DAEN: Deep Autoencoder Networks for Hyperspectral Unmixing

Yuanchao Su, *Student Member, IEEE*, Jun Li, *Senior Member, IEEE*, Antonio Plaza, *Fellow, IEEE*, Andrea Marinoni, *Senior Member, IEEE*, Paolo Gamba, *Fellow, IEEE*, and Somdatta Chakravortty

***Abstract*— Spectral unmixing is a technique for remotely sensed image interpretation that expresses each (possibly mixed) pixel as a combination of pure spectral signatures (endmembers) and their fractional abundances. In this paper, we develop a new technique for unsupervised unmixing which is based on a deep autoencoder network (DAEN). Our newly developed DAEN consists of two parts. The first part of the network adopts stacked autoencoders (SAEs) to learn spectral signatures, so as to generate a good initialization for the unmixing process. In the second part of the network, a variational autoencoder (VAE) is employed to perform blind source separation, aimed at obtaining the endmember signatures and abundance fractions simultaneously. By taking advantage from the SAEs, the robustness of the proposed approach is remarkable as it can unmix data sets with outliers and low signal-to-noise ratio. Moreover, the multihidden layers of the VAE ensure the required constraints (nonnegativity and sum-to-one) when estimating the abundances. The effectiveness of the proposed method is evaluated using both synthetic and real hyperspectral data. When compared with other unmixing methods, the proposed approach demonstrates very competitive performance.**

***Index Terms*— Deep autoencoder network (DAEN), deep learning, endmember identification, hyperspectral unmixing, variational autoencoder (VAE).**

Manuscript received July 18, 2018; revised November 4, 2018, December 19, 2018 and December 29, 2018; accepted December 29, 2018. Date of publication January 28, 2019; date of current version June 24, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61771496, in part by the Guangdong Provincial Natural Science Foundation under Grant 2016A030313254, and in part by the National Key Research and Development Program of China under Grant 2017YFB0502900. *(Corresponding author: Jun Li.)*

Y. Su is with the Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: suych3@mail2.sysu.edu.cn).

J. Li is with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China, and also with the Key Laboratory of Visual Perception and Artificial Intelligence of Hunan Province, Changsha 410082, China (e-mail: lijun48@mail.sysu.edu.cn).

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, 10071 Cáceres, Spain (e-mail: aplaza@unex.es).

A. Marinoni is with the Earth Observation Group, Centre for Integrated Remote Sensing and Forecasting for Arctic Operations, Department of Physics and Technology, UiT–The Arctic University of Norway, 9037 Tromsø, Norway (e-mail: andrea.marinoni@unipv.it).

P. Gamba is with the Telecommunications and Remote Sensing Laboratory, Department of Electrical, Computer and Biomedical Engineering, University of Pavia, 27100 Pavia, Italy (e-mail: paolo.gamba@unipv.it).

S. Chakravortty is with the Department of Information Technology, Maulana Abul Kalam Azad University of Technology, Kolkata 700064, India (e-mail: csomdatta@rediffmail.com).

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

Digital Object Identifier 10.1109/TGRS.2018.2890633

I. Introduction

Hyperspectral images have been extensively used in a wide variety of applications, including classification, target detection, environmental monitoring, precision agriculture, and so on [1]–[3]. However, due to the relatively low spatial resolution of hyperspectral images, many pixels are mixed by several materials, which brings difficulties to the characterization of hyperspectral data, leading to inaccuracies in the understanding and quantification of the considered scenes [4], [5].

In order to deal with mixed pixels, many algorithms have been proposed for hyperspectral unmixing [1]. Hyperspectral unmixing refers to any process that separates the pixel spectra from a hyperspectral image into a collection of pure constituent spectra, called endmembers, and a set of abundance fractions. At each pixel, the endmembers are generally assumed to represent the pure materials, and the abundances represent the percentage of each endmember [6].

Unmixing algorithms rely on specific mixing models, which can be characterized as either linear or nonlinear [7]. Linear spectral mixing (LSM) holds when the mixing scale is macroscopic and the photon reaching the sensor has interacted with just one material [8]. In contrast, non-LSM (NLSM) considers physical interactions between the light scattered by multiple materials in the scene. These interactions may happen at a multilayered level or at a microscopic (or intimate) level [9], [10]. In recent years, a new trend is to incorporate linear and nonlinear models for unmixing. For instance, in [11] and [12], nonlinearity is regarded as a penalty term of a linear objective function. In [13], the hyperspectral data are projected on hyperplanes representing the nonlinearities. Although NLSM is more faithful to the imaging mechanism of instrument, there are many challenges remaining. For instance, retrieving a reliable estimate of the actual areal proportions of the endmembers in the NLSM requires the use of complex metrics, which needs to consider the prior knowledge of the scene [1]. Furthermore, the high computational cost is beyond the reach of many techniques [1]. Therefore, most unmixing algorithms belong to the LSM category. Despite its simplicity, it provides very good approximations of the light scattering mechanisms in many real scenarios [14].

Regardless of the availability of training samples, most unmixing methods can be classified into two categories, i.e., handcrafted and learning based. In this paper, we use handcrafted to denote traditional algorithms, which derive

estimations from the image itself without learning. Handcrafted approaches are subdivided into pure pixel based and nonpure pixel based. Pure pixel-based methods, such as N-FINDR [15] and vertex component analysis (VCA) [16], assume that pure pixels made up of a single endmember are present in the observed scene. These methods often project and impose an orthogonality condition onto endmembers in their estimation [14]. However, in real scenarios, the pure pixel-based assumption may not hold for all endmembers. This is particularly the case when the data have relatively low spatial resolution [17]. In this context, the pure pixel assumption may collapse with the lack of purity of one or more materials in data. In order to overcome this problem, nonpure pixel-based methods such as minimum volume-constrained nonnegative matrix factorization (MVCNMF) [18], piecewise convex multiple-model endmember detection (PCOMMEND) [19], minimum volume simplex analysis [20], robust collaborative nonnegative matrix factorization [21], and others [22]–[25] focus on exploiting the data structure via some geometrical and statistical assumptions. Handcrafted methods, as the main trend in the past two decades, have shown to be very effective in the unmixing of hyperspectral data, for scenarios of outliers [26], [27], noisy conditions [28], highly mixed cases [29], [30], and so on.

With advances in computer technology, learning-based approaches for unmixing have achieved a fast development in the past few years. Joint Bayesian unmixing is a typical example of learning-based approaches, which leads to good abundance estimates due to the incorporation of full additivity (i.e., sum-to-one) and nonnegativity constraints [31], [32]. However, the randomness of sample probabilities may lead to random local estimations for the extracted endmembers, which may not properly represent the global characteristics of the scene [33]. On the other hand, approaches based on artificial neural networks have also been developed for the learning of abundance fractions, assuming the prior knowledge of the endmember signatures [34]–[36]. These approaches exhibit better performance when compared with handcrafted methods, but they assume that endmembers are known in advance and, therefore, need to incorporate endmember extraction algorithms to perform unmixing. More recently, as a common tool for deep learning, autoencoders have achieved a fast development in unmixing applications. Nonnegative sparse autoencoder and denoising autoencoder were employed to obtain the endmember signatures and abundance fractions simultaneously for unmixing, with advanced denoising and intrinsic self-adaptation capabilities [37]–[39]. Sparse autoencoder [40] and multiple hidden layer autoencoder [41], [42] are further employed for hyperspectral unmixing, demonstrating a good potential of autoencoders in the tasks of unmixing. However, their strength is in the aspect of noise reduction and they exhibit limitations when dealing with outliers. Due to the fact that outliers likely lead to initialization problems, their presence can bring strong interference to the unmixing solutions. In [43], we proposed a stacked nonnegative sparse autoencoder (SNSA) to address the issue of outliers. For LSM-based hyperspectral unmixing, the physical meaning of the model implies the sum-to-one on abundance fractions when every material in a pixel can be identified [14], [39], [44]. However, similar to the NMF-based approaches, SNSA adopts an additivity penalty on the abundance coefficients. The additivity penalty denotes that a penalty coefficient is used for controlling approximation of the sum-to-one. As this is not a hard constraint, the sum-to-one constraint is not necessarily ensured [43].

In this paper, we develop a new deep autoencoder network (DAEN) for hyperspectral unmixing, which specifically addresses the presence of outliers in hyperspectral data. The proposed DAEN has two main steps. In the first step, we learn the spectral signatures via the stacked autoencoders (SAEs), aiming at generating good initializations for the network. In the second step, we employ a variational autoencoder (VAE) to perform unmixing for the estimation of the endmembers and abundances. VAE combines variational inference to perform unsupervised learning and inherit autoencoder architecture which can be trained with gradient descent [45], [46]. Different from conventional autoencoders, VAEs include a reparameterization which strictly ensures the abundance sum-to-one constraint during unmixing. Compared with other NMF-based algorithms, our newly proposed DAEN has three main advantages: 1) with the use of SAEs, we can effectively tackle the problem of outliers and generate a good initialization of the unmixing network; 2) with the adoption of a VAE, we can ensure the nonnegativity and sum-to-one constraints, resulting in the good performance on abundance estimation; and 3) the endmember signatures and abundance fractions are obtained simultaneously. Finally, it should be noted that the basic concept of DAEN was introduced in [47]. In comparison, this paper presents an optimized scheme for DAEN in detail. Specifically, the SAEs are used for initialization in this paper, while in [47], it was for outlier detection. Comprehensive validation by using different simulated and real hyperspectral data is presented in this paper.

The remainder of this paper is organized as follows. Section II describes the proposed DAEN approach. In Section III, synthetic data sets are used for evaluation. This allows us to conduct a quantitative comparison with other methods. In Section IV, experiments with two real hyperspectral data sets are conducted for further validation. Finally, Section V concludes this paper with some remarks and hints at plausible future research lines.

