Efficient Weighted-Adaptive Sparse Constrained Nonnegative Tensor Factorization for Hyperspectral Unmixing

Ping Yang, Ting-Zhu Huang, Jie Huang, Member, IEEE, and Jin-Ju Wang

Abstract—Hyperspectral unmixing aims to separate pure materials and their corresponding proportions that constitute the mixed pixels of hyperspectral imagery (HSI). Recently, the matrix-vector nonnegative tensor factorization (MV-NTF) has attracted wide attention in this field due to its natural third-tensor representation of HSI. However, the NTF-based unmixing approaches are limited to the nonunique solution and long computational time. To solve these problems, we consider the low-rank and sparse priors of each abundance map simultaneously and propose a new unmixing model adopting a weighted nuclear norm and an $L_{1/2}$ norm under the MV-NTF framework. Instead of using low-rank matrix decomposition of MV-NTF, this model imposes the low-rank property on the whole abundance map, which avoids determining the rank of the abundance map in advance. Observing that each abundance map is different, we build an adaptive update mechanism to treat each low-rank and sparse constraint differently in the model. Furthermore, a new multiplicative iterative algorithm is designed to solve the proposed model. Specially, the algorithm designed for this tensor model is simplified by using the equivalence relation with nonnegative matrix factorization. Experiments demonstrate that the proposed method is effective in improving both the unmixing effect and the solving speed.

Index Terms—Hyperspectral unmixing (HU), low-rank property, nonnegative tensor factorization (NTF), sparsity.

I. INTRODUCTION

HYPERSPECTRAL imagery (HSI) is 3-D image data with rich spatial and spectral information, which has been widely studied and applied in many fields [1], [2], [3], [4]. However, due to the limitation of low spatial resolution and the complex ground distribution [5], the pixels in HSI are mostly consisted of mixtures of several different materials. Therefore, hyperspectral unmixing (HU), which decomposes mixed pixels into several pure materials (i.e., endmembers) and their corresponding fractions (i.e., abundances), is of great significance in improving the utility of HSI.

Generally, unmixing can be categorized as linear [2] or non-linear, including bilinear [6], [7] and multilinear [8] unmixing, relied on different interpretations of forming the mixed pixels in HSI during photon propagation. Linear mixing model (LMM) assumes that photons arriving at the sensor only interact with one material so that the pixels in HSI are linear combinations of endmembers, weighted by their corresponding abundances. The structure of LMM is simple and tractable, thus a variety of unmixing methods adopting it have emerged [9], [10], [11]. One kind of LMM-based methods is in a semisupervised fashion, which needs to get the information of endmembers in advance. Geometry-based methods have been studied to extract endmembers from HSI, such as pixel purity index [12], N-FINDR [13], and vertex component analysis (VCA) [14]. Then, the extracted endmembers are use to invert the abundances by fully constrained least squares (FCLS) [15]. These methods are premised on the presence of pure pixels in HSI. Meanwhile, the performance of extracted endmembers also affects the accuracy of subsequent abundance estimation. Sparse unmixing is another kind of semisupervised method [16], [17], which requires that the spectral library covering all endmembers in HSI is available.

Instead, there is not so much information available in real HSI scenes, thus the unsupervised methods of considering unmixing as the blind source separation came into being. Typically, nonnegative matrix factorization (NMF) [18] decomposes HSI data into the product of two nonnegative matrices representing endmember and abundance, respectively. Because of its nonnegativity and clear physical meaning, NMF and its extensions have been widely used for HU [19], [20], [21]. Unfortunately, the objective functions of NMF-based methods are usually nonconvex, which tend to fall into the local optimum. For remedying this defect, an effective way is to introduce the priors of HSI as the regular constraints of NMF. The common abundance constraints include the abundance nonnegative constraint (ANC), and the abundance sum-to-one constraint (ASC). Furthermore, sparsity, the low-rank property, and smoothness of abundances have been taken into account. Sparsity results from the fact that pixels are assumed to be mixed by a subset of endmembers so that a large number of entries in the corresponding abundance matrix are zero. Some works utilize $L_0$, $L_1$, and $L_{1/2}$ norms under the NMF.
framework to enhance the sparsity of abundances [22], [23], [24], [25]. Owing to the spatial correlation among the spectral signatures of pixels, the corresponding abundance vectors are linear dependent. This fact reveals a low-rank behavior in the abundance matrix, which is often characterized by minimizing the nuclear norm [26], [27]. Besides, HSI data always vary smoothly in spatial domain, so a total variation (TV) term is exploited to promote the smoothness of abundances [28]. Moreover, there are improved NMF methods by imposing the minimum volume [29], minimum dispersion [30], or endmember dissimilarity [31] constraints upon endmembers.

These regularized NMF methods consider internal properties of HSI under matrix factorization framework. Additionally, tensor is a high-dimensional extension of matrix [32]. It is a more natural representation of HSI by directly expressing it as a third-order tensor. Therefore, more and more research works consider HU as a nonnegative tensor factorization (NTF) problem [33], [34], [35], [36]. At present, there are three typical tensor factorizations: canonical polyadic decomposition [37], Tucker decomposition [38], and matrix-vector nonnegative tensor factorization (MV-NTF) [39]. Among them, MV-NTF is based on block term decomposition (BTD) [40], which regards HSI as the sum of finite third-order tensors. And each component of it is generated by the outer product of a matrix (i.e., abundance) and a vector (i.e., endmember). This model satisfies the linear spectral mixing mechanism and has attracted much attention. It is shown in [39] that the unmixing performance of MV-NTF is superior to that of the advanced NMF-based methods. However, HSI data are susceptible to noise and one limitation of MV-NTF is the lack of constraints to ensure that HSI is correctly decomposed into endmembers and abundances. Thus, Xiong et al. [41] apply the TV term to abundance maps to suppress the influence of noise. Feng et al. [42] integrate sparsity, minimum volume, and nonlinearity constraints into MV-NTF. Li et al. [43] couple NMF into MV-NTF, and impose the sparse constraint on the abundance matrix. Zheng et al. [44] add the nuclear norm and the \( L_1 \) norm to enhance the low-rank property and sparsity of abundance maps, respectively. Wang et al. [45] construct an NTF-based unmixing method with an independent constraint of endmembers. All above-mentioned methods promote the unmixing performance. However, it is still necessary to explore more detailed information to establish a more effective unmixing approach.

Recently, owing to the powerful learning and data fitting ability of deep learning strategy, various state-of-the-art learning models have been adopted for hyperspectral remote sensing [46], [47], [48]. Particularly, deep learning-based unmixing methods have developed rapidly, including using autoencoder [49], convolutional neural network [50], [51], or generative adversarial network [52], etc. In addition, they also combine additional networks to extract the characteristics of endmembers and abundances to promote unmixing performance, such as EGU-Net [53] and SeCoDe [54]. Deep learning-based unmixing methods have good applicability to the actual unmixing scene through data training. In view of data dependence and running time, however, we mainly study the traditional unmixing methods here.

