

# Hyperspectral Data Geometry-Based Estimation of Number of Endmembers Using $p$ -Norm-Based Pure Pixel Identification Algorithm

ArulMurugan Ambikapathi, *Member, IEEE*, Tsung-Han Chan, *Member, IEEE*,  
Chong-Yung Chi, *Senior Member, IEEE*, and Kannan Keizer

**Abstract**—Hyperspectral endmember extraction is a process to estimate endmember signatures from the hyperspectral observations, in an attempt to study the underlying mineral composition of a landscape. However, estimating the number of endmembers, which is usually assumed to be known *a priori* in most endmember estimation algorithms (EEAs), still remains a challenging task. In this paper, assuming hyperspectral linear mixing model, we propose a hyperspectral data geometry-based approach for estimating the number of endmembers by utilizing successive endmember estimation strategy of an EEA. The approach is fulfilled by two novel algorithms, namely geometry-based estimation of number of endmembers—convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm. The GENE-CH and GENE-AH algorithms are based on the fact that all the observed pixel vectors lie in the convex hull and affine hull of the endmember signatures, respectively. The proposed GENE algorithms estimate the number of endmembers by using the Neyman–Pearson hypothesis testing over the endmember estimates provided by a successive EEA until the estimate of the number of endmembers is obtained. Since the estimation accuracies of the proposed GENE algorithms depend on the performance of the EEA used, a reliable, reproducible, and successive EEA, called  $p$ -norm-based pure pixel identification (TRI-P) algorithm is then proposed. The performance of the proposed TRI-P algorithm, and the estimation accuracies of the GENE algorithms are demonstrated through Monte Carlo simulations. Finally, the proposed GENE and TRI-P algorithms are applied to real AVIRIS hyperspectral data obtained over the Cuprite mining site, Nevada, and some conclusions and future directions are provided.

**Index Terms**—Endmember identifiability, estimation of number of endmembers, hyperspectral imaging, hyperspectral unmixing (HU), pure pixel, reproducibility, successive endmember extraction.

Manuscript received July 20, 2011; revised March 19, 2012; accepted July 10, 2012. Date of publication September 24, 2012; date of current version April 18, 2013. This work was presented in part at the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Prague, Czech Republic May 22–27, 2011. This work was supported by the National Science Council (R.O.C.) under Grants NSC 99-2221-E-007-003-MY3 and NSC 96-2628-E-007-003-MY3.

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

C.-Y. Chi is with 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>.

Digital Object Identifier 10.1109/TGRS.2012.2213261

## I. INTRODUCTION

**H**YPERSPECTRAL remote sensing is a powerful technique to study and analyze an object of interest, through a set of images that are obtained over hundreds of narrowly spaced spectral bands. The object of interest could be geographical areas of the Earth [1], extra terrestrial objects [2], counterfeited tablets [3], and many others. In this paper, we consider the hyperspectral images that are taken with a purpose of material identification and quantification in a given geographical location. Herein, the electromagnetic scattering patterns are recorded over hundreds of continuous spectral bands from a range of visible to near-infrared wavelength, thereby yielding a data cube containing spatial and spectral information of the substances within the area of interest. In analyzing the obtained data cube, hyperspectral unmixing (HU) [1] that decomposes the data cube into endmember signatures (which correspond to the reflection coefficients of the materials present in the scene) and their corresponding abundance maps (which correspond to the proportional distribution of those endmembers present in the scene), plays a prominent role, and many new HU algorithms [4], [5] are being introduced for this purpose. HU primarily involves three processing procedures, namely dimension reduction, endmember extraction, and abundance estimation [1]. The HU results can only be completely interpretable, when the number of substances (or endmembers) present in that area is given *a priori*, which however is unknown in practice.

Although a number of early efforts have been made in developing algorithms to estimate the number of endmembers (also known as rank estimation [6], [7] or model order selection [8], [9]), the problem of estimating the number of endmembers remains one of the greatest challenges. The vast majority of the existing methods for estimating the number of endmembers can be classified into two categories: information theoretic criteria-based methods and eigenvalue thresholding methods. Methods falling into the group of information theoretic criteria includes Akaike's information criterion (AIC) [10], minimum description length (MDL) [11], and Bayesian information criterion (BIC) [12], [13], to name a few. These criteria generally consist of two additive terms: a negative data log-likelihood term and a penalty term. As the number of endmembers increases, the value of the negative data log-likelihood term decreases, whereas the value of the penalty term increases. The best estimate of the number of endmembers is the one that yields

the minimum value of the criteria. In AIC, MDL and BIC, the data log-likelihoods in the criteria are identical, but how each method penalizes the overestimation of the number of endmembers makes the difference. Since the criteria require the prior knowledge of the mixture model or likelihood function, the estimation results may suffer from model mismatch errors resulting from incorrect prior information. It has been shown in [14] that the results of AIC and MDL when applied to hyperspectral data are seriously overestimated due to the invalid Gaussian distribution assumption made on the abundances [15].

Methods belonging to the group of eigenvalue thresholding (either explicitly or implicitly) include principal component analysis (PCA)-based approaches [16], Neyman–Pearson detection theory-based method [17] (also referred to as virtual dimensionality (VD) in [14]), and hyperspectral signal subspace identification by minimum error (Hysime) [18], to name a few. PCA-based approaches aim to determine the cutoff threshold between the eigenvalues caused by signals and noise, but the variation between the two adjacent eigenvalues may not be significant in a practical scenario (more discussions about the threshold selection for PCA can be found in [19]), thereby easily leading to estimation errors in number of endmembers estimation, as demonstrated in [18]. The Neyman–Pearson detection theory-based method was first proposed by Harsanyi, Farrand, and Chang (HFC) in 1993, and is termed as HFC method [17]. The HFC method uses Neyman–Pearson detector for a binary hypothesis testing problem, built on the differences in eigenvalues of the sample correlation and sample covariance matrices. The HFC method was later revisited by incorporating the concepts of VD and noise-prewhitening step [14]. An alternative to VD, namely second moment linear dimensionality has been recently reported in [20]. Hysime [18] utilizes a minimum mean square error criterion to estimate the signal subspace in hyperspectral images. It starts with estimating the signal and the noise correlation matrices, and then selects the subset of eigenvectors that best represent the signal subspace in the least squared error sense. In addition to the two aforementioned categories, two other methods are worth mentioning. A Markov chain Monte Carlo (MCMC)-based approach was proposed in [21] to estimate the number of endmembers. While the applicability of the algorithm is restricted to the data with a smaller number of endmembers/pixels due to high computational demand, it provides Bayesian estimation of the number of endmembers, with theoretical basis. An extension of the unmixing algorithm namely, iterated constrained endmember (ICE) [22] has been reported in [23], where iterative endmember pruning through sparsity-promoting priors has been used to simultaneously estimate the number of endmembers and to unmix the hyperspectral data.

In this paper, we consider the linear mixing model (which is the most prominent model in HU [1]) and propose two hyperspectral data geometry-based algorithms for estimating the number of endmembers, namely geometry-based estimation of number of endmembers—convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm. The proposed GENE algorithms (GENE-CH and GENE-AH) exploit successive estimation property of a pure-pixel-based endmember extraction algorithm (EEA), and aim to decide when the EEA should stop

estimating the next endmember signature. The GENE-CH and GENE-AH algorithms are devised based on the data geometry fact that all the observed pixel vectors should lie in the convex hull (CH) and affine hull (AH) of the endmember signatures, respectively. Since the EEAs identify endmember estimates from the set of observed pixel vectors, the fact pertaining to the data geometry also implies that the current endmember estimate should lie in the CH/AH of the previously found endmembers when the current endmember estimate is obtained for an overly estimated number of endmembers. In the noisy scenario, the decision of whether the current endmember estimate is in the CH/AH of the previously found endmembers can be formulated as a binary hypothesis testing problem, which we solve by Neyman–Pearson detection theory. The performances of the proposed GENE algorithms depend on the accuracy of pure pixel indices identification of an EEA. Existing EEAs include pure pixel search algorithms [24], [25], algorithms based on Winter’s belief [26]–[31], and spatial-information-based algorithms [32], [33], to name a few. However, EEAs that can be used in conjunction with the GENE algorithms are preferred to have the following properties for better estimation accuracy and efficiency.

- Reliability—The EEA can reliably find a set of true endmembers provided that pure pixels exist in the hyperspectral data. Specifically, for the noiseless case, its endmember identifiability can be guaranteed.
- Reproducibility—The EEA provides reproducible endmember estimates for a given hyperspectral data set without need of any random initialization.
- Successive estimation—The EEA estimates the endmembers successively.
- Computational efficiency—The EEA has low computational complexity (as the overall complexity of the GENE algorithms also depends on the complexity of the EEA employed).

Therefore, we herein propose a reliable, reproducible, and computationally efficient, successive EEA, called  $p$ -norm-based pure pixel identification (TRI-P, abbreviated for Triple-P) algorithm. The TRI-P algorithm basically consists of two processes: First, the data are projected onto a subspace orthogonal to already found endmember signatures (affine transformation), and secondly, maximum  $p$ -norm of the observed pixel vectors is used to identify a new pure pixel (1-D pixel search). The notion of orthogonal subspace projection has been used in EEAs such as VCA [25], automatic target generation process (ATGP) [34], and successive volume maximization (SVMAX) [31], to avoid repeated identification of an endmember. While ATGP and SVMAX were developed from target detection and Winter’s volume maximization framework, respectively, they incidentally turn out to be special cases of the pure pixel search-based TRI-P algorithm, for  $p = 2$ . As will be seen in Section V, TRI-P algorithm shows improved performance over all the benchmark EEAs under test. Some Monte Carlo simulations and experiments with AVIRIS Cuprite data are also presented to demonstrate the merits of the proposed GENE and TRI-P algorithms over some existing benchmark methods.

We should also mention that very recent developments by Chang *et al.* [35], [36] have explored the feasibility of using EEAs to determine the number of endmembers. In [35], two stopping criteria using gradient/difference of the projection error power, and in [36], a Neyman–Pearson detector based on the probability density function (pdf) of the maximum projection error power have been reported. In short, these methods [35], [36] estimate the number of endmembers by implicitly examining if the current endmember estimate is in the range space of the previously found endmembers or not, whereas the proposed GENE algorithms make good use of CH/AH geometry of the hyperspectral data to estimate the number of endmembers, with appealing estimation accuracies.

The following is the organization of this paper. Section II presents the linear mixing model and some general assumptions. The dimension reduction technique, namely affine set fitting is also discussed, following which the convex geometry of the dimension-reduced hyperspectral data is presented in that section. The proposed GENE-CH and GENE-AH algorithms for estimating the number of endmembers in a hyperspectral data cube are presented in Section III. Section IV presents the TRI-P algorithm. In Sections V and VI, the efficacy of the proposed methods is demonstrated through Monte Carlo simulations for various scenarios and by real data experiments, respectively. Finally, some concluding remarks and future directions are given in Section VII.

The notations used in this paper are briefed as follows:  $\mathbb{R}^M$  and  $\mathbb{R}^{M \times N}$  represent the set of real  $M \times 1$  vectors and  $M \times N$  matrices, respectively,  $\mathbf{1}_N$  represents the  $N \times 1$  all-one vector,  $\mathbf{0}$  represents an all-zero vector of proper dimension, and  $\mathbf{I}_N$  is the  $N \times N$  identity matrix.  $\|\mathbf{x}\|_p$  represents the  $p$ -norm of a vector  $\mathbf{x}$ . A Gaussian distribution with mean vector  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$  is denoted as  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ .  $\mathbf{a} \succeq \mathbf{b}$  means that every component in  $\mathbf{a}$  is larger than or equal to the corresponding component in  $\mathbf{b}$ .  $\text{range}[\mathbf{A}]$  denotes the range space of the matrix  $\mathbf{A}$ , and  $\text{diag}(\lambda_1, \dots, \lambda_M)$  denotes a  $\mathbb{R}^{M \times M}$  diagonal matrix whose diagonal elements are  $\lambda_1, \dots, \lambda_M$ .  $P(\cdot)$  denotes the probability function.

## II. SIGNAL MODEL AND DATA PRE PROCESSING

### A. Linear Mixing Model

The linear mixing model is commonly used in analyzing hyperspectral images [1], wherein the observed images are assumed to have undergone a single reflection, and is devoid of any interference due to scattering. In this paper, we consider the linear mixing model, which can also be viewed as a special case of the nonlinear mixing models proposed in [37] and [38]. Specifically, we consider the scenario in which a hyperspectral sensor with  $M$  spectral bands measures solar electromagnetic radiations reflecting from  $N$  distinct substances, over a scene of interest. Owing to low spatial resolution, each pixel vector of the measured hyperspectral image cube can be described by an  $M \times N$  linear mixing model [1], [25], [26], [39], [40]

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] + \mathbf{w}[n] = \sum_{i=1}^N s_i[n]\mathbf{a}_i + \mathbf{w}[n], \forall n = 1, \dots, L \quad (1)$$

where  $M$  is the number of spectral bands and  $N$  is the number of endmembers present in the scene. Further,  $\mathbf{x}[n] = [x_1[n], \dots, x_M[n]]^T$  is the  $n$ th pixel vector in the hyperspectral observation,  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$  denotes the endmember signature matrix whose  $i$ th column vector  $\mathbf{a}_i$  is the  $i$ th endmember signature (or simply endmember),  $\mathbf{s}[n] = [s_1[n], \dots, s_N[n]]^T \in \mathbb{R}^N$  is the  $n$ th abundance vector comprising  $N$  fractional abundances and  $L$  is the total number of observed pixel vectors. Since noise is inevitable in hyperspectral images,  $\mathbf{w}[n] = [w_1[n], \dots, w_M[n]]^T$  represents the zero-mean, uniform/non-uniform additive Gaussian noise vector with the distribution  $\mathcal{N}(\mathbf{0}, \mathbf{D})$ , where  $\mathbf{D} = \text{diag}(\sigma_1^2, \dots, \sigma_M^2)$  in which  $\sigma_i^2 > 0$  denotes the noise variance in the  $i$ th spectral band. If  $\sigma_i^2 = \sigma_j^2, \forall i \neq j$ , then it is called uniform Gaussian noise, else it is called non-uniform Gaussian noise.

The following are the general assumptions [31], [40] associated with (1):

- (A1) (Non-negativity condition)  $s_i[n] \geq 0 \forall i, n$ .
- (A2) (Full additivity condition)  $\sum_{i=1}^N s_i[n] = 1 \forall n$ .
- (A3)  $\min\{L, M\} \geq N$  and  $\mathbf{A}$  is of full column rank.
- (A4) (Pure pixel assumption) There exists at least an index set  $\{l_1, \dots, l_N\}$  such that  $\mathbf{x}[l_i] = \mathbf{a}_i + \mathbf{w}[l_i]$ , for  $i = 1, \dots, N$ .