II. NMF-Based Unmixing

Let $\mathbf{Y} \equiv [\mathbf{y}_1, \ldots, \mathbf{y}_n] \in \mathbb{R}^{d \times n}$ be matrix representation of a hyperspectral data set with n spectral vectors and d spectral bands. Under the LSM, we have [14], [48]

$$\begin{aligned} &\mathbf{Y} = \mathbf{W}\mathbf{H} + \mathbf{N} \\ &\text{s.t. } \mathbf{H} \geq 0, \quad \mathbf{1}_m^T \mathbf{H} = \mathbf{1}_n^T \end{aligned} \tag{1}$$

where $\mathbf{W} \equiv [\mathbf{w}_1, \ldots, \mathbf{w}_m] \in \mathbb{R}^{d \times m}$ is the mixing matrix containing m endmembers, $\mathbf{w}_i$ denotes the ith endmember, $\mathbf{H} \equiv [\mathbf{h}_1, \ldots, \mathbf{h}_n] \in \mathbb{R}^{m \times n}$ is the abundance matrix, $\mathbf{H} \geq 0$ and $\mathbf{1}_m^T \mathbf{H} = \mathbf{1}_n^T$ are the so-called abundance nonnegativity and sum-to-one constraints, which stem from a physical interpretation of the abundance vectors, and $\mathbf{1}_m = [1, 1, \ldots, 1]^T$ is a column

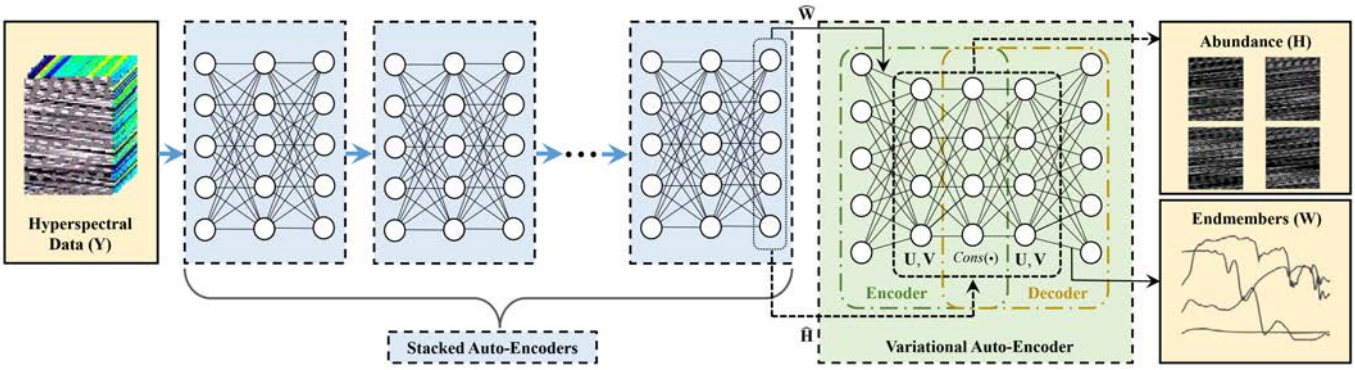

Fig. 1. Flowchart of the proposed DAEN, which includes two parts, i.e., SAEs and a VAE. The SAEs generate the initializations $\widehat{\mathbf{W}}$ and $\widehat{\mathbf{H}}$ for the VAE, while the VAE performs the NMF-based unmixing aiming at obtaining the endmembers $\mathbf{W}$ and abundances $\mathbf{H}$, respectively.

vector of size m (the notation $[\cdot]^T$ stands for vector or matrix transpose). Finally, $\mathbf{N} \in \mathbb{R}^{d\times n}$ is the error matrix that may affect the measurement process (e.g., noise).

For a given observation $\mathbf{Y}$, unmixing aims at obtaining the mixing matrix $\mathbf{W}$ and the abundance matrix $\mathbf{H}$. In this paper, we tackle the simultaneous estimation of $\mathbf{W}$ and $\mathbf{H}$ by seeking a solution with the following NMF-based optimization:

$$(\mathbf{W},\mathbf{H}) = \arg\min_{\mathbf{W},\mathbf{H}} \frac{1}{2}\|\mathbf{Y}-\mathbf{W}\mathbf{H}\|_F^2 + \mu\ f_1(\mathbf{W}) + \lambda f_2(\mathbf{H}) \quad (2)$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm, $f_1(\mathbf{W})$ and $f_2(\mathbf{H})$ are two regularizers on the mixing matrix $\mathbf{W}$ and the abundance fractions $\mathbf{H}$, respectively, with μ and λ being the regularization parameters.

III. Deep Autoencoder Network

In this section, we describe the proposed DAEN for unmixing (illustrated in Fig. 1), where $\mathbf{U}$ and $\mathbf{V}$ are the latent variables (LV) of the reparameterization of the VAE, respectively. As shown in Fig. 1, the endmember matrix $\mathbf{W}$ corresponds to the last weight matrix of the decoder in VAE, and the abundance $\mathbf{H}$ is estimated from the hidden layers of VAE, while $\widehat{\mathbf{W}}$ and $\widehat{\mathbf{H}}$ denote the initializations for VAE generated by SAEs, respectively. This section is organized as follows. In Section III-A, we first revisit the SAEs for initialization. Then, in Section III-B, we introduce the proposed NMF-VAE-based unmixing scheme.

A. SAEs for Initialization

Based on the geometry assumption that endmembers are generally located around the vertices of the data simplex, we use a pure pixel-based method to extract a set of candidate pixels as the training set for the SAEs. Specifically, we adopt VCA to obtain a set of k candidates, with $k > m$. As VCA considers random directions in the subspace projection [14], [16], we run it for p times, resulting in q candidates, with $q = p \cdot k$. These q candidates are then grouped into m training sets $\{\mathbf{C}_i\}_{i=1}^m$ based on the spectral angle distance (SAD) and clustering, with $\mathbf{C}_i = [\mathbf{c}_1,\ldots,\mathbf{c}_{i_n}] \in \mathbb{R}^{d\times i_n}$ and i_n is the number of samples in $\mathbf{C}_i$. Let $\mathbf{c}_{i_o}$ and $\mathbf{c}_{j_o}$ be the cluster centers of $\mathbf{C}_i$ and $\mathbf{C}_j$, respectively. For any candidate $\mathbf{c}_{i_s}$ in $\mathbf{C}_i$, for $i_s = 1,\ldots,i_n$, we have $\mathrm{SAD}(\mathbf{c}_{i_o},\mathbf{c}_{i_s}) \le \mathrm{SAD}(\mathbf{c}_{j_o},\mathbf{c}_{i_s})$, for any $j = 1,\ldots,m$ and $j \neq i$, where

$$\mathrm{SAD}(\mathbf{c}_{i_o},\mathbf{c}_{i_s}) = \arccos\left(\frac{[\mathbf{c}_{i_o},\mathbf{c}_{i_s}]}{\|\mathbf{c}_{i_o}\|\cdot\|\mathbf{c}_{i_s}\|}\right). \quad (3)$$

In this paper, for p and k, we empirically set $p = 30$ and $k = 3m$, respectively. By enforcing nonnegativity, the training of SAEs minimizes the reconstruction error (RE) as follows:

$$\min \sum_{s=1}^{i_n} \|\mathbf{c}_s - \widehat{\mathbf{w}}_i\|_2^2 \quad (4)$$

where $\widehat{\mathbf{w}}_i$ is the reconstructed signature of the ith endmember, and $\widehat{\mathbf{W}} = [\widehat{\mathbf{w}}_1,\ldots,\widehat{\mathbf{w}}_m]$ are the reconstructed endmember matrix. Following [49], the reconstructed signature is denoted as

$$\widehat{\mathbf{w}}_i = \mathbf{M}_i\mathbf{f}\left(\mathbf{M}_i^T\mathbf{C}_i\right) \quad (5)$$

where $\mathbf{M}_i$ is the matrix of weights between the input and hidden neurons or those from hidden to output neurons, and $\mathbf{f}(\cdot)$ is the activation function [49] given by

$$\mathbf{f}(\mathbf{g}_i) = \frac{1}{1+\exp(-\mathbf{a}_i .* \mathbf{g}_i - \mathbf{b}_i)} \quad (6)$$

where $\mathbf{g}_i = \mathbf{M}_i^T\mathbf{C}_i$, $\mathbf{a}_i$ and $\mathbf{b}_i$ are the parameters aimed at controlling the information transmission between neurons, and $.*$ is the dot product, i.e., elementwise operator. Note that the number of input neurons and output neurons is the same as the hidden neurons, while the number of hidden neurons here is set as the number of bands. Then, we can use a gradient rule to update $\mathbf{a}_i$ and $\mathbf{b}_i$ as follows:

$$\begin{cases} \Delta\mathbf{a}_i = \gamma\left(1-\left(2+\dfrac{1}{\tau}\right)\mathbf{f}_i + \dfrac{1}{\tau}\mathbf{f}_i^2\right) \\ \Delta\mathbf{b}_i = \gamma\dfrac{1}{\mathbf{b}_i} + \mathbf{g}_i\Delta\mathbf{a}_i \end{cases} \quad (7)$$

where γ and τ are hyperparameters in the learning process controlling the mean activity level of the desired output distribution. Following the empirical settings in [49], we set $\gamma = 0.0001$ and $\tau = 0.2$. With the aforementioned definition in hand, the learning reduces to the following update rule:

$$\Delta\mathbf{M}_i \Leftarrow \eta\Delta\widehat{\mathbf{w}}_i\mathbf{f}_i^T + |\mathbf{M}_i| \quad (8)$$

where $\Delta\widehat{\mathbf{w}}_i$ is the gradient of candidate i for update, $|\mathbf{M}_i|$ enforces the weight matrix to be nonnegative, and η is an adaptive learning rate. In this paper, following [49], we set $\eta = \hat{\eta}(\|\mathbf{f}_i\|^2 + \epsilon)^{-1}$ with $\hat{\eta} = 0.002$, where $\epsilon = 0.001$ is a small parameter to ensure the positivity of η.

Finally, let $\widehat{\mathbf{w}}_i^t$, $\widehat{\mathbf{w}}_i^{t+1}$ be the reconstructions from the tth and $(t+1)$th autoencoders, respectively. The SAEs ends when $\|\widehat{\mathbf{w}}_i^{t+1} - \widehat{\mathbf{w}}_i^t\|_2^2$ converges.

After the endmember matrix $\widehat{\mathbf{W}}$ is reconstructed, based on the linear mixing model (1), the abundances $\widehat{\mathbf{H}}$ can be obtained via the fully constrained least square (FCLS) [48]. In the learning of the VAE, $\widehat{\mathbf{W}}$ and $\widehat{\mathbf{H}}$ are used as initializations of $\mathbf{W}$ and $\mathbf{H}$, respectively.

B. VAE for Unmixing

First, let us recall the NMF-based objective function in (2), which contains two regularizers on the mixing matrix and abundance matrix, respectively. For the first regularizer $f_1(\mathbf{W})$ on the mixing matrix, following [20], we have

$$f_1(\mathbf{W}) = \text{MinVol}(\mathbf{W}) \tag{9}$$

where MinVol$(\cdot)$ is a function aiming at enclosing all the pixels into the simplex constructed by the endmembers. Specifically, following [20], we set MinVol$(\mathbf{W}) = \|\det(\mathbf{W})\|$, with $\|\det(\mathbf{W})\|$ being the volume defined by the origin and the columns of $\mathbf{W}$.

With respect to regularizer $f_2(\mathbf{H})$ on the abundance matrix, in order to ensure the nonnegativity and sum-to-one constraints, we employ the VAE to penalize the solution of $\mathbf{H}$, denoted as

$$f_2(\mathbf{H}) = \text{VAE}(\mathbf{H}) \tag{10}$$

where the neurons of all hidden layers are set as the number of endmembers, while the number of inputs and outputs corresponds to the number of pixels.

