

Bayesian Unmixing of Hyperspectral Image Sequence With Composite Priors for Abundance and Endmember Variability

Hongyi Liu^{ID}, *Member, IEEE*, Youkang Lu, *Student Member, IEEE*, Zebin Wu^{ID}, *Senior Member, IEEE*, Qian Du^{ID}, *Fellow, IEEE*, Jocelyn Chanussot^{ID}, *Fellow, IEEE*, and Zhihui Wei^{ID}

Abstract—A hyperspectral image sequence can be obtained at different time in the same region from a hyperspectral sensor. The environmental change usually leads to variation in endmember reflectance, which has an important influence on unmixing process. In this article, a Bayesian unmixing model considering spectral variability for hyperspectral sequence is proposed, in which composite prior distributions of abundance and endmember variability are developed. The abundance priors consider the continuity of abundance in the temporal and spatial domains, simultaneously. Specifically, in the spatial domain, a data-adaptive variance of the abundance prior distribution is put forward based on local spatial difference. Moreover, the priors of endmember variability in temporal continuity and spectral smoothness are also exploited. Finally, a joint posterior distribution is obtained by the likelihood function and the parameter prior distributions, which can be calculated by the Markov chain Monte Carlo (MCMC) algorithm. Experiments on synthetic and real data sets demonstrate the effectiveness of the proposed approach in terms of abundance, endmember, and its variability estimation accuracy.

Index Terms—Bayesian unmixing, endmember variability, hyperspectral (HS) image, Markov chain Monte Carlo (MCMC) method.

I. INTRODUCTION

HYPERSPECTRAL (HS) image has received an increasing interest due to its high spectral resolution [1]–[5]. Each pixel in an HS image is considered as a reflectance

spectrum, which can be regarded as the fingerprint of the corresponding material. However, due to the low spatial resolution of the HS image, pixels are often mixed. Unmixing, which decomposes a mixed pixel into a number of spectral signatures (pure endmembers) and their coefficients (abundance), has been widely investigated. In general, the unmixing approaches can be mainly divided into linear [6]–[9] or nonlinear models [10]–[12], which are incorporated by various constraints on endmember and abundance. The endmember is often regarded as l_2 limited energy [13]. For abundance, except the basic non-negativity prior [14], [15], other image priors, such as total variation (TV) [16], [17], sparsity [18], and low-rank [19], are also widely used. However, most of the unmixing methods assume that the endmember spectra are fixed, which neglect the fact that a material's reflectance may vary with illumination, atmosphere, season, and so on. Thus, unmixing considering endmember variability needs to be further studied [20]–[22]. Spectral variability involves several principles as mentioned in the literature [20]: iterative mixture analysis cycles, spectral feature selection, spectral weighting, spectral transformation, and spectral modeling. However, each of them concerns the cause of the variability from one aspect and omits the separation of the endmember variability from the endmember itself, which results in a suboptimal solution. Merging different conceptual approaches leads to reducing the variation influence effectively, and they can be divided into three main categories discussed later on.

One of the first attempts to address endmember variability is iterative mixture analysis by Roberts *et al.* [23], from which spectral bundles are developed. Based on the available spectral bundles, a standard signature of each pure endmember can be defined, and a finite number of spectral signatures can capture most of the spectral variability of each material [24], [25]. However, spectral bundles have the drawbacks of dependence on the diversity of extracted bundles. A simple solution is to randomly select spectra from a spectral data set, which suffers from diversity and lack of representation. Hence, dictionary learning method, as well as sparse representation, becomes an attractive way to give more accurate estimations of endmember and abundance [26], [27]. However, the extracted bundles sometimes may locate within the boundary of the simplex rather than in the vertices of the simplex. Hence, the original data are divided into three feature spaces along the

Manuscript received December 14, 2020; revised February 3, 2021 and February 25, 2021; accepted March 1, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61971223, Grant 61772274, Grant 61671243, Grant 61701238, Grant 11671205, and Grant 61871226; in part by the Jiangsu Provincial Natural Science Foundation of China under Grant BK20180018; and in part by the Fundamental Research Funds for the Central Universities under Grant 30917015104, Grant 30919011103, and Grant 30919011402. (*Corresponding author: Zebin Wu.*)

Hongyi Liu and Youkang Lu are with the School of Science, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: hylu@njust.edu.cn; luyoukang123@163.com).

Qian Du is with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS 39762 USA (e-mail: du@ece.msstate.edu).

Jocelyn Chanussot is with the Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France (e-mail: jocelyn@hi.is).

Zebin Wu and Zhihui Wei are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: wuzb@njust.edu.cn; gswei@njust.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TGRS.2021.3064708>.

Digital Object Identifier 10.1109/TGRS.2021.3064708

spectral dimension, and then the multiobjective particle swarm optimization algorithm is performed to identify the multiple endmember bundles with variability [28].

Another unmixing method is based on the linear mixing model (LMM) [29]. Considering the variation in endmember in each pixel, an extended linear mixing model (ELMM) is proposed in [30] and [31], in which the reflectance affected by illumination is simulated by multiplying a scaling ratio. However, the ELMM model assumes a fixed scaling ratio for different wavelengths, which lacks flexibility when endmembers are varying under more complex environments, for example, experimental measurements of vegetation spectrum show a significant difference with seasonal changing [32], [33]. To address the fact that endmember variation in each pixel may be caused by the illumination and intrinsic material, based on the prescribed ELMM model, an integrated unmixing chain is proposed [34]. The augmented linear mixing model (ALMM) [35] solves the spectral variability by applying data-driven learning strategy to the inverse problem of HS unmixing. Taking the spectral variation of each wavelength interval into consideration, generalized linear mixed model (GLMM) is developed [36]. The scaling factor of a spectral segmentation is used to make a new model adapt to the arbitrary changes in endmember spectrum. By assuming that spectral variations are caused by additive perturbations, the perturbed LMM (PLMM) [37] is established. In this model, endmember and its variability are considered separately, and different constraints or regularization can be imposed easily to model more complex environmental variation.

The third unmixing method relies on a statistical probability distribution, in which different distributions are added to constrain the abundance, endmember, and endmember variability. Then, a joint posterior distribution is inferred by optimization approaches, for example, expectation maximization, and Markov chain Monte Carlo (MCMC). A popular one is normal compositional model (NCM) [38], [39], in which endmembers are modeled as random vectors obeying a Gaussian distribution with different variances. A Beta compositional model (BCM) [40] is used to express distribution and spectral variability. In [41], it is argued that although each endmember obeys Gaussian distribution, its histogram peak may be different, so it cannot be modeled by a Gaussian or Beta distribution. Then, a Gaussian mixture model (GMM) with different mean and variance are used, resulting in good performance on spectral variability.

With the development of HS imaging, multitemporal HS images, which acquire a sequence of hyperspectral images (SHS) at different time in the same scene, have received attention in HS image processing [42], [43]. The unmixing of SHS can be traced to [44], in which a two-step unmixing method based on pruning the dictionary is proposed to improve the unmixing accuracy and save execution time. Another dynamical spectral unmixing is developed in [45], in which the endmember and abundance of multiframe HS are regarded as latent variables and emerged into the LMM unmixing model. Furthermore, the spectral variability is captured by a scale change and an additive noise terms at the same time. Based on the well-known PLMM, the unmixing problem

is formulated as a two-stage stochastic program that allows the model parameters to be estimated online [46], providing a relevant variability estimation on each image of the sequence. To ease the influence of parameters in the PLMM unmixing model, statistical methods based on PLMM are produced to represent temporal variability in case of SHS unmixing [47]. The prior of spectral variability along with temporal dimension are assumed to be smooth, and a Gaussian distribution is used to represent the variability prior, while the abrupt spectral variation is modeled as outliers. Moreover, in the Bayesian framework, three factors, namely, spectral, spatial, and temporal, are taken fully into consideration. For instance, abundance in temporal dimension is continuous, and endmember variability in multiframe images is also smooth [47]. Unfortunately, spatial and spectral dimension correlations of abundance and endmember variability are not fully exploited. In this article, by integrating the variable priors from three dimensions, the composite prior distributions for abundance and endmember variability are formulated and a joint probability distribution is established. The main contributions are summarized as follows.

First, the abundance prior is concerned not only on temporal dimension but also on spatial dimension. Moreover, the continuity of abundance in the spatial domain is well-modeled by a normal distribution with a data-adaptive variance. Second, the prior constraints of endmember variability take into account the temporal continuity and spectral smoothness. Specifically, the endmember perturbation caused by the same imaging situation is assumed to be consistent along different bands, which results in a new prior of endmember variability in spectral dimension. Finally, a joint posterior probability is derived and a Gibbs sampler is used to implement the Bayesian inference.

This article is organized as follows. The general Bayesian unmixing model [47] is introduced in Section II. In Section III, the new proposed priors of abundance and endmember variability are developed. Section IV investigates a Gibbs sampler to solve the resulting unmixing model. Section V shows the performance of the proposed approach on synthetic and real data compared with other state-of-the-art unmixing approaches. Finally, Section VI concludes the work.

II. RELATED WORKS

A. Observed Model

An observed SHS $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_T]$, which is acquired at T different time instants, shares most of the same endmembers. According to the PLMM unmixing model, \mathbf{Y}_t can be modeled as [46], [47]

$$\mathbf{Y}_t = (\mathbf{M}_t + \mathbf{dM}_t)\mathbf{A}_t + \mathbf{B}_t \quad (1)$$

where $t = 1, 2, \dots, T$. Each $\mathbf{Y}_t = [\mathbf{y}_{1,t}, \dots, \mathbf{y}_{N,t}]$ is an $L \times N$ matrix, where L denotes the number of bands, N is the total number of pixels in HS image, \mathbf{M}_t denotes an endmember matrix of size $L \times R$ with R pure materials (or endmembers) present in all observed scenes, \mathbf{A}_t is a $R \times N$ matrix composed of the abundance vector $\mathbf{a}_{n,t}$ at time t , \mathbf{dM}_t is an $L \times R$ matrix whose columns contain the variability inherent to the

t th image, and \mathbf{B}_t is an $L \times N$ matrix accounting for the noise at time t .

Considering that the materials in the scene are unchanged in a short-time interval, most of the endmembers are the unchanged in each image of the sequence, that is, $\mathbf{M} = \mathbf{M}_t$. Then, model (1) can be rewritten as

$$\mathbf{Y}_t = (\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t + \mathbf{B}_t. \quad (2)$$

B. Hierarchical Bayesian Model

In terms of Bayesian framework, PLMM-based unmixing is converted to maximize the joint posterior distribution

$$\begin{aligned} & p(\mathbf{M}, \mathbf{dM}, \mathbf{A} | \mathbf{Y}, \Phi) \\ &= \frac{p(\mathbf{Y} | \mathbf{M}, \mathbf{dM}, \mathbf{A}, \sigma^2) p(\mathbf{M} | \Phi) p(\mathbf{dM} | \Phi) p(\mathbf{A} | \Phi) p(\sigma^2 | \Phi)}{p(\mathbf{Y})} \end{aligned} \quad (3)$$

where \mathbf{M} , \mathbf{dM} , \mathbf{A} , σ^2 follow a given probability distribution, hyperparameters of these distributions are denoted as Φ , and $p(\mathbf{Y} | \mathbf{M}, \mathbf{dM}, \mathbf{A}, \sigma^2)$ is called the likelihood function. $p(\mathbf{M} | \Phi)$, $p(\mathbf{dM} | \Phi)$, $p(\mathbf{A} | \Phi)$, and $p(\sigma^2 | \Phi)$ are the priori distributions of endmember, endmember variability, abundance, and noise, respectively, which will be discussed later on.