Assumptions (A1)–(A3) are generally true in hyperspectral images [31], [39]. Assumption (A4) implies that there exists at least one (location unknown) pixel index for each endmember such that the associated observed pixel vector will be fully contributed by that single endmember, and it usually holds true for hyperspectral images taken with a reasonable spatial resolution [41].

### B. Dimension Reduction by Weighted Affine Set Fitting

As in conventional hyperspectral image analysis [1], we begin with dimension reduction of the observed pixel data with the dual aim of reducing the complexity of the ensuing unmixing process and reducing the impact of noise in the hyperspectral data to an extent. For a linear mixing model, it has been discussed in [42] that the true information about the mixing process is fully available in a lower dimensional subspace of dimension  $N - 1$ , which is usually much smaller than  $M$ . It has also been proved in noiseless case that the dimension reduction can be performed due to the following fact [43]:

$$\mathbf{x}[n] \in \text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \text{aff}\{\mathbf{x}[1], \dots, \mathbf{x}[L]\}, \forall n \quad (2)$$

where  $\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$  is the *affine hull* of  $\{\mathbf{a}_1, \dots, \mathbf{a}_N\} \subset \mathbb{R}^M$  and is defined as [44]

$$\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^N \theta_i \mathbf{a}_i \mid \mathbf{1}_N^T \boldsymbol{\theta} = 1, \boldsymbol{\theta} \in \mathbb{R}^N \right\} \quad (3)$$

in which  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T$ . Moreover, due to (A3), the endmember affine hull  $\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$  can be represented as [43]:

$$\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \mathcal{A}(\mathbf{C}, \mathbf{d}) \triangleq \{\mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \mid \boldsymbol{\alpha} \in \mathbb{R}^{N-1}\} \quad (4)$$

for some (non-unique) affine set parameter  $(\mathbf{C}, \mathbf{d}) \in \mathbb{R}^{M \times (N-1)} \times \mathbb{R}^M$ , where  $N - 1$  is the affine dimension

of  $\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ . An optimal solution for  $(\mathbf{C}, \mathbf{d})$  has been provided in [43], where

$$\mathbf{d} = \frac{1}{L} \sum_{n=1}^L \mathbf{x}[n], \quad (5)$$

$$\mathbf{C} = [\mathbf{q}_1 (\mathbf{U}_x \mathbf{U}_x^T), \dots, \mathbf{q}_{N-1} (\mathbf{U}_x \mathbf{U}_x^T)] \quad (6)$$

in which

$$\mathbf{U}_x = [\mathbf{x}[1] - \mathbf{d}, \dots, \mathbf{x}[L] - \mathbf{d}] \in \mathbb{R}^{M \times L} \quad (7)$$

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

However, in practice, the number of endmembers  $N$  is unknown and ought to be estimated. Therefore, by assuming  $N_{\max}$  to be a maximum bound on the number of endmembers, where  $N \leq N_{\max} \leq M$ , similar to the one in [43], we can obtain an affine set fitting parameter  $(\mathbf{C}, \mathbf{d}) \in \mathbb{R}^{M \times (N_{\max}-1)} \times \mathbb{R}^M$ , as below:

$$\mathbf{C} = [\mathbf{q}_1 (\mathbf{U}_x \mathbf{U}_x^T), \dots, \mathbf{q}_{N_{\max}-1} (\mathbf{U}_x \mathbf{U}_x^T)] \quad (8)$$

and  $\mathbf{d}$  is defined in (5). From (6) and (8), it can be easily verified that

$$\mathbf{x}[n] \in \mathcal{A}(\mathbf{C}, \mathbf{d}) \subseteq \mathcal{A}(\mathbf{C}, \mathbf{d}). \quad (9)$$

Then, by virtue of (4), and since  $\mathbf{C}$  given by (8) is semi-unitary, the dimension-reduced pixel vectors  $\tilde{\mathbf{x}}[n]$  can be obtained by the following affine transformation of  $\mathbf{x}[n]$

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^T (\mathbf{x}[n] - \mathbf{d}) \in \mathbb{R}^{N_{\max}-1}. \quad (10)$$

The dimension reduction using affine set fitting can also be viewed as that using PCA [16], though the former is the outcome of convex geometry, and the later is an approach based on second-order statistics of the data.

For the noisy scenario, as given in [39], the approximate affine set fitting parameter  $(\hat{\mathbf{C}}, \hat{\mathbf{d}})$  can be obtained as

$$\hat{\mathbf{d}} = \frac{1}{L} \sum_{n=1}^L \mathbf{x}[n] = \frac{1}{L} \sum_{n=1}^L \mathbf{A} \mathbf{s}[n] + \frac{1}{L} \sum_{n=1}^L \mathbf{w}[n] \cong \mathbf{d} \quad (11)$$

$$\hat{\mathbf{C}} = \left[ \mathbf{q}_1 (\mathbf{U}_x \mathbf{U}_x^T - L \hat{\mathbf{D}}), \dots, \mathbf{q}_{N_{\max}-1} (\mathbf{U}_x \mathbf{U}_x^T - L \hat{\mathbf{D}}) \right] \cong \mathbf{C} \quad (12)$$

where  $\hat{\mathbf{D}}$  is an estimate of the noise covariance matrix  $\mathbf{D}$ . As shown in [39], for given  $\hat{\mathbf{D}}$ , the affine set fitting solution  $(\hat{\mathbf{C}}, \hat{\mathbf{d}})$  [given by (11) and (12)] can be shown to be an approximation to the true  $(\mathbf{C}, \mathbf{d})$  and it asymptotically approaches the true  $(\mathbf{C}, \mathbf{d})$  for large  $L$ . In practical situations, the multiple-regression-analysis-based noise covariance estimation method reported in HySiMe [18] can be used to estimate  $\mathbf{D}$ .

Further, in the noisy scenario, by substituting (11) and (12) in (1) for  $\mathbf{d}$  and  $\mathbf{C}$ , and due to (1) and (A2), we have

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^N s_i[n] \alpha_i + \tilde{\mathbf{w}}[n], \quad n = 1, \dots, L \quad (13)$$

where

$$\alpha_i = \hat{\mathbf{C}}^T (\mathbf{a}_i - \hat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}, \quad i = 1, \dots, N \quad (14)$$

is the  $i$ th dimension-reduced endmember, and  $\tilde{\mathbf{w}}[n] \triangleq \hat{\mathbf{C}}^T \mathbf{w}[n] \sim \mathcal{N}(\mathbf{0}, \Sigma)$ , in which

$$\Sigma = \hat{\mathbf{C}}^T \mathbf{D} \hat{\mathbf{C}} \in \mathbb{R}^{(N_{\max}-1) \times (N_{\max}-1)}. \quad (15)$$

The relation between the dimension-reduced endmember  $\alpha_i$  and the true endmember  $\mathbf{a}_i$  is given by

$$\mathbf{a}_i = \hat{\mathbf{C}} \alpha_i + \hat{\mathbf{d}}, \quad i = 1, \dots, N. \quad (16)$$

It is worth mentioning that from (13) and under (A4), we have

$$\tilde{\mathbf{x}}[l_i] = \alpha_i + \tilde{\mathbf{w}}[l_i], \quad \forall i = 1, \dots, N \quad (17)$$

which is essential in the development of the proposed GENE algorithms in Section III.

### C. Convex Geometry of the Hyperspectral Data

In this subsection we will present the convex geometry of the noise-free dimension-reduced hyperspectral data given by (13) for which  $\tilde{\mathbf{w}}[n] = \mathbf{0}, \forall n$ . The convex geometry will lay a solid platform for the ensuing sections, though the presence of noise will be taken into account therein. Before proceeding further, let us introduce an important concept in convex analysis. The *convex hull* of a set of vectors  $\{\alpha_1, \dots, \alpha_N\} \subset \mathbb{R}^{N_{\max}-1}$  is defined as

$$\text{conv}\{\alpha_1, \dots, \alpha_N\} = \left\{ \mathbf{x} = \sum_{i=1}^N \theta_i \alpha_i \mid \mathbf{1}_N^T \boldsymbol{\theta} = 1, \boldsymbol{\theta} \succeq \mathbf{0} \right\} \quad (18)$$

where  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T$ . A convex hull,  $\text{conv}\{\alpha_1, \dots, \alpha_N\}$  is called an  $N-1$  dimensional simplex in  $\mathbb{R}^{N_{\max}-1}$  if  $\{\alpha_1, \dots, \alpha_N\} \subset \mathbb{R}^{N_{\max}-1}$  is affinely independent, and the simplex  $\text{conv}\{\alpha_1, \dots, \alpha_N\}$  has only  $N$  extreme points  $\alpha_1, \dots, \alpha_N$  [44]. Then, based on (13), we have the following facts.

(F1) In the noise-free case, by (A1)–(A4), any dimension-reduced pixel vectors  $\tilde{\mathbf{x}}[n]$  lie in the convex hull of the dimension-reduced endmember signatures, and

$$\text{conv}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\} = \text{conv}\{\alpha_1, \dots, \alpha_N\} \quad (19)$$

in which  $\text{conv}\{\alpha_1, \dots, \alpha_N\}$  is a simplex with  $N$  extreme points being  $\alpha_1, \dots, \alpha_N$ . A more general case of (F1) can be obtained by relaxing (A1) and (A4), as stated next.

(F2) In the noise-free case, by (A2) and (A3), any dimension-reduced pixel vectors  $\tilde{\mathbf{x}}[n]$  lie in the affine hull of the dimension-reduced endmember signatures, and

$$\text{aff}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\} = \text{aff}\{\alpha_1, \dots, \alpha_N\} \quad (20)$$

with the affine dimension equal to  $N-1$ .

A simple illustration of (F1) and (F2), for  $N=3$  case is shown in Fig. 1. These geometrical properties of the observed hyperspectral data play a significant role in the proposed algorithms for estimating the number of endmembers, which will be presented in the next section.

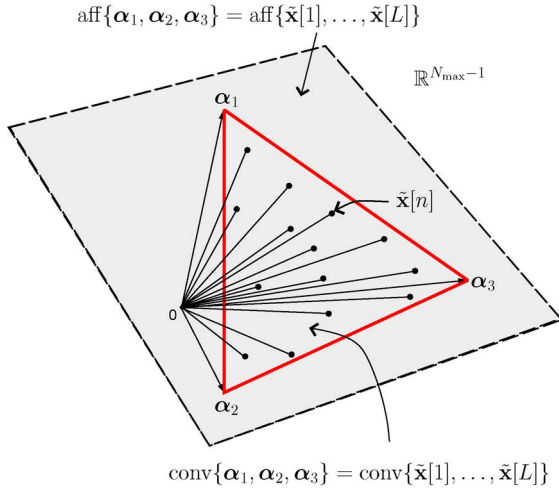


Fig. 1. Illustration of (F1) and (F2), for  $N = 3$  case.

### III. GEOMETRY-BASED ESTIMATION OF NUMBER OF ENDMEMBERS

In this section, we make use of the key geometric characteristics of the observed dimension-reduced hyperspectral data, i.e., (F1) and (F2) to systematically estimate the number of endmembers present in the hyperspectral data. Thus the proposed algorithms are aptly named as geometry-based estimation of number of endmembers (GENE) algorithms. In the first subsection, we propose the GENE-CH algorithm based on the convex hull geometry (F1), provided that (A4) holds true. However, for data with (A4) violated, the GENE-AH algorithm is proposed in the subsequent subsection. GENE-AH is based on (F2) and it turns out to be more robust than the GENE-CH algorithm against absence of pure pixels in the data (which will be confirmed by the simulations in Section V).

#### A. GENE-CH Algorithm

In this subsection we assume that (A4) holds true. Suppose that a reliable, successive EEA has found the pixel indices  $l_1, \dots, l_N, l_{N+1}, \dots, l_{k-1}, l_k$ , in which  $l_1, \dots, l_N$  are pure pixel indices and the rest are not. Here,  $l_k$  is the current pixel index estimate and  $\{l_1, l_2, \dots, l_{k-1}\}$  are the previously found pixel index estimates, and  $k \leq N_{\max}$ . Then by (13) and (17), it can be readily inferred that

$$\tilde{\mathbf{x}}[l_i] = \beta_i + \tilde{\mathbf{w}}[l_i], \quad i = 1, \dots, k \quad (21)$$

where

$$\beta_i = \begin{cases} \alpha_i, & i = 1, \dots, N, \\ \sum_{j=1}^N s_j[l_i] \alpha_j, & i = N+1, \dots, k. \end{cases} \quad (22)$$

To explain the idea behind GENE-CH, let us first consider the noise-free scenario, i.e.,  $\tilde{\mathbf{w}}[l_i] = \mathbf{0}, \forall i = 1, \dots, k$ , in (21). Recall from (F1) that the total number of extreme points in  $\text{conv}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\}$  is  $N$ . That is to say, if  $\tilde{\mathbf{x}}[l_k] = \beta_k$  cannot contribute a new extreme point to the  $\text{conv}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}], \tilde{\mathbf{x}}[l_k]\}$ , i.e.,

$$\begin{aligned} \text{conv}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}], \tilde{\mathbf{x}}[l_k]\} &= \text{conv}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}]\} \\ &= \text{conv}\{\beta_1, \dots, \beta_{k-1}\} \end{aligned}$$

or in other words, if  $\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , then it can be inferred by (22) that all the endmembers are already found, that is  $k \geq N + 1$ . Therefore, the smallest  $k$  such that  $\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , must take the value of  $N + 1$ , and thus  $N$  can be estimated as  $k - 1$ , provided that the smallest  $k$  can be reliably estimated. However, in a real scenario, since only noisy  $\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_k]$  are available (rather than  $\beta_1, \dots, \beta_k$ ), in the process of estimating the number of endmembers, the presence of noise in the  $\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_k]$  must be taken into account. To this end, we propose a Neyman–Pearson hypothesis [45] testing-based method to determine whether  $\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , or not, based on noisy  $\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_k]$ . The details are as follows.