With these definitions in mind, we obtain the following objective function:

$$(\mathbf{W}, \mathbf{H}) = \arg\min_{\mathbf{W},\mathbf{H}} \quad \frac{1}{2}\|\mathbf{Y} - \mathbf{W}\mathbf{H}\|_F^2 + \mu \text{ MinVol } (\mathbf{W}) + \lambda \text{VAE}(\mathbf{H}). \tag{11}$$

In the following, we present the VAE-based regularizer in detail. Let $\mathbf{U}$ and $\mathbf{V}$ be the LV, we define $f_2(\mathbf{H})$ as

$$f_2(\mathbf{H}(\mathbf{U}, \mathbf{V})) = \left\| \frac{1}{2n}(\mathbf{1}_{m\times n} + \ln \mathbf{V}^2 - \mathbf{U}^2 - \mathbf{V}^2)\mathbf{1}_n \right\|_2^2 \tag{12}$$

where $\mathbf{1}_{m\times n} \in \mathbb{R}^{m\times n}$ with all elements being 1, and vector $\mathbf{1}_n = [1, \ldots, 1]^T \in \mathbb{R}^n$, $\mathbf{U} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\} \in \mathbb{R}^{m\times n}$, $\mathbf{V} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\} \in \mathbb{R}^{m\times n}$. The derivation of (12) is shown in Appendix A. Following [45], let $\mathbf{u}_j = [u_{1,j}, \ldots, u_{m,j}]^T \in \mathbb{R}^m$ and $\mathbf{v}_j = [v_{1,j}, \ldots, v_{m,j}]^T \in \mathbb{R}^m$ be the reparameters of LV, we define $h_{i,j} = \text{Cons}(u_{i,j}, v_{i,j})$, where Cons$(\cdot)$ represents a decay function as follows:

$$\text{Cons}(u_{i,j}, v_{i,j}) = \begin{cases} u_{i,j} + \sigma v_{i,j}, & 0 < (u_{i,j} + \sigma v_{i,j}) < 1 \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

where σ is a parameter that, as indicated in [31], can be obtained via Monte Carlo (MC) sampling. In order to meet the abundance sum-to-one constraint, we have

$$h_{m,j} = 1 - \sum_{i=1}^{m-1} h_{i,j}. \tag{14}$$

The objective function in (11) is a combinational problem, which is nonconvex, therefore, it is difficult to solve. In this paper, we propose an iterative scheme to optimize $\mathbf{W}$ and $\mathbf{H}$, respectively, both of which are solved by a gradient descent method. The first-order derivatives of the objective function are computed as follows:

$$\begin{cases} \nabla_{\mathbf{U}}(\mathbf{W}, \mathbf{H}) = d(\mathbf{U}) - \dfrac{2\lambda}{n}\mathbf{z}(\mathbf{1}_n)^T .* \mathbf{U} \\ \nabla_{\mathbf{V}}(\mathbf{W}, \mathbf{H}) = d(\mathbf{V}) + \dfrac{2\lambda}{n}\mathbf{z}(\mathbf{1}_n)^T .* (\ln \mathbf{V}./\mathbf{V} - \mathbf{V}) \end{cases} \tag{15}$$

where $./$ is the dot division, $\mathbf{z} = (1/2n)(\mathbf{1}_{m\times n} + \ln \mathbf{V}^2 - \mathbf{U}^2 - \mathbf{V}^2)\mathbf{1}_n$. $d(\mathbf{U})$ and $d(\mathbf{V})$ are the gradients of reconstructed errors, which are

$$\begin{cases} d(\mathbf{U}) = \mathbf{W}^T(\mathbf{W}\mathbf{H} - \mathbf{Y})). * \mathbb{C}_{\text{cons}} \\ d(\mathbf{V}) = \sigma\mathbf{W}^T(\mathbf{W}\mathbf{H} - \mathbf{Y}). * \mathbb{C}_{\text{cons}} \end{cases} \tag{16}$$

where $\mathbb{C}_{\text{cons}}$ is an indicative function, $\mathbb{C}_{\text{cons}} = \mathbf{1}_{m\times n}\{0 < (\mathbf{U} + \sigma\mathbf{V}) < 1\}$. For more details, the derivation of (15) is given in Appendix B.

With respect to the updates of $\mathbf{H}$ and $\mathbf{W}$, we employ the gradient descent method for the solutions as follows:

$$\mathbf{H} \Leftarrow \mathbf{H} + \Delta\mathbf{H}$$

and

$$\mathbf{W} \Leftarrow \mathbf{W} + \Delta\mathbf{W}$$

where $\Delta\mathbf{H}$ and $\Delta\mathbf{W}$ are the gradients for $\mathbf{H}$ and $\mathbf{W}$, respectively. Specifically:

1) for $\mathbf{H}$, we have

$$\Delta\mathbf{H} = -\varphi(\nabla_{\mathbf{U}}(\mathbf{W}, \mathbf{H}) + \sigma\nabla_{\mathbf{V}}(\mathbf{W}, \mathbf{H})) \tag{17}$$

where φ is the learning rates that can be estimated by the Armijo rule [50];

2) for $\mathbf{W}$, we obtain $\Delta\mathbf{W}$ via Adadelta [51] as follows:

$$\Delta\mathbf{W} = -\frac{\text{RMS}[\Delta\mathbf{W}]}{\text{RMS}[\nabla\mathbf{W}(\mathbf{W}, \mathbf{H})]}\nabla_{\mathbf{W}}(\mathbf{W}, \mathbf{H}) \tag{18}$$

where RMS$[\cdot]$ is the root mean square [51]. The first-order derivatives of the objective function (11) are calculated as

$$\nabla_{\mathbf{W}}(\mathbf{W}, \mathbf{H}) = (\mathbf{W}\mathbf{H} - \mathbf{Y})\mathbf{H}^T + \mu d(\text{MinVol}(\mathbf{W})) \tag{19}$$

where $d(\text{MinVol}(\mathbf{W}))$ is the gradient for the volume function, which can be computed as the one in [18].

Finally, a pseudocode of the proposed DAEN is given in Algorithm 1. As shown in Algorithm 1, DAEN consists of two main parts, a set of SAEs for initialization and one VAE for unmixing. Specifically, in Line 1, $\mathbf{M}_i$ is randomly initialized. In Line 2, the hyperparameters are set following [49], while in Line 3, the candidate samples used for training are

Algorithm 1 DAEN for Hyperspectral Unmixing

Input: data set $\mathbf{Y}$.
Output: endmembers $\mathbf{W}$, abundances $\mathbf{H}$.
Step 1. */* SAE for initialization */*
1. **Initialization:** $\mathbf{M}_i$.
2. Set hyper-parameters following [49].
3. Obtain $p \times k$ candidates via VCA [16].
repeat
4. Update $\{\widehat{\mathbf{w}}_i\}_{i=1}^m$ in (5).
5. Update $\{\mathbf{M}_i\}_{i=1}^m$ in (8).
until convergence
6. Compute $\widehat{\mathbf{H}}$ via FCLS [48].
Step 2. */*VAE for unmixing*/*
7. **Initialization:** $\mathbf{U}$ and $\mathbf{V}$.
repeat
8. Update $\Delta\mathbf{H}$ in (17).
9. Update $\Delta\mathbf{W}$ in (18).
until convergence

generated via VCA. In Lines 4 and 5, $\{\widehat{\mathbf{w}}_i\}$ and $\{\mathbf{M}_i\}$ are iteratively updated until SAE terminates. In Line 6, it computes the abundance estimation $\widehat{\mathbf{H}}$ via FCLS. In Line 7, the LV variables, $\mathbf{U}$ and $\mathbf{V}$, are randomly initialized. Finally, in Lines 8 and 9, the endmember matrix $\mathbf{W}$ and the abundance matrix $\mathbf{H}$ are iteratively updated, respectively.

IV. Experiments With Synthetic Data

In this section, we use simulated hyperspectral data to evaluate the effectiveness of our newly proposed DAEN. The simulated data, with a maximum abundance purity of 0.8, are generated according to the LSM, where the pure spectral signatures (with 224 spectral bands covering the spectral range from 0.4 μm to 2.5 μm) are randomly selected from the United States Geological Survey (USGS) spectral library.[1] In our experiments, several different scenarios, with different numbers of pixels, different numbers of endmembers and outliers, have been considered. Note that all our experiments have been performed in a desktop personal computer with Intel Core I7 CPU and 16 GB of RAM.

Three indicators, i.e., SAD, RE, and root mean square error (RMSE) are used to measure the accuracy of the unmixing results, which are defined as follows:

$$\begin{cases} \mathrm{SAD}(\mathbf{w}_i, \widehat{\mathbf{w}}_i) = \arccos\left(\dfrac{[\mathbf{w}_i, \widehat{\mathbf{w}}_i]}{\|\mathbf{w}_i\| \cdot \|\widehat{\mathbf{w}}_i\|}\right) \\ \mathrm{RE}\big(\{\mathbf{y}_j\}_{j=1}^n, \{\widehat{\mathbf{y}}_j\}_{j=1}^n\big) = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\sqrt{\|\mathbf{y}_j - \widehat{\mathbf{y}}_j\|_2^2} \\ \mathrm{RMSE}(\widehat{\mathbf{h}}_j,\ \mathbf{h}_j) = \dfrac{1}{n}\displaystyle\sum_{j=1}^{n}\sqrt{\|\ \mathbf{h}_j - \widehat{\mathbf{h}}_j\|_2^2} \end{cases} \tag{20}$$

where $\widehat{\mathbf{w}}_i$ and $\mathbf{w}_i$ denote the extracted endmember and the library spectrum, $\widehat{\mathbf{y}}_j$ and $\mathbf{y}_j$ are the reconstruction and original signature of pixel j, and $\widehat{\mathbf{h}}_j$ and $\mathbf{h}_j$ are the corresponding estimated and actual abundance fractions, respectively.

[1] https://speclab.cr.usgs.gov/spectrallib.html

TABLE I
SADs (in Radians), RMSEs, and REs Obtained by the Proposed DAEN by Using Different Penalty Coefficients, for a Simulated Data Set With 3 Endmembers, 676 Pixels, SNR = 30 dB, and 3 Outliers

		SAD	RMSE	RE
$\mu = 0.1$	$\lambda = 0.1$	0.0048	0.4129	0.0096
	$\lambda = 0.2$	0.0044	0.4140	0.0106
	$\lambda = 0.3$	0.0092	0.4453	0.0132
	$\lambda = 0.4$	0.0102	0.5051	0.0146
	$\lambda = 0.5$	0.0115	0.4992	0.0155
$\mu = 0.2$	$\lambda = 0.1$	0.0039	0.4109	0.0101
	$\lambda = 0.2$	0.0045	0.4201	0.0117
	$\lambda = 0.3$	0.0100	0.4903	0.0149
	$\lambda = 0.4$	0.0115	0.4899	0.0157
	$\lambda = 0.5$	0.0122	0.4980	0.0164
$\mu = 0.3$	$\lambda = 0.1$	0.0089	0.4705	0.0136
	$\lambda = 0.2$	0.0103	0.4879	0.0151
	$\lambda = 0.3$	0.0112	0.4960	0.0149
	$\lambda = 0.4$	0.0136	0.5016	0.0162
	$\lambda = 0.5$	0.0162	0.5241	0.0177
$\mu = 0.4$	$\lambda = 0.1$	0.0102	0.4831	0.0154
	$\lambda = 0.2$	0.0107	0.4823	0.0160
	$\lambda = 0.3$	0.0128	0.5206	0.0169
	$\lambda = 0.4$	0.0151	0.5196	0.0173
	$\lambda = 0.5$	0.0149	0.5326	0.0180
$\mu = 0.5$	$\lambda = 0.1$	0.0114	0.4794	0.0482
	$\lambda = 0.2$	0.0132	0.5119	0.0502
	$\lambda = 0.3$	0.0164	0.5312	0.0519
	$\lambda = 0.4$	0.0159	0.5507	0.0541
	$\lambda = 0.5$	0.0173	0.5498	0.0552

Experiment 1 (Parameter Settings): In order to verify the performance of network parameters in (12), we test DAEN under different parameter settings. In this test, the data encompasses 3 endmembers, 676 pixels, and SNR = 30 dB. In addition, 3 outliers are included in the data. The obtained results are listed in Table I, which shows that better results can be obtained when the parameters are set to $\mu = 0.1$ and $\lambda = 0.1$. In addition, it can be observed the proposed algorithm is insensitive to the settings with $\mu < 0.3$ and $\lambda < 0.3$. Nevertheless, in the following experiments, we empirically set $\mu = 0.1$ and $\lambda = 0.1$, respectively.