Although MV-NTF considers the spatial information of HSI, the prior information of endmembers and abundances has not been fully utilized. Therefore, the focus of this study is to make full use of the spatial-spectral characteristics to obtain endmembers and abundances efficiently on the premise of protecting HSI structure. In this article, we propose a new nonnegative tensor unmixing method named efficient weighted-adaptive sparse constrained nonnegative tensor factorization (EWSP-NTF). Instead of using matrix factorization to impose a strict low-rank constraint on abundance maps in MV-NTF, we adopt a weighted nuclear norm regularization to constrain the low-rank property, which avoids estimating the rank of abundance map and preserves the details of HSI. Fig. 1 shows the low-rank and sparse characteristics of each abundance map under the MV-NTF framework. It can be seen that each abundance map has different low-rank and sparse properties. Inspired by this, we introduce the low-rank and sparse regularizations with self-adaptive weights. Differing from [43] and [44], the low-rank property and sparsity of each abundance map in this article are distinctively constrained conforming to the actual HSI scene. In lineage [45], it is effective to treat the low-rank property of each abundance map differently by minimizing the weighted nuclear norm, and we further consider the sparsity of abundance maps to improve the unmixing performance. In addition, both NMF and MV-NTF conform to the linear mixing mechanism. They are convertible when the abundance matrix of NMF is reshaped into abundance maps of MV-NTF. We take this link in the optimization process. The main contributions of this work are summarized as follows.

1) Due to the prior information that each abundance map of HSI has different low-rank and sparse properties, the weighted-adaptive low-rank and sparse regularizers are incorporated into the MV-NTF framework. Unlike the low-rank matrix decomposition used in MV-NTF, we introduce the weighted nuclear norm to constrain the low-rank property of each abundance map. And we use the \( L_{1/2} \) norm to promote the sparsity of each abundance map simultaneously. Moreover, the adaptive weights are taken to treat each low-rank and sparse constraint differently in our model, which is more consistent with the real scene of HSI and more inclined to obtain the correct solution.

2) We adopt an augmented multiplicative algorithm to solve the proposed optimization problem, which introduces auxiliary variables to make the variables independent and so the subproblems are easy to solve. In order to simplify the operation of tensor decomposition, we take advantage of the equivalence relation between MV-NTF and NMF. We design the optimization algorithm under NTF framework combined with NMF, which has high computational efficiency.

The rest of this article is organized as follows. Section II reviews the basic knowledge of tensor, LMM mechanism, NMF, and MV-NTF models. Section III introduces our proposed unmixing model, solving and updating processes in detail. Section IV reports the experimental results on synthetic and real-world datasets. Section V gives some conclusions and the prospects of future work.
YANG et al.: EFFICIENT WEIGHTED-ADAPTIVE SPARSE CONSTRAINED NONNEGATIVE TENSOR FACTORIZATION FOR HU

II. PRELIMINARIES

A. NOTATIONS AND DEFINITIONS

Some notations and definitions of tensor algebra involved in latter work are introduced here. Tensor is a multidimensional array. In this article, scalars which are zero-order tensors are denoted by lowercase letters, e.g., \( x \). Vectors which are first-order tensors are denoted by boldface lowercase letters, e.g., \( \mathbf{x} \). Matrices that are second-order tensors are denoted by script letters, e.g., \( \mathcal{X} \). Tensors that have three or more orders are identified using Euler notation. Notations and Definitions

Definition 1 (Mode, Fiber, and Slice): The dimension of a tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N} \) refers to the matricization of the tensor, which arranges the \( n \)th mode fibers to be the columns of the resulting matrix and is denoted as \( \mathbf{X}_{(n)} \). For a third-order tensor \( \mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3} \), there are three matrix unfolded forms, expressed as

\[
\begin{align*}
(Y_{(1)})_{i_1,j_2,j_3} & = Y_{i_2,j_3} \\
(Y_{(2)})_{i_1,j_2,j_3} & = Y_{i_1,j_3} \\
(Y_{(3)})_{i_1,j_2,j_3} & = Y_{i_1,j_2}
\end{align*}
\]

where each of them is transformed from \( I_1, I_2, \) and \( I_3 \) modes, respectively.

Definition 2 (Mode-n Unfolding): The mode-\( n \) unfolding of a tensor \( \mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N} \) refers to the matricization of the tensor, which arranges the \( n \)th mode fibers to be the columns of the resulting matrix and is denoted as \( \mathbf{X}_{(n)} \). For a third-order tensor \( \mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3} \), there are three matrix unfolded forms, expressed as

\[
\begin{align*}
(Y_{(1)})_{i_1,j_2,j_3} & = Y_{i_2,j_3} \\
(Y_{(2)})_{i_1,j_2,j_3} & = Y_{i_1,j_3} \\
(Y_{(3)})_{i_1,j_2,j_3} & = Y_{i_1,j_2}
\end{align*}
\]

where each of them is transformed from \( I_1, I_2, \) and \( I_3 \) modes, respectively.

Definition 3 (Outer Product): The outer product of two tensors \( \mathbf{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_P} \) and \( \mathbf{B} \in \mathbb{R}^{J_1 \times J_2 \times \ldots \times J_Q} \) obtains a higher order tensor \( \mathbf{F} \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_P \times J_3 \times J_2 \times \ldots \times J_Q} \), denoted as

\[
\mathbf{F} = \mathbf{A} \circ \mathbf{B}
\]

where each entry of \( \mathbf{F} \) is defined as the product of corresponding entries in \( \mathbf{A} \) and \( \mathbf{B} \).
its expression is

\[ \mathbf{y}_i = \sum_{r=1}^{R} \mathbf{e}_r \mathbf{c}_r + \mathbf{n}_i = \mathbf{C} \mathbf{e}_i + \mathbf{n}_i \]  

(1)

where \( \mathbf{C} = [c_1, c_2, \ldots, c_R] \in \mathbb{R}^{K \times R} \), \( \mathbf{e}_i = [e_{1i}, \ldots, e_{ri}, \ldots] \), \( \mathbf{e}_r \in \mathbb{R}^R \) denote the endmember matrix composed of the spectral signatures of \( R \) endmembers, the abundance vector containing the proportions of all endmembers in each pixel, and the noise vector, respectively. Considering \( y_i \) is one of the \( P \) pixels in the matrix \( \mathbf{Y} \in \mathbb{R}^{K \times P} \), the mixing model (1) can be rewritten as

\[ \mathbf{Y} = \mathbf{CE} + \mathbf{N} \]  

(2)

where \( \mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_i, \ldots, \mathbf{e}_P] \in \mathbb{R}^{R \times P} \) and \( \mathbf{N} \in \mathbb{R}^{K \times P} \) denote the abundance matrix containing the abundance fractions of all pixels and the noise matrix, respectively.

In the physical sense of unmixing, all above components are nonnegative. Besides, the abundance fractions of each pixel in all endmembers should be summed to one. Thus, the NMF optimization problem can be described as

\[
\min_{\mathbf{C}, \mathbf{E}} \frac{1}{2} \| \mathbf{Y} - \mathbf{CE} \|^2_F \\
\text{s.t. } \mathbf{E} \geq 0, \mathbf{C} \geq 0, 1^T_R \mathbf{E} = 1^T_P
\]  

(3)

where \( \| \cdot \|_F \) represents the Frobenius norm, and \( 1 \in \mathbb{R}^\square \) denotes an array with all ones, in which the subscript indicates its size.

Unlike NMF transforming the HSI cube into 2-D matrix for processing, MV-NTF is a decomposition model that directly builds on tensors. This model regards HSI \( \mathcal{Y} \in \mathbb{R}^{I \times J \times K} \) as a third-order tensor and decomposes the tensor into the sum of \( R \) component tensors. Each component tensor is defined as the outer product of a matrix (i.e., abundance) and a vector (i.e., endmember). It is formulated as

\[ \mathcal{Y} = \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r + \mathcal{N} = \sum_{r=1}^{R} \mathbf{A}_r \mathbf{B}_r^T \circ \mathbf{c}_r + \mathcal{N} \]  

(4)

where \( \mathbf{c}_r \in \mathbb{R}^K \) denotes the \( r \)-th endmember vector; \( \mathbf{E}_r \in \mathbb{R}^{I \times J} \) denotes its corresponding abundance map, which can be made up of two low-rank matrices \( \mathbf{A}_r \in \mathbb{R}^{I \times L_r} \) and \( \mathbf{B}_r \in \mathbb{R}^{L_r \times J} \) with \( L_r \leq \min(I, J) \); \( \mathcal{N} \in \mathbb{R}^{I \times J \times K} \) stands for the noise tensor.

