Chance-Constrained Robust Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing

ArulMurugan Ambikapathi, *Student Member, IEEE*, Tsung-Han Chan, *Member, IEEE*,
Wing-Kin Ma, *Member, IEEE*, and Chong-Yung Chi, *Senior Member, IEEE*

Abstract—**Effective unmixing of hyperspectral data cube under a noisy scenario has been a challenging research problem in remote sensing arena. A branch of existing hyperspectral unmixing algorithms is based on Craig's criterion, which states that the vertices of the minimum-volume simplex enclosing the hyperspectral data should yield high fidelity estimates of the endmember signatures associated with the data cloud. Recently, we have developed a minimum-volume enclosing simplex (MVES) algorithm based on Craig's criterion and validated that the MVES algorithm is very useful to unmix highly mixed hyperspectral data. However, the presence of noise in the observations expands the actual data cloud, and as a consequence, the endmember estimates obtained by applying Craig-criterion-based algorithms to the noisy data may no longer be in close proximity to the true endmember signatures. In this paper, we propose a robust MVES (RMVES) algorithm that accounts for the noise effects in the observations by employing chance constraints. These chance constraints in turn control the volume of the resulting simplex. Under the Gaussian noise assumption, the chance-constrained MVES problem can be formulated into a deterministic nonlinear program. The problem can then be conveniently handled by alternating optimization, in which each subproblem involved is handled by using sequential quadratic programming solvers. The proposed RMVES is compared with several existing benchmark algorithms, including its predecessor, the MVES algorithm. Monte Carlo simulations and real hyperspectral data experiments are presented to demonstrate the efficacy of the proposed RMVES algorithm.**

Index Terms—**Abundance map, chance-constrained optimization, convex analysis, endmember signature, hyperspectral imaging (HI), hyperspectral unmixing (HU), sequential quadratic programming (SQP).**

A. Ambikapathi and T.-H. Chan are with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: aareul@ieee.org; thchan@ieee.org).

W.-K. Ma is with the Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: wkma@ieee.org).

C.-Y. Chi is with the Institute of Communications Engineering and the Department of Electrical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan (e-mail: cychi@ee.nthu.edu.tw).

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

I. INTRODUCTION

HYPERSPECTRAL imaging (HI) is a crucial technique to identify the materials and their composition in an area by exploiting the spectral diversity of the observed hyperspectral data. The areas in which HI is employed are diverse, and they include mineral identification [1], space object identification [2], analytical chemistry [3], retinal analysis [4], and many others. In this paper, we focus on HI for mineral identification and quantification, wherein a hyperspectral sensor (usually located on an aircraft or satellite) records the electromagnetic scattering patterns of materials present in an area, over hundreds of spectral bands that range from visible to near-infrared wavelength region. The limited spatial resolution of the sensor used for HI and the presence of noise in the measured hyperspectral data demand an effective hyperspectral unmixing (HU) scheme to extract the underlying endmember signatures (or simply endmembers) and the associated abundance maps (or abundance fractions) distributed over a scene of interest. The endmember signature corresponds to the reflection pattern of a mineral in different wavelengths, and the abundance fraction is the fractional distribution of a mineral over the given scene. The design of HU algorithms often (but not always) involves dimension reduction as the preprocessing step. The dimension reduction algorithms are intended to reduce the complexity of the unmixing algorithm that will be used in the sequel, and to some extent, they also aid in mitigating the noise effects in the data cloud. Conventional dimension reduction algorithms for hyperspectral data typically include principal component analysis [5] and maximum noise fraction [6]. A detailed survey of various dimension reduction algorithms can be found in [7]. However, for these algorithms (and also for the unmixing algorithms), the number of endmembers must be known *a priori*. A Neyman–Pearson-detection-theory-based eigenvalue thresholding method, namely, virtual dimensionality (VD) [8], is generally used for the estimation of the number of endmembers in the hyperspectral observations. Other algorithms that can be used to estimate the number of endmembers include Akaike information criterion [9] and minimum description length [10]. Hyperspectral subspace identification (HySiMe) [11] is a recently proposed dimension reduction algorithm that can also estimate the number of endmembers.

Conventional HU algorithms are based on linear mixing model (to be explained later) and can be classified into two main categories. The algorithms in the first category are based on the existence of pure pixels (pixels that are fully contributed by
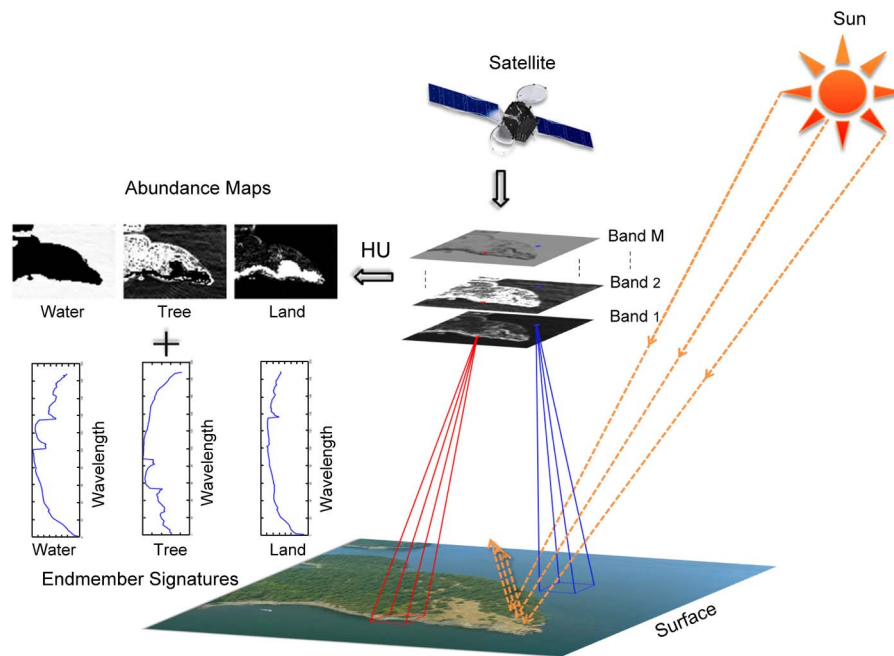
Fig. 1. Illustration of HU.

a single endmember) in the given hyperspectral observations, and those in the second category may not require the existence of pure pixels. Fig. 1 shows the notion of pure pixels and mixed pixels in an HI scenario. The red pixel corresponds to a mixed pixel (contributed by land, vegetation, and water), and the blue pixel corresponds to a pure pixel (contributed by only water). Pure-pixel-based algorithms include pixel purity index [12], N-finder (N-FINDR) [13], the simplex growing algorithm (SGA) [14], [15], vertex component analysis (VCA) [16], and alternating volume maximization [17], to name a few. All the pure-pixel-based algorithms aim to find the purest pixels in the given observations, either directly or indirectly, and hence can identify only the endmember signatures. However, the fully constrained least squares (FCLS) algorithm [18] can be used to estimate the abundance maps from the endmember estimates and the observations. The algorithms in the second category include minimum-volume transform [19], iterated constrained endmembers (ICEs) [20], piecewise convex endmember [21], alternating projected subgradients (APSs) [22], minimum-volume constrained nonnegative matrix factorization (MVC-NMF) [23], minimum-dispersion constrained nonnegative matrix factorization [24], minimum-volume simplex analysis (MVSA) [25], and minimum-volume enclosing simplex (MVES) [26], to name a few. The MVSA and MVES algorithms are directly based on Craig's criterion [19], which states that the vertices of the minimum-volume simplex enclosing the data cloud should yield high fidelity estimates of its endmember signatures. Apart from the aforementioned algorithms, there are certain other categories of algorithms that are based on spatial processing and morphological operators [27], [28]. For a recent survey on the various existing unmixing algorithms, please refer to [29].

Generally speaking, HU algorithms that do not require the existence of pure pixels in the data set are more computationally complicated than pure-pixel-based algorithms. On the other hand, the presence of pure pixels cannot be guaranteed for real hyperspectral data that are acquired under poor spatial

resolutions. Hence, there exists a tradeoff between the two categories of algorithms, in terms of computational tractability and accuracy. Nevertheless, the performance of all the HU algorithms degrades when the observations are noisy.

In recent years, several concurrent works on linear unmixing of noisy observations have been reported. Those works include the joint Bayesian algorithm (JBA) [30] and simplex identification via split augmented Lagrangian (SISAL) [31]. The JBA [30] is a statistical HU algorithm that assumes conjugate prior distributions for the abundances and endmember parameters. The endmember signatures are estimated by using a hierarchical Bayesian model, which generates the posterior distributions of both the abundances and the endmember parameters. The SISAL algorithm [31] is based on Craig's criterion, and it employs variable splitting and the augmented Lagrangian approach to estimate the minimum-volume simplex. Furthermore, it uses soft constraints to mitigate the effects of the outlier pixels.

In this paper, we propose a robust MVES (RMVES) algorithm for a general case, where uniform/nonuniform additive Gaussian noise is present in the observations. We begin with a linear mixing model and suitably modify an existing dimension reduction method, called affine set fitting [26], so as to account for the presence of noise. A toy example shown in Fig. 2 demonstrates how Craig's criterion may yield inaccurate endmember estimates for the noisy observations (shown as dots).

Clearly, the random noise expands the data cloud, and consequently, the volume (vertices) of the simplex formed by Craig's criterion (shown in dashed line) is larger than (far from) that of the true simplex (shown in solid line). Motivated by this, we formulate the RMVES problem by incorporating probability constraints or chance constraints [32] into Craig-criterion-based MVES problem [26], so as to deal with the effect of random noise. Under the Gaussian noise assumption, the chance-constrained RMVES problem can be conveniently formulated into a deterministic nonlinear program, where the
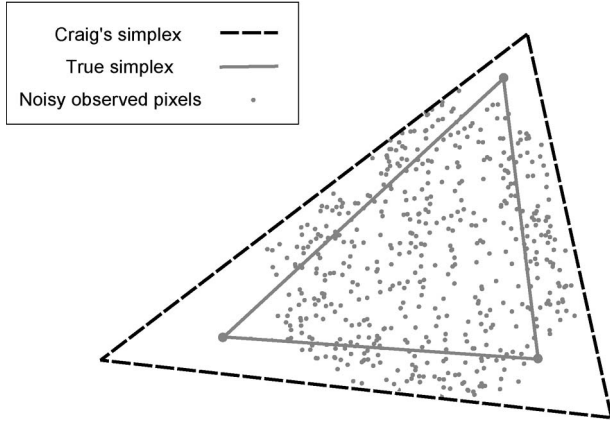
Fig. 2. Illustration of Craig's simplex for noisy observations, where the number of endmembers is equal to three.

preassigned probability given in the chance constraints when coupled with the noise covariance controls the volume of Craig's simplex. We then handle the resulting problem by alternating optimization, where each subproblem involved therein is specifically handled by readily available sequential quadratic programming (SQP) solvers. Some existing benchmark HU methods, including the MVES algorithm, are compared with the proposed RMVES algorithm through Monte Carlo simulations so as to demonstrate its efficacy. The RMVES algorithm, along with MVES and VCA, is applied to real hyperspectral data, and some inferences regarding the obtained endmember signatures and abundance maps are also discussed.

The forthcoming sections are organized as follows. The linear mixing model and the general assumptions are detailed in Section II. The dimension reduction technique employed for uniform and nonuniform Gaussian noises is discussed in Section III. In Section IV, we briefly summarize the MVES algorithm [26], which is developed based on Craig's criterion [19]. In Section V, we present the robust reformulation of the MVES algorithm by applying chance constraints on the noisy data, and then, we develop a robust HU algorithm, namely, the RMVES algorithm. Monte Carlo simulations and real hyperspectral data experiments are given in Section VI and Section VII, respectively. Finally, Section VIII provides some conclusions and future directions.

In the remainder of this paper, the following notations are employed. $\mathbb{R}^M$ ($\mathbb{R}^M_+$) and $\mathbb{R}^{M \times N}$ ($\mathbb{R}^{M \times N}_+$) represent a set of real (nonnegative real) $M \times 1$ vectors and real (nonnegative real) $M \times N$ matrices, respectively. $\mathbf{A}^\dagger$ represents pseudoinverse of a matrix $\mathbf{A}$. $\mathbf{1}_N$ and $\mathbf{0}$ represent an $N \times 1$ all-one vector and an all-zero column vector of proper dimension, respectively. A Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted as $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. The symbol A_i denotes the (i, i)th diagonal element of the matrix $\mathbf{A}$. $\|\mathbf{a}\|$ represents the Euclidean norm of vector $\mathbf{a}$, and $\mathbf{a} \succeq \mathbf{b}$ means that every component in $\mathbf{a}$ is larger than or equal to the corresponding component in $\mathbf{b}$. Finally, $\det(\mathbf{B})$ represents the determinant of a square matrix $\mathbf{B}$.

