{X}}} \ \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}} \times_2 \boldsymbol{A}\|_F^2 + \lambda_1 \|\boldsymbol{\mathcal{X}}\|_{\boldsymbol{Z},2,1} + \lambda_2 \|\boldsymbol{\mathcal{X}}\|_{\boldsymbol{W},*} \\ &\text{s.t. } \boldsymbol{\mathcal{X}} \geq \boldsymbol{0} \end{aligned} \quad (11)$$

where $\lambda_1$ and $\lambda_2$ are nonnegative regularization parameters and $\boldsymbol{Z}$ and $\boldsymbol{W}$ are two nonnegative weighting matrices.

### *B. NL-TSUn Algorithm*

To unmix each group of similar HSI patches $\boldsymbol{\mathcal{Y}}$, we now solve the model in (11). To begin, we let $\iota_\Omega$ be the indicator function of a set $\Omega$, i.e., $\iota_\Omega(x) = 0$ if $x \in \Omega$ and $\iota_\Omega(x) = +\infty$ otherwise. Combining with (9) and (10), the model in (11) becomes

$$\begin{aligned} \min_{\boldsymbol{\mathcal{X}}} \ \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{X}} \times_2 \boldsymbol{A}\|_F^2 + \lambda_1 \sum_{i=1}^{I} \|\boldsymbol{X}_{i::}\|_{\boldsymbol{z}_i,2,1} \\ + \lambda_2 \sum_{n=1}^{3} \|\boldsymbol{X}_{(n)}\|_{\boldsymbol{w}_n,*} + \iota_{\mathbb{R}^+}(\boldsymbol{\mathcal{X}}). \end{aligned} \quad (12)$$

Equivalently, we have

$$\begin{aligned} \min_{\boldsymbol{\mathcal{X}},\boldsymbol{\mathcal{M}}_1,\ldots,\boldsymbol{\mathcal{M}}_6} \ & \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{M}}_1 \times_2 \boldsymbol{A}\|_F^2 + \lambda_1 \sum_{i=1}^{I} \|\boldsymbol{M}_{2,i::}\|_{\boldsymbol{z}_i,2,1} \\ & + \lambda_2 (\|\boldsymbol{M}_{3(1)}\|_{\boldsymbol{w}_1,*} + \|\boldsymbol{M}_{4(2)}\|_{\boldsymbol{w}_2,*} \\ & \quad + \|\boldsymbol{M}_{5(3)}\|_{\boldsymbol{w}_3,*}) \\ & + \iota_{\mathbb{R}^+}(\boldsymbol{\mathcal{M}}_6) \\ \text{s.t. } & \boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{M}}_l, \quad l = 1, \ldots, 6 \end{aligned} \quad (13)$$

where $\boldsymbol{M}_{2,i::}$ is the $i$th horizontal slice of $\boldsymbol{\mathcal{M}}_2$ and $\boldsymbol{M}_{l(n)}$ is the mode-$n$ unfolding of $\boldsymbol{\mathcal{M}}_l$, for $l = 1, \ldots, 6$ and $n = 1, 2, 3$.

We now solve (13) under the ADMM framework. Define

$$\begin{aligned}
&\mathcal{L}_\mu(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{M}}_1, \ldots, \boldsymbol{\mathcal{M}}_6; \boldsymbol{\mathcal{D}}_1, \ldots, \boldsymbol{\mathcal{D}}_6) \\
&= \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{M}}_1 \times_2 \boldsymbol{A}\|_F^2 + \lambda_1 \sum_{i=1}^{I} \|\boldsymbol{M}_{2,i::}\|_{z_i,2,1} \\
&\quad + \lambda_2 \|\boldsymbol{M}_{3(1)}\|_{\boldsymbol{w}_1,*} + \lambda_2 \|\boldsymbol{M}_{4(2)}\|_{\boldsymbol{w}_2,*} + \lambda_2 \|\boldsymbol{M}_{5(3)}\|_{\boldsymbol{w}_3,*} \\
&\quad + \iota_{\mathbb{R}^+}(\boldsymbol{\mathcal{M}}_6) + \frac{\mu}{2} \sum_{l=1}^{6} \|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{M}}_l - \boldsymbol{\mathcal{D}}_l\|_F^2 \qquad (14)
\end{aligned}$$

where $\mu > 0$ is a penalty parameter. The ADMM framework is then derived

$$\begin{cases}
\{\boldsymbol{\mathcal{M}}_1^{t+1}, \ldots, \boldsymbol{\mathcal{M}}_6^{t+1}\} \\
\quad = \underset{\boldsymbol{\mathcal{M}}_1, \ldots, \boldsymbol{\mathcal{M}}_6}{\operatorname{argmin}} \ \mathcal{L}_\mu(\boldsymbol{\mathcal{X}}^t, \boldsymbol{\mathcal{M}}_1, \ldots, \boldsymbol{\mathcal{M}}_6; \boldsymbol{\mathcal{D}}_1^t, \ldots, \boldsymbol{\mathcal{D}}_6^t) \\
\boldsymbol{\mathcal{X}}^{t+1} = \underset{\boldsymbol{\mathcal{X}}}{\operatorname{argmin}} \ \mathcal{L}_\mu(\boldsymbol{\mathcal{X}}, \boldsymbol{\mathcal{M}}_1^{t+1}, \ldots, \boldsymbol{\mathcal{M}}_6^{t+1}; \boldsymbol{\mathcal{D}}_1^t, \ldots, \boldsymbol{\mathcal{D}}_6^t) \\
\boldsymbol{\mathcal{D}}_l^{t+1} = \boldsymbol{\mathcal{D}}_l^t - \boldsymbol{\mathcal{X}}^{t+1} + \boldsymbol{\mathcal{M}}_l^{t+1}, l = 1, \ldots, 6.
\end{cases} \qquad (15)$$

It is easy to see that the minimization problem about $\boldsymbol{\mathcal{M}}_l$, $l = 1, \ldots, 6$, can be decoupled and so we can solve these subproblems separately.

1) For $\boldsymbol{\mathcal{M}}_1$-subproblem, we have

$$\begin{aligned}
&\min_{\boldsymbol{\mathcal{M}}_1} \frac{1}{2}\|\boldsymbol{\mathcal{Y}} - \boldsymbol{\mathcal{M}}_1 \times_2 \boldsymbol{A}\|_F^2 + \frac{\mu}{2}\left\|\boldsymbol{\mathcal{X}}^t - \boldsymbol{\mathcal{M}}_1 - \boldsymbol{\mathcal{D}}_1^t\right\|_F^2 \\
&\quad = \min_{\boldsymbol{M}_{1(2)}} \frac{1}{2}\|\boldsymbol{Y}_{(2)} - \boldsymbol{A}\boldsymbol{M}_{1(2)}\|_F^2 \\
&\qquad + \frac{\mu}{2}\left\|\boldsymbol{X}_{(2)}^t - \boldsymbol{M}_{1(2)} - \boldsymbol{D}_{1(2)}^t\right\|_F^2.
\end{aligned}$$

This is a least squares problem and we obtain

$$\boldsymbol{M}_{1(2)}^{t+1} = (\boldsymbol{A}^T\boldsymbol{A} + \mu\boldsymbol{I})^{-1}\left(\boldsymbol{A}^T\boldsymbol{Y}_{(2)} + \mu\left(\boldsymbol{X}_{(2)}^t - \boldsymbol{D}_{1(2)}^t\right)\right)$$

where $\boldsymbol{I}$ is an $m \times m$ identity matrix. With the fold operator, we have

$$\boldsymbol{\mathcal{M}}_1^{t+1} = \text{fold}_2\left(\boldsymbol{M}_{1(2)}^{t+1}\right). \qquad (16)$$

2) For $\boldsymbol{\mathcal{M}}_2$-subproblem, we have

$$\min_{\boldsymbol{\mathcal{M}}_2} \ \lambda_1 \sum_{i=1}^{I} \|\boldsymbol{M}_{2,i::}\|_{z_i,2,1} + \frac{\mu}{2}\left\|\boldsymbol{\mathcal{X}}^t - \boldsymbol{\mathcal{M}}_2 - \boldsymbol{\mathcal{D}}_2^t\right\|_F^2.$$