According to the PLMM unmixing model (2) and the additive Gaussian noise, the likelihood function can be formulated as

$$p(\mathbf{Y} | \Theta) \propto \prod_{t=1}^T (\sigma_t^2)^{-NL/2} \exp\left\{-\frac{1}{2\sigma_t^2} \|\mathbf{Y}_t - (\mathbf{M} + \mathbf{dM}_t)\mathbf{A}_t\|_F^2\right\} \quad (4)$$

where $\|\cdot\|_F$ is the Frobenius norm, σ_t^2 represents the noise variance at time t , and Θ denotes all the parameters and hyperparameters defined in Sections II-B1–II-B4 and shown in Fig. 4, that is, $\Theta = \{\mathbf{M}, \mathbf{dM}_t, \mathbf{A}_t, \sigma_t^2, \Phi\}$.

1) *Abundance Prior in Temporal Domain*: When the prior of abundance is concerned, it is assumed that the coefficient of abundance is usually greater than zero and their summation is equal to 1, that is,

$$a_{n,t} \geq 0_R, a_{n,t}^T \mathbf{1}_R = 1 \quad \forall t \in \{1, \dots, T\} \quad (5)$$

where \geq denotes a term-wise inequality, 0_R represents the R -dimensional full zero vector, and $\mathbf{1}_R$ represents the R -dimensional full one vector.

For SHS, in a short time, the objection of the scene has slightly changed as shown in Fig. 1. Meanwhile, the abundance can also be regarded as less fluctuation within a small range in an adjacent time interval, which can be modeled by a normal distribution. So the priori distribution of the abundance in different time instant can be represented as [47]

$$p(\mathbf{a}_{n,t} | \mathbf{A}_{\setminus\{a_{n,t}\}}) \propto \exp\left\{-\frac{1}{2\varepsilon_n^2} \|\mathbf{a}_{n,t} - \mathbf{a}_{n,t-1}\|_2^2\right\} \mathbf{1}_{\mathbf{S}_R}(\mathbf{a}_{n,t}), \quad t \geq 2 \quad (6)$$

where ε_n^2 is the variance and usually it is a small fixed value (e.g., $\varepsilon_n^2 = 10^{-3}$), and $\mathbf{1}_{\mathbf{S}_R}$ is the indicator function of the set \mathbf{S}_R , which is defined as

$$\mathbf{S}_R = \{\mathbf{x} \in R^R | \forall i, x_i \geq 0 \text{ and } \mathbf{x}^T \mathbf{1}_R = 1\}. \quad (7)$$

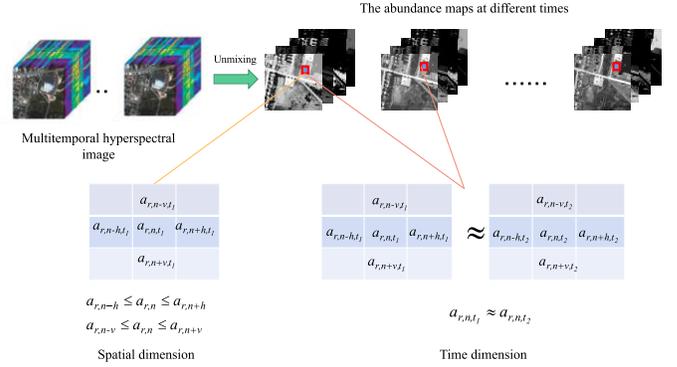


Fig. 1. Relationship of abundance coefficients in time and spatial dimensions.

While for $t = 1$, the uniform distribution is used as a priori distribution of the first moment of the abundance [27]

$$\mathbf{a}_{n,1} \sim U_{\mathbf{S}_R} \quad (8)$$

where $U_{\mathbf{S}_R}$ denotes the uniform distribution on the set \mathbf{S}_R .

2) *Endmember Prior in Transform Domain*: It is assumed that the pixels are composed by the linear combination of endmembers, which belong to a convex polytope simplex with the dimension of $K = R - 1$ [48]. Then, the dimension-reduced endmember \mathbf{e}_r can be obtained by means of principal component analysis (PCA) or robust PCA (rPCA) [49]. That is,

$$\mathbf{m}_r = \mathbf{U}\mathbf{e}_r + \check{\mathbf{y}}, \check{\mathbf{y}} = (\mathbf{I}_L - \mathbf{U}\mathbf{U}^T)\bar{\mathbf{y}}, \quad \mathbf{U}^T\mathbf{U} = \mathbf{I}_K \quad (9)$$

where \mathbf{U} denotes a basis of the subspace with K -dimension, $\check{\mathbf{y}}$ represents the projected variable, and $\bar{\mathbf{y}}$ denotes the average spectral signature obtained from \mathbf{Y} .

After the projection, the generated non-negative endmembers \mathbf{e}_r can be modeled as a truncated Gaussian prior [48]

$$\mathbf{e}_r \sim N_{\varepsilon_r}(\mathbf{0}_K, \zeta \mathbf{I}_K), \quad \text{for } r = 1, \dots, R \quad (10)$$

$$\boldsymbol{\varepsilon}_r = [e_{1,r}^-, e_{1,r}^+] \times \dots \times [e_{K,r}^-, e_{K,r}^+] \quad (11)$$

where ζ is set to a sufficiently large value to ensure an uninformative prior, and $e_{i,r}^-$ and $e_{i,r}^+$ are calculated according to the PCA transform, which ensures the non-negativity of endmember reconstructed by inverse PCA.

3) *Endmember Variability Prior in Temporal Dimension*: As proposed in PLMM, endmember variability is an additive perturbation term. By taking into account the non-negative nature of endmember (i.e., $\mathbf{m}_r + \mathbf{dM}_{r,t} \geq 0_L$), the prior distribution of endmember variability $\mathbf{dM}_{l,r,t}$ can be written as [47]

$$\mathbf{dM}_{r,1} | \mathbf{m}_r \sim N_{\zeta_r}(\mathbf{0}_L, v \mathbf{I}_L) \quad (12)$$

$$\mathbf{dM}_{l,r,t} | \mathbf{m}_{l,r}, \mathbf{dM}_{l,r,t-1}, \psi_{l,r}^2 \sim N_{\zeta_{l,r}}(\mathbf{dM}_{l,r,(t-1)}, \psi_{l,r}^2) \quad (13)$$

where $\zeta_r = \zeta_{1,r} \times \dots \times \zeta_{L,r}$, $\zeta_{l,r} \in [-m_{l,r}, +\infty)$ and $l = 1, \dots, L$, $r = 1, \dots, R$, $t = 1, \dots, T$, and v is a sufficiently small variance, and the variability variance Ψ^2 is associated with conjugate prior as shown below.

4) *Noise Variance and Hyperparameters Prior*: Inverse Gamma is the conjugate prior for a Gaussian likelihood with unknown variance. Then, the priors of noise variance σ_t^2 and

variability variance Ψ^2 for $t = 1, \dots, T$ and $l = 1, \dots, L$, $r = 1, \dots, R$ are satisfied as follows:

$$\sigma_t^2 \sim IG(a_\sigma, b_\sigma) \quad (14)$$

$$\psi_{l,r}^2 \sim IG(a_\psi, b_\psi) \quad (15)$$

where parameters $a_\sigma, b_\sigma, a_\psi, b_\psi$ are sufficiently small to ensure a weakly informative prior.

III. COMPOSITE PRIORS FOR ABUNDANCE AND ENDMEMBER VARIABILITY

The priors of abundance, endmember, and endmember variability as mentioned in Section II are well-modeled; however, the spatial structure of abundance, as well as the spectral smoothness of endmember variability, is omitted. Therefore, in this part, composited priors of abundance and endmember variability are both taken into consideration.

A. Abundance in Spatial Domain

The abundance prior in (6) only considers the time factor by omitting the spatial correlations in each HS image in an SHS as illustrated in Fig. 1. In fact, in a local spatial area, due to the consistency of the existing materials, the abundance coefficient $a_{r,n,t}$ is coincident with the surrounding adjacent points, and then Gaussian distribution prior can be used to model abundance $a_{r,n,t}$ in the spatial domain as

$$a_{r,n,t} \sim N(\mu, \sigma_a^2) \quad (16)$$

where μ is the mean of $a_{r,n,t}$ and σ_a^2 is the variance.

The average of the surrounding points of $a_{r,n,t}$ can be taken as the mean of the current point, which is denoted as $\bar{a}_{r,n,t}$. In addition, the abundance also has spatial smoothness, which means it is similar to the surrounding points with high probability at the current time. Therefore, a high reliable interval decided by data itself is needed to ensure that the sampled value located near to its adjacent points with a high probability. The definition of local difference makes it appropriate to represent the continuity and the smoothness of an image in a local spatial neighborhood. Typically, a smaller difference means better smoothness of the adjacent points. Hence, the local difference is used to define a spatial data-adaptive variance as

$$\sigma(a_{r,n,t}) = \alpha \sqrt{\left((\nabla_{r,n,t}^h(a_{r,n,t}))^2 + (\nabla_{r,n,t}^v(a_{r,n,t}))^2 \right) / 2} \quad (17)$$

where the operators $\nabla_{r,n,t}^h(a_{r,n,t})$ and $\nabla_{r,n,t}^v(a_{r,n,t})$ represent the horizontal and vertical first-order differences, respectively, and α is used to adjust the probability.

The prior distribution of $a_{r,n,t}$ in the spatial domain is written as

$$p(a_{r,n,t} | \mathbf{A}_{\setminus\{a_{r,n,t}\}}) \propto \exp\left\{ -\frac{\|a_{r,n,t} - \bar{a}_{r,n,t}\|_2^2}{2\sigma^2(a_{r,n,t})} \right\} \mathbf{1}_{\mathbf{S}_R}(\mathbf{a}_{n,t}). \quad (18)$$

Finally, by incorporating (6) and (18), a mixed prior distribution of the abundance can be formulated as

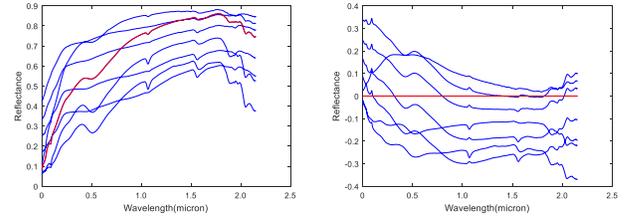


Fig. 2. (Left) Eight endmember diagrams of the Andesine obtained from USGS. (Right) Reference endmember (red) and the differences (blue).