II. LINEAR MIXING MODEL AND ASSUMPTIONS

The linear mixing model is the one in which each pixel in the observations is obtained via single reflection of the end-

members present in that location. Such a linear mixing model is conventionally used in HU [1], [13], [16], [17], [33], [34]. We consider a scenario in which a hyperspectral sensor measures solar electromagnetic radiation from N distinct substances. Owing to low spatial resolution, each observed pixel vector represents a mixture of multiple distinct substances. Hence, each pixel vector of the hyperspectral images measured over M spectral bands can be represented by the following $M \times N$ linear mixing model

$$\mathbf{y}[n] = \mathbf{x}[n] + \mathbf{w}[n] \tag{1}$$

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i \qquad \forall\, n = 1, \ldots, L. \tag{2}$$

In (1), $\mathbf{y}[n] = [y_1[n], \ldots, y_M[n]]^\mathrm{T}$ represents the nth observed pixel vector (or simply pixel for convenience) comprising M spectral bands, $\mathbf{x}[n] = [x_1[n], \ldots, x_M[n]]^\mathrm{T}$ corresponds to its noise-free counterpart, and $\mathbf{w}[n] = [w_1[n], \ldots, w_M[n]]^\mathrm{T}$ is the noise vector. In (2), $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$ denotes the endmember signature matrix with the ith column vector $\mathbf{a}_i$ being the ith endmember signature, $\mathbf{s}[n] = [s_1[n], \ldots, s_N[n]]^\mathrm{T} \in \mathbb{R}^N$ is the nth abundance vector comprising N fractional abundances, and L is the total number of observed pixels. The noise vector $\mathbf{w}[n]$ considered in the signal model (1) is zero-mean uniform/nonuniform additive Gaussian noise vector, i.e., $\mathcal{N}(\mathbf{0}, \mathbf{D})$, where $\mathbf{D} = \mathrm{diag}(\sigma_1^2, \ldots, \sigma_M^2) \in \mathbb{R}^{M \times M}_+$ is a diagonal matrix and σ_i^2 denotes the noise variance in the ith spectral band. If $\sigma_i^2 = \sigma_j^2, \forall i \neq j$, then it is called uniform Gaussian noise; else, it is called nonuniform Gaussian noise. The additive Gaussian noise is a reasonable assumption and is widely used in designing HU algorithms [1], [16], [30].

Assuming prior knowledge about the number of endmembers N, we aim to estimate the endmember signature matrix $\mathbf{A}$ and the abundances $\mathbf{s}[1], \ldots, \mathbf{s}[L]$ from the noisy pixels $\mathbf{y}[1], \ldots, \mathbf{y}[L]$ under the following general assumptions [17], [33].

(A1) The intensities of all the abundance vectors are nonnegative, i.e., $s_i[n] \geq 0$, for all i and n.

(A2) Abundance fractions are proportionally distributed in each observed pixel, i.e., $\sum_{i=1}^{N} s_i[n] = 1, \forall n$.

(A3) $\min\{L, M\} \geq N$, and the endmembers are linearly independent, i.e., $\mathbf{A}$ is of full column rank.

In the ensuing development, we employ two convex analysis concepts, namely, *affine hull* and *convex hull* [32]. While the affine hull is employed for dimension reduction, the convex hull is used to infer the geometry of the observations. For ease of later use, they are defined as follows.

- The *affine hull* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is defined as

$$\mathrm{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \;\middle|\; \mathbf{1}_N^\mathrm{T} \boldsymbol{\theta} = 1, \boldsymbol{\theta} \in \mathbb{R}^N \right\} \tag{3}$$

where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^\mathrm{T}$. An affine hull can be represented as

$$\mathrm{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \mathcal{A}(\mathbf{C}, \mathbf{d}) = \{\mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \mid \boldsymbol{\alpha} \in \mathbb{R}^P\} \tag{4}$$

for some (nonunique) $\mathbf{d} \in \mathbb{R}^M$ and $\mathbf{C} \in \mathbb{R}^{M \times P}$, where $P \leq N - 1$ is the affine dimension. If
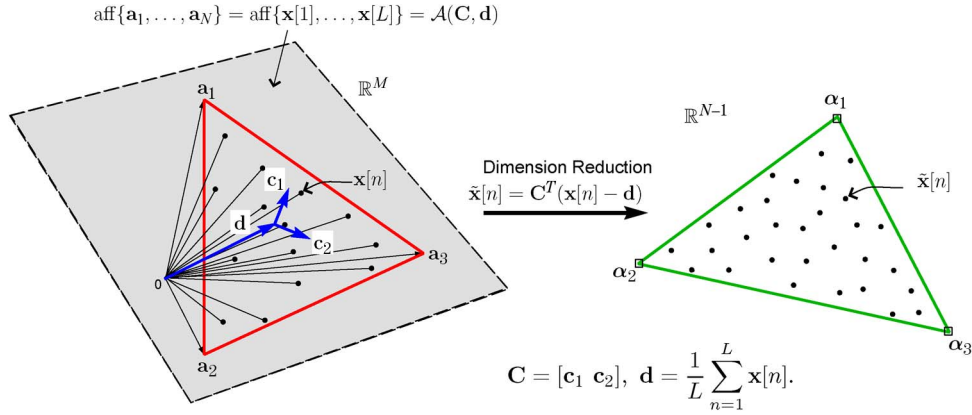
Fig. 3. Illustration of dimension reduction by affine set fitting for $N = 3$.

$\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ is affinely independent (i.e., the vectors $\mathbf{a}_1 - \mathbf{a}_N, \ldots, \mathbf{a}_{N-1} - \mathbf{a}_N$ are linearly independent), then the affine dimension $P = N - 1$.

- The *convex hull* of the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is defined as

$$\operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \,\middle|\, \mathbf{1}_N^{\mathrm{T}} \boldsymbol{\theta} = 1, \boldsymbol{\theta} \in \mathbb{R}_+^N \right\}. \tag{5}$$

A convex hull $\operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ is called an $N - 1$ dimensional simplex in $\mathbb{R}^M$ if $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is affinely independent and is called the *simplest simplex* when $M = N - 1$.

III. DATA PREPROCESSING

In this section, the dimension reduction technique employed for the observations corrupted by uniform and nonuniform Gaussian noise is discussed. For the noisy scenario, we employ the dimension reduction by using a variant of affine set fitting procedure introduced in [26]. To begin with, the affine set fitting procedure for noise-free data is reviewed in the following for convenience.

It can be readily inferred from the signal model (2), assumptions (A2) and (A3), and from (3) that $\mathbf{x}[n] \in \operatorname{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$. Furthermore, it is shown in [26] that the affine hull of the observations is the same as the affine hull of the endmember signatures, i.e., $\operatorname{aff}\{\mathbf{x}[1], \ldots, \mathbf{x}[L]\} = \operatorname{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$, which, in turn, can be represented by the affine set fitting parameters $\mathbf{C}$ and $\mathbf{d}$ in (4). Such representation aids in dimension reduction of the observed data. The dimension-reduced pixel vector $\tilde{\mathbf{x}}[n]$ can be obtained by the following affine transformation of $\mathbf{x}[n]$:

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^{\mathrm{T}} \left(\mathbf{x}[n] - \mathbf{d} \right) \in \mathbb{R}^{N-1} \qquad \forall n = 1, \ldots, L \tag{6}$$

where $(\mathbf{C}, \mathbf{d})$ is an affine set fitting solution given by [26]

$$\mathbf{d} = \frac{1}{L} \sum_{n=1}^{L} \mathbf{x}[n] \tag{7}$$

$$\mathbf{C} = \left[\boldsymbol{q}_1 \left(\mathbf{U}_x \mathbf{U}_x^{\mathrm{T}} \right), \ldots, \boldsymbol{q}_{N-1} \left(\mathbf{U}_x \mathbf{U}_x^{\mathrm{T}} \right) \right] \tag{8}$$

where

$$\mathbf{U}_x = [\mathbf{x}[1] - \mathbf{d}, \ldots, \mathbf{x}[L] - \mathbf{d}] \in \mathbb{R}^{M \times L} \tag{9}$$

is the mean-removed data matrix and $\boldsymbol{q}_i(\mathbf{R})$ denotes the unit-norm eigenvector associated with the ith principal eigenvalue of the matrix $\mathbf{R}$.

It should be mentioned that under (A2) and (A3), the aforementioned dimension reduction procedure is a lossless transformation in the absence of noise, and the dimension of the data is reduced from M to $N - 1$. An illustration of dimension reduction via the affine set fitting procedure is given in Fig. 3, where $N = 3$.

Since $\sum_{j=1}^{N} s_j[n] = 1$ (A2), it follows by substituting the noise-free signal model (2) into (6) that

$$\tilde{\mathbf{x}}[n] = \sum_{j=1}^{N} s_j[n] \boldsymbol{\alpha}_j \qquad \forall n = 1, \ldots, L \tag{10}$$

where

$$\boldsymbol{\alpha}_j = \mathbf{C}^{\mathrm{T}} (\mathbf{a}_j - \mathbf{d}) \in \mathbb{R}^{N-1} \tag{11}$$

is the jth dimension-reduced endmember signature. Moreover, due to $s_i[n] \geq 0$ (A1), it can be seen that

$$\tilde{\mathbf{x}}[n] \in \operatorname{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} \subset \mathbb{R}^{N-1} \qquad \forall n \tag{12}$$

and $\operatorname{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ is the simplest simplex. This can be regarded as the outcome of the fact that the affine transformation of a simplex is also a simplex, as one can infer from (2) that, for all n, $\mathbf{x}[n] \in \operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$, and $\operatorname{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is itself a simplex [by (A3)].

However, when the data are corrupted by noise, the observations can no longer lie in a single specific affine set. Accordingly, for the noisy scenario, the approximate affine set fitting parameters ($\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$) can be obtained as explained below.

Similar to (9), by (1), the mean-removed noisy data matrix is

$$\mathbf{U}_y = [\mathbf{y}[1] - \mathbf{d}, \ldots, \mathbf{y}[L] - \mathbf{d}] = \mathbf{U}_x + \mathbf{W} \in \mathbb{R}^{M \times L} \tag{13}$$

where $\mathbf{W} = [\mathbf{w}[1], \ldots, \mathbf{w}[L]]$. Then

$$\mathbf{U}_y \mathbf{U}_y^{\mathrm{T}} = \mathbf{U}_x \mathbf{U}_x^{\mathrm{T}} + \mathbf{U}_x \mathbf{W}^{\mathrm{T}} + \mathbf{W}^{\mathrm{T}} \mathbf{U}_x + \mathbf{W} \mathbf{W}^{\mathrm{T}}$$

$$\cong \mathbf{U}_x \mathbf{U}_x^{\mathrm{T}} + \mathbf{W} \mathbf{W}^{\mathrm{T}} = \mathbf{U}_x \mathbf{U}_x^{\mathrm{T}} + L \widehat{\mathbf{D}} \tag{14}$$

which is because the matrices $\mathbf{U}_x \mathbf{W}^{\mathrm{T}}$ and $\mathbf{W} \mathbf{U}_x^{\mathrm{T}}$ asymptotically approach zero matrix for large L (since the noise is zero

mean and independent of the noise-free observations), and $\widehat{\mathbf{D}}$ is defined as

$$\widehat{\mathbf{D}} = \frac{1}{L} \mathbf{W} \mathbf{W}^{\mathrm{T}} \qquad (15)$$

which is actually an estimate of the noise covariance matrix $\mathbf{D}$. By replacing $\mathbf{x}[n]$ with $\mathbf{y}[n]$ in (7) and replacing $\mathbf{U}_x \mathbf{U}_x^{\mathrm{T}}$ with $\mathbf{U}_y \mathbf{U}_y^{\mathrm{T}} - L\widehat{\mathbf{D}}$ [from (14)] in (8), respectively, we have the following equations for $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$ as approximations for $\mathbf{C}$ and $\mathbf{d}$, respectively

$$\widehat{\mathbf{d}} = \frac{1}{L} \sum_{n=1}^{L} \mathbf{y}[n] = \frac{1}{L} \sum_{n=1}^{L} \mathbf{x}[n] + \frac{1}{L} \sum_{n=1}^{L} \mathbf{w}[n] \cong \mathbf{d} \qquad (16)$$

$$\widehat{\mathbf{C}} = \left[\boldsymbol{q}_1 \left(\mathbf{U}_y \mathbf{U}_y^{\mathrm{T}} - L\widehat{\mathbf{D}} \right), \ldots, \boldsymbol{q}_{N-1} \left(\mathbf{U}_y \mathbf{U}_y^{\mathrm{T}} - L\widehat{\mathbf{D}} \right) \right]$$
$$\cong \mathbf{C}. \qquad (17)$$

Note that (16) and (17) hold since the noise is of zero mean such that $(1/L)\sum_{n=1}^{L} \mathbf{w}[n]$ asymptotically approaches zero vector for large L and $\widehat{\mathbf{D}}$ approaches the true $\mathbf{D}$ for large L, respectively. In this paper, the multiple-regression-analysis-based noise covariance estimation method reported in HySiMe [11] is employed to obtain $\widehat{\mathbf{D}}$. With the given $\widehat{\mathbf{D}}$, the affine set fitting solution $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ of the noisy data [given by (16) and (17)] serves as an approximation to the true $(\mathbf{C}, \mathbf{d})$, and it asymptotically approaches the true $(\mathbf{C}, \mathbf{d})$ for large L.

