Combinatorial Nonnegative Matrix-Tensor Factorization for Hyperspectral Unmixing Using a General ℓ_q Norm Regularization

Saeid Gholinejad and Alireza Amiri-Simkooei

***Abstract*—Hyperspectral unmixing (HU), an essential procedure for various environmental applications, has garnered significant attention within remote sensing communities. Among different groups of HU methods, nonnegative matrix factorization (NMF)-based ones have gained widespread popularity due to their high capability of simultaneously extracting endmembers and their corresponding abundances. However, converting a 3-D hyperspectral data cube into a 2-D matrix format leads to the loss of spatial and potential correlation information. Consequently, in recent years, nonnegative tensor factorization (NTF) methods, which preserve the 3-D nature of hyperspectral data cube, have been extensively embraced by numerous researchers. Nevertheless, incorporating prior information into NTF-based problems faces limitations owing to the inconsistency of such information, particularly concerning ℓ_1 norm sparsity and the abundance sum-to-one constraint (ASC). To address this limitation, our study introduces a novel general regularization term. This term leverages sparsity and ASC simultaneously, integrating it into a matrix-tensor factorization framework. Our proposed method, named a matrix-tensor-based HU method with general ℓ_q norm regularization (MTUHL$_q$), is established on the block term decomposition (BTD) paradigm, which ensures physical interpretability and simple implementation. To investigate the performance of the proposed MTUHL$_q$, a series of experiments on both synthetic and real hyperspectral datasets were conducted. The results of the implemented experiments indicated that the proposed method outperformed other state-of-the-art HU methods.**

***Index Terms*—Abundance sum-to-one constraint (ASC), hyperspectral unmixing (HU), nonnegative matrix factorization (NMF), nonnegative tensor factorization (NTF), sparse regularization.**

I. Introduction

CONTAINING a large number of narrow spectral bands, hyperspectral images (HSIs) are known as one of the most wealthy resources of information about the Earth's surface. A wide range of their applications can be enumerated in various fields including agricultural, environmental, geological, and mineralogical problems. Despite this excellent capability, the existence of mixed pixels in these images due to the low spatial resolution of the sensor, multiscattering, and microscopic mixing, has affected their use in remote sensing processes [1].

To deal with the abovementioned challenge, hyperspectral unmixing (HU) has been presented, which is also a preprocess for many other remote sensing analyses. HU is the procedure of decomposing a mixed pixel into its constituent endmembers and their corresponding abundances. Generally, according to the approach at hand, unmixing models can be divided into physics-based models and data-driven models [2]. Physics-based unmixing models, like the linear mixing model (LMM) [3], rely on established physical principles to simulate light-material interactions, while data-driven models, such as deep learning ones [4], [5], learn directly from data without explicit physical modeling, leveraging complex relationships within hyperspectral data. While data-driven models have become popular for their ability to autonomously extract intricate features from data, opting for a physics-based model in this context presents several compelling advantages, like interpretability, transparency, and few computational resources and training data. Furthermore, physics-based models can be customized to leverage domain-specific knowledge and prior information, enhancing their performance and adaptability to specific application needs.

Based on the mixing level, i.e., microscopic or macroscopic, different physics-based HU methods are categorized under two different models: 1) LMM [3], and 2) nonlinear mixing model (NLMM) [2]. LMM assumes that the mixing of pixels is due to the low spatial resolution of the hyperspectral sensor, while NLMM also considers multiscattering and microscopic mixing.

LMM is widely used in the remote sensing community due to its simplicity and proper performance. LMM-based HU methods can be divided into three main categories including geometrical, statistical, and sparse-regression-based methods [3]. Geometrical methods try to embed data points in a simplex in the feature space, and therefore, the procedure of endmember extraction using geometrical methods is to find the vertices of the simplex. Therefore, geometrical methods are only used in endmember extraction, and inevitably, abundances are estimated through some postprocessing algorithms such as fully constrained least squares (FCLS) [6]. Statistical methods, considered as blind source separation methods, simultaneously estimate endmembers and abundances. Nonnegative matrix factorization

Manuscript received 21 January 2024; revised 19 March 2024 and 8 April 2024; accepted 19 April 2024. Date of publication 22 April 2024; date of current version 8 May 2024. *(Corresponding author: Alireza Amiri-Simkooei.)*

Saeid Gholinejad is with the Department of Geomatics Engineering, Faculty of Civil Engineering and Transportation, University of Isfahan, Isfahan 8174673441, Iran (e-mail: saeedghj.1991@gmail.com).

Alireza Amiri-Simkooei is with the Department of Control and Operations, Faculty of Aerospace Engineering, Delft University of Technology, 2629 Delft, The Netherlands, and also with the Department of Geomatics Engineering, University of Isfahan, Isfahan 8174673441, Iran (e-mail: a.amirisimkooei@tudelft.nl).

Digital Object Identifier 10.1109/JSTARS.2024.3392497

(NMF)-based methods [7], [8] are the most famous ones in the statistical category of HU methods. Sparse regression-based methods [9] are semisupervised types of HU methods that utilize spectral libraries as the matrix of endmembers.

Compared with the other two categories, statistical methods, especially NMF-based methods, are more powerful in extracting both endmember and abundance matrices. Despite the great advantages of the NMF algorithm, the nonconvexity of its problem leads the results to local minima. To alleviate this challenge, different studies proposed several versions of the NMF problem by imposing some constraints on the main NMF objective function, having regularization effects. Moreover, many studies focus on the structure of the NMF problem to excel in the performance of the NMF algorithm. Readers are referred to [8] for more details on NMF-based methods.

To take into account spatial relationships between pixels of HSI, various spatial constraints are also added to the NMF problem. Abundance separation and smoothness constrained NMF (ASSNMF) was proposed by Liu et al. [10] to minimize the mutual information of the abundances of the different endmembers. It simultaneously reduced abrupt changes in the abundance of the adjacent pixels. Local neighborhood weights have been introduced into the NMF problem in [11]. Some other structured NMF methods were also proposed to incorporate spatial information into the problem [12], [13], [14]. In addition to these methods, a large number of spatially constrained NMF methods belong to the total variation (TV)-based algorithms. TV regularization promotes piecewise smoothness in the abundances matrix of the neighboring pixels for the same endmembers. It was first introduced in HU by the study of Iordache et al. [15] and followed by a series of other studies like double reweighted ℓ_1 minimization with TV regularizer [16], row-sparsity spectral unmixing via TV (RSSUn-TV) [17], improved collaborative nonnegative matrix factorization and TV algorithm (ICoNMF-TV) [18], a data-driven graph TV regularization-based NMF [19], robust double spatial regularization sparse unmixing (RDSRSU) [20], and a weighted TV regularized blind unmixing (wtvBU) based on the log-exp function [21].

In spite of adding several spatial constraints, loss of spatial information in the HU problem through NMF-based methods is still inevitable. This is due to the 3-D nature of the HSIs, which is distorted in the matricization process. To overcome this limitation, tensors [22] are the best alternatives for matrices, which can represent an HSI as a cube; as it really is. Accordingly, considering HSI as a third-order tensor, and regarding the HU as a nonnegative tensor factorization (NTF) problem is more appropriate to preserve the original 3-D mode of the HSI. To the best of our knowledge, the study of [23] was the first to introduc NTF into the HU problem. Subsequently, a wide array of studies has been presented in this field, which can be categorized into three main groups based on the tensor decomposition procedure: canonical polyadic decomposition (CPD), Tucker tensor decomposition (TTD), and block term decomposition (BTD) [24].

CPD decomposes a third-order tensor into a series of rank-one tensors. The abovementioned NTF-based study by Zhang et al. [23] applied CPD for the first time in HU. ULTRA [25] and its extended version ULTRA-V [26] were also other CPD-based methods that imposed low-rank regularization terms on the abundance and endmember tensors. Low-rank multifeature HU based on CPD was another study in this field [27]. Another decomposition method, TTD, decomposes a tensor into a set of matrices and one small core tensor. In this context, the study of Sun et al. [28] was the only one, in which a blind unmixing method, based on ℓ_1 norm and TTD was proposed.

Generally, CDP and TTD-based results in HU are not physically interpretable [29]. Consequently, there are not many studies in the HU literature that apply these two types of decomposition. In contrast, BTD, which decomposes an N-order tensor into R component tensors, has been extensively applied in the HU methods due to its high compatibility with the spectral mixing model.

A remarkable point in applying BTD for HU belongs to the study of Qian et al. [30], in which a matrix-vector nonnegative tensor decomposition (MV-NTF) was proposed. Despite the superiority of this method against many NMF-based algorithms, there are some challenges that reduce the accuracy of its outcomes. Suffering from some difficulties including low signal-to-noise ratio (SNR), bare identifiability, and ill-conditioning of tensor decomposition, along with the ignorance of detailed spatial structure information through BTD, are the most important challenges, which MV-NTF faces [31], [32]. Accordingly, several methods were presented, incorporating different constraints into the MV-NTF framework.

Xiong et al. [32] added the TV regularization term into the MV-NTF problem to exploit local spatial information. A combinatorial method of NMF and NTF, named SCNMTF, was also proposed in [33] to simultaneously use the advantages of both, i.e., preserving the intrinsic structure information by MV-NTF and exploiting detailed spatial information by pixelwise NMF. A weighted nonlocal low-rank tensor decomposition was proposed in [34], where TV regularization and collaborative sparseness were also imposed. Double-weighted sparse NTF (DWSNTF) was proposed in [35], in which combinatorial $\ell_{1,1}$ norm was incorporated into MV-NTF with a multiplicative two-part weight tensor. One of the parts enhanced sparsity in the abundance maps, and the other part was embedded to take into account the detailed spatial information.

Sparse and low-rank constraints were imposed on the MV-NTF framework by Zheng et al. [36], iteratively through ℓ_1 and nuclear norms. A combination of the TV and low-rank regularizers, together with a regularizer based on nonlocal tensor similarity was proposed in [37] to achieve smooth abundance maps with preserving edges. Weighted nuclear norm and $\ell_{1/2}$ sparsity regularizers were also applied on the MV-NTF framework in [38]. Feng et al. [39] provided a weighted group sparsity-constrained tensor decomposition, in which weighted combinatorial $\ell_{2,1}$ norm and also weighted TV norm were included in the MV-NTF problem. Sparsity enhanced convolutional decomposition, named SeCoDe [40], was another method based on BTD. This method employed a convolutional operation to establish the relationship between adjacent pixels, and also used a two-layer sparse regularization term to exploit structural information of abundance maps. To expedite the process of BTD and make it more compatible with structural prior information, a

method was proposed in [41] that reexpressed BTD in the form of a constrained matrix factorization problem. A hypergraph regularized BTD-based method was also proposed in [29] to better reflect the relationship of neighboring pixels in the HU process via NTF.

Although NTF-based HU methods exhibit impressive capabilities, limitations arise when attempting to incorporate supplementary prior information into their problem. Moreover, sometimes the obtained results are physically uninterpretable [29]. One of the main associated limitations is the imposition of the ℓ_1 norm sparsity constraint together with the strict ASC. This is due to the inconsistency of the mentioned constraints, so that the convergence of one leading to the nonconvergence of the other [42]. Therefore, as mentioned earlier, the extracted abundances lack physical significance, meaning they are neither sparse nor sum-to-one. This problem manifests more prominently in NTF-based methods, where ASC is applied on the main objective function as a separate term. To alleviate this problem, in this study, we introduce a new general ℓ_q norm regularization term, which simultaneously establish sparsity and ASC. Notably, the proposed term leverages a relaxed form of ASC, where the sum of abundances in a pixel is not strictly constrained to be equal to one. Then, we incorporate the mentioned regularization term into the combinatorial NMF-NTF problem, making it beneficiary of both NMF preservation of local spatial information and NTF avoiding structural information loss. Totally, the main contributions of our proposed method, called matrix-tensor-based HU method with general ℓ_q norm regularization (MTUHL$_q$), are given as follows.