Let us consider the following constrained least squares problem:

$$\theta^* = \arg \min_{\theta \succeq \mathbf{0}, \mathbf{1}_{k-1}^T \theta = 1} \left\| \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta \right\|_2^2 \quad (23)$$

where

$$\hat{\mathbf{A}}_{k-1} = [\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}]] \in \mathbb{R}^{(N_{\max}-1) \times (k-1)}. \quad (24)$$

The optimization problem in (23) is convex and can be solved by using available convex optimization solvers such as SeDuMi[46] and CVX[47]. We define the fitting error vector  $e \in \mathbb{R}^{N_{\max}-1}$  as below:

$$\begin{aligned} e &= \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta^* \\ &= \mu_k + \left( \tilde{\mathbf{w}}[l_k] - \sum_{i=1}^{k-1} \theta_i^* \tilde{\mathbf{w}}[l_i] \right), \quad (\text{by (21)}) \end{aligned} \quad (25) \quad (26)$$

where

$$\mu_k = \beta_k - \sum_{i=1}^{k-1} \theta_i^* \beta_i. \quad (27)$$

Then the following can be observed from (26):

- If  $\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , then it implies that  $\beta_k - \sum_{i=1}^{k-1} \theta_i^* \beta_i = \mathbf{0}$ , for some  $\theta' = [\theta'_1, \dots, \theta'_{k-1}]^T \succeq \mathbf{0}$ ,  $\mathbf{1}_{k-1}^T \theta' = 1$ . In the noise-free case, and when the endmembers are perfectly identified,  $\theta^* = \theta'$ . However, in real scenarios the optimal (in least-squares error sense) estimate  $\theta^*$  is a good approximation to  $\theta'$  for higher signal-to-noise ratio (SNR), which makes  $\mu_k \simeq \mathbf{0}$ . The approximation error increases as the SNR decreases. As will be seen in the simulations in Section V, this approximation holds reasonably well even for moderately low SNRs. Therefore,  $e$  can be approximated as a zero-mean Gaussian random vector, i.e.,  $e \sim \mathcal{N}(\mathbf{0}, \xi^* \Sigma)$ , where

$$\xi^* = 1 + \theta_1^{*2} + \theta_2^{*2} + \dots + \theta_{k-1}^{*2} \quad (28)$$

and  $\Sigma$  is given by (15).

- If  $\beta_k \notin \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , then  $e \sim \mathcal{N}(\mu_k, \xi^* \Sigma)$  is a non-zero mean Gaussian random vector.

Now define

$$r = e^T (\xi^* \Sigma)^{-1} e. \quad (29)$$

When  $\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}$ , it is easy to see that  $r$  can be approximated as a central Chi-square distributed random variable, and otherwise  $r$  is a non-central Chi-square distributed random variable [48]. In both cases, the degrees of freedom is  $N_{\max} - 1$ . Hence, we consider the following two hypotheses:

$$H_0 (\beta_k \in \text{conv}\{\beta_1, \dots, \beta_{k-1}\}) : r \sim f_{\chi^2}(x, N_{\max} - 1) \tag{30a}$$

$$H_1 (\beta_k \notin \text{conv}\{\beta_1, \dots, \beta_{k-1}\}) : r \sim f_{N\chi^2}(x, N_{\max} - 1, \|\mu_k\|_2^2). \tag{30b}$$

Here,  $f_{\chi^2}(x, N_{\max} - 1)$  is the pdf of central Chi-square distribution and is given by [48]

$$f_{\chi^2}(x, k) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{31}$$

where  $\Gamma(k/2)$  denotes the Gamma function. However, the non-central Chi-square pdf  $f_{N\chi^2}(x, N_{\max} - 1, \|\mu_k\|_2^2)$  is unknown, as  $\mu_k$  is unknown. Therefore, we use Neyman-Pearson classifier rule for the hypothesis testing problem:

$$\text{Decide } H_0 \quad \text{if } r < \eta \tag{32a}$$

$$\text{Decide } H_1 \quad \text{if } r > \eta \tag{32b}$$

where  $\eta$  can be found by minimizing the  $P(H_0|H_1)$  subject to  $P(H_1|H_0) \leq P_{\text{FA}}$ , in which  $P_{\text{FA}}$  is the preassigned acceptable false alarm rate. Obviously, the optimal value of  $\eta$  should satisfy [45]

$$\int_{\eta}^{\infty} f_{\chi^2}(x, N_{\max} - 1) dx = P_{\text{FA}}.$$

Hence, the decision rules in (32) can be equivalently written as

$$\text{Decide } H_0 \quad \text{if } \int_r^{\infty} f_{\chi^2}(x, N_{\max} - 1) dx > P_{\text{FA}} \tag{33a}$$

$$\text{Decide } H_1 \quad \text{if } \int_r^{\infty} f_{\chi^2}(x, N_{\max} - 1) dx < P_{\text{FA}}. \tag{33b}$$

The integral in (33) can be easily computed as follows:

$$\int_r^{\infty} f_{\chi^2}(x, N_{\max} - 1) dx = 1 - \frac{\gamma\left(\frac{r}{2}, \frac{N_{\max}-1}{2}\right)}{\Gamma\left(\frac{N_{\max}-1}{2}\right)} \tag{34}$$

where  $\gamma(x/2, (N_{\max} - 1)/2)$  is the lower incomplete Gamma function [49]. Once the integral is evaluated, one of the hypotheses should be true, based on (33). The entire procedure for GENE-CH is summarized in Table I.

**B. GENE-AH Algorithm**

Recall that the GENE-CH algorithm is based on the assumption that the pure pixels are present in the data (i.e., (A4) holds true). However, for practical hyperspectral data the presence of

TABLE I  
PSEUDOCODE FOR GENE-CH AND GENE-AH ALGORITHMS

---

<b>Given</b>	noisy hyperspectral data $\mathbf{x}[n]$ , maximum number of endmembers $N \leq N_{\max} \leq M$ , false alarm probability $P_{\text{FA}}$ , and estimate of noise covariance matrix $\hat{\mathbf{D}}$ .
<b>Step 1.</b>	Compute $(\hat{\mathbf{C}}, \hat{\mathbf{d}})$ given by (11) and (12).
<b>Step 2.</b>	Obtain the first pixel index $l_1$ by a successive EEA and compute $\tilde{\mathbf{x}}[l_1] = \hat{\mathbf{C}}^T(\mathbf{x}[l_1] - \hat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}$ . Set $k = 2$ .
<b>Step 3.</b>	Obtain the $k$ th pixel index $l_k$ using the successive EEA and compute $\tilde{\mathbf{x}}[l_k] = \hat{\mathbf{C}}^T(\mathbf{x}[l_k] - \hat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}$ and form $\hat{\mathbf{A}}_{k-1} = [\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}]] \in \mathbb{R}^{(N_{\max}-1) \times (k-1)}$ .
<b>Step 4.</b>	Solve the following:  GENE-CH : $\theta^* = \arg \min_{\substack{\theta \geq 0, \\ \mathbf{1}_{k-1}^T \theta = 1}} \ \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta\ _2^2$ ,  GENE-AH : $\theta^* = \arg \min_{\mathbf{1}_{k-1}^T \theta = 1} \ \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta\ _2^2$ ,  and calculate $e = \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta^*$ .
<b>Step 5.</b>	Compute $r = e^T (\xi^* \Sigma)^{-1} e$ , where $\xi^* = \mathbf{1} + \theta^{*T} \theta^*$ and $\Sigma = \hat{\mathbf{C}}^T \hat{\mathbf{D}} \hat{\mathbf{C}}$ .
<b>Step 6.</b>	Calculate $\psi = \int_r^{\infty} f_{\chi^2}(x, N_{\max} - 1) dx$ by (34).
<b>Step 7.</b>	If $\psi > P_{\text{FA}}$ , then output $k - 1$ as the estimate for number of endmembers, else $k := k + 1$ and if $k < N_{\max}$ go to <b>Step 3</b> .

---

pure pixels cannot be guaranteed. In this case, the dimension-reduced endmembers estimated by an EEA can be expressed in general as in (21), where

$$\beta_i = \sum_{j=1}^N s_j[l_i] \alpha_j, \quad \forall i = 1, \dots, k. \tag{35}$$

Therefore, GENE-CH may not provide an accurate estimate of the number of endmembers. A pictorial illustration is given in Fig. 2, where  $N = 3$  endmembers  $\alpha_1, \alpha_2, \alpha_3$  are not present in the noise-free hyperspectral data. For this case, the endmember estimates, denoted by  $\beta_i, i = 1, \dots, N_{\max} = 6$ , obtained by an EEA are shown in Fig. 2(a) and can be expressed as [by (21) and (35)]

$$\beta_i = \tilde{\mathbf{x}}[l_i] = \sum_{j=1}^3 s_j[l_i] \alpha_j, \quad i = 1, \dots, N_{\max} = 6 \tag{36}$$

where  $l_1, \dots, l_6$  are the pixel indices provided by the EEA under consideration. Then, as can be inferred from Fig. 2(a), for the  $\text{conv}\{\beta_1, \dots, \beta_6\}$ , there can be more than 3 extreme points which in fact is 6 in this case, i.e.,

$$\beta_k \notin \text{conv}\{\beta_1, \dots, \beta_{k-1}\}, \quad k = 2, 3, 4, 5, 6 \tag{37}$$

which means that the hypothesis  $H_1$  given by (30b) will be true even for  $k > N = 3$ . Hence, using the fact (F1) will obviously result in an overestimation of the number of endmembers for this case. However, from Fig. 2(b), it can be readily inferred that

$$\begin{aligned} \beta_k &\notin \text{aff}\{\beta_1, \dots, \beta_{k-1}\}, & k = 2, 3 \\ \beta_k &\in \text{aff}\{\beta_1, \dots, \beta_{k-1}\}, & k = 4, 5, 6. \end{aligned}$$

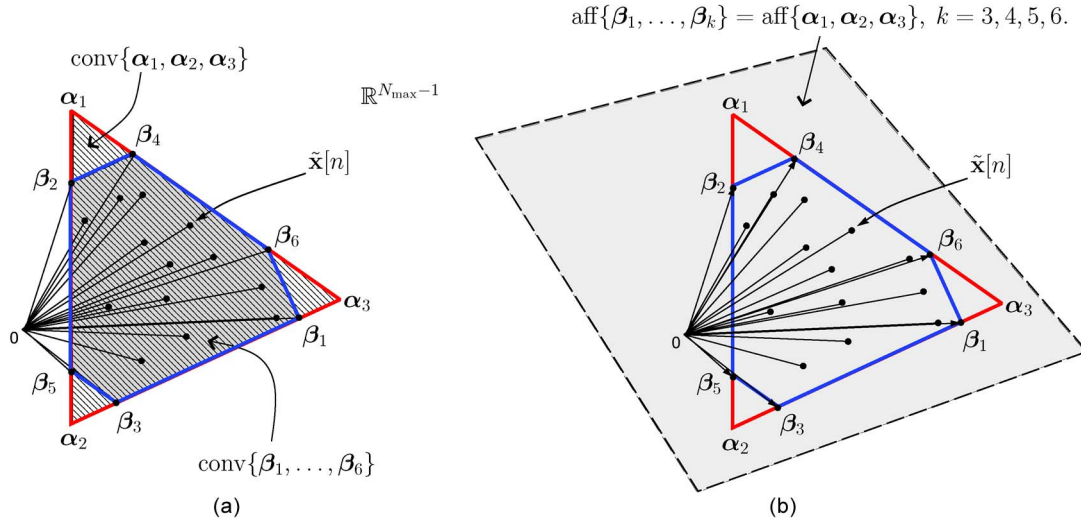


Fig. 2. Illustration of GENE-CH algorithm, when no pure pixel is present in the noise-free hyperspectral data ( $N = 3$  case). (a) The endmember estimates are denoted by  $\beta_i$ ,  $i = 1, \dots, N_{\max} = 6$ , but  $\text{conv}\{\beta_1, \dots, \beta_6\} \neq \text{conv}\{\alpha_1, \alpha_2, \alpha_3\}$  because the true endmembers  $\alpha_1, \alpha_2, \alpha_3$  are not present in the data cloud, whereas  $\text{aff}\{\beta_1, \dots, \beta_k\} = \text{aff}\{\alpha_1, \alpha_2, \alpha_3\}$ ,  $k = 3, 4, 5, 6$ , as shown in (b).

Motivated by the above illustration, we next propose the GENE-AH algorithm.

The GENE-AH algorithm uses the fact (F2), which states that in the noise-free case, the affine dimension of  $\text{aff}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\}$  is  $N - 1$ . This implies that in the noise-free case, if  $\tilde{\mathbf{x}}[l_k] = \beta_k$  cannot contribute an increment to the affine dimension of  $\text{aff}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}], \tilde{\mathbf{x}}[l_k]\}$ , i.e.,

$$\begin{aligned} \text{aff}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}], \tilde{\mathbf{x}}[l_k]\} &= \text{aff}\{\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_{k-1}]\} \\ &= \text{aff}\{\beta_1, \dots, \beta_{k-1}\} \end{aligned}$$

or in other words, if  $\beta_k \in \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ , then  $k \geq N + 1$ . Therefore, the smallest  $k$  such that  $\beta_k \in \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ , must take the value of  $N + 1$ , and thus  $N$  can be estimated as  $k - 1$ . As presented in Section III-A, again we use Neyman–Pearson hypothesis [45] testing to determine whether  $\beta_k \in \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ , or not, based on noisy  $\tilde{\mathbf{x}}[l_1], \dots, \tilde{\mathbf{x}}[l_k]$ . The details are as follows:

As in (23), we consider solving the following constrained least squares problem:

$$\theta^* = \arg \min_{\mathbf{1}_{k-1}^T \theta = 1} \left\| \tilde{\mathbf{x}}[l_k] - \hat{\mathbf{A}}_{k-1} \theta \right\|_2^2 \quad (38)$$

where  $\hat{\mathbf{A}}_{k-1}$  is defined in (24). Again, since (38) is convex,  $\theta^*$  can be obtained by available convex optimization solvers [46], [47]. By defining the fitting error vector  $e$  as in (25), we have the following inferences:

- if  $\beta_k \in \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ , then it can be approximated that  $e \sim \mathcal{N}(\mathbf{0}, \xi^* \Sigma)$ ;
- if  $\beta_k \notin \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ , then  $e \sim \mathcal{N}(\mu_k, \xi^* \Sigma)$ ,

where  $\mu_k$ ,  $\xi^*$ , and  $\Sigma$  are defined in (27), (28), and (15), respectively. Defining the random variable  $r$  as in (29), a similar Neyman–Pearson hypothesis testing procedure can be devised for GENE-AH to estimate the number of endmembers present in the data. As will be seen from the simulations (see Section V), the GENE-AH algorithm yields better performance

demonstrating its robustness against the absence of pure pixels in the data. The procedure for GENE-AH is also given in Table I and is similar to that of GENE-CH, except that in Step 4, the optimal  $\theta^*$  is obtained by solving (38).

Simply speaking, the GENE algorithms are based on convex geometry, which in turn is based on the standard assumptions (A1)–(A4). Though the validity of assumption (A4) is debated over the years, the other assumptions (A1) to (A3) are believed to hold true for hyperspectral data, and are extensively employed in EEAs and HU algorithms [1]. However, as detailed in [50], there may be some real situations in which both (A2) and (A4) could be violated. For such a scenario, the GENE-AH algorithm presented above can be modified so as to accurately estimate the number of endmembers. The idea behind the modified GENE-AH (GENE-AH-MOD) algorithm is based only on (A3), and is explained next.