Experiment 2 (Comparison With Other Methods): Comparisons with several widely used unmixing algorithms, including NFINDR [15], VCA [16], PCOMMEND [19], MVCNMF [18], Bayesian [31], and SNSA [43], are included for further evaluation. It should be noted that for the competitors, MVCNMF, Bayesian, and PCOMMEND provide estimations of endmembers and abundances, simultaneously, while for NFINDR and VCA, the abundances are obtained using the FCLS method.

Table II presents the obtained SADs, RMSEs, and REs for different scenarios. From the results reported in Table II, it can be observed that the proposed DAEN obtained competitive results when compared with other methods. Specifically, the proposed approach outperforms the other methods in terms of mean SAD, which reveals the high quality of the endmembers produced by our method. At the same time,

TABLE II

AVERAGE SADS (IN RADIANS), RES, AND RMSES, ALONG WITH THEIR STANDARD DEVIATIONS, OBTAINED FROM 10 MC RUNS BY DIFFERENT METHODS FOR THE CONSIDERED SIMULATED DATA UNDER DIFFERENT SCENARIOS. BEST RESULTS ARE IN BOLD

SNR test: 3 endmembers, 676 pixels, no outlier						
Method	SAD	RMSE	RE	SAD	RMSE	RE
	SNR=40dB			SNR=20dB		
N-FINDR	0.0428±0.04%	0.5074±0.91%	0.0182±0.01%	0.0816±0.94%	0.6403±0.93%	0.0426±0.53%
VCA	0.0430±0.09%	0.4297±0.10%	0.0157±0.03%	0.0550±0.14%	0.5726±0.28%	0.0352±0.62%
MVC-NMF	0.0059±0.03%	0.3920±0.03%	0.0120±0.01%	0.0102±0.72%	0.4496±0.74%	0.0269±0.21%
Bayesian	0.0157±0.05%	0.3819±0.09%	0.0133±0.03%	0.0166±1.13%	0.5014±4.09%	0.0288±0.35%
PCOMMEND	0.0113±1.66%	0.4175±3.67%	**0.0104**±0.27%	0.0327±2.73%	0.4869±5.30%	0.0276±0.29%
SNSA	0.0056±0.04%	0.3903±0.07%	0.0115±0.01%	0.0100±0.72%	0.4504±0.96%	0.0271±0.48%
DAEN	**0.0035**±0.03%	**0.3891**±0.02%	0.0116±0.01%	**0.0092**±0.64%	**0.4342**±0.30%	**0.0260**±0.28%
Outlier test: 676 pixels, 4 endmembers, SNR=30dB						
Method	SAD	RMSE	RE	SAD	RMSE	RE
	0 outliers			10 outliers		
N-FINDR	0.0621±0.05%	0.5454±0.36%	0.0239±0.02%	0.4330±3.24%	0.5881±3.05%	0.0304±0.17%
VCA	0.0533±0.03%	0.4877±0.42%	0.0274±0.07%	0.3856±6.34%	0.6413±1.29%	0.0291±0.14%
MVC-NMF	0.0066±0.09%	0.4387±0.25%	0.0148±0.06%	0.4392±1.95%	0.5977±0.37%	0.0189±0.10%
Bayesian	0.0112±0.12%	0.4728±0.19%	0.0145±0.08%	0.1734±3.82%	0.6420±0.79%	0.0196±0.12%
PCOMMEND	0.0119±0.27%	0.4636±1.05%	**0.0125**±0.03%	0.1736±8.85%	0.5686±8.33%	0.0232±0.14%
SNSA	0.0064±0.04%	0.4307±0.22%	0.0139±0.06%	0.0163±0.74%	0.4742±0.41%	0.0190±0.10%
DAEN	**0.0059**±0.04%	**0.4170**±0.20%	0.0128±0.03%	**0.0126**±0.68%	**0.4503**±0.26%	**0.0183**±0.08%
Endmembers test: 676 pixels, 5 outliers, SNR=30dB						
Method	SAD	RMSE	RE	SAD	RMSE	RE
	4 endmembers			8 endmembers		
N-FINDR	0.3005±0.64%	0.5268±1.04%	0.0818±0.05%	0.5749±0.58%	0.7421±0.31%	0.1153±0.10%
VCA	0.2503±0.49%	0.6986±0.97%	0.0949±0.03%	0.4615±0.72%	0.819±1.84%	0.1049±0.26%
MVC-NMF	0.2904±0.72%	0.5608±0.60%	0.0180±0.05%	0.4286±0.69%	0.7389±1.55%	0.0542±0.12%
Bayesian	0.1413±0.66%	0.7065±0.51%	0.0193±0.04%	0.3713±0.75%	0.8615±0.97%	0.0465±0.34%
PCOMMEND	0.2386±2.68%	0.6517±1.25%	0.0386±0.09%	0.4629±3.51%	0.7016±1.93%	0.0419±0.16%
SNSA	0.0137±0.35%	0.4629±0.58%	0.0151±0.03%	0.1164±0.53%	0.6609±0.87%	0.0294±0.10%
DAEN	**0.0104**±0.30%	**0.4473**±0.45%	**0.0142**±0.03%	**0.0914**±0.50%	**0.6205**±0.84%	**0.0263**±0.09%
Pixels test: 4 endmembers, 5 outliers, SNR=30dB						
Method	SAD	RMSE	RE	SAD	RMSE	RE
	3364 pixels			1156 pixels		
N-FINDR	0.3826±0.79%	0.6751±0.49%	0.0730±0.09%	0.3147±0.81%	0.6135±0.63%	0.0596±0.15%
VCA	0.2950±0.63%	0.7263±0.61%	0.0974±0.13%	0.2681±0.46%	0.5902±0.35%	0.0628±0.10%
MVC-NMF	0.3373±1.46%	0.5916±4.63%	0.0370±0.08%	0.3057±0.21%	0.5723±2.58%	0.0247±0.09%
Bayesian	0.1694±0.98%	0.5825±0.54%	0.0417±0.10%	0.1168±0.35%	0.4985±2.04%	0.0259±0.10%
PCOMMEND	0.3067±1.13%	0.7036±0.72%	**0.0248**±0.35%	0.2892±0.85%	0.5651±0.30%	0.0240±0.22%
SNSA	0.0203±0.34%	0.5029±0.35%	0.0256±0.07%	0.0152±0.28%	0.4749±0.22%	0.0231±0.09%
DAEN	**0.0192**±0.30%	**0.4935**±0.24%	0.0251±0.06%	**0.0140**±0.14%	**0.4624**±0.19%	**0.0228**±0.06%

DAEN is able to retrieve more stable abundance estimates relying on its accurate endmember identification step. The advantages are remarkable, especially for problems with low SNR, outliers, and a relatively large number of endmembers and pixels.

For illustrative purposes, Fig. 2 displays the obtained endmember signatures for a problem with 3364 pixels, 4 endmembers, 10 outliers, and SNR $=$ 20 dB, along with the references from the USGS library. It can be seen that the results obtained by the proposed DAEN are better than those produced by the other methods.

V. EXPERIMENTS WITH REAL DATA

In this section, the proposed DAEN approach is applied to two real hyperspectral images: Mangrove [52], Samson [53], and Cuprite [54] data sets for further evaluation. In these experiments, the parameters involved in the considered algorithms follow the settings in the simulated experiments, i.e., we use $\mu = 0.1$ and $\lambda = 0.1$, respectively.

A. Experiment With Mangrove Data Set

The Mangrove data is an EO-1 Hyperion (hyperspectral) image which has been obtained from the USGS Earth Resources Observation and Science Center through a data acquisition request to the satellite data provider [52], [55], and collected over the Henry Island of the Sunderban Biosphere Reserve of West Bengal, India. Atmospheric correction has converted the radiance of the data to reflectance units by using FLAASH model in ENVI software, and the endmembers (pure signatures of mangrove species) of the data have been

TABLE III

SADs (IN RADIANS) AND REs ALONG WITH THEIR STANDARD DEVIATIONS OBTAINED BY DIFFERENT METHODS FOR THE MANGROVE DATA FROM 10 MC RUNS, WHERE THE BEST RESULTS ARE IN BOLD

Mineral	N-FINDR	VCA	MVC-NMF	Bayesian	PCOMMEND	SNSA	DAEN
Avicennia	0.1495±0%	0.9602±4.65%	0.1973±2.18%	0.1001±3.05%	0.0992±2.81%	0.0995±4.26%	**0.0968**±2.58%
Bruguiera	0.8235±0%	0.8824±7.79%	0.9825±7.53%	1.5957±4.16%	0.9564±3.95%	0.1103±3.84%	**0.1025**±4.03%
Excoecaria	0.7361±0%	0.7377±1.50%	0.0904±4.08%	0.0963±2.96%	0.0876±5.73%	0.0880±3.27%	**0.0863**±1.96%
Phoenix	0.1306±0%	1.3063±0.12%	0.9782±3.93%	0.9624±3.93%	1.7065±8.23%	0.0711±4.41%	**0.0706**±3.72%
Mean SAD	0.4599	0.9717	0.5621	0.6886	0.7124	0.0922	**0.0890**
RE	0.0822±0%	0.0980±0.40%	0.0392±2.36%	0.0129±4.57%	0.0162±5.19%	0.0057±0.25%	**0.0050**±0.13%

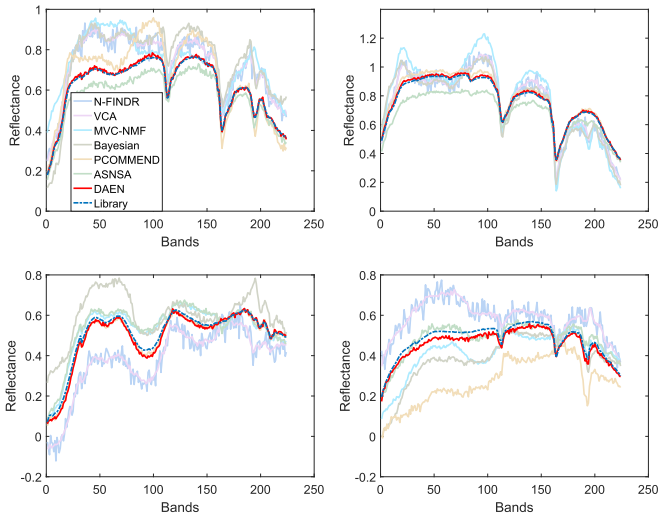

Fig. 2. Estimated endmember signatures by the proposed method for a problem with 4 endmembers, 3364 pixels, 10 outliers, and SNR = 20 dB, along with the references from the USGS library.