1) To take the advantages of two conflicting aspects, the sparsity and ASC, we have introduced a novel general regularization term capable of addressing both simultaneously. This regularization term relies on the ℓ_q norm, adaptable to various optimization scenarios, and encompasses a relaxed version of ASC, making it more suitable and feasible compared to a strict implementation.
2) To the best of our knowledge, while the ℓ_1 or other norm regularizers have been previously introduced in the NTF-based unmixing problem, this is the first instance of incorporating the ℓ_q norm into the problem of NTF.
3) While ℓ_q norm regularization-based unmixing is typically optimized in a complicated procedure using the majorization-minimization algorithm, we have implemented it in a much simpler form. We employ a straightforward optimization scheme using a multiplicative iterative algorithm, which largely enhances ease of implementation. Through extensive experiments conducted on synthetic and real hyperspectral datasets, our method demonstrates a high capability in accurately extracting endmembers and their corresponding abundances. Additionally, we analyze the impact of various parameters within our proposed model on synthetic hyperspectral data.

The rest of this article is organized as follows. First, in Section II, the preliminaries of the proposed method including tensor algebra and notations, LMM, and MV-NTF are explained. Our proposed objective function for HU and updating rules for its minimization are introduced in detail in Section III. Experiments on the synthetic and real HSIs are analyzed and discussed in Section IV. Finally, Section VI concludes this article.

II. Backgrounds

A. Tensor: Notations and Algebra

Tensor is a multidimensional array, defined by a parameter called order [22]. The order of a tensor is the number of its dimensions. In the following, a few definitions are provided. In this article, a scaler (zero-order tensor) is represented by a lowercase letter y, a vector (first-order tensor) is shown by a bold lowercase $\boldsymbol{y}$, a matrix (second-order tensor) is given by a bold capital letter $\boldsymbol{Y}$, and a tensor is denoted by a calligraphic letter $\mathcal{Y}$.

Definition 1. (*Tensor Mode*)*:* Tensor's different dimensions are called modes. Let suppose $\mathcal{Y} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is a Nth-order tensor. Then it has N modes. For example, a matrix has two modes: column mode and row mode.

Definition 2. (*Tensor Fiber*)*:* A fiber is a 1-D section of a tensor, extracted by fixing the indices of all the dimensions except one.

Definition 3. (*Tensor Slice*)*:* A slice is a 2-D section of a tensor, obtained by fixing all but two indices.

Definition 4. (*Mode-n Unfolding*)*:* Mode-n unfolding is the process of matricization of a tensor in such a way that constructs a matrix containing all the mode-n vectors of a tensor. For third-order tensor $\mathcal{Y} \in \mathbb{R}^{I \times J \times K}$, three unfolded matrices can be obtained as follows:

$$\begin{aligned}(\boldsymbol{Y}_{(1)})_{(j-1)K+k,i} &= y_{ijk}\\ (\boldsymbol{Y}_{(2)})_{(k-1)I+i,j} &= y_{ijk}\\ (\boldsymbol{Y}_{(3)})_{(i-1)J+j,k} &= y_{ijk}.\end{aligned} \tag{1}$$

Definition 5. (*Outer Product*)*:* The outer product of two tensors $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_P}$ and $\mathcal{B} \in \mathbb{R}^{J_1 \times \cdots \times J_Q}$ is a higher order tensor $\mathcal{C} \in \mathbb{R}^{I_1 \times \cdots \times I_P \times J_1 \times \cdots \times J_Q}$, whose elements are calculated as follows:

$$\mathcal{C}_{i_1 i_2 \cdots i_p j_1 j_2 \cdots j_Q} = (\mathcal{A} \circ \mathcal{B})_{i_1 i_2 \cdots i_p j_1 j_2 \cdots j_Q} = a_{i_1 i_2 \cdots i_p} b_{j_1 j_2 \cdots j_Q}. \tag{2}$$

Definition 6. (*Kronecker Product*)*:* The Kronecker product of two matrices $\boldsymbol{A} \in \mathbb{R}^{I \times J}$ and $\boldsymbol{B} \in \mathbb{R}^{K \times R}$ is an $IK \times JR$ block matrix, calculated as

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{bmatrix} a_{11}\boldsymbol{B} & \cdots & a_{1J}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{I1}\boldsymbol{B} & \cdots & a_{IJ}\boldsymbol{B} \end{bmatrix}. \tag{3}$$

Definition 7. (*Khatri-Rao Product*)*:* Let assume two matrices $\boldsymbol{A} \in \mathbb{R}^{I \times K}$ and $\boldsymbol{B} \in \mathbb{R}^{J \times K}$, which have the same number of columns. The Khatri–Rao product of these two matrices is matrix $\boldsymbol{C} \in \mathbb{R}^{IJ \times K}$, which is achieved as

$$\boldsymbol{C} = (\boldsymbol{A} \odot \boldsymbol{B}) = (\boldsymbol{a}_1 \otimes \boldsymbol{b}_1 \cdots \boldsymbol{a}_K \otimes \boldsymbol{b}_K). \tag{4}$$

Definition 8. ($\bar{\odot}$ *Operation*)*:* If $\boldsymbol{A} = [\boldsymbol{A}_1 \cdots \boldsymbol{A}_R]$ and $\boldsymbol{B} = [\boldsymbol{B}_1 \cdots \boldsymbol{B}_R]$ are two block matrices with the same number of

submatrices, then

$$\boldsymbol{A}\bar{\odot}\boldsymbol{B} = (\boldsymbol{A}_1 \otimes \boldsymbol{B}_1 \cdots \boldsymbol{A}_R \otimes \boldsymbol{B}_R). \tag{5}$$

B. Linear Mixing Model

As already mentioned, the LMM [3] considers the macroscopic factor in the mixture of pixels, i.e., the low resolution of the sensor. This model assumes that the reflectance value of a given pixel is a linear combination of its constituent spectral signatures and their corresponding abundances. Let $\boldsymbol{y} \in \mathbb{R}^{K\times 1}$ be a pixel of an HSI with K spectral bands, then under the LMM we have

$$\boldsymbol{y} = \boldsymbol{C}\boldsymbol{s} + \boldsymbol{\eta} \tag{6}$$

where $\boldsymbol{s} \in \mathbb{R}^{R\times 1}$ and $\boldsymbol{\eta} \in \mathbb{R}^{K\times 1}$ are the abundances and random noise vectors, respectively. Moreover, $\boldsymbol{C} \in \mathbb{R}^{K\times R}$ is the matrix of spectral signatures of the endmembers, called endmembers matrix. R is the number of endmembers, which is usually obtained using some subspace identification or intrinsic dimensionality estimation methods [43], [44], [45], [46]. Considering all pixels of the HSI, (6) can be extended as follows:

$$\boldsymbol{Y} = \boldsymbol{C}\boldsymbol{S} + \boldsymbol{H} \tag{7}$$

where $\boldsymbol{Y} \in \mathbb{R}^{K\times N}$, $\boldsymbol{S} \in \mathbb{R}^{R\times N}$, and $\boldsymbol{H} \in \mathbb{R}^{K\times N}$ are, respectively, HSI, abundances, and random noise matrices, and N is the total number of pixels.

Abundance nonnegativity constraint (ANC) and ASC are also imposed on the LMM problem. ANC ensures the nonnegativity of the abundances, while ASC enforces that the sum of the endmembers' abundances in a pixel is equal to one. Consequently, HU under the LMM is the process of finding matrices $\boldsymbol{C}$ and $\boldsymbol{S}$ as

$$\begin{aligned} \min_{\boldsymbol{C},\boldsymbol{S}} \quad & \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{C}\boldsymbol{S}\|_F^2 \\ \text{s.t.} \quad & \boldsymbol{C}, \boldsymbol{S} \geq 0 \\ & \sum_i \boldsymbol{S}_{ij} = 1 \; ; \; j = 1, \ldots, N \end{aligned} \tag{8}$$

where $\|.\|_F^2$ is the Frobenius norm of a matrix.

C. Matrix Vector Nonnegative Tensor Decomposition

MV-NTF decomposes a 3-D HSI tensor into a set of tensors, each of which is factorized as the outer product of a matrix and a vector, i.e., abundance matrix and endmember vector [30]. This can be defined as the following model:

$$\begin{aligned} \mathcal{Y} &= \sum_{r=1}^{R} \boldsymbol{E}_r \circ \boldsymbol{c}_r + \mathcal{H} \\ &= \sum_{r=1}^{R} \boldsymbol{A}_r \boldsymbol{B}_r^T \circ \boldsymbol{c}_r + \mathcal{H} \end{aligned} \tag{9}$$

where $\mathcal{Y} \in \mathbb{R}^{I\times J\times K}$ is a $I \times J$ HSI image in K spectral bands. $\boldsymbol{E}_r \in \mathbb{R}^{I\times J}$ and $\boldsymbol{c}_r \in \mathbb{R}^{K\times 1}$ are, respectively, abundance matrix and spectral signature vector of the rth endmember, and $\mathcal{H} \in \mathbb{R}^{I\times J\times K}$ is random noise tensor.

$\boldsymbol{A}_r \in \mathbb{R}^{I\times L_r}$ and $\boldsymbol{B}_r \in \mathbb{R}^{J\times L_r}$ are two factor matrices, used as auxiliary matrices in the process of HSI unmixing. Two block matrices $\boldsymbol{A} = [\boldsymbol{A}_1 \cdots \boldsymbol{A}_R]$ and $\boldsymbol{B} = [\boldsymbol{B}_1 \cdots \boldsymbol{B}_R]$ and endmember matrix $\boldsymbol{C} = [\boldsymbol{c}_1 \cdots \boldsymbol{c}_R]$ are also created using $\boldsymbol{A}_r$, $\boldsymbol{B}_r$, and $\boldsymbol{c}_r$ with $r = 1, \ldots, R$. The parameter L_r is the maximum rank of the abundance map, which is empirically formulated in the previous studies as [30]

$$L_r = \frac{2}{3}\min(I, J) \tag{10}$$

which is rounded to its nearest integer to ensure that L_r is an integer. To solve the model in (9), the minimization problem of the reconstruction error between $\mathcal{Y}$ and its R components is used as follows:

$$\min_{A,B,C} \frac{1}{2}\left\|\mathcal{Y} - \sum_{r=1}^{R} \boldsymbol{A}_r \boldsymbol{B}_r^T \circ \boldsymbol{c}_r\right\|_F^2 \qquad \text{s.t.} \;\; \boldsymbol{A}_r, \boldsymbol{B}_r, \boldsymbol{c}_r \geq 0. \tag{11}$$

As stated in [30], the above optimization problem is solved through the alternating least squares minimization algorithm, which decomposes the main problem into some subproblems. Accordingly, under the framework of the multiplicative updating rules, $\boldsymbol{A}$ and $\boldsymbol{B}$ can be obtained as

$$\boldsymbol{A} \leftarrow \boldsymbol{A} * (\boldsymbol{Y}_{(1)}^T \boldsymbol{M})/(\boldsymbol{A}\boldsymbol{M}^T\boldsymbol{M}) \tag{12}$$

$$\boldsymbol{B} \leftarrow \boldsymbol{B} * (\boldsymbol{Y}_{(2)}^T \boldsymbol{M})/(\boldsymbol{B}\boldsymbol{M}^T\boldsymbol{M}) \tag{13}$$

where in (12) $\boldsymbol{M} = \boldsymbol{B}\bar{\odot}\boldsymbol{C}$ and in (13) $\boldsymbol{M} = \boldsymbol{C}\bar{\odot}\boldsymbol{A}$. Two $*$ and $/$ symbols are also used for elementwise multiplication and division, respectively. Moreover, $\boldsymbol{C}$ is achieved as follows:

$$\boldsymbol{C} \leftarrow \boldsymbol{C} * (\boldsymbol{Y}_{(3)}^T \boldsymbol{M})/(\boldsymbol{C}\boldsymbol{M}^T\boldsymbol{M}) \tag{14}$$

where $\boldsymbol{M} = [(\boldsymbol{A}_1 \odot \boldsymbol{B}_1)1_{L_r} \cdots (\boldsymbol{A}_R \odot \boldsymbol{B}_R)1_{L_r}]$ and 1_{L_r} is a L_r-vector containing all ones.