Clearly, we can equivalently decouple the above minimization problem to $I$ parts, that is, for $i$th horizontal slice of $\boldsymbol{\mathcal{M}}_2$, we have

$$\min_{\boldsymbol{M}_{2,i::}} \ \lambda_1 \|\boldsymbol{M}_{2,i::}\|_{z_i,2,1} + \frac{\mu}{2}\left\|\boldsymbol{X}_{i::}^t - \boldsymbol{M}_{2,i::} - \boldsymbol{D}_{2,i::}^t\right\|_F^2 \qquad (17)$$

for $i = 1, \ldots, I$. Recall from [52] that each subproblem admits a unique block solution, and the $k$th row of the unique block solution $\boldsymbol{M}_{2,i::}$, i.e., $\boldsymbol{m}_{2,i::}^{[k]}$, of (17), can be written explicitly as

$$\left(\boldsymbol{m}_{2,i::}^{t+1}\right)^{[k]} = \textbf{vect-soft}_{\frac{\lambda_1}{\mu} z_{k,i}}\left(\left(\boldsymbol{x}_{i::}^t\right)^{[k]} - \left(\boldsymbol{d}_{2,i::}^t\right)^{[k]}\right) \qquad (18)$$

for $k = 1, \ldots, m$, $i = 1, \ldots, I$, where $(\boldsymbol{x}_{i::}^t)^{[k]}$ and $(\boldsymbol{d}_{2,i::}^t)^{[k]}$ are the $k$th rows of $\boldsymbol{X}_{i::}^t$ and $\boldsymbol{D}_{2,i::}^t$, respectively, and $\textbf{vect-soft}_\alpha(\cdot)$ is a nonlinear operator defined by

$$\textbf{vect-soft}_\alpha(\boldsymbol{x}) = \boldsymbol{x}\frac{\max\{\|\boldsymbol{x}\|_2 - \alpha, 0\}}{\max\{\|\boldsymbol{x}\|_2 - \alpha, 0\} + \alpha} \qquad (19)$$

for $\forall\ \boldsymbol{x} \in \mathbb{R}^K$ and $\alpha > 0$. Here, similarly as in JSpBLRU, we employ a reweighting strategy to enhance sparsity along rows in $\boldsymbol{M}_{2,i::}^{t+1}$ and for (18), set

$$z_{k,i} = \frac{1}{\left\|\left(\boldsymbol{x}_{i::}^t\right)^{[k]} - \left(\boldsymbol{d}_{2,i::}^t\right)^{[k]}\right\|_2 + \epsilon} \qquad (20)$$

where $\epsilon = 10^{-16}$ is a small constant added to avoid singularities. We note that the reweighting coefficient $z_{k,i}$ is an extension of weights for reweighted $\ell_1$ minimization [53]. It is shown in [22] that the abundance estimation accuracy has been substantially improved by using the reweighting coefficients. We will also show the effectiveness of $z_{k,i}$ in (20) in Section IV-B. Then, we get $\boldsymbol{\mathcal{M}}_2^{t+1}$ by stacking all the horizontal slices $\boldsymbol{M}_{2,i::}^{t+1}$s, whose $k$th row is obtained from (18).

3) To solve the $\boldsymbol{\mathcal{M}}_3$-subproblem, we first give some definitions. For any matrix $\boldsymbol{X}$, denote $r = \text{rank}(\boldsymbol{X})$ and let $\sigma_i$ denote the $i$th singular value of $\boldsymbol{X}$, for $i = 1, \ldots, r$. Denote elementwise $(\cdot)_+ = \max(\cdot, 0)$ and $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{Diag}(\sigma_1, \ldots, \sigma_r)\boldsymbol{V}^T$ be the singular value decomposition (SVD) of $\boldsymbol{X}$. Define the weighted singular value thresholding operator $\textbf{SVT}_{\tilde{\boldsymbol{w}},\beta}(\cdot)$ on $\boldsymbol{X}$ as

$$\begin{aligned}
&\textbf{SVT}_{\tilde{\boldsymbol{w}},\beta}(\boldsymbol{X}) \\
&\quad = \boldsymbol{U}\boldsymbol{Diag}((\sigma_1 - \beta\tilde{w}_1)_+, \ldots, (\sigma_r - \beta\tilde{w}_r)_+)\boldsymbol{V}^T
\end{aligned}$$

with $\tilde{\boldsymbol{w}} = [\tilde{w}_1, \ldots, \tilde{w}_r]$ being a nonnegative weighting vector. We note that the weighting coefficient $\tilde{w}_i = (1/(\sigma_i + \epsilon))$ has designed to increase the punishment on smaller singular values and decrease the punishment on large singular values simultaneously. This technique has been widely used for many practical problems; see, e.g., [22], [28], [51], [54], [55]. Consider the $\boldsymbol{\mathcal{M}}_3$-subproblem now. From (14) and (15), we get that

$$\min_{\boldsymbol{M}_{3(1)}} \ \lambda_2 \|\boldsymbol{M}_{3(1)}\|_{\boldsymbol{w}_1,*} + \frac{\mu}{2}\left\|\boldsymbol{X}_{(1)}^t - \boldsymbol{M}_{3(1)} - \boldsymbol{D}_{3(1)}^t\right\|_F^2. \qquad (21)$$

Then, from [54] and [55], the closed-form solution of (21) is

$$\boldsymbol{M}_{3(1)}^{t+1} = \textbf{SVT}_{\boldsymbol{w}_1, \frac{\lambda_2}{\mu}}\left(\boldsymbol{X}_{(1)}^t - \boldsymbol{D}_{3(1)}^t\right). \qquad (22)$$

Here, similarly as in [54], [28], [55], [22], and [51], we set $\boldsymbol{w}_1 = [w_{1,1}, \ldots, w_{r_1,1}]^T$ based on the singular values of the argument of the **SVT** operator in (22). That is, define $\sigma_{l,1}$ is the $l$th singular values of $\boldsymbol{X}_{(1)}^t - \boldsymbol{D}_{3(1)}^t$ and $r_1$ is the rank of $\boldsymbol{X}_{(1)}^t - \boldsymbol{D}_{3(1)}^t$, then

$$w_{l,1} = \frac{1}{\sigma_{l,1} + \epsilon}, \quad l = 1, \ldots, r_1.$$

Then we have

$$\boldsymbol{\mathcal{M}}_3^{t+1} = \text{fold}_1\left(\boldsymbol{M}_{3(1)}^{t+1}\right). \qquad (23)$$

4) Note that the minimization problems with respect to $\mathcal{M}_4$ and $\mathcal{M}_5$ are similar to that of $\mathcal{M}_3$. Thus, we obtain that

$$\mathcal{M}_4^{t+1} = \text{fold}_2\left(\mathbf{SVT}_{\boldsymbol{w}_2, \frac{\lambda_2}{\mu}}\left(\boldsymbol{X}_{(2)}^t - \boldsymbol{D}_{4(2)}^t\right)\right)$$
$$\mathcal{M}_5^{t+1} = \text{fold}_3\left(\mathbf{SVT}_{\boldsymbol{w}_3, \frac{\lambda_2}{\mu}}\left(\boldsymbol{X}_{(3)}^t - \boldsymbol{D}_{5(3)}^t\right)\right). \quad (24)$$

Here, we compute the weights $\boldsymbol{w}_2 = [w_{1,2}, \ldots, w_{r_2,2}]^T$ and $\boldsymbol{w}_3 = [w_{1,3}, \ldots, w_{r_3,3}]^T$ similarly as $\boldsymbol{w}_1$. That is

$$w_{l,n} = \frac{1}{\sigma_{l,n} + \epsilon}, \quad l = 1, \ldots, r_n \quad (25)$$

in which $\sigma_{l,n}$ is the $l$th singular value of $\boldsymbol{X}_{(n)}^t - \boldsymbol{D}_{n+2,(n)}^t$, the argument of the **SVT** operator in (24), and $r_n$ is the rank of $\boldsymbol{X}_{(n)}^t - \boldsymbol{D}_{n+2,(n)}^t$, for $n = 1, 2, 3$.

5) For $\mathcal{M}_6$-subproblem, we have

$$\mathcal{M}_6^{t+1} = \underset{\mathcal{M}_6}{\text{argmin}}\ \iota_{\mathbb{R}^+}(\mathcal{M}_6) + \frac{\mu}{2}\left\|\mathcal{X}^t - \mathcal{M}_6 - \mathcal{D}_6^t\right\|_F^2.$$

It is easy to obtain that

$$\mathcal{M}_6^{t+1} = \max\left(\mathcal{X}^t - \mathcal{D}_6^t, 0\right). \quad (26)$$

6) For $\mathcal{X}$-subproblem, it is to solve

$$\min_{\mathcal{X}} \frac{\mu}{2}\sum_{l=1}^{6}\left\|\mathcal{X} - \mathcal{M}_l^{t+1} - \mathcal{D}_l^t\right\|_F^2$$

and thus

$$\mathcal{X}^{t+1} = \frac{1}{6}\sum_{l=1}^{6}\left(\mathcal{M}_l^{t+1} + \mathcal{D}_l^t\right). \quad (27)$$

7) Finally, we update Lagrange multipliers $\mathcal{D}_l^{t+1}$, $l = 1, \ldots, 6$, according to (15).

To make it more clear, we summarize the proposed NL-TSUn in the following.