For the data with (A2) and (A4) not satisfied, any EEA should output the dimension reduced endmembers  $\{\beta_1, \dots, \beta_k\}$  [given by (35)], which may not be purest pixels, but they can ensure that  $\{\beta_1, \dots, \beta_k\} \subset \mathbb{R}^{N_{\max}-1}$  must be in the range space of  $[\alpha_1, \dots, \alpha_N] \in \mathbb{R}^{(N_{\max}-1) \times N}$ . Following the footsteps of GENE-AH, and assuming that only (A3) holds true, the aim now is to find the smallest  $k$  such that  $\beta_k \in \text{range}[\beta_1, \dots, \beta_{k-1}]$ , and thus  $N$  can be estimated as  $k - 1$ . Based on this idea, we next modify the GENE-AH algorithm by considering the following lemma.

**Lemma 1:** Suppose that there exists  $\beta_k \in \text{range}[\beta_1, \dots, \beta_{k-1}]$  and  $\beta_k \notin \text{aff}\{\beta_1, \dots, \beta_{k-1}\}$ . Then, it holds true that  $\text{range}[\beta_1, \dots, \beta_{k-1}] = \text{aff}\{\beta_1, \dots, \beta_{k-1}, \beta_k\}$ .

*Proof:* We first prove that  $\text{aff}\{\beta_1, \dots, \beta_{k-1}, \beta_k\} \subset \text{range}[\beta_1, \dots, \beta_{k-1}]$ . Let  $\mathbf{Q} = [\beta_1, \dots, \beta_{k-1}] \in \mathbb{R}^{(N_{\max}-1) \times (k-1)}$ , and let  $\beta_k = \mathbf{Q}\zeta$ , where  $\zeta \in \mathbb{R}^{k-1}$  satisfying  $\mathbf{1}_{k-1}^T \zeta \neq 1$ . Then, any vector  $\mathbf{q} \in \text{aff}\{\beta_1, \dots, \beta_{k-1}, \beta_k\}$  can be expressed as

$$\mathbf{q} = \mathbf{Q}\nu + (1 - \mathbf{1}_{k-1}^T \nu) \beta_k = \mathbf{Q}(\nu + \zeta - \zeta \mathbf{1}_{k-1}^T \nu) \quad (39)$$

$$= \mathbf{Q}[(\mathbf{I}_{k-1} - \zeta \mathbf{1}_{k-1}^T) \nu + \zeta], \quad (40)$$

where  $\boldsymbol{\nu} \in \mathbb{R}^{k-1}$ . Furthermore, since  $\mathbf{1}_{k-1}^T \boldsymbol{\zeta} \neq 1$ , based on Proposition 1 in [51], we can easily see that the only non-zero eigenvalue of  $\boldsymbol{\zeta} \mathbf{1}_{k-1}^T$  is not equal to unity, implying that the matrix  $\mathbf{I}_{k-1} - \boldsymbol{\zeta} \mathbf{1}_{k-1}^T$  is non-singular. Let  $\boldsymbol{\mu} = (\mathbf{I}_{k-1} - \boldsymbol{\zeta} \mathbf{1}_{k-1}^T) \boldsymbol{\nu} + \boldsymbol{\zeta}$  which is an invertible affine transformation from  $\boldsymbol{\nu}$ . By (40) the vector  $\mathbf{q} = \mathbf{Q}\boldsymbol{\mu}$  belongs to  $\text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}]$ , implying that  $\text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_k\} \subset \text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}]$ .

Next, we show that  $\text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}] \subset \text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_k\}$ . Any vector  $\mathbf{q} \in \text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}]$  can be represented by  $\mathbf{q} = \mathbf{Q}\boldsymbol{\mu}$  where  $\boldsymbol{\mu} \in \mathbb{R}^{k-1}$ , and such a  $\boldsymbol{\mu}$  can always find a one-to-one affine mapping from a  $\boldsymbol{\nu} \in \mathbb{R}^{k-1}$ ; i.e.,

$$\boldsymbol{\mu} = (\mathbf{I}_{k-1} - \boldsymbol{\zeta} \mathbf{1}_{k-1}^T) \boldsymbol{\nu} + \boldsymbol{\zeta} \quad (41)$$

for any  $\boldsymbol{\zeta}$  satisfying  $\mathbf{1}_{k-1}^T \boldsymbol{\zeta} \neq 1$ . Hence,  $\mathbf{q}$  can be expressed as

$$\mathbf{q} = \mathbf{Q}\boldsymbol{\nu} + (1 - \mathbf{1}_{k-1}^T \boldsymbol{\nu}) \mathbf{Q}\boldsymbol{\zeta}. \quad (42)$$

As  $\boldsymbol{\beta}_k = \mathbf{Q}\boldsymbol{\zeta}$ ,  $\mathbf{q}$  belongs to  $\text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_k\}$  [given by (39)], implying that  $\text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}] \subset \text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_k\}$ . Thus, it is proved that the two sets  $\text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}]$  and  $\text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}, \boldsymbol{\beta}_k\}$  are identical. ■

Lemma 1 implies that the smallest  $k$  such that  $\boldsymbol{\beta}_k \in \text{range}[\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{k-1}]$  is equivalent to finding the smallest  $t = k + 1$  such that  $\boldsymbol{\beta}_t \in \text{aff}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_{t-1}\}$ . Hence, following the Neyman–Pearson hypothesis testing procedure in GENE-AH,  $N$  can be estimated as  $k - 1 = t - 2$ , and for the sake of differentiation, the algorithm is named as GENE-AH-MOD. In summary, the estimated number of endmembers using GENE-AH-MOD corresponds to one less than that of GENE-AH.

Some conceptual distinctions between the proposed GENE algorithms and some existing benchmark algorithms are as follows. Existing algorithms for the estimation of number of endmembers, such as HySiMe [18] and orthogonal subspace projection (OSP) [35], are developed based on the projection power on the most apt range space of the hyperspectral data, which in turn is based only on (A3) (as the case with GENE-AH-MOD, discussed above). In other words, those methods consider the following fact:  $\tilde{\mathbf{x}}[l_i] \in \text{range}[\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N]$ . However, GENE-CH and GENE-AH involve convex hull and affine hull i.e.,

$$\begin{aligned} \tilde{\mathbf{x}}[l_i] &\in \text{conv}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\} \subset \text{aff}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\} \\ &\subset \text{range}[\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N] \end{aligned}$$

as they not only make use of (A3), but also (A2) for GENE-AH algorithm, and (A1), (A2), and (A4) for GENE-CH algorithm. The estimation accuracy of GENE-CH, GENE-AH, and GENE-AH-MOD algorithms on a given data, depends on the validity of the assumptions for the given data. In Section V, some sophisticated simulations will be performed to demonstrate the performances of GENE-CH, GENE-AH, and GENE-AH-MOD algorithms, by testing them using simulation data generated with realistic combinations of these assumptions.

It should be noted that the estimation accuracies of the GENE algorithms depend on the performance of the EEA used. Hence,

in the next section, we propose a reliable (with theoretical support for endmember identifiability), reproducible (without any initialization), and successive EEA, namely  $p$ -norm-based pure pixel identification (TRI-P) algorithm.

#### IV. $p$ -NORM-BASED PURE PIXEL IDENTIFICATION (TRI-P)

The proposed TRI-P algorithm aims to find the pure pixel indices (and thereby the endmembers) from the hyperspectral observations. Throughout the derivation of the TRI-P algorithm, we focus on a noise-free signal model by means of which the endmember identifiability of the TRI-P algorithm can be theoretically proved (noisy scenarios will be considered in simulations in Section V). We begin by incorporating the assumption (A2) in (13) so that we have the following augmented dimension-reduced data:

$$\bar{\mathbf{x}}[n] = \begin{bmatrix} \tilde{\mathbf{x}}[n] \\ 1 \end{bmatrix} = \sum_{i=1}^N s_i[n] \bar{\boldsymbol{\alpha}}_i \in \mathbb{R}^{N_{\max}} \quad (43)$$

where

$$\bar{\boldsymbol{\alpha}}_i = [\boldsymbol{\alpha}_i^T \mathbf{1}]^T, \quad i = 1, \dots, N \quad (44)$$

are the augmented dimension-reduced endmembers.

We now find the first dimension-reduced endmember by  $p$ -norm maximization procedure as follows: Considering the  $p$ -norm of all the pixel vectors in the augmented dimension-reduced data cloud  $\bar{\mathbf{X}} = [\bar{\mathbf{x}}[1], \dots, \bar{\mathbf{x}}[L]]$ , by the triangle inequality, (A1), and (A2), one can infer from (43) that for all  $n$

$$\|\bar{\mathbf{x}}[n]\|_p \leq \sum_{i=1}^N s_i[n] \cdot \|\bar{\boldsymbol{\alpha}}_i\|_p \leq \max_{i=1, \dots, N} \{\|\bar{\boldsymbol{\alpha}}_i\|_p\} \quad (45)$$

where  $p \geq 1$ . The equality in (45) holds if and only if  $n = l_i$  (by (A4)) for any  $i \in \arg \max_{k=1, \dots, N} \{\|\bar{\boldsymbol{\alpha}}_k\|_p\}$ . Thus, a dimension-reduced endmember can be identified by

$$\boldsymbol{\alpha}_1 = \tilde{\mathbf{x}}[l_1] \text{ for any } l_1 \in \arg \max_{n=1, \dots, L} \{\|\bar{\mathbf{x}}[n]\|_p\}. \quad (46)$$

Once the first endmember is found, the other endmembers  $\boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_N$  can be obtained successively by the following general procedure: Suppose that the augmented dimension-reduced endmembers [given by (44)]  $\bar{\boldsymbol{\alpha}}_1, \dots, \bar{\boldsymbol{\alpha}}_{k-1}$  (where  $k - 1 < N$ ) are already identified. Let

$$\mathbf{Q} = [\bar{\boldsymbol{\alpha}}_1, \dots, \bar{\boldsymbol{\alpha}}_{k-1}] \in \mathbb{R}^{N_{\max} \times (k-1)}. \quad (47)$$

To find a new endmember different from  $\bar{\boldsymbol{\alpha}}_1, \dots, \bar{\boldsymbol{\alpha}}_{k-1}$ , we consider the following orthogonal complement subspace projection:

$$\mathbf{P}_{\mathbf{Q}}^\perp \bar{\mathbf{x}}[n] = \sum_{i=k}^N s_i[n] \mathbf{P}_{\mathbf{Q}}^\perp \bar{\boldsymbol{\alpha}}_i, \quad \forall n, \quad (\text{by (43)}) \quad (48)$$



where  $\mathbf{P}_Q^\perp = \mathbf{I}_{N_{\max}} - \mathbf{Q}(\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T$  is the orthogonal complement projector of  $\mathbf{Q}$ . Again, by the triangle inequality, (A1), (A2), and (48), we have

$$\|\mathbf{P}_Q^\perp \bar{\mathbf{x}}[n]\|_p \leq \sum_{i=k}^N s_i[n] \cdot \|\mathbf{P}_Q^\perp \bar{\alpha}_i\|_p \leq \max_{i=k, \dots, N} \left\{ \|\mathbf{P}_Q^\perp \bar{\alpha}_i\|_p \right\}. \quad (49)$$

The equality in (49) holds if and only if  $n = l_j$  (by (A4)) for any  $j \in \arg \max_{i=k, \dots, N} \{\|\mathbf{P}_Q^\perp \bar{\alpha}_i\|_p\}$ . Therefore, one can find a new dimension-reduced endmember as

$$\alpha_j = \bar{\mathbf{x}}[l_j] \text{ for any } l_j \in \arg \max_{n=1, \dots, L} \left\{ \|\mathbf{P}_Q^\perp \bar{\mathbf{x}}[n]\|_p \right\} \quad (50)$$

and  $\alpha_j = \bar{\mathbf{x}}[l_j] \notin \{\alpha_1, \dots, \alpha_{k-1}\}$ .

The above endmember estimation methodology is called the TRI-P algorithm, and it can identify all the dimension-reduced endmembers, as stated in the following lemma.

*Lemma 2:* Under (A1)–(A4), with  $N$  known, and in the absence of noise, TRI-P algorithm yields  $\{\alpha_1, \dots, \alpha_N\}$  such that the simplex  $\text{conv}\{\alpha_1, \dots, \alpha_N\} = \text{conv}\{\bar{\mathbf{x}}[1], \dots, \bar{\mathbf{x}}[L]\}$ .

Once the dimension-reduced endmembers are found, the corresponding endmembers can be obtained by (16).

At this juncture, it is worthwhile to point out some characteristics of the proposed TRI-P algorithm.

- Lemma 2 is valid only if (A4) is satisfied and the number of endmembers  $N$  is perfectly known. However, if (A4) is satisfied and  $N$  is unknown (which is the case in the proposed GENE-CH algorithm), then  $N_{\max}$  pixel indices can be obtained by TRI-P algorithm, where the first  $N$  pixel indices will be a set of pure pixel indices. In this case, an interesting question is: What will be the pixel index  $l_k$  obtained by TRI-P algorithm when  $k > N$ ? It can be shown from (13) and (A3) that in the noiseless case  $\bar{\mathbf{x}}[n] \in \text{range}\{\bar{\alpha}_1, \dots, \bar{\alpha}_N\}$ . Therefore, under (A1)–(A4), when finding  $\alpha_k$  for  $k > N$ , we have by (47) that  $\|\mathbf{P}_Q^\perp \bar{\mathbf{x}}[n]\|_p = 0, \forall n$ . Equation (50) will therefore yield a pixel index  $l_k$  for which

$$\bar{\mathbf{x}}[l_k] = \sum_{j=1}^N s_j[l_k] \alpha_j, \quad \forall k > N \quad (51)$$

can be any pixel in the data cloud.