For $t = 1$

$$p(a_{r,n,1} | \mathbf{A}_{\setminus\{a_{r,n,1}\}}) \propto \exp\left\{ -\frac{1}{2\sigma^2(a_{r,n,1})} \|a_{r,n,1} - \bar{a}_{r,n,1}\|_2^2 \right\} \times \mathbf{1}_{\mathbf{S}_R}(\mathbf{a}_{n,1}). \quad (19)$$

For $t > 1$

$$p(a_{r,n,t} | \mathbf{A}_{\setminus\{a_{r,n,t}\}}) \propto \exp\left\{ -\frac{1}{2\varepsilon_n^2} \|a_{r,n,t} - a_{r,n,t-1}\|_2^2 \right\} * \exp\left\{ -\frac{1}{2\sigma^2(a_{r,n,t})} \|a_{r,n,t} - \bar{a}_{r,n,t}\|_2^2 \right\} \mathbf{1}_{\mathbf{S}_R}(\mathbf{a}_{n,t}). \quad (20)$$

From the defined composed priors, the abundance value of each pixel at a spatial position is correlated with its neighbors, which may lead to a more complex posterior distribution. To simplify the form of the posterior distribution, the abundance of the previous iteration is used to calculate the mean value and variance. After several iterations, the results can approximate the complex posterior distribution.

B. Endmember Variability in Spectral Dimension

Fig. 2 shows the endmember curves of the same material (Andesine) and their difference curves. One of the curves (the red one) is selected randomly as a reference curve, and the differences between this one and others are shown in the left-hand side of Fig. 2. The differences show that the endmember variability tends to be a smooth variation with the bands. Moreover, the curve of variability is not necessarily as random as the noise does, which indicates not only the consistency in time (12) and (13) but also the smoothness along with bands. These observations help a lot to formulate the prior of the endmember variability. Fig. 3 illustrates both the characters of the endmember variability. Specifically, in the spectral domain, the endmember variability of the adjacent bands has the same trend at the same time, resulting a new hypothesis of the smoothness of endmember variability in spectral dimension. Finally, a normal distribution is used to model the prior of $dm_{l,r,t}$ as

$$dm_{l,r,t} | m_{l,r}, dm_{l\pm 1,r,t}, s_{l,r}^2 \sim N_{\zeta_{l,r}}(dm_{l\pm 1,r,t}, s_{l,r}^2), \quad l = 1, L \quad (21)$$

$$dm_{1,r,t} | m_{1,r}, dm_{l\pm 1,r,t}, s_{l,r}^2 \sim N_{\zeta_{l,r}}((dm_{l-1,r,t} + dm_{l+1,r,t})/2, s_{l,r}^2) \quad (22)$$

where $s_{l,r}^2$ is the variance and usually it is a small fixed value.

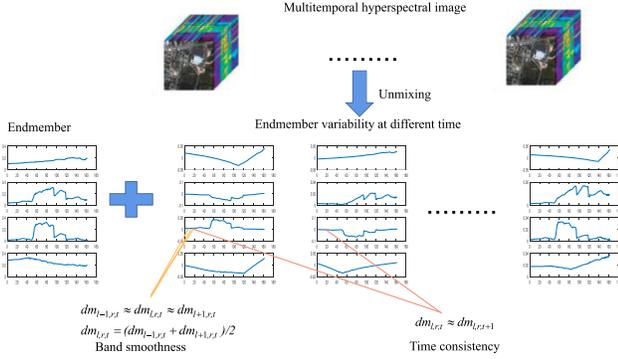


Fig. 3. Endmember curves and possible endmember variability curves.

Finally, by incorporating (13) and (22), a composite prior distribution of the endmember variability is formulated as follows:

For $l = 1, L; t = 1$

$$\begin{aligned}
 & p(dm_{1,r,t} | dm_{l\pm 1,r,t}, m_{l,r}) \\
 & \propto \exp\left\{-\frac{1}{2v} \|dm_{1,r,t} - 0\|_2^2\right\} \\
 & * \exp\left\{-\frac{1}{2s_{1,r}^2} \|dm_{1,r,t} - dm_{l\pm 1,r,t}\|_2^2\right\}. \quad (23)
 \end{aligned}$$

For $l = 1, L; t \geq 2$

$$\begin{aligned}
 & p(dm_{1,r,t} | dm_{l\pm 1,r,t}, dm_{l,r,t-1}, m_{l,r}) \\
 & \propto \exp\left\{-\frac{1}{2v} \|dm_{1,r,t} - dm_{l,r,t-1}\|_2^2\right\} \\
 & * \exp\left\{-\frac{1}{2s_{1,r}^2} \|dm_{1,r,t} - dm_{l\pm 1,r,t}\|_2^2\right\}. \quad (24)
 \end{aligned}$$

For $2 \leq l \leq L - 1, t = 1$

$$\begin{aligned}
 & p(dm_{l,r,t} | dm_{l\pm 1,r,t}, dm_{l,r,t-1}, m_{l,r}, \psi_{l,r}^2) \\
 & \propto \exp\left\{-\frac{1}{2v} \|dm_{l,r,t} - 0\|_2^2\right\} \\
 & * \exp\left\{-\frac{1}{2s_{l,r}^2} \left\|dm_{l,r,t} - \frac{dm_{l-1,r,t} + dm_{l+1,r,t}}{2}\right\|_2^2\right\}. \quad (25)
 \end{aligned}$$

For $2 \leq l \leq L - 1, t \geq 2$

$$\begin{aligned}
 & p(dm_{l,r,t} | dm_{l\pm 1,r,t}, dm_{l,r,t-1}, m_{l,r}, \psi_{l,r}^2) \\
 & \propto \exp\left\{-\frac{1}{2\psi_{l,r}^2} \|dm_{l,r,t} - dm_{l,r,t-1}\|_2^2\right\} \\
 & * \exp\left\{-\frac{1}{2s_{l,r}^2} \left\|dm_{l,r,t} - \frac{dm_{l-1,r,t} + dm_{l+1,r,t}}{2}\right\|_2^2\right\}. \quad (26)
 \end{aligned}$$

IV. GIBBS SAMPLER STRATEGY FOR SHS UNMIXING

By incorporating the proposed priors into the Bayesian framework, a hierarchical Bayesian (HB) unmixing model is proposed as illustrated in Fig. 4. It results in the maximization posterior as follows:

$$p(\Theta | \mathbf{Y}) \propto p(\mathbf{Y} | \Theta) p(\mathbf{dM} | \mathbf{M}, \Psi^2) p(\mathbf{M}) p(\mathbf{A}) p(\sigma^2) p(\Psi^2) \quad (27)$$

An MCMC method is consequently adopted to sample the posterior (27) and build estimators of the parameters involved

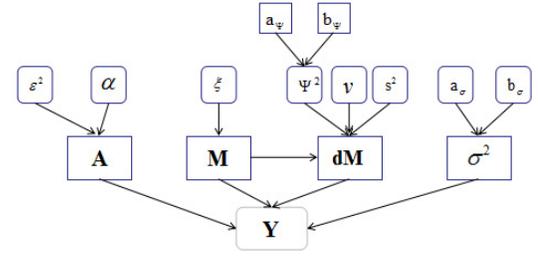


Fig. 4. Directed acyclic graph associated with the proposed Bayesian model.

Algorithm 1 Gibbs Sampler

Data: $N_{bi}, N_{MC}, \mathbf{M}^{(0)}, \mathbf{A}^{(0)}, \mathbf{dM}^{(0)}, (\sigma^2)^{(0)}, (\psi^2)^{(0)}$

Begin

For $q = 1: N_{MC}$

Sampling the endmembers $\mathbf{M}^{(q)}$ according to Eq. (35);

Sampling the endmember variability terms $\mathbf{dM}^{(q)}$ according to Eq. (37);

Sampling the abundances $\mathbf{A}^{(q)}$ according to Eq. (31);

Sampling the noise variances $\sigma^{2(q)}$ according to Eq. (42);

Sampling the variability variances $\psi^{2(q)}$ according to Eq. (43).

End

End

Result: $\{\mathbf{M}^{(q)}, \mathbf{dM}^{(q)}, \mathbf{A}^{(q)}, \sigma^{2(q)}, \psi^{2(q)}\}_{q=1}^{N_{MC}}$

in the proposed Bayesian model using the generated samples. The main steps of the proposed Gibbs sampler are summarized in Algorithm 1. After N_{bi} burn-in iterations, the minimum mean square error (MMSE) estimator of the SHS is used as the expected value of the unknown parameters \mathbf{M} , \mathbf{dM}_t and \mathbf{A}_t

$$\hat{\mathbf{M}}^{MMSE} \approx \frac{1}{N_{MC} - N_{bi}} \sum_{q=N_{bi}+1}^{N_{MC}} \mathbf{M}^{(q)} \quad (28)$$

$$\hat{\mathbf{dM}}^{MMSE} \approx \frac{1}{N_{MC} - N_{bi}} \sum_{q=N_{bi}+1}^{N_{MC}} \mathbf{dM}^{(q)} \quad (29)$$

$$\hat{\mathbf{A}}^{MMSE} \approx \frac{1}{N_{MC} - N_{bi}} \sum_{q=N_{bi}+1}^{N_{MC}} \mathbf{A}^{(q)} \quad (30)$$

where q represents the iterations of sampling and N_{MC} is the maximum iteration number.

A. Sampling the Abundance \mathbf{A}

The likelihood function (4) combined with the prior (19) and (20) leads to the following conditional distribution for the abundances:

$$\mathbf{a}_{n,t} | \mathbf{y}_{n,t}, \Theta \setminus \{\mathbf{a}_{n,t}\} \sim N_{S_R}(\mu_{n,t}^{(\mathbf{A})}, \Sigma_{n,t}^{(\mathbf{A})}) \quad (31)$$

$$\begin{aligned}
 & (\Sigma_{n,t}^{(\mathbf{A})})^{-1} \\
 & = \frac{1}{\sigma_t^2} \mathbf{M}_t^T \mathbf{M}_t + \frac{1}{\varepsilon_n^2} ([t > 1] + [t < T]) \mathbf{I}_R \\
 & + \text{diag}\left(\frac{1}{\sigma^2(a_{1,n,t})}, \dots, \frac{1}{\sigma^2(a_{R,n,t})}\right) \quad (32)
 \end{aligned}$$

$$\mathbf{M}_t = \mathbf{M} + \mathbf{dM}_t \quad (33)$$

$$\mu_{n,t}^{(\mathbf{A})} = \sum_{n,t}^{(\mathbf{A})} \left(\frac{1}{\sigma_t^2} \mathbf{M}_t^T \mathbf{y}_{n,t} + \frac{1}{\varepsilon_n^2} ([t > 1] \mathbf{a}_{n,t-1} + [t < T] \mathbf{a}_{n,t+1}) + \left(\frac{\bar{a}_{1,n,t}}{\sigma^2(a_{1,n,t})}, \dots, \frac{\bar{a}_{R,n,t}}{\sigma^2(a_{R,n,t})} \right)^T \right) \quad (34)$$

where $N_{S_R}(\mu, \mathbf{A})$ denotes a Gaussian distribution truncated to the set \mathbf{S}_R . The symbol $[p]$ denotes the logical operation, namely, it returns to 1 when p is true; otherwise, it returns to 0.

