# Constrained Nonnegative Matrix Factorization for Blind Hyperspectral Unmixing Incorporating Endmember Independence

E. M. M. B. Ekanayake, *Graduate Student Member, IEEE*, H. M. H. K. Weerasooriya, D. Y. L. Ranasinghe, S. Herath, B. Rathnayake, *Graduate Student Member, IEEE*, G. M. R. I. Godaliyadda, *Senior Member, IEEE*, M. P. B. Ekanayake, *Senior Member, IEEE*, and H. M. V. R. Herath, *Senior Member, IEEE*

***Abstract*—Hyperspectral unmixing (HU) has become an important technique in exploiting hyperspectral data since it decomposes a mixed pixel into a collection of endmembers weighted by fractional abundances. The endmembers of a hyperspectral image (HSI) are more likely to be generated by independent sources and be mixed in a macroscopic degree before arriving at the sensor element of the imaging spectrometer as mixed spectra. Over the past few decades, many attempts have focused on imposing auxiliary regularizes on the conventional nonnegative matrix factorization (NMF) framework in order to effectively unmix these mixed spectra. As a promising step toward finding an optimum regularizer to extract endmembers, this article presents a novel blind HU algorithm, referred to as kurtosis-based smooth nonnegative matrix factorization (KbSNMF) which incorporates a novel regularizer based on the statistical independence of the probability density functions of endmember spectra. Imposing this regularizer on the conventional NMF framework promotes the extraction of independent endmembers while further enhancing the parts-based representation of data. Experiments conducted on diverse synthetic HSI datasets (with numerous numbers of endmembers, spectral bands, pixels, and noise levels) and three standard real HSI datasets demonstrate the validity of the proposed KbSNMF algorithm compared to several state-of-the-art NMF-based HU baselines. The proposed algorithm exhibits superior performance especially in terms of extracting endmember spectra from hyperspectral data; therefore, it could uplift the performance of recent deep learning HU methods which utilize the endmember spectra as supervisory input data for abundance extraction.**

***Index Terms*—Blind source separation, constrained, endmember independence, Gaussianity, hyperspectral unmixing (HU), kurtosis, nonnegative matrix factorization (NMF).**

Manuscript received August 10, 2021; accepted October 26, 2021. Date of publication November 9, 2021; date of current version December 2, 2021. *(Corresponding author: H. M. H. K. Weerasooriya.)*

E. M. M. B. Ekanayake is with the Department of Electrical and Computer Systems Engineering, Monash University, Clayton, VIC 3800, Australia, and also with the Office of Research and Innovation Services, Sri Lanka, Technological Campus, CO 10500, Sri Lanka (e-mail: mevan.ekanayake@monash.edu).

H. M. H. K. Weerasooriya, D. Y. L. Ranasinghe, S. Herath, G. M. R. I. Godaliyadda, M. P. B. Ekanayake, and H. M. V. R. Herath are with the Department of Electrical and Electronic Engineering, University of Peradeniya, KY 20400, Sri Lanka (e-mail: kavingaweerasooriya@eng.pdn.ac.lk; e14273@eng.pdn.ac.lk; sanjaya.h@eng.pdn.ac.lk; roshangodd@ee.pdn.ac.lk; mpb.ekanayake@ee.pdn.ac.lk; vijitha@eng.pdn.ac.lk).

B. Rathnayake is with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: rathnb@rpi.edu).

Digital Object Identifier 10.1109/JSTARS.2021.3126664

## I. Introduction

HYPERSPECTRAL image (HSI) technology has become a leading imaging technology in many fields including medical imaging, food quality assessment, forensic sciences, surveillance, and remote sensing [1]. However, due to the insufficient spatial resolution of spectrometers and homogeneous mixture of distinct macroscopic materials in imaging scenes, the observed reflectance spectrum at each pixel of an HSI could easily be a mixture of spectra belonging to a set of constituent members (also called endmembers). This mixing phenomenon constitutes a major concern with regard to many applications. As a remedy to this complication, various methods of hyperspectral unmixing (HU) have been implemented to extract endmember spectra along with their fractional composition (also called abundances). HU is a study of three subproblems, i.e., determining the number of endmembers, extracting the endmember spectra, and realizing their abundances [2].

In the past, many algorithms have been introduced in order to solve the HU problem [3]–[15] and these algorithms can be categorized under three main schemes according to the basic computational approaches [16]: 1) statistical algorithms, 2) geometric algorithms, and 3) sparse regression based unmixing algorithms. Statistical algorithms interpret a mixed pixel by utilizing statistical representations. These representations are commonly analytical expressions based on the probability density functions (pdf) of the underlying mixed pixel spectra. Bayesian self-organizing maps [17], independent component analysis (ICA) [18], [19], independent factor analysis (IFA) [20], dependent component analysis [21], automated morphological endmember extraction [22], nonnegative matrix factorization (NMF) [23], and spatial–spectral endmember extraction algorithm [24] are some of the popular statistical algorithms utilized for HU. Geometric algorithms exploit the geometric orientation of HSI data in an $n$-dimensional space, where $n$ is the number of spectral bands captured by the imaging spectrometer. Vertex component analysis [25], minimum volume transform [26], simplex identification via split augmented Lagrangian [27], optical real-time adaptive spectral identification system [28], and iterative error analysis [29] are some of the geometric algorithms frequently utilized for HU. Sparse regression based approaches

utilize known libraries. The unmixing problem is formulated as a sparse linear regression problem which is based on the assumption that every feature can be linearly created by few elements extracted from known libraries [30]–[33].

In the recent past, several approaches have been introduced where deep learning (DL) is utilized for HU. In [34], a Hopfield neural network machine learning approach is utilized to solve the semi-NMF problem, which has illustrated promising performance with regard to abundance extraction when reliable endmember spectra are given as supervisory input data. In [35], an artificial neural network (ANN) is utilized to inverse the pixel spectral mixture in Landsat imagery. Here, to train the network, a spectral library had been created, consisting of endmember spectra collected from the image and simulated mixed spectra. In [36], a two-staged ANN architecture has been introduced in which the first stage reduces the dimension of the input vector utilizing endmember spectra as input data. Thus, some of the current DL-based methods for HU utilizes endmember spectra as supervisory input data in order to extract the abundances. In [37], a deep autoencoder network is utilized to extract endmembers and abundances simultaneously from nonlinearly mixed data. Unlike [34]–[36], no supervisory data is required for the architecture, thus acting as an unsupervised network. In [38], the autoencoder architecture is extended to a convolutional autoencoder network to benefit from the feature detection ability of convolutional layers. Here, weights of the output layer and the activations from the layer before the output layer provide abundances and endmembers simultaneously. In [39], long short-term memory network structure is included to capture the spectral correlation information and a spatial regulator is used to improve the spatial continuity of the results. In [40], a two-stream architecture termed TANet is introduced for HU. Here, the first stream tries to learn a mapping from pseudo-pixels to abundances, while the second stream tries to minimize the reconstruction errors.

Originally introduced by Lee and Seung [23], NMF is a mathematical tool which is utilized to decompose a nonnegative data matrix into the product of two other nonnegative matrices of lower rank based on the optimization of a particular objective function. Since the non-negativity criterion does not accommodate any negative elements in resultant matrices [41], the non-negativity objective of HU is satisfied automatically. Driven by this parts-based representation of the NMF framework, NMF-based algorithms are often utilized to solve the HU problem. However, NMF is an ill-posed geometric algorithm; therefore, it does not possess a unique solution [8]. The nonconvex objective functions utilized for NMF compel its solution space to be wide. Thus, many researchers have introduced novel NMF algorithms by adding different auxiliary regularizes to the conventional NMF framework in order to improve the uniqueness of its solution with respect to the HU setting. $l_{1/2}$-sparsity constrained NMF ($l_{1/2}$-NMF) [8], spatial group sparsity regularized NMF (SGSNMF) [42], minimum volume rank deficient NMF (Min-vol NMF) [43], manifold regularized sparse NMF [7], double constrained NMF [44], total variation regularized reweighted sparse NMF [45], subspace clustering constrained sparse NMF [46], nonsmooth NMF [47], robust collaborative NMF (R-CoNMF) [48], subspace structure regularized NMF (SSRNMF) [49], graph regularized NMF [50], and projection-based NMF [51] are some customary NMF-based baselines utilized for HU. Furthermore, a new architecture has recently emerged for blind unmixing under the premise nonnegative tensor factorization (NTF). The algorithm presented in the article Matrix-Vector Nonnegative Tensor Factorization for Blind Unmixing of Hyperspectral Imagery (MVNTF) [52] preserves the spectral and spatial information in that the factorization is performed on hyperspectral 3D cubes instead of unwrapped image datasets.

In HU, the endmembers are typically macroscopic objects in the HSI scene, such as soil, water, vegetation, etc. [2]. In a broader sense, HU attempts to find these macroscopic objects by utilizing the observations of signals that have already interacted (or mixed) with other objects in the scene before arriving at the sensing element of the imaging spectrometer. It is pragmatic to assume that the endmembers are consequences of different physical processes; hence, they are statistically independent[1] [18]. If a particular methodology promotes maximizing the independence of endmembers, each of the endmember spectra extracted utilizing that particular method will be more independent than the mixed pixel spectra. Therefore, such a method would be a progression toward the extraction of more realistic endmember spectra belonging to independent macroscopic objects. Even though the frequently associated abundance sum-to-one constraint [53] in HU does not accommodate the concept of independent endmembers, algorithms such as ICA [18], [19], IFA [20], and independent innovation analysis [54] are popular algorithms utilized in HU which consider this concept. Also, several attempts have been taken previously in order to incorporate the independence of endmembers onto the conventional NMF framework. The authors of [55] have proposed a novel initialization method based on statistical independence between NMF components. In [56], an attempt has been made to initialize NMF with a modified ICA method. In [57], a novel effective method has been introduced unifying independent vector analysis and NMF. Our previous work [58] discusses the suitability of utilizing the fundamental notions of kurtosis-based ICA to enhance the conventional NMF algorithm.

Inspired by the interpretable parts-based representations and simplicity of imposing auxiliary regularizes of the conventional NMF framework and motivated by our previous work [58]–[64], this study proposes a novel regularizer to the conventional NMF framework named average kurtosis regularizer. Incorporating this regularizer along with an abundance smoothing mechanism, we present a novel blind HU algorithm named kurtosis-based smooth nonnegative matrix factorization (KbSNMF) along with its two variants KbSNMF-fnorm and KbSNMF-div. The motivation of the proposed work is to promote the independence of endmembers while extracting them in accordance with the parts-based representations of the conventional NMF framework, thereby attempting to extract the most realistic endmember spectra from a given HSI. The contributions of this article are summarized as follows:

[1]Throughout the rest of this article, we refer to the "statistical independence" of endmembers as the "independence" of endmembers.

1) introduction of a novel regularizer for HU, based on kurtosis, which promotes the independence of endmembers of an HSI;
2) computation of the gradient of the aforesaid regularizer w.r.t. the factors of the conventional NMF framework, and the establishment of a blind HU algorithm named KbSNMF, which effectively promotes the independence of endmembers while maintaining the smoothness of abundance maps.

We also implement and evaluate the performance of the proposed algorithm in comparison with several selected state-of-the-art NMF-based HU baselines. Experiments are conducted on diverse synthetic HSI datasets (with numerous numbers of endmembers, spectral bands, pixels, and noise levels) as well as on three standard real HSI datasets. These experiments substantiate that the proposed algorithm outperforms other state-of-the-art NMF-based blind HU algorithms in many instances, especially in extracting endmember spectra. This observation is understandable since the proposed algorithm tries to improve upon the pragmatic characteristics of the endmember spectra, rather than trying to improve upon the pragmatic characteristics of the abundance maps. Thus, in an unsupervised setting where there is the luxury of utilizing a DL-based method for abundance extraction, the proposed algorithm would provide a viable counterpart to generate endmember spectra as supervisory input data to the DL-based method.

The rest of this article is arranged as follows. Section II provides the background related to the proposed algorithm. In Section III, the novel kurtosis-based regularizer is developed along with its derivatives. In Section IV, the novel KbSNMF algorithm is introduced. Section V discusses some key issues related to the implementation of the proposed algorithm. Section VI is devoted for experimental results and the article is concluded in Section VII.