- On the other hand, if (A4) is not satisfied and  $N$  is perfectly known. Then the set of pixel indices  $\{l_1, \dots, l_N\}$  corresponds to a set of so-called “purest pixels” available in the data cloud.
- Finally, if both (A4) is not satisfied and  $N$  is unknown (which is the case in the proposed GENE-AH algorithm), then still  $N_{\max}$  pixel indices can be obtained, where the first  $N$  pixel indices will be indices corresponding to a set of so-called “purest pixels” in the data cloud, whereas pixel index  $l_k$  when  $k > N$  can be any pixel in the data cloud, i.e.,

$$\bar{\mathbf{x}}[l_k] = \sum_{j=1}^N s_j[l_k] \alpha_j, \quad \forall N < k \leq N_{\max}. \quad (52)$$

TABLE II  
p-NORM-BASED PURE PIXEL (TRI-P) ALGORITHM

<b>Given</b>	dimension-reduced observations $\bar{\mathbf{x}}[n]$ , $\bar{\mathbf{x}}[n]$ given by (43), and maximum number of endmembers $N_{\max}$ .
<b>Step 1.</b>	Set $k = 1$ . Obtain $\bar{\alpha}_1 = \bar{\mathbf{x}}[l_1]$ for any $l_1 \in \arg \max_n \{\ \bar{\mathbf{x}}[n]\ _p\}$ . Let $\mathbf{Q} = \bar{\alpha}_1$ .
<b>Step 2.</b>	Update $k := k + 1$ and obtain $\bar{\alpha}_k = \bar{\mathbf{x}}[l_k]$ for any $l_k \in \arg \max_n \{\ \mathbf{P}_Q^\perp \bar{\mathbf{x}}[n]\ _p\}$ .
<b>Step 3.</b>	Update $\mathbf{Q} := [\mathbf{Q} \ \bar{\alpha}_k] \in \mathbb{R}^{N_{\max} \times k}$ and go to <b>Step 2</b> until $k = N_{\max} - 1$ .
<b>Output</b>	the pixel indices $\{l_1, \dots, l_{N_{\max}}\}$

The augmented dimension-reduced data [given by (43)] used in the development of TRI-P algorithm is not only to account for (A2) but also to ensure that in total  $N_{\max}$  purest pixel indices can be identified via orthonormal projection of the  $N_{\max} - 1$  dimension-reduced data. By contrast, using the dimension-reduced data [given by (13)] in TRI-P algorithm will yield only  $N_{\max} - 1$  purest pixel indices due to the lack of degrees of freedom to identify the  $N_{\max}$ th purest pixel.

- While other existing pure-pixel-based EEAs such as pixel purity index (PPI) [24], vertex component analysis (VCA) [25], N-FINDR [27]–[29], and simplex growing algorithm (SGA) [30] require initializations (though the outcome of SGA is insensitive to initialization [30]), TRI-P does not require initialization of any form, and hence the solution is unique for a given hyperspectral data. Therefore, the endmember estimates of the TRI-P algorithms are reproducible even in the presence of noise, i.e., they always yield the same endmember estimates for a given hyperspectral data.
- Incidentally, the SVMAX algorithm [31] which is developed based on Winter’s unmixing criterion, and ATGP algorithm [34] which is based on target detection, turn out to be the special cases of the pixel search based TRI-P algorithm with  $p = 2$ . This also shows the interrelation between the EEAs based on Winter’s unmixing criterion and the pure pixel search strategy.

Though the notion of orthogonal projections has already been employed in VCA [25] to avoid repeated identification of a previously found endmember, the insensitivity to initialization, theoretical support for endmember identifiability, and superior performance (as will be demonstrated in Section V) of the proposed EEA, are some of the special characteristics of the proposed TRI-P algorithm.

The entire TRI-P algorithm is summarized in Table II. Though the TRI-P algorithm in Table II alone can yield a set of pixel indices  $\{l_1, \dots, l_{N_{\max}}\}$ , if used in conjunction with the GENE algorithms (presented in Table I), the pixel index  $l_1$  needed in Step 2 of Table I is provided by Step 1 of Table II, while  $l_k$  needed in Step 3 of Table I for any  $1 < k \leq N_{\max}$  is successively provided by Step 2 of Table II, for each  $k > 1$ . In other words, TRI-P algorithm with the above characteristics serves as a good candidate for successively providing the pixel indices to feed the proposed GENE algorithms in a synchronous fashion.

## V. SIMULATIONS

In this section, various Monte Carlo simulations are performed to analyze the performances of the proposed TRI-P, GENE-CH, and GENE-AH algorithms.<sup>1</sup> In the first subsection, the effectiveness of the proposed TRI-P (for  $p = 1, 2,$  and  $\infty$ ) algorithm is studied. Algorithms that are considered for comparison with TRI-P algorithm are VCA [25], iterative N-FINDR (I-N-FINDR) [27], successive N-FINDR (SC-N-FINDR), sequential N-FINDR (SQ-N-FINDR) [28], SGA [30], and alternating volume maximization (AVMAX) [31]. The algorithmic details for those EEAs under test are as follows: Affine set fitting [43] is employed for dimension reduction in I-N-FINDR, SC-N-FINDR, SQ-N-FINDR, and AVMAX, while VCA uses either singular value decomposition (SVD) or PCA based on the signal-to-noise ratio (SNR). To have fair complexity comparison with other methods, VCA is supplied with the SNR value (instead of letting the VCA algorithm estimate the SNR value). The convergence tolerance for I-N-FINDR, SQ-N-FINDR, and AVMAX is set to  $10^{-6}$ .

In the second subsection, simulations are performed to study the effectiveness of GENE-CH and GENE-AH algorithms in various scenarios. Algorithms that are considered for comparison are HySiMe [18], HFC, NWHFC [14], ATGP-NPD [36], and MINMAX-SVD [7]. The GENE algorithms, HFC, NWHFC, and ATGP-NPD are evaluated for the following false alarm probability:  $10^{-3}, 10^{-4}, 10^{-5},$  and  $10^{-6}$ , and for GENE, NWHFC, and ATGP-NPD algorithms, the true noise covariance matrix is supplied for each simulated data set.

In both subsections, for all the scenarios under consideration, 100 Monte Carlo runs are performed. The average root-mean-square (rms) spectral angle [25], [43] between the true and the estimated endmember signatures is used as the performance index for evaluating the performances of the EEAs under test. The root-mean-square (rms) spectral angle  $\phi$  is defined below for convenience

$$\phi = \min_{\pi \in \Pi_N} \sqrt{\frac{1}{N} \sum_{i=1}^N \left[ \arccos \left( \frac{\mathbf{a}_i^T \widehat{\mathbf{a}}_{\pi_i}}{\|\mathbf{a}_i\| \|\widehat{\mathbf{a}}_{\pi_i}\|} \right) \right]^2} \quad (53)$$

where  $\widehat{\mathbf{a}}_i$  denotes the estimated endmember signature,  $\boldsymbol{\pi} = [\pi_1, \dots, \pi_N]^T$ , and  $\Pi_N = \{\boldsymbol{\pi} \in \mathbb{R}^N | \pi_i \in \{1, 2, \dots, N\}, \pi_i \neq \pi_j \text{ for } i \neq j\}$  is the set of all the permutations of  $\{1, 2, \dots, N\}$ . Lower spectral angle  $\phi$  corresponds to better performance of the EEA. For performance comparison of the algorithms under test for estimating the number of endmembers, the mean and standard deviation of the estimated number of endmembers are calculated.

In the simulations, the endmembers are chosen from the USGS library [52]. The endmembers considered in our simulations are from the following pool: Alunite, Andradite, Buddingtonite, Chalcedony, Desert Varnish, Goethite,

Halloysite, Kaolinite, Montmorillonite, Muscovite, Nontronite, Pyrope, Ammonium Smectite, Calcite, Dickite, Dumortierite, Hematite, Jarosite, Opal, and Paragonite, with  $M = 224$ . The abundance vectors  $\mathbf{s}[n]$ ,  $n = 1, \dots, L$  are generated by following the Dirichlet distribution [25], [43], which ensures that the assumptions (A1) and (A2) hold true for the simulated hyperspectral data. In addition to the number of endmembers  $N$ , and the number of pixels  $L$ , there are two other parameters that define a particular scenario: purity level  $\rho$  of the data set and SNR. A data set with purity level  $\rho$  implies the purity factor  $1/\sqrt{N} \leq \rho_n = \|\mathbf{s}[n]\|_2 \leq \rho$  [43] for each of its abundance vectors  $\mathbf{s}[n]$ . The value of  $\rho_n$  defines the quantitative dominance of an endmember  $\mathbf{a}_i$  in the observed pixel vector  $\mathbf{x}[n] = \sum_{i=1}^N s_i[n] \mathbf{a}_i, \forall n$ . The SNR of a data set with noise variance  $\sigma^2$  is defined as

$$\text{SNR} = \frac{1}{ML\sigma^2} \sum_{n=1}^L \|\mathbf{x}[n]\|_2^2. \quad (54)$$

### A. Evaluation of TRI-P Algorithm in Uniform Gaussian Noise Scenario

Here, the first eight endmembers are considered ( $N = 8$ ) from the aforementioned pool and additive white Gaussian noise is added to the noise-free data and the noisy observations are generated as per (1). The noise in each band is assumed to be independent and identically distributed Gaussian with zero mean and variance  $\sigma^2$ . The EEAs are tested for different purity levels and for different SNRs. The average rms spectral angles  $\phi$  for the EEAs under test for SNRs ranging from 10 to 40 dB, and for the noise-free case (SNR =  $\infty$ ), with  $\rho = 0.8, 0.9, 1,$  and  $L = 1000$  pixels are shown in Table III. Though the EEAs are designed for  $\rho = 1$ , we consider cases with different purity levels so as to study the performances of the algorithms when the pure pixel assumption is violated. The bold-faced numbers in Table III correspond to the minimum average rms spectral angle for a specific pair of ( $\rho, \text{SNR}$ ), over all the algorithms under test. It can be observed from Table III that for  $\rho = 1$ , and 0.9, TRI-P ( $p = 2$ ) wins in almost all situations, for  $\rho = 0.8$  (highly mixed case), TRI-P ( $p = 1$ ) performs well in many cases. The average computation time  $T_c$  (over all the scenarios under consideration) of each algorithm implemented in Matlab R2008a and running in a desktop computer equipped with Core i7-930 CPU with speed 2.80 GHz, and 12 GB memory, is also shown Table III. It can be observed from Table III that, the TRI-P algorithm, besides better performance, also offers the highest computational efficiency.

### B. Evaluation of GENE-CH and GENE-AH Algorithms

The GENE-CH, GENE-AH, and GENE-AH-MOD (in Scenario 6 below) algorithms introduced in Section III are tested on the simulated hyperspectral data. Since, out of all the EEAs considered in Section V-A, TRI-P (with  $p = 2$ ) offered the best performance, the pixel indices required for GENE algorithms (see Table I) are obtained from the TRI-P (with  $p = 2$ ) algorithm. There are totally six scenarios under consideration.

<sup>1</sup>We have provided practical implementations of TRI-P and GENE algorithms at [http://www.ee.nthu.edu.tw/cychi/source\\_code\\_download-e.html](http://www.ee.nthu.edu.tw/cychi/source_code_download-e.html). The source codes were written in MATLAB, and are based on a reliable convex optimization software CVX[47]. Readers are encouraged to use the codes for their research purpose.

TABLE III  
AVERAGE  $\phi$  (DEGREES) AND AVERAGE COMPUTATION TIME  $T_c$  (SECS) OVER THE VARIOUS EEAS FOR DIFFERENT PURITY LEVELS ( $\rho$ ) AND SNRS (UNIFORM GAUSSIAN NOISE CASE),  $L = 1000$ ,  $N = 8$

Methods	$\rho$	$\phi$ (degrees)								$T_c$ (secs)
		SNR (dB)								
		10	15	20	25	30	35	40	$\infty$	
VCA	0.8	9.34	6.85	5.28	4.98	4.55	4.26	4.39	4.03	0.0413
	0.9	8.88	5.43	3.65	3.14	2.66	2.60	2.47	2.03	
	1	8.65	4.28	2.15	1.57	0.84	0.53	0.32	0.02	
I-N-FINDR	0.8	9.24	6.40	4.93	4.70	4.54	4.54	4.60	4.50	0.1212
	0.9	8.69	4.88	3.32	2.81	2.66	2.68	2.62	2.49	
	1	8.39	3.87	1.91	1.06	0.60	0.35	0.23	0.12	
SC-N-FINDR	0.8	9.57	7.14	5.94	5.44	5.27	5.04	4.94	4.80	0.0552
	0.9	9.11	5.84	3.76	3.14	2.98	2.86	2.79	2.79	
	1	8.55	4.32	2.07	1.23	0.71	0.38	0.31	0.12	
SQ-N-FINDR	0.8	9.25	6.37	4.88	4.70	4.50	4.51	4.63	4.51	0.1361
	0.9	8.98	4.94	3.32	2.81	2.66	2.68	2.62	2.49	
	1	8.23	3.97	1.91	1.06	0.60	0.35	0.23	0.12	
SGA	0.8	8.71	6.31	5.03	4.59	4.61	4.56	4.43	4.25	0.1886
	0.9	8.51	5.01	3.10	2.66	2.40	2.39	2.37	2.12	
	1	<b>7.96</b>	3.76	1.78	0.99	0.57	0.34	0.22	<b>0.01</b>	
AVMAX	0.8	9.34	6.52	5.13	4.77	4.60	4.62	4.55	4.17	0.0207
	0.9	8.91	5.26	3.34	2.82	2.67	2.68	2.66	2.36	
	1	8.56	4.00	1.90	1.06	0.60	0.35	0.23	<b>0.01</b>	
TRI-P ( $p = 1$ )	0.8	9.46	6.25	<b>4.59</b>	<b>4.22</b>	<b>4.05</b>	<b>3.94</b>	<b>4.01</b>	4.02	<b>0.0166</b>
	0.9	9.07	4.93	3.16	2.54	2.29	2.18	2.19	<b>1.97</b>	
	1	8.64	3.95	1.88	1.00	0.56	0.34	0.22	<b>0.01</b>	
TRI-P ( $p = 2$ )	0.8	9.02	<b>6.22</b>	4.63	4.35	4.30	4.25	4.19	<b>3.95</b>	0.0170
	0.9	8.59	<b>4.54</b>	<b>2.64</b>	<b>2.17</b>	<b>2.03</b>	<b>2.04</b>	<b>2.04</b>	<b>1.97</b>	
	1	8.10	<b>3.74</b>	<b>1.75</b>	<b>0.95</b>	<b>0.55</b>	<b>0.33</b>	<b>0.21</b>	<b>0.01</b>	
TRI-P ( $p = \infty$ )	0.8	8.48	6.28	4.98	4.72	4.63	4.62	4.57	4.30	0.0172
	0.9	<b>8.36</b>	4.87	3.11	2.64	2.41	2.39	2.37	2.13	
	1	<b>8.04</b>	3.92	1.80	0.99	0.56	<b>0.33</b>	0.22	<b>0.01</b>	

TABLE IV  
MEAN  $\pm$  STANDARD DEVIATION OF THE ESTIMATED NUMBER OF ENDMEMBERS FOR GENE ALGORITHMS OVER 100 INDEPENDENT RUNS, WITH DIFFERENT FALSE ALARM PROBABILITIES  $P_{FA}$ —UNIFORM GAUSSIAN NOISE CASE SNR = 20 dB AND 40 dB, FOR  $N_{max} = 10, 20, 30,$  AND  $50$ . TRUE  $N = 8, L = 5000, M = 224,$  AND  $\rho = 1$