B. Sampling the Endmember \mathbf{M}

The projected endmembers \mathbf{e}_r , for $r = 1, R$ have a truncated multivariate Gaussian distribution as shown in (10) and (11), and then the conditional posterior distribution is formulated as

$$\mathbf{e}_r | \mathbf{Y}, \Theta_{\setminus \{\mathbf{e}_r\}} \sim N_{e_r}(\mu_r^{(\mathbf{E})}, \Sigma_r^{(\mathbf{E})}) \quad (35)$$

$$(\Sigma_r^{(\mathbf{E})})^{-1} = \left(\frac{1}{\xi} + \sum_{n,t} \frac{a_{r,n,t}^2}{\sigma_t^2} \right) \mathbf{I}_{R-1} \quad (36)$$

$$\mu_r^{(\mathbf{E})} = \Sigma_r^{(\mathbf{E})} \mathbf{U}^T \left(\sum_{t,n} \frac{1}{\sigma_t^2} \left(\mathbf{y}_{n,t} - \mathbf{dM}_t \mathbf{a}_{n,t} - \check{\mathbf{y}}_{l,r} a_{r,n,t} - \sum_{j \neq r} a_{j,n,t} \mathbf{m}_j \right) a_{r,n,t} \right). \quad (37)$$

C. Sampling the Variability Term \mathbf{dM}

Similarly, the combination of the likelihood function (4) and the priors (23)–(26) leads to

$$dm_{l,r,t} \sim N_{[-m_{l,r}, +\infty)}(\mu_{l,r,t}^{(\mathbf{dM})}, \eta_{l,r,t}^2) \quad (38)$$

$$\frac{1}{\eta_{l,r,t}^2} = \frac{1}{\sigma_t^2} \sum_n a_{r,n,t}^2 + \frac{1}{v} [t = 1] + \frac{1}{\psi_{l,r}^2} (1 + [1 < t < T]) + \frac{1}{2s_{l,r}^2} ([l > 1] + 1 + [l < L]) \quad (39)$$

$$\mu_{l,r,t}^{(\mathbf{dM})} = \left(\frac{1}{\sigma_t^2} (\check{\mathbf{y}}_{l,t} - \mathbf{d}\check{\mathbf{m}}_{l,\setminus r,t} \mathbf{A}_{\setminus r,t} - \check{\mathbf{m}}_l \mathbf{a}_{n,t}) \check{\mathbf{a}}_{r,t}^T + \frac{1}{\psi_{l,r}^2} ([t < T] dm_{l,r,t+1} + [t > 1] dm_{l,r,t-1}) + \frac{1}{s_{l,r}^2} \times \frac{1}{4} \{ (2 + [l > 1] + [l > 2]) dm_{l-1,r,t} + (2 + [l < L - 1] + [l < L - 2]) dm_{l+1,r,t} - [l > 2] dm_{l-2,r,t} - [l < L - 2] dm_{l+2,r,t} \} \right) \times \eta_{l,r,t}^2; dm_{(0 \text{ or } L+1),r,t} = dm_{(1 \text{ or } L),r,t} \quad (40)$$

where $\mathbf{d}\check{\mathbf{m}}_{l,\setminus r,t}$ denotes the l th row of \mathbf{dM}_t whose r th element has been removed, $\check{\mathbf{m}}_l$ is the l th row of \mathbf{M} , $\check{\mathbf{y}}_{l,t}$ represents the n -dimensional vector of the l -band at t -time, $\check{\mathbf{a}}_{r,t}$ is the n -dimensional vector of r -endmember at t -time, and $\mathbf{A}_{\setminus r,t}$ is the matrix \mathbf{A}_t without its r th row. The rows of each

variability matrix \mathbf{dM}_t can be sampled in parallel to reduce the computational time of the sampler. According to the method in [50]–[52], the truncated multivariate Gaussian distribution of unknown parameters \mathbf{M} , \mathbf{dM}_t and \mathbf{A}_t is sampled.

D. Sampling the Noise Variances σ^2 and the Variability Variance Ψ^2

Using the likelihood function (4) and priors (14) and (15), the sampled noise and variability variances are represented as

$$\sigma_t^2 | \mathbf{Y}_t, \Theta_{\setminus \{\sigma_t^2\}} \sim IG \left(a_\sigma + \frac{LN}{2}, b_\sigma + \frac{1}{2} \|\mathbf{Y}_t - (\mathbf{M} + \mathbf{dM}_t) \mathbf{A}_t\|_F^2 \right) \quad (41)$$

$$\psi_{l,r}^2 | \Theta_{\setminus \{\psi_{l,r}^2\}} \sim IG \left(a_\psi + \frac{T-1}{2}, b_\psi + \frac{1}{2} \sum_{t=2}^T (dm_{l,r,t} - dm_{l,r,t-1})^2 \right). \quad (42)$$

V. EXPERIMENTAL RESULTS

In this section, both the synthetic and real data are conducted to verify the effectiveness of the proposed method. The compared unmixing methods include the classical unmixing methods, vertex component analysis (VCA)/fully constrained least squares (FCLS) [13] and simplex identification via split augmented Lagrangian (SISAL)/FCLS [8], in which spectral variability is not considered. Unmixing methods consider spectral variability, that is, PLMM [37], ELMM [30], online unmixing (OU) [46], and HB [47], in which the first two are proposed by improving the LMM, the OU method is an improved PLMM, while the HB method analyzes the SHS unmixing problem from the point of Bayesian view.

For more comprehensive comparison, indexes in terms of estimation error of endmember, abundance, endmember variability, and the total reconstruction error (RE) are computed, which are defined as follows.

- 1) aSAM: endmember estimation through the average spectral angle mapper

$$\text{aSAM}(\mathbf{M}) = \frac{1}{R} \sum_{r=1}^R \arccos \left(\frac{\mathbf{m}_r^T \hat{\mathbf{m}}_r}{\|\mathbf{m}_r\|_2 \|\hat{\mathbf{m}}_r\|_2} \right). \quad (43)$$

- 2) Abundance and variability estimation through the global mean square errors (GMSEs)

$$\text{GMSE}(\mathbf{A}) = \frac{1}{\text{TRN}} \sum_{t=1}^T \left\| \mathbf{A}_t - \hat{\mathbf{A}}_t \right\|_F^2 \quad (44)$$

$$\text{GMSE}(\mathbf{dM}) = \frac{1}{\text{TLR}} \sum_{t=1}^T \left\| \mathbf{dM}_t - \hat{\mathbf{dM}}_t \right\|_F^2. \quad (45)$$

- 3) The total RE

$$\text{RE} = \frac{1}{\text{TLN}} \sum_{t=1}^T \left\| \mathbf{Y}_t - \hat{\mathbf{Y}}_t \right\|_F^2. \quad (46)$$

The parameters of the compared methods are adjusted manually to optimal in terms of aSAM of endmember and GMSE of abundance, simultaneously. In the proposed method,

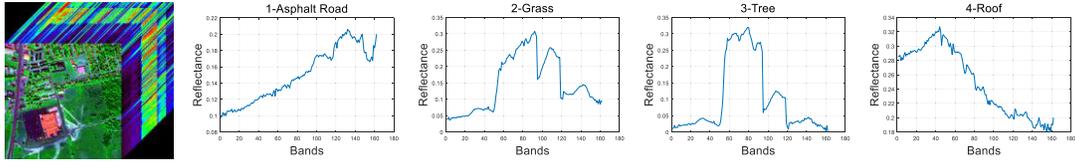
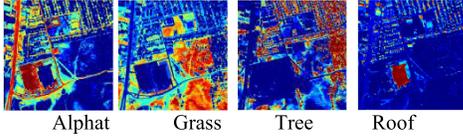
Fig. 5. *Urban* data set and the reference four endmember curves.Fig. 6. Corresponding abundance maps of *Urban* data set with four endmembers.

TABLE I
FIXED PARAMETERS AND INITIAL VARIABLES

Parameters	Synthetic data	Real data
ϵ_n^2	10^{-3}	10^{-2}
ξ	1	1
ν	10^{-3}	10^{-5}
a_ν, a_σ	10^{-3}	10^{-3}
b_ν, b_σ	10^{-3}	10^{-3}
N_{bi}	350	450
N_{MC}	400	500
σ_t^2	10^{-4}	10^{-4}
ψ_{lr}^2	10^{-3}	10^{-2}
s_{lr}^2	10^{-4}	10^{-5}
α	1.2	1

most of the parameters, which also appeared in the HB method, are fixed as that of HB as shown in Table I. While for the newly occurring parameters, α and s_{lr}^2 in our method, they are adjusted according to parameter analysis given later on. Moreover, to alleviate the effects of the random sample, the final result is the average of multiple iterations, which also benefits the setting of the parameters and hyperparameters.

A. Synthetic Data

In the simulated experiments, the data set comes from *Urban* data set with size $307 \times 307 \times 162$, which has the ground truth with 4, 5, and 6 endmembers (*Asphalt, Road, Grass, Tree, Roof, Metal, and Dirt*) and the corresponding abundances. Figs. 5 and 6 give the *Urban* data set and the ground truth of $R = 4$ endmembers and their abundance maps, respectively. It should be noted that the images of the experiments do not satisfy the pure pixel assumption, which provides an assessment of the proposed method in more challenging situations.

The synthetic SHS is produced in the following way. First, the real abundance map is regarded as a reference, and then other abundance maps are generated by multiplying the reference maps with trigonometric functions (e.g., $\cos((\pi/100) + t \times (36\pi/100))$ and $\sin((\pi/100) + t \times (36\pi/100))$), with $t = 1, \dots, T$ ($T = 6$) to ensure that the evolution of time is smooth enough. Besides, the temporal evolution of the

last abundance map is obtained by leveraging the sum-to-one condition. On the other hand, the endmember variability is modeled by multiplying the reference endmembers with randomly piecewise affine functions as mentioned in [17]. Finally, the abundance maps multiple the summation of the given endmembers and their variabilities, and then in addition to the additive white Gaussian noises (signal-to-noise ratio around 25 dB), the synthetic SHS is obtained.

In addition, unmixing experiments on different endmember numbers, namely, 4, 5, and 6, are conducted, respectively. Figs. 7 and 8 show the visual unmixing results of $R = 4$ endmembers. Table I reports the performance assessment for the data set associated with mixtures of 4, 5, and 6 endmembers.

1) *Endmember and Its Variability Estimation*: Fig. 7 shows the endmembers and their variability with four endmembers. As can be seen from Fig. 7, the traditional VCA- and SISAL-based methods perform in a similar way; both of them can detection the endmember well, but cannot distinguish the endmember variability. The resulted curves are far away from that of the ground truth, especially for the fourth endmember, which is caused by the absence of the non-negative constraints of the endmember. It should be noted that some indexes of SISAL are smaller than zero, which may be caused by omitting the non-negative of endmember matrix in the model. Although PLMM and ELMM concern the spectral variability, the estimated endmember variability curves are too compacted to represent the variation in spectrum. It may be due to the regularization constraints of the endmember variability in the models, in which only energy minimization is involved. Moreover, the estimated endmember curves of PLMM and ELMM differ greatly from the actual curves both in amplitude and trajectory, especially for the first, second, and the fourth endmembers. For the OU method, the first endmember approximates the original one well; however, the second and the fourth endmembers have bigger deviation from the original curves. On the other hand, the Bayesian-based methods, including HB and the proposed method, have a strong ability to extract the endmember and endmember variability. The proposed method shows more accurate approximation to the endmember variability than that of the HB method. Obviously difference between the two methods can be seen from the middle parts in the first endmember curves and the highest peaks in the third endmembers curves, respectively. Owing to the composited prior of the endmember variability, the proposed method provides competitive unmixing results while allowing temporal endmember variability to be estimated for each endmember. Furthermore, the indexes in Table I also confirm the superiority of the proposed method.