Similarly, considering the ANC and ASC in the model, the MV-NTF optimization model is

\[
\min_{\mathbf{A}, \mathbf{B}, \mathbf{c}, \mathbf{e}} \frac{1}{2} \| \mathcal{Y} - \sum_{r=1}^{R} (\mathbf{A}_r \mathbf{B}_r^T) \circ \mathbf{c}_r \|^2_F \\
+ \frac{\delta}{2} \| 1_{I \times J} - \sum_{r=1}^{R} \mathbf{A}_r \mathbf{B}_r^T \|^2_F \\
\text{s.t. } \mathbf{A}_r \geq 0, \mathbf{B}_r \geq 0, \mathbf{c}_r \geq 0
\]  

(5)

where \( \delta \) is a tradeoff parameter reflecting the impact of the ASC.

We note that MV-NTF conforms to the LMM in a tensor form. There is an equivalent numerical relationship between (2) and (4), when taking \( \mathbf{E}_r \) as a whole. The \( \mathbf{Y} \) in (2) is just the mode-3 unfolding \( \mathcal{Y} \) of the tensor \( \mathcal{Y} \) when \( P = I \times J \). Besides, the abundance matrix \( \mathbf{E} \) in (2) can be reblocked and recorded as \( \mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_r, \ldots, \mathbf{e}_P]^T \in \mathbb{R}^{R \times P} \) along row, then the vector \( \mathbf{e}_r \in \mathbb{R}^P \) is reshaped by the abundance map \( \mathbf{E}_r \) in (4). This link will be used to simplify the solving of the tensor model proposed later.

III. EFFICIENT WEIGHTED-ADAPTIVE SPARSE CONstrained NONNEGATIVE TENSOR FACTORIZATION

A. Motivation

MV-NTF is a special case of BTD, and its solution is hard to be unique unless strict conditions are met [55]. Therefore, introducing the priors of data as constraints is a direct and effective way to reduce the solution space and get the unmixing results physically meaningful. There are low-rank and sparse characteristics of abundance maps under the MV-NTF framework.

We take Jasper Ridge data as an example. As shown in Fig. 1, the low-rank property and sparsity of each abundance map are reflected by counting the number of nonzero singular values and nonzero entries, respectively. It is well known that the number of nonzero singular values is the rank of the matrix. As can be seen from the marked points in the figure, the number of singular values does not exceed or approach half when the cumulative probability of singular values reaches 95%. It shows a low-rank behavior of each abundance map. Likewise, when the cumulative probability of pixel values reaches 95%, there are only less than half of nonzero entries. It shows that most entries in the abundance map are zero conversely, which indicates a sparsity of the abundance map. Furthermore, Fig. 1 shows that each abundance map corresponds to a different value when the cumulative probability distribution reaches 95%. It reflects that the low-rank property and sparsity of each abundance map are not completely the same in degree. For making full use of this detailed abundance information, we introduce the weighted-adaptive low-rank and sparse regularizers in our model. Besides, the forceful low-rank constraint of MV-NTF is replaced to protect the details of HSI.

B. Proposed EWSP-NTF Model

Based on above facts, we adopt a weighted nuclear norm and an \( L_{1/2} \) norm with adaptive weights to constrain the characteristics. The proposed model is established as follows:

\[
\min_{\mathbf{E}_r, \mathbf{c}_r} \frac{1}{2} \| \mathcal{Y} - \sum_{r=1}^{R} \mathbf{E}_r \circ \mathbf{c}_r \|^2_F + \frac{\delta}{2} \sum_{r=1}^{R} \| \mathbf{E}_r - 1_{I \times J} \|^2_F \\
+ \lambda \sum_{r=1}^{R} \alpha_r \| \mathbf{E}_r \|_{w, r, \alpha} + \rho \sum_{r=1}^{R} \beta_r \| \mathbf{E}_r \|_{1/2} \\
\text{s.t. } \mathbf{E}_r \geq 0, \mathbf{c}_r \geq 0
\]  

(6)

where \( 1_{I \times J} \) denotes an \( I \times J \) matrix with all entries as one; \( \delta, \lambda, \) and \( \rho \) are parameters of regularizations; \( \alpha_r \) and \( \beta_r \) are self-adaptive parameters, which vary with the low-rank property
and sparsity of each abundance map. The definitions of weighted nuclear norm and $L_{1/2}$ norm are given as follows.

For a matrix $X \in \mathbb{R}^{I_1 \times I_2}$, the rank constraint is usually replaced by the nuclear norm that represents the sum of singular values, i.e., $\|X\|_* = \sum_i \sigma_i(X)$, where $\sigma_i(X)$ is the $i$th singular value of $X$. Since the larger singular value indicates the more important information in image data, the weighted nuclear norm is introduced to treat singular values differently [56]. It is defined as

$$
\|X\|_{w,*} = \sum_i w_i \sigma_i(X)
$$

(7)

where $w_i$ is the nonnegative weight coefficient of $\sigma_i(X)$. The $L_{1/2}$ norm is used to describe the sparsity of data [24], and it is defined as

$$
\|X\|_{1/2} = \left( \sum_{i_1, i_2 = 1}^{I_1, I_2} (x_{i_1, i_2})^2 \right)^{1/2}
$$

(8)

where $x_{i_1, i_2}$ denotes the entry in row $i_1$ and column $i_2$ of matrix $X$.

C. Updating Rules

The augmented multiplicative algorithm is used to solve (6). We introduce auxiliary variables $U_r$ and $V_r$ for $E_r$, then the equivalent form can be obtained as follows:

$$
\min_{E_r, c_r, U_r, V_r} \frac{1}{2} \left\{ \|Y - \sum_{r=1}^{R} E_r \circ c_r \|_F^2 + \delta \left\| \sum_{r=1}^{R} E_r - 1_{I \times J} \right\|_F^2 + \lambda \sum_{r=1}^{R} \alpha_r \|U_r\|_{w,*} + \rho \sum_{r=1}^{R} \beta_r \|V_r\|_{1/2} \right. \\
\left. + \mu \frac{1}{2} \|E_r - U_r\|_F^2 + \frac{1}{2} \|E_r - V_r\|_F^2 + Tr(\psi E) + Tr(\Theta C) \right.
$$

(9)

s.t. $E_r \geq 0, c_r \geq 0, E_r = U_r, E_r = V_r$.

Subsequently, constructing the Lagrange function, (9) can be reformulated as

$$
L(E, C, U_r, V_r) = \frac{1}{2} \left\{ \|Y - \sum_{r=1}^{R} E_r \circ c_r \|_F^2 + \delta \left\| \sum_{r=1}^{R} E_r - 1_{I \times J} \right\|_F^2 + \lambda \sum_{r=1}^{R} \alpha_r \|U_r\|_{w,*} + \rho \sum_{r=1}^{R} \beta_r \|V_r\|_{1/2} \\
+ \frac{1}{2} \|E_r - U_r\|_F^2 + \frac{1}{2} \|E_r - V_r\|_F^2 + Tr(\psi E) + Tr(\Theta C) \right.
$$

(10)

where $E = [e_1, e_2, \ldots, e_R]^T \in \mathbb{R}^{R \times I \times J}$ and $C = [c_1, c_2, \ldots, c_R] \in \mathbb{R}^{K \times R}$ are consistent with the expression in Section II-B: $\Psi$ and $\Theta$ are Lagrange multipliers indicating the nonnegative variable constraints in (9), and $\mu > 0$ is a penalty parameter. Now, the optimization problem can be separated into independent subproblems. The variables $E_r, C, U_r$, and $V_r$ are updated alternately in the following.