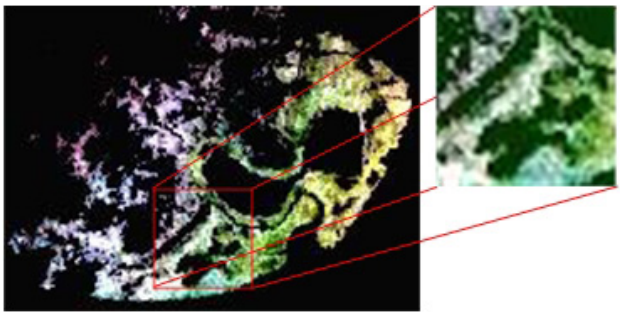

Fig. 3. 45 × 45 pixel subscene of the Mangrove data used in our experiment.

identified by a ground survey of the study area, including *Avicennia*, *Bruguiera*, *Excoecaria*, and *Phoenix*. The Mangrove data, as shown in Fig. 3, include 137 × 187 pixels and 155 bands, with a spatial resolution of 30 m. For detailed information of the Mangrove data, we refer to [52].

In our experiment, a subscene with 45 × 45 pixels of the Mangrove data has been used to further evaluate the proposed DAEN. Following [52], the considered subscene contains 4 endmembers, i.e., $m = 4$.

Table III presents the obtained quantitative results from the Mangrove data. It can be seen that the proposed DAEN achieved very promising results for the four considered mangrove spices. However, the other competitors ended up with

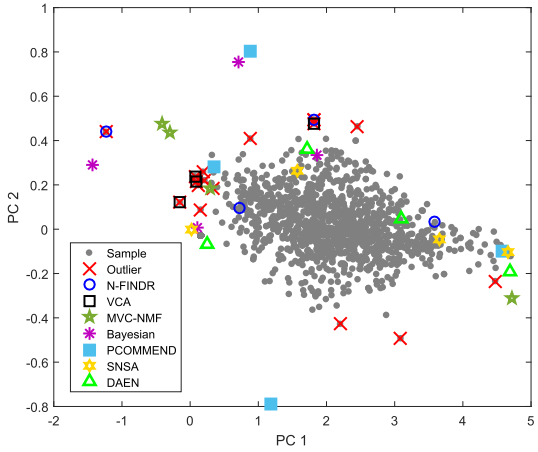

Fig. 4. Unmixing results for the subscene of the Mangrove data, where the data are projected onto the first two principal components (PCs).

errors when detecting or estimating the endmembers. This is due to the fact that according to our observation, the Mangrove scene contains many outliers across the whole image, which brings a lot of difficulties for general unmixing methods. This point was verified by our experiment, in which we detected a total of 17 outliers. For illustrative purposes, Fig. 4 scatterplots the unmixing results obtained by the considered methods, in which the detected outliers are also illustrated. In Fig. 4, we can observe that the proposed DAEN produced good unmixing results for this data set, while all the other methods resulted in problems.

Finally, for illustrative purposes, the estimated endmember signatures, along with their ground references, and the corresponding abundance maps obtained by the proposed DAEN are shown in Fig. 5. Effective results can be observed from Fig. 5.

In summary, our experiments with this challenging Mangrove data set demonstrate the effectiveness of the proposed DAEN for real scenarios with outliers, which is a general situation in real problems.

B. Experiment With the Samson Data Set

In this experiment, we use the Samson data set that includes 156 bands covering the wavelengths from 0.401 to 0.889 μm and 95 × 95 pixels, as shown in Fig. 6, for validation [53]. There are three endmembers, including soil, tree, and water, in the ground truth image.

Table IV demonstrates the obtained quantitative results for the considered methods. It can be observed that the

TABLE IV

SADs (IN RADIANS) AND REs ALONG WITH THEIR STANDARD DEVIATIONS OBTAINED BY DIFFERENT METHODS FOR THE SAMSON DATA FROM 10 MC RUNS, WHERE THE BEST RESULTS ARE IN BOLD

Mineral	N-FINDR	VCA	MVC-NMF	Bayesian	PCOMMEND	SNSA	DAEN
Soil	0.0713±0%	0.0627±1.85%	**0.0402**±3.87%	0.1062±7.25%	0.2849±4.35%	0.0410±5.02%	0.0405±2.76%
Tree	0.0495±0%	0.0501±7.82%	0.0261±3.62%	0.0610±8.34%	0.0505±6.14%	0.0205±2.89%	**0.0196**±3.52%
Water	0.0408±0%	0.0273±3.74%	0.0304±5.29%	0.0364±2.48%	0.0716±4.34%	0.0291±2.59%	**0.0279**±3.83%
Mean SAD	0.0539	0.0467	0.0322	0.0679	0.1357	0.0302	**0.0293**
RMSE	0.9572±0%	0.8926±1.35%	0.6430±0.98%	0.7501±1.63%	0.9439±2.35%	0.6143±3.37%	**0.6097**±3.62%
RE	0.0129±0%	0.0116±0.19%	0.0075±0.86%	0.0103±0.42%	**0.0057**±1.15%	0.0066±0.25%	0.0062±0.85%

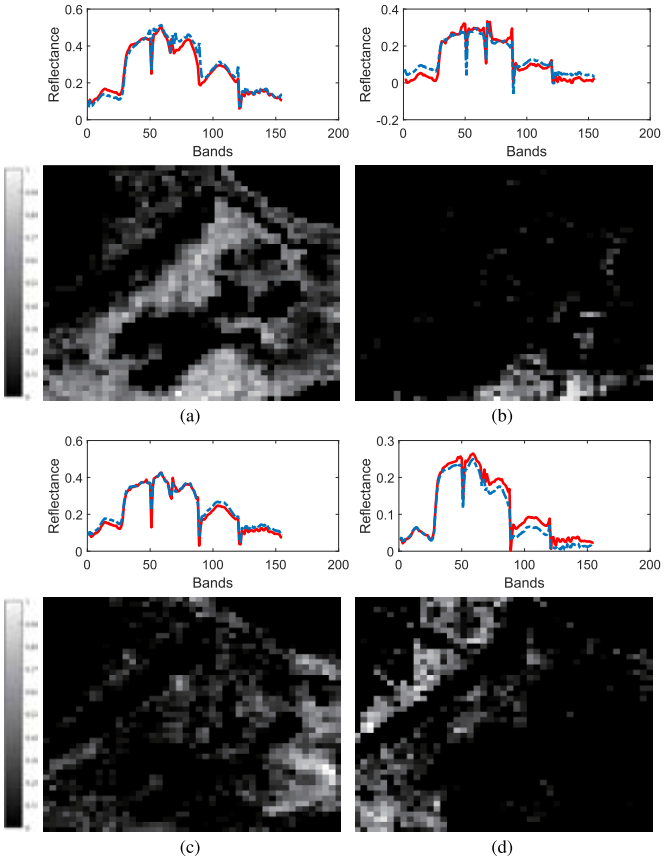

Fig. 5. Estimated endmember signatures (in red), along with the ground reference (in blue) and their corresponding abundance maps by the proposed DAEN. (a) Avicennia. (b) Bruguiera. (c) Excoecaria. (d) Phoenix.

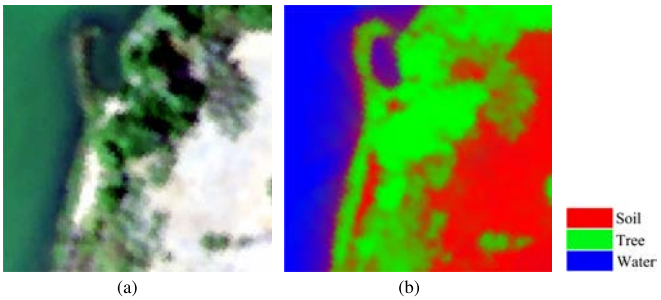

Fig. 6. (a) Samson image and its (b) corresponding ground truth.

proposed DAEN obtained the best mean SAD and RMSE. For illustrative purposes, the endmembers signatures and the estimated abundances are shown in Fig. 7. Fig. 7 reveals that

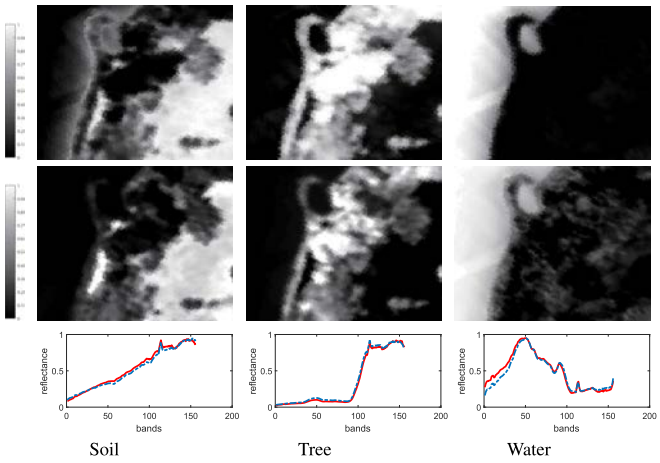

Fig. 7. Results obtained by the proposed DAEN on the Samson data set. (Top) Ground truth abundance maps on Samson data. (Middle) Estimated abundance maps from the proposed DAEN. (Bottom) Estimated endmember signatures (in red) along with their corresponding reference signatures (in blue).

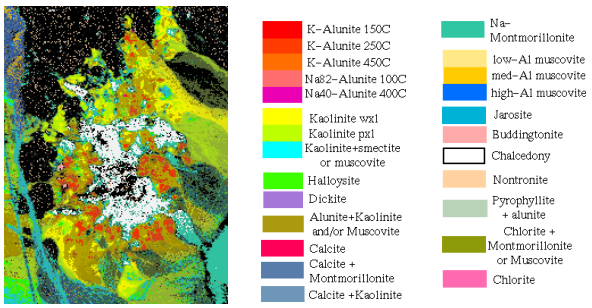

Fig. 8. Cuprite data set (Tricorder 3.3 product), showing the location of different minerals in the scene.

the endmembers and abundances, estimated from DAEN, have good matches with regard to the corresponding ones in the ground truth.