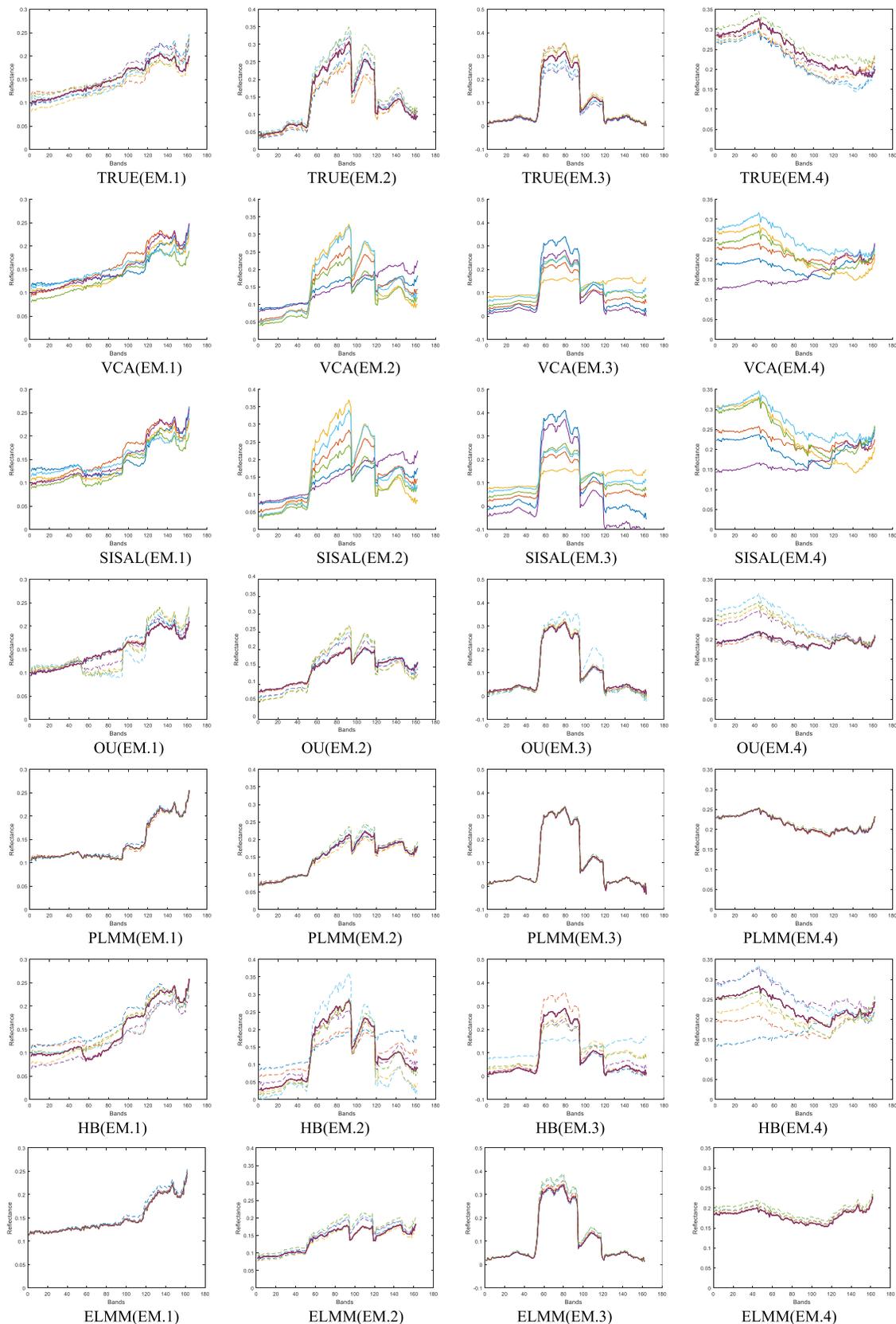


Fig. 7. Endmembers (\mathbf{m}_r , purple lines) and their endmember variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, dotted lines) recovered from synthetic data by different methods.

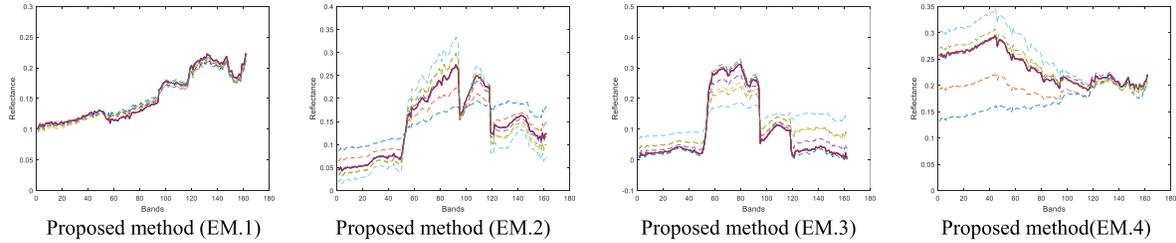


Fig. 7. (Continued.) Endmembers (\mathbf{m}_r , purple lines) and their endmember variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, dotted lines) recovered from synthetic data by different methods.

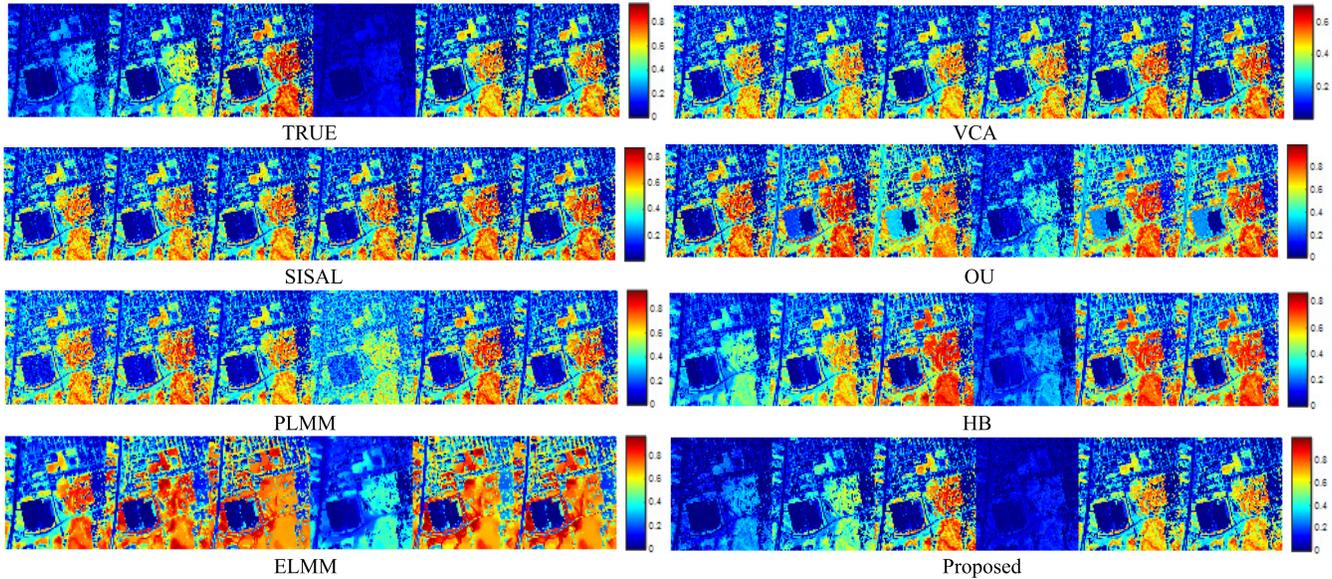


Fig. 8. Change in the second endmember's abundance graph at different time extracted from synthetic data.

2) *Abundance Estimation*: The corresponding abundance maps of Fig. 7 are shown in Fig. 8. For the second endmember, the abundance maps obtained by VCA, SISAL, PLMM, ELMM, and OU methods are in heavy colors than the ground truth with large abundance coefficients, which also coincides with the endmember estimation in Fig. 7. Meanwhile, the worst cases are $t = 1$ and $t = 4$ for overestimating the coefficients. Comparing SISAL and VCA, the abundance maps of them look alike; the difference is the SISAL abundance is more dark than that of VCA, which means the coefficients of the SISAL are slightly greater than that of VCA. Both VCA and SISAL perform better at $t = 3, 5,$ and 6 . PLMM offers improvement but may also overestimate at $t = 1$ and $t = 4$. ELMM produces excessive abundance estimates. The abundance maps obtained by HB and the proposed method show an improved resemblance as ground truth. At $t = 1$, both produce a great improvement over others; at $t = 4$, the proposed method obtains the best approximation; at $t = 2, 3, 5,$ and 6 , at the left-hand side of the images, the color of blue is closer to the color of ground truth, while HB shows heavy orange color at the same area. Furthermore, the whole abundance values of HB are greater than that of the ground truth. In general, the proposed method gives the abundance maps which are closest to the real ones. Moreover, the abundance is smoother than other methods in the spatial domain.

3) *Total Performance*: Table II shows the evaluation indexes in different numbers of endmember for different approaches. The performance for each methods are consistent for $R = 4, 5,$ and 6 . The indexes of VCA and SISAL are higher than other methods due to lack of endmember variability in the models. The proposed method has the best values in SAM and GMSE(A), which demonstrates that the proposed method has a strong ability to distinguish the endmember and achieve an accurate abundance approximation. In terms of GMSE(dM), PLMM, ELMM, and OU methods yield more accurate numerical results than other methods. It is the fact that PLMM, ELMM, and OU methods establish a corresponding term for each pixel. However, this does not necessarily lead to a better visual approximation of the endmember variability (Fig. 7). On the contrary, the Bayesian-based methods only establish a general \mathbf{dm}_t term for all pixels at the same time to reduce the time consumption. Therefore, they admit higher values in GMSE of \mathbf{dm} . In terms of RE, PLMM attempts to minimize the energy function, which is consistent with the RE, so PLMM achieves the best RE value. The proposed method takes the second place in RE value, implying better unmixing results than other methods. Compared with HB, the proposed method shows its superiority not only on abundance and endmember variability but also on endmember and RE, which further proves the effectiveness of the improved prior assumptions of endmember variability and abundance.

TABLE II
QUANTITATIVE RESULTS OF SYNTHETIC DATA

Endmember	methods	aSAM(M)	GMSE(A)	GMSE(dM)	RE
$R = 4$	VCA/FCLS	8.98	0.0428	/	6.04e-5
	SISAL/FCLS	9.97	0.0382	/	5.87e-5
	OU	7.56	<u>0.0264</u>	7.1e-4	4.92e-5
	PLMM	6.82	0.0334	4.50e-4	8.02e-6
	HB	<u>4.02</u>	0.0315	0.0013	4.92e-5
	ELMM	7.94	0.0433	<u>6.03e-4</u>	5.17e-5
	Proposed method	2.82	0.0246	0.0011	<u>4.89e-5</u>
$R = 5$	VCA/FCLS	10.32	0.0342	/	8.36e-5
	SISAL/FCLS	9.83	0.0331	/	7.07e-5
	OU	4.34	0.0312	0.0017	6.37e-5
	PLMM	7.13	0.0380	0.0011	8.58e-6
	HB	<u>3.83</u>	<u>0.0257</u>	0.0024	4.94e-5
	ELMM	6.52	0.0414	<u>0.0014</u>	5.24e-5
	Proposed method	3.40	0.0204	0.0016	<u>4.80e-5</u>
$R = 6$	VCA	11.21	0.0418	/	7.21e-5
	SISAL	10.32	0.0398	/	5.91e-5
	OU	6.43	0.0330	0.0011	4.97e-5
	PLMM	8.55	0.0384	9.08e-4	6.01e-6
	HB	7.02	<u>0.0221</u>	0.0036	4.88e-5
	ELMM	8.07	0.0388	<u>0.0010</u>	6.24e-5
	Proposed method	6.08	0.0198	0.0017	<u>4.79e-5</u>



Fig. 9. Sequence of real images taken in the same area at different time.