Update $E$: With fixed $C, U_r$, and $V_r$, the suboptimization problem for $E$ is

$$
\min_{E_r} \frac{1}{2} \left\{ \|Y - \sum_{r=1}^{R} E_r \circ c_r \|_F^2 + \delta \left\| \sum_{r=1}^{R} E_r - 1_{I \times J} \right\|_F^2 + \mu \frac{1}{2} \|E_r - U_r\|_F^2 + \frac{1}{2} \|E_r - V_r\|_F^2 + Tr(\psi E) \right. \\
\left. + Tr(\Theta C) \right\}
$$

(11)

Here, we take advantage of the equivalence relation between MV-NTF and NMF pointed out in Section II-B to unify the variables. Considering this convex subproblem (11) with Karush–Kuhn–Tucker conditions, then we have

$$
C^T (CE - Y) + \delta 1_R (1_R^T E - 1_{I \times J}^T) \\
+ \mu (E - U) + \mu (E - V) + \Psi = 0.
$$

(12)

When $\Psi, * E = 0$, the problem (11) gets its optimal value. Hence, the update rule for $E$ is derived as

$$
E^{(t+1)} \leftarrow E^{(t)} \left( C^{(t)}^T Y + \delta 1_R 1_R^T + \mu U^{(t)} + \mu V^{(t)} \right) \\
/ \left( C^{(t)}^T C^{(t)} E^{(t)} + \delta 1_R 1_R^T E^{(t)} + 2 \mu E^{(t)} \right).
$$

(13)

Update $C$: With fixed $E, U_r$, and $V_r$, the suboptimization problem for $C$ is

$$
\min_{C} \frac{1}{2} \left\{ \|Y - \sum_{r=1}^{R} E_r \circ c_r \|_F^2 + Tr(\Theta C) \right. \\
\left. + Tr(\Psi E) \right\}
$$

(14)

where $E$ and $C$ are the same as (11). Similarly, we have the linear equation

$$
(CE - Y) E^T + \Theta = 0
$$

(15)

where $\Theta, * C = 0$, and the update rule for $C$ is obtained

$$
C^{(t+1)} \leftarrow C^{(t)}, * Y E^{(t+1)} / (C^{(t)} E^{(t+1)} E^{(t+1)^T}).
$$

(16)

Update $U_r$: With fixed $E, C$, and $V_r$, the suboptimization problem for $U_r$ is

$$
\min_{U_r} \lambda \alpha_r \|U_r\|_{w,*} + \frac{1}{2} \|E_r - U_r\|_F^2.
$$

(17)

We solve it by using soft-thresholding operator of the singular values. Define the shrinkage operator as

$$
SHR (X, \Delta) = L \text{diag} ( \max \{ \Sigma - \Delta, 0 \} ) R^T
$$

(18)

where $[L, \Sigma, R] = \text{SVD}(X)$ is the singular value decomposition (SVD) of a matrix $X$; $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_J)$ refers to the singular value matrix formed by the singular values of $X$ as the diagonal entries; $L$ and $R$ are orthogonal matrices. Then we
The unmixing performance use spectral angle distance (SAD), root mean square (RMSE), image reconstruction error (RE) as measured metrics. The SAD evaluates the dissimilarity between the real \( r \)th endmember signature \( \hat{e}_r \) and its estimated signature \( c_r \). It is defined as

\[
\text{SAD}_r = \arccos \left( \frac{c_r^T \hat{e}_r}{\|c_r\| \|\hat{e}_r\|} \right).\tag{25}
\]

The RMSE quantifies the difference between the real abundance map \( \hat{E}_r \) of the \( r \)th endmember and its estimated map \( \hat{E}_r \). It is defined as

\[
\text{RMSE}_r = \frac{1}{\sqrt{N}} \| E_r - \hat{E}_r \|_F.\tag{26}
\]

where \( N = I \times J \) is the number of all pixels in HSI. Generally, there are no referenced endmembers and abundances in real data. The RE is used to measure the difference between the observed HSI \( Y \) and the reconstructed HSI \( \hat{Y} \). It is defined as

\[
\text{RE} = \frac{1}{\sqrt{K \times N}} \| Y - \hat{Y} \|_F.\tag{27}
\]
where \( K \) and \( N \) are the number of bands and the number of pixels in HSI, respectively. For all the above indices, the lower the value, the better the result.

A. Synthetic Data Experiments

1) Data Generation: The synthetic data are generated by a similar way as in [39]. First, we select six pure spectral signatures from the United States Geological Survey digital spectral library as endmembers. Specifically, these are Carnallite, Ammonijarosite, Almandine, Brucite, Axinite, and Chlorite, with 224 spectral bands ranging from 0.38 to 2.5 \( \mu \)m. The signatures of these selected endmembers are shown in Fig. 2.

Then, we generate the abundance as the following steps.

Step 1: Generate an image with size \( z^2 \times z^2 \) pixels divided into \( z \times z \) blocks. Each block has \( z \times z \) pixels.

Step 2: Full each pixel of a block with two kinds of randomly selected endmembers, and set their mixing ratios as \( \theta \) and \( 1 - \theta \), respectively.

Step 3: Use a \( (2z + 1) \times (2z + 1) \) low-pass filter to process the image, thus obtaining mixed pixels.

Step 4: Rescale the fractions of all endmembers in each pixel to satisfy the ANC and ASC constraints.

With endmembers and abundance maps constructed above, we can obtain a clean synthetic HSI by LMM. In order to further evaluate the robustness to additive noise, zero-mean white Gaussian noise is added to the clean HSI. Signal-to-noise ratio (SNR), the measure of noise level, is defined as follows:

\[
\text{SNR} = 10 \log_{10} \frac{E[y^T y]}{E[e^T e]} \tag{28}
\]

where \( y \) and \( e \) represent the clean data and noise data of a pixel, respectively. \( E[\cdot] \) denotes the expectation operator. A series of generated noisy images with SNR will be used in later experiments.

2) Parameter Setting: In the proposed algorithm, there are four parameters \( \delta, \mu, \lambda, \) and \( \rho \) to be determined. In order to reflect the effect of each parameter on the model, we analyze the change of the parameter in a variation range by fixing other parameters for the clean synthetic data (where \( z = 8 \) and \( \theta = 0.8 \)).

First, we set \( \delta \in \{0, 0.2, 0.4, 0.6, 0.8, 1, 3, 5, 8, 10\} \), which gives the impact of ASC on the unmixing performance. Fig. 3(a) shows the index results of SAD and RMSE when \( \delta \) changes. From the figure, it can be seen that \( \delta \) has a stable result between 0.6 and 1. We choose the optimal value \( \delta = 0.8 \) in the following experiments.

Then, we discuss the balance parameter \( \mu \) within the range of \( \{20, 30, 40, 50, 60, 80, 100, 150\} \), which controls the convergence rate of our proposed unmixing algorithm. Fig. 3(b) shows the index results of SAD and RMSE when \( \mu \) changes. According to the figure, both SAD and RMSE obtain the best results when \( \mu = 50 \), thus this value is set in the following experiments.