**Algorithm 1** Pseudocode of NL-TSUn

**Input:** $\mathcal{Y}_o$ and $\boldsymbol{A}$
Construct groups of similar HSI patches $\{\mathcal{Y}^{(g)}\}_{g=1}^G$
**for** each group $\mathcal{Y}^{(g)}$ **do**
  //Solve the problem (13) by ADMM
  Initialize: $\mathcal{X}^0$, $\mathcal{D}_l^0$, $\mathcal{M}_l^0$, $l = 1, \ldots, 6$, and set $t = 0$
  Select parameters: $\lambda_1$, $\lambda_2$, $\mu$, maximum iterations
  **while** not convergence **do**:
    Compute $\mathcal{M}_1^{t+1}$ by (16)
    Compute $\mathcal{M}_2^{t+1}$ by (18)
    Compute $\mathcal{M}_3^{t+1}$ by (23)
    Compute $\mathcal{M}_4^{t+1}$ and $\mathcal{M}_5^{t+1}$ by (24)
    Compute $\mathcal{M}_6^{t+1}$ by (26)
    Compute $\mathcal{X}^{t+1}$ by (27)
    Update Lagrange multipliers $\mathcal{D}_l^{t+1}$ by (15)
    $t = t + 1$
  **end while**
  Calculate $\mathcal{X}^{(g)} = \mathcal{X}^{t+1}$
**end for**
Aggregate $\{\mathcal{X}^{(g)}\}_{g=1}^G$ to form the estimated abundance tensor $\hat{\mathcal{X}}$ for $\mathcal{Y}_o$

## IV. Experiments on Simulated Data

In this section, we illustrate the unmixing performance of the proposed NL-TSUn algorithm on simulated hyperspectral data sets. We will compare NL-TSUn with five state-of-the-art algorithms: SUnSAL [11], CLSUnSAL [12], SUnSAL-TV[1] [13], ADSpLRU[2] [28], and JSpBLRU[3] [22]. Our tests were done by using MATLAB R2016a on a PC with 3.7 GHz Intel Core i7 and 64 GB memory. The floating-point precision is $10^{-16}$.

We use the signal-to-reconstruction error (SRE) and the root mean square error (RMSE) to evaluate the performance of unmixing results. They are defined as

$$\text{SRE (dB)} = 10\log_{10}\left(\frac{\frac{1}{N}\sum_{i=1}^N \|\boldsymbol{x}_i\|_2^2}{\frac{1}{N}\sum_{i=1}^N \|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|_2^2}\right)$$
$$\text{RMSE} = \sqrt{\frac{1}{mN}\sum_{i=1}^N \|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|_2^2}$$

where $N$ is the number of pixels, $m$ is the number of endmembers, and $\hat{\boldsymbol{x}}_i$ and $\boldsymbol{x}_i$ are estimated and exact abundance vectors of the $i$th pixel, respectively. Generally, the higher SRE values and lower RMSE values, the high quality of the unmixing results. For all algorithms, the regularization parameters are tuned to get maximum SRE values. We select optimal regularization parameters for all compared algorithms from the following sequence:

$$\{0,\ 0.0005,\ 0.001,\ 0.005,\ 0.01,\ 0.05,\ 0.1,\ 0.5,\ 1,\ 5,\ 10,\ 50\}$$

and choose the optimal penalty parameter $\mu$ from $\{0.001, 0.01, 0.1, 1\}$.

We empirically set the HSI patch size $25 \times 25$, i.e., $I = 625$, and $K = 10$ in NL-TSUn for all the experiments. When the ADMM is applied to solve each group of similar HSI patches in NL-TSUn, we adopt the termination criterion similarly as in ADSpLRU and JSpBLRU. That is, define

$$r^t = \sqrt{\sum_{l=1}^6 \left\|\mathcal{X}^t - \mathcal{M}_l^t\right\|_F^2}, \quad d^t = \mu\sqrt{\sum_{l=1}^6 \left\|\mathcal{M}_l^t - \mathcal{M}_l^{t-1}\right\|_F^2}$$

to be the Frobenius norms of the primal residual and the dual residual at the $t$th ADMM iteration, respectively. Then, we stop the iteration if both the following two conditions hold:

$$r^t \le \zeta, \quad d^t \le \zeta$$

or when the number of iterations reached 500. Here, we set $\zeta = ((3m + J)IK)^{1/2}\zeta^{\text{rel}}$ and the positive relative tolerance $\zeta^{\text{rel}}$ is empirically set to $5{\cdot}10^{-6}$. Similarly as the compared algorithms SUnSAL, CLSUnSAL, SUnSAL-TV, ADSpLRU, and JSpBLRU, we initialize $\mathcal{X}^0 = \mathcal{M}_l^0 = \text{fold}_2((\boldsymbol{A}^T\boldsymbol{A} + \mu\boldsymbol{I})^{-1}\boldsymbol{A}^T\boldsymbol{Y}_{(2)})$ and $\mathcal{D}_l^0 = \boldsymbol{0}$, for $l = 1, \ldots, 6$.

[2]Available online: http://members.noa.gr/parisg/demo_splr_unmixing.zip
[3]Available online: https://liangjiandeng.github.io

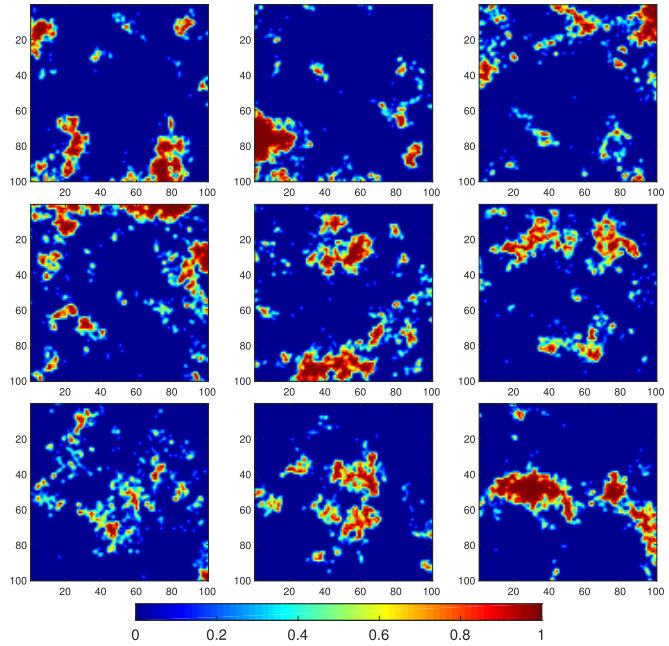



Fig. 3. True abundance maps of selected endmembers for Example 1. (Top Row) Endmembers #1–#3. (Middle Row) Endmembers #4–#6. (Bottom Row) Endmembers #7–#9.

### A. Comparison With Different Algorithms

We now compare NL-TSUn with different unmixing algorithms on the following two widely used simulated data cubes.

*Example 1:* The first data cube has $100 \times 100$ pixels with 100 bands per pixel. The spectral library $\boldsymbol{A}_1 \in \mathbb{R}^{100\times 120}$ is a subset of a library of 262 spectral signatures with 100 spectral bands generally found on satellites, from the National Aeronautics and Space Administration (NASA) Johnson Space Center (JSC) Spacecraft Materials Spectral Database. According to LMM, nine endmember signatures are randomly selected from $\boldsymbol{A}_1$ and the true corresponding abundance maps are shown in Fig. 3. We note that the exact abundance vector for each selected endmember is obtained by stacking the corresponding abundance map. After the above procedure, the true data cube is contaminated by white Gaussian i.i.d. noise with SNR $=$ 20, 30, and 40 dB, respectively.

*Example 2:* The data cube contains $128 \times 128$ pixels with 224 spectral bands. We use the spectral library $\boldsymbol{A}_2 \in \mathbb{R}^{224\times 240}$: a randomly selected subset of the U.S. Geological Survey (USGS) spectral library (splib06a),[4] which comprises 498 spectral signatures with reflectance values measured in 224 spectral bands, distributed uniformly ranging from 0.4 to 2.5 $\mu m$. We randomly choose five endmembers from $\boldsymbol{A}_2$ to generate the true data cube by LMM. For illustrative purposes, corresponding true abundance maps are shown in Fig. 4. Then, the true data cube is contaminated by white Gaussian i.i.d. noise with the same SNR values adopted for Example 1.

For different unmixing algorithms, we show true and estimated abundance maps of endmembers #7 and #9 for Example 1 in Figs. 5 and 6, respectively. Similar results

[4]Available online: http://speclab.cr.usgs.gov/spectral.lib06

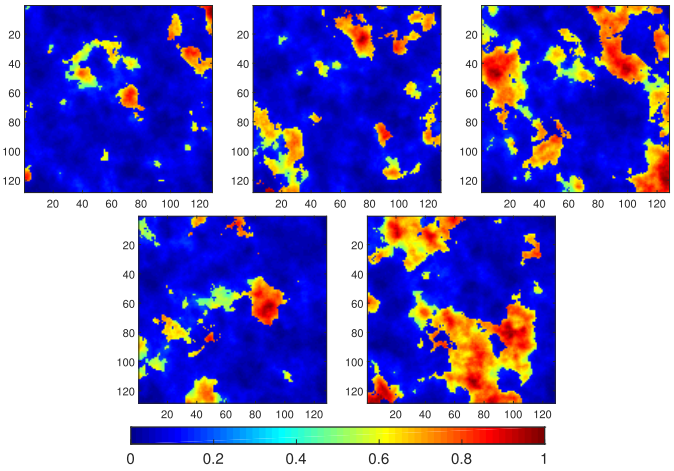



Fig. 4. True abundance maps of selected endmembers for Example 2. (Top Row) Endmembers #1–#3. (Bottom Row) Endmembers #4 and #5.

for endmembers #3 and #5 for Example 2 are shown in Figs. 7 and 8, respectively. Other abundance maps show a similar conclusion, so we omit here for space considerations. In Figs. 5–8, we see that SUnSAL and CLSUnSAL give less accurate estimations for SNR $=$ 20 and 30 dB. It shows that only exploiting sparsity of the abundance is hard to obtain estimations with higher accuracy. We note that, however, each of SUnSAL and CLSUnSAL has only one regularization parameter and therefore costs much less computational time to choose the optimal parameter than other compared algorithms. SUnSAL-TV gives smoother estimations than SUnSAL and CLSUnSAL, but the estimated maps are over-smooth for SNR $=$ 20 dB. ADSpLRU and JSpBLRU both delineate the regions with high fractional abundance, but it is hard for them to remove noise in the background for the high noise level SNR $=$ 20 dB. Clearly, NL-TSUn not only delineates the high fractional abundance regions but also produces a smooth background for all examined noise levels. The advantage is especially clear under heavy noise. This is in line with the observation from the HSI denoising problem that the use of spatial nonlocal similarity helps to achieve better denoising performance. For further comparisons, we list the SRE and RMSE values by different unmixing algorithms in Table II. It shows that NL-TSUn produces the best SRE and RMSE values, which is consistent with the observation from Figs. 5–8.