Methods	$P_{FA}$	SNR=20 dB				SNR=40 dB			
		$N_{max}$				$N_{max}$			
		10	20	30	50	10	20	30	50
GENE-CH (TRI-P, $p = 2$ )	$10^{-3}$	8.02 $\pm$ 0.17	8.13 $\pm$ 0.36	8.16 $\pm$ 0.39	8.19 $\pm$ 0.37	8.00 $\pm$ 0	8.06 $\pm$ 0.23	8.17 $\pm$ 0.45	8.18 $\pm$ 0.45
	$10^{-4}$	8.01 $\pm$ 0.09	8.02 $\pm$ 0.14	8.02 $\pm$ 0.17	7.98 $\pm$ 0.31	8.00 $\pm$ 0	8.02 $\pm$ 0.17	8.02 $\pm$ 0.17	8.06 $\pm$ 0.27
	$10^{-5}$	8.01 $\pm$ 0.09	8.01 $\pm$ 0.09	8.01 $\pm$ 0.09	7.93 $\pm$ 0.25	8.00 $\pm$ 0	8.01 $\pm$ 0.09	8.02 $\pm$ 0.14	8.00 $\pm$ 0
	$10^{-6}$	8.01 $\pm$ 0.09	8.00 $\pm$ 0	7.98 $\pm$ 0.14	7.88 $\pm$ 0.32	8.00 $\pm$ 0	8.00 $\pm$ 0	8.02 $\pm$ 0.14	8.00 $\pm$ 0
GENE-AH (TRI-P, $p = 2$ )	$10^{-3}$	8.00 $\pm$ 0	8.02 $\pm$ 0.14	8.07 $\pm$ 0.25	8.03 $\pm$ 0.37	8.00 $\pm$ 0	8.00 $\pm$ 0	8.03 $\pm$ 0.19	8.06 $\pm$ 0.23
	$10^{-4}$	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	7.94 $\pm$ 0.23	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
	$10^{-5}$	8.00 $\pm$ 0	8.00 $\pm$ 0	7.99 $\pm$ 0.09	7.93 $\pm$ 0.25	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
	$10^{-6}$	7.99 $\pm$ 0.09	7.98 $\pm$ 0.14	7.92 $\pm$ 0.30	7.85 $\pm$ 0.35	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0

*Scenario 1:* In the first scenario, the effect of the chosen  $N_{max}$  value on the estimation accuracy of the GENE algorithms is investigated. The endmembers used are the same as used in the previous subsection and the data generation parameters are  $N = 8, M = 224, L = 5000,$  and  $\rho = 1$ . The  $N_{max}$  is varied as 10, 20, 30, and 50. Uniform Gaussian noise is added to the data and the SNR values considered are 20 dB (low SNR) and 40 dB (high SNR). The obtained mean  $\pm$  standard deviation of the estimated number of endmembers over 100 independent runs are shown in Table IV. It can be seen from Table IV that for both SNRs the estimation accuracy of the GENE algorithms is considerably robust to the  $N_{max}$  values, and the closer the  $N_{max}$  value is to the true  $N$ , the better will be the estimation accuracy.

*Scenario 2:* The endmembers used and the data generation parameters are the same as in Scenario 1, and  $N_{max} = 25$ . As

in Section V-A, uniform Gaussian noise was added to produce noisy hyperspectral data for SNR values of 15, 25, 35 and 45 dB. The mean and the standard deviation of the estimated number of endmembers over 100 independent runs, for the algorithms under test are shown in Table V (left half). From Table V, it can be readily observed that for this scenario, the estimation accuracies of the proposed algorithms (GENE-CH in particular) are the best for low SNR (15 dB), and for other SNRs, GENE-CH and GENE-AH with  $P_{FA} = 10^{-6}$  perform well and so does HySiMe. It should be noted that for SNR  $\geq 25$  dB, as the  $P_{FA}$  decreases, the standard deviations almost reduce to zero. The performances of the other algorithms under test are below par, and that of HFC (for SNR  $\geq 25$  dB) and NWHFC (for all SNRs) are independent of SNR.

*Scenario 3:* In this scenario, we study the performances of the algorithms under test for the case when the data are

TABLE V  
MEAN  $\pm$  STANDARD DEVIATION OF THE ESTIMATED NUMBER OF ENDMEMBERS FOR VARIOUS ALGORITHMS OVER 100 INDEPENDENT RUNS, WITH DIFFERENT FALSE ALARM PROBABILITIES  $P_{FA}$  (WHENEVER APPLICABLE) AND SNRS—UNIFORM AND NON-UNIFORM GAUSSIAN NOISE CASE.  $N_{\max} = 25$ , TRUE  $N = 8$ ,  $L = 5000$ ,  $M = 224$ , AND  $\rho = 1$

Methods	$P_{FA}$	Uniform Gaussian noise				Non-uniform Gaussian noise ( $\tau = 36$ )			
		SNR (dB)				SNR (dB)			
		15	25	35	45	15	25	35	45
GENE-CH (TRI-P, $p = 2$ )	$10^{-3}$	8.26 $\pm$ 1.07	8.27 $\pm$ 0.56	8.16 $\pm$ 0.36	8.16 $\pm$ 0.41	9.07 $\pm$ 1.11	8.44 $\pm$ 0.68	8.51 $\pm$ 0.84	8.26 $\pm$ 0.52
	$10^{-4}$	7.56 $\pm$ 0.67	8.01 $\pm$ 0.09	8.02 $\pm$ 0.17	8.05 $\pm$ 0.21	8.51 $\pm$ 0.74	8.19 $\pm$ 0.42	8.15 $\pm$ 0.41	8.10 $\pm$ 0.34
	$10^{-5}$	7.21 $\pm$ 0.51	8.00 $\pm$ 0	8.00 $\pm$ 0	8.01 $\pm$ 0.09	8.27 $\pm$ 0.60	8.07 $\pm$ 0.25	8.03 $\pm$ 0.19	8.02 $\pm$ 0.20
	$10^{-6}$	6.99 $\pm$ 0.41	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.13 $\pm$ 0.44	8.05 $\pm$ 0.21	8.02 $\pm$ 0.17	8.00 $\pm$ 0
GENE-AH (TRI-P, $p = 2$ )	$10^{-3}$	7.78 $\pm$ 0.75	8.07 $\pm$ 0.29	8.06 $\pm$ 0.23	8.03 $\pm$ 0.19	8.26 $\pm$ 0.48	8.10 $\pm$ 0.30	8.08 $\pm$ 0.27	8.09 $\pm$ 0.28
	$10^{-4}$	7.27 $\pm$ 0.67	8.00 $\pm$ 0	8.02 $\pm$ 0.14	8.00 $\pm$ 0	8.11 $\pm$ 0.35	8.03 $\pm$ 0.19	8.00 $\pm$ 0	8.01 $\pm$ 0.09
	$10^{-5}$	6.90 $\pm$ 0.65	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	7.95 $\pm$ 0.50	8.01 $\pm$ 0.09	8.00 $\pm$ 0	8.00 $\pm$ 0
	$10^{-6}$	6.52 $\pm$ 0.70	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	7.91 $\pm$ 0.47	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
HYSIME	–	6.10 $\pm$ 0.30	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	6.09 $\pm$ 0.28	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
HFC	$10^{-3}$	4.05 $\pm$ 0.21	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	3.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0
	$10^{-4}$	4.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	3.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0
	$10^{-5}$	3.99 $\pm$ 0.09	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	3.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0
	$10^{-6}$	3.99 $\pm$ 0.09	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	2.99 $\pm$ 0.09	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0
NW-HFC	$10^{-3}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.04 $\pm$ 0.19	6.14 $\pm$ 0.34	6.13 $\pm$ 0.37	6.13 $\pm$ 0.33
	$10^{-4}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.03 $\pm$ 0.17	6.04 $\pm$ 0.19	6.04 $\pm$ 0.19	6.04 $\pm$ 0.19
	$10^{-5}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.01 $\pm$ 0.10	6.03 $\pm$ 0.17	6.03 $\pm$ 0.17	6.03 $\pm$ 0.17
	$10^{-6}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.99 $\pm$ 0.09	6.03 $\pm$ 0.17	6.03 $\pm$ 0.17	6.03 $\pm$ 0.17
ATGP-NPD	$10^{-3}$	18.14 $\pm$ 1.52	30.86 $\pm$ 1.92	34.95 $\pm$ 1.70	35.89 $\pm$ 1.91	24.60 $\pm$ 1.82	33.30 $\pm$ 1.83	35.41 $\pm$ 1.65	35.75 $\pm$ 1.75
	$10^{-4}$	14.78 $\pm$ 1.18	25.91 $\pm$ 1.68	29.92 $\pm$ 1.53	30.68 $\pm$ 1.62	20.40 $\pm$ 1.39	28.29 $\pm$ 1.49	30.71 $\pm$ 1.69	30.93 $\pm$ 1.74
	$10^{-5}$	12.60 $\pm$ 0.97	22.41 $\pm$ 1.30	26.29 $\pm$ 1.27	27.02 $\pm$ 1.50	17.52 $\pm$ 1.15	24.90 $\pm$ 1.25	26.77 $\pm$ 1.28	26.97 $\pm$ 1.43
	$10^{-6}$	10.94 $\pm$ 0.87	19.91 $\pm$ 1.21	24.25 $\pm$ 1.19	24.93 $\pm$ 0.89	15.41 $\pm$ 1.08	22.54 $\pm$ 1.22	24.59 $\pm$ 1.15	24.81 $\pm$ 0.98
MINMAX-SVD	–	3.31 $\pm$ 1.81	2.66 $\pm$ 1.48	2.85 $\pm$ 1.31	3.47 $\pm$ 1.12	3.65 $\pm$ 1.68	3.10 $\pm$ 1.70	3.15 $\pm$ 1.62	3.30 $\pm$ 1.11

corrupted by non-uniform Gaussian noise, while maintaining the other parameters used in the previous scenario. The noise in each band is considered to be uncorrelated, but with different variances in each band. The noise variances  $\sigma_i^2$  in each of the  $M$  spectral bands follow a Gaussian shape that is centered at the  $(M/2)$ th band. The noise variances  $\sigma_i^2$  in each band are given by [25], [43]

$$\sigma_i^2 = \sigma^2 \frac{\exp(-(i-M/2)^2/2\tau^2)}{\sum_{j=1}^M \exp(-(j-M/2)^2/2\tau^2)}, \quad \forall i=1, \dots, M \quad (55)$$

where  $\sigma^2$  is defined in (54),  $\exp(\cdot)$  represents exponential function, and  $\tau$  controls the variance of the Gaussian shape among  $\sigma_1^2, \dots, \sigma_M^2$ . It corresponds to uniform Gaussian noise for  $\tau = \infty$ , and one-band noise for  $\tau = 0$ . The  $\tau$  value in the simulations is set to 36 and the purity level  $\rho$  is fixed to be 1. Values of mean  $\pm$  standard deviation of the number of endmembers estimated by the algorithms under test are also shown in Table V (right half). Here, again, for low SNR (15 dB), the proposed GENE algorithms perform well. For other SNRs, GENE-AH with  $P_{FA} = 10^{-6}$  and HySiMe yield the best performance. It is worthwhile to mention that contrary to the GENE algorithms, the performance of HySiMe algorithm is almost independent of the noise types (uniform or non-uniform Gaussian noise).

*Scenario 4:* The purity level  $\rho$  of the generated hyperspectral data is allowed to vary while maintaining  $N = 8$ . The data are corrupted by uniform Gaussian noise with SNR = 30 dB. For the case with  $N = 8$ ,  $M = 224$ ,  $L = 5000$ , and  $N_{\max} = 25$ , values of mean  $\pm$  standard deviation of the number of endmembers estimated by the algorithms under test are tabulated in Table VI (left half). It can be readily seen from Table VI

that when purity level is smaller, GENE-CH overestimates the number of endmembers which is consistent with the discussions in Section III-B and the illustration in Fig. 2. On the other hand, GENE-AH with  $P_{FA} = 10^{-6}$  and HySiMe correctly estimates the number of endmembers.

*Scenario 5:* In the data generation, the number of endmember  $N$  is allowed to vary as 8, 12, 16 and 20, while maintaining  $\rho = 1$ ,  $M = 224$ , and  $L = 5000$ . Here, again, the data is corrupted by uniform Gaussian noise with SNR = 30 dB. Values of mean  $\pm$  standard deviation of the number of endmembers estimated by the algorithms under test, are also tabulated in Table VI (right half). It can be observed from Table VI that for higher number of endmember  $N = 16, 20$  GENE-CH yields the best performance followed by GENE-AH. For  $N = 8, 12$  both GENE-AH, GENE-CH, and HySiMe yield the best performance.

*Scenario 6:* The final scenario is to study the performance of the GENE-AH-MOD and other algorithms on data with (A2) and (A4) violated. The synthetic data are generated by Dirichlet distribution except that the sum of the abundances are not constrained to unity. The parameters are set to  $N = 8$ ,  $M = 224$ ,  $L = 5000$ ,  $\rho = 0.8$  and  $N_{\max} = 25$ . Uniform Gaussian noise was added and the SNR values considered are 15, 25, 35 and 45 dB. The mean  $\pm$  standard deviation of the number of endmembers estimated by the algorithms under test, over 100 independent runs, are shown in Table VII. From Table VII, it can be readily seen that for lower SNR (15 dB) GENE-AH-MOD and GENE-AH offer the best estimate of  $N$ , and for other SNRs both GENE-AH-MOD and HySiMe yield the best estimates of  $N$ . Whenever there is uncertainty in a data regarding the validity of assumptions (A2) and (A4), GENE-AH-MOD could be used to estimate the number of endmembers, irrespective of the SNR.