Finally, as \( \lambda \) and \( \rho \) are the mutually influential parameters in our proposed model, which represent the low-rank and sparse constraints of the abundance map on the unmixing performance, respectively, we use a grid method with \( \lambda \in \{0.001, 0.01, 0.1, 1, 50, 100, 250, 500\} \) and \( \rho \in \{0.001, 0.01, 0.1, 0.5, 10, 50, 100, 250, 500\} \) to jointly search for their optimal values. Fig. 4(a) and (b) displays the changes of SAD and RMSE values, respectively. In these two figures, the change of \( \rho \) is more gentle than that of \( \lambda \), and both of them have slight change between 0.001 and 1. For the following experiments, we set the coordinates of the concave point \( \lambda = 1 \) and \( \rho = 50 \) illustrated in the figures as their parameter values, which contribute to the best unmixing result.

3) Scene Variation: According to the synthesis method, there are three factors that can be debugged to generate the experimental simulated data, which are noise level SNR, image size \( z \), and endmembers’ mixing ratio \( \theta \). In addition, the number of endmembers is usually unknown in the actual scene. Thus, we discuss these four variables within the appropriate scopes in the following, and test the performance of the seven unmixing methods. The parameters of all the used methods are adjusted to their optimal values for experiments.
First, we reveal the robustness of different methods under different noise levels. The synthetic data are generated by fixing $z = 8$, $\theta = 0.8$ with SNR varied from $\{15, 25, 35, 45, \text{inf}\}$ dB. Fig. 5(a) shows the results of all methods in terms of SAD and RMSE. It can be seen that with the increase of noise level, the index values also increase. NMF gets the worst result because it has no regular constraint and loses the spectral-spatial structure of HSI. Our proposed method considers the actual natures of abundance maps upon tensor framework; thus, it obtains the lowest index results that outperform other methods.

Next, we measure the influence of different methods under different image sizes. The synthetic data are generated by fixing SNR $= 25$ dB, $\theta = 0.8$ with the image size selected as $\{36 \times 36, 49 \times 49, 64 \times 64, 81 \times 81, 100 \times 100\}$, corresponding to $z \in \{6, 7, 8, 9, 10\}$. Fig. 5(b) shows the index results of all methods when $z$ changes. This figure reflects the phenomenon that more pixels contain more abundant information, which is more conducive to unmixing. It is worth noting that our proposed method gets the best results in both SAD and RMSE of all the set image sizes.

Then, we evaluate the performance of different methods under different mixing levels. The mixing level is represented by ratio $\theta$, where a smaller $\theta$ implies a higher mixing level. We construct the synthetic data by fixing SNR $= 25$ dB, $z = 8$, and $\theta$ is chosen from $\{0.5, 0.6, 0.7, 0.8, 0.9\}$. Fig. 5(c) presents the experimental results of all methods varying with the mixing level. As shown in this figure, all methods perform better when $\theta$ increases. Our proposed method has better results at higher $\theta$ values compared with other methods, which indicates that its advantages are enhanced in highly correlated data.

Finally, we analyze the impact of different methods under different estimated number of endmembers. The data of $z = 8$, $\theta = 0.8$, SNR $= 25$ dB, and the true number of endmembers $R = 6$ are tested. We select the estimated number of endmembers in range of $\{4, 5, 6, 7, 8\}$ for experiments. Fig. 5(d) shows the SAD and RMSE results of all methods adopting different estimates. It can be seen that the number of endmembers gets the best result at the exact value. The misestimation of the number of endmembers affects the unmixing performance of all methods.

However, our method can still keep the advantage in a small range of the misestimation.

4) **Unmixing Performance**: We select a set of unmixing results for the synthetic data with the size of $64 \times 64 \times 224$, $\theta = 0.8$, and SNR $= 25$ dB for demonstration. Table I gives the numerical index results, in which the bold value corresponds to the best result of each endmember. Compared with these methods, our method performs better for endmembers #2, #3, #5, and #6 in terms of SAD. Besides, the endmembers except endmembers #2 and #4 also get the best RMSE. The mean values of SAD and RMSE indicate that our method is superior to other methods on the overall performance. Figs. 6 and 7 visually show the advantages of the endmember curves and abundance maps restored by our method.

5) **Initialization Option**: We test the unmixing methods under two different initialization strategies: random initialization and VCA-FCLS initialization for the synthetic data with $z = 8$, $\theta = 0.8$, and SNR $= 25$ dB. All the methods have 20 trials at the same initial value. The means and standard deviations of SAD and RMSE are shown in Fig. 8. This figure shows that all the methods have the better numerical results by VCA-FCLS strategy. And our method achieves the best performance on both initialization strategies. The advantage of our method in adopting random initialization is more obvious, which indicates that it is robust in initial selection.

6) **Ablation Analysis**: Here, we design two ablation experiments to discuss the role of regular terms and adaptive parameters in our model. First, the ablation experiment of regularizers is performed by ignoring the sum-to-one constraint, the low-rank constraint, and the sparse constraint in our proposed model, respectively. Corresponding to the following three loss functions:

\[
\min_{\mathbf{E}_r, c_r} \frac{1}{2} \left\| \mathbf{Y} - \sum_{r=1}^{R} \mathbf{E}_r \circ c_r \right\|_F^2 + \lambda \sum_{r=1}^{R} \alpha_r \|\mathbf{E}_r\|_{w, r} + \rho \sum_{r=1}^{R} \beta_r \|\mathbf{E}_r\|_{1/2}
\]  

\[
\min_{\mathbf{E}_r, c_r} \frac{1}{2} \left\| \mathbf{Y} - \sum_{r=1}^{R} \mathbf{E}_r \circ c_r \right\|_F^2 + \lambda \sum_{r=1}^{R} \alpha_r \|\mathbf{E}_r\|_{w, r} + \rho \sum_{r=1}^{R} \beta_r \|\mathbf{E}_r\|_{1/2}
\]  

\[
\min_{\mathbf{E}_r, c_r} \frac{1}{2} \left\| \mathbf{Y} - \sum_{r=1}^{R} \mathbf{E}_r \circ c_r \right\|_F^2 + \lambda \sum_{r=1}^{R} \alpha_r \|\mathbf{E}_r\|_{w, r} + \rho \sum_{r=1}^{R} \beta_r \|\mathbf{E}_r\|_{1/2}
\]

Fig. 5. Performance with respect to different scenes. (a) Noise levels. (b) Image sizes. (c) Mixing levels. (d) Endmember estimates.
TABLE I
MEANS AND STANDARD DEVIATIONS OF SAD AND RMSE BY DIFFERENT UNMIXING METHODS FOR THE SYNTHETIC DATA

| Method     | SAD      | RMSE     |
|------------|----------|----------|
| NMF        | 0.1499±0.50% | 0.1107±0.63% |
| L1/2-NMF   | 0.1372±4.26% | 0.1011±2.25% |
| MRS-NMF    | 0.1246±3.20% | 0.0799±1.86% |
| TV-RSNMF   | 0.1264±4.29% | 0.0914±1.85% |
| MV-NTF     | 0.0941±4.72% | 0.0670±1.59% |
| MV-NTF-TV  | 0.0964±4.01% | 0.0707±1.74% |
| EWSP-NTF   | 0.0565±4.59% | 0.0355±1.65% |

\[ \min_{E_r, c_r} \frac{1}{2} \left\| Y - \sum_{r=1}^{R} E_r \circ c_r \right\|_F^2 + \frac{\delta}{2} \sum_{r=1}^{R} \left\| E_r - 1_1 \times J \right\|_F^2 + \rho \sum_{r=1}^{R} \beta_r \left\| E_r \right\|_{1/2} \quad (30) \]

\[ \min_{E_r, c_r} \frac{1}{2} \left\| Y - \sum_{r=1}^{R} E_r \circ c_r \right\|_F^2 + \frac{\delta}{2} \sum_{r=1}^{R} \left\| E_r - 1_1 \times J \right\|_F^2 + \frac{\rho}{2} \sum_{r=1}^{R} \beta_r \left\| E_r \right\|_{1/2} + \lambda \sum_{r=1}^{R} \alpha_r \left\| E_r \right\|_{w_r, *}, \quad (31) \]

Fig. 6. Reference and estimated spectral curves by different unmixing methods for the synthetic data. From top to bottom: endmembers #1 to #6. Orange lines: reference endmembers. Blue lines: estimated endmembers.
We test them and the original loss function (6) for the synthetic data with $z = 8$, $\theta = 0.8$, and SNR = 25 dB. Table II gives the SAD and RMSE values of 20 trials. It can be seen that the proposed method with all constraints has the best results. Figs. 9 and 10 show the endmember curves and abundance maps for endmember #3 obtained by the four models, respectively.