4) *Real Data*: The proposed algorithm has been applied to real HS images acquired by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) over the Lake Tahoe region (CA, USA) between 2014 and 2015. Water absorption bands were removed from the 224 spectral bands, leading to 169 exploitable spectral bands. The subscene of interest (50×50) is partly composed of a lake and a nearby field (Fig. 9), and it is suggested with three endmembers (*Water*, *Soil*, and *Vegetation*) [22]. The parameters used for the proposed approach are given in Table I. The original HS image sequence is shown in Fig. 9, and RE is obtained as an evaluation metric.

5) *Endmember Estimation*: Fig. 10 shows the spectral curve and its variability at different time for three endmembers. In general, the spectral curves of the compared methods reflect that the endmembers are varied, in which VCA and SISAL have more messy curves, while others are rather regular. For the second endmember (*Water*), it should also be mentioned that the PLMM and OU methods may produce negative endmembers with perturbed variability. The curves for HB and the proposed methods at $t = 3$ have large values than that of other methods, which means a big variability occurs. It is coincident with the real image shown in Fig. 9, in which the color of *Water* in image $t = 3$ is very different from others. For endmember of *Veg.* and *Soil*, the same conclusion can be drawn from Figs. 9 and 10.

6) *Abundance Estimation*: In the real data (Fig. 9), at $t = 1, 2,$ and 5 , the second endmember (*Water*) has a large percent in the images. At $t = 3, 4,$ and 6 , the first and the third endmembers (*Veg.* and *Soil*) are mixed more deeply with large spectral variation.

TABLE III
QUANTITATIVE RESULTS OF REAL DATA

Methods	RE
VCA/FCLS	1.12e-4
SISAL/FCLS	1.21e-4
OU	2.1e-5
PLMM	9.92e-6
HB	2.01e-5
ELMM	8.77e-5
Proposed method	1.95e-5

The abundance maps for three endmembers are shown in Figs. 11–13. The estimated abundances of VCA and SISAL are very different from the ground truth. For the abundance of *Water*, it is supposed to have large values at $t = 1, 2,$ and 5 according to the ground truth. However, Fig. 12 shows relatively small values at these points. The same case happens for other two endmembers; in Figs. 11 and 13, there are mess estimated errors at $t = 1$ for *Veg.* and *Soil*, respectively. ELMM fails to discern *Water*, especially at $t = 3$ and 5 (Fig. 12). Meanwhile, it mixes *Soil* with *Veg.* together, resulting in higher abundance of *Soil* and lower abundance of *Veg.* at all time instants in Figs. 13 and 11, respectively. The OU method has the same disadvantage in distinguishing *Veg.* and *Water*, as shown in Figs. 11–13, especially at $t = 1, 2,$ and 3 . In Fig. 13, the PLMM, HB and the proposed methods have very similarity abundance maps of *Soil*. However, PLMM fails to estimate the abundance coefficient accurately at the top right hand at $t = 6$ in Fig. 13. Compared with HB, the proposed method exhibits more smooth and accurate abundance. The same conclusion can also be drawn from Table III; PLMM, HB and the proposed methods have relatively low RE than others. Moreover, the proposed method gets the second place and is smaller than that of HB.

B. Discussions

1) *Parameter Analysis*: It should be mentioned that in the Bayesian-based method, there are a lot of parameters

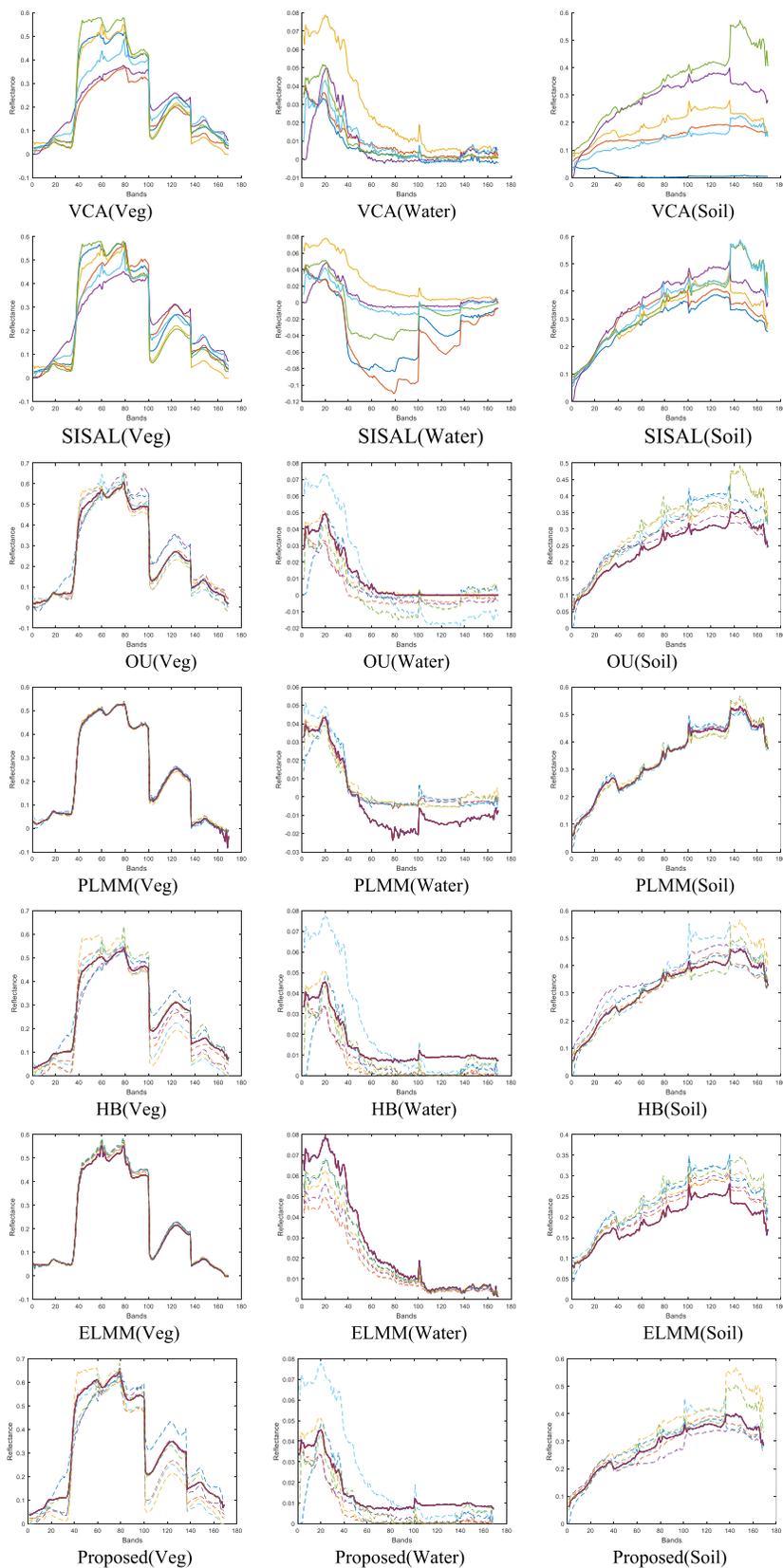


Fig. 10. Endmembers (\mathbf{m}_r , purple lines) and their endmember variability ($\mathbf{m}_r + \mathbf{dm}_{r,t}$, dotted lines) recovered from real data by different methods.

to be decided. However, the Bayesian-based method has better generalization ability, which results in a wide range of parameters. More precisely, the smaller the variance in

Gaussian distributions, the more prior information will be given. Therefore, variances of abundance, noise, and end-member variability, including ε_n^2 , σ_t^2 , v , $\psi_{l,r}^2$, and $s_{l,r}^2$, are

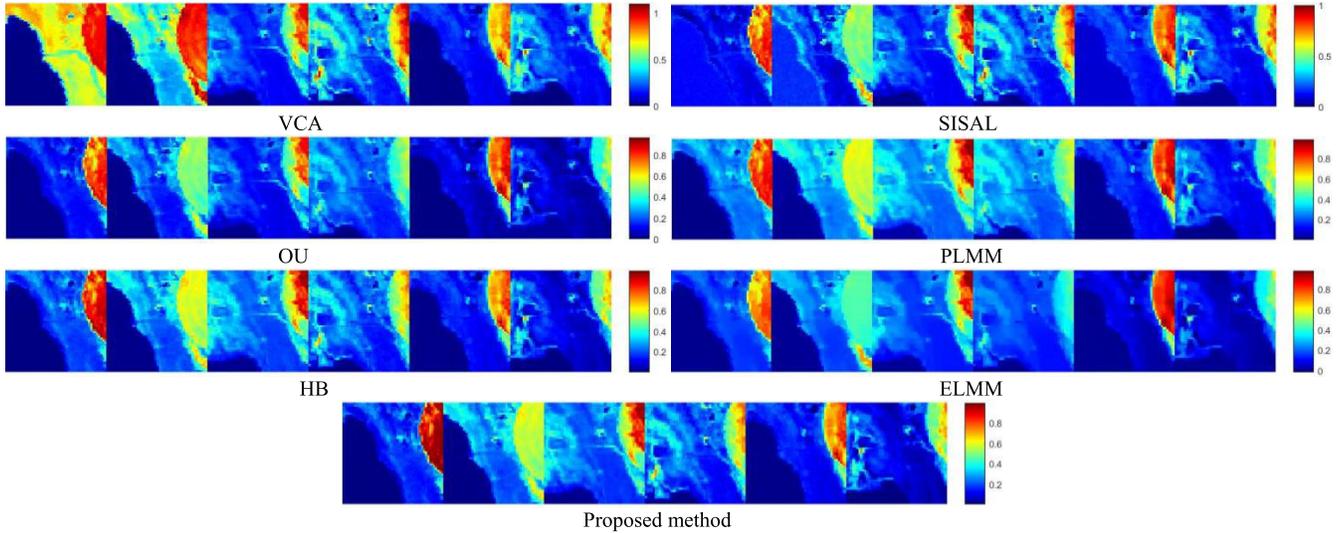


Fig. 11. Change in the first endmember's (*Veg.*) abundance map at different time extracted from real data (50×50).