TABLE VI  
 MEAN  $\pm$  STANDARD DEVIATION OF THE ESTIMATED NUMBER OF ENDMEMBERS FOR VARIOUS ALGORITHMS OVER 100 INDEPENDENT RUNS,  
 WITH DIFFERENT FALSE ALARM PROBABILITIES  $P_{FA}$  (WHENEVER APPLICABLE), FOR VARIOUS PURITY LEVELS AND NUMBER OF  
 ENDMEMBERS. UNIFORM GAUSSIAN NOISE CASE, SNR = 30 dB,  $N_{max} = 25$ ,  $L = 5000$ , AND  $M = 224$

Methods	$P_{FA}$	$N = 8$ SNR=30 dB				$\rho = 1$ SNR=30 dB			
		Purity Level $\rho$				Number of Endmembers $N$			
		0.8	0.85	0.9	0.95	8	12	16	20
GENE-CH (TRI-P, $p = 2$ )	$10^{-3}$	14.06 $\pm$ 3.68	12.79 $\pm$ 2.80	9.98 $\pm$ 1.44	8.80 $\pm$ 0.93	8.19 $\pm$ 0.40	12.28 $\pm$ 0.62	15.86 $\pm$ 0.51	19.82 $\pm$ 0.55
	$10^{-4}$	12.82 $\pm$ 2.94	10.97 $\pm$ 2.38	9.28 $\pm$ 1.28	8.33 $\pm$ 0.58	8.02 $\pm$ 0.17	12.04 $\pm$ 0.19	15.77 $\pm$ 0.46	19.78 $\pm$ 0.50
	$10^{-5}$	12.13 $\pm$ 2.70	10.23 $\pm$ 1.92	8.85 $\pm$ 0.98	8.17 $\pm$ 0.45	8.00 $\pm$ 0	12.03 $\pm$ 0.17	15.77 $\pm$ 0.46	19.74 $\pm$ 0.52
	$10^{-6}$	11.65 $\pm$ 2.69	9.85 $\pm$ 1.71	8.61 $\pm$ 0.92	8.11 $\pm$ 0.38	8.00 $\pm$ 0	12.02 $\pm$ 0.14	15.72 $\pm$ 0.47	19.70 $\pm$ 0.52
GENE-AH (TRI-P, $p = 2$ )	$10^{-3}$	8.10 $\pm$ 0.31	8.05 $\pm$ 0.21	8.14 $\pm$ 0.34	8.09 $\pm$ 0.28	8.06 $\pm$ 0.23	12.02 $\pm$ 0.14	14.99 $\pm$ 0.38	18.01 $\pm$ 0.46
	$10^{-4}$	8.02 $\pm$ 0.14	8.01 $\pm$ 0.09	8.00 $\pm$ 0	8.01 $\pm$ 0.09	8.00 $\pm$ 0	12.00 $\pm$ 0	14.76 $\pm$ 0.42	17.75 $\pm$ 0.50
	$10^{-5}$	8.00 $\pm$ 0	8.01 $\pm$ 0.09	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	12.00 $\pm$ 0	14.57 $\pm$ 0.49	17.51 $\pm$ 0.54
	$10^{-6}$	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	12.00 $\pm$ 0	14.32 $\pm$ 0.46	17.17 $\pm$ 0.66
HYSIME	–	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	12.00 $\pm$ 0	14.00 $\pm$ 0	16.15 $\pm$ 0.35
HFC	$10^{-3}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	7.81 $\pm$ 0.44	8.35 $\pm$ 0.49	5.01 $\pm$ 0.61
	$10^{-4}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	7.14 $\pm$ 0.68	8.66 $\pm$ 0.27	4.19 $\pm$ 0.63
	$10^{-5}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.44 $\pm$ 0.53	7.93 $\pm$ 0.25	3.67 $\pm$ 0.60
	$10^{-6}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.10 $\pm$ 0.46	7.76 $\pm$ 0.47	3.23 $\pm$ 0.52
NW-HFC	$10^{-3}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	7.79 $\pm$ 0.53	9.39 $\pm$ 0.54	7.01 $\pm$ 0.74
	$10^{-4}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	7.18 $\pm$ 0.70	9.15 $\pm$ 0.35	6.23 $\pm$ 0.69
	$10^{-5}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	6.46 $\pm$ 0.62	8.97 $\pm$ 0.30	5.46 $\pm$ 0.77
	$10^{-6}$	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.00 $\pm$ 0	5.96 $\pm$ 0.58	8.80 $\pm$ 0.42	4.78 $\pm$ 0.70
ATGP-NPD	$10^{-3}$	35.34 $\pm$ 1.97	35.32 $\pm$ 2.06	34.73 $\pm$ 1.80	34.10 $\pm$ 1.85	33.88 $\pm$ 1.73	44.22 $\pm$ 2.09	49.30 $\pm$ 2.40	55.59 $\pm$ 2.65
	$10^{-4}$	29.75 $\pm$ 1.63	29.50 $\pm$ 1.62	29.43 $\pm$ 1.57	29.07 $\pm$ 1.45	28.68 $\pm$ 1.56	37.28 $\pm$ 1.55	41.69 $\pm$ 2.15	47.16 $\pm$ 1.80
	$10^{-5}$	25.84 $\pm$ 1.45	25.51 $\pm$ 1.39	25.55 $\pm$ 1.27	25.38 $\pm$ 1.22	25.25 $\pm$ 1.25	32.83 $\pm$ 1.60	36.33 $\pm$ 2.05	40.92 $\pm$ 1.89
	$10^{-6}$	22.98 $\pm$ 1.42	22.96 $\pm$ 1.15	23.17 $\pm$ 1.23	22.90 $\pm$ 1.20	22.93 $\pm$ 1.14	28.64 $\pm$ 1.61	31.90 $\pm$ 1.74	36.01 $\pm$ 1.76
MINMAX-SVD	–	3.20 $\pm$ 1.88	2.90 $\pm$ 1.81	2.91 $\pm$ 1.67	3.15 $\pm$ 1.81	2.73 $\pm$ 1.64	3.57 $\pm$ 2.01	4.33 $\pm$ 2.01	3.87 $\pm$ 1.98

TABLE VII  
 MEAN  $\pm$  STANDARD DEVIATION OF THE ESTIMATED NUMBER OF ENDMEMBERS FOR VARIOUS ALGORITHMS OVER 100 INDEPENDENT RUNS,  
 WITH DIFFERENT FALSE ALARM PROBABILITIES  $P_{FA}$  (WHENEVER APPLICABLE) AND SNRS—UNIFORM GAUSSIAN NOISE CASE,  
 FOR DATA WITH (A2) AND (A4) VIOLATED.  $N_{max} = 25$ , TRUE  $N = 8$ ,  $L = 5000$ ,  $M = 224$ , AND  $\rho = 0.8$

Methods	$P_{FA}$	Uniform Gaussian noise			
		SNR (dB)			
		15	25	35	45
GENE-CH (TRI-P, $p = 2$ )	$10^{-3}$	9.69 $\pm$ 1.12	10.33 $\pm$ 1.72	10.95 $\pm$ 2.31	12.56 $\pm$ 4.17
	$10^{-4}$	9.19 $\pm$ 0.88	9.80 $\pm$ 1.19	10.23 $\pm$ 1.59	11.53 $\pm$ 3.73
	$10^{-5}$	8.98 $\pm$ 0.69	9.74 $\pm$ 1.06	10.00 $\pm$ 1.31	11.47 $\pm$ 3.09
	$10^{-6}$	8.85 $\pm$ 0.63	9.64 $\pm$ 0.95	9.85 $\pm$ 1.19	11.40 $\pm$ 3.02
GENE-AH (TRI-P, $p = 2$ )	$10^{-3}$	8.73 $\pm$ 0.56	9.02 $\pm$ 0.17	9.05 $\pm$ 0.21	9.02 $\pm$ 0.17
	$10^{-4}$	8.51 $\pm$ 0.54	9.00 $\pm$ 0	9.00 $\pm$ 0	9.00 $\pm$ 0
	$10^{-5}$	8.35 $\pm$ 0.50	9.00 $\pm$ 0	9.00 $\pm$ 0	9.00 $\pm$ 0
	$10^{-6}$	8.27 $\pm$ 0.49	9.00 $\pm$ 0	9.00 $\pm$ 0	9.00 $\pm$ 0
GENE-AH-MOD (TRI-P, $p = 2$ )	$10^{-3}$	7.73 $\pm$ 0.56	8.02 $\pm$ 0.17	8.05 $\pm$ 0.21	8.02 $\pm$ 0.17
	$10^{-4}$	7.51 $\pm$ 0.54	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
	$10^{-5}$	7.36 $\pm$ 0.50	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
	$10^{-6}$	7.28 $\pm$ 0.49	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
HYSIME	–	6.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0	8.00 $\pm$ 0
HFC	$10^{-3}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-4}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-5}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-6}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
NW-HFC	$10^{-3}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-4}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-5}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
	$10^{-6}$	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0	1.00 $\pm$ 0
ATGP-NPD	$10^{-3}$	14.77 $\pm$ 1.42	19.04 $\pm$ 1.38	19.86 $\pm$ 1.49	19.90 $\pm$ 1.54
	$10^{-4}$	13.40 $\pm$ 1.20	17.54 $\pm$ 1.44	18.25 $\pm$ 1.47	18.31 $\pm$ 1.42
	$10^{-5}$	12.31 $\pm$ 1.07	16.23 $\pm$ 1.39	16.99 $\pm$ 1.42	17.03 $\pm$ 1.36
	$10^{-6}$	11.50 $\pm$ 1.08	15.31 $\pm$ 1.19	16.10 $\pm$ 1.37	16.18 $\pm$ 1.27
MINMAX-SVD	–	2.01 $\pm$ 0.97	1.71 $\pm$ 1.52	1.60 $\pm$ 1.14	1.55 $\pm$ 1.35

In summary, it can be inferred that as long as  $N_{max} > N$ , the estimation of the number of endmembers do not suffer much. It can also be concluded from the above simulation results that the GENE-CH algorithm is more suitable for data with pure pixels

(i.e., for data with (A4) satisfied) and larger number of endmembers. While the GENE-AH is the better choice when (A4) is violated, GENE-AH-MOD serves as the preferred choice (irrespective of the SNR) when both (A2) and (A4) are violated.

TABLE VIII  
NUMBER OF ENDMEMBERS ESTIMATED BY VARIOUS ALGORITHMS,  
WHERE NA DENOTES “NON-APPLICABLE” AND \* DENOTES  
“OUT OF MEMORY” ENCOUNTERED IN MATLAB

Algorithms	$P_{FA}$	Estimated $N$
GENE-CH	$10^{-8}$	76
GENE-AH	$10^{-8}$	27
GENE-AH-MOD	$10^{-8}$	26
HySiMe	NA	21
HFC	$10^{-8}$	11
NWHFC	$10^{-8}$	11
ATGP-NPD-Gaussian Noise	$10^{-8}$	61
ATGP-NPD-Laplacian Noise	$10^{-8}$	36
ATGP-NPD-Modified Gaussian Noise	$10^{-8}$	44
MINMAX-SVD	NA	*

## VI. REAL DATA EXPERIMENTS

In this section, the proposed GENE-CH and GENE-AH algorithms using TRI-P ( $p = 2$ ), and some other algorithms for estimating the number of endmembers, are tested with AVIRIS real hyperspectral data obtained over the Cuprite Nevada site [53]. The AVIRIS data are well studied in the recent years [25], [43] and the availability of a structured library of endmember signature [52], [54] aids in identification of the mineral maps. The static nature of the Cuprite Nevada site over the recent years, together with the availability of a standard library of minerals makes the data appropriate for conducting real data experiments, so as to validate the algorithms under test. The algorithms under test are GENE-CH, GENE-AH, GENE-AH-MOD, HySiMe, HFC, NWHFC, ATGP-NPD, and MINMAX-SVD. The hyperspectral data over the Cuprite Nevada contains 224 bands with better SNRs in most of the bands [55]. Among the 224 bands, bands 1–2, 104–113, 148–167, and 221–224 were removed due to strong noise or dense water-vapor content in those bands. In our experiment, we considered a  $200 \times 200$  sub-image (region of interest) of the hyperspectral data, with 188 bands (after removing the bands with poor information).

The estimated number of endmembers obtained by the algorithms under test are given in Table VIII. For the GENE algorithms  $N_{max}$  is set to 100 and for all the algorithms  $P_{FA}$  is set to  $10^{-8}$ , wherever applicable, in order to get a reliable estimate. The noise covariance matrix for this real data is estimated by multiple regression analysis [18] and is supplied for GENE, NWHFC, and ATGP-NPD algorithms. As can be seen from Table VIII, the estimated numbers of endmembers are different for each algorithm. For the Cuprite data set, initially it was concluded that there are about 13 minerals (endmembers) in the site, and later it was increased to nearly 70 mineral compounds (endmembers) [55]. Hence, it is difficult to comment on the estimation accuracies of the algorithms under test. Assuming (A2) to be true for this data set, the good estimation accuracy of GENE-AH algorithm (as inferred from Tables V and VI), makes 27 (see Table VIII), a reasonable estimate for this data set. On the other hand, GENE-CH overestimates the number of endmembers. This may be attributed to the fact that in the absence of pure pixels, GENE-CH indeed overestimates the number of endmembers (cf. Table VI).

To show the applicability of TRI-P ( $p = 2$ ) algorithm in real data, a quantitative measurement of the endmembers estimated

TABLE IX  
MEAN-REMOVED SPECTRAL ANGLES  $\bar{\phi}$  (DEGREES) BETWEEN LIBRARY  
SPECTRA AND ENDMEMBERS ESTIMATED BY TRI-P ( $p = 2$ )

Minerals	$\bar{\phi}$ (degrees)
Alunite	17.91
Andradite	18.03 (19.16)
Buddingtonite	26.33
Chalcedony	19.85 (30.29)
Desert Varnish	11.92
Dumortierite	26.13 (31.43)
Kaolinite	24.34 (32.85)
Montmorillonite#1	18.79
Montmorillonite#2	20.34 (20.61)
Muscovite	34.01 (37.71)
Nontronite#1	23.89 (24.07) (25.38) (25.12)
Nontronite#2	15.35 (23.07) (26.02) (27.01)
Nontronite#3	22.76
Paragonite	36.73
Pyrope	12.20
Average $\bar{\phi}$	24.12

by TRI-P ( $p = 2$ ) algorithm for the considered real data with  $N = 27$ , namely the mean removed spectral angle between the estimated signature  $\mathbf{a}_{est}$  and the corresponding library signature  $\mathbf{a}_{lib}$  is considered. The mean removed spectral angle is defined in [43] and [56] as

$$\bar{\phi} = \arccos \left( \frac{(\mathbf{a}_{est} - \mathbf{m}(\mathbf{a}_{est}))^T (\mathbf{a}_{lib} - \mathbf{m}(\mathbf{a}_{lib}))}{\|\mathbf{a}_{est} - \mathbf{m}(\mathbf{a}_{est})\| \cdot \|\mathbf{a}_{lib} - \mathbf{m}(\mathbf{a}_{lib})\|} \right) \quad (56)$$

where  $\mathbf{m}(\mathbf{a}) = (\mathbf{1}_M^T \mathbf{a}) \mathbf{1}_M / M$  for any vector  $\mathbf{a} \in \mathbb{R}^M$ . The value of  $\bar{\phi}$  for the various minerals identified by the TRI-P ( $p = 2$ ), is given in Table IX, and the numbers in the parentheses correspond to the values of  $\bar{\phi}$  for repeatedly identified materials. The abundance maps corresponding to the endmembers estimated by TRI-P ( $p = 2$ ) are obtained by fully constrained least squares (FCLS) algorithm [57] and are shown in Fig. 3. Note that the minerals in Table IX and Fig. 3 are arranged in alphabetical order and the minerals were identified by the visual comparison of the obtained abundance maps with the ones available in [43], [54], [56], [58], and [59].