III. Matrix-Tensor HU With General ℓ_q Norm Regularizer (MTHUL$_q$)

In this section, first the proposed regularization term to simultaneously impose sparseness and ASC is introduced. Then, MTHUL$_q$ model for HU is explained. Afterward, the corresponding updating rules and implementation issues are discussed.

A. Proposed Regularization Term

As mentioned in the previous section, sparseness and ASC terms are not consistent [42]. This means that it is not possible to simultaneously fulfill both terms. Therefore, the extracted abundance maps from these methods are neither sparse enough nor the sum-to-one constraint is applied. To tackle this problem and enhance extracted abundance maps, we define a new regularization term that simultaneously applies sparseness and ASC on the abundance maps.

According to ASC, the sum of elements in each column of the abundance matrix $\boldsymbol{S}$ should be equal to one. This means

$$\sum_{i=1}^{R} \boldsymbol{S}_{i,j} = 1 \quad ; \quad j = [1, \ldots, N]. \tag{15}$$

Let us also assume the following constraint:

$$\sum_{i=1}^{R} \boldsymbol{S}_{i,j}^2 = 1 \quad ; \quad j = [1, \ldots, N]. \tag{16}$$

Ideally, to establish the aforementioned relation, each column of the matrix $\boldsymbol{S}$ should contain only one element equal to one, and the rest being zeros. This is exactly the solution of the ℓ_0 norm minimization, which makes a completely sparse solution. Therefore, the constraint in (16) has an extremely sparse property while keeping also the ASC property. A general and more relaxed form of (16) is as

$$\sum_{i=1}^{R} \boldsymbol{S}_{i,j}^q = \alpha \quad ; \quad j = [1, \ldots, N] \tag{17}$$

where $0 < q \leq 2$ and $a \leq 1$. Let suppose $\boldsymbol{S}(:,t) = [\boldsymbol{S}_{1,t}, \ldots, \boldsymbol{S}_{R,t}]^T$ be a sample column of matrix $\boldsymbol{S}$. Then, applying the constraint in (17) on this column gives

$$\boldsymbol{S}_{1,t}^q + \cdots + \boldsymbol{S}_{R,t}^q = \alpha. \tag{18}$$

Two above equations are nonlinear, which require linearization for integration into our formulation. We use the Taylor series expansion for this purpose. Using the first two terms of the Taylor expansion, the above formula can be rewritten in the following form:

$$\begin{aligned}(\boldsymbol{S}_{1,t}^0)^q + q(t_{1,t}^0)^{q-1}(\boldsymbol{S}_{1,t} - \boldsymbol{S}_{1,t}^0) + \cdots + (\boldsymbol{S}_{R,t}^0)^q \\ + q(\boldsymbol{S}_{R,t}^0)^{q-1}(\boldsymbol{S}_{R,t} - \boldsymbol{S}_{R,t}^0) = \alpha\end{aligned} \tag{19}$$

where $\boldsymbol{S}_0(:,t) = [\boldsymbol{S}_{1,t}^0, \ldots, \boldsymbol{S}_{R,t}^0]$ is an initial value for the tth column of $\boldsymbol{S}$. Using simple mathematical operations, (19) can be written as follows:

$$(1-q)f_1 + qf_2 = \alpha \tag{20}$$

where

$$\begin{aligned}f_1 &= (\boldsymbol{S}_{1,t}^0)^q + \cdots + (\boldsymbol{S}_{R,t}^0)^q \\ f_2 &= (\boldsymbol{S}_{1,t}^0)^{q-1}\boldsymbol{S}_{1,t} + \cdots + (\boldsymbol{S}_{R,t}^0)^{q-1}\boldsymbol{S}_{R,t}.\end{aligned} \tag{21}$$

To simplify the above equations, we can reform f_1 as

$$\begin{aligned}f_1 &= \begin{bmatrix}(\boldsymbol{S}_{1,t}^0)^{q-1} & \cdots & (\boldsymbol{S}_{R,t}^0)^{q-1}\end{bmatrix} \begin{bmatrix}\boldsymbol{S}_{1,t}^0 \\ \vdots \\ \boldsymbol{S}_{R,t}^0\end{bmatrix} \\ &= ((\boldsymbol{S}_0(:,t))^{q-1})^T \boldsymbol{S}_0(:,t)\end{aligned} \tag{22}$$

and f_2 as

$$\begin{aligned}f_2 &= \begin{bmatrix}(\boldsymbol{S}_{1,t}^0)^{q-1} & \cdots & (\boldsymbol{S}_{R,t}^0)^{q-1}\end{bmatrix} \begin{bmatrix}\boldsymbol{S}_{1,t} \\ \vdots \\ \boldsymbol{S}_{R,t}\end{bmatrix} \\ &= ((\boldsymbol{S}_0(:,t))^{q-1})^T \boldsymbol{S}(:,t).\end{aligned} \tag{23}$$

By generalizing the above relationships to all columns of matrix $\boldsymbol{S}$, we have

$$f_1^g = \left((\boldsymbol{S}_0^{q-1})^T \boldsymbol{S}_0\right) * \boldsymbol{I}_N \tag{24}$$

and

$$f_2^g = \left((\boldsymbol{S}_0^{q-1})^T \boldsymbol{S}\right) * \boldsymbol{I}_N \tag{25}$$

where f_1^g and f_2^g are the generalized forms of f_1 and f_2, and $\boldsymbol{I}_N$ is the identity matrix of size N. Thus, (20) can be rewritten as

$$(1-q)f_1^g + qf_2^g = \alpha \boldsymbol{I}_N. \tag{26}$$

Then, the above equation can subsequently be simplified as

$$\boldsymbol{F} * \boldsymbol{I}_N = 0 \tag{27}$$

where

$$\boldsymbol{F} = (1-q)\left((\boldsymbol{S}_0^{q-1})^T \boldsymbol{S}_0\right) + q\left((\boldsymbol{S}_0^{q-1})^T \boldsymbol{S}\right) - \alpha \boldsymbol{I}_N. \tag{28}$$

Finally, the regularization term on the abundance matrix $\boldsymbol{S}$ can be defined as

$$\Phi(\boldsymbol{S}) = \|\boldsymbol{F} * \boldsymbol{I}_N\|_F^2. \tag{29}$$

B. MTHUL$_q$ Model

Inspired by the study of Li et al. [33], a matrix-tensor model for HU can be considered as

$$\begin{aligned}\min_{A,B,C} \ & \frac{1}{2}\|\mathcal{Y} - \sum_{r=1}^{R} \boldsymbol{A_r}\boldsymbol{B_r^T} \circ \boldsymbol{c_r}\|_F^2 + \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{CS}\|_F^2 \\ & + \frac{\delta}{2}\|\mathcal{S} - \sum_{r=1}^{R} (\boldsymbol{A_r}\boldsymbol{B_r}^T) \circ \boldsymbol{i_r}\|_F^2 \\ & \text{s.t.} \ \ \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{S} \geq 0 \ \ ; \ \ \boldsymbol{1}_R^T \boldsymbol{S} = \boldsymbol{1}_N^T.\end{aligned} \tag{30}$$

In the above equation, $\boldsymbol{i}_r$ is a column vector of the identity matrix $\boldsymbol{I} = [\boldsymbol{i}_1 \cdots \boldsymbol{i}_R] \in \mathbb{R}^{R \times R}$. δ controls the similarity between abundances, obtained by NMF and NTF, which are, respectively, in $\boldsymbol{S}$ and $\sum_{r=1}^{R} \boldsymbol{A_r}\boldsymbol{B_r}^T$. Moreover, $\mathcal{S} \in \mathbb{R}^{I \times J \times R}$ is the third-order tensor, obtained as follows:

$$\mathcal{S}(:,:,r) = \text{Mat}(\boldsymbol{S}(r,:)) \quad ; \quad r = [1, \ldots, R] \tag{31}$$

where Mat(.) denotes the matricization operator.

As is clear in (30), ASC is taking into account the constraints of the problem. To impose sparseness into the objective function of matrix-tensor HU in such a way that it does not conflict with the ASC, the proposed regularization term in (29) is added to the problem in (30) as

$$\begin{aligned}\min_{A,B,C} \ & \frac{1}{2}\|\mathcal{Y} - \sum_{r=1}^{R} \boldsymbol{A_r}\boldsymbol{B_r^T} \circ \boldsymbol{c_r}\|_F^2 + \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{CS}\|_F^2 \\ & + \frac{\delta}{2}\|\mathcal{S} - \sum_{r=1}^{R} (\boldsymbol{A_r}\boldsymbol{B_r}^T) \circ \boldsymbol{I_r}\|_F^2 + \frac{\lambda}{2}\Phi(\boldsymbol{S}) \\ & \text{s.t.} \ \ \boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{S} \geq 0 \ \ ; \ \ \boldsymbol{1}_R^T \boldsymbol{S} = \boldsymbol{1}_N^T.\end{aligned} \tag{32}$$

C. Update Rules

To solve the proposed model in (32), first, the main problem should be split into some subproblems, and then, optimize each of them through the multiplicative iterative algorithm.

Update $\boldsymbol{A}$ and $\boldsymbol{B}$: Assuming that other variables are fixed, the subproblem for optimizing $\boldsymbol{A}$ is as follows:

$$\mathcal{J}(\boldsymbol{A}) = \frac{1}{2}\|\boldsymbol{Y}_{(1)} - \boldsymbol{M}\boldsymbol{A}^{\boldsymbol{T}}\|_F^2 + \frac{\delta}{2}\|\boldsymbol{S}_{(1)} - (\boldsymbol{Q}\boldsymbol{A}^{\boldsymbol{T}})\|_F^2 + \mathtt{Tr}(\boldsymbol{\Gamma}\boldsymbol{A}) \tag{33}$$

where $\boldsymbol{M} = \boldsymbol{B}\bar{\odot}\boldsymbol{C}$ and $\boldsymbol{Q} = \boldsymbol{B}\bar{\odot}\boldsymbol{I_R}$. $\boldsymbol{S_{(k)}}$ and $\boldsymbol{\Gamma}$ are also mode-k matricization of the tensor $\mathcal{S}$ and Lagrange multiplier, respectively. According to the Karush–Kuhn–Tucker (KKT) conditions, the following equations should be satisfied to optimize the model in (33):

$$-\boldsymbol{Y_{(1)}}^T\boldsymbol{M} + \boldsymbol{A}\boldsymbol{M}^T\boldsymbol{M} - \delta\boldsymbol{S_{(1)}}^T\boldsymbol{Q} + \delta\boldsymbol{A}\boldsymbol{Q}^T\boldsymbol{Q} + \boldsymbol{\Gamma} = 0 \tag{34}$$

$$\boldsymbol{\Gamma} * \boldsymbol{A} = 0. \tag{35}$$

Substituting (35) into (34) gives

$$\boldsymbol{A} \leftarrow \boldsymbol{A} * \left(\boldsymbol{Y_{(1)}}^T\boldsymbol{M} + \delta\boldsymbol{S_{(1)}}^T\boldsymbol{Q}\right) / \left(\boldsymbol{A}\boldsymbol{M}^T\boldsymbol{M} + \delta\boldsymbol{A}\boldsymbol{Q}^T\boldsymbol{Q}\right). \tag{36}$$

Similarly, the update rule for $\boldsymbol{B}$ is achieved as

$$\boldsymbol{B} \leftarrow \boldsymbol{B} * \left(\boldsymbol{Y_{(2)}}^T\boldsymbol{M} + \delta\boldsymbol{S_{(2)}}^T\boldsymbol{Q}\right) / \left(\boldsymbol{B}\boldsymbol{M}^T\boldsymbol{M} + \delta\boldsymbol{B}\boldsymbol{Q}^T\boldsymbol{Q}\right) \tag{37}$$

where $\boldsymbol{M} = \boldsymbol{C}\bar{\odot}\boldsymbol{A}$ and $\boldsymbol{Q} = \boldsymbol{I_R}\bar{\odot}\boldsymbol{A}$.