The obtained $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$ are then used to obtain the dimension-reduced noisy observations, as given by

$$\tilde{\mathbf{y}}[n] = \widehat{\mathbf{C}}^{\mathrm{T}} \left(\mathbf{y}[n] - \widehat{\mathbf{d}} \right) = \widehat{\mathbf{C}}^{\mathrm{T}} \left(\mathbf{x}[n] - \widehat{\mathbf{d}} \right) + \widehat{\mathbf{C}}^{\mathrm{T}} \mathbf{w}[n]$$
$$\cong \tilde{\mathbf{x}}[n] + \tilde{\mathbf{w}}[n] \qquad (18)$$

where

$$\tilde{\mathbf{w}}[n] = \widehat{\mathbf{C}}^{\mathrm{T}} \mathbf{w}[n] \qquad (19)$$

and the approximation in (18) is because the affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ for the noisy data serve as a good approximation to the true $(\mathbf{C}, \mathbf{d})$ [see (16) and (17)].

IV. BRIEF REVIEW OF MVES ALGORITHM

As mentioned in Section I, for the hyperspectral data with noise, the aim is to fit a simplex for the observed data, whose volume is less than that of the simplex volume obtained by Craig's criterion, so that the resultant simplex will be close to the true simplex. For ease of continuity and understanding, the robust algorithm to be presented in Section V will be derived from Craig-criterion-based MVES algorithm [19], which is briefly summarized in this section. The MVES algorithm aims to solve the following optimization problem [17], [26]:

$$\begin{aligned} \min_{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \text{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\} \qquad \forall n. \end{aligned} \qquad (20)$$

Here, $\tilde{\mathbf{x}}[n]$ is the dimension-reduced noise-free observation, $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N$ correspond to the dimension-reduced endmember

signatures, and $V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N)$ is the volume of the simplex $\text{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$, given by [35]

$$V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N) = \frac{|\det(\mathbf{B})|}{(N-1)!} \qquad (21)$$

where $(N-1)!$ represents the factorial of $N-1$ and

$$\mathbf{B} = [\boldsymbol{\beta}_1 - \boldsymbol{\beta}_N, \ldots, \boldsymbol{\beta}_{N-1} - \boldsymbol{\beta}_N] \in \mathbb{R}^{(N-1)\times(N-1)}. \qquad (22)$$

By (22) and (5), the constraints of (20) can be expressed as

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n] \boldsymbol{\beta}_i = \mathbf{B} \mathbf{s}'[n] + \boldsymbol{\beta}_N \qquad \forall n \qquad (23)$$

where

$$\mathbf{s}'[n] = [s_1[n], \ldots, s_{N-1}[n]]^{\mathrm{T}} \in \mathbb{R}_+^{N-1} \qquad (24a)$$
$$s_N[n] = 1 - \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{s}'[n] \geq 0 \qquad (24b)$$

are the abundance fractions of the nth pixel in the observations. Note that (24a) and (24b) jointly enforce the nonnegativity and full-additivity constraints on the abundances. Equation (23) can be rewritten as

$$\mathbf{s}'[n] = \mathbf{B}^{-1} \left(\tilde{\mathbf{x}}[n] - \boldsymbol{\beta}_N \right) = \mathbf{H} \tilde{\mathbf{x}}[n] - \mathbf{g} \qquad \forall n \qquad (25)$$

where

$$\mathbf{H} = \mathbf{B}^{-1} \qquad (26a)$$
$$\mathbf{g} = \mathbf{B}^{-1} \boldsymbol{\beta}_N. \qquad (26b)$$

Substituting (25) into (24) gives rise to the following equivalent constraints:

$$s_i[n] = \mathbf{h}_i^{\mathrm{T}} \tilde{\mathbf{x}}[n] - g_i \geq 0,$$
$$\forall i = 1, \ldots, N-1; \quad n = 1, \ldots, L \qquad (27a)$$
$$s_N[n] = 1 - \mathbf{1}_{N-1}^{\mathrm{T}} \left(\mathbf{H} \tilde{\mathbf{x}}[n] - \mathbf{g} \right) \geq 0,$$
$$\forall n = 1, \ldots, L \qquad (27b)$$

where $\mathbf{h}_i^{\mathrm{T}}$ is the ith row vector of $\mathbf{H}$ and g_i is the ith element of $\mathbf{g}$. Hence, the minimization problem in (20) is equivalent to the following maximization problem:

$$\max_{\substack{\mathbf{H} \in \mathbb{R}^{(N-1)\times(N-1)}, \\ \mathbf{g} \in \mathbb{R}^{N-1}}} \quad |\det(\mathbf{H})| \qquad (28a)$$
$$\text{s.t.} \quad \mathbf{h}_i^{\mathrm{T}} \tilde{\mathbf{x}}[n] - g_i \geq 0 \qquad \forall i, n \qquad (28b)$$
$$1 - \mathbf{1}_{N-1}^{\mathrm{T}} \left(\mathbf{H} \tilde{\mathbf{x}}[n] - \mathbf{g} \right) \geq 0 \qquad \forall n. \qquad (28c)$$

After some simplifications, the nonconvex optimization problem (28) is then handled via alternating optimization, where each and every subproblem is convex (in fact, linear program), and can be solved by using available convex optimization solvers such as SeDuMi [36] and CVX [37]. Let us emphasize that the MVES algorithm in [26] considers a noise-free scenario, and hence, for noisy hyperspectral data, the solution, i.e., the simplex $\text{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$, obtained by MVES may be far from the true one.

V. RMVES ALGORITHM

In this section, we develop a robust HU algorithm that aims to account for the presence of noise in the observations. The presence of noise in the observations expands the data cloud (even after dimension reduction). Hence, for noisy data, the minimum-volume simplex estimated by (28) will be larger than the true simplex, and therefore, its vertices (endmember signatures) will be away from those of the true simplex (as shown in Fig. 2). Given the noisy observations, our aim now is to find a minimum-volume simplex that can be closer to the true simplex. This means that we allow some pixels in the noisy data cloud to be outside the estimated simplex.

From (18), we have $\tilde{\mathbf{x}}[n] \cong \tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]$. Substituting this for $\tilde{\mathbf{x}}[n]$ in (28), we have

$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} |\det(\mathbf{H})| \tag{29a}$$

$$\text{s.t.} \quad \mathbf{h}_i^{\mathrm{T}}(\tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]) - g_i \geq 0 \qquad \forall i, n \tag{29b}$$

$$1 - \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}(\tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]) + \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{g} \geq 0 \\ \forall n. \tag{29c}$$

However, $\tilde{\mathbf{w}}[n] = \widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{w}[n]$ is an unknown random vector. We therefore need to consider this uncertainty in the constraints of (29). In the ensuing development, we will make use of chance constraints or probabilistic constraints [32] that can account for the randomness in the observations. A chance-constrained counterpart of problem (29) is proposed as follows:

$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} |\det(\mathbf{H})| \tag{30a}$$

$$\text{s.t.} \quad \Pr\left(u_i[n] \leq \mathbf{h}_i^{\mathrm{T}}\tilde{\mathbf{y}}[n] - g_i\right) \geq \eta \qquad \forall i, n \tag{30b}$$

$$\Pr\left(z[n] \leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}\tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{g}\right) \geq \eta \\ \forall n \tag{30c}$$

where $\Pr(\cdot)$ represents the probability; $u_i[n] \triangleq \mathbf{h}_i^{\mathrm{T}}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{w}[n]$ is a random variable with the distribution $\mathcal{N}(0, \mathbf{h}_i^{\mathrm{T}}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{D}\widehat{\mathbf{C}}\mathbf{h}_i)$, $i = 1, \ldots, N-1$; $z[n] \triangleq -\mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{w}[n]$ is a random variable with the distribution $\mathcal{N}(0, \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{D}\widehat{\mathbf{C}}\mathbf{H}^{\mathrm{T}}\mathbf{1}_{N-1})$; and $\eta \in (0, 1)$ is a design parameter. Note that, in (30), the hard constraints in (29), i.e., (29b) and (29c) are replaced by the soft constraints (30b) and (30c), respectively. Specifically, in (30), we only require that the constraints (29b) and (29c) hold true with probability not less than η.

The chance constraints in (30) can be further simplified by normalizing the random variables involved [32]. Specifically, for a random variable $\varepsilon \sim \mathcal{N}(\mu, \delta^2)$ and $t \in \mathbb{R}$, $\Pr(\varepsilon \leq t) \geq \eta$ is true as $t \geq \delta\Phi^{-1}(\eta) + \mu$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal random variable (Gaussian random variable with zero mean and unit variance), and $\Phi^{-1}(\cdot)$ is the inverse of $\Phi(\cdot)$. Applying this procedure to the constraints of (30), we have the following RMVES problem:

$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} |\det(\mathbf{H})| \tag{31a}$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \leq \mathbf{h}_i^{\mathrm{T}}\tilde{\mathbf{y}}[n] - g_i \qquad \forall i, n \tag{31b}$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{Q}\mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}\tilde{\mathbf{y}}[n] \\ + \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{g} \qquad \forall n \tag{31c}$$

where $\mathbf{Q} = \mathbf{H}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{D}\widehat{\mathbf{C}}\mathbf{H}^{\mathrm{T}}$ and $Q_i = \mathbf{h}_i^{\mathrm{T}}\widehat{\mathbf{C}}^{\mathrm{T}}\mathbf{D}\widehat{\mathbf{C}}\mathbf{h}_i$. Let the constraint set of (31) be

$$\mathcal{F}(\eta) = \left\{(\mathbf{H}, \mathbf{g}) | \Phi^{-1}(\eta)\sqrt{Q_i} \leq \mathbf{h}_i^{\mathrm{T}}\tilde{\mathbf{y}}[n] - g_i \forall i, n, \Phi^{-1}(\eta) \right. \\ \left. \times \sqrt{\mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{Q}\mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{H}\tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^{\mathrm{T}}\mathbf{g}, \forall n \right\}. \tag{32}$$

As $\Phi^{-1}(\eta)$ is a monotone increasing function of η, one can infer that

$$\mathcal{F}(\eta_1) \subseteq \mathcal{F}(\eta_2) \qquad \forall \eta_1 \geq \eta_2. \tag{33}$$

Thus, the optimal $|\det(\mathbf{H})|$ will be larger for smaller η. Also note from (26a) that an increased value of $|\det(\mathbf{H})|$ corresponds to a decreased value of $|\det(\mathbf{B})| = 1/|\det(\mathbf{H})|$ (i.e., a smaller simplex). It can be readily seen that, when $\eta = 0.5$, the constraint set $\mathcal{F}(0.5)$ of (31) is identical to that of (28) (as $\Phi^{-1}(\eta = 0.5) = 0$) with $\tilde{\mathbf{x}}[n]$ replaced by $\tilde{\mathbf{y}}[n]$. Simply speaking, for noisy observations, our aim is to find a simplex whose volume is minimized further when compared to that of the simplex obtained by Craig-criterion-based MVES algorithm (see Fig. 2). This, together with (33), confirms that the appropriate range of η for robust design should be between 0 and 0.5. It is also worthwhile to mention that, when $\eta = 0$, $|\det(\mathbf{H})|$ becomes unbounded as $\mathcal{F}(0) = \mathbb{R}^{(N-1)\times(N-1)} \times \mathbb{R}^{N-1}$ and, hence, $|\det(\mathbf{B})|$ becomes zero, and for the other extreme case of $\eta = 1$, problem (31) becomes infeasible (as could be inferred from (32)) since $\mathcal{F}(1)$ is an empty set.

Note that the values of η affect the feasible set of (31). When $\eta > 0.5$ (i.e., $\Phi^{-1}(\eta) > 0$), the constraints (31b) and (31c) are second-order cone constraints (convex). These constraints ensure that the corresponding abundances of each noisy observation are positive, rendering all the observations to be inside the estimated simplex. However, from (33), one can infer that the estimated simplex when $\eta > 0.5$ will be larger than the one estimated when $\eta = 0.5$ (simplex estimated by MVES). On the other hand, if $\eta < 0.5$ (i.e., $\Phi^{-1}(\eta) < 0$), the constraints (31b) and (31c) become nonconvex in $(\mathbf{H}, \mathbf{g})$. However, these constraints allow some negativity in the abundance fractions, by means of which some noisy pixels can be left out of the estimated simplex. Recall that, for the robust design, some noisy pixels must be left out of the simplex such that the estimated simplex can be close to the true simplex (as could be inferred from Fig. 2). Thus, we again conclude that the apt range of the design parameter η should be between 0 and 0.5.

Fig. 4 shows a scatter plot (for $N = 3$) of the dimension-reduced noisy observations and optimal simplexes $\text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ of the RMVES problem (31) for different values of η. As elaborately discussed earlier, one can see in Fig. 4 that, when $\eta < 0.5$, the solution of RMVES problem indeed approaches the true simplex (as some of the noisy pixels are left outside the simplex). When $\eta = 0.5$, the minimum-volume simplex tightly encloses all the observations, and
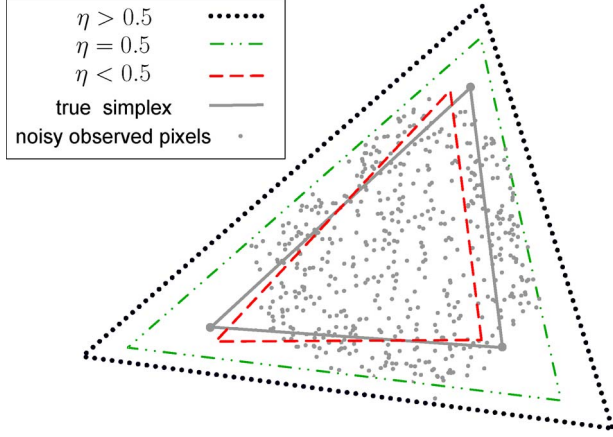
Fig. 4. Scatter plot of the dimension-reduced pixels for $N = 3$, illustrating the solutions of RMVES for different values of η.

when $\eta > 0.5$, the simplex expands but still encloses all the observations. Note that, for $\eta \geq 0.5$, the estimated simplex is away from the true simplex.