**TABLE II**

| Index | Problem (8)       | Problem (29)      | Problem (30)      | Problem (31)      |
|-------|-------------------|-------------------|-------------------|-------------------|
| SAD   | $0.0565 \pm 4.50\%$ | $0.5445 \pm 3.82\%$ | $0.1035 \pm 3.90\%$ | $0.1225 \pm 3.95\%$ |
| RMSE  | $0.0355 \pm 1.65\%$ | $0.0971 \pm 3.36\%$ | $0.0465 \pm 1.33\%$ | $0.0483 \pm 1.86\%$ |
Comparing them with the references, we can draw the conclusion that each regularizer in our model has a positive role on the unmixing performance.

Then, we discuss the effect of the low-rank and sparse adaptive parameters $\alpha_r$ and $\beta_r$ in our model (6). Fig. 11 shows the changes of adaptive parameters during the iteration for the synthetic data. It can be seen that both the two adaptive parameters of each abundance map tend to different stable values as iteration increases. This behavior confirms that each abundance map has different properties and it is necessary to treat it differently. Four cases are considered in the ablation experiment of adaptive parameters: with no adaptive parameters (when $\alpha_r = 1$, $\beta_r = 1$), with sparse adaptive parameter (when $\alpha_r = 1$), with low-rank adaptive parameter (when $\beta_r = 1$), and with both low-rank and sparse adaptive parameters (the original model). We test these cases for the 20 groups synthetic data with $z = 8$, $\theta = 0.8$, and SNR = 25 dB. Table III gives the mean and standard deviation of SAD and RMSE. The results show that it is advantageous to have adaptive parameters.

### B. Real-World Data Experiments

Four real-world datasets are used to test the performance of the proposed method: Urban, Jasper Ridge, Samson, and Washington DC Mall. Different datasets were collected by different sensors with different sizes and endmembers, which can reflect the applicability of the method. Unlike the simulated data, these real-world datasets originally have no real ground truth. For visual comparison with other methods, we follow the same approach in [39] to get the referenced endmembers and abundances: the strategy of initialization and parameter debugging is as well as that adopted in the synthetic data experiments.

1) **Urban Data:** Urban dataset is one of the most widely used data in HU study, which is acquired by the hyperspectral digital imagery collection experiment (HYDICE) sensor over an urban area. This image has $307 \times 307$ pixels with 210 spectral bands, whose wavelength ranges from 0.4 to 2.5 $\mu$m. Affected by the imaging environment, we remove some low-SNR bands and water-vapor bands (1–4, 76, 87, 101–111, 136–153, and 198–210) from the original data, leaving 162 bands for experiments. Besides, we assume that Urban data are consisted of four endmembers: asphalt, grass, tree, and roof.

Table IV presents the SAD and RMSE results of the seven unmixing methods. From Table IV, we see that MRS-NMF and TV-RSNMF obtain the best SAD for the endmember asphalt and the endmember grass, respectively. MV-NTF have the best RMSE for the endmember grass. And our proposed method achieves the best results in other endmembers, and the mean value of SAD and RMSE is far superior to other methods. Figs. 12 and 13 show the results of endmembers and abundances obtained by unmixing the Urban data. Fig. 12 displays that TV-RSNMF is closest to the reference curve for the endmember grass, and our method performs well in the remaining endmembers. According to Fig. 13, all methods can restore the general information compared with the referenced abundance maps. However, our method is superior in details since more properties of abundances are considered.

2) **Jasper Ridge Data:** Jasper Ridge dataset is another commonly used data for HSI processing, which is collected by the airborne visible/infrared imaging spectrometer sensor over the Jasper Ridge Natural Reserve in California, USA. The original data contain $512 \times 614$ pixels with 224 bands ranging from 0.38 to 2.5 $\mu$m. In order to ensure the efficiency of the experiments, we remove the useless bands (1–3, 108–112, 154–166, and 220–224) and intercept the data of $100 \times 100$ pixels with 198 bands for use. Four endmembers are assumed to exist in this subimage: tree, water, soil, and road.

Table V presents the SAD and RMSE results for Jasper Ridge data obtained by different unmixing methods. From Table V, we
### TABLE IV
**MEANS AND STANDARD DEVIATIONS OF SAD AND RMSE BY DIFFERENT UNMIXING METHODS FOR URBAN DATA**

| Method  | SAD | RMSE |
|---------|-----|------|
|        | NMF | $L_{1/2}$-NMF | MRS-NMF | TV-RSNMF | MV-NTF | MV-NTF-TV | EWSP-NMF |
| Asphalt | 0.447±0.80% | 0.368±9.76% | 0.2500±6.23% | 0.2826±8.57% | 0.4330±3.51% | 0.4434±3.48% | 0.3087±10.35% |
| Grass   | 0.350±9.61% | 0.3157±2.24% | 0.1799±2.54% | 0.1451±5.29% | 0.3752±3.83% | 0.3887±4.51% | 0.2214±8.08% |
| Tree    | 0.115±1.36% | 0.093±1.67%  | 0.1328±3.46% | 0.1312±1.04% | 0.1414±1.23% | 0.1402±2.74% | 0.0923±9.49% |
| Roof    | 0.077±5.56% | 0.058±5.70%  | 0.026±2.53%  | 0.024±1.18%  | 0.0472±6.59% | 0.0489±5.23% | 0.0870±2.47% |
| Mean    | 0.2469±4.29% | 0.2087±4.39% | 0.1922±2.43% | 0.2004±2.18% | 0.3356±3.48% | 0.3653±4.20% | 0.1697±5.92% |

### TABLE V
**MEANS AND STANDARD DEVIATIONS OF SAD AND RMSE BY DIFFERENT UNMIXING METHODS FOR JASPER RIDGE DATA**

| Method  | SAD | RMSE |
|---------|-----|------|
|        | NMF | $L_{1/2}$-NMF | MRS-NMF | TV-RSNMF | MV-NTF | MV-NTF-TV | EWSP-NTF |
| Tree    | 0.784±7.85% | 0.308±4.19% | 0.1241±8.06% | 0.1564±8.94% | 0.1915±3.32% | 0.2023±5.54% | 0.2203±4.39% |
| Water   | 0.135±5.46% | 0.123±6.51%  | 0.145±3.46%  | 0.158±6.73%  | 0.1704±3.53% | 0.1674±2.32% | 0.1049±5.39% |
| Soil    | 0.261±8.4%  | 0.200±7.38%  | 0.171±5.48%  | 0.285±10.63% | 0.182±6.59% | 0.198±14.89% | 0.1667±7.31% |
| Road    | 0.245±5.11% | 0.229±6.75%  | 0.445±9.42%  | 0.248±7.26%  | 0.245±6.07% | 0.250±3.15% | 0.207±4.41% |
| Mean    | 0.266±5.8%  | 0.215±5.08%  | 0.221±3.29%  | 0.212±4.14%  | 0.197±3.22% | 0.204±4.03% | 0.177±4.67% |