C. Experiment With the Cuprite Data Set

The third real hyperspectral data, namely Cuprite, was collected by the Airborne Visible Infra-Red Imaging Spectrometer over the Cuprite mining site, Nevada, in 1997 [54]. This scene has 224 spectral bands over a wavelength from 0.4 μm to 2.5 μm, with a nominal spectral resolution of 10 nm. Prior to the analysis, bands 1–2, 105–115, 150–170, and 223–224 were removed due to low SNR and water

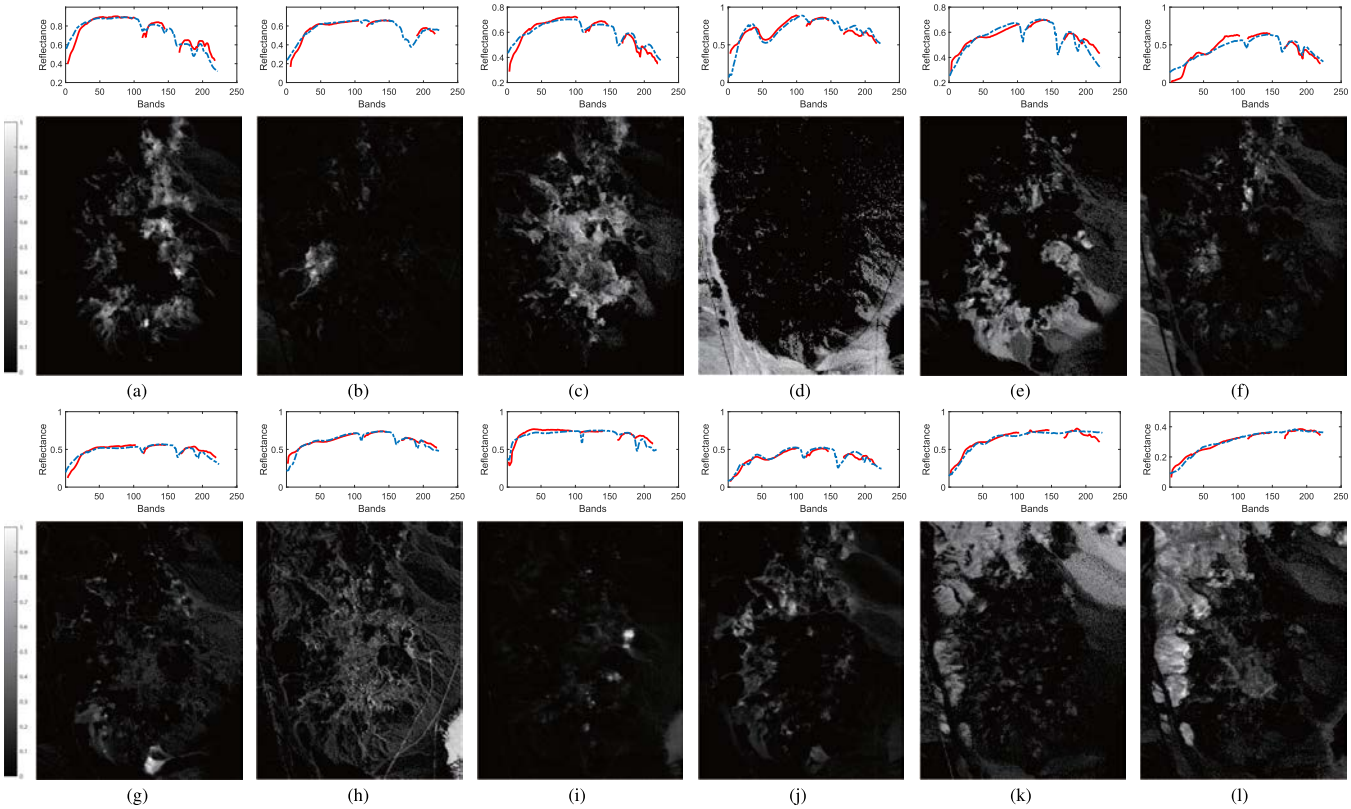

Fig. 9. Estimated endmembers (in red), along with the references from the USGS library (in blue), and their corresponding abundance maps obtained by the proposed DAEN. (a) Alunite, (b) Buddingtonite, (c) Chalcedony, (d) Jarosite, (e) Kaolin/Smect, (f) Kaolinite wx1, (g) Kaolinite px1, (h) Montmorillonite, (i) Muscovite, (j) Nontronite, (k) Pyrope, and (l) Sphene.

TABLE V

SADs (IN RADIANS) AND REs ALONG WITH THEIR STANDARD DEVIATIONS OBTAINED AFTER 10 MC RUNS BY DIFFERENT METHODS FOR THE CUPRITE DATA, WHERE THE BEST RESULTS ARE IN BOLD

Mineral	N-FINDR	VCA	MVC-NMF	Bayesian	PCOMMEND	SNSA	DAEN
Alunite	**0.0992**±0%	0.1019±2.65%	0.1304±3.94%	0.1197±6.12%	0.1450±4.37%	0.1201±3.12%	0.1097±4.29%
Buddingtonite	0.1639±0%	0.1208±1.54%	0.0992±4.01%	0.1140±5.38%	0.1208±3.41%	0.1062±2.95%	**0.0939**±5.06%
Chalcedony	0.1687±0%	0.1704±3.47%	0.1108±1.32%	0.1297±4.16%	0.1315±2.92%	**0.1097**±4.10%	0.1104±4.29%
Jarosite	0.1893±0%	0.1490±3.47%	**0.1085**±2.65%	0.1264±3.32%	0.1164±3.29%	0.1131±3.62%	0.1095±2.81%
Kaolin/Smect	0.0804±0%	**0.0764**±4.01%	0.0826±2.28%	0.1003±5.16%	0.0936±3.63%	0.0819±5.43%	0.0801±4.15%
Kaolinite wx1	0.0792±0%	0.0805±1.78%	0.0813±2.84%	0.0909±3.49%	0.0794±4.28%	0.0794±3.92%	**0.0786**±2.15%
Kaolinite px1	0.0108±0%	0.0992±3.12%	**0.0937**±3.56%	0.0982±3.52%	0.1123±5.04%	0.1079±3.54%	0.1086±4.01%
Montmorillonite	0.0651±0%	0.0587±1.95%	0.0635±4.56%	0.0591±3.67%	0.0798±4.72%	0.0616±3.83%	**0.0583**±2.54%
Muscovite	0.1437±0%	0.1302±2.14%	0.1291±1.97%	0.1284±3.25%	0.1351±3.17%	**0.1252**±2.75%	0.1303±3.12%
Nontronite	0.0803±0%	0.0816±4.41%	0.0804±3.85%	0.0887±4.63%	0.0914±3.94%	0.0825±3.32%	**0.0796**±3.97%
Pyrope	0.0532±0%	0.0607±5.18%	0.0574±5.18%	**0.0528**±4.90%	0.0741±3.26%	0.0638±4.47%	0.0593±3.82%
Sphene	0.1496±0%	0.1541±3.95%	**0.1380**±4.21%	0.1589±3.64%	0.1442±5.18%	0.1396±3.91%	0.1407±4.45%
Mean SAD	0.1070	0.1095	0.0979	0.1056	0.1103	0.0993	**0.0966**
RE	0.0095±0%	0.0054±0.07	0.0038±0.04	0.0042±0.08	**0.0030**±0.06	0.0038±0.07	0.0032±0.05

absorption in those bands, leaving a total of 188 spectral bands. The portion of the data used in the experiment has 191×250 pixels. The Cuprite site is well understood mineralogically and has several exposed minerals of interest, all included in the USGS library considered in the experiment. Fig. 8 shows a mineral map produced in 1995 by USGS, in which the Tricorder 3.3 software product was used to map different minerals present in the Cuprite mining district. It should be noted that following the work in [22] and [53], the number of endmembers considered in the experiment is set to 12.

For illustrative purposes, the estimated endmember signatures, along with the references from the USGS library and their corresponding abundance maps, are depicted in Fig. 9. From this figure, we can observe that the estimated endmember signatures generally match well with the references. These results indicate the effectiveness of the proposed DAEN for hyperspectral unmixing in real scenarios.

Finally, Table V presents the obtained SADs and REs from all the considered methods. From this table, we can further observe that the proposed DAEN obtained better or competitive results when compared with other methods, in terms of both SADs and REs, which reveals the quality of the endmembers and fractional abundances estimated by our method in this particular scenario.

VI. Conclusion

In this paper, we have introduced a new hyperspectral unmixing approach based on a DAEN which includes two parts, i.e., a set of SAEs for initialization and a VAE for nonnegative matrix factorization, aimed at obtaining the endmember signatures and the abundance fractions, simultaneously. Our experimental results, conducted using both synthetic and real data sets, indicate that the proposed DAEN can particularly handle problems with significant outliers. This is an important contribution since the presence of outliers is a general situation in real problems and traditional unmixing algorithms are often misguided by outliers that can be also understood as endmembers due to their singularity. This peculiarity of our method has been investigated and substantiated via experiments with both simulated and real hyperspectral data sets. As the proposed DAEN is based on the linear model, future works will investigate the application of DAEN to nonlinear unmixing scenarios, in order to achieve enhanced characterizations via the combination of endmembers in hyperspectral data sets.

Appendix

A. Objective Function $f_2(\mathbf{H}(\mathbf{U},\mathbf{V}))$

Let vector $\mathbf{h}_j$ represents the LV. Following [45], the prior is set to multivariate Gaussian $p(\mathbf{h}_j) = N(\mathbf{0},\mathbf{I})$, and the posterior is given by $q(\mathbf{h}_j|\mathbf{y}_j) = N(\mathbf{u}_j, \mathbf{v}_j{}^2\mathbf{I})$. The objective of the original VAE is defined as

$$D_{\mathrm{KL}} \quad (q(\mathbf{h}_j|\mathbf{y}_j) \parallel p(\mathbf{h}_j)) - E_q \ (\ln\ p(\mathbf{y}_j|\mathbf{h}_j)) \tag{21}$$

where $D_{\mathrm{KL}}(\cdot)$ denotes the Kullback–Leibler divergence, which is also called the expected RE [46]. $-E_q(\ln\ p(\mathbf{y}_j|\mathbf{h}_j)$ is regarded as RE. Following the work in [45], the Frobenius norm of $D_{KL}(q(\mathbf{h}_j|\mathbf{y}_j) \parallel p(\mathbf{h}_j))$ can be written as

$$\|D_{KL}(q(\mathbf{h}_j|\mathbf{y}_j) \parallel p(\mathbf{h}_j))\|_F^2 \tag{22}$$

$$= \left\| \int q(\mathbf{h}_j|\mathbf{y}_j)\ln\ p(\mathbf{h}_j)d\mathbf{h}_j - \int q(\mathbf{h}_j|\mathbf{y}_j)\ln q(\mathbf{h}_j|\mathbf{y}_j)d\mathbf{h}_j \right\|_F^2. \tag{23}$$

For convenience, we split the right-hand side of (23) into the following two terms:

$$\begin{cases} \Phi_1 = \int q(\mathbf{h}_j|\mathbf{y}_j)\ln\ p(\mathbf{h}_j)d\mathbf{h}_j \\ \Phi_2 = \int q(\mathbf{h}_j|\mathbf{y}_j)\ln q(\mathbf{h}_j|\mathbf{y}_j)d\mathbf{h}_j. \end{cases} \tag{24}$$