While the feasible set of (31) could be convex or nonconvex (depending on η), the objective function of (31) is always nonconvex. Inspired by our previous works [26], [38]–[40], we utilize the alternating optimization methodology, by virtue of which we form subproblems that are hopefully "less nonconvex" than (31). We consider the cofactor expansion for $\det(\mathbf{H})$ along the ith row of $\mathbf{H}$

$$\det(\mathbf{H}) = \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}) \qquad (34)$$

where $\boldsymbol{\mathcal{H}}_{ij}$ is the submatrix of $\mathbf{H}$ with the ith row and the jth column removed. One can observe from (34) that $\det(\mathbf{H})$ is linear in each $\mathbf{h}_i^{\mathrm{T}}$, which enables us to update $\mathbf{h}_i^{\mathrm{T}}$ and g_i while fixing the other rows of $\mathbf{H}$ and the other entries of $\mathbf{g}$. Then, the partial maximization of (31) with respect to $\mathbf{h}_i^{\mathrm{T}}$ and g_i can be formulated as

$$\max_{\mathbf{h}_i^{\mathrm{T}}, g_i} \left| \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}) \right| \qquad (35\mathrm{a})$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i \qquad \forall n \qquad (35\mathrm{b})$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{g}$$
$$\forall n. \quad (35\mathrm{c})$$

The objective function in (35) is still nonconvex but can be handled by breaking it into the following two optimization problems:

$$p^{\star} = \max_{\mathbf{h}_i^{\mathrm{T}}, g_i} \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij})$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \leq \mathbf{h}_i^{\mathrm{T}} \mathbf{y}[n] - g_i \qquad \forall n \qquad (36\mathrm{a})$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{1}_{N-1}}$$
$$\leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{g} \qquad \forall n$$

$$q^{\star} = \min_{\mathbf{h}_i^{\mathrm{T}}, g_i} \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij})$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \leq \mathbf{h}_i^{\mathrm{T}} \tilde{\mathbf{y}}[n] - g_i \qquad \forall n \qquad (36\mathrm{b})$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{Q} \mathbf{1}_{N-1}}$$
$$\leq 1 - \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^{\mathrm{T}} \mathbf{g} \qquad \forall n.$$

As discussed earlier, the apt range of the design parameter is $\eta < 0.5$, and hence, unlike in [26], [38], and [40], the subproblems here (36a) and (36b) are nonconvex, which make problem (31) yet difficult to solve. SQP [41] is a well-known method to handle nonconvex problems, and here, we employ the available MATLAB-based SQP solver, namely, *fmincon* [42], to handle the subproblems (36a) and (36b). It should be noted that although the subproblems (36a) and (36b) are nonconvex, the objective functions are now linear (convex). From our simulation experience, we found that this strategy substantially aids in mitigating the local optimality issue associated with the problem (31) (this issue will be addressed later in Section VI-C). The optimal solution of (35), denoted by $(\mathbf{h}_i^{\star \mathrm{T}}, g_i^{\star})$, is chosen as an optimal solution of the maximization problem (36a) if $|p^{\star}| > |q^{\star}|$ and that of the minimization problem (36b) if $|q^{\star}| > |p^{\star}|$. This rowwise maximization is conducted cyclically (i.e., $i := (i \bmod (N-1)) + 1$ via each row update of $\mathbf{H}$). We define one iteration as one full update of the matrix $\mathbf{H}$ and the vector $\mathbf{g}$. At each iteration, if the relative change in $|\det(\mathbf{H})|$ between the current and the previous iteration exceeds a given threshold, then we continue with the next updating iteration; else, the updated $\mathbf{H}$ and $\mathbf{g}$ are the obtained estimates.

Suppose that a solution $(\mathbf{H}^{\star}, \mathbf{g}^{\star})$ is obtained by the aforementioned alternating maximization method. From (11), (26a), and (26b), the endmember signatures can then be recovered by $\hat{\mathbf{a}}_i = \mathbf{C} \hat{\boldsymbol{\alpha}}_i + \mathbf{d}$ for $i = 1, \ldots, N$, where

$$\hat{\boldsymbol{\alpha}}_N = (\mathbf{H}^{\star})^{-1} \mathbf{g}^{\star}, \qquad (37)$$

$$[\hat{\boldsymbol{\alpha}}_1, \ldots, \hat{\boldsymbol{\alpha}}_{N-1}] = \hat{\boldsymbol{\alpha}}_N \mathbf{1}_{N-1}^{\mathrm{T}} + (\mathbf{H}^{\star})^{-1}. \qquad (38)$$

Since some abundance fractions $s_i[n]$ can be negative for those pixels outside the simplex $\mathrm{conv}\{\hat{\boldsymbol{\alpha}}_1, \ldots, \hat{\boldsymbol{\alpha}}_N\}$, the abundance estimates $\hat{\mathbf{s}}[1], \ldots, \hat{\mathbf{s}}[L]$ are therefore obtained by using the FCLS algorithm [18], which can ensure the nonnegativity and full additivity of the estimated abundances. The previously illustrated procedure is collectively termed as the RMVES algorithm.

The proposed RMVES algorithm uses the well-known VCA [16] for the initialization of (35) [33]. The endmember estimates obtained by VCA are first expanded until all the dimension-reduced data are well within the simplex formed by the expanded endmember estimates. The expanded endmember estimates are then used to get the initial estimates of $\mathbf{H}$ [by (26a)] and $\mathbf{g}$ [by (26b)]. The pseudocode of the expanded VCA, which is used for RMVES initialization [25], is given in Table I. The pseudocode of the proposed RMVES algorithm is given in Table II.

We conclude this section by summarizing the key distinctions between the MVES algorithm [26] and the proposed RMVES algorithm. In contrast to the MVES algorithm that only considers hard constraints assuming the absence of noise, the RMVES

TABLE I
PSEUDOCODE FOR THE EXPANDED VCA INITIALIZATION [25]

Given	The affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ and the endmember matrix estimated by VCA [16], say $\mathbf{A}_0$.
Step 1.	Obtain the dimension-reduced VCA endmember estimates $\mathbf{E}_0 = \widehat{\mathbf{C}}^T(\mathbf{A}_0 - \widehat{\mathbf{d}}\mathbf{1}_N^T)$.
Step 2.	Subtract the mean vector $\boldsymbol{\mu}_{E_0} = \mathbf{E}_0\mathbf{1}_N(1/N)$ from $\mathbf{E}_0$, i.e., $\bar{\mathbf{E}} = \mathbf{E}_0 - \boldsymbol{\mu}_{E_0}\mathbf{1}_N^T$.
Step 3.	Repeat $\mathbf{E}_0 := \mathbf{E}_0 + k\bar{\mathbf{E}}$ until all the elements of $\mathbf{E}_0^\dagger\mathbf{y}[n]$ are non-negative for all $n = 1, \ldots, L$, where k is a constant (say $k = 5$).
Step 4.	The output $\mathbf{E}_0 \triangleq [\boldsymbol{\beta}_1^{(0)}, \ldots, \boldsymbol{\beta}_N^{(0)}]$ can then be used to obtain the initial $\mathbf{H}^{(0)} = [\boldsymbol{\beta}_1^{(0)} - \boldsymbol{\beta}_N^{(0)}, \ldots, \boldsymbol{\beta}_{N-1}^{(0)} - \boldsymbol{\beta}_N^{(0)}]^{-1}$ and $\mathbf{g}^{(0)} = \mathbf{H}^{(0)}\boldsymbol{\beta}_N^{(0)}$.

TABLE II
PSEUDOCODE FOR RMVES ALGORITHM

Given	The noisy observed data $\mathbf{y}[n]$ for $n = 1, \ldots, L$, the number of endmembers N, a design parameter η, and a convergence tolerance $\varepsilon > 0$.
Step 1.	Estimate the noise covariance matrix (denoted by $\widehat{\mathbf{D}}$) using multiple regression analysis [11].
Step 2.	**Dimension reduction**: Obtain the dimension-reduced pixels: $\tilde{\mathbf{y}}[n] = \widehat{\mathbf{C}}^T(\mathbf{y}[n] - \widehat{\mathbf{d}})$ for all n, with the affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ given by

$$\widehat{\mathbf{d}} = \frac{1}{L}\sum_{n=1}^{L}\mathbf{y}[n],$$
$$\widehat{\mathbf{C}} = [\,\boldsymbol{q}_1(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}}), \ldots, \boldsymbol{q}_{N-1}(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}})\,],$$

where $\mathbf{U}_y = [\,\mathbf{y}[1] - \widehat{\mathbf{d}}, \ldots, \mathbf{y}[L] - \widehat{\mathbf{d}}\,] \in \mathbb{R}^{M \times L}$ and $\boldsymbol{q}_i(\mathbf{R})$ denotes the unit-norm eigenvector associated with the ith principal eigenvalue of the matrix $\mathbf{R}$.

Step 3.	**Initialization**: Let $\mathbf{H} = \mathbf{H}^{(0)}$ and $\mathbf{g} = \mathbf{g}^{(0)}$, where $\mathbf{H}^{(0)}$ and $\mathbf{g}^{(0)}$ are given by the expanded VCA (as explained in Table I). Set $i := 1$ and $\varrho := |\det(\mathbf{H})|$.
Step 4.	Define $\mathbf{Q} = \mathbf{H}\widehat{\mathbf{C}}^T\widehat{\mathbf{D}}\widehat{\mathbf{C}}\mathbf{H}^T$ and let $\boldsymbol{\mathcal{H}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ denote the submatrix of $\mathbf{H}$ with the ith row and jth column removed. Then, handle the following problems

$$p^\star = \max_{\mathbf{h}_i^T, g_i} \ \sum_{j=1}^{N-1}(-1)^{i+j}h_{ij}\det(\boldsymbol{\mathcal{H}}_{ij})$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \le \mathbf{h}_i^T\tilde{\mathbf{y}}[n] - g_i, \ \forall n,$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^T\mathbf{Q}\mathbf{1}_{N-1}} \le 1 - \mathbf{1}_{N-1}^T\mathbf{H}\tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T\mathbf{g}, \ \forall n,$$

$$q^\star = \min_{\mathbf{h}_i^T, g_i} \ \sum_{j=1}^{N-1}(-1)^{i+j}h_{ij}\det(\boldsymbol{\mathcal{H}}_{ij})$$

$$\text{s.t.} \quad \Phi^{-1}(\eta)\sqrt{Q_i} \le \mathbf{h}_i^T\tilde{\mathbf{y}}[n] - g_i, \ \forall n,$$

$$\Phi^{-1}(\eta)\sqrt{\mathbf{1}_{N-1}^T\mathbf{Q}\mathbf{1}_{N-1}} \le 1 - \mathbf{1}_{N-1}^T\mathbf{H}\tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T\mathbf{g}, \ \forall n,$$

	using SQP and obtain their solutions, denoted by $(\bar{\mathbf{h}}_i^T, \bar{g}_i)$ and $(\breve{\mathbf{h}}_i^T, \breve{g}_i)$, respectively.
Step 5.	If $|p^\star| > |q^\star|$, update $(\mathbf{h}_i^T, g_i) := (\bar{\mathbf{h}}_i^T, \bar{g}_i)$, otherwise $(\mathbf{h}_i^T, g_i) := (\breve{\mathbf{h}}_i^T, \breve{g}_i)$.
Step 6.	If $(i \text{ modulo } (N-1)) \ne 0$, then $i := i + 1$, and go to **Step 4**, else
	If $|\max\{|p^\star|, |q^\star|\} - \varrho|/\varrho < \varepsilon$, then $\mathbf{H}^\star = \mathbf{H}$ and $\mathbf{g}^\star = \mathbf{g}$. Otherwise, set $\varrho := \max\{|p^\star|, |q^\star|\}$, $i := 1$, and go to **Step 4**.
Step 7.	Calculate the dimension reduced endmember estimates: $\widehat{\boldsymbol{\alpha}}_N = (\mathbf{H}^\star)^{-1}\mathbf{g}^\star$ and $[\widehat{\boldsymbol{\alpha}}_1, \ldots, \widehat{\boldsymbol{\alpha}}_{N-1}] = \widehat{\boldsymbol{\alpha}}_N\mathbf{1}_{N-1}^T + (\mathbf{H}^\star)^{-1}$.
Step 8.	Calculate the actual endmember estimates: $\widehat{\mathbf{a}}_i = \widehat{\mathbf{C}}\widehat{\boldsymbol{\alpha}}_i + \widehat{\mathbf{d}}$ for $i = 1, \ldots, N$.
Step 9.	Estimate the abundance vectors $\widehat{\mathbf{s}}[1], \ldots, \widehat{\mathbf{s}}[L]$ by using the FCLS algorithm [18].

algorithm employs soft constraints (chance constraints) to account for the presence of noise in the observed data. In addition, the RMVES algorithm uses the chance constraint parameter η to control the volume of the estimated simplex such that some noisy dimension-reduced data $\tilde{\mathbf{y}}[n]$ naturally fall outside the estimated simplex for $\eta < 0.5$, and it yields the same endmember estimates with the MVES algorithm for $\eta = 0.5$. While both algorithms employ the alternating optimization approach, the RMVES algorithm resorts to a series of nonlinear and nonconvex programs (handled via SQP) to obtain the desired endmember estimates, while the MVES algorithm solves a series of linear programs. Finally, in RMVES algorithm, the endmember estimates, along with the hyperspectral data, are used to estimate the abundances via FCLS algorithm, whereas the MVES algorithm yields endmember estimates and abundance estimates simultaneously.