Fig. 12. Reference and estimated spectral curves by different unmixing methods for Urban data. From top to bottom: asphalt, grass, tree, and roof. Orange lines: reference endmembers. Blue lines: estimated endmembers.

| Method  | SAD | RMSE |
|---------|-----|------|
|        | NMF | $L_{1/2}$-NMF | MRS-NMF | TV-RSNMF | MV-NTF | MV-NTF-TV | EWSP-NTF |
| Tree    | 0.185±4.33% | 0.154±7.07% | 0.168±7.61% | 0.253±10.32% | 0.169±6.31% | 0.177±5.33% | 0.145±6.42% |
| Water   | 0.254±9.33% | 0.230±11.34% | 0.219±7.31% | 0.205±5.33% | 0.219±7.33% | 0.203±5.33% | 0.2001±8.91% |
| Soil    | 0.218±7.79% | 0.2254±5.45% | 0.340±5.39% | 0.237±8.17% | 0.2662±2.12% | 0.2660±2.04% | 0.183±5.51% |
| Road    | 0.302±6.30% | 0.306±5.41% | 0.306±6.60% | 0.317±4.28% | 0.1650±0.84% | 0.157±2.11% | 0.232±3.49% |
| Mean    | 0.240±2.61% | 0.230±2.33% | 0.257±4.15% | 0.249±1.79% | 0.209±2.42% | 0.201±2.07% | 0.188±3.52% |
Fig. 13. Reference and estimated abundance maps by different unmixing methods for Urban data. From top to bottom: asphalt, grass, tree, and roof.

Fig. 14. Reference and estimated spectral curves by different unmixing methods for Jasper Ridge data. From top to bottom: tree, water, soil, and road. Orange lines: reference endmembers. Blue lines: estimated endmembers.

see that MRS-NMF gets the best SAD for the endmember tree, and TV-MV-NTF achieves the best RMSE for the endmember road. Apart from these, the best results of other endmembers are obtained by our proposed method. The mean value of SAD and RMSE indicates that our method is better than other methods on overall unmixing performance. Figs. 14 and 15 visually show the difference between estimated and referenced endmembers and abundance maps for Jasper Ridge data, respectively. Fig. 14 displays that the endmember curves of MRS-NMF recovered well for endmembers tree, water, and soil, but the curve for the endmember road is quite different from the referenced curve. The recovery effect of our method for four endmembers is relatively stable. According to Fig. 15, the abundance maps restored by our proposed method are more consistent with the references, especially for endmembers tree, water, and soil.

3) Samson Data: Samson hyperspectral dataset is collected by Florida Environment Research Institute using Samson sensor. The full image contains 952 × 952 pixels with 156 bands, and its wavelength ranges from 0.401 to 0.889 μm. In our experiments, there is a part of Samson data with the size of 95 × 95 pixels for unmixing. This subimage is set to include three endmembers: soil, tree, and water.

Table VI gives the SAD and RMSE values of the seven unmixing methods for Samson data. It can be seen that our
proposed method still outperforms other methods for this data, especially for endmembers tree and water. Figs. 16 and 17 show the unmixing performance of different methods by images. Fig. 16 gives the spectral curves that are referenced and restored by unmixing methods. It can be found that the curves obtained by our method have the smallest difference with the references. From the restored and corresponding reference abundance maps shown in Fig. 17, our method gets the closest result to the references. It illustrates that the low-rank and sparse constraints added by our method to the abundance maps are effective.

Moreover, the RE is considered to compare the difference between the real ground data $Y$ and $\hat{Y} = \sum_{r=1}^{R} \hat{E}_r \circ \hat{c}_r$ reconstructed by unmixing components here. The index does not require the referenced endmembers and abundance maps in real data. Table VII presents the RE results for the above three real data. As can be seen from the table, our method gets the minimum value among all the three data. It reflects that our proposed method has the best unmixing performance because of the least information loss of the original data $Y$.

4) Washington DC Mall Data: Washington DC Mall data are an HSI captured by HYDICE sensor over the Washington DC Mall. The original data are $1208 \times 307$ pixels with 210 bands, whose spectral range is from 0.4 to 2.4 $\mu$m. We crop a subimage of $319 \times 292$ pixels with 191 bands for experiment. The six

![Image of Fig. 15](image-url)  
Fig. 15. Reference and estimated abundance maps by different unmixing methods for Jasper Ridge data. From top to bottom: tree, water, soil, and road.

| TABLE VI |
|-------------------------------|
| | **MEANS AND STANDARD DEVIATIONS OF SAD AND RMSE BY DIFFERENT UNMIXING METHODS FOR SAMSON DATA** |
| | **SAD** |
| | $\text{NMF}$ | $\text{L}_{1/2}$-$\text{NMF}$ | $\text{MRS}$-$\text{NMF}$ | $\text{TV-RS}$-$\text{NMF}$ | $\text{MV}$-$\text{NTF}$ | $\text{MV}$-$\text{NTF}$-$\text{TV}$ | $\text{EWSP}$-$\text{NTF}$ |
| Soil | 0.1949±5.49% | 0.1317±6.11% | 0.1464±4.69% | **0.0882±4.61%** | 0.1034±5.68% | 0.1022±5.31% | 0.0948±5.09% |
| Tree | 0.092±0.86% | 0.078±0.49% | 0.066±1.25% | 0.084±1.55% | 0.082±0.5% | 0.0826±0.33% | **0.0419±1.28%** |
| Water | 0.166±1.16% | 0.126±2.56% | 0.046±0.24% | 0.062±3.34% | 0.081±1.62% | 0.0496±3.38% | 0.0195±0.96% |
| Mean | 0.1519±1.88% | 0.1120±2.41% | 0.084±3.18% | 0.0897±2.89% | 0.0791±3.04% | 0.0781±2.92% | **0.0521±2.36%** |

| **RMSE** |
| $\text{NMF}$ | $\text{L}_{1/2}$-$\text{NMF}$ | $\text{MRS}$-$\text{NMF}$ | $\text{TV-RS}$-$\text{NMF}$ | $\text{MV}$-$\text{NTF}$ | $\text{MV}$-$\text{NTF}$-$\text{TV}$ | $\text{EWSP}$-$\text{NTF}$ |
| Soil | 0.0991±3.20% | 0.0976±4.76% | 0.1138±2.82% | 0.1036±2.51% | 0.0796±3.84% | 0.0743±2.54% | **0.0729±2.76%** |
| Tree | 0.1359±2.16% | 0.126±2.56% | 0.0781±3.93% | 0.0847±3.71% | 0.0914±3.12% | 0.1008±3.24% | **0.0593±2.17%** |
| Water | 0.1749±2.76% | 0.149±3.82% | 0.1187±2.43% | 0.0914±1.27% | 0.0923±2.86% | 0.0850±2.83% | **0.0732±1.16%** |
| Mean | 0.1366±1.38% | 0.1244±2.47% | 0.1035±1.90% | 0.0932±3.19% | 0.0878±2.25% | 0.0867±1.90% | **0.0685±1.64%** |

| TABLE VII |
|-------------------------------|
| | **RE VALUES AND STANDARD DEVIATIONS BY DIFFERENT UNMIXING METHODS FOR THE REAL DATASETS** |
| | **NMF** | $\text{L}_{1/2}$-$\text{NMF}$ | $\text{MRS}$-$\text{NMF}$ | $\text{TV-RS}$-$\text{NMF}$ | $\text{MV}$-$\text{NTF}$ | $\text{MV}$-$\text{NTF}$-$\text{TV}$ | $\text{EWSP}$-$\text{NTF}$ |
| Urban | 0.0205±0.63% | 0.0181±0.55% | 0.0140±0.49% | 0.0129±0.36% | 0.0146±0.34% | 0.0135±0.30% | **0.0098±0.19%** |
| Jasper Ridge | 0.0170±0.57% | 0.0164±0.51% | 0.0134±0.30% | 0.0158±0.45% | 0.0131±0.28% | 0.0127±0.23% | **0.0069±0.06%** |
| Samson | 0.0162±0.46% | 0.0150±0.43% | 0.0178±0.24% | 0.0070±0.15% | 0.0125±0.29% | 0.0115±0.27% | **0.0043±0.04%** |
endmembers considered in this scene are grass, tree, road, roof, water, and trail.