Update $\boldsymbol{C}$: The subproblem of optimizing $\boldsymbol{C}$ is as follows:

$$\mathcal{J}(\boldsymbol{C}) = \frac{1}{2}\|\boldsymbol{Y_{(3)}} - \boldsymbol{M}\boldsymbol{C}^T\|_F^2 + \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{C}\boldsymbol{S}\|_F^2 + \mathtt{Tr}(\boldsymbol{\Gamma}\boldsymbol{C}) \tag{38}$$

where $\boldsymbol{M} = [(\boldsymbol{A}_1 \odot \boldsymbol{B}_1)\boldsymbol{1}_L \cdots (\boldsymbol{A}_R \odot \boldsymbol{B}_R)\boldsymbol{1}_L]$ and $\boldsymbol{1}_L$ is the summation matrix of size L containing all ones. Hence, the update rule for $\boldsymbol{C}$ is as follows:

$$\boldsymbol{C} \leftarrow \boldsymbol{C} * \left(\boldsymbol{Y_{(3)}}^T\boldsymbol{M} + \boldsymbol{Y}\boldsymbol{S}^T\right) / \left(\boldsymbol{C}\boldsymbol{M}^T\boldsymbol{M} + \boldsymbol{C}\boldsymbol{S}\boldsymbol{S}^T\right). \tag{39}$$

Update $\boldsymbol{S}$: Three terms in the main problem of the MTHUL$_q$ in (32) are functions of $\boldsymbol{S}$. The first term is the NMF problems, whose optimization process has been extensively investigated in the previous studies [8].

The second part, which controls the similarity between NMF and NTF extracted abundances, can be seen as

$$\|\boldsymbol{S} - \boldsymbol{S}_{\text{NTF}}\|_F^2 \tag{40}$$

where

$$\boldsymbol{S}_{\text{NTF}}(r,:) = \text{vec}(\boldsymbol{A}_r\boldsymbol{B}_r^T) \;\; ; \;\; r = 1, \ldots, R \tag{41}$$

where $\mathtt{vec}$ is the vectorization operator.

The third part of optimizing $\boldsymbol{S}$ is the minimization of the proposed regularization term, $\Phi(\boldsymbol{S})$. Accordingly, the subproblem of optimizing $\boldsymbol{S}$ is as

$$\mathcal{J}(\boldsymbol{S}) = \frac{1}{2}\|\boldsymbol{Y}_f - \boldsymbol{C}_f\boldsymbol{S}\|_F^2 + \frac{\delta}{2}\|\boldsymbol{S} - \boldsymbol{S}_{\text{NTF}}\|_F^2 + \frac{\lambda}{2}\Phi(\boldsymbol{S}) + \mathtt{Tr}(\boldsymbol{\Gamma}\boldsymbol{S}) \tag{42}$$

in which

$$\boldsymbol{Y}_f = \begin{bmatrix} \boldsymbol{Y} \\ \beta\boldsymbol{1}_N^T \end{bmatrix}, \quad \boldsymbol{C}_f = \begin{bmatrix} \boldsymbol{C} \\ \beta\boldsymbol{1}_R^T \end{bmatrix}. \tag{43}$$

Along with the proposed $\Phi(\boldsymbol{S})$, the two above assumptions promote ASC. Taking the partial derivative of $\mathcal{J}(\boldsymbol{S})$ with respect to $\boldsymbol{S}$

$$\nabla_{\boldsymbol{S}}\mathcal{J}(\boldsymbol{S}) = -\boldsymbol{C}_f^T\boldsymbol{Y}_f + \boldsymbol{C}_f^T\boldsymbol{C}_f\boldsymbol{S} - \delta\boldsymbol{S}_{\text{NTF}} + \delta\boldsymbol{S} + (\lambda f * \boldsymbol{I}_N)(q\boldsymbol{S}_0^{q-1})^T + \boldsymbol{\Gamma}. \tag{44}$$

Substituting f from (28) into the above equation and applying KKT conditions, the update rule for $\boldsymbol{S}$ is achieved as

$$\boldsymbol{S} \leftarrow \boldsymbol{S} * \left(\boldsymbol{C}_f^T\boldsymbol{Y}_f + \delta\boldsymbol{S}_{\text{NTF}} + \lambda\alpha(q\boldsymbol{S}_0^{q-1})^T\right) / \left(\boldsymbol{C}_f^T\boldsymbol{C}_f\boldsymbol{S} + \delta\boldsymbol{S} + \lambda\boldsymbol{G}\right) \tag{45}$$

where

$$\boldsymbol{G} = \left[\left((1-q)(\boldsymbol{S}_0^{q-1})^T\boldsymbol{S}_0 + q(\boldsymbol{S}_0^{q-1})^T\boldsymbol{S}\right) * \boldsymbol{I}_N\right](q\boldsymbol{S}_0^{q-1})^T. \tag{46}$$

Algorithm 1: Pseudo-Code of the Proposed MTHUL$_q$ Algorithm.

Input: HSI in 2D and 3D versions, $\boldsymbol{Y} \in \mathbb{R}^{K\times N}$ and $\mathcal{Y} \in \mathbb{R}^{I\times J\times K}$; Number of endmembers R; Regularization parameters λ, δ, and β; Parameter of the proposed regularization term, q and α.
Output: $\boldsymbol{C}$ and $\boldsymbol{S}$.
1: Initialize $\boldsymbol{C}$ and $\boldsymbol{S}$ using VCA-FCLS [47], and $\boldsymbol{A}$ and $\boldsymbol{B}$ using NMF.
2: **while** stopping criteria are not met **do**
3: Update $\boldsymbol{A}$ with Eq. (36).
4: Update $\boldsymbol{B}$ with Eq. (37).
5: Update $\boldsymbol{C}$ with Eq. (39).
6: Update $\boldsymbol{S}$ with Eq. (45).
7: **end while**

D. Implementation Issue and Computational Complexity

Two stopping criteria are also considered including the number of iterations and norm of relative changes of the abundance matrix as

$$\epsilon = \|\boldsymbol{S}^{i+1} - \boldsymbol{S}^i\|_2 \tag{47}$$

where $\boldsymbol{S}^i$ and $\boldsymbol{S}^{i+1}$ are the extracted abundance matrices in the ith and $(i+1)$th iterations, respectively. In this study, the maximum number of iterations is equal to 2000 and $\epsilon = 0.005$.

The complete pseudocode of the proposed MTHUL$_q$ method is presented in Algorithm 1. Clearly, this algorithm is primarily driven by matrix multiplications, which are smaller in scale. The parts involving exponentiating a matrix are constant. Totally, there are four updating steps including (36), (37), (39), and (45), which are the cores

of the process in each iteration. The number of floating-point operations, needed for (36), is $RL(4I + 2IJK + 2JKRL + 4IRL + JK + 2IJR + 2JR^2\,L + JR)$. For (37), this number is $RL(4\,J + 2IJK + 2IKRL + 4JRL + IK + 2IJR + 2IR^2\,L + IR)$. Moreover, for (39) the number is $R(4\,K + 4IJK + 4IJR + 4KR + 3IJL)$, and it is $R(6IJ + 2IJK + 2KR + 2R + 2IJR + I^2\,J^2R + I^2\,J^2) + I^3\,J^3$. To sum up, the overall computational complexity of MTHUL$_q$ is $\mathcal{O}(IJKRL + IKR^2\,L^2 + I^2\,J^2R^2 + I^3\,J^3)$.

IV. Experiments

To examine the performance of the proposed method, we have conducted a series of experiments in this section. The first part of the experiments has been applied to the synthetic data to choose the optimum values for the MTHUL$_q$ parameters including λ, q, and α. The proposed method also have two other parameters δ and β. However, these two parameters have been set as in [33] to have a fair comparison.

The robustness of the MTHUL$_q$ against different SNR values and the number of endmembers has been also evaluated in this part of our experiments. In the last part, real HSIs have been utilized to represent the performance of the proposed method in real-world scenarios.

A number of state-of-the-art HU algorithms including graph $\ell_{1/2}$ sparse NMF (GLNMF) [48], MV-NTF [30] and its ASC included version (MV-NTF-S), SCNMTF [33], and EWSP-NTF [38] have been implemented for comparison. GLNMF is a powerful NMF-based algorithm, which has been used as a benchmark method in different HU studies. MV-NTF is the basis of many NTF-based HU methods and has been used in many studies as a competing method, along with MV-NTF-S. SCNMTF is a combinatorial NMF and NTF-based method, whose objective function is very similar to the objective function of our MTHUL$_q$ with just one difference, which is the sparse regularization term. EWSP-NTF is a recently introduced NTF-based method. This method incorporates weighted nuclear norm and $\ell_{1/2}$ sparse regularizers into the MV-NTF problem. It should be noted that all tunable parameters of the aforementioned competing methods have been set according to their original references.

Besides the visual evaluation, the results of the different methods have been quantitatively compared using two metrics, i.e., spectral angle distance (SAD) and root-mean-squares error (RMSE). SAD is a criterion to examine the performance of HU algorithms in endmembers estimation and is calculated as

$$\text{SAD}_r = \arccos\left(\frac{\boldsymbol{c}_r^T\hat{\boldsymbol{c}}_r}{\|\boldsymbol{c}_r\|\|\hat{\boldsymbol{c}}_r\|}\right) \tag{48}$$

where $\boldsymbol{c}_r$ and $\hat{\boldsymbol{c}}_r$ are true and estimated spectral signatures for the rth endmember, respectively. RMSE is also a criterion to determine the dissimilarity between the true and estimated abundance maps. It is obtained as follows:

$$\text{RMSE} = \left(\frac{1}{N}|\boldsymbol{S} - \hat{\boldsymbol{S}}|^2\right)^{\frac{1}{2}} \tag{49}$$

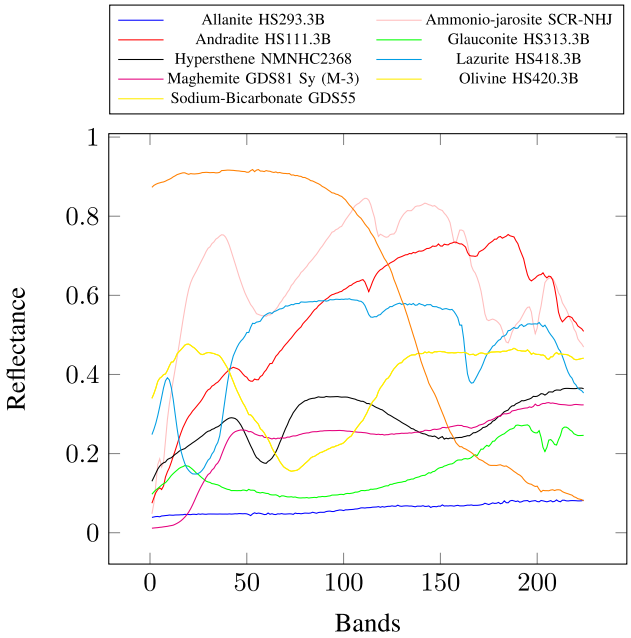

Fig. 1. Endmembers selected from the USGS library to generate synthetic data.

where $\boldsymbol{S}$ and $\hat{\boldsymbol{S}}$ are, respectively, true and estimated abundance maps.