VI. SIMULATIONS

In this section, the efficacy of the proposed RMVES algorithm is demonstrated. The results shown in this section are averaged over 50 Monte Carlo runs for each considered scenario. To show the robustness of the proposed algorithm, two different noisy scenarios are considered. The first scenario is for the observations corrupted with uniform Gaussian noise (Section VI-A). The performance is studied under different purity levels and for different signal-to-noise ratios (SNRs). The second scenario is that the observations are corrupted by nonuniform Gaussian noise (Section VI-B). In this case, we fix the purity level and evaluate the performance of the algorithms under test for different distributions of the nonuniform Gaussian noise variance over various SNR levels. Finally, in Section VI-C, we study the local optimality issue associated with handling the nonconvex problems (31) and (35) (via RMVES algorithm) to show the substantial performance improvement of the latter over the former. Several existing algorithms are used to compare the performance of the proposed RMVES algorithm. The pure-pixel-based methods used are N-FINDR [13], SGA [14], and VCA [16], and then, FCLS [18] is used to find the abundances associated with the obtained endmember estimates. The other methods used are ICE [20], APS [22], MVC-NMF [23], MVSA [25], SISAL [31], and MVES [26]. The parameter settings and algorithmic details of the algorithms under test are summarized in Table III (for other parameters of the algorithms not listed in the table, their default values mentioned in their respective references are used). The root-mean-square (rms) spectral angles, denoted as ϕ_{en} (ϕ_{ab}) between the true endmembers (abundance maps) and estimated endmembers (abundance maps) (which have been widely used in HU [1], [16], [26]), are used as the performance indices, which are defined as follows:

$$\phi_{en} = \min_{\boldsymbol{\pi}\in\Pi_N}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\arccos\left(\frac{\mathbf{a}_i^T\widehat{\mathbf{a}}_{\pi_i}}{\|\mathbf{a}_i\|\,\|\widehat{\mathbf{a}}_{\pi_i}\|}\right)\right]^2} \tag{39}$$

$$\phi_{ab} = \min_{\boldsymbol{\pi}\in\Pi_N}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\arccos\left(\frac{\boldsymbol{s}_i^T\widehat{\boldsymbol{s}}_{\pi_i}}{\|\boldsymbol{s}_i\|\,\|\widehat{\boldsymbol{s}}_{\pi_i}\|}\right)\right]^2} \tag{40}$$

TABLE III
IMPLEMENTATION DETAILS FOR THE ALGORITHMS UNDER TEST

Algorithm	Algorithmic details
NFINDR	Dimension reduction: PCA; Initialization: Randomly selected data points; Abundance estimation: FCLS.
SGA	Dimension reduction: variable PCA; Initialization: Randomly selected target pixel; Abundance estimation: FCLS.
VCA	Dimension reduction: SVD/PCA; Abundance estimation: FCLS.
ICE	Dimension reduction: None; Initialization: VCA and FCLS; Convergence tolerance: 10^{-5}; Maximum iterations: 1000 Regularization parameter: 0.01.
APS	Dimension reduction: None; Initialization: VCA and FCLS; Convergence tolerance: 2×10^{-4}; Maximum iterations: 500; Number of subgradient updates: 20; Regularization parameter: 0.001.
MVC-NMF	Dimension reduction: SVD; Initialization: VCA and FCLS; Convergence tolerance: 10^{-6}; Regularization parameter: 0.001.
MVSA	Dimension reduction: PCA; Initialization: Expanded VCA; Regularization parameter: 10^{-6}.
SISAL	Dimension reduction: PCA; Initialization: Expanded VCA; Regularization parameter: 0.010 to 0.035 (depending on SNR).
MVES	Dimension reduction: Affine set fitting; Initialization: Solving feasibility problem; Convergence tolerance: 10^{-8}.
RMVES	Dimension reduction: Modified affine set fitting; Initialization: Expanded VCA; Convergence tolerance: 10^{-6}; Design parameter η: 0.001.

where $\hat{a}_i$ denotes the estimated endmember signature, $\hat{s}_i = [\hat{s}_i[1], \ldots, \hat{s}_i[L]]^{\mathrm{T}}$ denotes the estimated abundance map, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)^{\mathrm{T}}$, and $\Pi_N = \{\boldsymbol{\pi} \in \mathbb{R}^N | \pi_i \in \{1, 2, \ldots, N\}, \pi_i \neq \pi_j \text{ for } i \neq j\}$ is the set of all the permutations of $\{1, 2, \ldots, N\}$. The average computation time (averaged over all the scenarios and independent runs) of all the algorithms (i.e., the average time required to estimate both the endmember signatures and the abundances), when implemented in Matlab R2008a and running in a desktop computer equipped with Core i7-930 CPU with 2.80-GHz speed and 12-GB memory, is considered as the computational complexity measure.

In the simulations, 8 endmembers, namely, Alunite, Andradite, Buddingtonite, Calcite, Chalcedony, Chlorite, Desert Varnish, and Halloysite, with 224 spectral bands selected from the U.S. Geological Survey library [43], [44] are used to generate the synthetic observations. The abundance vectors were generated by following the Dirichlet distribution [16][26], which can ensure the assumptions on the abundance vectors [i.e., (A1) and (A2)]. The purity parameter ρ_n of the nth abundance vector $\mathbf{s}[n]$ is defined as $1/\sqrt{N} \leq \rho_n = \|\mathbf{s}[n]\| \leq 1$ (where the bounds are due to (A1) and (A2) [26]), which indicates the quantitative dominance of an endmember $\mathbf{a}_i$ in the observed pixel vector $\mathbf{x}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i$. It should be clear that the purity of the observed pixel $\mathbf{x}[n]$ is higher for larger ρ_n. In the simulations, we define a parameter ρ, which means that the ρ_n for all the abundance vectors lies between $1/\sqrt{N}$ and ρ. In each simulation scenario, a pool of $10\,000$ pixels was generated randomly by following the Dirichlet distribution, and then, 1000 vectors ($L = 1000$) with $\rho_n \in [1/\sqrt{N}, \rho]$ were taken from the generated pool, for a given purity parameter ρ. The purity levels considered in our simulations are $\rho = 0.6, 0.8,$ and 1.

As previously inferred for the RMVES algorithm, the apt value of η must be less than 0.5, and from our extensive simulation experience, we found that η should lie in the range of 0.01 to 0.0001. Therefore, in each simulation and real data experiment (see Section VII), the η value for the RMVES algorithm is set to 0.001 (constant) for all SNRs. The reasons behind fixing an η value can be explained by scrutinizing the constraints of problem (35), where Q_i and $\mathbf{Q}$ in (35b) and (35c), respectively, are functions of the noise covariance matrix $\mathbf{D}$. This implies that, for a fixed η value, the noise covariance matrix indirectly determines the lower bound of the abundances, which, in turn, controls the volume of the estimated simplex. Furthermore, since (35) is nonconvex, for each given data set (under a given scenario), the RMVES algorithm is executed ten times with ten different expanded VCA initializations (since each time VCA yields different endmember estimates for noisy observations), and the endmember signature and abundance fractions associated with the largest $|\det(\mathbf{H})|$ are chosen as the optimal endmember and abundance estimates for the data under consideration. Such a technique has been applied before in handling nonconvex problems, e.g., [45].

A. Uniform Gaussian Noise Case

For the uniform Gaussian noise case, the noisy data were obtained by adding independent and identically distributed

TABLE IV
AVERAGE ϕ_{en} AND ϕ_{ab} (IN DEGREES) AND AVERAGE COMPUTATION TIME T (IN SECONDS) OVER THE VARIOUS UNMIXING METHODS FOR DIFFERENT PURITY LEVELS (ρ) AND SNRS—UNIFORM GAUSSIAN NOISE CASE

Methods	ρ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
N-FINDR	0.6	9.35	8.98	8.83	9.21	9.13	8.93	50.79	48.38	47.02	47.40	47.79	47.40	
	0.8	7.22	6.08	5.45	5.00	4.97	4.87	39.41	29.38	21.81	19.64	19.33	18.28	1.49
	1	5.37	2.31	1.48	1.06	0.51	0.32	31.84	13.38	8.18	5.77	2.93	1.78	
SGA	0.6	8.99	8.62	8.76	8.59	8.70	8.58	51.77	48.05	47.48	47.40	47.23	47.28	
	0.8	6.76	4.73	4.23	4.16	3.95	3.93	41.76	28.36	21.98	20.56	19.09	18.83	1.35
	1	5.28	2.13	**1.22**	0.72	0.52	0.40	32.01	13.93	8.27	5.11	3.42	2.44	
VCA	0.6	9.35	8.97	8.33	8.05	8.06	7.82	51.41	47.64	45.98	44.93	45.10	44.69	
	0.8	7.16	5.86	5.38	4.63	4.30	4.30	41.91	30.53	29.25	24.14	20.30	18.90	1.50
	1	5.29	2.39	3.02	1.17	0.57	0.39	34.28	15.28	16.24	6.81	3.49	1.99	
ICE	0.6	8.70	8.19	8.10	8.41	8.14	8.22	**39.98**	33.57	32.12	30.16	30.11	30.03	
	0.8	**5.33**	3.96	3.83	3.74	3.73	3.71	**28.77**	19.40	16.62	14.48	14.50	12.68	420.61
	1	**4.17**	2.53	2.22	2.02	1.96	1.95	**23.65**	14.71	10.96	9.11	8.47	8.28	
APS	0.6	10.67	8.99	9.26	9.77	9.81	10.28	44.45	37.67	34.61	30.71	28.75	26.50	
	0.8	9.00	5.09	7.21	6.46	5.75	5.65	36.33	23.65	24.00	17.52	15.21	14.69	16.76
	1	7.11	**1.96**	5.08	**0.59**	0.70	0.55	32.49	**13.29**	14.43	**4.60**	3.51	2.31	
MVC-NMF	0.6	11.88	10.69	10.91	10.25	9.89	9.72	45.25	36.07	34.96	33.41	32.11	32.04	
	0.8	11.06	6.04	6.69	4.89	3.69	5.35	37.32	22.93	23.70	14.70	10.41	9.93	72.25
	1	10.27	4.50	4.64	1.12	**0.39**	**0.16**	30.47	14.17	12.12	4.82	**2.58**	**1.50**	
MVSA	0.6	14.35	13.18	10.59	8.01	5.40	3.51	45.63	36.31	24.33	15.98	10.17	5.81	
	0.8	14.22	12.72	10.39	7.78	5.02	3.62	40.85	29.42	21.02	13.79	8.56	5.68	3.04
	1	14.47	12.30	10.63	8.06	5.14	3.44	38.49	27.04	19.51	13.08	8.04	4.82	
SISAL	0.6	**8.10**	7.33	6.11	5.25	3.82	2.67	46.93	38.75	26.38	20.21	13.68	8.49	
	0.8	6.48	**3.53**	**1.54**	**0.80**	**0.58**	**0.41**	36.56	20.06	**9.80**	**5.20**	**3.22**	**2.13**	1.91
	1	5.15	2.41	1.77	1.19	0.72	0.42	29.51	13.54	**7.89**	4.76	2.82	1.68	
MVES	0.6	15.01	12.37	9.70	7.44	5.18	3.53	42.33	33.07	21.90	14.76	9.46	5.76	
	0.8	14.20	12.36	10.25	7.68	5.16	3.68	43.28	32.14	22.41	15.00	9.55	5.72	46.44
	1	14.63	12.59	10.30	7.72	5.28	3.65	44.05	32.90	23.19	15.41	10.15	6.46	
RMVES	0.6	9.14	**5.72**	**3.26**	**2.31**	**1.58**	**1.09**	41.49	**26.82**	**13.13**	**7.90**	**4.73**	**2.84**	
	0.8	8.72	5.13	3.42	2.43	1.68	1.11	35.63	**18.48**	10.30	6.49	4.18	2.66	118.97
	1	8.74	5.63	4.41	3.55	2.59	1.77	32.94	16.83	10.85	7.73	6.10	3.61	

zero-mean Gaussian noise vector to the noise-free data for different SNRs, where $SNR = \sum_{n=1}^{L} \|\mathbf{x}[n]\|_2^2/(\sigma^2 ML)$. To maintain nonnegativity of the observed pixels, we artificially set the negative values of the noisy pixels to zero. Table IV shows the rms endmember spectral angle ϕ_{en} (in degrees), rms abundance spectral angle ϕ_{ab} (in degrees), and the overall average computation time (over all the scenarios) T (in seconds) of all the algorithms under test, for the observations corrupted by uniform Gaussian noise, with SNRs between 15 and 40 dB, and $\rho = 0.6$, 0.8, and 1. For each scenario, the best (minimum ϕ_{en} and ϕ_{ab}) out of all the algorithms is highlighted as a bold-faced number. As one can infer from Table IV, the proposed RMVES algorithm achieves the best performance for the highly mixed case ($\rho = 0.6$) when $SNR \geq 20$ dB, and for all the values of ρ, the RMVES algorithm performs better than its predecessor, the MVES algorithm. For the moderately mixed case ($\rho = 0.8$), SISAL shows the best performance for $SNR \geq 20$ dB, followed by RMVES for $SNR \geq 25$ dB. As expected, when pure pixels exist in the data ($\rho = 1$), the performances of the pure-pixel-based algorithms such as N-FINDR, VCA, and SGA are comparable with each other, and in some scenarios, ICE and APS perform well. As far as the average computation time of the algorithms is concerned, Table IV confirms that the values of the average computation time for pure-pixel-based algorithms (N-FINDR, VCA, and SGA) are less than those of other algorithms. Among the other algorithms, SISAL costs the least average computation time, and the average computation time of RMVES algorithm is less than that of ICE algorithm.

