

# A ROBUST ALTERNATING VOLUME MAXIMIZATION ALGORITHM FOR ENDMEMBER EXTRACTION IN HYPERSPECTRAL IMAGES

ArulMurugan Ambikapathi<sup>†</sup>, Tsung-Han Chan<sup>†</sup>, Wing-Kin Ma<sup>†\*</sup> and Chong-Yung Chi<sup>†</sup>

<sup>†</sup>Inst. Commun. Eng., National Tsing Hua Univ.  
Hsinchu, Taiwan

E-mail: (tsunghan@mx, cychi@ee).nthu.edu.tw

\*Dept. Electronic Eng., Chinese Univ. Hong Kong  
Shatin, N.T., Hong Kong

E-mail: wkma@ieee.org

## ABSTRACT

Accurate estimation of endmember signatures and the associated abundances of a scene from its hyperspectral observations is at present, a challenging research area. Many of the existing hyperspectral unmixing algorithms are based on Winter's belief, which states that the vertices of the maximum volume simplex inside the data cloud (observations) will yield high fidelity estimates of the endmember signatures if pure-pixels exist. Based on Winter's belief, we recently proposed a convex analysis based alternating volume maximization (AVMAX) algorithm. In this paper we develop a robust version of the AVMAX algorithm. Here, the presence of noise in the hyperspectral observations is taken into consideration with the original deterministic constraints suitably reformulated as probabilistic constraints. The subproblems involved are convex problems and they can be effectively solved using available convex optimization solvers. Monte Carlo simulations are presented to demonstrate the efficacy of the proposed RAVMAX algorithm over several existing pure-pixel based hyperspectral unmixing methods, including its predecessor, the AVMAX algorithm.

**Index Terms**— Hyperspectral unmixing, Convex analysis, Chance constraints, Second-order cone program

## 1. INTRODUCTION

A hyperspectral sensor records the electromagnetic scattering patterns of distinct materials over hundreds of spectral bands that range from visible to near-infrared wavelength region. The limited spatial resolution of the sensor used for hyperspectral imaging demands an effective hyperspectral unmixing (HU) scheme to extract the underlying endmember signatures (or simply endmembers) and the associated abundance maps distributed over a scene of interest [1]. Conventional HU algorithms based on a linear mixing model (to be explained later) are based on the assumption that in a given set of hyperspectral observations, there exists a pure-pixel for each endmember, namely, pixels that are fully contributed by a single endmember. HU algorithms based on the pure-pixel assumption includes pixel purity index (PPI) [2], N-finder (N-FINDR) [3] and vertex component analysis (VCA) [4]. A recent addition to this group is the AVMAX [5] algorithm, which is a convex analysis based optimization algorithm based on Winter's belief. It employs an alternating linear programming approach to solve the optimization problem. Nevertheless, the performance of the above mentioned algorithms is degraded when the observations are corrupted by noise.

This work was supported by the National Science Council (R.O.C.) under Grant NSC 96-2628-E-007-003-MY3, and by General Research Funds of Hong Kong Research Grant Council (Project Nos. CUHK415509, CUHK415908).

In this paper, we extend our AVMAX [5] algorithm to a robust AVMAX (RAVMAX) algorithm that accounts for the noise effects by employing chance constraints. We first reformulate the AVMAX subproblems into equivalent problems to which the chance constraint can be suitably applied, and then reformulate them as second-order cone problems. Hence, the chance constrained problem can be efficiently solved by any convex optimization solvers in an alternating fashion. We finally show some Monte-Carlo simulations to demonstrate the efficacy of the proposed RAVMAX algorithm, in comparison with the conventional pure-pixel based algorithms, including its predecessor.