A. Synthetic Data Experiments

In this part of the experiments, several synthetic data were generated to analyze the accuracy of the experimental methods' outcomes. We designed three groups of experiments on the synthetic data, each of which followed a specific goal. The first group of experiments is established to analyze the impact of regularization parameter λ, and two regularization term's parameters, q and α. Then, experiments are conducted on the synthetic data with different numbers of endmembers. Finally, the performance of experimental methods in the different noise levels is investigated in the third group of experiments on synthetic data. In each group of the experiments, the noise level and the number of endmembers are assigned based on their goals. Nevertheless, the same procedure is used to generate endmembers with a given noise level and the number of endmembers.

To generate synthetic data, a number of pure signatures are selected from the USGS library [49] as endmembers, which cover wavelengths in the range of 0.38–2.5 μm in 224 spectral bands with 10 nm bandwidth. Selected endmembers from the USGS library are shown in Fig. 1. In each of the following experiments, some or all of these endmembers have been utilized according to the problem in hand.

Abundance maps, and subsequently synthetic image, are generated using the procedure stated in [50] as in the following steps.

1) An image with size $z^2 \times z^2$ is divided into z^2 blocks, each of which is randomly assigned an endmember.
2) A low-pass filter with $z + 1$ window size is used to mix the pixels of the image.

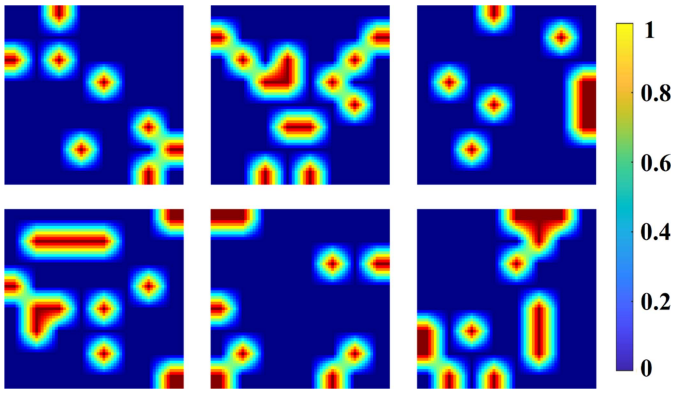

Fig. 2. Abundance maps of a sample synthetic data with six endmembers.

3) To prevent the existence of the pure pixels and provide a highly mixed scenario, pixels with abundance fractions larger than a predefined value are replaced by a mixture of all endmembers with equal proportions.
4) Finally clean image obtained from the previous steps is polluted by zero-mean white Gaussian noise whose SNR is defined as

$$\text{SNR} = 10\log_{10}\frac{E[y^T y]}{E[e^T e]} \tag{50}$$

where y and e are, respectively, the clean signal and the noise, and $E[.]$ is the expectation operator.

The noise-free abundance maps of a sample synthetic data with 6 endmembers, generated through the above procedure, are illustrated in Fig. 2.

1) Parameters Setting: In this part of our experiments, we analyze the impact of the parameters λ, q, and α on the unmixing outcomes. The two other regularization parameters have been set according to [33] as $\delta = 10$ and $\beta = 10$ for a fair comparison. The tests of this part were performed on the synthetic data with six endmembers and SNR $= \{30, 50\}$ dB. Different SNR values have been applied to prevent influencing noise level.

First, we analyze the regularization parameter λ. To this end, we set $q = 2$ and $\alpha = 1$, and select λ from the set $\{10, 5, 1, 0.5, 0.1, 0.05, 0.01\}$. The results of this test are illustrated in Fig. 3. As shown in the figure, at SNR $= 50$, $\lambda = \{5, 10\}$ had the best performance in estimating endmembers. However, a general examination of the two graphs in this figure indicates that the proposed method had proper results when λ is less than 1. Finally, according to the total results of the experiments, we can consider $\lambda = 0.1$ to be a suitable choice in the remaining experiments on the synthetic data.

After optimizing λ, in this part, to investigate the effect of norm value q on the results of the proposed MTHUL$_q$, we conduct the experiments on the previously mentioned data with $q \in \{\frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{3}{2}, 2\}$, while keeping $\lambda = 0.1$ and $\alpha = 1$. The results obtained from these experiments are shown in Fig. 4. As is clear in this figure, all the q values have achieved acceptable SAD and RMSE values, and the difference between results is mostly negligible. However, the best results have been gained when $q = \{0.25, 0.5\}$. Although $q = 0.25$ has had better SAD values among these two values, its corresponding RMSE values in the case with SNR $= 50$ dB are the highest. Consequently, we select $q = 0.5$ as the best norm value in the synthetic datasets' experiments.

Finally, to determine the optimum α value, a set of experiments have been conducted, in which α has been selected from the range $[0 : 0.1 : 1]$. In these experiments, $\lambda = 0.1$ and $q = 0.5$ have been applied according to the previous investigations. SAD and RMSE values obtained from these experiments are shown in Fig. 5. As is clear, the SAD and RMSE values do not have the same trends when α increases. For example, $\alpha = 0$ yields the best SAD values while having the highest RMSEs. Totally, it is inferred that α values in the range $[0.5 - 0.9]$ can be proper choices. Therefore, the middle point of this range, i.e., $\alpha = 0.7$, which is more relaxed than $\alpha = 1$ is selected for synthetic data.

2) Number of Endmembers Analysis: To examine the performance of the experimental methods in changing the number of endmembers, synthetic data using 3, 6, and 9 endmembers and SNR $= \{30, 60\}$ dB were generated. Two different SNR values were considered to prevent the influence of noise level. The results of this experiment have been shown in Tables I and II.

Table I shows the mean SAD values of the extracted endmembers. As is reported in this table, SCNMTF outperformed other methods in extracting endmembers in all cases. Our proposed method had also proper results in comparison with four other methods. However, as is shown in Table II, the main achievement of the proposed method is high-precision abundance maps, where their corresponding RMSE values are lower than all the methods in all cases.

A remarkable point in the results of Table II is the superiority of MV-NTF-S over MV-NTF. This could be attributed to the inclusion of the ASC constraint in MV-NTF-S, which clearly demonstrates the impact of this constraint on the quality of the extracted abundance maps.

3) Noise Analysis: To investigate the impact of noise level on the performance of the proposed method along with other experimental ones, synthetic data were generated using six endmembers with SNR values from the set $\{30, 40, 50, 60\}$ dB. Endmembers used in this part are also the same as those applied in experiments in Section IV-A1. Moreover, the parameters proposed for the method are $\lambda = 0.1$, $q = 0.5$, and $\alpha = 0.7$, as optimized in the previous experiments for synthetic data. Results obtained from implementing different methods on the synthetic data with different SNR values are illustrated in Fig. 6.

As is clear from Fig. 6(a), although MTHUL$_q$ has higher SAD values than other methods in some cases, the decreasing trend shows its high capability in extracting endmembers, to the extent that in SNR $= 60$ dB, it had the best accuracy along with SCNMTF. For the SAD values, it can be said that GLNMF had the best performance, which is due to using a manifold structure designed to reduce its sensitivity to noise.

In the RMSE metric, as shown in Fig. 6(b), the results of the proposed method are almost always more accurate than the others. As previously mentioned, we focus on the quality of the abundance maps and impose a new regularization parameter on them. Therefore, we expected that the main progress of our

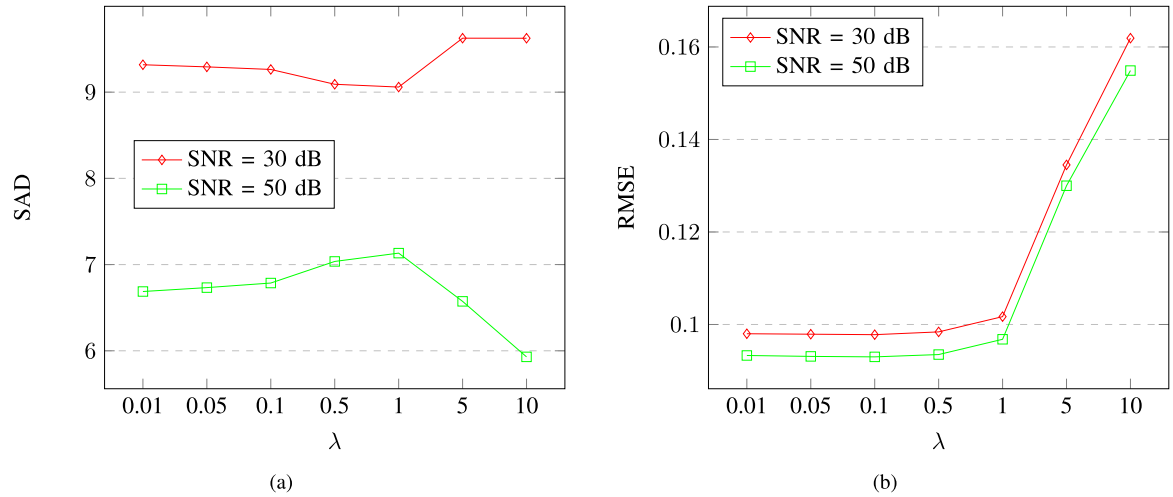

Fig. 3. Analysis of the MTHUL$_q$'s performance against different λ regularization parameter values.

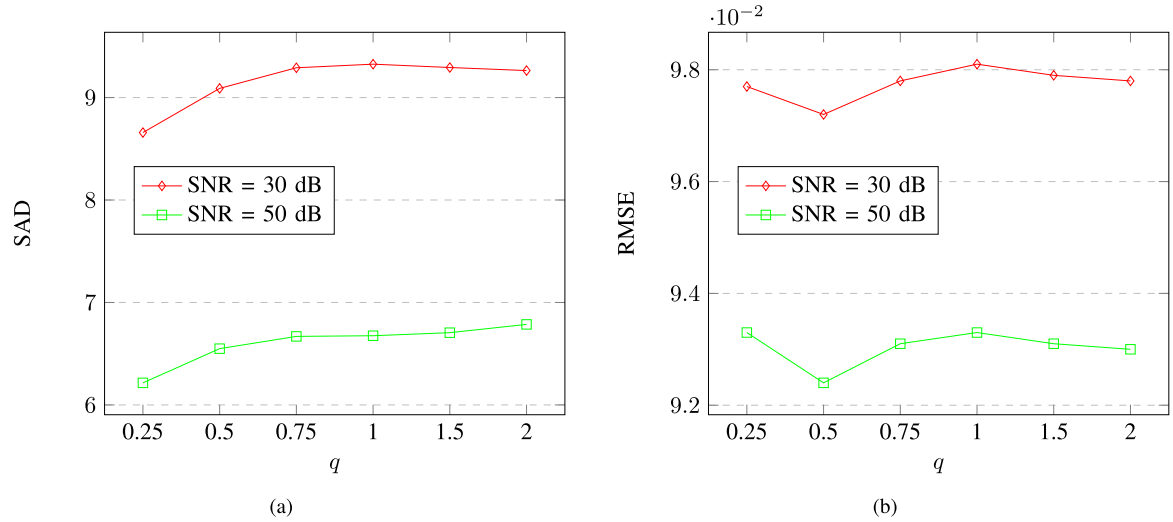

Fig. 4. Analysis of the MTHUL$_q$'s performance against different norm values q.