The first term in (24) can be obtained as

$$\begin{aligned} \Phi_1 &= \int q(\mathbf{h}_j|\mathbf{y}_j)\ln\ p(\mathbf{h}_j)d\mathbf{h}_j \\ &= E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}(\ln\ p(\mathbf{h}_j)) \\ &= E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)} \quad \left(\ln\ \left(\frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{\mathbf{h}_j{}^2}{2}\right)\right)\right) \\ &= E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}\left(-\frac{1}{2}\ln\ (2\pi)\right) - \frac{1}{2}E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}\left(\mathbf{h}_j^2\right). \end{aligned} \tag{25}$$

Let $\mathrm{Var}(\mathbf{h}_j)$ denote the variance of $\mathbf{h}_j$

$$\begin{aligned} \mathrm{Var}(\mathbf{h}_j) &= E(\mathbf{h}_j - E(\mathbf{h}_j))^2 \\ &= E\left(\mathbf{h}_j^2 - 2\mathbf{h}_j E(\mathbf{h}_j) + (E(\mathbf{h}_j))^2\right) \\ &= E\left(\mathbf{h}_j^2\right) - 2(E(\mathbf{h}_j))^2 + (E(\mathbf{h}_j))^2 \\ &= E\left(\mathbf{h}_j^2\right) - (E(\mathbf{h}_j))^2 \end{aligned} \tag{26}$$

where $E(\cdot) = E_{\mathbf{h}_j\sim D(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I})}(\cdot)$, we can obtain

$$\begin{aligned} E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}\left(\mathbf{h}_j^2\right) &= \mathrm{Var}(\mathbf{h}_j) + (E(\mathbf{h}_j))^2 \\ &= \mathbf{u}_j^2 + \mathbf{v}_j^2. \end{aligned} \tag{27}$$

Therefore, we have

$$\begin{aligned} \Phi_1 &= E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}\left(-\frac{1}{2}\ln\ (2\pi)\right) - \frac{1}{2}E_{\mathbf{h}_j\sim D\left(\mathbf{u}_j,\mathbf{v}_j^2\mathbf{I}\right)}\left(\mathbf{h}_j^2\right) \\ &= -\frac{1}{2}\ln\ 2\pi - \frac{1}{2n}\sum_{j=1}^{n}\left(\mathbf{u}_j^2 + \mathbf{v}_j^2\right). \end{aligned} \tag{28}$$

Similarly, we can obtain the second term Φ_2 in (24) as follows:

$$\begin{aligned} \Phi_2 &= \int q(\mathbf{h}_j|\mathbf{y}_j)\ln q(\mathbf{h}_j|\mathbf{y}_j)d\mathbf{h}_j \\ &= -\frac{1}{2}\ln 2\pi - \frac{1}{2n}\sum_{j=1}^{n}\left(\ln\mathbf{v}_j^2 + 1\right). \end{aligned} \tag{29}$$

By introducing (28) and (29) into (24), we can obtain

$$\begin{aligned} &\|D_{KL}(q(\mathbf{h}_j|\mathbf{y}_j) \parallel p(\mathbf{h}_j))\|_2^2 \\ &= \sum_{i=1}^{m-1}\left(\frac{1}{2n}\sum_{j=1}^{n}(1 + \ln(v_{i,j})^2 - (u_{i,j})^2 - (v_{i,j})^2)\right)^2. \end{aligned} \tag{30}$$

Let $p_{i,j} = 1 + \ln(v_{i,j})^2 - (u_{i,j})^2 - (v_{i,j})^2$, then we have

$$\mathbf{P} = \mathbf{1}_{m\times n} + \ln\mathbf{V}^2 - \mathbf{U}^2 - \mathbf{V}^2 \tag{31}$$

where $\mathbf{1}_{m\times n}$ is a matrix of 1s, and the matrices of reparameters are $\mathbf{U} = [\mathbf{u}_1,\ldots,\mathbf{u}_n]$ and $\mathbf{V} = [\mathbf{v}_1,\ldots,\mathbf{v}_n]$, respectively.

Let $z_i = (1/2n)\sum_{j=1}^{n} p_{i,j}$, $\mathbf{z} = [z_1,\ldots,z_m]^T \in \mathbb{R}^m$, $i = 1,\ldots m$, we have

$$\begin{aligned} \mathbf{z} &= \frac{1}{2n}\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m-11} & p_{m-12} & \cdots & p_{m-1n} \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \\ &= \frac{1}{2n}(\mathbf{1}_{m\times n} + \ln\mathbf{V}^2 - \mathbf{U}^2 - \mathbf{V}^2)\mathbf{1}_n \end{aligned} \tag{32}$$

where $\mathbf{1}_n = [1,\ldots,1]^T \in \mathbb{R}^n$. According to (30) and (32), we can obtain $f_2(\mathbf{H})$ with $\mathbf{U}$ and $\mathbf{V}$ as follows:

$$f_2(\mathbf{H}(\mathbf{U},\mathbf{V})) = \left\|\frac{1}{2n}(\mathbf{1}_{\mathbf{m}\times\mathbf{n}} + \ln\mathbf{V}^2 - \mathbf{U}^2 - \mathbf{V}^2)\mathbf{1}_n\right\|_2^2. \tag{33}$$

B. Partial Derivatives of Reparameters

The first-order derivative of $f_2(\mathbf{H}(\mathbf{U},\mathbf{V}))$ for $u_{i,j}$ can be calculated as

$$\frac{\partial f_2}{\partial u_{i,j}} = \frac{\partial f_2}{\partial z_i}\cdot\frac{\partial z_i}{\partial u_{i,j}} = -2z_i\left(\frac{1}{n}u_{i,j}\right). \tag{34}$$

Similarly, the first-order derivative for $v_{i,j}$ can be obtained as

$$\frac{\partial f_2}{\partial v_{i,j}} = \frac{\partial f_2}{\partial z_i}\cdot\frac{\partial z_i}{\partial v_{i,j}} = \frac{2}{n}z_i\left(\frac{\ln v_{i,j}}{v_{i,j}} - v_{i,j}\right). \tag{35}$$

Then, the problem converts to the following one:

$$\begin{cases}\dfrac{\partial f_2}{\partial \mathbf{U}} = -\dfrac{2}{n}\mathbf{z}(\mathbf{1_n})^T * \mathbf{U}\\ \dfrac{\partial f_2}{\partial \mathbf{V}} = \dfrac{2}{n}\mathbf{z}(\mathbf{1}_n)^T * (\ln \mathbf{V}./\mathbf{V} - \mathbf{V}).\end{cases} \tag{36}$$

Let $R_0 = (1/2)\|\mathbf{Y}-\mathbf{WH}\|_F^2$, and as

$$\frac{1}{2}\|\mathbf{Y}-\mathbf{WH}\|_F^2 = \frac{1}{2}\sum_j^n\sum_k^l\left(y_{k,j} - \sum_i^m w_{b,i}h_{i,j}\right)^2$$

where d is the number of bands, for $k = 1,\ldots d$, the first-order derivative of R_0 is then defined as

$$\frac{\partial R_0}{\partial h_{i,j}} = -\sum_k^d (y_{k,j} - w_{k,i}h_{i,j})w_{k,i}. \tag{37}$$

For $u_{i,j}$ and $v_{i,j}$, the first-order derivatives of the abundance fraction are given as follows, respectively:

$$\frac{\partial h_{i,j}}{\partial u_{i,j}} = \begin{cases}1, & 0 < (u_{i,j} + \sigma v_{i,j}) < 1\\ 0, & \text{otherwise}\end{cases} \tag{38}$$

and

$$\frac{\partial h_{i,j}}{\partial u_{i,j}} = \begin{cases}\sigma, & 0 < (u_{i,j} + \sigma v_{i,j}) < 1\\ 0, & \text{otherwise.}\end{cases} \tag{39}$$

By introducing (38) and (39) into (37), we have

$$\begin{cases}\dfrac{\partial R_0}{\partial \mathbf{U}} = \mathbf{W}^T(\mathbf{WH}-\mathbf{Y})).*\mathbb{C}_{\text{cons}}\\ \dfrac{\partial R_0}{\partial \mathbf{V}} = \sigma\mathbf{W}^T(\mathbf{WH}-\mathbf{Y}).*\mathbb{C}_{\text{cons}}\end{cases} \tag{40}$$

where $\mathbb{C}_{\text{cons}}$ is an indicative function, $\mathbb{C}_{\text{cons}} = \mathbf{1}_{m\times n}\{0 < (\mathbf{U}+\sigma\mathbf{V}) < 1\}$. Then, we can obtain the first-order derivatives of the objective function as follows:

$$\begin{cases}\nabla\mathbf{U} = \dfrac{\partial R_0}{\partial \mathbf{U}} + \dfrac{\partial f_2}{\partial \mathbf{U}}\\ \nabla\mathbf{V} = \dfrac{\partial R_0}{\partial \mathbf{V}} + \dfrac{\partial f_2}{\partial \mathbf{V}}.\end{cases} \tag{41}$$

Finally, the objective function (15) is derived by combining (36), (40), and (41).

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their constructive suggestions which significantly improved the clarity of this paper.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. 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] X. Jin and Y. Gu, "Superpixel-based intrinsic image decomposition of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 8, pp. 4285–4295, Aug. 2017.

[3] Y. Gu, J. Chanussot, X. Jia, and J. A. Benediktsson, "Multiple Kernel learning for hyperspectral image classification: A review," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6547–6565, Nov. 2017.

[4] W. Wei, L. Zhang, Y. Jiao, C. Tian, C. Wang, and Y. Zhang, "Intr-acluster structured low-rank matrix analysis method for hyperspectral denoising," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2018.2862384.

[5] L. Zhang, W. Wei, C. Bai, Y. Gao, and Y. Zhang, "Exploiting clustering manifold structure for hyperspectral imagery super-resolution," *IEEE Trans. Image Process.*, vol. 27, no. 12, pp. 5969–5982, Dec. 2018.

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

[7] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.

[8] 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.

[9] A. Marinoni and P. Gamba, "A novel approach for efficient p-linear hyperspectral unmixing," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1156–1168, Sep. 2015.

[10] J. Cui, X. Li, and L. Zhao, "Nonlinear spectral mixture analysis by determining per-pixel endmember sets," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1404–1408, Aug. 2014.

[11] F. Zhu and P. Honeine, "Biobjective nonnegative matrix factorization: Linear versus kernel-based models," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 4012–4022, Jul. 2016.

[12] 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.

[13] B. Yang, B. Wang, Z. Wu, and Q. Lu, "Bilinear mixture models based unsupervised nonlinear unmixing using constrained nonnegative matrix factorization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 582–585.

[14] 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.

[15] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE Image Symp. V*, vol. 3753, 1999, pp. 266–277.

[16] J. M. P. Nascimento and J. M. Bioucas-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.

[17] V. F. Haertel and Y. E. Shimabukuro, "Spectral linear mixing model in low spatial resolution image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 11, pp. 2555–2562, Nov. 2005.

[18] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[19] A. Zare, P. Gader, O. Bchir, and H. Frigui, "Piecewise convex multiple-model endmember detection and spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2853–2862, May 2013.

[20] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, "Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5067–5082, Sep. 2015.