### B. Reweighting Coefficient Choice and Effectiveness of the Proposed Tensor-Based Regularizers

In the proposed model (11), we impose the low-rank constraint via the weighted nuclear norm on each mode-$n$ unfolding of $\boldsymbol{\mathcal{X}}$ and also impose the joint sparsity via the weighted $\ell_{2,1}$-norm on each horizontal slice of $\boldsymbol{\mathcal{X}}$. In the following, we will first show the efficiency of reweighting coefficients in the tensor-based regularizers and then illustrates the low-rank and joint-sparse characters of the estimated abundance tensors for groups of similar HSI patches.

*1) Reweighting Coefficient Efficiency:* Herein, we aspire to demonstrate the merits emerging from the utilization of the reweighting coefficients $w_{l,n}$ in (25) and $z_{k,i}$ in (20), which are used to enhance the sparsity of the singular values of the

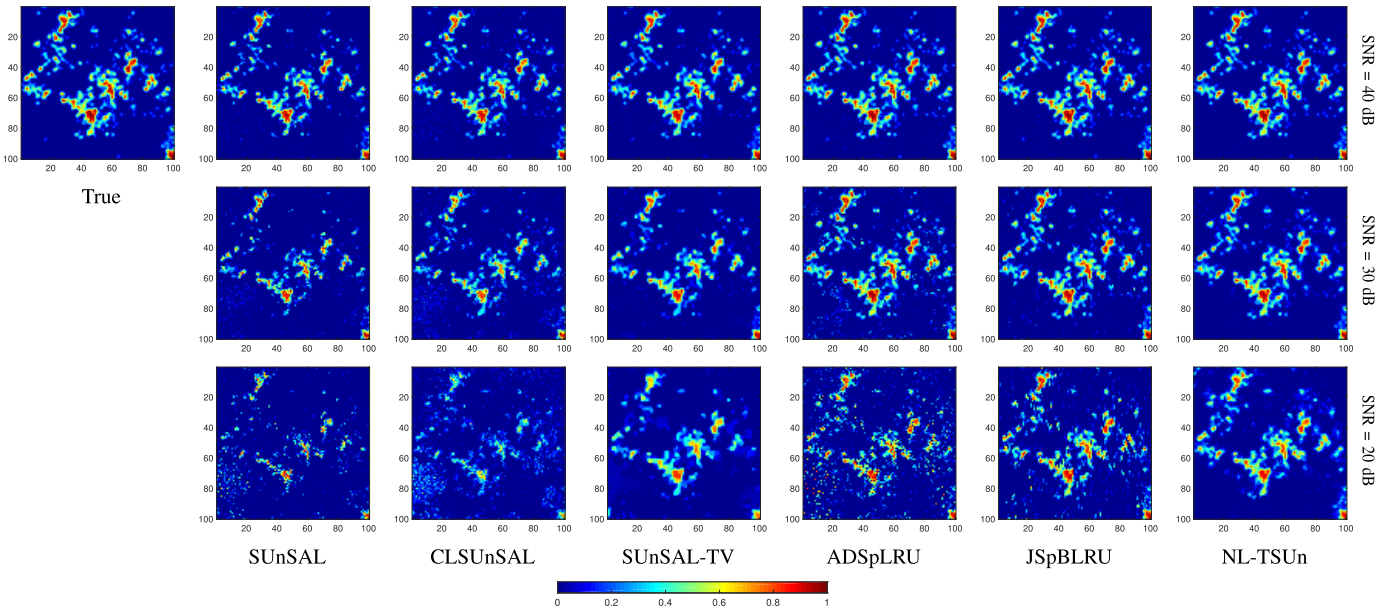



Fig. 5. True and estimated abundance maps for endmember #7 by different unmixing algorithms for Example 1.

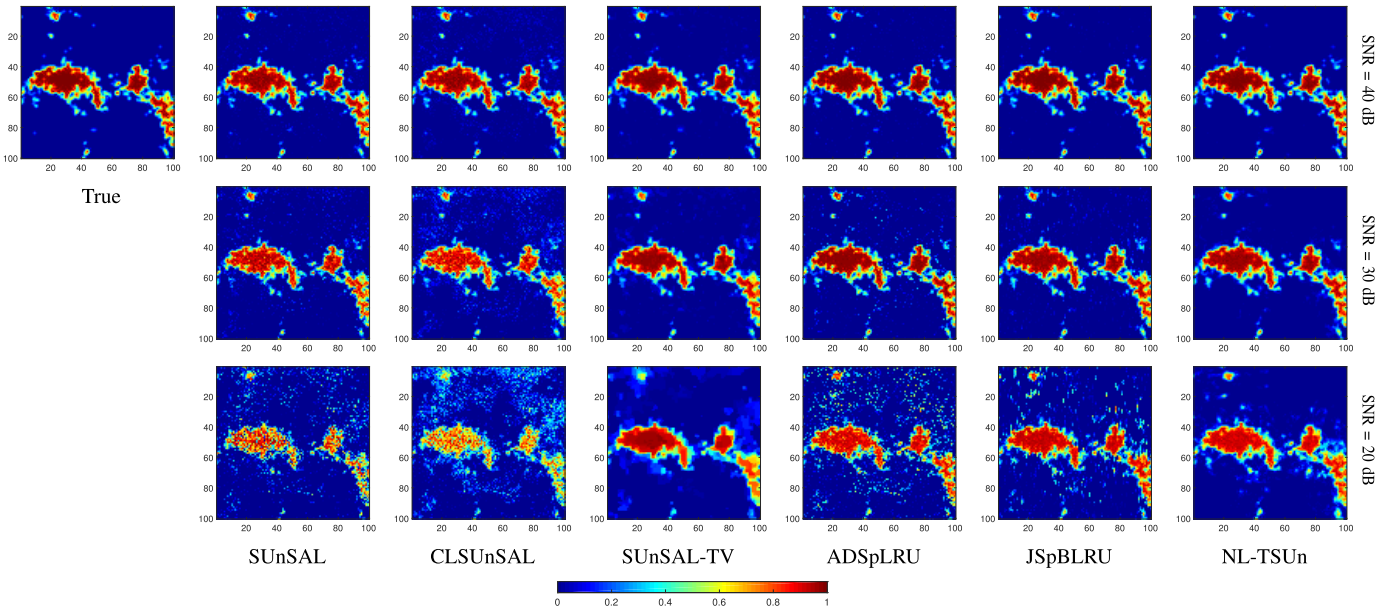



Fig. 6. True and estimated abundance maps for endmember #9 by different unmixing algorithms for Example 1.

mode-$n$ unfolding $\boldsymbol{X}_{(n)}$ and the joint sparsity of the horizontal slice $\boldsymbol{X}_{i::}$, respectively. To this end, four cases are considered as below.

NL-TSUn

1) with reweighting $w_{l,n}$ and $z_{k,i}$.
2) with only reweighting $z_{k,i}$, i.e., $z_{k,i}$ as in (20) but $w_{l,n} = 1$.
3) with only reweighting $w_{l,n}$, i.e., $w_{l,n}$ as in (25) but $z_{k,i} = 1$.
4) without reweighting $w_{l,n}$ and $z_{k,i}$, i.e., $z_{k,i} = w_{l,n} = 1$.

Table III lists the SRE (dB) and the RMSE values of the above four cases for Example 1 with SNR $=$ 20 dB. It shows the efficiency of the reweighting coefficients $z_{k,i}$ and $w_{l,n}$ in NL-TSUn. Recall that in the second stage of NL-TSUn, we unmix each group of similar HSI patches under the ADMM framework. Here, for further illustration purpose, we study the abundance estimation accuracy when unmixing 12 groups of similar HSI patches. In particular, we show the SRE values obtained by NL-TSUn with different choices of weights $z_{k,i}$ and $w_{l,n}$ in Fig. 9. It is clear that NL-TSUn with reweighting both $w_{l,n}$ and $z_{k,i}$ gives optimal SRE values for all considered groups, which is in line with the observation from Table III. We also note that other simulated tests show a similar conclusion so we omit the details for space consideration.