As can be seen from Table IV, the computational complexity of RMVES algorithm is the price for the accuracy, particularly when the hyperspectral data are highly mixed.

B. Nonuniform Gaussian Noise Case

For the nonuniform Gaussian noise case, the noise variances σ_i^2 of the M spectral bands, following a Gaussian shape centered at the $(M/2)$th band, as used in [16] and [26], are given by

$$\sigma_i^2 = \sigma^2 \frac{\exp\left(-\frac{(i-M/2)^2}{2\tau^2}\right)}{\sum_{j=1}^{M} \exp\left(-\frac{(j-M/2)^2}{2\tau^2}\right)} \qquad \forall i = 1, \ldots, M \qquad (41)$$

where τ controls the variance of the Gaussian shape among $\sigma_1^2, \ldots, \sigma_M^2$. It corresponds to uniform Gaussian noise for $\tau = \infty$ and one-band noise for $\tau = 0$. Table V shows the average rms spectral angles and the average computation time of all the algorithms under test, for SNR from 15 to 40 dB, $\rho = 0.7$, and $\tau = \infty$, 18, and 9. For the nonuniform Gaussian noise case, in most of the scenarios under consideration, RMVES performs best, followed by SISAL and ICE. It is worth mentioning that, for $\tau = 9$, some performance drop of RMVES may be due to the estimation of the noise covariance matrices in such scenarios. Here again, in all the scenarios, the RMVES algorithm performs better than its predecessor, the MVES algorithm. The

TABLE V
AVERAGE ϕ_{en} AND ϕ_{ab} (IN DEGREES) AND AVERAGE COMPUTATION TIME T (IN SECONDS) OVER THE VARIOUS UNMIXING METHODS FOR DIFFERENT τ VALUES AND SNRs—NONUNIFORM GAUSSIAN NOISE CASE, $\rho = 0.7$

Methods	τ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
N-FINDR	∞	8.14	7.77	7.43	7.23	7.38	7.34	46.89	41.50	37.13	35.12	36.99	34.41	
	18	10.08	8.41	7.19	7.41	7.25	7.28	53.02	47.10	41.89	36.13	33.72	33.62	0.34
	9	11.27	8.96	7.71	7.66	7.60	7.43	56.13	50.63	44.96	40.66	35.64	34.21	
SGA	∞	8.06	6.98	6.50	6.43	6.45	6.45	47.86	40.82	37.96	37.27	37.10	36.73	
	18	10.04	8.02	7.30	6.55	6.47	6.59	53.53	49.58	42.08	37.58	37.26	36.00	1.41
	9	11.37	8.72	7.83	6.85	6.57	6.54	57.33	51.34	46.95	41.95	37.14	35.89	
VCA	∞	8.38	7.68	6.54	6.38	6.62	6.29	47.76	40.46	38.97	36.26	36.61	36.54	
	18	9.89	8.22	6.85	6.49	6.34	6.17	56.48	49.81	41.35	39.59	36.69	35.69	1.43
	9	10.91	**8.68**	7.56	6.78	6.41	6.37	59.34	53.39	46.54	40.46	36.58	36.36	
ICE	∞	**6.02**	4.89	4.93	4.62	4.56	4.76	**31.60**	22.25	19.54	17.26	15.99	17.87	
	18	10.95	**6.65**	4.83	4.71	4.70	4.78	**47.79**	37.31	19.12	17.97	17.71	17.47	470.17
	9	14.05	8.95	**5.22**	4.67	4.92	4.65	**51.20**	**44.73**	26.00	16.75	18.77	17.43	
APS	∞	10.37	8.33	9.08	9.22	9.81	9.57	39.78	32.09	27.98	25.33	25.06	23.47	
	18	17.40	11.71	10.42	9.47	9.15	9.67	50.20	41.77	32.22	26.72	25.01	23.03	24.36
	9	20.64	15.06	12.76	9.82	9.68	9.97	53.21	46.37	37.55	28.35	25.50	24.03	
MVC-NMF	∞	11.94	9.79	9.61	9.31	7.98	7.85	40.63	32.54	32.96	26.58	23.13	21.75	
	18	11.19	11.75	13.05	9.55	9.11	8.20	51.33	40.48	33.32	29.39	25.23	23.91	44.62
	9	13.05	11.44	17.72	10.83	8.58	8.69	54.36	47.72	39.38	29.93	26.40	23.53	
MVSA	∞	14.37	12.87	10.38	7.56	4.90	3.43	44.05	32.82	22.60	14.14	8.72	5.57	
	18	15.54	14.62	13.27	8.71	5.54	3.74	59.47	50.44	33.29	17.00	9.44	5.55	2.96
	9	15.35	14.42	12.93	11.91	6.41	4.20	61.96	56.46	45.67	28.32	11.02	6.38	
SISAL	∞	7.37	5.84	3.87	**2.17**	1.83	1.91	41.23	30.66	17.97	8.84	6.79	5.52	
	18	**9.76**	8.13	5.92	**2.94**	**1.38**	1.59	52.81	43.75	30.08	12.16	6.07	4.49	2.41
	9	**10.68**	8.93	7.28	4.90	**1.50**	**1.39**	58.75	49.39	38.27	23.55	6.27	4.70	
MVES	∞	14.26	12.15	10.17	7.46	5.15	3.42	42.90	32.38	22.26	14.80	9.33	5.66	
	18	14.98	14.31	13.08	8.30	5.30	3.77	52.34	44.75	32.05	16.42	9.43	5.56	38.58
	9	15.92	15.17	12.55	11.17	6.35	4.25	53.70	49.77	40.40	26.81	11.64	6.83	
RMVES	∞	8.52	**4.85**	**3.23**	2.44	**1.48**	**1.17**	37.45	**20.63**	**11.44**	**6.99**	**4.23**	**2.63**	
	18	10.44	7.89	**4.17**	2.97	1.63	**1.08**	52.07	**32.96**	**14.57**	**8.13**	**4.11**	**2.56**	159.44
	9	11.73	9.07	6.57	**3.13**	2.24	1.44	63.06	45.89	**25.71**	**8.97**	**5.33**	**3.23**	

TABLE VI
AVERAGE ϕ_{en} AND ϕ_{ab} (IN DEGREES) AND AVERAGE COMPUTATION TIME T (IN SECONDS) WHILE DIRECTLY APPLYING SQP TO (31), FOR DIFFERENT ρ VALUES AND SNRs—UNIFORM GAUSSIAN NOISE CASE

Method	ρ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
Applying SQP to (31)	0.6	21.02	20.34	20.86	20.48	20.75	20.14	48.69	46.00	45.95	43.48	43.34	41.90	
	0.8	19.41	20.10	20.47	17.89	16.98	17.11	52.88	49.61	48.48	44.49	40.90	41.10	5.64
	1	18.86	18.91	16.56	15.49	15.39	15.44	55.46	52.25	46.09	44.16	44.18	44.23	

observations on the average computation time of the algorithms under test for the uniform Gaussian noise case in Section VI-A also apply to the nonuniform Gaussian noise case here.

C. Local Optimality Issues

In this section, let us investigate the local optimality issues while attempting to solve the nonconvex problems (31) and (35) (via RMVES algorithm). Table VI shows the average ϕ_{en} and ϕ_{ab} for the uniform Gaussian noise case and the average computation time T (in seconds), when (31) is directly handled using SQP. Here again, the η value used is 0.001, and (31) is executed ten times with ten different VCA initializations (as in Table I). The estimates associated with the maximum value of $|\det(\mathbf{H})|$ are chosen as the solution for (31). As one can infer by comparing Tables IV and VI, attempting to solve (35) in a cyclic optimization fashion (RMVES algorithm) significantly improves the performance of the algorithm. The numerical results from these tables indicate that cyclic optimization is not

as susceptible to local optimality issues as attempting to solve (31) directly by SQP.

Another interesting observation is the variation in the average endmember spectral angle when handling (35) (via RMVES algorithm) with one VCA initialization (as in Table I) and with ten different VCA initializations. In Fig. 5, the asterisks correspond to the average endmember spectral angle (averaged over 50 independent runs) obtained by executing the RMVES algorithm with a single VCA initialization in each of the 50 independent runs, for $\rho = 0.6$ and SNR (uniform Gaussian noise) varying from 15 to 40 dB with a step size of 5 dB. The triangles in the figure represent the average endmember spectral angles associated with the RMVES algorithm for the same scenarios, but with ten VCA initializations in each of the 50 independent runs, and choosing the spectral angle associated with the maximum $|\det(\mathbf{H})|$ (i.e., those ϕ_{en}'s shown in Table IV). As one can infer, the differences between the average endmember spectral angles obtained by RMVES with one VCA initialization and by RMVES with ten VCA initializations for
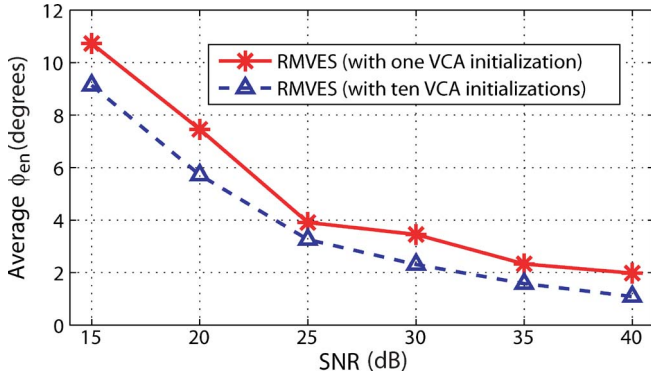
Fig. 5. Average endmember spectral angles obtained by the RMVES algorithm for $\rho = 0.6$ with one VCA initialization and ten VCA initializations and choosing the spectral angle associated with the maximum $|\det(\mathbf{H})|$.

TABLE VII
var$\{\phi_{en}\}$ AND mean$\{\phi_{en}\}$ OF VARIOUS HU ALGORITHMS,
WHEN $\rho = 0.6$ AND $SNR = 25$ dB, FOR 50 INDEPENDENT DATA SETS.
FOR EACH DATA SET, THE ALGORITHM IS APPLIED TEN TIMES, EACH
TIME WITH ITS INHERENT INITIALIZATION

Algorithm	var$\{\phi_{en}\}$ (in degrees^2)	mean$\{\phi_{en}\}$ (in degrees)
NFINDR	1.16	9.06
SGA	0	8.52
VCA	1.23	8.17
ICE	1.64	8.38
APS	0	9.26
MVC-NMF	5.04	10.69
MVSA	3.02	10.81
SISAL	2.38	6.74
MVES	0	9.71
RMVES	3.77	4.29

all the scenarios are within two degrees. This indicates that the local optimality issue associated with (35) is not severe, but still exists (even after linearizing the objective function of (31) via cofactor expansion), since the constraint sets are nonconvex.

Analyzing the local optimality issue of the existing HU algorithms in a fair fashion is nontrivial due to their distinct characteristics and can be a separate research topic. However, to illustrate the existence of local optimality in other HU methods, we proceed with the following simulations. A total of 50 data sets have been independently generated with $\rho = 0.6$, under uniform Gaussian noise case with an SNR of 25 dB, and the HU algorithms were applied ten times (*each time with a different initialization*) to each of the data set. Then, for the ith data set, the variance and mean of ϕ_{en} (over the ten executions of each algorithm on the ith data set), denoted by var$\{\phi_{en}\}_i$ and mean$\{\phi_{en}\}_i$, respectively, are noted and the average variance and mean over 50 independent runs are computed, respectively, as

$$\text{var}\{\phi_{en}\} = \frac{1}{50} \sum_{i=1}^{50} \text{var}\{\phi_{en}\}_i$$

$$\text{mean}\{\phi_{en}\} = \frac{1}{50} \sum_{i=1}^{50} \text{mean}\{\phi_{en}\}_i.$$

Table VII shows the var$\{\phi_{en}\}$ and mean$\{\phi_{en}\}$ of the various algorithms under test. The mean$\{\phi_{en}\}$ values in Table VII are comparable with the corresponding values shown in Table IV,

where the average values of ϕ_{en} are over 50 independent runs without using 10 different initializations for each independent run. As can be seen from Table VII, in addition to SGA [14] and APS [22], the var$\{\phi_{en}\}$ for MVES algorithm is also zero because the MVES algorithm is initialized by solving the feasibility problem [26] (that is inherent in the MVES algorithm), which, being convex, always yields the same initial estimates for each data set (i.e., the same initialization is yielded for the same data set).