[21] 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.

[22] 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.

[23] T.-H. Chan, A. Ambikapathi, W.-K. Ma, and C.-Y. Chi, "Robust affine set fitting and fast simplex volume max-min for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 3982–3997, Jul. 2013.

[24] L. Zhang, W. Wei, Y. Zhang, C. Shen, A. van den Hengel, and Q. Shi, "Cluster sparsity field: An internal hyperspectral imagery prior for reconstruction," *Int. J. Comput. Vis.*, vol. 126, no. 8, pp. 797–821, 2018.
[25] G. Gao and Y. Gu, "Tensorized principal component alignment: A unified framework for multimodal high-resolution images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 46–61, Jan. 2019.
[26] J. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. IEEE GRSS Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Aug. 2009, pp. 1–4.
[27] Y. Altmann, S. McLaughlin, and A. Hero, "Robust linear spectral unmixing using anomaly detection," *IEEE Trans. Comput. Imag.*, vol. 1, no. 2, pp. 74–85, Jun. 2015.
[28] S. Zhang, A. Agathos, and J. Li, "Robust minimum volume simplex analysis for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6431–6439, Nov. 2017.
[29] J. M. P. Nascimento and J. M. Bioucas-Dias, "Hyperspectral unmixing algorithm via dependent component analysis," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2007, pp. 4033–4036.
[30] J. M. P. Nascimento and J. M. Bioucas-Dias, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 863–878, Mar. 2012.
[31] 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.
[32] N. Dobigeon, S. Moussaoui, J.-S. Tourneret, and C. Carteret, "Bayesian separation of spectral sources under non-negativity and full additivity constraints," *Signal Process.*, vol. 89, no. 12, pp. 2657–2669, Dec. 2009.
[33] F. Schmidt, A. Schmidt, E. Tréguier, M. Guiheneuf, S. Moussaoui, and N. Dobigeon, "Implementation strategies for hyperspectral unmixing using Bayesian source separation," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4003–4013, Nov. 2010.
[34] J. Plaza, A. Plaza, R. Perez, and P. Martinez, "On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images," *Pattern Recognit.*, vol. 42, no. 11, pp. 3032–3045, 2009.
[35] J. Plaza and A. Plaza, "Spectral mixture analysis of hyperspectral scenes using intelligently selected training samples," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 371–375, Apr. 2010.
[36] G. A. Licciardi and F. D. 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] R. Guo, W. Wang, and H. Qi, "Hyperspectral image unmixing using autoencoder cascade," in *Proc. 7th Workshop Hyperspectral Image Signal Process., Evol. Remote, Sens. (WHISPERS)*, Jun. 2015, pp. 1–4.
[38] Y. Su, A. Mariononi, J. Li, A. Plaza, and P. Gamba, "Nonnegative sparse autoencoder for robust endmember extraction from remotely sensed hyperspectral images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 205–208.
[39] Y. Qu and H. Qi, "uDAS: An untied denoising autoencoder with sparsity for spectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: 10.1109/TGRS.2018.2868690.
[40] S. Ozkan, B. Kaya, and G. B. Akar, "EndNet: Sparse AutoEncoder network for endmember extraction and hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 1, pp. 482–496, Jan. 2019.
[41] F. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Neural network hyperspectral unmixing with spectral information divergence objective," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 755–758.
[42] B. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Hyperspectral unmixing using a neural network autoencoder," *IEEE Access*, vol. 6, pp. 25646–25656, 2018.
[43] Y. Su, A. Marinoni, J. Li, J. Plaza, and P. Gamba, "Stacked nonnegative sparse autoencoders for robust hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 9, pp. 1427–1431, Sep. 2018.
[44] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*. Cambridge, U.K.: Cambridge Univ. Press, 1993.
[45] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Dec. 2014, pp. 1–14.
[46] C. Doersch. (Jun. 2016). "Tutorial on variational autoencoders." [Online]. Available: https://arxiv.org/abs/1606.05908
[47] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and Y. Huang, "Deep auto-encoder network for hyperspectral image unmixing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2018, pp. 6400–6403.
[48] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[49] A. Lemme, R. F. Reinhart, and J. J. Steil, "Online learning and generalization of parts-based image representations by non-negative sparse autoencoders," *Neural Netw.*, vol. 33, pp. 194–203, Sep. 2012.
[50] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer, 2006, pp. 33–36.
[51] M. D. Zeiler. (Dec. 2012). "ADADELTA: An adaptive learning rate method." [Online]. Available: https://arxiv.org/abs/1212.5701
[52] S. Chakravortty, J. Li, and A. Plaza, "A technique for subpixel analysis of dynamic mangrove ecosystems with time-series hyperspectral image data," *IEEE J., Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1244–1252, Apr. 2018.
[53] 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.
[54] R. O. Green *et al.*, "Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 227–248, Sep. 1998.
[55] S. Chakravortty and D. Sinha, "Performance of pure pixel extraction algorithms on hyperspectral data for species level classification of mangroves," in *Proc. 4th Int. Conf. Emerg. Appl. Inf. Technol. (EAIT)*, Dec. 2014, pp. 209–214.

Yuanchao Su (S'16) received the B.S. and M.Sc. degrees from the Xi'an University of Science and Technology, Xi'an, China, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China.

Since 2018, he has been a Visiting Researcher with the Advanced Imaging and Collaborative Information Processing Group, Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA. His research interests include hyperspectral unmixing, target detection, neural network, deep learning, and image processing.

Mr. Su is a member of the IEEE Geoscience and Remote Sensing Society and serves as a reviewer for several IEEE journals and SPIE journals.

Jun Li (M'12–SM'16) received the B.S. degree from Hunan Normal University, Changsha, China, in 2004, the M.Sc. degree in remote sensing and photogrammetry from Peking University, Beijing, China, in 2007, and the Ph.D. degree in electrical and computer engineering from the Instituto Superior Tecnico, Technical University of Lisbon, Lisbon, Portugal, in 2011.

From 2011 to 2012, she was a Post-Doctoral Researcher with the Department of Technology of Computers and Communications, University of Extremadura, Badajoz, Spain. From 2013 to 2018, she was a Professor with the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China, where she founded her own research group on hyperspectral image analysis. She is currently a Professor with the College of Electrical and Information Engineering, Hunan University, Changsha. Since 2013, she has obtained several prestigious funding grants at the national and international level. Her research interests include remotely sensed hyperspectral image analysis, signal processing, supervised/semisupervised learning, and active learning.

Dr. Li is an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING and the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, and has been a reviewer for several journals.

Antonio Plaza (M'05–SM'07–F'15) received the M.Sc. and Ph.D. degrees in computer engineering from the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is currently the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He is currently one of the top cited authors in Spain and in the University of Extremadura. He has guest edited 10 special issues on hyperspectral remote sensing for different journals. His research interests include remotely sensed hyperspectral image analysis, signal processing, and efficient implementations of large-scale scientific problems on high-performance computing architectures, including commodity Beowulf clusters, heterogeneous networks of computers and clouds, and specialized computer architectures such as field-programmable gate arrays or graphical processing units.

Dr. Plaza was a member of the Editorial Board of the IEEE Geoscience and Remote Sensing Newsletter from 2011 to 2012 and *IEEE Geoscience and Remote Sensing Magazine* in 2013. He was also a member of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He is a fellow of IEEE for the contributions to hyperspectral data processing and parallel computing of Earth observation data. He was a recipient of a recognition as an Outstanding Associate Editor of the IEEE ACCESS in 2017. He is a reviewer of 500 manuscripts for over 50 different journals. He served as the Editor-in-Chief for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING from 2013 to 2017. He is an Associate Editor of the IEEE ACCESS. He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) from 2011 to 2012 and the President of the Spanish Chapter of the IEEE GRSS from 2012 to 2016.

Andrea Marinoni (S'07–M'11–SM'16) received the B.S., M.Sc. *(cum laude)*, and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 2005, 2007, and 2011, respectively.

From 2013 to 2018, he was a Research Fellow with the Telecommunications and Remote Sensing Laboratory, Department of Electrical, Computer, and Biomedical Engineering, University of Pavia. From 2015 to 2017, he was a Visiting Researcher with the Earth and Planetary Image Facility, Ben-Gurion University of the Negev, Beersheba, Israel; the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China; the School of Computer Science, Fudan University, Shanghai, China; the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China; and the Instituto de Telecomunicacoes, Instituto Superior Tecnico, Universidade de Lisboa, Lisbon, Portugal. He is currently an Associate Professor with the Earth Observation Laboratory, Department of Physics and Technology, UiT–The Arctic University of Norway, Tromsø, Norway. His research interests include efficient nonlinear signal processing applied to multimodal remote sensing, Earth observation interpretation, and Big Data mining, analysis, and management for human–environment interaction assessment.

Dr. Marinoni serves as a reviewer for several major journals and conferences sponsored by the IEEE Geoscience and Remote Sensing Society, the IEEE Communications Society, the IEEE Signal Processing Society, the IEEE Information Theory Society, the International Association of Pattern Recognition, the International Society on Urban Health, and the American Medical Informatics Association.

Paolo Gamba (SM'00–F'13) received the Laurea *(cum laude)* and Ph.D. degrees in electronic engineering from the University of Pavia, Pavia, Italy, in 1989 and 1993, respectively.

He is currently a Professor of telecommunications with the University of Pavia, where he leads the Telecommunications and Remote Sensing Laboratory and serves as a Deputy Coordinator of the Ph.D. School in Electronics and Computer Science. He has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, EO data, and risk management.

Dr. Gamba has been the organizer and Technical Chair of the biennial GRSS/ISPRS Joint Workshops on Remote Sensing and Data Fusion over Urban Areas since 2001. He was the Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from 2005 to 2009. He serves as a GRSS Vice President for Professional Activities, the Chair of the Chapters Committee, and the Chair of the Symposium Awards Committee. He was the Editor-in-Chief of IEEE GEOSCIENCE AND REMOTE SENSING LETTERS from 2009 to 2013. He has been the Guest Editor of special issues of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN REMOTE SENSING APPLICATIONS, the *ISPRS Journal of Photogrammetry and Remote Sensing*, the *International Journal of Information Fusion*, and the *Pattern Recognition Letters* on the topics of Urban Remote Sensing, Remote Sensing for Disaster Management, and Pattern Recognition in Remote Sensing Applications.

Somdatta Chakravortty received the B.Tech. degree in civil engineering from Kanpur University, Kanpur, India, in 1997, the M.Tech. degree in information technology from Bengal Engineering and Science University, Sibpur, India, in 2002, and the Ph.D. degree in computer science and engineering from Calcutta University, Kolkata, India, in 2017.

She is currently an Assistant Professor with the Department of Information Technology, Maulana Abul Kalam Azad University of Technology, Kolkata, India. She has several research projects funded by the Department of Science and Technology, University Grants Commission, and All India Council of Technical Education, Government of India, which focuses on research related to hyperspectral image analysis. Her research interests include hyperspectral image analysis, algorithm development, and image fusion.

Dr. Chakravortty is a member of professional societies such as the Indian Society of Remote Sensing, the Computer Society of India, and the Institute of Engineers.