## II. Background

### A. Linear Mixture Model

Linear mixture model (LMM) is the most frequently utilized model for HU, and its implications had been widely discussed in many previous works [4], [6], [8], [25], [65]. This model highly depends on the assumption that the incident light waves reflect only once from the underlying macroscopic objects and are captured by the sensing element of the imaging spectrometer without being subjected to scattering. In the LMM, the spectrum at each pixel is represented as a linear combination of the endmember spectra as follows:

$$\mathbf{x}_j = \sum_{i=1}^{r} S_{ij}\mathbf{a}_i + \mathbf{e}_j \tag{1}$$

where $\mathbf{x}_j \in \mathbb{R}_+^{n\times 1}$ is the $j$th pixel spectrum, $S_{ij}$ is the fractional composition occupied by the $i$th endmember in the $j$th pixel, $\mathbf{a}_i \in \mathbb{R}_+^{n\times 1}$ is the spectrum of the $i$th endmember of the HSI, $\mathbf{e}_j \in \mathbb{R}^{n\times 1}$ is an additive Gaussian noise associated with modeling errors, and $r$ is the number of endmembers in the HSI. All spectra are measured in reflectance values; hence, there is nonnegativity in $\mathbf{x}_j$'s and $\mathbf{a}_i$'s. The nonnegativity constraint $S_{ij} \geq 0$ and the sum-to-one constraint $\sum_{i=1}^{r} S_{ij} = 1$ are implied in order to guarantee that the fractional compositions representing the endmembers are nonnegative and the abundance summation equals 1 at each pixel. The LMM can be reformulated in matrix notations as follows:

$$\mathbf{X} = \mathbf{A} \times \mathbf{S} + \mathbf{E} \tag{2}$$

where $\mathbf{X} \in \mathbb{R}_+^{n\times m}$ is the HSI data matrix, $n$ being the number of spectral bands and $m$ being the number of pixels of the HSI, $\mathbf{A} \in \mathbb{R}_+^{n\times r}$ is the endmember matrix whose columns represent the spectra of each of the $r$ endmembers, $\mathbf{S} \in \mathbb{R}_+^{r\times m}$ is the abundance matrix whose columns represent the fractional compositions at each of the $m$ pixels, and $\mathbf{E} \in \mathbb{R}^{n\times m}$ is the noise matrix. This formulation casts the HU problem as a BSS problem, i.e., simultaneous extraction of the endmember spectra and their abundances at each pixel while utilizing the HSI as the input.

### B. Nonnegative Matrix Factorization

NMF is a low-rank approximation of nonnegative matrices widely utilized in the fields of computer vision, clustering, data compression, etc. [55], [66]–[69]. NMF was first introduced by Lee and Seung [23] as a parts-based representation technique which permits the data in a nonnegative matrix to be decomposed into two other nonnegative matrices. Given a matrix $\mathbf{V} \in \mathbb{R}_+^{n\times m}$, NMF tries to find nonnegative matrices $\mathbf{W} \in \mathbb{R}_+^{n\times r}$ (known as the source matrix) and $\mathbf{H} \in \mathbb{R}_+^{r\times m}$ (known as the mixing matrix) which satisfy the following approximation:

$$\mathbf{V} \approx \mathbf{WH}. \tag{3}$$

However, there are infinite number of $\mathbf{W}, \mathbf{H}$ solution pairs which satisfy the above approximation. For instance, it is possible to write $\mathbf{WH} = (\mathbf{W\Gamma}^{-1})(\mathbf{\Gamma H})$ for any invertible $\mathbf{\Gamma} \in \mathbb{R}_+^{r\times r}$. The conventional procedure to achieve (3) is by defining an objective function which quantifies the quality of the approximation between $\mathbf{V}$ and $\mathbf{WH}$ and implementing an optimization algorithm to minimize the defined objective function w.r.t. $\mathbf{W}$ and $\mathbf{H}$. One of the most commonly utilized objective function is the square of the Frobenius norm between $\mathbf{V}$ and $\mathbf{WH}$ as in the following equation:

$$\|\mathbf{V} - \mathbf{WH}\|_F^2 = \sum_{ij} \left(V_{ij} - (WH)_{ij}\right)^2. \tag{4}$$

The above expression is lower bounded by zero and distinctly vanishes if and only if $\mathbf{V} = \mathbf{WH}$. Another popular objective function is the divergence[2] of $\mathbf{V}$ from $\mathbf{WH}$ as in the following equation:

$$D(\mathbf{V}\|\mathbf{WH}) = \sum_{ij} \left( V_{ij} \log \frac{V_{ij}}{(WH)_{ij}} - V_{ij} + (WH)_{ij} \right). \tag{5}$$

[2]Unlike the Frobenius norm, the divergence cannot be designated as a "distance" since it is not symmetric in $\mathbf{V}$ and $\mathbf{WH}$. Thus, it is common practice to refer to it as the "divergence of $\mathbf{V}$ from $\mathbf{WH}$."

Similar to the Frobenius norm, the divergence is also lower bounded by zero and vanishes if and only if $\mathbf{V} = \mathbf{WH}$. Even though (4) and (5) functions are convex in $\mathbf{W}$ and $\mathbf{H}$ alone, they are not convex in $\mathbf{W}$ and $\mathbf{H}$ together [23]. Hence, it is not possible to analytically find global minima of these functions w.r.t. $\mathbf{W}$ and $\mathbf{H}$. However, it is possible to find local minima utilizing numerical optimization methods. Lee and Seung [23] have proposed the below (6) and (7) multiplicative update rules to find local minima of the above (4) and (5) functions, respectively.

$$\mathbf{W} \leftarrow \mathbf{W} \circ \frac{\mathbf{V}\mathbf{H}^T}{\mathbf{W}\mathbf{H}\mathbf{H}^T}, \qquad \mathbf{H} \leftarrow \mathbf{H} \circ \frac{\mathbf{W}^T\mathbf{V}}{\mathbf{W}^T\mathbf{W}\mathbf{H}} \tag{6}$$

$$\mathbf{W} \leftarrow \mathbf{W} \circ \frac{\frac{\mathbf{V}}{\mathbf{W}\mathbf{H}}\mathbf{H}^T}{\mathbf{1}_{n\times m}\mathbf{H}^T}, \qquad \mathbf{H} \leftarrow \mathbf{H} \circ \frac{\mathbf{W}^T\frac{\mathbf{V}}{\mathbf{W}\mathbf{H}}}{\mathbf{W}^T\mathbf{1}_{n\times m}}. \tag{7}$$

Lee and Seung have further proven the convergence of both the above update rules utilizing an auxiliary function analogous to the proof of convergence of the expectation–maximization algorithm [23].

The LMM model transforms the HU problem into the form of a conventional NMF problem. If $\mathbf{V}$ is the HSI data matrix $\mathbf{X}$, then source matrix $\mathbf{W}$ is the endmember matrix $\mathbf{A}$ and mixing matrix $\mathbf{H}$ is the abundance matrix $\mathbf{S}$. Thus, given $\mathbf{X}$, solving the blind HU problem for $\mathbf{A}$ and $\mathbf{S}$ utilizing the conventional NMF problem can be formulated as in (8) and (9) for Frobenius norm and divergence-based objective functions, respectively:

$$\arg\min_{\mathbf{A},\mathbf{S}} \|\mathbf{X} - \mathbf{AS}\|_F^2, \quad \text{s.t. } \mathbf{A}, \mathbf{S} \succeq 0 \tag{8}$$

$$\arg\min_{\mathbf{A},\mathbf{S}} D(\mathbf{X}\|\mathbf{AS}), \quad \text{s.t. } \mathbf{A}, \mathbf{S} \succeq 0. \tag{9}$$

In order to solve the above problems while improving the uniqueness, many previous works have incorporated additional auxiliary regularizes on $\mathbf{A}$ and $\mathbf{S}$ [8], [10], [42], [43], [48], [50].

## III. Average Kurtosis Regularizer

### A. Kurtosis of a Signal

Central moments are often utilized in signal processing in order to characterize the spread of the pdf of a signal [18]. A normalized version of the fourth central moment, given by (10), is called the Kurtosis of a signal. Here, $y$ denotes the signal, $\overline{y}$ denotes the mean of the signal, and $\mathbb{E}$ is the expectation operator. Intuitively, Kurtosis provides a measure of the "peaky"ness of the shape of the pdf of a signal. Excess kurtosis is a measure that compares the kurtosis of a given pdf with the kurtosis of a Gaussian distribution. Since the kurtosis of a Gaussian distribution equals 3, the excess kurtosis can be defined as in (11)

$$\text{kurtosis} = \frac{\mathbb{E}[(y-\overline{y})^4]}{(\mathbb{E}[(y-\overline{y})^2])^2} \tag{10}$$

$$\text{excess kurtosis} = \text{kurtosis} - 3. \tag{11}$$

Based on the value of excess kurtosis, distributions are categorized under three main types. *Mesokurtic* distribution is close to a Gaussian distribution; has an excess kurtosis closer to zero. *Leptokurtic* (also known as super-Gaussian) distribution has a higher and sharper central peak; tails are longer and flatter; has positive excess kurtosis. *Platykurtic* (also known as sub-Gaussian) distribution has a lower and broader central peak, tails are shorter and thinner, and has negative excess kurtosis.

Central limit theorem (CLT) ensures that a mixture of signals is approximately Gaussian irrespective of the distributions of the underlying source signals. Even though the converse of CLT is not assured, i.e., it is not certain that any Gaussian signal is a mixture of non-Gaussian signals, in practical scenarios, Gaussian signals do consist of a mixture of non-Gaussian signals [18]. Thus, to extract the underlying source signals from a signal mixture, it is common practice in BSS to define a measure of non-Gaussianity and implement an algorithm which maximizes the defined measure as Fig. 1 illustrates. Subsequently, excess kurtosis seems to be a suitable candidate for this purpose as it is a measure of non-Gaussianity. If the excess kurtosis value of a signal is close to zero, it tempts to be Gaussian and if the excess kurtosis value of a signal is away from zero, it tempts to be non-Gaussian (super- or sub-Gaussian). Since there are two types of non-Gaussian distributions, it is common practice in most BSS methods to assume that source signals are super-Gaussian [18]. Hence, in this work, we consider the constituent spectra of an HSI to have super-Gaussian distributions. Hence, from a given HSI data matrix $\mathbf{X}$, we aim to extract an endmember matrix $\mathbf{A}$, whose columnwise average kurtosis is maximized, utilizing an NMF framework. Thus, we introduce a novel constrained NMF algorithm which incorporates the maximization of the average kurtosis of endmembers.

### B. Average Kurtosis

Obeying the notations introduced in Section II-A, $\mathbf{A} \in \mathbb{R}_+^{n\times r}$ is the endmember matrix whose columns represent the spectra of each of the $r$ endmembers of the HSI. Thus, it is possible to extract the $i$th endmember utilizing a simple matrix manipulation as follows:

$$\mathbf{a}_i = \mathbf{A}\mathbf{\Phi}_i \tag{12}$$

where $\mathbf{a}_i$ is spectrum of the $i$th endmember from matrix $\mathbf{A}$ (or the $i$th column of matrix $\mathbf{A}$) and $\mathbf{\Phi}_i \in \mathbb{R}^{r\times 1}$ is a column vector whose all elements are zeros except for the $i$th element which equals 1. If the kurtosis of the $i$th endmember is $K_i$, it can be expressed as follows according to (10):

$$K_i = \frac{\mathbb{E}[(\mathbf{a}_i - \overline{\mathbf{a}_i})^4]}{(\mathbb{E}[(\mathbf{a}_i - \overline{\mathbf{a}_i})^2])^2} \tag{13}$$

where $\overline{\mathbf{a}_i}$ is the average of the corresponding $i$th endmember. Thus, the average kurtosis through all $r$ endmembers $\overline{K}$ can be expressed as follows utilizing (12) and (13):

$$\begin{aligned}\overline{K} &= \frac{1}{r}\sum_{q=1}^{r} K_q \\ &= \frac{1}{r}\sum_{q=1}^{r} \frac{\mathbb{E}[(\mathbf{A}\mathbf{\Phi}_q - \overline{\mathbf{A}\mathbf{\Phi}_q})^4]}{(\mathbb{E}[(\mathbf{A}\mathbf{\Phi}_q - \overline{\mathbf{A}\mathbf{\Phi}_q})^2])^2}.\end{aligned} \tag{14}$$

Thus, it is seen that $\overline{K}$ is a function of $\mathbf{A}$; therefore, it can be written as $\overline{K}(\mathbf{A})$. We try to maximize $\overline{K}$ so that the extracted

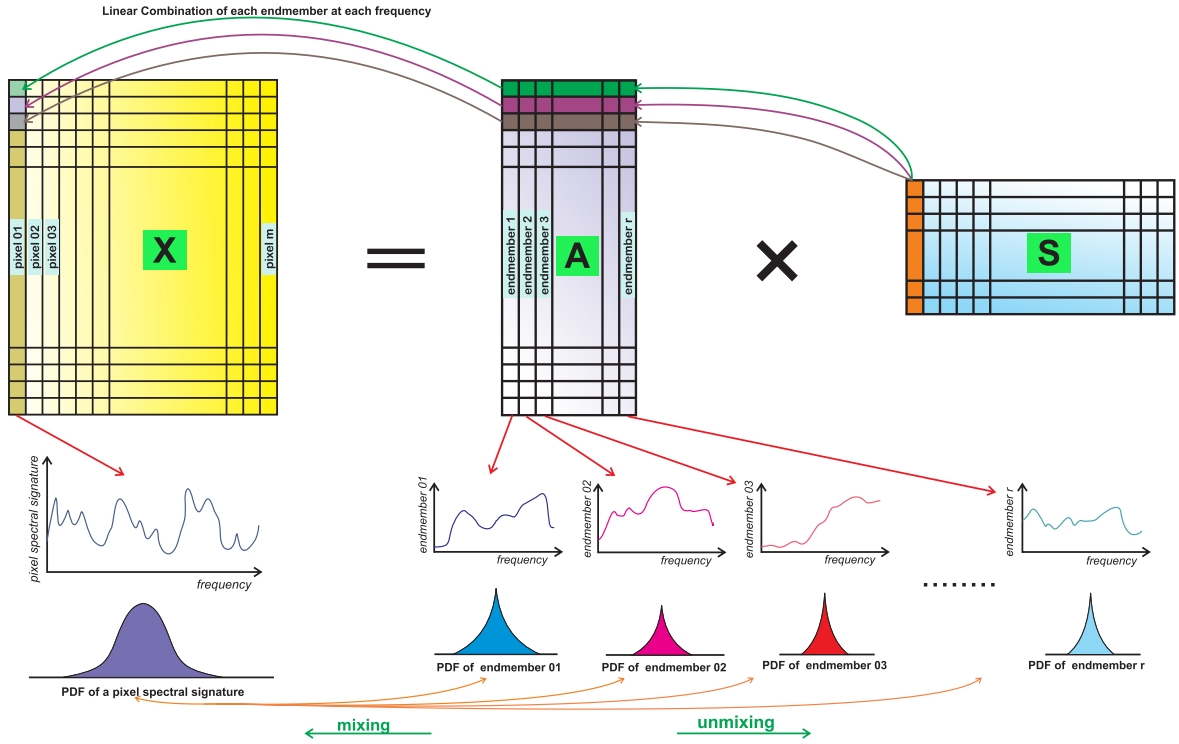



Fig. 1. Underlying mechanism of the proposed algorithm. According to the unmixing strategy, it is discernible that every pixel is a linear combination of several independent endmembers. The proposed method promotes independence of endmembers by increasing the super-Gaussianity. The algorithm concept is a derivative of the central limit theorem.

endmembers will have a higher average kurtosis, i.e., they will be closer to super-Gaussian signals. Hence, the proposed framework would favorably influence the extraction of more realistic endmember spectra from the underlying HSI.