In the remainder of the paper, the following notations are employed.  $\mathbb{R}^M$  and  $\mathbb{R}^{M \times N}$  represent the sets of all real  $M \times 1$  vectors and  $M \times N$  matrices, respectively. The symbol  $\succeq$  denotes componentwise inequality.  $\mathbf{1}_N$ ,  $\mathbf{I}_M$  and  $\mathbf{0}$  respectively represent the  $N \times 1$  all one vector, the  $M \times M$  identity matrix and an all-zero vector of proper dimension. A Gaussian distribution with mean vector  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$  is denoted as  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . The notation  $\text{sign}(\mathbf{b})$  denotes a vector whose elements are the signs of the elements in the vector  $\mathbf{b}$ ,  $|\mathbf{b}|$  denotes a column vector whose elements are the absolute values of the individual elements in  $\mathbf{b}$ , and  $\text{diag}(\mathbf{b})$  denotes a diagonal matrix with the elements of  $\mathbf{b}$  as its diagonal elements. The symbol  $[\mathbf{a}]_i$  and  $A_{ij}$  denote the  $i$ th element of the vector  $\mathbf{a}$  and  $(i, j)$ th element of matrix  $\mathbf{A}$ , respectively. Finally, the symbol  $[\mathbf{A}]_{i,:}$  corresponds to the  $i$ th row vector of  $\mathbf{A}$ .

## 2. SIGNAL MODEL AND ASSUMPTIONS

Consider a scenario in which a hyperspectral sensor measures solar electromagnetic radiation from  $N$  distinct substances. Each pixel of the hyperspectral images measured over  $M$  spectral bands can be represented by the following  $M \times N$  linear mixing model [1, 3–6]

$$\mathbf{y}[n] = \mathbf{x}[n] + \mathbf{w}[n], \quad (1)$$

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{i=1}^N s_i[n]\mathbf{a}_i, \quad \forall n = 1, \dots, L. \quad (2)$$

In (1),  $\mathbf{y}[n] = [y_1[n], \dots, y_M[n]]^T$  represents the  $n$ th observed pixel vector comprising  $M$  spectral bands and  $\mathbf{x}[n] = [x_1[n], \dots, x_M[n]]^T$  corresponds to its noise-free counterpart. The noise vector,  $\mathbf{w}[n] = [w_1[n], \dots, w_M[n]]^T$  in (1), is zero-mean, uniform white Gaussian noise vector (i.e.,  $\mathcal{N}(\mathbf{0}, \mathbf{D})$ , where  $\mathbf{D} = \sigma^2\mathbf{I}_M$  and  $\sigma$  denotes the standard deviation of the noise). In (2),  $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$  denotes the endmember signature matrix with the  $i$ th column vector  $\mathbf{a}_i$  being the  $i$ th endmember signature,  $\mathbf{s}[n] = [s_1[n], \dots, s_N[n]]^T \in \mathbb{R}^N$  is the  $n$ th abundance vector comprising  $N$  fractional abundances and  $L$  is the total number of observed pixels. Assuming prior knowledge of the number of

endmembers  $N$ , we aim to estimate the endmember signature matrix  $\mathbf{A}$  and the abundances  $\mathbf{s}[1], \dots, \mathbf{s}[L]$  from the given noisy pixels  $\mathbf{y}[1], \dots, \mathbf{y}[L]$ , under the following general assumptions [5, 6]:

- (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) There exists an index set  $\{\ell_1, \ell_2, \dots, \ell_N\}$ , such that  $\mathbf{x}[\ell_i] = \mathbf{a}_i$  for  $i = 1, \dots, N$  (i.e., the pure-pixel assumption).

For ease of later use, the convex hull [7] of the vectors  $\mathbf{a}_1, \dots, \mathbf{a}_N \in \mathbb{R}^M$  is defined as

$$\text{conv}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^N \theta_i \mathbf{a}_i \mid \mathbf{1}_N^T \boldsymbol{\theta} = 1, \boldsymbol{\theta} \succeq \mathbf{0} \right\}, \quad (3)$$

where  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_N]^T$ . A convex hull,  $\text{conv}\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$  is called an  $(N-1)$ -dimensional simplex in  $\mathbb{R}^M$  if  $\{\mathbf{a}_1, \dots, \mathbf{a}_N\} \subset \mathbb{R}^M$  is affinely independent.