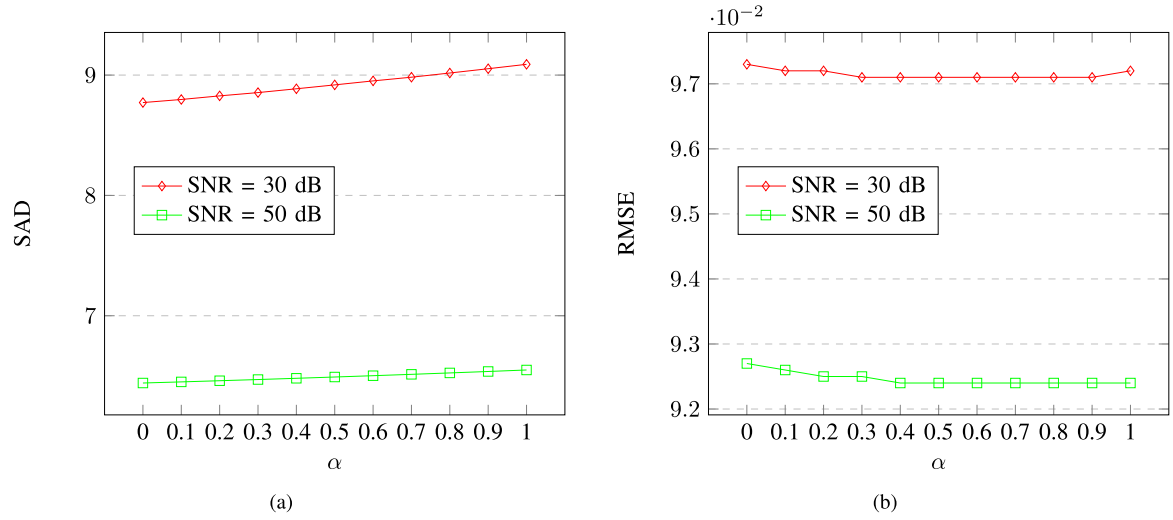

Fig. 5. Analysis of the MTHUL$_q$'s performance against different α values.

TABLE I
SAD VALUES WITH RESPECT TO THE DIFFERENT NUMBER OF ENDMEMBERS

SNR	R	GLNMF	MV-NFT	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
30 dB	3	0.9067	3.8799	6.5728	**0.3077**	0.7016	0.6500
	6	0.9517	6.5909	5.5578	**0.3657**	1.3888	0.7019
	9	1.1231	6.3275	6.6608	**0.4306**	1.4562	0.8240
60 dB	3	0.8637	1.9460	4.9652	**0.2424**	0.6100	0.3397
	6	0.9007	9.3899	8.8868	**0.2227**	0.7338	0.3200
	9	0.8708	14.6741	18.2533	**0.2411**	1.1124	0.5977

The bold values show the smallest SAD values among different methods.

TABLE II
RMSE VALUES WITH RESPECT TO THE DIFFERENT NUMBER OF ENDMEMBERS

SNR	R	GLNMF	MV-NFT	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
30 dB	3	0.0759	0.1458	0.0914	0.0556	0.0638	**0.0503**
	6	0.0532	0.0995	0.0718	0.0361	0.0881	**0.0341**
	9	0.0425	0.0713	0.0544	0.0293	0.0615	**0.0287**
60 dB	3	0.0691	0.1648	0.0996	0.0543	0.0762	**0.0499**
	6	0.0553	0.1629	0.1415	0.0430	0.0674	**0.0405**
	9	0.0441	0.1578	0.1617	0.0301	0.0612	**0.0284**

The bold values show the smallest RMSE values among different methods.

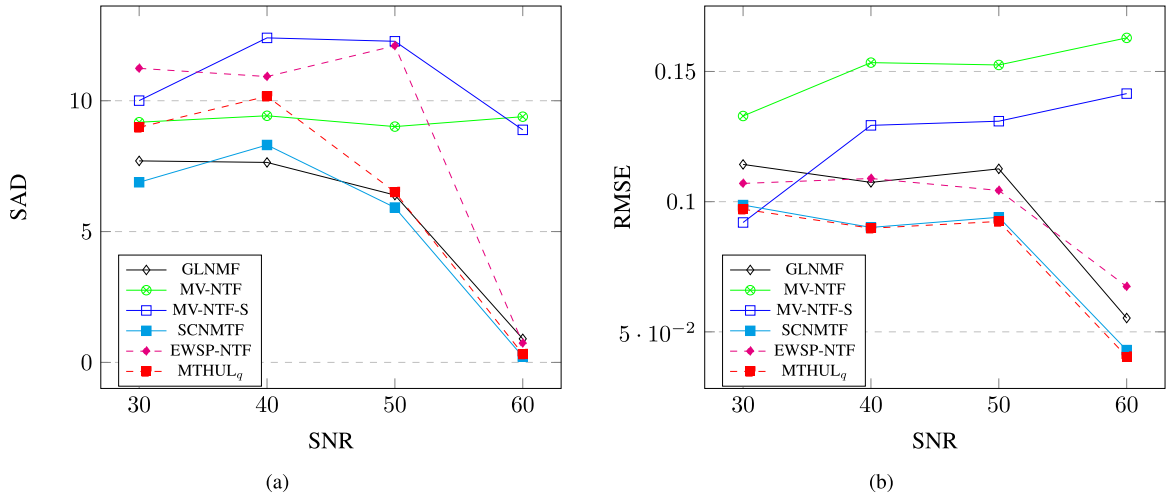

Fig. 6. Analysis of the performance of different experimental methods with respect to the different noise levels.

proposed method would be in improving the accuracy of the final abundance maps, which has been met.

B. Real Data Experiments

To investigate the performance of the proposed method in real-world scenarios, experiments on the three real hyperspectral datasets datasets, including Samson dataset, Jasper Ridge dataset, and Urban dataset [51], have been conducted in the following. Regarding parameters, all of them, except λ, have been set as supposed or tuned in the synthetic data experiments; i.e., $\delta = 10$, $\beta = 10$, $q = 0.5$, and $\alpha = 0.7$. The regularization parameter λ has been set to 5 in the real scenarios.

1) Samson Dataset: Samson dataset is a real hyperspectral dataset, whose real endmembers' signatures and abundance maps are available. It is originally of size 952×952 and has 156 spectral bands. Since the size of this image is too large, in this study, a subset of it, which contains 95×95 pixels, has been applied. There are three endmembers in the Samson dataset including "rock," "tree," and "water." A false-color composite of this dataset is shown in Fig. 7(a).

Results obtained through the implementation of six different experimental methods on the Samson dataset have been reported in Tables III and IV. Table III shows that the EWSP-NTF outperformed other methods for extracting all the endmembers. However, as in Table IV, our proposed MTHUL$_q$ has had the best performance in extracting abundance maps. Extracted endmembers and abundance maps generated by different methods have been, respectively, represented in Figs. 8 and 9. The results of this figure is also illustrate the superiority of EWSP-NTF in endmember extraction, and our proposed MTHUL$_q$ in abundance estimation.

2) Jasper Ridge Dataset: Jasper Ridge dataset is the second real hyperspectral data used to analyze the performance of the experimental methods in real case scenarios. It is a popular hyperspectral dataset in different HU studies, originally of size

TABLE III
SAD VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE SAMSON DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Rock	0.0217	0.0712	0.1060	0.0173	**0.0167**	0.0273
Tree	0.0350	0.0828	0.0885	**0.0293**	0.0509	0.0309
water	0.1882	0.2235	0.2332	0.1661	**0.1052**	0.2195
Average	0.0816	0.1258	0.1426	0.0709	**0.0576**	0.0926

The bold values show the smallest SAD values among different methods.

TABLE IV
RMSE VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE SAMSON DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Rock	**0.1278**	0.1655	0.1900	0.1364	0.1441	0.1384
Tree	0.1422	0.1705	0.1796	0.1293	0.1477	**0.1182**
Water	0.2414	0.3757	0.2874	0.2232	0.2537	**0.1848**
Average	0.1705	0.2372	0.2190	0.1630	0.1818	**0.1472**

The bold values show the smallest RMSE values among different methods.

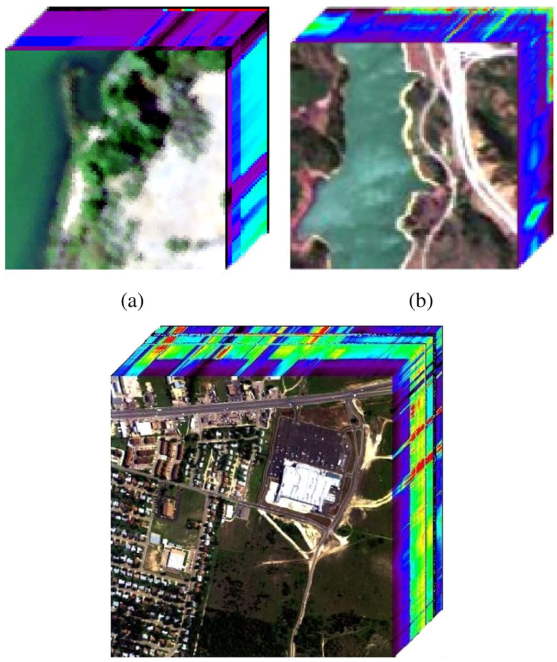

Fig. 7. False color composite of the real hyperspectral datasets. (a) Samson dataset. (b) Jasper Ridge dataset. (c) Urban dataset.

512 × 614 pixels in 224 bands. Since there is no ground truth for the whole surface of the HSI, a 100 × 100 subscene of this dataset is commonly used in different HU problems. Moreover, atmospherically affected bands and water vapor absorption ones, including bands 1–3, 108–112, 154–166, 220, and 224, have been removed from this dataset, remaining 198 bands. There are four endmembers in the Jasper Ridge dataset, including "tree," "dirt," "water," and "road." An illustration of the false color composite of this dataset is shown in Fig. 7(b).

SAD and RMSE values, obtained from the implementation of different experimental methods on the Jasper Ridge dataset, are reported in Tables V and VI. Estimated endmembers by different methods on the Jasper Ridge and their corresponding abundance maps are also illustrated in Figs. 10 and 11, respectively.

Results in Table V show that EWSP-NTF, in total, performed best in extracting endmembers' signatures. A closer examination of this table shows that the main reason for this superiority is the lowest SAD value for the "road" endmember. But, a remarkable point is that this SAD values are sometimes misleading. This is revealed by comparing the results of Table V with Figs. 10 and 11. For example, also EWSP-NTF has the best SAD value in extracting "road" endmember, figures show that this method has been unsuccessful in accurately extracting both the endmember and abundances of "road" when compared to other methods. Therefore, neglecting this endmember, MV-NTF and our proposed MTHUL$_q$ outperformed others in endmember extraction from Jasper Ridge dataset. In RMSE values, as shown in Table VI, MTHUL$_q$ overally had the best results. Fig. 11 also proves the higher quality of the MTHUL$_q$ abundance maps, compared to other competing methods.

3) Urban Dataset: One of the most commonly used datasets in the HU literature is the Urban dataset, characterized by its extensive spatial coverage and complex land cover distribution. Acquired by the HYDICE sensor, the dataset boasts a 2-m spatial resolution, consisting of 307 × 307 pixels. Comprising 210 spectral bands within the 400–2500 nm range and a spectral resolution of 10 nm, the dataset provides comprehensive spectral information. Bands 1–4, 76, 87, 101–111, 136–153, and 198–210, were usually removed due to water vapor absorption and atmospheric effects. Consequently, the applied dataset comprises 162 spectral bands. There are four endmembers in the scene of this dataset: "asphalt," "grass," "tree," and "roof." A false color composite of this dataset has been represented in Fig. 7(c).

Quantitative results obtained from applying experimental methods to urban data are presented in Tables VII and VIII. In addition, Figs. 12 and 13 depict the extracted endmembers and their corresponding abundances, respectively.