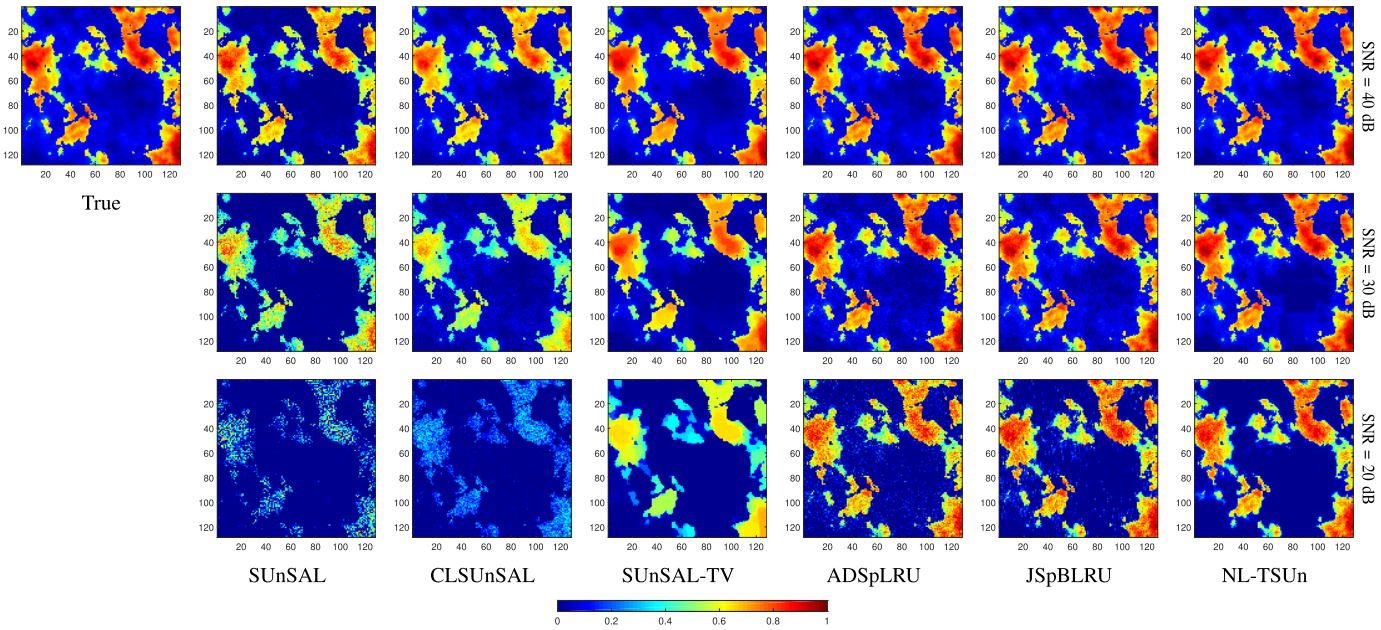



Fig. 7. True and estimated abundance maps for endmember #3 by different unmixing algorithms for Example 2.

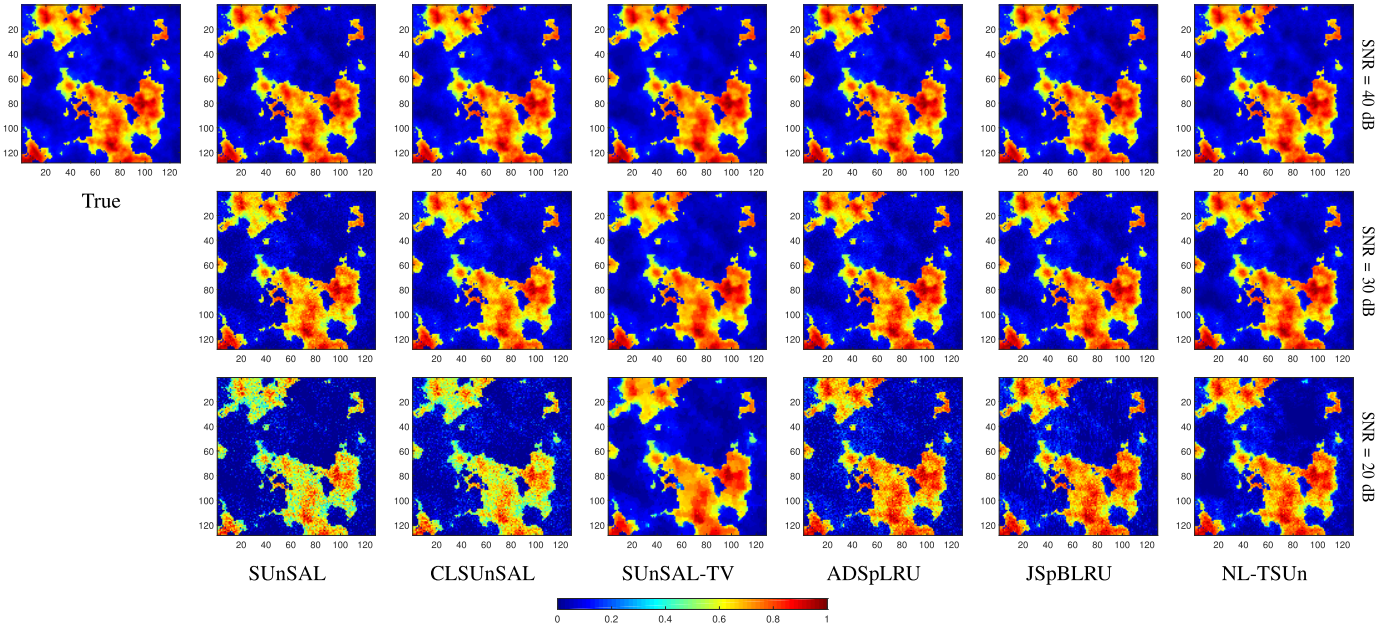



Fig. 8. True and estimated abundance maps for endmember #5 by different unmixing algorithms for Example 2.

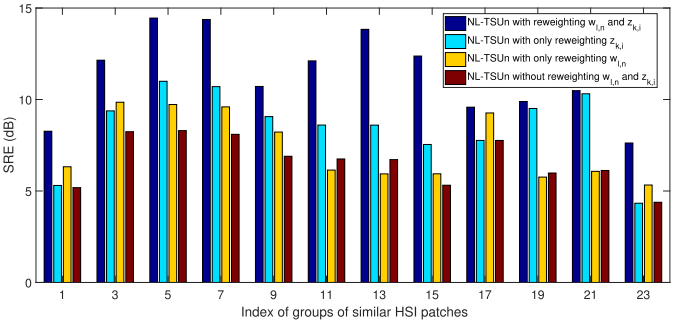



Fig. 9. Performance of NL-TSUn with different weighting coefficients.

*2) Low-Rankness Regularizer:* For the low-rank regularization term, we plot the singular values of the mode-$n$ unfoldings of a true third-order abundance tensor and its estimation by NL-TSUn for Example 1 with SNR $=$ 20 dB in Fig. 10. As shown in the figure that the low-rank property is clear in the unfolding matrices, and moreover, the estimated singular values of each mode-$n$ unfolding are close to the exact ones. It shows the rationality and the effectiveness of the low-rankness constraint on the abundance tensor.

*3) Joint-Sparsity Regularizer:* Considering the joint-sparse regularization, we show eight horizontal slices of a true third-order abundance tensor and its estimations by NL-TSUn for Example 1 with SNR $=$ 20 dB in Fig. 11. Recall that each group of similar abundance patches is a tensor of size $625 \times 120 \times 10$ with $25 \times 25$ pixels and 120 endmembers and ten similar patches. We note that we list eight horizontal

TABLE II
SRE (dB) AND RMSE BY DIFFERENT UNMIXING ALGORITHMS FOR EXAMPLES 1 AND 2

| Example 1 | | | | | | | |
|---|---|---|---|---|---|---|---|
| SNR | Criteria | SUnSAL | CLSUnSAL | SUnSAL-TV | ADSpLRU | JSpBLRU | NL-TSUn |
| 20 dB | SRE | 3.27 | 2.46 | 8.09 | 7.62 | 10.03 | **11.91** |
| | RMSE | 0.0497 | 0.0546 | 0.0285 | 0.0302 | 0.0228 | **0.0183** |
| 30 dB | SRE | 8.25 | 5.90 | 13.26 | 16.59 | 18.37 | **18.75** |
| | RMSE | 0.0280 | 0.0368 | 0.0158 | 0.0107 | 0.0087 | **0.0084** |
| 40 dB | SRE | 13.75 | 11.60 | 19.68 | 26.41 | 28.72 | **29.20** |
| | RMSE | 0.0149 | 0.0191 | 0.0075 | 0.0035 | 0.0027 | **0.0025** |
| Example 2 | | | | | | | |
| SNR | Criteria | SUnSAL | CLSUnSAL | SUnSAL-TV | ADSpLRU | JSpBLRU | NL-TSUn |
| 20 dB | SRE | 4.92 | 5.64 | 12.13 | 13.02 | 14.52 | **15.52** |
| | RMSE | 0.0262 | 0.0241 | 0.0114 | 0.0103 | 0.0087 | **0.0077** |
| 30 dB | SRE | 11.90 | 14.59 | 19.31 | 21.20 | 22.93 | **24.11** |
| | RMSE | 0.0117 | 0.0086 | 0.0050 | 0.0040 | 0.0033 | **0.0029** |
| 40 dB | SRE | 19.07 | 22.49 | 28.24 | 31.66 | 32.73 | **33.44** |
| | RMSE | 0.0051 | 0.0035 | 0.0018 | 0.0012 | 0.0011 | **0.0010** |