### 3. REVIEW OF AVMAX ALGORITHM

As in many other HU algorithms, we begin with the dimension reduction of the observations. In our work, we employ the affine set fitting procedure in [5] to perform dimension reduction. To begin with, we start with the noise-free signal model, given by (2). The affine set fitting procedure is summarized as follows:

**Lemma 1. (Dimension reduction by affine set fitting [5, 6])** Under (A2) and (A3), a dimension-reduced pixel vector  $\tilde{\mathbf{x}}[n]$  can be obtained by an affine transformation of  $\mathbf{x}[n]$ :

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^T (\mathbf{x}[n] - \mathbf{d}) \in \mathbb{R}^{N-1}, \quad (4)$$

where  $(\mathbf{C}, \mathbf{d})$  is the affine set fitting solution given by

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

$$\mathbf{C} = [\mathbf{q}_1(\mathbf{U}\mathbf{U}^T), \mathbf{q}_2(\mathbf{U}\mathbf{U}^T), \dots, \mathbf{q}_{N-1}(\mathbf{U}\mathbf{U}^T)], \quad (6)$$

where  $\mathbf{U} = [\mathbf{x}[1] - \mathbf{d}, \dots, \mathbf{x}[L] - \mathbf{d}] \in \mathbb{R}^{M \times L}$ , and  $\mathbf{q}_i(\mathbf{R})$  denotes the orthonormal eigenvector associated with the  $i$ th principal eigenvalue of  $\mathbf{R}$ .

The affine set fitting solution  $(\hat{\mathbf{C}}, \hat{\mathbf{d}})$  for noisy observations is obtained by replacing  $\mathbf{x}[n]$  in (5) and (6) with  $\mathbf{y}[n]$ . In the noisy case,  $(\hat{\mathbf{C}}, \hat{\mathbf{d}})$  serves as a best least-squares approximation to true  $(\mathbf{C}, \mathbf{d})$  and the former asymptotically approaches the latter for large  $L$ . By (2), (4) and under (A1)-(A3), it has been proved in [8] that

$$\tilde{\mathbf{x}}[n] \in \text{conv}\{\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N\}, \quad (7)$$

where  $\boldsymbol{\gamma}_j = \mathbf{C}^T (\mathbf{a}_j - \mathbf{d}) \forall j = 1, \dots, N$  correspond to dimension reduced endmembers.

Now, the main problem is how we estimate  $\boldsymbol{\gamma}_1, \dots, \boldsymbol{\gamma}_N$  from  $\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]$ . Winter [3] proposed a belief that under (A4) the vertices of the maximum volume simplex inside the data cloud (observations) yield high fidelity estimates of the endmember signatures. Based on that, the unmixing problem [5] can be written as:

$$\begin{aligned} & \max_{\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} V(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) \\ & \text{s.t. } \boldsymbol{\nu}_i \in \text{conv}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\}, \forall i, \end{aligned} \quad (8)$$

where  $V(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N)$  is the volume of the  $(N-1)$ -dimensional simplex  $\text{conv}\{\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N\}$  in  $\mathbb{R}^{N-1}$  and is given by [9],

$$V(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) = \frac{|\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N))|}{(N-1)!}, \quad (9)$$

where

$$\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) = \begin{bmatrix} \boldsymbol{\nu}_1 & \dots & \boldsymbol{\nu}_N \\ 1 & \dots & 1 \end{bmatrix}.$$

By letting  $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]] \in \mathbb{R}^{(N-1) \times L}$  and by (3), problem (8) can be expressed as

$$\begin{aligned} & \max_{\boldsymbol{\nu}_i \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_N \in \mathbb{R}^L} |\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N))| \\ & \text{s.t. } \boldsymbol{\nu}_i = \tilde{\mathbf{X}} \boldsymbol{\theta}_i, \boldsymbol{\theta}_i \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_i = 1 \forall i. \end{aligned} \quad (10)$$