## VII. CONCLUSION

In this paper, we have considered the estimation of number of endmembers in hyperspectral images, which has been a challenging problem prevailing in the field of hyperspectral image analysis. To this end, we have presented two convex geometry-based algorithms, namely GENE-CH and GENE-AH algorithms, based on the fact that the observed dimension-reduced observations lie in the convex hull and affine hull of the endmember signatures, respectively. For data in which (A2) does not hold true, the GENE-AH algorithm is modified as GENE-AH-MOD, which relies only on (A3). The GENE algorithms employ a Neyman–Pearson hypothesis testing strategy to estimate the true number of endmembers. Any successive EEA can be employed in conjunction with the GENE algorithms, therefore the performances of the GENE algorithms depend on the successive EEA employed, implying that a reliable, reproducible and computationally efficient EEA is desirable. Furthermore, we have presented the successive TRI-P algorithm (with its endmember identifiability proven) which serves as a good EEA for the proposed GENE algorithms.

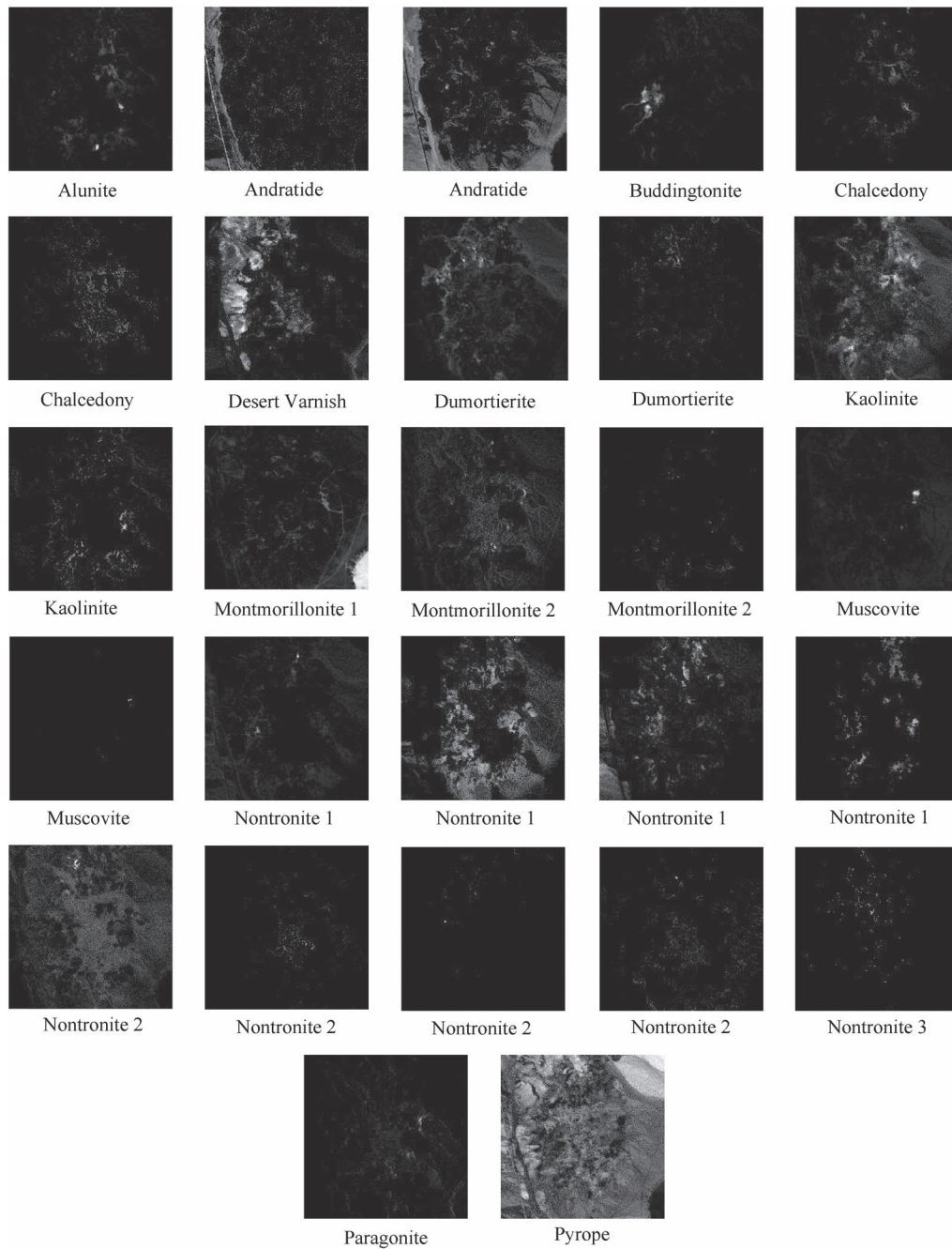


Fig. 3. Abundance maps estimated based on endmembers obtained by TRI-P (with  $p = 2$ ).

Simulation results confirm the superior efficacy of TRI-P ( $p = 2$ ) and TRI-P ( $p = 1$ ), for purer (e.g.  $\rho = 1, 0.9$ ) and more mixed (e.g.  $\rho = 0.8$ ) hyperspectral data, respectively. It is also shown via simulations that to estimate the number of endmembers, GENE-CH is preferred when pure pixels are present and the number of endmembers is large. For data without pure pixels, GENE-AH should be the advisable choice. If both (A2) and (A4) are violated in a given data, irrespective of the SNR, GENE-AH-MOD is the preferred choice for estimating the number of endmembers. Simulations also reveal that the HySiMe algorithm which is independent of assumptions (A2) and (A4) and does not require any tuning parameter, has performance comparable to that of the proposed GENE algorithms, in many of the scenarios. The real hyperspectral data experi-

ment also exhibits the practical applicability of the proposed GENE-AH algorithm and TRI-P algorithm.

Though real data are obtained with high SNRs, the presence of outliers in the data could affect the estimation of number of endmembers. Algorithms for estimation of number of endmembers, which are robust against the presence of outliers and against model mismatch errors (which may occur when multiple reflections are recorded by the hyperspectral sensor), will be our future research.

#### ACKNOWLEDGMENT

The authors would like to place on record their hearty gratitude to Prof. J. M. B. Dias for providing us with the

source codes of the HySiMe algorithm, and Prof. C.-I. Chang for providing us with the source codes of HFC, NWHFC, and ATGP-NPD algorithms.

## 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," *Elsevier J. 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, Feb. 2010.
- [4] J. M. P. Nascimento and J. M. B. Dias, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 863–878, Mar. 2012.
- [5] Y. Qian, S. Jia, J. Zhou, and A. R. Kelly, "Hyperspectral unmixing via L-1/2 sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.
- [6] P. Stoica, "Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 3069–3078, Dec. 1996.
- [7] C. Kuybeda, D. Malah, and M. Barzohar, "Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5579–5592, Dec. 2007.
- [8] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36–47, Jul. 2004.
- [9] K. P. Burnham, *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. New York: Springer-Verlag, 2002.
- [10] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. AC-19, no. 6, pp. 716–723, Dec. 1974.
- [11] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, Sep. 1978.
- [12] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.
- [13] M. W. Graham and D. J. Miller, "Unsupervised learning of parsimonious mixtures on large spaces with integrated feature and component selection," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1289–1303, Apr. 2006.
- [14] 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.
- [15] M. Wax and T. Kailath, "Detection of signals by information criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, pp. 387–392, Apr. 1985.
- [16] P. R. P. Neto, D. A. Jackson, and K. M. Somers, "How many principal components? stopping rules for determining the number of non-trivial axes revisited," *J. Comput. Stat. Data Anal.*, vol. 49, no. 4, pp. 974–997, Jun. 2005.
- [17] J. Harsanyi, W. Farrand, and C.-I. Chang, "Determining the number and identity of spectral endmembers: An integrated approach using Neyman–Pearson Eigenthresholding and iterative constrained RMS error minimization," in *Proc. 9th Thematic Conf. Geol. Remote Sens.*, Feb. 1993.
- [18] 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.
- [19] P. Bajorski, *Statistics for Imaging, Optics, and Photonics*. Hoboken, NJ: Wiley, 2011.
- [20] P. Bajorski, "Second moment linear dimensionality as an alternative to virtual dimensionality," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 672–678, Feb. 2011.
- [21] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 3, pp. 582–591, Jun. 2010.
- [22] 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.
- [23] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 446–450, Jul. 2007.
- [24] 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.
- [25] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [26] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE Conf. Imag. Spectrom.*, Pasadena, CA, Oct. 1999, pp. 266–275.
- [27] M. E. Winter, "A proof of the N-FINDR algorithm for the automated detection of endmembers in a hyperspectral image," in *Proc. SPIE Conf. Algorithms Technol. Multispectr., Hyperspectr., Ultraspectr. Imagery*, Aug. 2004, vol. 5425, pp. 31–41.
- [28] C.-C. Wu, S. Chu, and C.-I. Chang, "Sequential N-FINDR algorithms," in *Proc. SPIE*, Aug. 2008, vol. 7086, pp. 70860C-1–70860C-12.
- [29] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "Variants of N-FINDR algorithm for endmember extraction," in *Proc. SPIE*, Oct. 2008, vol. 7109, pp. 71090G-1–71090G-8.
- [30] 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.
- [31] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, Nov. 2011.
- [32] 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.
- [33] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.
- [34] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 39, no. 4, pp. 1232–1249, Oct. 2003.
- [35] C.-I. Chang, W. Xiong, W. Liu, M.-L. Chang, C.-C. Wu, and C.-C. Chen, "Linear spectral mixture analysis based approaches to estimation of virtual dimensionality in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 3960–3979, Nov. 2010.
- [36] C.-I. Chang, W. Xiong, H.-M. Chen, and J.-W. Chai, "Maximum orthogonal subspace projection approach to estimating the number of spectral signal sources in hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 504–520, Jun. 2011.
- [37] J. M. B. Dias and J. M. P. Nascimento, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image Signal Process. Remote Sens. XV*, vol. 7477, SPIE, 2009, no. 1, pp. 74770I-1–74770I-8.
- [38] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
- [39] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "Chance constrained robust minimum volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4194–4209, Nov. 2011.
- [40] 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.
- [41] M. T. Eismann and R. C. Hardie, "Application of the stochastic mixing model to hyperspectral resolution enhancement," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 9, pp. 1924–1933, Sep. 2004.
- [42] J. W. Boardman, "Analysis, understanding and visualization of hyperspectral data as convex sets in n-space," in *Proc. SPIE Conf. Imag. Spectrom.*, Orlando, FL, 1995, pp. 14–22.
- [43] 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.
- [44] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [45] L. C. Ludeman, *Random Processes Filtering, Estimation, and Detection*. New York: Wiley-Interscience, 2003.
- [46] 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.
- [47] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," CVX Research, Inc., Austin, TX, Oct. 2010, ver. 1.21.
- [48] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, 2nd ed. Hoboken, NJ: Wiley-Interscience, 1993.



[49] G. Arfken and H. Weber, *Mathematical Methods for Physicists*. New York: Harcourt/Academic Press, 2000.

[50] J. B. Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[51] S. M. Osnaga, "On rank one matrices and invariant subspaces," *Balkan J. Geom. Appl.*, vol. 10, no. 1, pp. 145–148, 2005.

[52] 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., Washington, DC, USGS Open File Rep., pp. 93–592, 1993.

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

[54] Cuprite, Nevada Research Papers, Tech. Rep. [Online]. Available: <http://speclab.cr.usgs.gov/cuprite.html>

[55] R. N. Clark 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.

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

[57] 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.

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

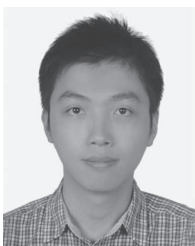
[59] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.



**ArulMurugan Ambikapathi** (S'02–M'11) received the B.E. degree in electronics and communication engineering from Bharathidasan University, Tiruchirappalli, India, in 2003, the M.E degree in communication systems from Anna University, Chennai, India, in 2005, and the Ph.D. degree from the Institute of Communications Engineering (ICE), National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2011. He secured top university ranks in both his B.E and M.E programs.

He is currently a Postdoctoral Research Fellow with the ICE, NTHU. His research interests are in convex analysis and optimization for blind source separation, hyperspectral unmixing, and recently, biomedical image analysis.

Dr. Ambikapathi was the recipient of Gold and Silver medals for academic excellence in his B.E and M.E programs, respectively. He is also the recipient of the NTHU Outstanding Student Scholarship award for two consecutive years (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.

He is currently a Research Scientist at Advanced Digital Sciences Center, Singapore. He was a visiting Doctoral Graduate Research Assistant with Virginia Polytechnic Institute and State University, Arlington, in 2008, and a Postdoctoral Research Fellow with

NTHU from 2009 to 2012. His research interests are in signal processing, image processing, and convex optimization, with a recent emphasis on hyperspectral remote sensing, dynamic medical imaging, and computer vision applications.



**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 has published more than 190 technical papers, including more than 65 journal papers (mostly in the IEEE TRANSACTIONS ON SIGNAL PROCESSING), two book chapters, and more than 120 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 co-sponsored workshops, symposiums, and conferences on signal processing and wireless communications, including Co-organizer and General Co-chairman of 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Co-Chair of Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008, and Lead Co-Chair of SPC Symposium, ChinaCOM 2009. He was an Associate Editor of IEEE TRANSACTIONS ON SIGNAL PROCESSING (5/2001–4/2006), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (1/2006–12/2007), IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (1/2008–12/2009), Associate Editor of the IEEE SIGNAL PROCESSING LETTERS (6/2006–5/2010), and a member of Editorial Board of EURASIP *Signal Processing Journal* (6/2005–5/2008), and an editor (7/2003–12/2005) as well as a Guest Editor (2006) of EURASIP *Journal on Applied Signal Processing*. He was a member of IEEE Signal Processing Committee on Signal Processing Theory and Methods (2005–2010). Currently, he is a member of the IEEE Signal Processing Committee on Signal Processing for Communications and Networking, and an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING.



**Kannan Keizer** received the B.E. degree in electronics and communications engineering from Madurai Kamaraj University, Madurai, India, in 2004. From 2004–2006, he was a Junior Research Fellow (JRF) at MEPCO Schlenk Engineering College, Sivakasi, India.

During his tenure as JRF, he has been working on industrial computerized tomography image denoising using wavelet transform. From 2006–2009, he worked as an Imaging Specialist at Network Systems and Technologies, Trivandrum, India. During this

period, he has been involved in research and product development in the area of biomedical image analysis. He has completed his Master degree in the Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan. His research interests are in the area of biomedical image analysis and hyperspectral unmixing.