VII. REAL DATA EXPERIMENTS

In this section, we demonstrate the performance of the proposed RMVES algorithm using the AVIRIS data taken over the Cuprite Nevada site [46], as a good library of endmember signature is available [43], [44] and thus can be used for performance evaluation. The AVIRIS data are challenging because of two reasons. One reason is that the spatial resolution is not so high (about 17 m per pixel), and the other reason is that the true total number of endmembers and the associated minerals are yet to be accurately identified. While, initially, it was concluded that there are about 13 minerals (endmembers) in the site, the number kept increasing, and later, it was reported that there are around 70 mineral compounds (endmembers) [47] present in the site and the number is still expected to rise. In our experiment, we considered a 200×200 subimage of the hyperspectral data as our region of interest (ROI), with 224 spectral bands. The bands 1–2, 104–113, 148–167, and 221–224 are less significant (due to strong noise or dense water-vapor content) and were removed. A total of 188 bands is therefore considered. Estimation of the number of endmembers present in a given scene of interest is an important issue, because both the dimension reduction algorithm and the HU algorithm are in need of this number. To begin with, the eigenvalue distribution (signal energy distribution) of the ROI's data covariance matrix shows that the number of sources according to the principal eigenvalues is approximately 4 ($N = 4$). Obviously, it is an underestimate of N, as the ground truth [43] reports more than four endmembers in the ROI. Applying HySiMe [11] to the ROI data yields $N = 18$. On the other hand, for a very similar ROI, it is reported in [16] and [26] that VD [8] estimates the number of endmembers as 14. Also, Miao and Qi reported in their work [23] that, for a similar ROI, $N = 9$, and they further claimed that the overestimation of N by other methods can be attributed to the nonlinear mixing of the observations in a real scenario. However, for the sake of consistency with our previous work [26], here we consider $N = 14$. Two existing algorithms, namely, MVES and VCA, were also tested with the same real data and compared with the proposed RMVES algorithm. It should be noted that the nonpure-pixel-based algorithms are sensitive to initialization, and hence, their performances vary with the initialization. In order to have a fair comparison with our RMVES algorithm, the MVES algorithm was applied to the hyperspectral data with ten different expanded VCA initializations (which are also feasible initializations for MVES), and the estimates associated with the least simplex volume is chosen as the results of MVES algorithm.

In order to remove some pixels corresponding to inactive constraints and thereby speed up the algorithms, data subsampling is needed. Since the pixels at the periphery (the
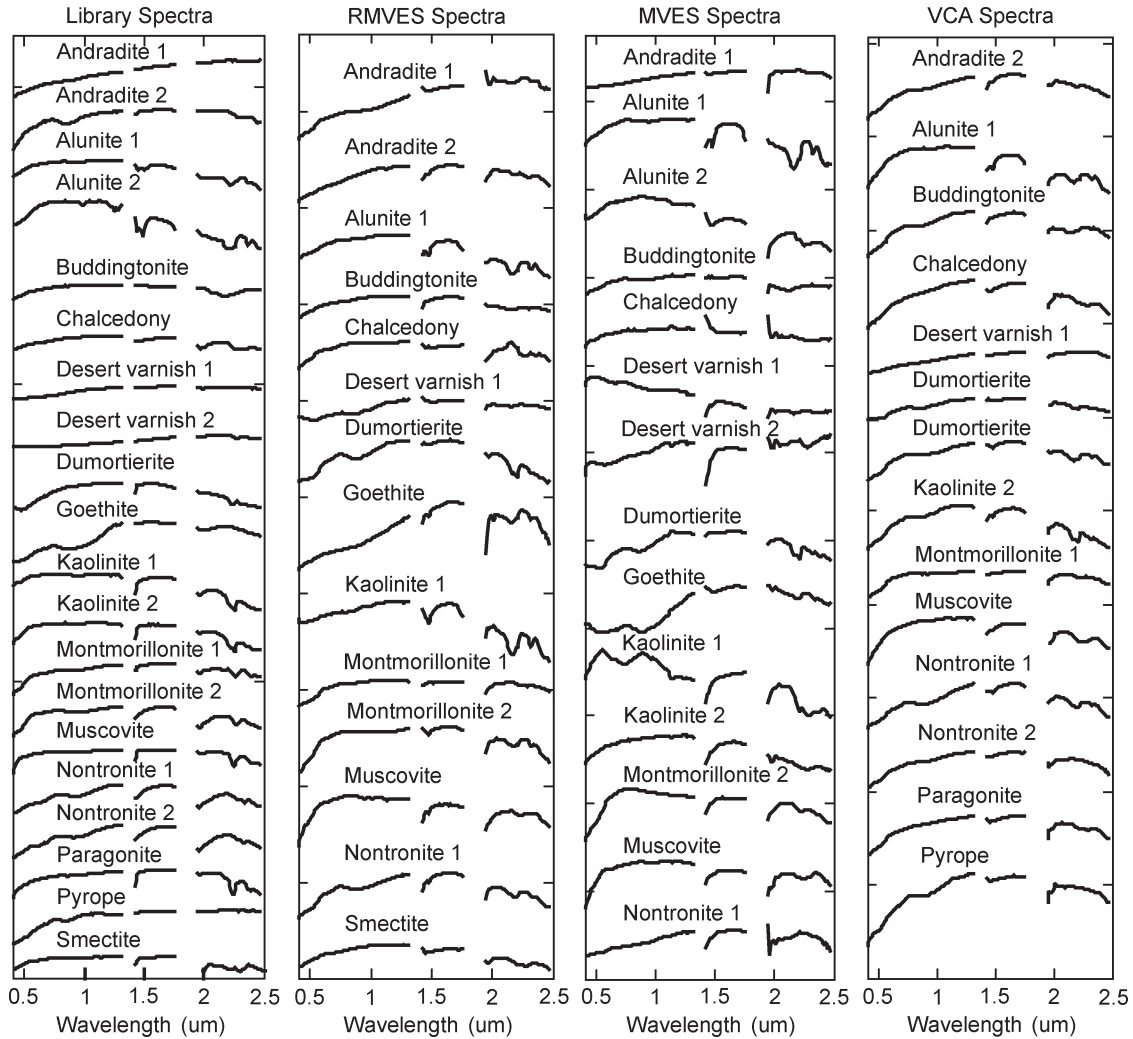
Fig. 6.   Endmember signatures taken from the library and the ones estimated by RMVES, MVES, and VCA.

pixels associated with active constraints) of the dimension-reduced data cloud determine the simplex volume, we first perform some subsampling of the dimension-reduced hyperspectral data $\tilde{\mathbf{y}}[i]$ by using the following convex projection procedure: We consider solving the following optimization problem for all $i = 1, \ldots, L$:

$$\min_{\boldsymbol{\theta} \in \mathbb{R}^{L-1}} \quad \left\| \tilde{\mathbf{y}}[i] - \tilde{\mathbf{Y}}_i \boldsymbol{\theta} \right\|^2$$
$$\text{s.t.} \qquad \mathbf{1}_{L-1}^{\mathrm{T}} \boldsymbol{\theta} = 1, \boldsymbol{\theta} \succeq \mathbf{0}. \qquad (42)$$

where $\tilde{\mathbf{Y}}_i = [\tilde{\mathbf{y}}[1], \tilde{\mathbf{y}}[2], \ldots, \tilde{\mathbf{y}}[i-1], \tilde{\mathbf{y}}[i+1], \ldots, \tilde{\mathbf{y}}[L]]$. Problem (42) is a convex problem and can be solved using available convex optimization solvers, such as SeDuMi [36] and CVX [37]. The idea is to check each and every pixel in the dimension-reduced data cloud for the case whether it belongs to the convex hull of all the other remaining pixels (i.e., if the optimal value of (42) is zero) or not. If yes, the pixel is discarded, and if not, it is retained (as they correspond to a peripheral pixel). In spite of its initial cost, this procedure of data subsampling significantly aids in speeding up the algorithms under test. By doing so, we were able to identify 17 965 peripheral pixels out of the 40 000 pixels in the ROI,

and we have used them in our experiments for RMVES and MVES algorithms. Finally, it should be mentioned that, in all the algorithms under test, the complete hyperspectral data set (without subsampling) is used for the estimation of the abundance maps. When applied to real hyperspectral data, the computational times for RMVES and MVES (both with subsampled data) and VCA (with full data) are around 145 min (which is around 8 hours on average for the complete data), 119 min, and 4 min, respectively. Since we need to perform ten executions with different expanded VCA initializations and choose the solution corresponding to the maximum $|\det(\mathbf{H})|$ as the endmember estimates of the RMVES algorithm, the subsampled data are used.

The endmember signatures obtained via RMVES, MVES, and VCA, along with the corresponding library signatures, are shown in Fig. 6. The minerals were identified by the visual comparison of the obtained abundance maps with the ones available in [23], [26], [43], [48], and [49]. The materials identified are arranged in alphabetical order for ease of visual comparison. Owing to the space limitation, only the abundance maps obtained by the RMVES algorithm are shown in Fig. 7. It should be mentioned that the difference in the materials identified by the three algorithms could possibly be due to the working nature of the respective algorithms and their sensitivity
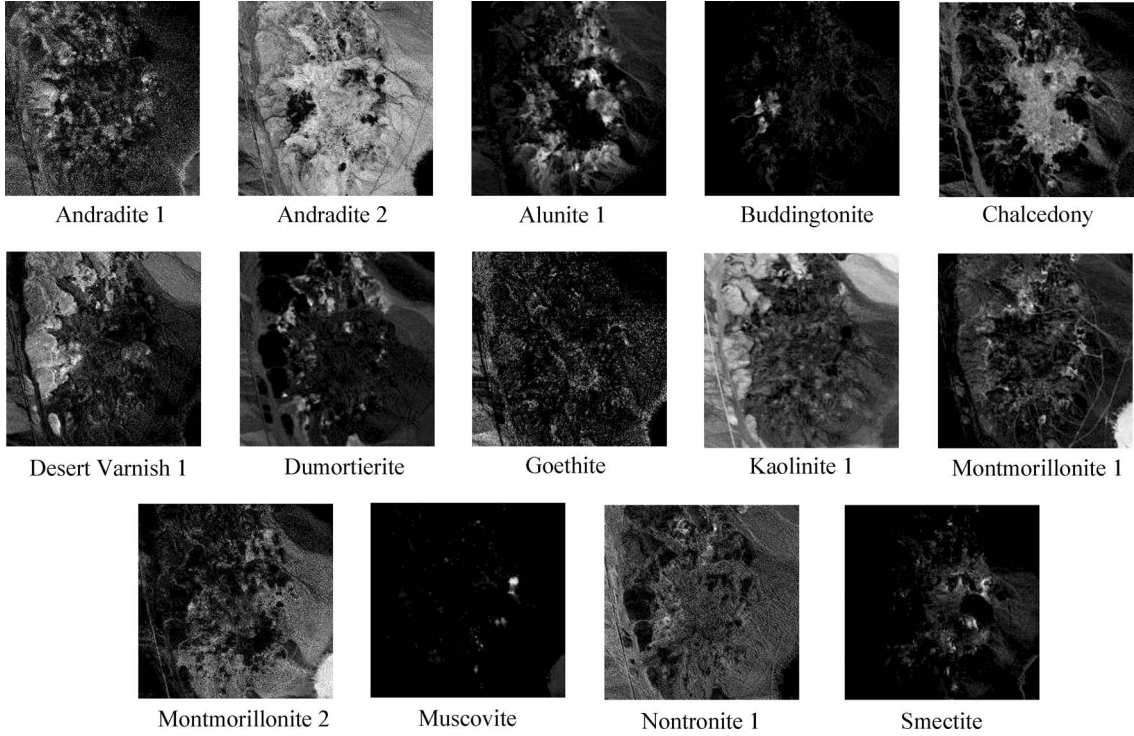
Fig. 7. Abundance maps obtained by the RMVES algorithm.

to initializations. The mean-removed spectral angle between the estimated signature $\mathbf{a}_{\text{est}}$ and the corresponding library signature $\mathbf{a}_{\text{lib}}$[26], [48] defined as

$$\phi = \arccos\left(\frac{(\mathbf{a}_{\text{est}} - \mathbf{m}(\mathbf{a}_{\text{est}}))^{\text{T}}(\mathbf{a}_{\text{lib}} - \mathbf{m}(\mathbf{a}_{\text{lib}}))}{\|\mathbf{a}_{\text{est}} - \mathbf{m}(\mathbf{a}_{\text{est}})\| \cdot \|\mathbf{a}_{\text{lib}} - \mathbf{m}(\mathbf{a}_{\text{lib}})\|}\right) \quad (43)$$

was used as the performance measure, where $\mathbf{m}(\mathbf{a}) = (\mathbf{1}_M^{\text{T}} \mathbf{a})\mathbf{1}_M(1/M)$ for any vector $\mathbf{a} \in \mathbb{R}^M$. The values of ϕ for the various minerals identified by the algorithms under test are given in Table VIII. Again, the least ϕ value for an endmember is highlighted as a bold-faced number, and the number in parenthesis is the ϕ value for the repeatedly identified endmember. It can be seen from Table VIII that, for the materials identified, the proposed RMVES algorithm performs best (with the minimum average ϕ), and furthermore, it mostly yields better endmember estimates (minimum ϕ) than its predecessor, i.e., the MVES algorithm, and the pure-pixel-based VCA algorithm. Note that few of the mean-removed spectral angles for endmembers (e.g., Buddingtonite and Kaolinite#1) identified by RMVES algorithm are marginally higher than that of the ones obtained by MVES. This could be attributed to some orientation of the simplex obtained by RMVES, with respect to the simplex of the true endmembers. In our simulations and real data experiments, HySiMe [11] was used to estimate the noise covariance matrix. Better estimation of the noise covariance matrix should further enhance the performance of the RMVES algorithm.