Though the constraints of (10) are convex, the non-convexity of the objective function makes the problem difficult to solve. The problem may be handled in a convenient manner by the idea of cofactor expansion and alternating optimization. The cofactor expansion of the objective function in (10) along the  $j$ th column is given by

$$\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N)) = \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}), \quad (11)$$

where  $\mathbf{b}_j = [(-1)^{i+j} \det(\mathbf{V}_{ij})]_{i=1}^{N-1} \in \mathbb{R}^{N-1}$  and the term  $\mathbf{V}_{ij} \in \mathbb{R}^{(N-1) \times (N-1)}$  is a submatrix of  $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N)$  with the  $i$ th row and  $j$ th column removed. We then consider the partial maximization of (10) with respect to  $\boldsymbol{\nu}_j$  and  $\boldsymbol{\theta}_j$ , while fixing  $\boldsymbol{\nu}_i$  and  $\boldsymbol{\theta}_i$  for all  $i \neq j$ . The problem (10) then becomes

$$\begin{aligned} & \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \left| \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \right| \\ & \text{s.t. } \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \quad (12)$$

The partial maximization problem (12) can be decomposed into the following two linear programs:

$$\begin{aligned} p^* &= \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \\ & \text{s.t. } \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1, \end{aligned} \quad (13)$$

$$\begin{aligned} q^* &= \min_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \\ & \text{s.t. } \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \quad (14)$$

The optimal solution of (12) is that of (13) if  $|p^*| > |q^*|$ , and that of (14) if  $|q^*| > |p^*|$ . This procedure of alternating optimization is performed for all the  $N$  columns (one iteration) and the relative change in the volume of the updated  $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N)$  is compared with a given threshold. If it exceeds the threshold, we continue with the next updating iteration, else we conclude that the current updated  $\boldsymbol{\nu}_j$ s are optimum. Once the optimal solution of (10), denoted by  $\boldsymbol{\nu}_1^*, \dots, \boldsymbol{\nu}_N^*$  is obtained, the endmember estimates can be recovered by using,  $\hat{\mathbf{a}}_i = \mathbf{C} \boldsymbol{\nu}_i^* + \mathbf{d}$  (by virtue of (4)) for all  $i$ . Next, we aim to make AVMAX more robust against noise effects.

### 4. ROBUST AVMAX FORMULATION AND ALGORITHM

In this section, we first do some reformulation to (13) and (14) so that chance constraints can be incorporated into the unmixing problem. Then, we move on to develop a robust version of AVMAX.

#### 4.1. Restructuring the AVMAX algorithm

Now, let  $\mathbf{B} = \text{diag}(\text{sign}(\mathbf{b}_j))$  and  $\mathbf{G} = -\mathbf{B}$ . Then, we can have  $\mathbf{G}\mathbf{G} = \mathbf{B}\mathbf{B} = \mathbf{I}_{N-1}$ ,  $\mathbf{b}_j^T \mathbf{B} = |\mathbf{b}_j|^T$  and  $\mathbf{b}_j^T \mathbf{G} = -|\mathbf{b}_j|^T$ . The subproblems (13) and (14) can then be equivalently written as:

$$\begin{aligned} p^* &= \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \mathbf{b}_j^T \mathbf{B} \boldsymbol{\nu}_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \\ & \text{s.t. } \mathbf{B} \boldsymbol{\nu}_j = \tilde{\mathbf{B}} \boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1, \end{aligned} \quad (15)$$

$$q^* = \min_{\nu_j \in \mathbb{R}^{N-1}, \theta_j \in \mathbb{R}^L} \mathbf{b}_j^T \mathbf{G} \mathbf{G} \nu_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \quad (16)$$

s.t.  $\mathbf{G} \nu_j = \mathbf{G} \tilde{\mathbf{X}} \theta_j$ ,  $\theta_j \succeq \mathbf{0}$ ,  $\mathbf{1}_L^T \theta_j = 1$ .

Then, by change of variables, we let

$$\alpha_j = \mathbf{B} \nu_j \quad (17)$$

and

$$\beta_j = \mathbf{G} \nu_j. \quad (18)$$

To facilitate the application of chance constraints to the AVMAX problem, we relax the first equality constraints of (15) and (16) and thus the corresponding subproblems are given as below:

$$p^* = \max_{\alpha_j \in \mathbb{R}^{N-1}, \theta_j \in \mathbb{R}^L} |\mathbf{b}_j|^T \alpha_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \quad (19)$$

s.t.  $\alpha_j \preceq \mathbf{B} \tilde{\mathbf{X}} \theta_j$ ,  $\theta_j \succeq \mathbf{0}$ ,  $\mathbf{1}_L^T \theta_j = 1$ ,

$$q^* = \min_{\beta_j \in \mathbb{R}^{N-1}, \theta_j \in \mathbb{R}^L} -|\mathbf{b}_j|^T \beta_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \quad (20)$$

s.t.  $\beta_j \preceq \mathbf{G} \tilde{\mathbf{X}} \theta_j$ ,  $\theta_j \succeq \mathbf{0}$ ,  $\mathbf{1}_L^T \theta_j = 1$ .