TABLE III
SRE (dB) AND RMSE BY NL-TSUN WITH DIFFERENT WEIGHTING COEFFICIENTS FOR EXAMPLE 1 WITH SNR = 20 dB

| Criteria | reweighting $w_{l,n}$ and $z_{k,i}$ | only reweighting $z_{k,i}$ | only reweighting $w_{l,n}$ | no reweighting |
|---|---|---|---|---|
| SRE | 11.91 | 10.85 | 7.76 | 7.17 |
| RMSE | 0.0183 | 0.0208 | 0.0297 | 0.0317 |

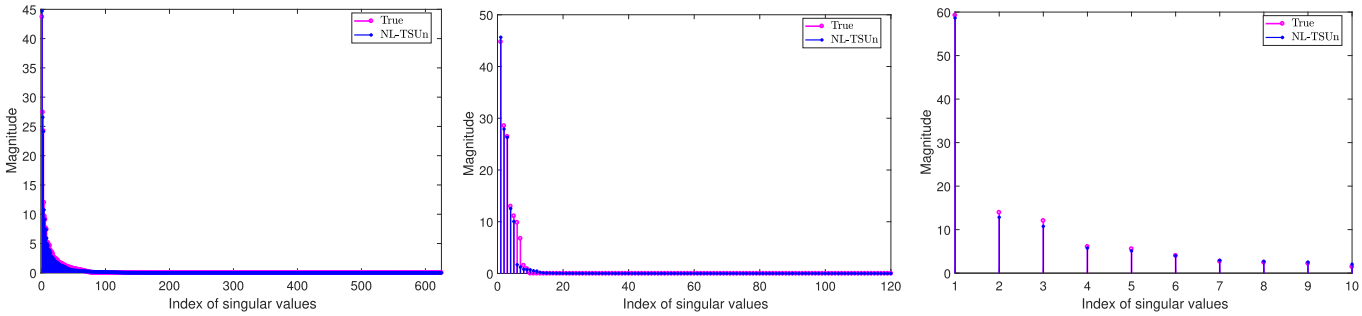



Fig. 10. Singular values of the mode-$n$ unfoldings of a true third-order abundance tensor and its estimation by NL-TSUn for Example 1 with SNR = 20 dB. (a) Mode-1. (b) Mode-2. (c) Mode-3.

slices for the 15th abundance tensor and these horizontal slices are selected corresponding to the fullband pixels located in the first, in the middle, and in the last rows and columns in the HSI patches. The results of other horizontal slices and abundance tensors show a similar conclusion, so we omit them here for space considerations. We clearly see from Fig. 11 that each horizontal slice has respective joint sparsity behavior. NL-TSUn maintains the behavior and provides satisfactory estimations. As a result, both the low-rank property and the joint sparsity characterize the abundance tensor. It therefore highlights the significance of simultaneous incorporation of the low-rank representation and the joint sparsity on the abundance tensors for hyperspectral unmixing problem.

### C. Convergence Analysis

Herein, we numerically show the convergence analysis when the ADMM is applied to solve the optimization problem in (13) for each group of similar HSI patches. We note that it is hard to guarantee the theoretical convergence analysis since the adoption of the reweighting strategies for $w_{l,n}$ and

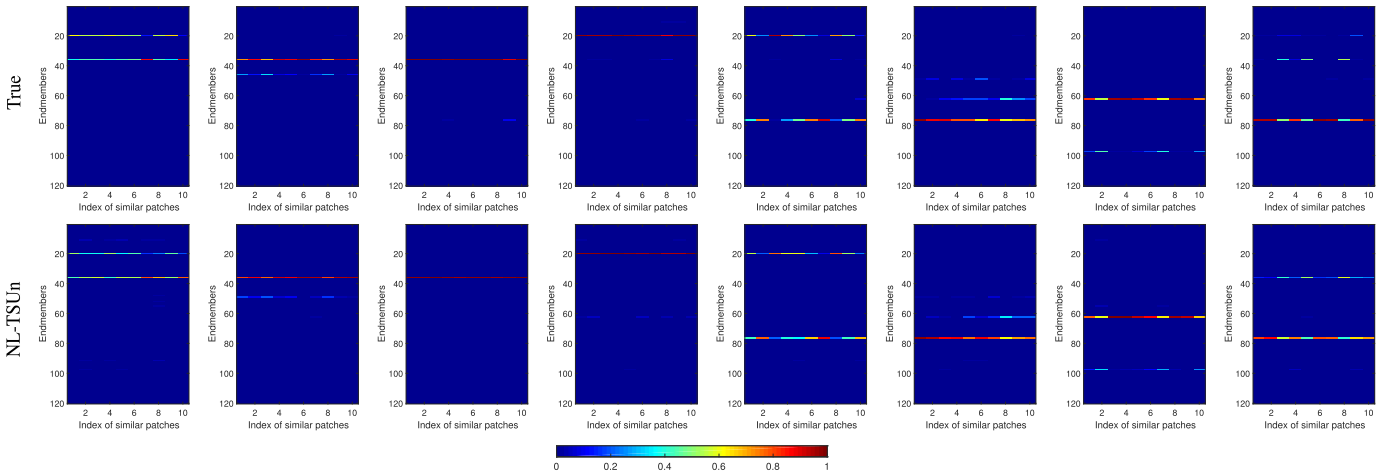



Fig. 11. Horizontal slices of (top row) a true third-order abundance tensor and (bottom row) its estimation by NL-TSUn for Example 1 with SNR = 20 dB. The slice numbers (from left to right) are: 1, 13, 25, 313, 325, 601, 613, and 625.

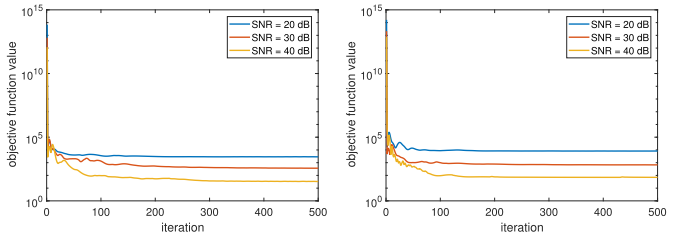



Fig. 12. Objective function value versus iteration for Examples 1 and 2 when the ADMM is applied to solve (13).

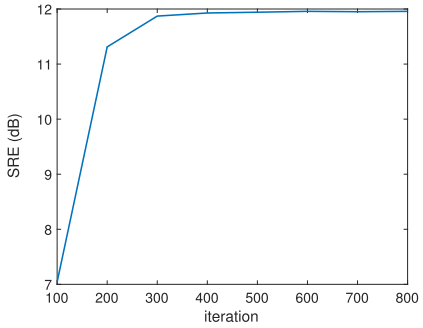



Fig. 13. Plot of SRE (dB) against iteration of NL-TSUn for Example 1 with SNR = 20 dB.

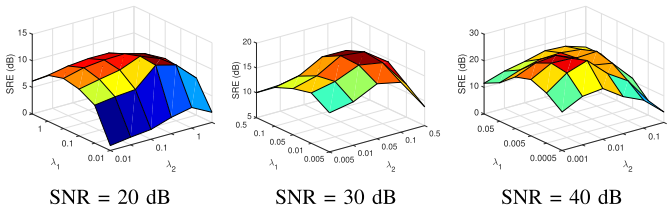



Fig. 14. SRE (dB) as a function of regularization parameters $\lambda_1$ and $\lambda_2$ in NL-TSUn for Example 1.

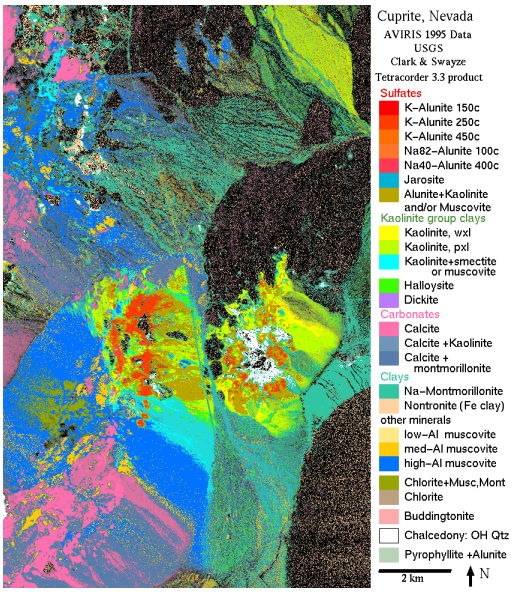



Fig. 15. USGS map showing the location of different minerals in the Cuprite mining district in Nevada.