### C. Derivative of Average Kurtosis

In order to incorporate the average kurtosis regularizer onto the conventional NMF framework, it is essential to find the gradient (or the partial derivative) of $\overline{K}$ w.r.t. $\mathbf{A}$ and $\mathbf{S}$, i.e., $\nabla_{\mathbf{A}}\overline{K} \in \mathbb{R}^{n\times r}$ and $\nabla_{\mathbf{S}}\overline{K} \in \mathbb{R}^{r\times m}$. Since $\overline{K}$ is not a function of $\mathbf{S}$, $\nabla_{\mathbf{S}}\overline{K} = \mathbf{0} \in \mathbb{R}^{n\times r}$. In this section, we provide a detailed explanation on finding $\nabla_{\mathbf{A}}\overline{K}$. Since $\mathbf{A}$ is the endmember matrix, we denote each of its elements by the notation $A_{ki}$, with the meaning of the reflectance value belonging to the $k$th spectral band of the $i$th endmember. Thus, the $(k,i)$th element of $\nabla_{\mathbf{A}}\overline{K}$ can be written as follows implementing an elementwise derivative:

$$\begin{aligned}\nabla_{\mathbf{A}}\overline{K}_{ki} &= \frac{\partial \overline{K}}{\partial A_{ki}} \\ &= \frac{1}{r}\sum_{q=1}^{r}\frac{\partial K_q}{\partial A_{ki}}\end{aligned} \tag{15}$$

where

$$\frac{\partial K_q}{\partial A_{ki}} = \begin{cases} \frac{\partial K_i}{\partial A_{ki}}, & \text{if } q = i \\ 0 & \text{otherwise.}\end{cases} \tag{16}$$

For the convenience of simplifying, we assume that each of the endmember spectra vectors have unit variance, i.e., $(\mathbb{E}[(\mathbf{a}_i - \overline{\mathbf{a}_i})^2])^2 = 1,\ \forall i$. In order to rectify the effects of this assumption, a normalization step is carried out as discussed in Section V-B. As a result, we obtain a simplified version of $\nabla_{\mathbf{A}}\overline{K}_{ki}$ as follows:

$$\begin{aligned}\nabla_{\mathbf{A}}\overline{K}_{ki} &= \frac{1}{r}\frac{\partial\left[\mathbb{E}[(\mathbf{a}_i - \overline{\mathbf{a}_i})^4]\right]}{\partial A_{ki}} \\ &= \frac{1}{nr}\sum_{p=1}^{n}\frac{\partial (A_{pi} - \mu_i)^4}{\partial A_{ki}}\end{aligned} \tag{17}$$

where $A_{pi}$ is the reflectance value belonging to the $p$th spectral band of the $i$th endmember, and $\mu_i$ is the mean reflectance of the $i$th endmember. As can be seen, $\nabla_{\mathbf{A}}\overline{K}_{ki}$ is a summation of $n$ more partial derivative terms for which the solutions can be obtained by utilizing the chain rule in calculus

$$\frac{\partial}{\partial A_{ki}}(A_{pi} - \mu_i)^4 = \begin{cases} -4(A_{ki} - \mu_i)^3\left(\frac{1}{n} - 1\right), & \text{if } p = k \\ -4(A_{pi} - \mu_i)^3\left(\frac{1}{n}\right) & \text{otherwise.}\end{cases} \tag{18}$$

Thus, the partial derivative term $\nabla_{\mathbf{A}}\overline{K}_{ki}$ in (17) can be written as follows:

$$\nabla_{\mathbf{A}}\overline{K}_{ki} = \frac{-4}{nr}\left[S_i - (A_{ki} - \mu_i)^3\right] \tag{19}$$

where $S_i = \frac{1}{n}\sum_{p=1}^{n}(A_{pi} - \mu_i)^3$ represents a normalized version of the third central moment (skewness) of the $i$th endmember. However, in this work, we do not explore the implication

of skewness within the derivative of the average kurtosis. Concatenating the elementwise derivatives, we then express $\nabla_{\mathbf{A}}\overline{K}$ as the difference between two matrices as follows:

$$\nabla_{\mathbf{A}}\overline{K} = \frac{-4}{nr}\left(\mathbf{P}-\mathbf{Q}\right) \tag{20}$$

where $P_{ki} = S_i$ and $Q_{ki} = (A_{ki}-\mu_i)^3$. Then, $\mathbf{Q}$ and $\mathbf{P}$ can be written as in (21) and (22), respectively, for the convenience of incorporating $\nabla_{\mathbf{A}}\overline{K}$ in the NMF framework

$$\mathbf{Q} = \left[\mathbf{A} - \frac{1}{n}\mathbf{1}_{n\times n}\mathbf{A}\right]^{\circ 3} = \left[\mathbf{N}\mathbf{A}\right]^{\circ 3} \tag{21}$$

$$\mathbf{P} = \frac{1}{n}\mathbf{1}_{n\times n}\mathbf{Q} = \frac{1}{n}\mathbf{1}_{n\times n}\left[\mathbf{N}\mathbf{A}\right]^{\circ 3} \tag{22}$$

where $\mathbf{N} = (\mathbf{I} - \frac{1}{n}\mathbf{1}_{n\times n})$ and $\mathbf{1}\in\mathbb{R}^{n\times n}$ denotes a matrix whose all elements are ones. $[.]^{\circ 3}$ denotes the Hadamard (element-by-element) power by 3. Finally, from (20), $\nabla_{\mathbf{A}}\overline{K}$ can be written as follows:

$$\begin{aligned}\nabla_{\mathbf{A}}\overline{K} &= \frac{-4}{nr}\left[\frac{1}{n}\mathbf{1}_{n\times n}\left[\mathbf{N}\mathbf{A}\right]^{\circ 3} - \left[\mathbf{N}\mathbf{A}\right]^{\circ 3}\right] \\ &= \frac{4}{nr}\left[\mathbf{N}\left[\mathbf{N}\mathbf{A}\right]^{\circ 3}\right].\end{aligned} \tag{23}$$

## IV. Kurtosis-Based Smooth Nonnegative Matrix Factorization

In this section, we propose a novel blind HU algorithm which not only promotes the independence of endmembers via the kurtosis regularizer but also promotes the smoothness of the abundance maps by integrating a smoothing matrix to the conventional NMF framework. Hence, we denominate the proposed algorithm as KbSNMF. In the proceeding sections, we discuss two variants of KbSNMF depending on the objective function utilized for approximation.

### A. KbSNMF-Fnorm

Here, we present KbSNMF based on Frobenius norm (KbSNMF-fnorm). The general optimization problem for KbSNMF-fnorm is as follows:

$$\arg\min_{\mathbf{A},\mathbf{S}} \left\{\|\mathbf{X}-\mathbf{A}\mathbf{M}\mathbf{S}\|_F^2 - \gamma\overline{K}(\mathbf{A})\right\}, \quad \text{s.t. } \mathbf{A},\mathbf{S}\succeq 0. \tag{24}$$

Here, $\gamma\in\mathbb{R}_+$ is a parameter which establishes the tradeoff between approximation error and non-Gaussianity of the endmembers rendered by $\overline{K}$, and $\mathbf{M}\in\mathbb{R}_+^{r\times r}$ is a symmetric matrix called the smoothing matrix which is defined as follows:

$$\mathbf{M} = (1-\theta)\mathbf{I} + \frac{\theta}{r}\mathbf{1}_{r\times 1}\mathbf{1}_{r\times 1}^T \tag{25}$$

where $\mathbf{I}$ is the identity matrix, $\mathbf{1}\in\mathbb{R}^{r\times 1}$ is a vector whose all elements are ones, and $\theta$ is a parameter which satisfies $0\le\theta\le 1$ and controls the extent of smoothness. Enforcing smoothness onto the abundance matrix can be interpreted as $\mathbf{Y} = \mathbf{M}\mathbf{S}$, where $\mathbf{Y}$ is the smoothness-enforced abundance matrix. When $\theta = 0$, $\mathbf{M} = \mathbf{I}$; hence, $\mathbf{Y} = \mathbf{S}$ and no smoothing has occurred in $\mathbf{S}$. As $\theta\to 1$, $\mathbf{Y}$ tends to become smoother and reaches the smoothest possible at $\theta = 1$. The matrix $\mathbf{M}$ is an adaptive smoothing matrix whose window length changes with the parameter $\theta$. The $\theta$ parameter will affect the minute details of the abundance maps such that as $\theta$ is increased, noise will be removed from the abundance map at the expense of fine information of the image. Fig. 2 demonstrates the effects of smoothing parameter $\theta$ on the abundance maps.

In order to find a solution for (24), we consider the objective function below

$$\mathbb{L}(\mathbf{A},\mathbf{S}) = \|\mathbf{X}-\mathbf{A}\mathbf{M}\mathbf{S}\|_F^2 - \gamma\overline{K}(\mathbf{A}). \tag{26}$$

In order to make the algorithm much simpler, the variable matrices $\mathbf{A}$ and $\mathbf{S}$ are updated in turns. In each iteration, first, $\mathbf{A}$ is updated while $\mathbf{S}$ is kept constant and then $\mathbf{S}$ is updated while $\mathbf{A}$ is kept constant. This scheme is called a block-coordinate descent approach and is widely utilized in NMF-based algorithms [70]. The updates rules can be primarily written as follows:

$$\begin{aligned}\mathbf{A} &\leftarrow \mathbf{A} - \eta_{\mathbf{A}}\circ\nabla_{\mathbf{A}}\mathbb{L} \\ \mathbf{S} &\leftarrow \mathbf{S} - \eta_{\mathbf{S}}\circ\nabla_{\mathbf{S}}\mathbb{L}\end{aligned} \tag{27}$$

where $\circ$ denotes the Hadamard (element-by-element) product. Updating $\mathbf{A}$ and $\mathbf{S}$ directly accounts to computing the partial derivatives $\nabla_{\mathbf{A}}\mathbb{L}\in\mathbb{R}_+^{n\times r}$ and $\nabla_{\mathbf{S}}\mathbb{L}\in\mathbb{R}_+^{r\times m}$ and finding suitable learning rates $\eta_{\mathbf{A}}\in\mathbb{R}_+^{n\times r}$ and $\eta_{\mathbf{S}}\in\mathbb{R}_+^{r\times m}$.