VIII. CONCLUSION

In this paper, we have presented a robust HU algorithm, namely, RMVES (as shown in Table II), for effective unmixing of mixed hyperspectral data corrupted by uniform or nonuniform Gaussian noise. The dimension reduction via affine set

TABLE VIII
MEAN-REMOVED SPECTRAL ANGLES ϕ (IN DEGREES) BETWEEN LIBRARY SPECTRA AND ENDMEMBERS ESTIMATED BY RMVES, MVES, AND VCA

	RMVES	MVES	VCA
Andradite#1	**9.36**	25.61	-
Andradite#2	24.52	-	**18.49**
Alunite#1	**15.92**	21.85	17.74
Alunite#2	-	**17.72**	-
Buddingtonite	23.54	**22.98**	27.25
Chalcedony	**27.74**	38.25	31.9
Desert Varnish#1	20.99	18.64	**12.12**
Desert Varnish#2	-	**43.04**	-
Dumortierite	**20.77**	29.32	31.95 (32.01)
Goethite	**17.71**	19.05	-
Kaolinite#1	27.25	**26.50**	-
Kaolinite#2	-	**21.09**	32.49
Montmorillonite#1	22.99	-	**18.06**
Montmorillonite#2	**24.34**	26.00	-
Muscovite	39.63	44.19	**32.7**
Nontronite#1	**22.95**	28.83	24.66
Nontronite#2	-	-	**21.51**
Paragonite	-	-	**35.91**
Pyrope	-	-	**25.59**
Smectite	**22.53**	-	-
Average ϕ	**22.87**	27.11	25.73

fitting procedure has been suitably modified for noisy hyperspectral observations. The randomness caused by noise has been dealt with by incorporating chance constraints in the unmixing problem formulation with a design parameter η. A detailed analysis on the role of η has been presented, and it was concluded that η must be less than 0.5, which, along with the objective function, results in a nonconvex optimization problem. In an effort to minimize the effect of nonconvexity of the objective function, an alternating optimization concept has been utilized. The partial maximization problems involved are handled by using available SQP solvers. Finally, Monte Carlo

simulations and real data experiments presented in Sections VI and VII, respectively, demonstrate that RMVES algorithm provides superior performance over several existing benchmark methods, including its predecessor, the MVES algorithm. Some future research directions are as follows: 1) From our simulation results and real data experiments (where we have set $\eta = 0.001$), the RMVES algorithm performs well; however, other more appropriate choices for η that can still enhance the performance of the RMVES algorithm for different scenarios are worth for investigations; 2) an algorithm that can account for the orientation of the estimated endmember simplex in addition to the volume shrinkage of the estimated simplex should further improve the performance of the RMVES algorithm; and 3) an algorithm that is not only robust against noise but also robust against outliers in the hyperspectral data will be an interesting, but challenging, future direction.

REFERENCES

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

[2] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Appl.*, vol. 416, no. 1, pp. 29–47, Jul. 2006.

[3] M. B. Lopes, J. C. Wolff, J. B. Dias, and M. Figueiredo, "NIR hyperspectral unmixing based on a minimum volume criterion for fast and accurate chemical characterisation of counterfeit tablets," *Anal. Chem.*, vol. 82, no. 4, pp. 1462–1469, Jan. 2010.

[4] W. R. Johnson, D. W. Wilson, and W. Fink, "Snapshot hyperspectral imaging in ophthalmology," *J. Biomed. Opt.*, vol. 12, no. 1, p. 014 036, Jan./Feb. 2007.

[5] M. O. Smith, P. E. Johnson, and J. B. Adams, "Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis," *J. Geophys. Res.*, vol. 90, no. 2, pp. C797–C804, Feb. 1985.

[6] A. A. Green, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 1, pp. 65–74, May 1988.

[7] I. K. Fodor, "A survey of dimension reduction techniques," Lawrence Livermore Nat. Lab., Livermore, CA, May 2002, Tech. Rep.

[8] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.

[9] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.

[10] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.

[11] J. M. B. Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.

[12] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. Summ. JPL Airborne Earth Sci. Workshop*, Pasadena, CA, Dec. 9–14, 1995, pp. 23–26.

[13] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE—Imaging Spectrometry V*, Pasadena, CA, Oct. 1999, pp. 266–275.

[14] C. I. Chang, C. C. Wu, W. M. Liu, and Y. C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, Oct. 2006.

[15] C. I. Chang, C. C. Wu, C. S. Lo, and M. L. Chang, "Real-time simplex growing algorithms for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1834–1850, Apr. 2010.

[16] 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.

[17] T.-H. Chan, C.-Y. Chi, W.-K. Ma, and A. Ambikapathi, "Hyperspectral unmixing from a convex analysis and optimization perspective," in *Proc. 1st IEEE WHISPERS*, Grenoble, France, Aug. 26–28, 2009, pp. 1–4.

[18] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.

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

[20] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.

[21] A. Zare and P. Gader, "PCE: Piecewise convex endmember detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2620–2632, Jun. 2010.

[22] A. Zymnis, S.-J. Kim, J. Skaf, M. Parente, and S. Boyd, "Hyperspectral image unmixing via alternating projected subgradients," in *Proc. 41st Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 4–7, 2007, pp. 1164–1168.

[23] 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.

[24] A. Huck, M. Guillaume, and J. B. Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2590–2602, Jun. 2010.

[25] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Boston, MA, Aug. 8–12, 2008, vol. 4, pp. 2369–2371.

[26] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.

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

[28] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.

[29] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," in *Proc. 2nd IEEE WHISPERS*, Reykjavik, Iceland, Jun. 14–16, 2010, pp. 1–4.

[30] 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.

[31] J. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. 1st IEEE WHISPERS*, Grenoble, France, Aug. 26–28, 2009, pp. 1–4.

[32] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[33] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "A robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," in *Proc. IEEE ICASSP*, Dallas, TX, Mar. 14–19, 2010, pp. 1202–1205.

[34] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "A robust alternating volume maximization algorithm for endmember extraction in hyperspectral images," in *Proc. IEEE WHISPERS*, Reykjavik, Iceland, Jun. 14–16, 2010, pp. 1–4.

[35] G. Strang, *Linear Algebra and Its Applications*, 4th ed. Belmont, CA: Thomson, 2006.

[36] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11-12, pp. 625–653, 1999.

[37] M. Grant and S. Boyd, CVX: Matlab Software for disciplined convex programming, Oct. 2010. [Online]. Available: http://cvxr.com/cvx

[38] T.-H. Chan, W.-K. Ma, C.-Y. Chi, and Y. Wang, "A convex analysis framework for blind separation of non-negative sources," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 5120–5134, Oct. 2008.

[39] W.-K. Ma, T.-H. Chan, C.-Y. Chi, and Y. Wang, "Convex analysis for nonnegative blind source separation with application in imaging," in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2010, ch. 7.

[40] F.-Y. Wang, C.-Y. Chi, T.-H. Chan, and Y. Wang, "Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 875–888, May 2010.

[41] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta Numerica*, vol. 4, pp. 1–51, 1996.

[42] MATLAB Optimization Toolbox—Version 7.6 (R2008a), MathWorks, Inc., Natick, MA.

[43] Tech. Rep. [Online]. Available: http://speclab.cr.usgs.gov/cuprite.html

[44] R. N. Clark, G. A. Swayze, A. Gallagher, T. V. King, and W. M. Calvin, "The U.S. Geological Survey digital spectral library: Version 1: 0.2 to 3.0," U. S. Geol. Surv., Denver, CO, Open File Rep. 93-592, 1993.

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

[46] AVIRIS Data Products. [Online]. Available: http://aviris.jpl.nasa. gov/html/aviris.freedata.html

[47] R. N. Clarke and G. A. Swayze, "Evolution in imaging spectroscopy analysis and sensor signal-to-noise: An examination of how far we have come," in *Proc. 6th Annu. JPL Airborne Earth Sci. Workshop*, Mar. 4–8, 1996, pp. 49–53.

[48] T. H. Chan, "Convex analysis based non-negative blind source separation for biomedical and hyperspectral image analysis," Ph.D. dissertation, Nat. Tsing Hua Univ., Hsinchu, Taiwan, 2009.

[49] J. M. P. Nascimento, "Unsupervised hyperspectral unmixing," Ph.D. dissertation, Tech. Univ. Lisbon, Lisbon, Portugal, 2006.

ArulMurugan Ambikapathi (S'02) received the B.E. degree from Bharathidasan University, Tiruchirappalli, India, in 2003 and the M.E degree from Anna University, Chennai, India, in 2005. He secured University ranks in both these programs. He is currently working toward the Ph.D. degree in the Institute of Communications Engineering, National Tsing Hua University (NTHU), Hsinchu, Taiwan.

His research interests are in signal processing, convex optimization, and hyperspectral unmixing.

Mr. Ambikapathi was the recipient of the NTHU Outstanding Student Scholarship award for two consecutive years (in 2009 and 2010).
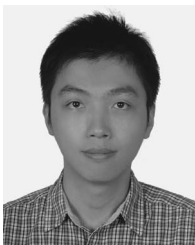
Tsung-Han Chan (S'08–M'09) received the B.S. degree from the Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan, in 2004 and the Ph.D. degree from the Institute of Communications Engineering, National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2009.

In 2008, he was a Visiting Doctoral Graduate Research Assistant at Virginia Polytechnic Institute and State University, Arlington. He is currently a Postdoctoral Research Fellow with the Institute of Communications Engineering, NTHU. His research interests are in signal processing, convex optimization, and pattern analysis, with a recent emphasis on dynamic medical imaging and hyperspectral remote sensing applications.

Wing-Kin Ma (S'96–M'01) received the B.Eng. (with first-class honors) degree in electrical and electronic engineering from the University of Portsmouth, Portsmouth, U.K., in 1995 and the M.Phil. and Ph.D. degrees in electronic engineering from the Chinese University of Hong Kong (CUHK), Hong Kong, in 1997 and 2001, respectively. His Ph.D. dissertation was commended to be "of very high quality and well deserved honorary mentioning" by the Faculty of Engineering, CUHK, in 2001.

He is currently an Assistant Professor with the Department of Electronic Engineering, CUHK. From 2005 to 2007, he was also an Assistant Professor with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan, where he is still holding an adjunct position. Prior to becoming a faculty member, he held various research positions with McMaster University, Hamilton, ON, Canada; CUHK; and the University of Melbourne, Melbourne, Australia. His research interests are in signal processing and communications, with a recent emphasis on MIMO, blind signal processing, and convex optimization techniques.

Dr. Ma currently serves as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS. He has also served as a Guest Editor of IEEE SIGNAL PROCESSING MAGAZINE on the Special Issue titled Convex Optimization for Signal Processing in May 2010.

Chong-Yung Chi (S'83–M'83–SM'89) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1983.

From 1983 to 1988, he was with the Jet Propulsion Laboratory, Pasadena, CA. He has been a Professor with the Department of Electrical Engineering since 1989 and the Institute of Communications Engineering (ICE) since 1999 (also the Chairman of ICE during 2002–2005), National Tsing Hua University, Hsinchu, Taiwan. He was a Member of the Editorial Board of *EURASIP Signal Processing Journal* (in June 2005–May 2008) and an Editor (in July 2003–December 2005) as well as a Guest Editor (in 2006) of *EURASIP Journal on Applied Signal Processing*. He has published more than 170 technical papers, including more than 60 journal papers (mostly in IEEE TRANSACTIONS ON SIGNAL PROCESSING), 2 book chapters, and more than 100 peer-reviewed conference papers, as well as a graduate-level textbook, "Blind Equalization and System Identification" (Springer-Verlag, 2006). His current research interests include signal processing for wireless communications, convex analysis and optimization for blind source separation, and biomedical and hyperspectral image analysis.

Dr. Chi has been a Technical Program Committee member for many IEEE sponsored and cosponsored workshops, symposiums, and conferences on signal processing and wireless communications, including Co-organizer and General Cochairman of the 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications, Cochair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008, and Lead Cochair of SPC Symposium, ChinaCOM 2009. He was an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING (in May 2001–April 2006), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (in January 2006–December 2007), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (in January 2008–December 2009), and IEEE SIGNAL PROCESSING LETTERS (in June 2006–May 2010). He was a member of the IEEE Signal Processing Committee on Signal Processing Theory and Methods (in 2005–2010). Currently, he is a member of the IEEE Signal Processing Committee on SPC and Networking.