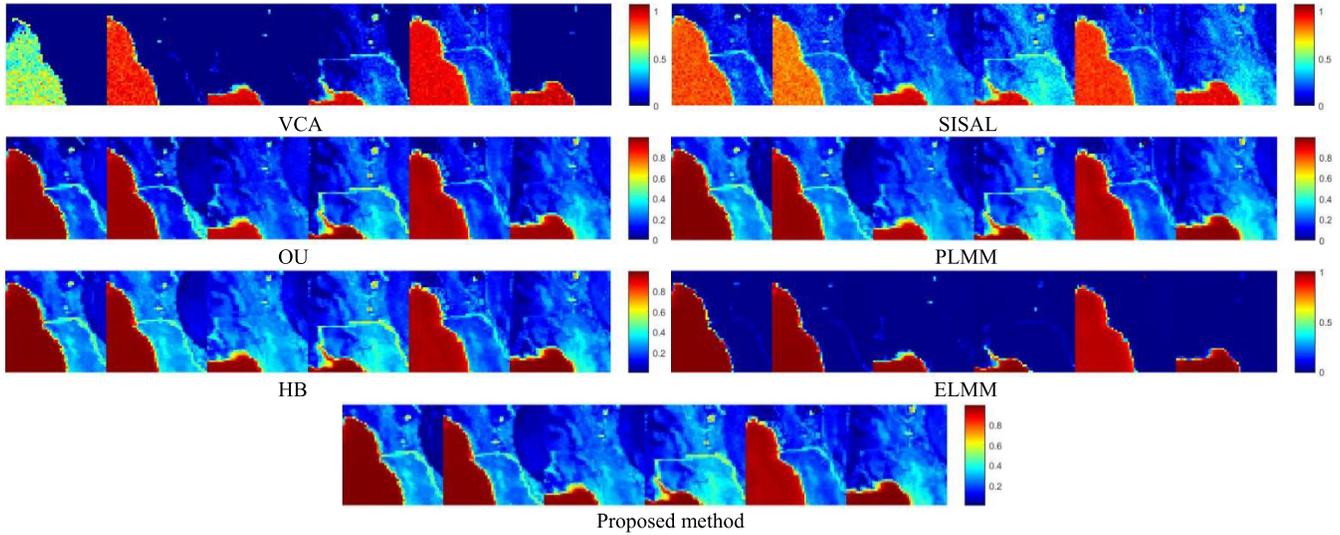


Fig. 12. Change in the second endmember's (*Water*) abundance map at different time extracted from real data.

small enough. In addition, according to the observation that the abundance in the time domain on the synthetic data has less variation than that of the real data, the variance of abundance on the real data is larger than that of the stimulated data to decrease the prior constraints in the temporal domain.

Furthermore, to reduce the dependence on the prior, the so-called noninformative prior is encouraged. Different from the Gaussian distribution, in the inverse Gamma distribution, the provided information by the parameters is proportional to the parameters' values, which results in more information being offered by the likelihood term associated with a small parameter value. Hence, the noise variance, as well as hyperparameters, is fixed to sufficient small values, namely, $a_\sigma = b_\sigma = a_\psi = b_\psi = 10^{-3}$ both for the synthetic and real data. The same rule is also applied to endmember parameter ξ that equals 1 in the experiments. In terms of the covariance and mean in the posterior distribution (36) and (37),

$\sum_{n,t} (a_{r,n,t}^2 / \sigma_t^2)$ is greatly larger than $(1/\xi)$ when $\sigma_t^2 = 10^{-4}$, $(1/\xi) = 1$, which results a noninformation prior.

The abundance parameter α in (17) controls the scale of the local difference, and it is a vital parameter in the proposed method. Therefore, the stability of the proposed method in terms of α is illustrated in Fig. 14. It can be seen that when α varies from 0.5 to 2 with step 0.1, the deviation of GMSE of abundance is about 10^{-3} , which fully proves that the proposed method is relatively robust and stable to the choice of parameters.

2) *Time Complexity Analysis*: Finally, the time consumption is given in Table IV both on the stimulated and real data sets. It spends much time in stimulated experiment than the real one due to the large size of the SHS image. In addition, with the number of the endmembers increasing, the computational cost of the Bayesian method increases greatly. Therefore, the mixture of six endmembers takes highest time computation. VCA and SISAL are faster than other algorithms, and PLMM and

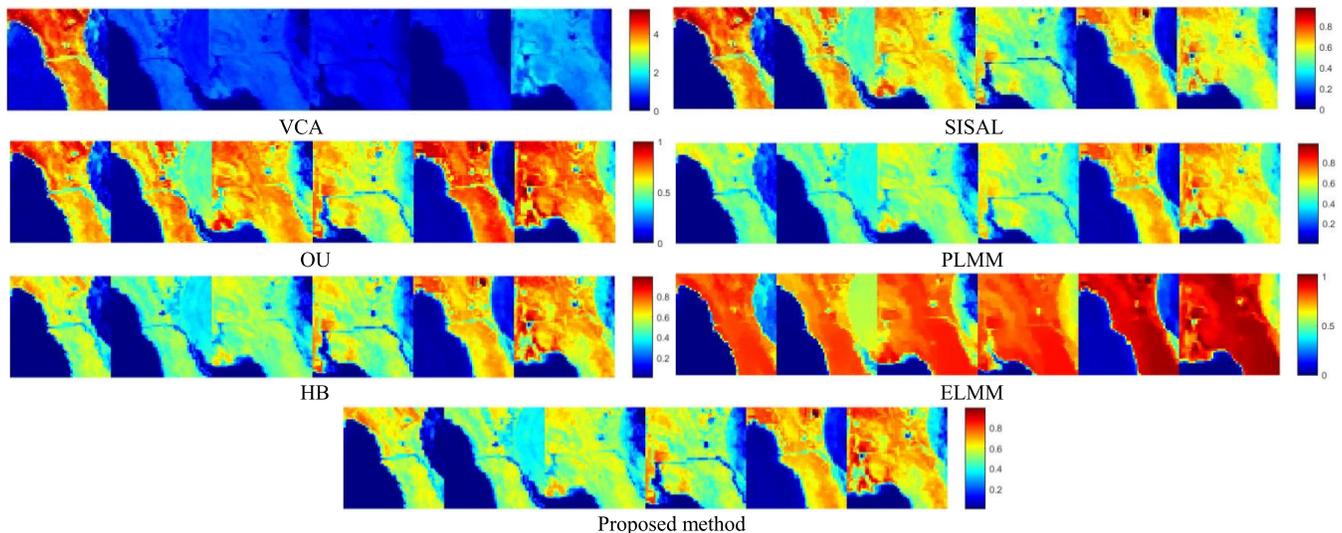


Fig. 13. Change in the third endmember's (*Soil*) abundance map at different time extracted from real data.

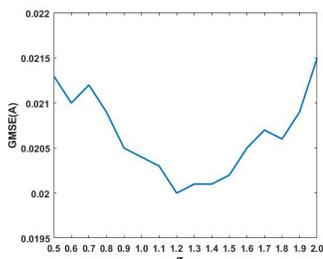


Fig. 14. Sensitivity analysis of parameter α .

TABLE IV
TIME COMPLEXITY (M)

Methods	Synthetic data			Real data
	$R=4$	$R=5$	$R=6$	
VCA/FCLS	0.28	0.3	0.32	0.013
SISAL/FCLS	<u>0.9</u>	<u>0.93</u>	<u>1.0</u>	<u>0.025</u>
OU	12.85	14.87	18.42	1.033
PLMM	333.57	431.23	520.77	15.53
HB	687.23	909.95	1055.92	23.72
ELMM	127.22	153.72	181.57	20.28
Proposed method	755.0	921.63	1104.05	24.11

ELMM have more time computation due to the iteration of the energy function minimization. OU has an obviously efficient improvement compared with PLMM. High time complexity is one of the disadvantages of the Bayesian-based methods due to the limitation of marginalizing over the parameter space, which leads to a long time to burn-in for a Markov Chain. Therefore, HB and the proposed methods take more time to converge to the stable solution. So how to obtain efficient and fast algorithm for the Bayesian-based method is an open and hot topic to be studied further.

Form Tables II–IV, although RE, GMSE of endmember variability, and the time complexity of the proposed method are inferior to some of the compared methods, the average indexes of the method achieve best, especially on endmember and abundance. Moreover, the visual effects also demonstrated

the advantages of the proposed method on estimating the abundance and endmember variability. In general, despite the issues that some of the indexes of the proposed method are not optimal, and the high cost of the computation, the proposed method still provides competitive unmixing results. Owing to the composite priors of abundance and endmember variability, the proposed method exhibits the strong ability of discerning the pure endmembers exactly, distinguishing the endmember variations robustly, and estimating the abundance maps accurately.

VI. CONCLUSION

In this article, priors of parameters and hyperparameter in unmixing of HS image sequence are established according to the HB unmixing model. To approximate abundances accurately, a new mixed prior distribution is proposed, which takes account of temporal continuity and spatial smoothness of abundance. For endmember variability, a mixed prior distribution considering temporal continuity and spectral smoothness is proposed. The MCMC algorithm is proposed to sample the posteriori of the proposed HB and generalized sampling is used to establish the estimators of the latent model parameters. The experimental results show that the proposed method can obtain accurate abundance estimation, as well as robust endmember variability, which in turn improves the accuracy of the endmember. The robustness of the proposed method has applications to change and anomaly detection, also interesting prospects for accounting for spatial variability. Moreover, a fast and efficient unmixing procedure is also under consideration.

REFERENCES

- [1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [2] Y. Xu, Z. Wu, J. Chanussot, and Z. Wei, "Nonlocal patch tensor sparse representation for hyperspectral image super-resolution," *IEEE Trans. Image Process.*, vol. 28, no. 6, pp. 3034–3047, Jun. 2019.