Computing the partial derivatives of $\mathbb{L}$ w.r.t. $\mathbf{A}$ and $\mathbf{S}$ can be seen as two parts, i.e., partial derivatives of $\|\mathbf{X}-\mathbf{A}\mathbf{M}\mathbf{S}\|_F^2$ term and $\gamma\overline{K}(\mathbf{A})$ term. We refer the readers to [47] and [70] for detailed explanation of the partial derivative of $\|\mathbf{X}-\mathbf{A}\mathbf{M}\mathbf{S}\|_F^2$. Incorporating the result in (23), we can present the partial derivatives of $\mathbb{L}$ as follows:

$$\begin{aligned}\frac{\partial\mathbb{L}}{\partial\mathbf{A}} &= -2\mathbf{X}\mathbf{S}^T\mathbf{M}^T + 2\mathbf{A}\mathbf{M}\mathbf{S}\mathbf{S}^T\mathbf{M}^T + 2\gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3} \\ \frac{\partial\mathbb{L}}{\partial\mathbf{S}} &= -2\mathbf{M}^T\mathbf{A}^T\mathbf{X} + 2\mathbf{M}^T\mathbf{A}^T\mathbf{A}\mathbf{M}\mathbf{S}\end{aligned} \tag{28}$$

where $\gamma'$ is the scalar quantity which equals $\frac{-2\gamma}{nr}$. By substituting $\frac{\partial\mathbb{L}}{\partial\mathbf{A}}$ and $\frac{\partial\mathbb{L}}{\partial\mathbf{S}}$ in the original block-coordinate descent equations in (27), we can obtain the following update rules for KbSNMF-fnorm:

$$\begin{aligned}\mathbf{A} &\leftarrow \mathbf{A} - \eta_{\mathbf{A}}\circ(-2\mathbf{X}\mathbf{S}^T\mathbf{M}^T + 2\mathbf{A}\mathbf{M}\mathbf{S}\mathbf{S}^T\mathbf{M}^T + 2\gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3}) \\ \mathbf{S} &\leftarrow \mathbf{S} - \eta_{\mathbf{S}}\circ(-2\mathbf{M}^T\mathbf{A}^T\mathbf{X} + 2\mathbf{M}^T\mathbf{A}^T\mathbf{A}\mathbf{M}\mathbf{S}).\end{aligned} \tag{29}$$

However, due to the subtracting terms in the gradients, the update rules (29) can enforce $\mathbf{A}$ and $\mathbf{S}$ to contain negative elements, which contradicts with the parts-based representation of the NMF framework as well as the HU setting. Thus, following a methodology similar to that proposed by Lee and Seung [23], we define data-adaptive learning rates $\eta_{\mathbf{A}}$ and $\eta_{\mathbf{S}}$ as follows in order to ensure all positive elements in $\mathbf{A}$ and $\mathbf{S}$ at each update step:

$$\begin{aligned}\eta_{\mathbf{A}} &= \frac{\mathbf{A}}{2\mathbf{A}\mathbf{M}\mathbf{S}\mathbf{S}^T\mathbf{M}^T + 2\gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3}}. \\ \eta_{\mathbf{S}} &= \frac{\mathbf{S}}{2\mathbf{M}^T\mathbf{A}^T\mathbf{A}\mathbf{M}\mathbf{S}}.\end{aligned} \tag{30}$$

The fraction line denotes element-by-element division. This results in the multiplicative update rules for the proposed

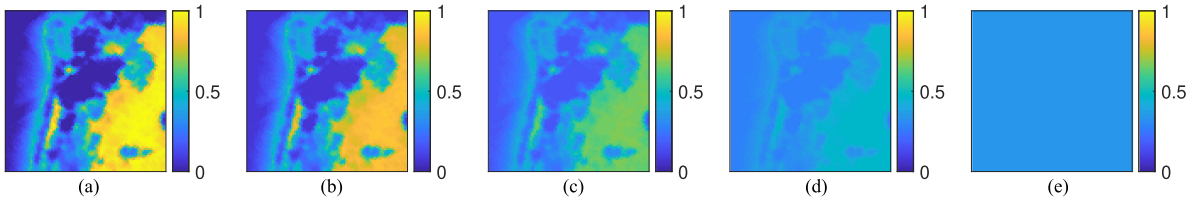



Fig. 2. Effects of the smoothing parameter demonstrated on a ground truth abundance map ("Soil") of a real HSI dataset ("Samson"). (a) $\theta = 0$ (no smoothing). (b) $\theta = 0.2$. (c) $\theta = 0.5$. (d) $\theta = 0.8$. (e) $\theta = 1$ (maximum smoothing).

KbSNMF-fnorm algorithm as follows:

$$\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\mathbf{X}\mathbf{S}^T\mathbf{M}^T}{\mathbf{A}\mathbf{M}\mathbf{S}\mathbf{S}^T\mathbf{M}^T + \gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3}}$$
$$\mathbf{S} \leftarrow \mathbf{S} \circ \frac{\mathbf{M}^T\mathbf{A}^T\mathbf{X}}{\mathbf{M}^T\mathbf{A}^T\mathbf{A}\mathbf{M}\mathbf{S}}. \tag{31}$$

For convenience, we reconfigure the placement of matrices. Therefore, the final update rules for the proposed KbSNMF-fnorm algorithm will be as follows:

$$\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\mathbf{X}(\mathbf{M}\mathbf{S})^T}{\mathbf{A}(\mathbf{M}\mathbf{S})(\mathbf{M}\mathbf{S})^T + \gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3}}$$
$$\mathbf{S} \leftarrow \mathbf{S} \circ \frac{(\mathbf{A}\mathbf{M})^T\mathbf{X}}{(\mathbf{A}\mathbf{M})^T(\mathbf{A}\mathbf{M})\mathbf{S}}. \tag{32}$$

It can be seen that choosing the data-adaptive learning rates in the form of (30) to avoid subtraction has enforced $\mathbf{A}$ and $\mathbf{S}$ to contain nonnegative elements throughout the block-coordinate descent approach, given initial nonnegative $\mathbf{A}$ and $\mathbf{S}$.

### B. KbSNMF-Div

Analogously, we present the following optimization problem for KbSNMF based on divergence (KbSNMF-div):

$$\arg\min_{\mathbf{A},\mathbf{S}} \left\{ D(\mathbf{X}\|\mathbf{A}\mathbf{M}\mathbf{S}) - \gamma\overline{K}(\mathbf{A}) \right\}, \quad \text{s.t.} \mathbf{A}, \mathbf{S} \succeq 0. \tag{33}$$

Following a similar procedure as in Section IV-A, the following multiplicative update rules can be derived for KbSNMF-div algorithm:

$$\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\frac{\mathbf{X}}{\mathbf{A}(\mathbf{M}\mathbf{S})}(\mathbf{M}\mathbf{S})^T}{\mathbf{1}_{n\times m}(\mathbf{M}\mathbf{S})^T + \gamma'\mathbf{N}[\mathbf{N}\mathbf{A}]^{\circ 3}}$$
$$\mathbf{S} \leftarrow \mathbf{S} \circ \frac{(\mathbf{A}\mathbf{M})^T\frac{\mathbf{X}}{(\mathbf{A}\mathbf{M})\mathbf{S}}}{(\mathbf{A}\mathbf{M})^T\mathbf{1}_{n\times m}} \tag{34}$$

where $\mathbf{1} \in \mathbb{R}^{n\times m}$ is a matrix whose all elements are one, and the other notations are the same as defined previously.

## V. Algorithm Implementation

In this section, we will discuss several points related to the implementation of the proposed algorithm.

### A. Initialization

Many algorithms had been designed in the past to enhance the initialization of the conventional NMF problem. In this work, we utilize the nonnegative double singular value decomposition (NNDSVD) algorithm [71] in order to initialize the matrices $\mathbf{A}$ and $\mathbf{S}$. NNDSVD takes the HSI $\mathbf{X}$ and the number of endmembers $r$ as the input and generates a pair of $\mathbf{A}$ and $\mathbf{S}$ matrices. The basic NNDSVD algorithm is based on two singular value decomposition (SVD) processes, first, approximating the data matrix and the, second, approximating positive sections of the resulting partial SVD factors incorporating the properties of unit rank matrices. Extensive evidence can be found to suggest that NNDSVD promotes the rapid convergence of the NMF algorithm.

### B. Normalization

To avoid the complexity of computing $\nabla_{\mathbf{A}}\overline{K}$, the endmember spectra are considered as signals of unit variance (see Section III-C), which is not always true in HU setting. In order to rectify this premise, at the beginning of each iteration of the proposed algorithm, we normalize the endmember spectra by their individual variances (see Algorithm 1: lines 5 and 13). Thus, the resulting algorithm follows the essence of projected gradient descent methods which are often utilized in signal processing applications [18].

### C. Convergence

Fig. 3 demonstrates the convergence of KbSNMF over number of iterations. Here, we have fixed the parameters $\gamma$ and $\theta$ at 3 and 0.4, respectively, for KbSNMF-fnorm and at 8 and 0.4, respectively, for KbSNMF-div. Selection of suitable $\gamma$ and $\theta$ and their effects on the unmixing performance are extensively discussed in Section VI-C1. Observing Fig. 3(a) and (b), it is evident that KbSNMF converges to a local minimum w.r.t. $\mathbf{A}$ and $\mathbf{S}$. Also, our primary objective of maximizing $\overline{K}$ has been achieved and can be clearly seen in Fig. 3(c) and (d). In the meantime, as seen in Fig. 3(e) and (f), Frobenius norm and divergence, respectively, converges to local minima w.r.t. $\mathbf{A}$ and $\mathbf{S}$ which ensures the quality of approximation between $\mathbf{X}$ and $\mathbf{A}\mathbf{M}\mathbf{S}$.

### D. Termination

In this work, we utilize two stopping criteria, one based on the maximum number of iterations and the other based on the rate of change in the objective function. We choose a maximum number of iterations, $t_{\max}$ and a minimum rate of change in the objective function $C_{\min}$. The algorithm is terminated either if the present iteration $t$ reaches $t_{\max}$ or if the present rate of

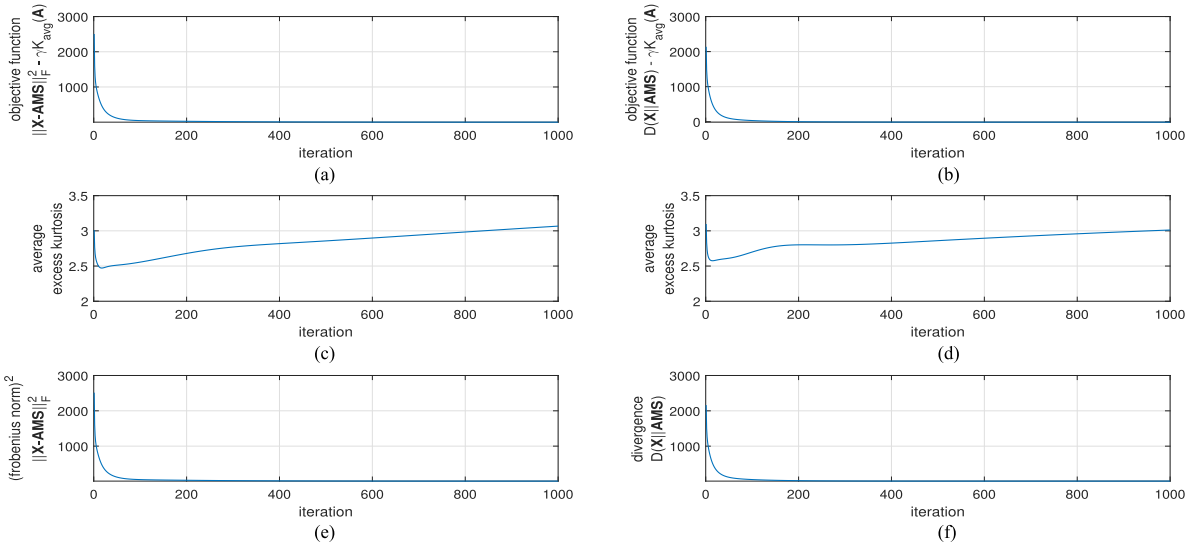



Fig. 3. Convergence of KbSNMF. Variation of (a) objective function, (c) average excess kurtosis of the extracted endmembers, and (e) square of Frobenius norm between X and AMS, over number iteration utilizing KbSNMF-fnorm algorithm. Variation of (b) objective function, (d) average excess kurtosis of the extracted endmembers, and (f) divergence of X from AMS, over number of iteration utilizing KbSNMF-div algorithm.

change in the objective function $C(t)$ falls below $C_{\min}$. Here, $C(t) = \frac{|L(t-1)-L(t)|}{|L(t-1)|}$, where $L(t)$ is the value of the objective function at the $t$th iteration. The selection of suitable $t_{\max}$ and $C_{\min}$ is discussed in Section V-E.

### E. Parameter Selection

Observing Fig. 3(a) and (b), it is evident that both variants of KbSNMF algorithm have converged to local minima by the 1000th iteration. Thus, we fix $t_{\max}$ at 1000 preserving a reasonable allowance. Also, it is seen that the percentage change in the objective function around the 1000th iteration is in the order of $10^{-4}$. Thus, we fix $C_{\min}$ at $10^{-5}$ to ensure convergence. Determining optimum control parameters $\gamma$ and $\theta$ is discussed in Section VI-C1 via experiment.

The matrix $\mathbf{M}$ removes minute details in the abundance map while preserving general spatial attributes and the $\theta$ balances noise removal and fine information retainment in the abundance maps. Hence, by removing noise in the abundance maps, the matrix $\mathbf{M}$ indirectly removes noise from the spectral signatures which has an impact on the endmember extraction process. Then the parameter $\gamma$ controls the effect from the kurtosis regularization for super-Gaussianity maximization. As $\gamma$ is increased, the thrust on super-Gaussianity is increased which affects the extracted endmembers. In the update rules, the term "$\mathbf{M}$" appears in both the endmember estimation and abundance estimation update rules for both the variants, while the effect of $\gamma$ is only applicable to endmember extraction update rules. Hence, it was apt to use spectral angle distance (SAD) of the endmembers for parametric sweep.

Adhering to all the implementing issues discussed above, the proposed KbSNMF algorithm can be summarized as in Algorithm 1.

### F. Computational Complexity

Computational complexity was calculated for the number of mathematical operations required for matrix operations. The computation of matrix N is only performed once, whereas rest of the operations are given for an iteration. Table I describes the computational complexity of the KbSNMF algorithms.

## VI. Experiments and Discussions

### A. Performance Criteria

In order to evaluate the performance of the proposed KbSNMF algorithm and assess its competitiveness with the other state-of-the-art algorithms, we utilize two performance criteria, which are commonly adopted in HU performance evaluation, i.e., SAD and root mean square error (RMSE). In most of the previous literature on HU, SAD had been utilized to compare the extracted endmember spectra with the ground truth endmember spectra, while RMSE had been utilized to compare the extracted abundance maps with the ground truth abundance maps. In our work, $\text{SAD}_i$, as in (35), measures the spectral angle between the $i$th ground truth endmember spectrum $\mathbf{a_i}$ and the corresponding extracted endmember spectrum $\widehat{\mathbf{a}}_\mathbf{i}$, in radians; $\text{RMSE}_i$ as in (36) measures the error between the $i$th ground truth abundance map $\mathbf{S_i}$ and the corresponding extracted abundance map $\widehat{\mathbf{S}}_\mathbf{i}$

$$\text{SAD}_i = \cos^{-1}\left(\frac{\widehat{\mathbf{a}}_i^T \mathbf{a}_i}{\|\widehat{\mathbf{a}}_i\|_2 \|\mathbf{a}_i\|_2}\right) \tag{35}$$

$$\text{RMSE}_i = \sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\mathbf{S}_{ij} - \widehat{\mathbf{S}}_{ij}\right)^2}. \tag{36}$$

Unless otherwise noted, in all experiments, SAD and RMSE are average values over all extracted endmember spectra and abundance maps, respectively.