$z_{k,i}$ renders the model (13) nonconvex. Fig. 12 depicts the objective function value versus iteration for Examples 1 and 2. Here for each noise level, the plot of one group of similar HSI patches is presented. Other groups show similar behavior so we omit them here. We see from Fig. 12 that, despite the fact that the convergence is not yet theoretically guaranteed, the convergence behavior becomes stable as the iteration number increases.

### *D. Parameter Discussion*

In this section, we first test different maximum iteration numbers and then show the role of regularization parameters $\lambda_1$ and $\lambda_2$ in NL-TSUn. Fig. 13 plots the SRE value against iteration for Example 1 with SNR = 20 dB. We see from Fig. 13 that setting the maximum iteration number to be 500 is enough to obtain satisfactory estimations. Fig. 14 plots SRE (dB) values as a function of parameters $\lambda_1$ and $\lambda_2$ for Example 1. We see from the figure that both optimal $\lambda_1$ and $\lambda_2$ decrease as SNR gets higher. Optimal values for

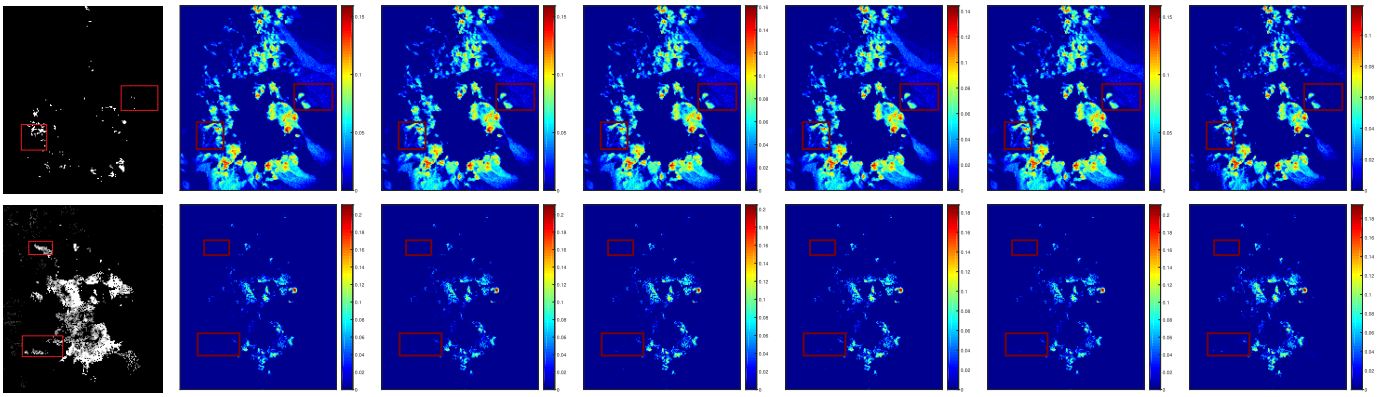


Fig. 16. Qualitative comparison between the classification maps of the AVIRIS Cuprite subscene produced by Tetracorder 4.4 and the abundance maps estimated by different unmixing algorithms for (top row) Alunite and (bottom row) Chalcedony. From left to right: Tetracorder 4.4, SUnSAL, CLSUnSAL, SUnSAL-TV, ADSpLRU, JSpBLRU, and NL-TSUn.

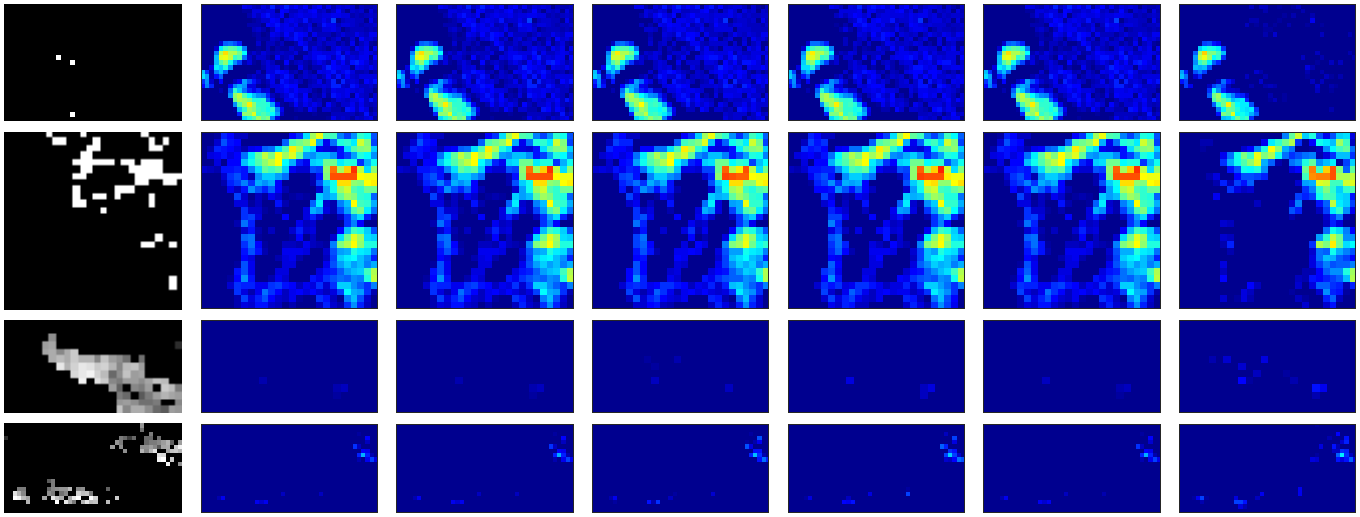


Fig. 17. Classification maps of the regions in red boxes from Fig. 16 for (the first two rows) Alunite and (the last two rows) Chalcedony. From left to right: Tetracorder 4.4, SUnSAL, CLSUnSAL, SUnSAL-TV, ADSpLRU, JSpBLRU, and NL-TSUn.

$\lambda_2$ are greater than or equal to those for $\lambda_1$. We also see that optimal parameter values are positive for all SNR levels, saying again the effectiveness of both low-rank and joint-sparse-based regularizations in NL-TSUn for hyperspectral unmixing.

## V. Experiment on Real Data

In this example, we test different unmixing algorithms on the widely used Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Cuprite data set.[5] We use a subscene containing $200 \times 170$ pixels with 188 spectral bands. The spectral library $\boldsymbol{A}_3 \in \mathbb{R}^{188\times 240}$ is generated from the USGS library which includes all exposed minerals of interest. Fig. 15 shows a mineral map produced in 1995 by USGS. Since the detailed ground truth information is unavailable, we just qualitatively assess obtained abundance maps by different unmixing algorithms, with reference to the Tetracorder 4.4 software product[6] [56]. Similarly as in [11]–[13], [22], [28], the regularization parameters for all algorithms are set to 0.001 except for CLSUnSAL with $\lambda = 0.01$.

Fig. 16 shows estimated abundance maps by different unmixing algorithms and Tetracorder 4.4 software for *Alunite* and *Chalcedony*. Particularly, the regions in red boxes are zoomed in and displayed in Fig. 17. We see from the two figures that all unmixing algorithms delineate the regions considered as respective materials. Nevertheless, NL-TSUn provides fractional abundance maps with more clear backgrounds and captures more spatial details. Generally, by exploiting nonlocal spatial similarity of HSI and imposing the low-rankness constraint and the joint sparsity on fractional abundance tensors, NL-TSUn is effective for unmixing real hyperspectral data.

## VI. Conclusion and Future Work

In this article, we have proposed a nonlocal tensor-based unmixing algorithm. We propose to exploit nonlocal spatial information in the HSI, instead of that in abundance maps. In this vein, we first group similar patches in the HSI and then introduce a novel unmixing model for these groups. Particularly, we introduce the weighted tensor $\ell_{2,1}$-norm to describe

[5]Available online: http://aviris.jpl.nasa.gov/html/aviris.freedata.html
[6]Available online: https://speclab.cr.usgs.gov/PAPERS/tetracorder/

the joint sparsity of third-order abundance tensors, i.e., the groups of similar abundances. Simultaneously, we exploit the low-rank property of abundance tensors through the weighted tensor trace norm. The reweighting strategies have been employed in the weighted norms to enhance the sparsity on singular values and the row sparsity in each horizontal slice. Simulated and real-data experiments have demonstrated the effectiveness of the proposed algorithm, compared with other related state-of-the-art unmixing techniques. The derivation of more computationally efficient nonlocal techniques is currently under investigation.

## ACKNOWLEDGMENT

The authors would like to thank the editors and the three anonymous reviewers for their constructive, detailed, and helpful advice regarding this article.

## 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] 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.

[3] J. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. JPL Airborne Earth Sci. Workshop*, vol. 1, 1995, pp. 23–26.

[4] J. Li and J. Bioucas-Dias, "Minimum vol. simplex, analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE IGARSS*, vol. 3, Jul. 2008, pp. III-250–III-253.

[5] J. M. P. Nascimento and J. M. B. 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.

[6] A. Plaza, G. Martín, J. Plaza, M. Zortea, and S. Sánchez, "Recent developments in endmember extraction and spectral unmixing," in *Proc. Opt. Remote Sens., Adv. Signal Process. Exploitation Techn.*, S. Prasad, L. Bruce, and J. Chanussot, Eds. Berlin, Germany: Springer, 2011, pp. 235–267.