Since the deep learning-based unmixing methods widely use this data for test, we also compare the performance of our method and three deep learning-based methods UnDIP [50], CyCUNet [49], MiSiCNet [51] here. Figs. 18 and 19 show the spectral curves and abundance maps obtained by different unmixing methods for endmembers tree and roof, respectively. It can be seen that UnDIP obtains the most similar unmixing result to the reference. Although our method lacks the specific training of data, it is still competitive compared with these deep learning-based methods and is superior to the same type of traditional unmixing methods for this data.
C. Discussion

Not only does EWSP-NTF improve the unmixing performance, but its novelty lying in the combination of the equivalence with NMF also opens up a new algorithm for the MV-NTF-based methods. Therefore, in this section, we record the running time of all methods to further compare their computational efficiency. And the stability of the proposed method under different experimental data is also demonstrated.

1) Computing Time: In order to tolerate the differences among methods, the stopping conditions of all methods keep their original optimal strategies in our experiments. Thus, for the fairness of time comparison, we record the per iteration time of methods under the same computer equipment. Specifically, Table VIII displays the time cost per iteration of different methods for the experimental datasets. As seen in Table VIII, NMF takes the shortest time because it has no constraints. Although our proposed method adds the abundance constraints to the basic MV-NTF framework, even the low-rank constraint needs the SVD, it still has a great speed-up among these MV-NTF-based methods. This time advantage can be explained by comparing their algorithm complexity. The simplification of our method is to avoid the low-rank matrix decomposition, which can be done without the rank of the abundance map.

2) Convergence Analysis: For further convincing the stability of our proposed method, we numerically analyze its convergence by recording the objective function values during the iteration both for the synthetic data and real-world data. Fig. 20(a) depicts the convergence curve for the synthetic data with $z = 8, \theta = 0.8$, and SNR = 25 dB. And Fig. 20(b) shows the convergence curve for Samson data. It can be observed from the figures that their objective function values are monotonously decreasing and gradually becoming stable as the iteration number increases. Besides, other data have the similar behavior in the experiment, so they are omitted here.

V. CONCLUSION

In this article, we propose a new unmixing method named EWSP-NTF for the spectral unmixing problem. This method not only retains the lossless third-order tensor representation of HSI but also considers the characteristics of each abundance map, which are constrained by adaptive low-rank and sparse regular terms simultaneously. Compared with the existing methods, EWSP-NTF distinguishes to treat each abundance map, which is more in line with the actual scene. Furthermore, abandoning the low-rank matrix decomposition of MV-NTF but considering the abundance map as a whole also allows it to simplify the tensor operation by linking with NMF. Experiments on synthetic data and real-world data confirm the effectiveness of our proposed method. EWSP-NTF achieves the better unmixing performance and has advantages in time consumption against the same type of tensor-based unmixing methods. Besides, our method has a competitive performance with the deep learning-based unmixing methods. In the future, we will explore the potential properties of endmembers and abundances to further improve the unmixing performance. Moreover, we will study the unmixing application of NTF in nonlinear mixing mode.

ACKNOWLEDGMENT

The authors would like to thank the editors and the four anonymous reviewers for their constructive, detailed, and helpful advice regarding this article.
[45] J.-J. Wang, D.-C. Wang, T.-Z. Huang, J. Huang, X.-L. Zhao, and L.-J. Deng, “Endmember independence constrained hyperspectral unmixing via nonnegative tensor factorization,” Knowl.-Based Syst., vol. 216, 2021, Art. no. 106657.

[46] D. Hong et al., “More diverse means better: Multimodal deep learning meets remote-sensing imagery classification,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 5, pp. 4340–4354, May 2021.

[47] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, “Graph convolutional networks for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 7, pp. 5966–5978, Jul. 2021.

[48] D. Hong et al., “Spectralformer: Rethinking hyperspectral image classification with transformers,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5518615.

[49] L. Gao, Z. Han, D. Hong, B. Zhang, and J. Chanussot, “CyCU-Net: Cycle-consistency unmixing network by learning cascaded autoencoders,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5503914.

[50] B. Rasti, B. Koirala, P. Scheunders, and P. Ghamisi, “UnDIP: Hyperspectral unmixing using deep image prior,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5504615.

[51] B. Rasti, B. Koirala, P. Scheunders, and J. Chanussot, “MiSiCNet: Minimum simplex convolutional network for deep hyperspectral unmixing,” IEEE Trans. Geosci. Remote Sens., vol. 60, Jan. 2022, Art. no. 5522815.

[52] N. Audebert, B. Le Saux, and S. Lefevre, “Generative adversarial networks for realistic synthesis of hyperspectral samples,” in Proc. IEEE Int. Geosci. Remote Sens. Symp., 2018, pp. 4359–4362.

[53] D. Hong et al., “Endmember-guided unmixing network (EGU-Net): A general deep learning framework for self-supervised hyperspectral unmixing,” IEEE Trans. Neural Netw. Learn. Syst., vol. 33, no. 11, pp. 6518–6531, Nov. 2022.

[54] J. Yao, D. Hong, L. Xu, D. Meng, J. Chanussot, and Z. Xu, “Sparsity-enhanced convolutional decomposition: A novel tensor-based paradigm for blind hyperspectral unmixing,” IEEE Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 5505014.

[55] L. De Lathauwer and D. Nion, “Decompositions of a higher-order tensor in block terms—Part II: Definitions and uniqueness,” SIAM J. Matrix Anal. Appl., vol. 30, no. 3, pp. 1033–1066, 2008.

[56] 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.

Ping Yang received the B.S. degree in formation and computing science from Southwest Petroleum University, Chengdu, China, in 2020. She is currently working toward the M.S. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China. Her research interests include remote sensing and image processing.

Jin-Ju Wang received the B.S. degree in mathematics and applied mathematics from the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China, in 2017. She is currently working toward the Ph.D. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China.

Jie Huang (Member, IEEE) received the Ph.D. degree in applied mathematics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2013. She is currently an Associate Professor with the School of Mathematical Sciences, UESTC. Her research interests include remote sensing and image processing.

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

Ping Yang

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 currently 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 for The Scientific World Journal, Advances in Numerical Analysis, the Journal of Applied Mathematics, and the Journal of Electronic Science and Technology, China.

Jie Huang (Member, IEEE) received the Ph.D. degree in applied mathematics from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2013. She is currently an Associate Professor with the School of Mathematical Sciences, UESTC. Her research interests include remote sensing and image processing.

Jin-Ju Wang received the B.S. degree in mathematics and applied mathematics from the Southwest University of Science and Technology, Mianyang, China, in 2017. She is currently working toward the Ph.D. degree with the School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China.