**Algorithm 1:** KbSNMF Algorithm for HU

**Input:** $\mathbf{X}$, $r$, $\gamma$, $\theta$, $t_{max}$, and $C_{min}$, *Algorithm variant*

Initialize $\mathbf{A}$ and $\mathbf{S}$ utilizing NNDSVD algorithm [72];
Compute $\mathbf{M}$ utilizing (25);
Compute $\mathbf{N}$ as in (21);
Compute $\gamma'$ as in (28);
Normalize each column of $\mathbf{A}$ w.r.t its variance;
**while** $t \leq t_{max} \wedge C \geq C_{min}$ **do**
  Update $\mathbf{A}$:
  **switch** *Algorithm variant* **do**
    **case** *KbSNMF-fnorm* **do**
      $\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\mathbf{X}(\mathbf{MS})^T}{\mathbf{A}(\mathbf{MS})(\mathbf{MS})^T + \gamma' \mathbf{N}[\mathbf{NA}]^{\circ 3}}$;
    **end**
    **case** *KbSNMF-div* **do**
      $\mathbf{A} \leftarrow \mathbf{A} \circ \frac{\frac{\mathbf{X}}{\mathbf{A}(\mathbf{MS})}(\mathbf{MS})^T}{\mathbf{1}_{n \times m}(\mathbf{MS})^T + \gamma' \mathbf{N}[\mathbf{NA}]^{\circ 3}}$;
    **end**
  **end**
  Normalize each column of $\mathbf{A}$ w.r.t its variance;
  Update $\mathbf{S}$:
  **switch** *Algorithm variant* **do**
    **case** *KbSNMF-fnorm* **do**
      $\mathbf{S} \leftarrow \mathbf{S} \circ \frac{(\mathbf{AM})^T \mathbf{X}}{(\mathbf{AM})^T (\mathbf{AM}) \mathbf{S}}$
    **end**
    **case** *KbSNMF-div* **do**
      $\mathbf{S} \leftarrow \mathbf{S} \circ \frac{(\mathbf{AM})^T \frac{\mathbf{X}}{(\mathbf{AM})\mathbf{S}}}{(\mathbf{AM})^T \mathbf{1}_{n \times m}}$
    **end**
  **end**
  Normalize each column of $\mathbf{S}$ according to abundances sum-to-one constraint
**end**

**Output:** Extracted $\mathbf{A}$ and $\mathbf{S}$

### B. Experimental Setting

The proposed algorithm is tested on simulated as well as real hyperspectral datasets (see Fig. 4). Also, we compare the performance of our proposed algorithm with the popular state-of-the-art NMF-based HU baselines: $l_{1/2}$-NMF [8], SGSNMF [42], Min-vol NMF [43], R-CoNMF [48], SS-RNMF [49], and MVNTF [52]. To ensure that the evaluations are done on common grounds, we utilize the same initializing procedure and stopping criteria as mentioned in Sections V-A and V-D, respectively, for all the competing algorithms except MVNTF algorithm which is initialized with random values.

Simulated HSI data were generated utilizing the hyperspectral imagery synthesis toolbox (HSIST)[3] in order to conduct experiments. HSIST consists of the full USGS spectral library[4] which contains hundreds of endmember spectra including minerals, organic and volatile compounds, vegetation, and man-made materials. The corresponding abundance maps were generated incorporating a spherical Gaussian field [72].

To assess the performance of the proposed method in real environments, we conduct experiments on real hyperspectral data. The Samson dataset, the Urban dataset, and the Cuprite

[3][Online]. Available: http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB

[4][Online]. Available: https://www.usgs.gov/labs/spec-lab

TABLE I
COMPUTATIONAL COMPLEXITY OF THE PROPOSED ALGORITHMS

| Description | No. of Operations | |
|---|---|---|
| | KbSNMF-fnorm | KbSNMF-div |
| N matrix | additions : 1 | additions : 2 |
| numerator update - endmember | multiplications : $m + r$<br>additions : $m + r - 2$ | multiplications : $m + 3r + 1$<br>additions : $m + 3r - 4$ |
| denominator update - endmember | multiplications : $2m + 2r + 2n - 6$<br>additions : $2m + 2r + 2n - 6$ | multiplications : $m + r + 2n$<br>additions : $m + r + 2n - 4$ |
| mathematical operations | multiplications : 3<br>additions : 1 | multiplications : 3<br>additions : 2 |
| normalization | multiplications : $r + n$<br>additions : $2n$ | multiplications : $r + 2$<br>additions : $2n$ |
| numerator update - abundance | multiplications : $r + n$<br>additions : $r + n - 2$ | multiplications : $3r + n + 1$<br>additions : $3r + n - 4$ |
| denominator update - abundance | multiplications : $2r$<br>additions : $2r - 2$ | multiplications : $r + n$<br>additions : $r + n - 2$ |
| mathematical operations | multiplications : 2 | multiplications : 2 |
| normalization | multiplications : 1<br>additions : $r$ | multiplications : 1<br>additions : $r$ |

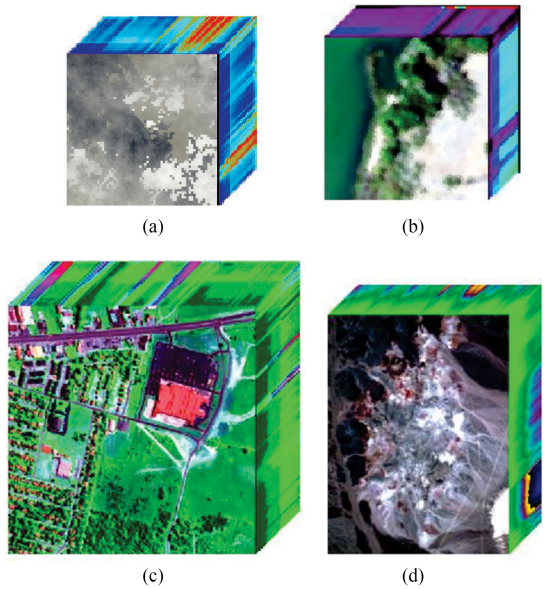


Fig. 4. RGB representations of the HSI datasets. (a) Simulated dataset. (b) Samson dataset. (c) Urban dataset. (d) Cuprite dataset.

TABLE II
UNMIXING PERFORMANCE COMPARISON IN TERMS OF SAD FOR THE SIMULATED DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Seawater | *0.3000* | **0.2746** | 0.6855 | 0.5416 | 0.8891 | 1.7408 | 0.3017 | 0.3905 |
| Clintonite | **0.1078** | *0.1119* | 0.2990 | 0.2288 | 0.2803 | 0.7109 | 0.1410 | 0.1326 |
| Sodiumbicarbonate | 0.1508 | 0.1311 | 0.0168 | **0.0090** | 0.0188 | 0.0614 | *0.0160* | 0.0567 |
| Average | 0.1862 | *0.1725* | 0.3337 | 0.2598 | 0.3961 | 0.8377 | **0.1529** | 0.1933 |

The best performances are in bold typeface; the second best performances are italicized; and the third best performances are underlined.

TABLE III
UNMIXING PERFORMANCE COMPARISON IN TERMS OF RMSE FOR THE SIMULATED DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Seawater | 0.3049 | 0.3190 | 0.2488 | 0.3579 | *0.1052* | 0.1753 | **0.0817** | 0.4088 |
| Clintonite | 0.1195 | 0.1200 | **0.0573** | 0.2906 | 0.0854 | 0.1473 | *0.0774* | 0.1506 |
| Sodiumbicarbonate | 0.2797 | 0.2985 | 0.2394 | 0.1123 | **0.0745** | 0.1475 | *0.0768* | 0.1324 |
| Average | 0.2347 | 0.2458 | 0.1818 | 0.2538 | *0.0884* | 0.1569 | **0.0786** | 0.2306 |

The best performances are in bold typeface; the second best performances are italicized; and the third best performances are underlined.

dataset have been widely utilized for performance evaluation and comparison in recent HU studies [3], [46], [73]. The Samson dataset's each pixel is recorded at 156 spectral channels covering wavelengths in the range of 401–889 nm with a spectral resolution of 3.13 nm. The Urban dataset's each pixel is recorded at 210 spectral channels originally; however, due to dense water vapor and atmospheric effects, several bands are customarily removed prior to analysis, resulting in 162 spectral bands ranging from 400 to 2500 nm, with a spectral resolution of 10 nm. The Urban dataset possesses several ground truth versions; here, we utilize the one with five endmembers. The Cuprite dataset is the widely used benchmark dataset for HU and each of its pixels is recorded at 188 spectral channels covering wavelengths in the range of 370–2480 nm.

The ground truths for all real datasets are worked out utilizing a procedure similar to that of [4] and [74]. First, the virtual dimensionality (VD) algorithm [75] is utilized to determine the number of endmembers of the HSI. Second, the pixels that contain pure endmember spectra are chosen manually in accordance with the USGS mineral spectral library. Finally, the corresponding abundances are computed utilizing the CVX optimization Toolbox in MATLAB. Accordingly generated ground truths are often utilized in HU method evaluation and comparison and are readily available.[5]

### C. Experiments on Simulated Data

*1) Sensitivity to Control Parameters:* We conduct experiments to find optimum values for $\gamma$ and $\theta$ for KbSNMF-fnorm and KbSNMF-div. We increase $\gamma$ from 0 to 25 in steps of 1, increase $\theta$ from 0 to 1 in steps of 0.1, and evaluate the unmixing performance at each step. It is seen that SAD reaches minimum around $\gamma = 3$ and $\theta = 0.4$ in Fig. 5(a) and around $\gamma = 8$ and $\theta = 0.4$ in Fig. 5(b). Thus, we fix $\gamma$ and $\theta$ at 3 and 0.4, respectively, for KbSNMF-fnorm and at 8 and 0.4, respectively, for KbSNMF-div.

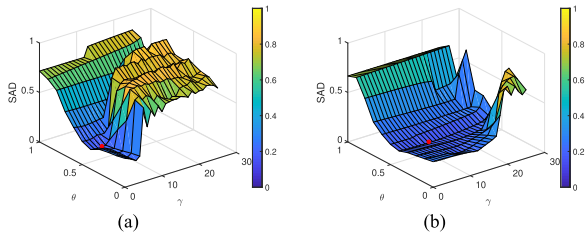



Fig. 5. Variation of unmixing performance in terms of SAD with $\gamma$ and $\theta$ for (a) KbSNMF-fnorm and (b) KbSNMF-div. The minimum SAD value in each 3-D surface is marked in red.

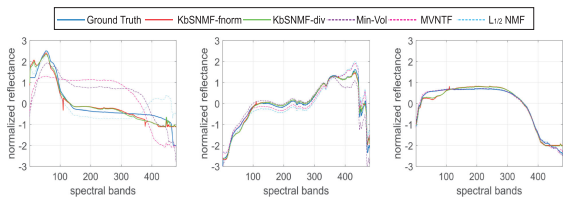



Fig. 6. Endmember spectra extracted utilizing KbSNMF algorithms and other comparison algorithms of the simulated dataset: "Seawater," "Clintonite," and "Sodiumbicarbonate," respectively.

*2) Unmixing Performance:* Under this experiment, we compare the unmixing performance of KbSNMF with the other HU algorithms. Table II shows SAD values for each of the extracted endmember spectra and Table III shows RMSE values for each of the extracted abundance maps, under the different methods. It is clearly seen that the KbSNMF under its both variants dominates the other competing algorithms in terms of SAD while signifying competitive performance in terms of RMSE. Figs. 6– 8, respectively, illustrate the endmember spectra, abundance maps, and mean reconstruction error matrix extracted utilizing KbSNMF algorithms along with their ground truths. However, for the optimization of the abundance estimation to converge within a fewer number of iterations, the initialization of the abundances

[5][Online]. Available: http://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes

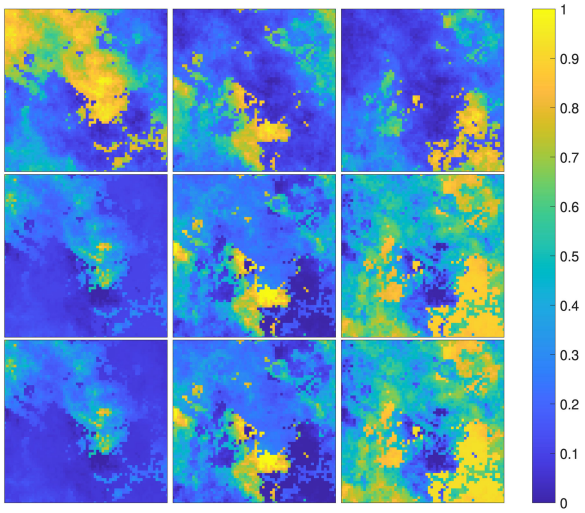


Fig. 7. Abundance maps extracted utilizing KbSNMF. Top row: ground truth abundance maps. Middle row: extracted abundance maps by KbSNMF-fnorm. Bottom row: extracted abundance maps by KbSNMF-div.