[7] M. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE, Imag. Spectrometry V*, Denver, CO, USA, vol. 3753, Oct. 1999, pp. 266–275.

[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] F. Xiong, Y. Qian, J. Zhou, and Y. Y. Tang, "Hyperspectral unmixing via total variation regularized nonnegative tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2341–2357, Apr. 2019.

[11] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[12] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.

[13] 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.

[14] G. Zhang, Y. Xu, and F. Fang, "Framelet-based sparse unmixing of hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1516–1529, Apr. 2016.

[15] 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.

[16] Y. E. Salehani, S. Gazor, and M. Cheriet, "Sparse hyperspectral unmixing via heuristic $\ell_p$-norm approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1191–1202, Apr. 2018.

[17] J. Sigurdsson, M. O. Ulfarsson, and J. R. Sveinsson, "Hyperspectral unmixing with $\ell_q$ regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6793–6806, Nov. 2014.

[18] Y. Xu, F. Fang, and G. Zhang, "Similarity-guided and $\ell_p$-regularized sparse unmixing of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2311–2315, Nov. 2015.

[19] S. Zhang, J. Li, Z. Wu, and A. Plaza, "Spatial discontinuity-weighted sparse unmixing of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5767–5779, Oct. 2018.

[20] X. Fu, W.-K. Ma, J. M. Bioucas-Dias, and T.-H. Chan, "Semiblind hyperspectral unmixing in the presence of spectral library mismatches," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5171–5184, Sep. 2016.

[21] Z. Shi, T. Shi, M. Zhou, and X. Xu, "Collaborative sparse hyperspectral unmixing using $\ell_0$ norm," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5495–5508, Sep. 2018.

[22] J. Huang, T.-Z. Huang, L.-J. Deng, and X.-L. Zhao, "Joint-sparse-blocks and low-rank representation for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2419–2438, Apr. 2019.

[23] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, Jul. 2013.

[24] X.-L. Zhao, W. Wang, T. Zeng, T.-Z. Huang, and M. K. Ng, "Total variation structured total least squares method for image restoration," *SIAM J. Sci. Comput.*, vol. 35, no. 6, pp. 1304–1320, 2013.

[25] X. Li, J. Huang, L.-J. Deng, and T.-Z. Huang, "Bilateral filter based total variation regularization for sparse hyperspectral image unmixing," *Inf. Sci.*, vol. 504, pp. 334–353, Dec. 2019.

[26] J.-J. Wang, T.-Z. Huang, J. Huang, H.-X. Dou, L.-J. Deng, and X.-L. Zhao, "Row-sparsity spectral unmixing via total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5009–5022, Dec. 2019.

[27] Q. Qu, N. M. Nasrabadi, and T. D. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4404–4423, Jul. 2014.

[28] P. V. Giampouras, K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4775–4789, Aug. 2016.

[29] Y. Zhong, R. Feng, and L. Zhang, "Non-local sparse unmixing for hyperspectral remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1889–1909, Jun. 2014.

[30] R. Feng, Y. Zhong, and L. Zhang, "An improved nonlocal sparse unmixing algorithm for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 915–919, Apr. 2015.

[31] R. Feng, Y. Zhong, and L. Zhang, "Adaptive non-local Euclidean medians sparse unmixing for hyperspectral imagery," *ISPRS J. Photogram. Remote Sens.*, vol. 97, pp. 9–24, Nov. 2014.

[32] R. Wang, H.-C. Li, W. Liao, X. Huang, and W. Philips, "Centralized collaborative sparse unmixing for hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1949–1962, May 2017.

[33] Y. Zheng, F. Wu, H. J. Shim, and L. Sun, "Sparse unmixing for hyperspectral image with nonlocal low-rank prior," *Remote Sens.*, vol. 11, no. 24, p. 2897, Dec. 2019.

[34] L. Sun, F. Wu, T. Zhan, W. Liu, J. Wang, and B. Jeon, "Weighted nonlocal low-rank tensor decomposition method for sparse unmixing of hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 1174–1188, 2020.

[35] Q. Xie *et al.*, "Multispectral images denoising by intrinsic tensor sparsity regularization," in *Proc. IEEE CVPR*, Las Vegas, NV, USA, Jun. 2016, pp. 1692–1700.

[36] J. Xue, Y. Zhao, W. Liao, and S. G. Kong, "Joint spatial and spectral low-rank regularization for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 4, pp. 1940–1958, Apr. 2018.

[37] X. Bai, F. Xu, L. Zhou, Y. Xing, L. Bai, and J. Zhou, "Nonlocal similarity based nonnegative tucker decomposition for hyperspectral image denoising," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 701–712, Mar. 2018.

[38] J. Xue, Y. Zhao, W. Liao, and J. C.-W. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5174–5189, Jul. 2019.

[39] W. He, Q. Yao, C. Li, N. Yokoya, and Q. Zhao, "Non-local meets global: An integrated paradigm for hyperspectral denoising," in *Proc. IEEE CVPR*, Long Beach, CA, USA, Jun. 2019, pp. 6868–6877.
[40] H. A. L. Kiers, "Towards a standardized notation and terminology in multiway analysis," *J. Chemometrics*, vol. 14, no. 3, pp. 105–122, 2000.
[41] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.
[42] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.
[43] J.-H. Yang, X.-L. Zhao, T.-Y. Ji, T.-H. Ma, and T.-Z. Huang, "Low-rank tensor train for tensor robust principal component analysis," *Appl. Math. Comput.*, vol. 367, Feb. 2020, Art. no. 124783.
[44] 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.
[45] L.-J. Deng, M. Feng, and X.-C. Tai, "The fusion of panchromatic and multispectral remote sensing images via tensor-based sparse modeling and hyper-Laplacian prior," *Inf. Fusion*, vol. 52, pp. 76–89, Dec. 2019.
[46] Y. Qian, F. Xiong, S. Zeng, J. Zhou, and Y. Y. Tang, "Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 3, pp. 1776–1792, Mar. 2017.
[47] F. Xiong, J. Chen, J. Zhou, and Y. Qian, "Superpixel-based nonnegative tensor factorization for hyperspectral unmixing," in *Proc. IEEE IGARSS*, Jul. 2018, pp. 6392–6395.
[48] B. Feng and J. Wang, "Constrained nonnegative tensor factorization for spectral unmixing of hyperspectral images: A case study of urban impervious surface extraction," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 583–587, Apr. 2019.
[49] T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, "Low-rank tensor modeling for hyperspectral unmixing accounting for spectral variability," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 3, pp. 1833–1842, Mar. 2020.
[50] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 208–220, Jan. 2013.
[51] J. Xue, Y. Zhao, W. Liao, J. C.-W. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Dec. 24, 2020, doi: 10.1109/TNNLS.2019.2956153.
[52] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.
[53] E. J. Candès, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted $\ell_1$ minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.
[54] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3618–3632, Aug. 2014.
[55] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.
[56] R. N. Clark *et al.*, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS tetracorder and expert systems," *J. Geophys. Res., Planets*, vol. 108, no. E12, pp. 5131–5135, Dec. 2003.

**Jie Huang** (Member, IEEE) received the Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2013.

She is an Associate Professor with the School of Mathematical Sciences, UESTC. Her research interests include remote sensing and image processing.

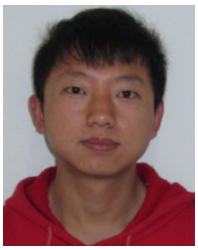
**Ting-Zhu Huang** received the B.S., M.S., and Ph.D. degrees in computational mathematics from the Department of Mathematics, Xian Jiaotong University, Xi'an, China, in 1986, 1992, and 2001, respectively.

He is a Professor with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. His research interests include scientific computation and applications, numerical algorithms for image processing, numerical linear algebra, preconditioning technologies, and matrix analysis with applications.

Dr. Huang is an Editor of *The Scientific World Journal*, *Advances in Numerical Analysis*, the *Journal of Applied Mathematics*, the *Journal of Pure and Applied Mathematics: Advances in Applied Mathematics*, and the *Journal of Electronic Science and Technology*, China.

**Xi-Le Zhao** received the M.S. and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2009 and 2012, respectively.

He is a Professor with the School of Mathematical Sciences, UESTC. His research interest mainly focuses on model-driven and data-driven methods for image processing problems. His homepage is https://zhaoxile.github.io/.



**Liang-Jian Deng** (Member, IEEE) received the B.S. and Ph.D. degrees in applied mathematics from the School of Mathematical Sciences, University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2010 and 2016, respectively.

From 2013 to 2014, he was a Joint-Training Ph.D. Student with Case Western Reserve University, Cleveland, OH, USA. In 2017, he was a Postdoc with Hong Kong Baptist University (HKBU), Hong Kong. In addition, he also stayed with the Isaac Newton Institute for Mathematical Sciences, Cambridge University, Cambridge, U.K., and HKBU for short visits. He is an Associate Professor with the School of Mathematical Sciences, UESTC. His research interests include the use of partial differential equations (PDE), optimization modeling, and deep learning to address several tasks in image processing, and computer vision, e.g., resolution enhancement and restoration.