As evident in Table VII, MTHUL$_q$ has demonstrated the most favorable results in endmember estimation. In three out of four cases, the proposed MTHUL$_q$ exhibits the lowest SAD values. While SCNMTF achieved the best SAD value for the "roof" endmember, a detailed analysis of Fig. 12 reveals its failure to consistently extract this specific endmember. Consequently,

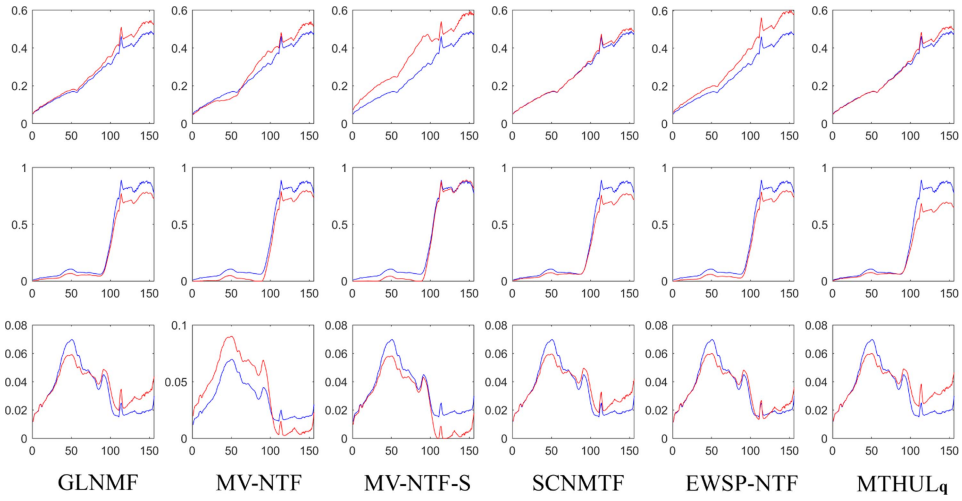

Fig. 8. Estimated endmembers (depicted in red) and true endmembers (depicted in blue) of the Samson dataset. The rows, arranged from top to bottom, represent rock, tree, and water, respectively.

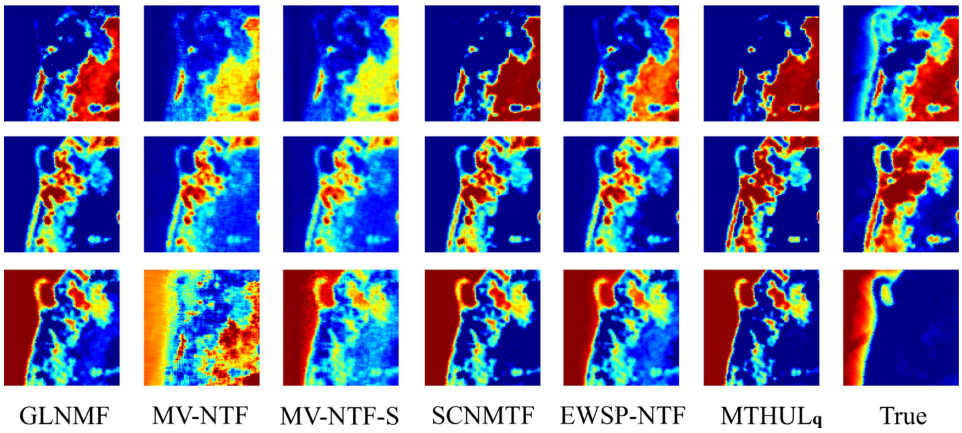

Fig. 9. Abundance maps generated by experimental methods, applied to the Samson dataset, alongside the corresponding true abundance maps. Rows from top to bottom belong, respectively, to rock, tree, and water.

TABLE V
SAD VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE JASPER RIDGE DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Tree	0.0983	0.1961	0.2033	**0.0502**	0.0517	0.0574
Water	0.0629	0.2515	0.3109	0.0721	0.0479	**0.0381**
Dirt	0.1141	0.1819	0.2720	0.1024	0.1118	**0.088**
Road	0.7670	0.2663	0.2891	0.7550	**0.1193**	0.7944
Average	0.2606	0.2240	0.2688	0.2449	**0.0827**	0.2445

The bold values show the smallest SAD values among different methods.

TABLE VI
RMSE VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE JASPER RIDGE DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Tree	0.0614	0.1465	0.1312	0.0517	0.0772	**0.0381**
Water	0.0420	0.1670	0.0896	0.0479	0.0410	**0.0407**
Dirt	0.1100	0.1846	**0.1091**	0.1118	0.1173	0.1104
Road	0.1285	0.2841	**0.0760**	0.1193	0.0955	0.1178
Average	0.0855	0.1955	0.1015	0.0827	0.0828	**0.0768**

The bold values show the smallest RMSE values among different methods.

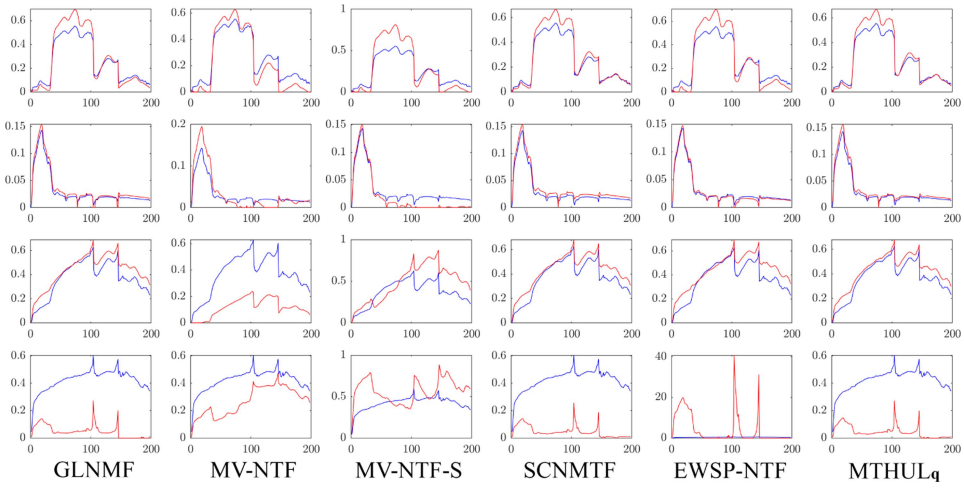

Fig. 10. Estimated endmembers (depicted in red) and true endmembers (depicted in blue) of the Jasper Ridge dataset. The rows, arranged from top to bottom, represent tree, water, dirt, and road, respectively.

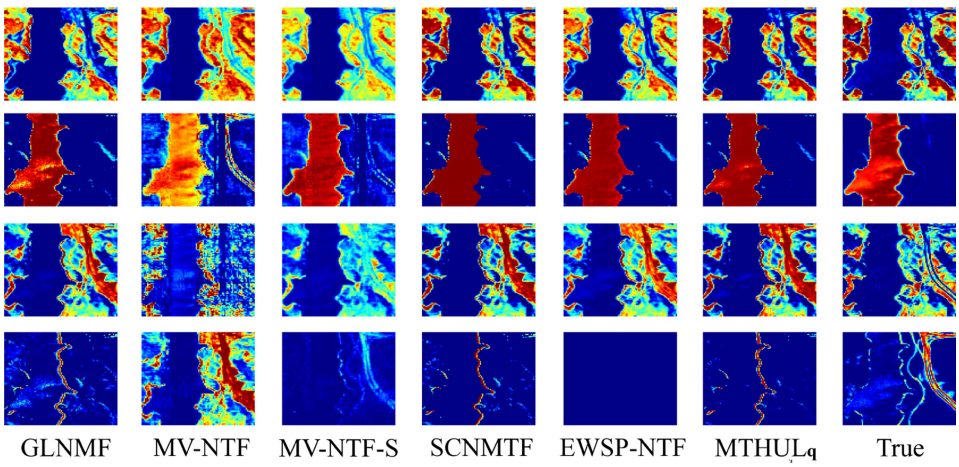

Fig. 11. Abundance maps generated by experimental methods, applied to the Jasper Ridge dataset, alongside the corresponding true abundance maps. Rows from top to bottom belong, respectively, to tree, water, dirt, and road.

TABLE VII
SAD VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE URBAN DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Asphalt	0.1285	0.2947	0.2060	1.0813	0.0792	**0.0725**
Grass	0.1895	0.1062	0.1437	0.0908	0.1977	**0.0477**
Tree	0.1590	0.1469	0.1522	0.1256	0.0892	**0.0608**
Roof	0.7707	0.6741	0.7150	**0.329**	0.7295	0.7893
Average	0.3119	0.3055	0.3042	0.4067	0.2739	**0.2426**

The bold values show the smallest SAD values among different methods.

TABLE VIII
RMSE VALUES OBTAINED BY IMPLEMENTATION OF DIFFERENT METHODS ON THE URBAN DATASET

Endmember	GLNMF	MV-NTF	MV-NTF-S	SCNMTF	EWSP-NTF	MTHUL$_q$
Asphalt	0.2105	**0.1717**	0.2076	0.2583	0.2134	0.1948
Grass	0.1503	0.2113	**0.1583**	0.2526	0.2039	0.2081
Tree	**0.0949**	0.6205	0.1908	0.1147	0.1412	0.1561
Roof	0.1087	0.0963	0.1117	0.0990	0.0836	**0.0741**
Average	**0.1411**	0.2750	0.1671	0.1812	0.1605	0.1583

The bold values show the smallest RMSE values among different methods.

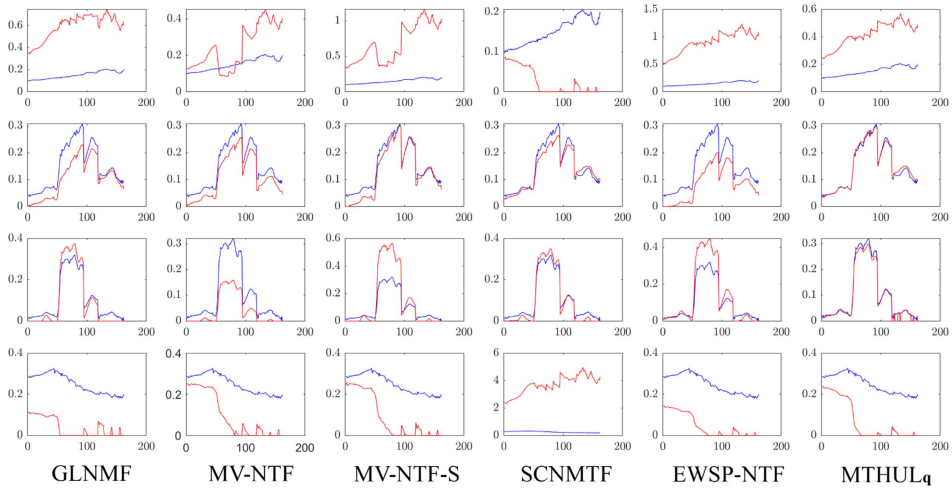

Fig. 12. Estimated endmembers (depicted in red) and true endmembers (depicted in blue) of the Urban dataset. The rows, arranged from top to bottom, represent asphalt, grass, tree, and roof, respectively.

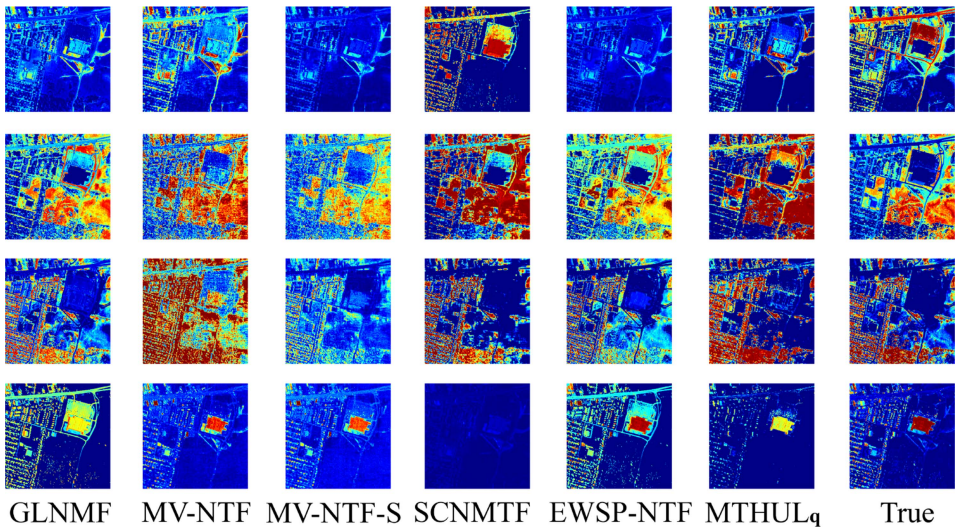

Fig. 13. Abundance maps generated by experimental methods, applied to the Urban dataset, alongside the corresponding true abundance maps. Rows from top to bottom belong, respectively, to asphalt, grass, tree, and roof.

a comprehensive examination of both quantitative and visual results substantiates that MTHUL$_q$ surpasses other methods in effectively extracting endmembers from the Urban dataset.