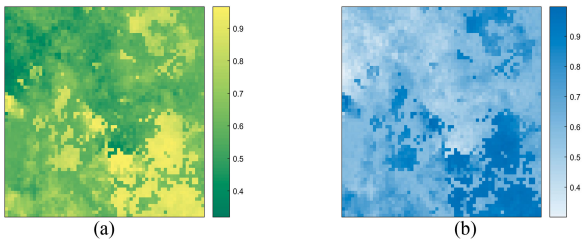


Fig. 8. Mean reconstruction error of the simulated dataset. (a) KbSNMF-fnorm. (b) KbSNMF-div.

must reside near, at least, a local minimum. Furthermore, since the gradient for the abundance estimation does not contain the $\gamma$ parameter, it can be concluded that the effect of the kurtosis regularization does not affect the abundance estimation.

*3) Robustness to Noise:* In this experiment, we aim to analyze how the proposed algorithm performs in noisy environments. We add zero-mean white Gaussian noise to the original noise-free simulated dataset with a predetermined signal-to-noise ratio (SNR) given by the following equation:

$$\text{SNR} = 10\log_{10}\frac{\mathbb{E}(\mathbf{x}^T\mathbf{x})}{\mathbb{E}(\mathbf{n}^T\mathbf{n})} \tag{37}$$

where $\mathbf{x}$ is the pixel spectrum vector, $\mathbf{n}$ is the noise signal vector, and $\mathbb{E}$ is the expectation operator. We conduct the experiment under 11 SNR levels: 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 dB, and the results are illustrated in Fig. 9 in terms of SAD and RMSE. Although SSRNMF and MVNTF show high immunity to large noise in terms of SAD values, it is discernible that KbSNMF-fnorm and KbSNMF-div report the best performance showing superior performance over all competing algorithms at noise levels in the range of 15–50 dB. They also show robustness to noise up until 30 dB. In terms of RMSE, both KbSNMF-fnorm and KbSNMF-div show robustness to noise up until 20 dB and gradually deteriorate in performance thereafter. However, both KbSNMF-fnorm and KbSNMF-div outperform SSRNMF at all

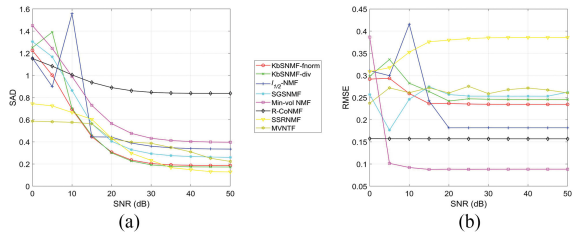



Fig. 9. Variation of (a) SAD and (b) RMSE with the noise level.

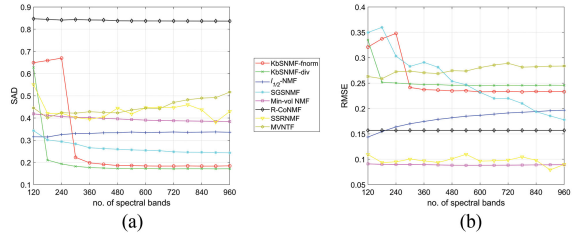



Fig. 10. Variation of (a) SAD and (b) RMSE with the number of spectral bands.

noise levels in terms of RMSE. The superior performance of KbSNMF-fnorm and KbSNMF-div in terms of SAD is due to the novel auxiliary regularizers on the endmember matrix, thereby attempting to extract the most realistic endmember spectra. The increase in RMSE values for low SNR values is due to the matrix M. Since the matrix M performs a compromise between noise removal and information retention, when the noise power is dominant, more portion of the high frequency components will be removed from the image. Hence, the abundance map will mostly contain general attributes of that hyperspectral dataset. Also, when the extracted abundance maps deviate from the actual abundance map, it will vitiate the endmember extraction process via the reconstruction error of the hyperspectral dataset.

*4) Sensitivity to Number of Spectral Bands:* Here we vary the number of spectral bands of the endmembers and observe the unmixing performance of the algorithms. The results are shown in Fig. 10. KbSNMF-fnorm and KbSNMF-div outperform all the competing algorithms in terms of SAD for the number of spectral bands in the range of 300–960. However, the performance of KbSNMF-fnorm and KbSNMF-div deteriorate drastically in terms of SAD for very low number of spectral bands, i.e., around 200 spectral bands. In terms of RMSE, KbSNMF-fnorm and KbSNMF-div outperform MVNTF at a high number of spectral bands, specifically more than 180, and outperform SGSNMF at a low number of spectral bands, i.e., below 480 spectral bands. When spectral bands are removed from the original signature, it damages the super-Gaussian nature of the endmember. But, as long as the super-Gaussian nature of the endmember is preserved, the algorithm will perform adequately. And once the endmember loses its super-Gaussian nature of the endmember, the algorithm's performance will worsen. To elaborate more, the kurtosis values of the sea-water endmember as the number of spectral bands is reduced are given in Fig. 11.

*5) Sensitivity to Number of Endmembers:* In this experiment, we vary the number of endmembers and investigate the

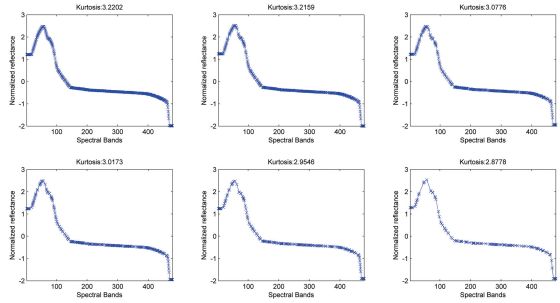



Fig. 11. Effect of number of spectral bands on signature kurtosis.

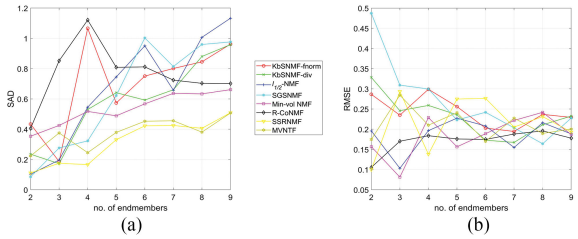



Fig. 12. Variation of (a) SAD and (b) RMSE with the number of endmembers.

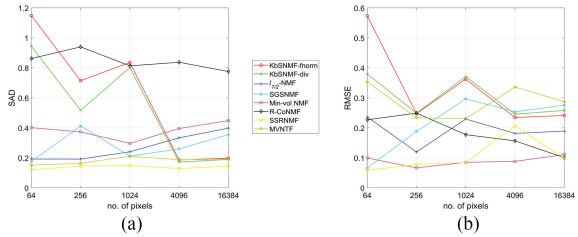



Fig. 13. Variation of (a) SAD and (b) RMSE with the number of pixels.

performance of the algorithms. The results are illustrated in Fig. 12. All the algorithms have the tendency to deteriorate the performance in terms of SAD with the number of endmembers. KbSNMF-fnorm and KbSNMF-div outperform R-CoNMF when the number of endmembers are low, i.e., below seven endmembers, and outperform SGSNMF when the number of endmembers is high, i.e., above five endmembers. In terms of RMSE, KbSNMF-fnorm and KbSNMF-div outperform SGSNMF when the number of endmembers is low, i.e., below four endmembers.

*6) Sensitivity to Number of Pixels:* Within this experiment, we illustrate how the proposed algorithm performs under simulated HSI datasets against different number of pixels. The number of pixels in an HSI is a major concern since it denotes the amount of statistical information in the input to the algorithm. The amount of statistical information presented to a numerical algorithm determines the tendency of an algorithm to be trapped in a local minima [76]. Fig. 13 illustrates the results in terms of SAD and RMSE. The unmixing performance of KbSNMF-fnorm and KbSNMF-div improves in terms of SAD when the number of pixels is increased and even outperforms all competing algorithms except MVNTF and SSRNMF when the number of pixels is very high, i.e., $64\times64$ and $128\times128$ pixels. In terms of RMSE, KbSNMF-fnorm and KbSNMF-div outperform SGSNMF when the number of pixels is very high, i.e., $64\times64$ and $128\times128$ pixels. Therefore, the more realizations we have of each endmember signature, the better error performances will record for endmember extraction and abundance estimation and vice versa as the algorithm uses the ensemble averaging process.

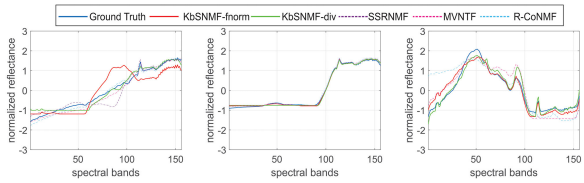



Fig. 14. Endmember spectra extracted utilizing KbSNMF algorithms and other comparison algorithms of the Samson dataset: "Soil," "Tree," and "Water," respectively.

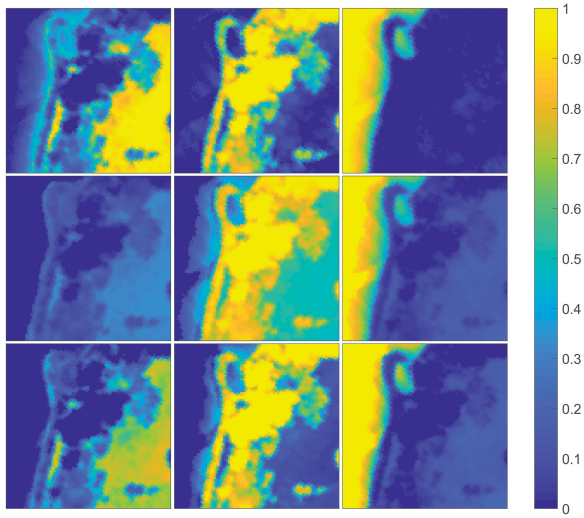


Fig. 15. Abundance maps extracted utilizing KbSNMF of the Samson dataset. Top row: ground truth abundance maps. Middle row: extracted abundance maps by KbSNMF-fnorm. Bottom row: extracted abundance maps by KbSNMF-div.

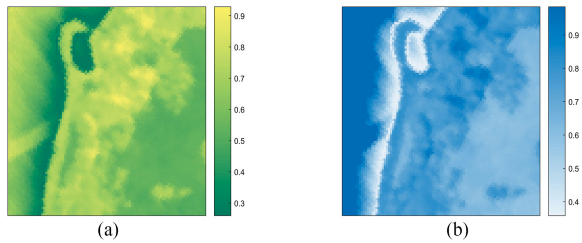



Fig. 16. Mean reconstruction error of the Samson dataset. (a) KbSNMF-fnorm. (b) KbSNMF-div.

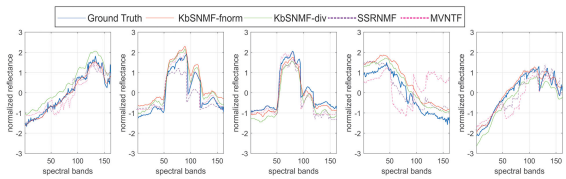



Fig. 17. Endmember spectra extracted utilizing KbSNMF algorithms and other comparison algorithms of the Urban dataset: "Asphalt," "Grass," "Tree," "Roof," and "Dirt," respectively.

TABLE IV
UNMIXING PERFORMANCE COMPARISON IN TERMS OF SAD FOR THE SAMSON DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Soil | 0.4975 | 0.2078 | 0.3455 | 0.3743 | 0.3463 | **0.1253** | 0.6091 | *0.1488* |
| Tree | *0.0456* | 0.0647 | 0.1433 | 0.1721 | 0.2315 | **0.0105** | 0.0750 | 0.0944 |
| Water | 0.2771 | 0.2014 | 0.3513 | 0.2941 | 0.2429 | 0.3219 | *0.1624* | **0.0887** |
| Average | 0.2734 | 0.1580 | 0.2800 | 0.2802 | 0.2736 | *0.1526* | 0.2822 | **0.1106** |

The best performances are in bold typeface; the second best performances are italicized; and the third best performances are underlined.

TABLE V
UNMIXING PERFORMANCE COMPARISON IN TERMS OF RMSE FOR THE SAMSON DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Soil | 0.3429 | 0.1574 | 0.4217 | *0.0532* | 0.0967 | **0.0431** | 0.3832 | 0.3517 |
| Tree | 0.2673 | 0.0911 | 0.0432 | 0.0882 | 0.1245 | **0.0118** | *0.0325* | 0.2454 |
| Water | *0.0910* | 0.0927 | 0.2359 | 0.1432 | **0.0432** | 0.5321 | 0.1654 | 0.4162 |
| Average | 0.2337 | 0.1137 | 0.2336 | *0.0949* | **0.0881** | 0.1957 | 0.1937 | 0.3378 |

The best performances are in bold typeface; the second best performances are italicized, and the third best performances are underlined.

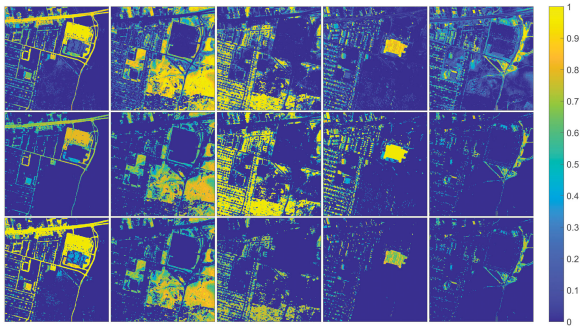



Fig. 18. Abundance maps extracted utilizing KbSNMF of the Urban dataset. Top row: ground truth abundance maps. Middle row: extracted abundance maps by KbSNMF-fnorm. Bottom row: extracted abundance maps by KbSNMF-div.