Although we relax the first constraints of the subproblems, we still can show that the optimal solutions of (19) and (20) are equivalent to that of (13) and (14), respectively, as proved in the following lemma:

**Lemma 2. (Equivalence of subproblems)** *The subproblems (19) and (20) are equivalent to (13) and (14), respectively.*

The proof of Lemma 2 is given in Appendix.

#### 4.2. Robust AVMAX Algorithm

Now we consider the unmixing problem with noisy observations given by (1). Substituting (1) into (4), we get the following dimension reduced observations

$$\tilde{\mathbf{y}}[n] \triangleq \hat{\mathbf{C}}^T (\mathbf{y}[n] - \hat{\mathbf{d}}) \cong \tilde{\mathbf{x}}[n] + \hat{\mathbf{C}}^T \mathbf{w}[n], \quad (21)$$

where  $\hat{\mathbf{C}}^T \mathbf{w}[n]$  is a random vector following  $\mathcal{N}(\mathbf{0}, \hat{\mathbf{C}}^T \mathbf{D} \hat{\mathbf{C}})$ . In matrix form (considering all the pixels,  $n = 1, \dots, L$ ), we can write the above equation as:  $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} + \hat{\mathbf{C}}^T \mathbf{W}$ , where  $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]]$ ,  $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]]$ , and  $\mathbf{W} = [\mathbf{w}[1], \dots, \mathbf{w}[L]]$ . The first inequality constraint in (19) now becomes:

$$\alpha_j \preceq \mathbf{B}(\tilde{\mathbf{Y}} - \hat{\mathbf{C}}^T \mathbf{W}) \theta_j = \mathbf{B} \tilde{\mathbf{Y}} \theta_j + \mathbf{z}_j, \quad (22)$$

where  $\mathbf{z}_j \triangleq -\mathbf{B} \hat{\mathbf{C}}^T \mathbf{W} \theta_j \sim \mathcal{N}(\mathbf{0}, \mathbf{B} \hat{\mathbf{C}}^T \mathbf{D} \hat{\mathbf{C}} \mathbf{B}^T \|\theta_j\|_2^2)$ .

Since the noise-induced vector  $\mathbf{z}_j$  is random and unknown, our approach is to replace (22) with chance constraints, as shown below

$$\Pr\{[\alpha_j]_i - [\mathbf{B}]_{i,:} \tilde{\mathbf{Y}} \theta_j \leq [\mathbf{z}_j]_i\} \geq \eta, \forall i = 1, \dots, N-1, \quad (23)$$

where  $0 < \eta < 1$  is a design parameter. A similar equation can be written for the first inequality constraint of (20), i.e.,

$$\Pr\{[\beta_j]_i - [\mathbf{G}]_{i,:} \tilde{\mathbf{Y}} \theta_j \leq [\mathbf{z}_j]_i\} \geq \eta, \forall i = 1, \dots, N-1. \quad (24)$$

The second-order cone equivalence of a chance constraint has been discussed in [7]. Specifically, for a random variable  $\varepsilon \sim \mathcal{N}(\mu, \delta^2)$  and  $t \in \mathbb{R}$ , one can show that  $\Pr(\varepsilon \leq t) \geq \eta$  is true as  $t \geq \delta \Phi^{-1}(\eta) + \mu$ , where  $\Phi^{-1}$  is the inverse of the cumulative distribution function of the standard normal random variable.