- [3] H. Guo, Q. Shi, B. Du, L. Zhang, D. Wang, and H. Ding, "Scene-driven multitask parallel attention network for building extraction in high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, early access, Aug. 19, 2020, doi: [10.1109/TGRS.2020.3014312](https://doi.org/10.1109/TGRS.2020.3014312).
- [4] Q. Shi *et al.*, "Domain adaption for fine-grained urban village extraction from satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1430–1434, Aug. 2020.
- [5] W. Sun, J. Peng, G. Yang, and Q. Du, "Fast and latent low-rank subspace clustering for hyperspectral band selection," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 3906–3915, Jun. 2020.
- [6] J. M. Bioucas-Dias and A. Plaza, "An overview on hyperspectral unmixing: Geometrical, statistical, and sparse regression based approaches," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2011, pp. 1135–1138.
- [7] F. Xiong, Y. Qian, J. Zhou, and Y. Y. Tang, "Hyperspectral unmixing via total variation regularized nonnegative tensor factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2341–2357, Apr. 2019.
- [8] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Jun. 2010, pp. 1–4.
- [9] W. Sun, J. Ma, G. Yang, B. Du, and L. Zhang, "A Poisson nonnegative matrix factorization method with parameter subspace clustering constraint for endmember extraction in hyperspectral imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 128, pp. 27–39, Jun. 2017.
- [10] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.
- [11] 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.
- [12] 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.
- [13] 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.
- [14] 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.
- [15] J. Peng, Y. Zhou, W. Sun, Q. Du, and L. Xia, "Self-paced nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1501–1515, Feb. 2021.
- [16] H. Liu, P. Sun, Q. Du, Z. Wu, and Z. Wei, "Hyperspectral image restoration based on low-rank recovery with a local neighborhood weighted spectral-spatial total variation model," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1409–1422, Mar. 2019.
- [17] Y. Yuan, Z. Zhang, and Q. Wang, "Improved collaborative non-negative matrix factorization and total variation for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 4, pp. 998–1010, Mar. 2020.
- [18] J.-J. Wang, T.-Z. Huang, J. Huang, H.-X. Dou, L.-J. Deng, and X.-L. Zhao, "Row-sparsity spectral unmixing via total variation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5009–5022, Dec. 2019.
- [19] B. Du, M. Zhang, L. Zhang, R. Hu, and D. Tao, "PLTD: Patch-based low-rank tensor decomposition for hyperspectral images," *IEEE Trans. Multimedia*, vol. 19, no. 1, pp. 67–79, Jan. 2017.
- [20] L. Drumetz, J. Chanussot, and C. Jutten, "Variability of the endmembers in spectral unmixing: Recent advances," in *Proc. 8th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens.*, Aug. 2017, pp. 1–5.
- [21] A. Zare and K. C. Ho, "Endmember variability in hyperspectral imagery analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.
- [22] R. A. Borsoi *et al.*, "Spectral variability in hyperspectral data unmixing: A comprehensive review," 2020, [arXiv:2001.07307](https://arxiv.org/abs/2001.07307). [Online]. Available: <http://arxiv.org/abs/2001.07307>
- [23] D. A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, "Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, Sep. 1998.
- [24] T. R. Meyer, L. Drumetz, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral unmixing with material variability using social sparsity," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2187–2191.
- [25] B. Somers, M. Zortea, A. Plaza, and G. P. Asner, "Automated extraction of image-based endmember bundles for improved spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 396–408, Apr. 2012.
- [26] L. Drumetz, T. R. Meyer, J. Chanussot, A. L. Bertozzi, and C. Jutten, "Hyperspectral image unmixing with endmember bundles and group sparsity inducing mixed norms," *IEEE Trans. Image Process.*, vol. 28, no. 7, pp. 3435–3450, Jul. 2019.
- [27] T. Uezato, R. J. Murphy, A. Melkumyan, and A. Chlingaryan, "A novel spectral unmixing method incorporating spectral variability within endmember classes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 2812–2831, May 2016.
- [28] R. Liu and X. Zhu, "Endmember bundle extraction based on multi-objective optimization," *IEEE Trans. Geosci. Remote Sens.*, early access, Nov. 24, 2020, doi: [10.1109/TGRS.2020.3037249](https://doi.org/10.1109/TGRS.2020.3037249).
- [29] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [30] L. Drumetz, M.-A. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Aug. 2016.
- [31] M. A. Veganzones *et al.*, "A new extended linear mixing model to address spectral variability," in *Proc. 6th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Lausanne, Switzerland, Jun. 2014, pp. 1–4.
- [32] Z. G. Gao and L. Q. Zhang, "Multi-seasonal spectral characteristics analysis of coastal salt marsh vegetation in Shanghai, China," *Estuarine, Coastal Shelf Sci.*, vol. 69, nos. 1–2, pp. 217–224, Aug. 2006.
- [33] P. Lukes, P. Stenberg, M. Rautiainen, M. Mõttus, and K. M. Vanhatalo, "Optical properties of leaves and needles for boreal tree species in Europe," *Remote Sens. Lett.*, vol. 4, no. 7, pp. 667–676, Jul. 2013.
- [34] L. Drumetz, J. Chanussot, C. Jutten, W.-K. Ma, and A. Iwasaki, "Spectral variability aware blind hyperspectral image unmixing based on convex geometry," *IEEE Trans. Image Process.*, vol. 29, pp. 4568–4582, 2020.
- [35] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.
- [36] T. Imbiriba, R. A. Borsoi, and J. C. M. Bermudez, "Generalized linear mixing model accounting for endmember variability," 2017, [arXiv:1710.07723](https://arxiv.org/abs/1710.07723). [Online]. Available: <http://arxiv.org/abs/1710.07723>
- [37] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Jan. 2016.
- [38] A. Zare, P. Gader, and G. Casella, "Sampling piecewise convex unmixing and endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 3, pp. 1655–1665, Mar. 2013.
- [39] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tourneret, "Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1403–1413, Jun. 2010.
- [40] X. Du, A. Zare, P. Gader, and D. Dranishnikov, "Spatial and spectral unmixing using the beta compositional model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1994–2003, Jun. 2014.
- [41] Y. Zhou, A. Rangarian, and P. D. Gader, "A Gaussian mixture model representation of endmember variability in hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2242–2256, May 2018.
- [42] G. Tochon, D. Pauwels, M. D. Mura, and J. Chanussot, "Unmixing-based gas plume tracking in LWIR hyperspectral video sequences," in *Proc. 8th Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Aug. 2016, pp. 1–5.
- [43] Y. Xu, Z. Wu, J. Chanussot, M. D. Mura, A. L. Bertozzi, and Z. Wei, "Low-rank decomposition and total variation regularization of hyperspectral video sequences," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1680–1694, Mar. 2018.
- [44] S. Henrot, J. Chanussot, and C. Jutten, "Dynamical spectral unmixing of multitemporal hyperspectral images," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3219–3232, Jul. 2016.
- [45] M.-D. Iordache, L. Tits, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "A dynamic unmixing framework for plant production system monitoring," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2016–2034, Jun. 2014.
- [46] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "Online unmixing of multitemporal hyperspectral images accounting for spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 9, pp. 3979–3990, Sep. 2016.

- [47] P.-A. Thouvenin, N. Dobigeon, and J.-Y. Tourneret, "A hierarchical Bayesian model accounting for endmember variability and abrupt spectral changes to unmix multitemporal hyperspectral images," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 32–45, Mar. 2018.
- [48] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.
- [49] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *J. ACM*, vol. 58, no. 3, pp. 1–37, 2009.
- [50] Y. Altmann, S. McLaughlin, and N. Dobigeon, "Sampling from a multivariate Gaussian distribution truncated on a simplex: A review," in *Proc. 17th Workshop Statist. Signal Process.*, Jun. 2014, pp. 113–116.
- [51] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a Hamiltonian Monte Carlo algorithm," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2663–2675, Jun. 2014.
- [52] A. Pakman and L. Paninski, "Exact Hamiltonian Monte Carlo for truncated multivariate Gaussians," *J. Comput. Graph. Statist.*, vol. 23, no. 2, pp. 518–542, Apr. 2014.



Hongyi Liu (Member, IEEE) received the M.S. degree in computational mathematics from Northwest University, Xi'an, China, in 2003 and the Ph.D. degree in pattern recognition and artificial intelligence from the Nanjing University of Science and Technology, Nanjing, China, in 2011.

She is an Associate Professor with the School of Science, Nanjing University of Science and Technology. Her research interests include partial differential equation-based image processing, multiscale analysis, sparse representation, and hyperspectral image processing.



Youkang Lu (Student Member, IEEE) was born in Anhui, China, in 1995. He received the B.Sc. degree in applied statistics from the School of Science, Nanjing University of Science and Technology, Nanjing, China, in 2021, where he is pursuing the M.S. degree.

His research interests include hyperspectral image unmixing.



Zebin Wu (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees in computer science and technology from the Nanjing University of Science and Technology, Nanjing, China, in 2003 and 2007, respectively.

He was a Visiting Scholar with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain, from June 2014 to June 2015. He was a Visiting Scholar with the Department of Mathematics, University of California, Los Angeles, CA, USA, from August 2016 to September 2016, and from July 2017 to August 2017. He was a Visiting Scholar with the GIPSA-Lab, Grenoble INP, the Univ. Grenoble Alpes, Grenoble, France, from August 2018 to September 2018. He is a Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology. He has authored over 76 publications, including 30 Journal Citation Reports (JCR) journal papers (22 in IEEE journals), and over 31 peer-reviewed conference proceeding papers (25 in IEEE conferences). He has reviewed more than 60 articles for over 15 different journals. His research interests include hyperspectral image processing, parallel computing, and remotely sensed big data processing.

Dr. Wu served as the Vice Chair for IEEE Geoscience and Remote Sensing Society (GRSS) Nanjing Chapter during 2016–2020. He received the Best Reviewers of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (IEEE JSTARS). He serves as an Associate Editor of IEEE JSTARS and the Chair of IEEE GRSS Nanjing Chapter.



Qian Du (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of Maryland–Baltimore County, Baltimore, MD, USA, in 2000.

She is a Bobby Shackouls Professor with the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. Her research interests include hyperspectral remote sensing image analysis and applications, pattern classification, data compression, and neural networks.

Dr. Du is a Fellow of the SPIE—International Society for Optics and Photonics. She was a recipient of the 2010 Best Reviewer Award from the IEEE Geoscience and Remote Sensing Society (GRSS). She was a Co-Chair of the Data Fusion Technical Committee of the IEEE GRSS from 2009 to 2013. She was the Chair of the Remote Sensing and Mapping Technical Committee of the International Association for Pattern Recognition from 2010 to 2014. She was the General Chair of the fourth IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing held at Shanghai, China, in 2012. She served as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (JSTARS), the *Journal of Applied Remote Sensing*, and the IEEE SIGNAL PROCESSING LETTERS. She was the Editor-in-Chief of the IEEE JSTARS in 2016–2020. She serves for the IEEE Technical Activities Board (TAB) Periodicals Review and Advisory Committee (PRAC) and SPIE Publications Committee.



Jocelyn Chanussot (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995 and the Ph.D. degree from the Université de Savoie, Annecy, France, in 1998.

Since 1999, he has been with Grenoble INP, where he is a Professor of signal and image processing. He has been a Visiting Scholar at Stanford University, Stanford, CA, USA, KTH Royal Institute of Technology, Stockholm, Sweden and National University of Singapore, Singapore. Since 2013, he is an Adjunct Professor with the University of Iceland, Reykjavik, Iceland. In 2015–2017, he was a Visiting Professor with the University of California, Los Angeles (UCLA), Los Angeles, CA, USA. He holds the AXA Chair in remote sensing and is an Adjunct Professor with the Chinese Academy of Sciences, Aerospace Information research Institute, Beijing, China. His research interests include image analysis, hyperspectral remote sensing, data fusion, machine learning, and artificial intelligence.

Dr. Chanussot is the Founding President of IEEE Geoscience and Remote Sensing French chapter (2007–2010) which received the 2010 IEEE GRSS Chapter Excellence Award. He has received multiple outstanding paper awards. He was the Vice-President of the IEEE Geoscience and Remote Sensing Society, in charge of meetings and symposia (2017–2019). He was the General Chair of the first IEEE Geoscience and Remote Sensing Society (GRSS) Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009–2011) and Co-Chair of the GRS Data Fusion Technical Committee (2005–2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE TRANSACTIONS ON IMAGE PROCESSING and the PROCEEDINGS OF THE IEEE. He was the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING (2011–2015). In 2014, he served as a Guest Editor for the *IEEE Signal Processing Magazine*. He is a member of the Institut Universitaire de France (2012–2017) and a Highly Cited Researcher (Clarivate Analytics/Thomson Reuters, since 2018).



Zhihui Wei was born in Jiangsu, China, in 1963. He received the B.Sc. and M.Sc. degrees in applied mathematics and the Ph.D. degree in communication and information system from South East University, Nanjing, China, in 1983, 1986, and 2003, respectively.

He is a Professor and a Doctoral Supervisor with the Nanjing University of Science and Technology (NUST), Nanjing. His research interests include inverse problems in image processing, machine learning, intelligent information processing, and scientific computing.