### D. Experiments on Real Data

We compare the unmixing performance of KbSNMF with the other competing methods in terms of SAD and RMSE for the Samson and Urban datasets. But for the Cuprite dataset, only the SAD values are tabulated.

*1) Samson Dataset:* Table IV shows SAD values for each of the extracted endmember spectra and Table V shows RMSE values for each of the extracted abundance maps, under the different methods. In terms of average SAD, MVNTF and R-CoNMF outperform all methods. However, KbSNMF-fnorm and KbSNMF-div outperform the rest of the other methods. Also, KbSNMF-div reports the third best performance in terms of SAD in extracting each endmember. In terms of RMSE, Min-vol NMF and SGSNMF outperform all methods. KbSNMF-div reports the third best average performance in terms of RMSE. The endmember spectra extracted utilizing KbSNMF-fnorm and KbSNMF-div as shown in Fig. 14. Also, the abundance maps extracted by KbSNMF-fnorm and KbSNMF-div are shown in Fig. 15, and it is evident that KbSNMF-div has managed to accurately extract the spatial variation of the abundance maps. As given in Fig. 14, the extracted signature for the "soil" endmember with the f-norm variant is disparate with the ground-truth signature while the difference between that and the signature from the div variant is comparatively less. Since the soil endmember is not super-Gaussian but sub-Gaussian, such a difference can be visible in the abundance map. Moreover, Fig. 16 illustrates the mean reconstruction error matrices for the KbSNMF algorithms.

*2) Urban Dataset:* Table VI shows SAD values for each of the extracted endmember spectra under the different methods. In terms of SAD, KbSNMF-fnorm outperforms all methods and KbSNMF-div outperforms the rest of the methods except for Min-vol NMF. Also, KbSNMF-fnorm reports the best performance and KbSNMF-div reports the second best performance in extracting the spectra of the endmembers "Tree" and "Roof." The endmember spectra extracted utilizing KbSNMF-fnorm and KbSNMF-div are shown in Fig. 17, and it can be observed that they closely follow their ground truth spectra. Table VII reports RMSE values for each of the extracted abundance maps, under the different methods. In terms of RMSE, Min-vol NMF outperforms all methods, followed by R-CoNMF and MVNTF. KbSNMF-fnorm reports the best performance in extracting the spectra of the endmembers "Grass" and "Roof," and KbSNMF-div reports the second best performance in extracting the spectra of the endmember "Asphalt." The abundance maps extracted utilizing KbSNMF-fnorm and KbSNMF-div are shown in Fig. 18, and it can be observed that they closely follow their ground truth abundance maps.

*3) Cuprite Dataset:* Table VIII shows SAD values for each of the extracted endmember spectra under the different methods. In terms of SAD, KbSNMF-div sits at the second place while SSRNMF stands at the top. The both KbSNMF forms report compelling results as they show best performance in extracting several endmembers, i.e., "Andradite," "Kaolinite 2," "Muscovite," and "Sphene." The endmember spectra extracted utilizing KbSNMF-fnorm and KbSNMF-div are shown in Fig. 19, and it can be observed that they closely follow their ground truth spectra.

TABLE VI
UNMIXING PERFORMANCE COMPARISON IN TERMS OF SAD FOR THE URBAN DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Asphalt | *0.1178* | 0.1252 | 0.2966 | 0.4173 | 0.1849 | 0.2017 | **0.0599** | 0.1721 |
| Grass | *0.1299* | 0.2821 | 0.4993 | 0.3434 | **0.1045** | 0.2786 | 0.1780 | 0.2080 |
| Tree | **0.1069** | *0.1498* | 0.1603 | 0.1499 | 0.1798 | 0.2125 | 0.1795 | 0.2310 |
| Roof | **0.1044** | *0.1621* | 0.2518 | 0.3822 | 0.1930 | 0.2478 | 0.4217 | 0.3941 |
| Dirt | 0.2999 | 0.1742 | 0.3379 | 0.3359 | **0.1521** | 0.2435 | *0.1534* | 0.3689 |
| Average | **0.1518** | 0.1787 | 0.3092 | 0.3257 | *0.1628* | 0.2368 | 0.1985 | 0.2748 |

The best performances are in bold typeface; the second best performances are italicized; and the third best performances are underlined.

TABLE VII
UNMIXING PERFORMANCE COMPARISON IN TERMS OF RMSE FOR THE URBAN DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Asphalt | 0.5425 | *0.1654* | 0.3197 | 0.5578 | **0.1376** | 0.1980 | 0.3237 | 0.3354 |
| Grass | **0.1296** | 0.2354 | 0.2140 | 0.2545 | *0.1490* | 0.3010 | 0.2437 | 0.3164 |
| Tree | 0.5470 | 0.4240 | 0.4607 | 0.2692 | **0.1064** | *0.1258* | 0.2660 | 0.3183 |
| Roof | **0.1528** | 0.4584 | 0.5350 | 0.2256 | 0.2130 | 0.1986 | 0.2851 | *0.1858* |
| Dirt | 0.3691 | 0.5789 | 0.4293 | 0.5610 | 0.5409 | 0.5210 | *0.3589* | **0.2004** |
| Average | 0.3482 | 0.3724 | 0.3917 | 0.3736 | **0.2294** | *0.2689* | 0.3121 | 0.2713 |

The best performances are in bold typeface; the second best performances are italicized, and the third best performances are underlined.

TABLE VIII
UNMIXING PERFORMANCE COMPARISON IN TERMS OF SAD FOR THE CUPRITE DATASET

| Methods | KbSNMF fnorm | KbSNMF div | $l_{1/2}$-NMF | SGSNMF | Min-vol NMF | R-CoNMF | SSRNMF | MVNTF |
|---|---|---|---|---|---|---|---|---|
| Alunite | 0.4960 | *0.3162* | 0.9145 | 0.7456 | 0.4747 | 0.5745 | **0.2521** | 0.3938 |
| Andradite | **0.0953** | 0.1977 | 0.5638 | 0.5926 | *0.1161* | 0.2054 | 0.1760 | 0.3118 |
| Buddingtonite | 0.5837 | 0.2731 | 0.8468 | **0.1182** | 0.1908 | 0.1936 | *0.1885* | 0.5634 |
| Dumortierite | 0.4210 | *0.1864* | 0.7002 | 0.7599 | **0.1316** | 0.2515 | 1.0213 | 0.3161 |
| Kaolinite 1 | 0.4682 | 0.2842 | *0.2366* | 0.5824 | 0.4536 | 0.5683 | **0.1375** | 0.4766 |
| Kaolinite 2 | 0.5530 | **0.1459** | 0.5869 | 0.7393 | 0.4702 | 0.4079 | *0.1625* | 0.4595 |
| Muscovite | 0.2352 | **0.2086** | 0.4558 | 0.5179 | 0.3358 | 0.3534 | *0.2240* | 0.2555 |
| Montmorillonite | 0.2926 | 0.4222 | **0.1127** | 0.5160 | 0.1651 | 0.2266 | *0.1311* | 0.3130 |
| Nontronite | 0.3374 | 0.5450 | 1.0758 | 0.2102 | *0.1845* | 0.4106 | **0.1249** | 0.2899 |
| Pyrope | *0.1654* | 0.2630 | 0.9745 | 0.4753 | 0.4885 | 0.4181 | **0.0595** | 0.3846 |
| Sphene | **0.1590** | 0.3174 | 1.0424 | 0.4875 | *0.2821* | 0.3770 | 0.6160 | 0.4048 |
| Chalcedony | 0.2723 | 0.4818 | *0.2291* | 0.6865 | 0.3511 | 0.5228 | **0.2394** | 0.3648 |
| Average | 0.3399 | *0.3035* | 0.6449 | 0.5360 | 0.3037 | 0.3758 | **0.2777** | 0.3778 |

The best performances are in bold typeface; the second best performances are italicized, and the third best performances are underlined.

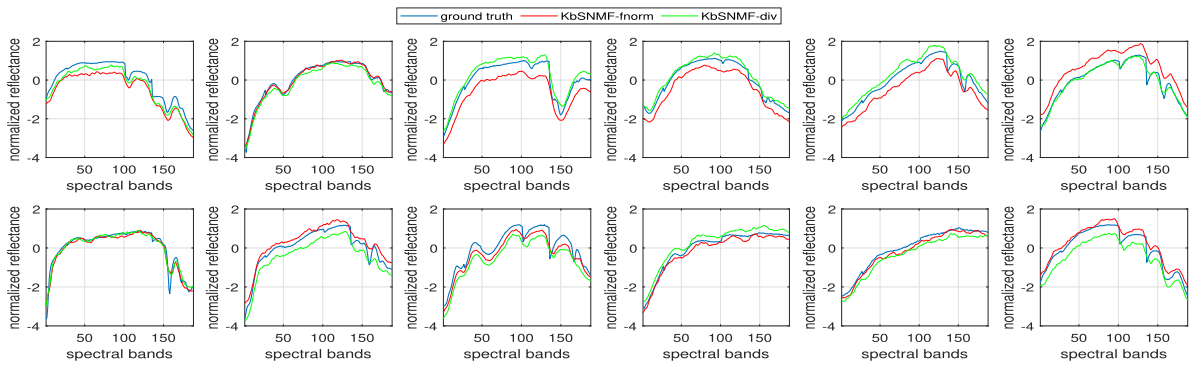



Fig. 19. Endmember spectra extracted utilizing KbSNMF algorithms of the Cuprite dataset: "Alunite," "Andradite," "Buddingtonite," "Dumortierite," "Kaolinite 1," "Kaolinite 2," "Muscovite," "Montmorillonite," "Nontronite," "Pyrope," "Sphene," and "Chalcedony," respectively.

## VII. Conclusion

This article proposed a blind HU algorithm called KbSNMF, which is based on incorporating the independence of endmembers to the conventional NMF framework. This was done by introducing a novel kurtosis regularizer based on the fourth central moment of a signal which signifies the statistical independence of the underlying signal. We illustrated a comprehensive derivation of the proposed algorithm in this article along with its performance evaluation in simulated as well as real environments (diverse simulated HSI datasets and three standard real HSI datasets). We have assessed the sensitivity of the proposed algorithm to control parameters, noise levels, number of spectral bands, number of pixels, and number of endmembers of the HSI. We have also provided performance comparisons of the proposed algorithm with the state-of-the-art NMF-based blind HU baselines. Moreover, experimental results verify that dominant performance in endmember extraction can be obtained through the novel algorithm. Hence, the proposed algorithm can be effectively utilized to extract accurate endmembers which can thereafter be passed through as supervisory input data to modern DL methods.

## References


[1] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.

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

[3] 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.

[4] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161–173, Jan. 2009.

[5] S. Yang, X. Zhang, Y. Yao, S. Cheng, and L. Jiao, "Geometric nonnegative matrix factorization (GNMF) for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2696–2703, Jun. 2015.

[6] F. Zhu, Y. Wang, B. Fan, S. Xiang, G. Meng, and C. Pan, "Spectral unmixing via data-guided sparsity," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5412–5427, Dec. 2014.

[7] 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.

[8] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.

[9] A. Halimi, Y. Altmann, N. Dobigeon, and J. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.

[10] C. Févotte and N. Dobigeon, "Nonlinear hyperspectral unmixing with robust nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4810–4819, Dec. 2015.

[11] H. Han *et al.*, "Hyperspectral unmixing via nonconvex sparse and low-rank constraint," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5704–5718, 2020, doi: 10.1109/JSTARS.2020.3021520.

[12] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 528–537, Feb. 2012.

[13] 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.

[14] Y. E. Salehani and S. Gazor, "Smooth and sparse regularization for NMF hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3677–3692, Aug. 2017.

[15] J. Yao, D. Meng, Q. Zhao, W. Cao, and Z. Xu, "Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 2991–3006, Jun. 2019.

[16] 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.

[17] L. Liu, B. Wang, L. Zhang, and J. Q. Zhang, "Decomposition of mixed pixels using Bayesian self-organizing map (BSOM) neural networks," in*Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 2014–2017.

[18] J. V. Stone, *Independent Component Analysis: A Tutorial Introduction*. Cambridge, MA, USA: MIT Press, 2004.

[19] J. Wang and C. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2601–2616, Sep. 2006.

[20] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[21] J. M. P. Nascimento and J. M. Bioucas-Dias, "Hyperspectral unmixing algorithm via dependent component analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2007, pp. 4033–4036.

[22] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.

[23] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst. 13 - Proc. Conf. Adv. Neural Inf. Process. Syst. Neural Inf. Process. Syst. Found.*, 2001, pp. 535–541 .

[24] D. Rogge, B. Rivard, J. Zhang, A. Sanchez, J. Harris, and J. Feng, "Integration of spatial-spectral information for the improved extraction of endmembers," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 287–303, 2007.

[25] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[26] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 3, pp. 542–552, May 1994.

[27] J. M. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in *Proc. 1st Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2009, pp. 1–4.