By letting  $\mathbf{Q} = \mathbf{B} \hat{\mathbf{C}}^T \mathbf{D} \hat{\mathbf{C}} \mathbf{B}^T \in \mathbb{R}^{(N-1) \times (N-1)}$  and applying the above mentioned chance constraint procedure to (23) and (24), we have

$$\sqrt{Q_{ii}} \|\theta_j\|_2 \Phi^{-1}(1 - \eta) \geq [\alpha_j]_i - [\mathbf{B}]_{i,:} \tilde{\mathbf{Y}} \theta_j, \quad (25)$$

and

$$\sqrt{Q_{ii}} \|\theta_j\|_2 \Phi^{-1}(1 - \eta) \geq [\beta_j]_i - [\mathbf{G}]_{i,:} \tilde{\mathbf{Y}} \theta_j. \quad (26)$$

for all  $i = 1, \dots, N-1$ . By replacing the first constraints of (19) and (20) with (25) and (26), respectively, the robust AVMAX problem can then be written as:

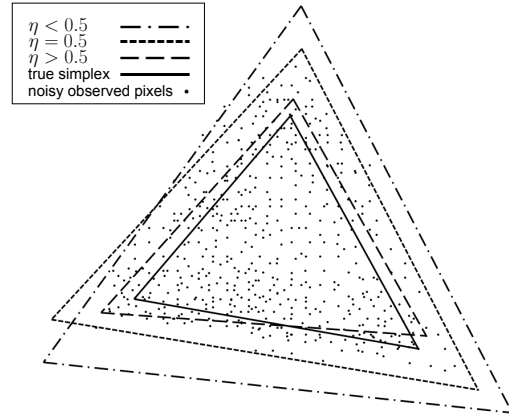
$$p^* = \max_{\alpha_j \in \mathbb{R}^{N-1}, \theta_j \in \mathbb{R}^L} |\mathbf{b}_j|^T \alpha_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \quad (27)$$

s.t.  $\sqrt{Q_{ii}} \|\theta_j\|_2 \Phi^{-1}(1 - \eta) \geq [\alpha_j]_i - [\mathbf{B}]_{i,:} \tilde{\mathbf{Y}} \theta_j$ ,  
 $\theta_j \succeq \mathbf{0}$ ,  $\mathbf{1}_L^T \theta_j = 1$ ,  $\forall i = 1, 2, \dots, N-1$ .

$$q^* = \min_{\beta_j \in \mathbb{R}^{N-1}, \theta_j \in \mathbb{R}^L} -|\mathbf{b}_j|^T \beta_j + (-1)^{N+j} \det(\mathbf{V}_{Nj}) \quad (28)$$

s.t.  $\sqrt{Q_{ii}} \|\theta_j\|_2 \Phi^{-1}(1 - \eta) \geq [\beta_j]_i - [\mathbf{G}]_{i,:} \tilde{\mathbf{Y}} \theta_j$ ,  
 $\theta_j \succeq \mathbf{0}$ ,  $\mathbf{1}_L^T \theta_j = 1$ ,  $\forall i = 1, 2, \dots, N-1$ .

The values of  $\eta$  affect the feasible sets of (27) and (28); specifically, their convexity. The following are the three possible cases: When  $\eta > 0.5$  (i.e.,  $\Phi^{-1}(1 - \eta) < 0$ ), the first constraints of both subproblems are second-order cone constraints and hence subproblems (27) and (28) are convex. If  $\eta = 0.5$  (i.e.,  $\Phi^{-1}(1 - \eta) = 0$ ), the first constraints of both subproblems reduce to those of the original AVMAX problem (as in (13) and (14)), i.e., the constraints become linear (convex). Finally, if  $\eta < 0.5$  (i.e.,  $\Phi^{-1}(1 - \eta) > 0$ ), the constraints become non-convex. The effect of  $\eta$  is illustrated in Figure 1. From our extensive numerical experiences we found that for satisfactory performance, the  $\eta$  value should lie between 0.9 and 1, in which case the subproblems (27) and (28) are convex.