As deduced from Table VIII, our proposed MTHUL$_q$ exhibited the most favorable overall average RMSE after GLNMF. The superiority of GLNMF primarily stems from the RMSE value associated with the "tree" endmember. Nevertheless, a closer inspection of the abundance maps for this endmember in Fig. 13 reveals the highest quality in the MTHUL$_q$ abundance map. Overall, the analysis of this figure underscores the high quality of the abundance maps extracted by this method.

V. Discussion

Based on the experimental section, our proposed method demonstrates a high capability for extracting accurate results, particularly in abundance estimation. As evidenced by the results obtained from both synthetic and real data, MTHUL$_q$ achieved the lowest RMSE values in most of the experiments. This outcome aligns with our expectations, as we incorporated an additional regularization term on the abundance maps. The pivotal aspect of these findings lies in the interpretability of the results generated by our proposed term. Our combinatorial term not only facilitated the application of the ASC to a large extent but also enforced sparsity on the abundance estimates at the same time. Moreover, our exploration of a general norm allowed us to attain the optimal type of the sparsity on our datasets. Another noteworthy point is that we relaxed the stringent ASC norm, commonly applied in NTF-based algorithms, yet still achieved a high degree of consistency in ensuring that the sum of pixels' abundances equals one.

In extracting endmembers, although our proposed MTHUL$_q$ was not the best in some experiments, it had proper results in all experiments. The notable point is that, although other

methods have shown better results than our proposed method in some experiments, they have exhibited significant weaknesses in certain cases. For example, the powerful EWSP-NTF method, which outperformed our method in some cases based on the SAD values, practically failed in extracting the road endmember in the Jasper Ridge dataset (see Fig. 10), leading to a failure in extracting the corresponding abundance map (see Fig. 11). Another example is the result of the state-of-the-art SCNMTF in extracting the roof endmember from the Urban dataset. As evident from Fig. 12, SCNMTF failed to properly extract the spectral signature of the roof, leading to poor results in abundance estimation, as illustrated in Fig. 13. These extremely weak results are not evident in the outcomes of our MTHUL$_q$, and our approach has consistently yielded acceptable results, even if it has not always achieved the best outcomes.

VI. Conclusion

To enhance the results of the HU process, especially extracted abundance maps, a new general ℓ_q norm-based regularization term was proposed in this study, which was incorporated into the combinatorial matrix-tensor-based HU. Unlike conventional sparse regularization terms, our proposed term is consistent with ASC, and even essentially ensures it. Our method, named MTHUL$_q$, was established on the BTD paradigm. Quantitative and qualitative investigations using both real and synthetic hyperspectral datasets demonstrate the powerful performance of the proposed MTHUL$_q$, particularly in extracting accurate abundance maps. Like any other project, the present study also faces certain challenges that can serve as guidelines for future research. While our MTHUL$_q$ approach effectively improves the extracted abundance maps, there were cases where the corresponding endmembers did not achieve sufficient accuracy, especially in high-level noise scenarios. Therefore, we suggest future research projects focus on refining the endmember extraction aspect of HU as a key area of investigation. In addition, exploring the generalization of the model's parameters can also be a valuable direction for future studies.

References

[1] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002, doi: 10.1109/79.974727.

[2] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014, doi: 10.1109/jstars.2014.2320576.

[3] 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.

[4] J. S. Bhatt and M. V. Joshi, "Deep learning in hyperspectral unmixing: A review," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 2189–2192.

[5] V. S. S, V. S. Deshpande, and J. S. Bhatt, "A practical approach for hyperspectral unmixing using deep learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022.

[6] D. Heinz and Chein-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, doi: 10.1109/36.911111.

[7] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999, doi: 10.1038/44565.

[8] X.-R. Feng, H.-C. Li, R. Wang, Q. Du, X. Jia, and A. J. Plaza, "Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4414–4436, 2022.

[9] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011, doi: 10.1109/tgrs.2010.2098413.

[10] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.

[11] J. Liu, J. Zhang, Y. Gao, C. Zhang, and Z. Li, "Enhancing spectral unmixing by local neighborhood weights," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1545–1552, Oct. 2012, doi: 10.1109/jstars.2012.2199282.

[12] H. Zhang, H. Zhai, L. Zhang, and P. Li, "Spectralspatial sparse subspace clustering for hyperspectral remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3672–3684, Jun. 2016, doi: 10.1109/tgrs.2016.2524557.

[13] X. Wang, Y. Zhong, L. Zhang, and Y. Xu, "Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6287–6304, Nov. 2017, doi: 10.1109/tgrs.2017.2724944.

[14] S. Khoshsokhan, R. Rajabi, and H. Zayyani, "Clustered multitask nonnegative matrix factorization for spectral unmixing of hyperspectral data," *J. Appl. Remote Sens.*, vol. 13, no. 2, p. 1, May 2019, Art. no. 026509, doi: 10.1117/1.jrs.13.026509.

[15] 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, doi: 10.1109/tgrs.2012.2191590.

[16] R. Wang, H.-C. Li, A. Pizurica, J. Li, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using double reweighted sparse regression and total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1146–1150, Jul. 2017, doi: 10.1109/lgrs.2017.2700542.

[17] 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, doi: 10.1109/jstars.2019.2950700.

[18] Y. Yuan, Z. Zhang, and Q. Wang, "Improved collaborative non-negative matrix factorization and total variation for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 998–1010, 2020, doi: 10.1109/jstars.2020.2977399.

[19] J. Qin et al., "Blind hyperspectral unmixing based on graph total variation regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 4, pp. 3338–3351, Apr. 2021, doi: 10.1109/tgrs.2020.3020810.

[20] F. Li, S. Zhang, C. Deng, B. Liang, J. Cao, and S. Wang, "Robust double spatial regularization sparse hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 12569–12582, 2021, doi: 10.1109/jstars.2021.3132164.

[21] H. Song, X. Wu, A. Zou, Y. Liu, and Y. Zou, "Weighted total variation regularized blind unmixing for hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/lgrs.2021.3094826.

[22] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[23] Q. Zhang, H. Wang, R. Plemmons, and V. P. Pauca, "Spectral unmixing using nonnegative tensor factorization," in *Proc. 45th Annu. Southeast Regional Conf.*, 2007, pp. 531–532, doi: 10.1145/1233341.1233449.

[24] M. Wang et al., "Tensor decompositions for hyperspectral data processing in remote sensing: A comprehensive review," *IEEE Geosci. Remote Sens. Mag.*, vol. 11, no. 1, pp. 26–72, Mar. 2023.

[25] T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, "A low-rank tensor regularization strategy for hyperspectral unmixing," in *Proc. IEEE Stat. Signal Process. Workshop*, 2018, pp. 373–377.

[26] 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.

[27] M. Jouni, M. Dalla Mura, L. Drumetz, and P. Comon, "MultiHU-TD: Multi-feature hyperspectral unmixing based on tensor decomposition," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–21, 2023.

[28] L. Sun and H. Guo, "Blind unmixing of hyperspectral images based on l 1 norm and tucker tensor decomposition," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2021.
[29] P. Zheng, H. Su, H. Lu, and Q. Du, "Adaptive hypergraph regularized multilayer sparse tensor factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, pp. 1–18, 2023.
[30] 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.
[31] A. Cichocki et al., "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015, doi: 10.1109/msp.2013.2297439.
[32] 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.
[33] H.-C. Li, S. Liu, X.-R. Feng, and S.-Q. Zhang, "Sparsity-constrained coupled nonnegative matrixtensor factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5061–5073, 2020, doi: 10.1109/jstars.2020.3019706.
[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] H.-C. Li, S. Liu, X.-R. Feng, R. Wang, and Y.-J. Sun, "Double weighted sparse nonnegative tensor factorization for hyperspectral unmixing," *Int. J. Remote Sens.*, vol. 42, no. 8, pp. 3180–3191, 2021.
[36] P. Zheng, H. Su, and Q. Du, "Sparse and low-rank constrained tensor factorization for hyperspectral image unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1754–1767, 2021, doi: 10.1109/jstars.2020.3048820.
[37] Y. Yuan, L. Dong, and X. Li, "Hyperspectral unmixing using nonlocal similarity-regularized low-rank tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–14, 2022.
[38] P. Yang, T.-Z. Huang, J. Huang, and J.-J. Wang, "Efficient weighted-adaptive sparse constrained nonnegative tensor factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 10113–10130, 2022.
[39] X. Feng, L. Han, and L. Dong, "Weighted group sparsity-constrained tensor factorization for hyperspectral unmixing," *Remote Sens.*, vol. 14, no. 2, Jan. 2022, Art. no. 383, doi: 10.3390/rs14020383.
[40] 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, pp. 1–14, 2022.
[41] M. Ding, X. Fu, and X.-L. Zhao, "Fast and structured block-term tensor decomposition for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1691–1709, 2023.
[42] J. Sigurdsson, M. O. Ulfarsson, and J. R. Sveinsson, "Hyperspectral unmixing with ℓ_q regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6793–6806, Nov. 2014.
[43] J. M. Bioucas-Dias and J. M. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
[44] B. Rasti, M. O. Ulfarsson, and J. R. Sveinsson, "Hyperspectral subspace identification using sure," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2481–2485, Dec. 2015.
[45] S. Gholinejad, R. Shad, H. S. Yazdi, and M. Ghaemi, "Improving signal subspace identification using weighted graph structure of data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 6, pp. 831–835, Jun. 2017.
[46] X. Zhu, Y. Kang, and J. Liu, "Estimation of the number of endmembers via thresholding ridge ratio criterion," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 637–649, Jan. 2020.
[47] J. Nascimento and J. 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, doi: 10.1109/tgrs.2005.844293.
[48] 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, doi: 10.1109/tgrs.2012.2213825.
[49] R. Clark et al., "USGS digital spectral library splib06a: US geological survey, digital data series 231, 2007," 2007. [Online]. Available: http://speclab.cr.usgs.gov/spectral.lib06/ds231/index.html
[50] 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.
[51] F. Zhu, "Hyperspectral unmixing: Ground truth labeling, datasets, benchmark performances and survey," 2017, *arXiv:1708.05125*.

Saeid Gholinejad received the Ph.D. degree in remote sensing from the University of Isfahan, Isfahan, Iran, in 2020.

Since 2021, he has been a Postdoctoral Researcher in remote sensing with the University of Isfahan. His research interests include hyperspectral image processing, geometric processing of remotely sensed images, and pattern recognition.

Alireza Amiri-Simkooei received the Ph.D. degree in mathematical geodesy from the Delft University of Technology, Delft, The Netherlands, in 2007.

He is currently an Associate Professor with the Department of Control and Operations, Faculty of Aerospace Engineering, Delft University of Technology. He has authored or coauthored more than 95 research peer-reviewed journal articles in ISI journals. His areas of interest include aviation acoustics, machine learning, advanced estimation and approximation methods, precise positioning applications using space systems, and assessment of noise characteristics of geoscience data series. His research area also focused on the least-squares variance component estimation with applications to GPS data.

Dr. Amiri-Simkooei has been the Associate Editor of the *Journal of Surveying Engineering* (ASCE) since 2015.