[28] J. H. Bowles, P. J. Palmadesso, J. A. Antoniades, M. M. Baumback, and L. J. Rickard, "Use of filter vectors in hyperspectral data analysis," in *Proc. Infrared Spaceborne Remote Sensing III*, vol. 2553, M. Strojnik and B. F. Andresen, Eds., SPIE International Society for Optics and Photonics, 1995, pp. 148–157.

[29] R. A. Neville, K. Staenz, T. Szeredi, J. Lefebvre, and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," in *Proc. 4th Int. Airborne Remote Sens. Conf. Exhib./21st Can. Symp. Remote Sens.*, 1999, pp. 21–24.

[30] 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.

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

[32] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3256–3274, Jun. 2014.

[33] W. Tang, Z. Shi, and Y. Wu, "Regularized simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5271–5288, Sep. 2014.

[34] J. Li, X. Li, B. Huang, and L. Zhao, "Hopfield neural network approach for supervised nonlinear spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 7, pp. 1002–1006, Jul. 2016.

[35] Z. Mitraka, F. Del Frate, and F. Carbone, "Spectral unmixing of urban landsat imagery by means of neural networks," in *Proc. Joint Urban Remote Sens. Event*, 2015, pp. 1–4.

[36] G. A. Licciardi and F. Del Frate, "Pixel unmixing in hyperspectral data by means of neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4163–4172, Nov. 2011.

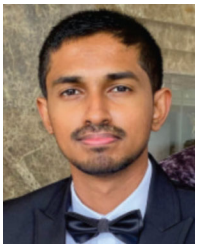
[37] M. Wang, M. Zhao, J. Chen, and S. Rahardja, "Nonlinear unmixing of hyperspectral data via deep autoencoder networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1467–1471, Sep. 2019.

[38] Y. Ranasinghe *et al.*, "Convolutional autoencoder for blind hyperspectral image unmixing," in *Proc. IEEE 15th Int. Conf. Ind. Inf. Syst.*, 2020, pp. 174–179.
[39] M. Zhao, L. Yan, and J. Chen, "LSTM-DNN based autoencoder network for nonlinear hyperspectral image unmixing," *IEEE J. Sel. Topics Signal Process.*, vol. 15, no. 2, pp. 295–309, Feb. 2021.
[40] Q. Jin, Y. Ma, X. Mei, and J. Ma, "Tanet: An unsupervised two-stream autoencoder network for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*,2021, doi: 10.1109/TGRS.2021.3094884.
[41] 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, 1999.
[42] 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.
[43] V. Leplat, A. M. S. Ang, and N. Gillis, "Minimum-volume rank-deficient nonnegative matrix factorizations," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 3402–3406.
[44] X. Lu, H. Wu, and Y. Yuan, "Double constrained NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2746–2758, May 2014.
[45] W. He, H. Zhang, and L. Zhang, "Total variation regularized reweighted sparse nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 3909–3921, Jul. 2017.
[46] X. Lu, L. Dong, and Y. Yuan, "Subspace clustering constrained sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3007–3019, May 2020.
[47] A. Pascual-Montano, J. M. Carazo, K. Kochi, D. Lehmann, and R. D. Pascual-Marqui, "Nonsmooth nonnegative matrix factorization (NSNMF)," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 403–415, Mar. 2006.
[48] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6076–6090, Oct. 2016.
[49] L. Zhou *et al.*, "Subspace structure regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4257–4270, Jul. 2020. doi: 10.1109/JSTARS.2020.3011257.
[50] D. Cai, X. He, J. Han, and T. S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011.
[51] Y. Yuan, Y. Feng, and X. Lu, "Projection-based NMF for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2632–2643, Jun. 2015. doi: 10.1109/JSTARS.2015.2427656.
[52] 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.
[53] D. C. 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.
[54] F. Geng, Z. Shi, Z. Jiang, and J. Yin, "Independent innovation analysis for hyperspectral imagery unmixing," in *Proc. 4th Int. Conf. Natural Comput.*, 2008, vol. 3, pp. 226–230.
[55] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 9, pp. 1622–1637, Sep. 2016.
[56] D. Benachir, S. Hosseini, Y. Deville, M. S. Karoui, and A. Hameurlain, "Modified independent component analysis for initializing non-negative matrix factorization: An approach to hyperspectral image unmixing," in *Proc. IEEE 11th Int. Workshop Electron., Control, Meas., Signals Appl. Mechatronics*, 2013, pp. 1–6.
[57] D. Kitamura, N. Ono, H. Sawada, H. Kameoka, and H. Saruwatari, "Determined blind source separation unifying independent vector analysis and nonnegative matrix factorization," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 24, no. 9, pp. 1626–1641, Sep. 2016.
[58] E. M. M. B. Ekanayake *et al.*, "Enhanced hyperspectral unmixing via non-negative matrix factorization incorporating the end member independence," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 2256–2259.
[59] B. Rathnayake, E. M. M. B. Ekanayake, K. Weerakoon, G. M. R. I. Godaliyadda, M. P. B. Ekanayake, and H. M. V. R. Herath, "Graph-based blind hyperspectral unmixing via nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 9, pp. 6391–6409, Sep. 2020.
[60] E. M. M. B. Ekanayake *et al.*, "A semi-supervised algorithm to map major vegetation zones using satellite hyperspectral data," in *Proc. 9th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, 2018, pp. 1–5.
[61] S. Vithana *et al.*, "Adaptive hierarchical clustering for hyperspectral image classification: Umbrella clustering," *J. Spectral Imag.*, vol. 8, no. 1, pp. a1, 2019. [Online]. Available: https://www.impopen.com/jsi-abstract/I08_a11
[62] E. Ekanayake *et al.*, "Mapping ilmenite deposit in Pulmudai, Sri Lanka using a hyperspectral imaging-based surface mineral mapping method," *J. Nat. Sci. Found. Sri Lanka*, vol. 47, no. 3, pp. 271–284, 2019.
[63] D. Y. L. Ranasinghe *et al.*, "Hyperspectral imaging based method to identify potential limestone deposits," in *Proc. 14th Conf. Ind. Inf. Syst.*, 2019, pp. 135–140.
[64] E. M. M. B. Ekanayake, W. G. C. Bandara, G. W. K. Prabhath, G. M. R. I. Godaliyadda, H. M. V. R. Herath, and M. P. B. Ekanayake, "Feature extraction using minor scatter directions of data to distinguish between classes with minute differences of a hyperspectral image," in *Proc. 14th Conf. Ind. Inf. Syst.*, 2019, pp. 130–134.
[65] 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.
[66] L. Weixiang, Z. Nanning, and Y. Qubo, "Nonnegative matrix factorization and its applications in pattern recognition," *Chin. Sci. Bull.*, vol. 51, no. 1, pp. 7–18, 2006.
[67] A. G. K. Janecek and W. N. Gansterer, *Utilizing Nonnegative Matrix Factorization for Email Classification Problems*. Hoboken, NJ, USA: Wiley, 2010, ch. 4, pp. 57–80.
[68] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Comput. Statist. Data Anal.*, vol. 52, no. 1, pp. 155–173, 2007.
[69] Z. Yang, G. Zhou, S. Xie, S. Ding, J. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Trans. Image Process.*, vol. 20, no. 4, pp. 1112–1125, Apr. 2011.
[70] J. J. Burred, "Detailed derivation of multiplicative updaterules for NMF," 2014. [Online]. Available: https://www.jjburred.com/research/pdf/jjburred_nmf_updates.pdf
[71] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognit.*, vol. 41, no. 4, pp. 1350–1362, 2008.
[72] C. Shi and L. Wang, "Linear spatial spectral mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3599–3611, Jun. 2016.
[73] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Convolutional autoencoder for spectral-spatial hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 535–549, Jan. 2021.
[74] S. Jia and Y. Qian, "Spectral and spatial complexity-based hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 12, pp. 3867–3879, Dec. 2007.
[75] C. -I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
[76] W. Wang, Y. Qian, and Y. Y. Tang, "Hypergraph-regularized sparse NMF for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 681–694, Feb. 2016.

**E. M. M. B. Ekanayake** (Graduate Student Member, IEEE) received the B.Sc. Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2019. He is currently working toward the Ph.D. degree in magnetic resonance image reconstruction using deep learning with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia.

He pursued his research on hyperspectral image processing as a Research Assistant at the Office of Research and Innovation Services (ORIS), Sri Lanka Technological Campus, Padukka, Sri Lanka, during 2019–2020. His previous works have been published in the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and several other IEEE-GRSS conferences including WHISPERS and IGARSS. His research interests include hyperspectral image processing, remote sensing, machine learning, and computer vision.

Mr. Ekanayake was the recipient of the Merit Award for Manamperi Award (Engineering) for his undergraduate research on hyperspectral image processing in 2019, which is presented to the best undergraduate Engineering research project submitted to a Sri Lankan University.



**H. M. H. K. Weerasooriya** received the B.Sc. degree in electrical and electronic engineering (with a first-class honors) from the University of Peradeniya, Peradeniya, Sri Lanka in 2020.

He is currently an Instructor with the Department of Electronic and Electrical Engineering, University of Peradeniya. He is presently involved in the researches on hyperspectral imaging for remote sensing and agriculture applications, and he has numerous publications in IEEE conferences. His research interests include image processing, signal processing, communication, machine learning and deep learning.

**D. Y. L. Ranasinghe** received the B.Sc.Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2020.

Immediately after receiving the B.Sc. Engineering degree, he joined the School of Engineering, Sri Lanka Technological Campus, Padukka, Sri Lanka, as a Research Assistant. His research interests include hyperspectral and multispectral image analysis and processing, blind source separation, and deep learning. He has numerous publications in IEEE conferences.

**S. Herath** received the B.Sc.Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2020.

Immediately after receiving the B.Sc.Eng. degree, he joined the Department of Engineering Mathematics, University of Peradeniya as a Teaching Instructor. His research interests include computer vision, image and signal processing, pattern recognition, blind source separation, and machine learning. He has numerous publications in IEEE conferences.

**B. Rathnayake** (Graduate Student Member, IEEE) received the B.Sc.Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2017.

Immediately after receiving the B.Sc. Engineering degree, he joined the Office of Research and Innovation Services (ORIS), Sri Lanka Technological Campus, Padukka, Sri Lanka, as a Research Assistant. He is currently working as a Research Assistant with Rensselaer Polytechnic Institute, Troy, NY, USA. His research interests include hyperspectral image analysis and processing, graph signal processing, and blind source separation. His previous works have been published in IEEE-TGRS and several other IEEE-GRSS conferences including IGARSS.

**G. M. R. I. Godaliyadda** (Senior Member, IEEE) received the B.Sc.Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2005, and the Ph.D. degree in electrical and computer engineering from the National University of Singapore, Singapore, in 2011.

He is currently a Senior Lecturer with the Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya. His current research interests include image and signal processing, pattern recognition, computer vision, machine learning, smart grid, bio-medical and remote sensing applications and algorithms. His previous works have been published in IEEE-TGRS and several other IEEE-GRSS conferences including WHISPERS and IGARSS. He also has numerous publications in many other IEEE transactions and Elsevier and IET journals.

Dr. Godaliyadda is a recipient of the Sri Lanka President's Award for Scientific Publications for 2018 and 2019. He is the recipient of multiple grants through the National Science Foundation (NSF) for research activities. He is the recipient of multiple best paper awards from international conferences for his work.

**M. P. B. Ekanayake** (Senior Member, IEEE) received the B.Sc.Eng. degree in electrical and electronic engineering from the University of Peradeniya, Peradeniya, Sri Lanka, in 2006, and the Ph.D. degree in mathematics from Texas Tech University, Lubbock, TX, USA, in 2011.

He is currently a Senior Lecturer with the University of Peradeniya. His current research interests include applications of signal processing and system modeling in remote sensing, hyperspectral imaging, and smart grid. His previous works have been published in IEEE-TGRS and several other IEEE-GRSS conferences including WHISPERS and IGARSS. He also has multiple publications in many IEEE transactions and Elsevier and IET journals.

Dr. Ekanayake is a recipient of the Sri Lanka President's Award for Scientific Publications in 2018 and 2019. He has obtained several grants through the National Science Foundation (NSF) for research projects. He has been awarded several best paper awards in international conferences.

**H. M. V. R. Herath** (Senior Member, IEEE) received the B.Sc.Eng. degree in electrical and electronic engineering with first class honours from the University of Peradeniya, Peradeniya, Sri Lanka, in 1998, the M.Sc. degree in electrical and computer engineering from the University of Miami, Miami, FL, USA, in 2002, and the Ph.D. degree in electrical engineering from the University of Paderborn, Paderborn, Germany, in 2009.

In 2009, he joined the Department of Electrical and Electronic Engineering, University of Peradeniya, as a Senior Lecturer. His current research interests include hyperspectral imaging for remote sensing, multispectral imaging for food quality assessment, coherent optical communications, and integrated electronics. He was a member of one of the teams that, for the first time, successfully demonstrated coherent optical transmission with QPSK and polarization multiplexing. His previous works have been published in IEEE-TGRS and several other IEEE-GRSS conferences including WHISPERS and IGARSS.

Dr. Herath is a member of the Institution of Engineers, Sri Lanka and The Optical Society. He was the General Chair of the IEEE International Conference on Industrial and Information Systems (ICIIS), 2013, held in Kandy, Sri Lanka. He is a recipient of the paper award in the ICTer 2017 conference held in Colombo, Sri Lanka, and of Sri Lanka President's Award for scientific research in 2013. He is the recipient of the award of academic merit from the University of Miami, USA.