**Fig. 1.** Illustration of the effect of  $\eta$  in RAVMAX for  $N = 3$ .

Some more technical aspects are discussed as follows. The subproblems (27) and (28) are solved in an alternating fashion, similar to the original AVMAX explained in Section 3, except for the following difference: after each execution of the subproblems, the corresponding  $\nu_j = \mathbf{B} \alpha_j$  (by (17)) is obtained if  $|p^*| > |q^*|$ , else  $\nu_j = \mathbf{G} \beta_j$  (by (18)) is obtained. The proposed RAVMAX algorithm uses any  $N$  pixel vectors in  $\tilde{\mathbf{X}}$  for its initialization. As the subproblems are convex (for our desired choice,  $\eta > 0.5$ ), they can be solved effectively by using available convex optimization solvers.

## 5. SIMULATIONS

This section demonstrates the efficacy of the proposed RAVMAX through comparison with other pure-pixel based HU algorithms. The algorithms considered are N-FINDR [3], VCA [4] and AVMAX [5].

**Table 1.** Average  $\phi_{en}$  and  $\phi_{ab}$  (degrees) over the various unmixing methods for different purity levels ( $\rho$ ) and SNRs.

Methods	$\rho$	$\phi_{en}$				$\phi_{ab}$					
		SNR (dB)				SNR (dB)					
		20	25	30	40	20	25	30	40		
N-FINDR	0.7	5.45	5.31	5.24	5.11	5.16	22.54	21.86	21.63	19.76	19.82
	0.85	2.65	2.67	2.66	2.65	2.61	9.60	8.37	8.03	7.93	7.77
	1	1.15	0.58	0.33	0.18	0.10	6.14	3.59	2.13	1.24	0.72
VCA	0.7	5.77	5.56	5.64	5.56	5.50	31.57	29.97	29.71	28.54	28.38
	0.85	2.79	2.70	2.67	2.71	2.61	10.83	9.45	9.00	8.89	8.82
	1	1.12	0.61	0.32	0.18	0.11	6.00	3.45	2.05	1.23	0.76
AVMAX	0.7	5.50	5.36	5.39	5.13	5.10	24.60	21.94	20.95	18.77	<b>16.48</b>
	0.85	2.77	2.64	2.65	2.69	2.65	9.15	7.96	<b>7.10</b>	<b>6.70</b>	<b>6.48</b>
	1	1.14	0.61	0.33	0.18	0.10	6.39	3.66	2.13	1.22	0.70
RAVMAX ( $0.9 < \eta < 1$ )	0.7	<b>4.87</b>	<b>4.87</b>	<b>4.88</b>	<b>4.83</b>	<b>4.90</b>	<b>18.95</b>	<b>18.15</b>	<b>18.13</b>	<b>17.83</b>	17.94
	0.85	<b>2.54</b>	<b>2.48</b>	<b>2.56</b>	<b>2.52</b>	<b>2.51</b>	<b>8.56</b>	<b>7.68</b>	7.44	7.39	7.34
	1	<b>0.79</b>	<b>0.43</b>	<b>0.24</b>	<b>0.14</b>	<b>0.08</b>	<b>4.34</b>	<b>2.60</b>	<b>1.56</b>	<b>0.98</b>	<b>0.59</b>

In all these algorithms, FCLS [10] is used to get the abundance maps. The performance of the algorithms under test is evaluated by performing 50 Monte Carlo runs for various purity levels ( $\rho$ ) and SNRs [8]. The simulation settings are  $L = 1000$  (number of pixels),  $N = 6$  (number of endmembers) and  $M = 417$  (number of observations). In each run, 1000 noise-free observed pixel vectors were synthetically generated following the signal model in (2), and the 6 endmembers (i.e., Alunite, Buddingtonite, Calcite, Copiapite, Kaolinite, and Muscovite) with 417 bands are selected from USGS library [11], and the abundance vectors  $\mathbf{s}[n]$  were generated following Dirichlet distribution  $D(\mathbf{s}[n], \boldsymbol{\mu})$  with  $\boldsymbol{\mu} = \mathbf{1}_N/N$  [4], for purity levels  $\rho = 0.7, 0.85, 1$ . The synthetic data for different SNRs were obtained by adding independent and identically distributed zero-mean Gaussian noise to the noise-free data generated, as per (1), and the SNR is defined as  $\sum_{n=1}^L \|\mathbf{x}[n]\|_2^2 / \sigma^2 ML$ . In our simulations, the noise covariance matrix is estimated from the observations, using the procedure elaborated in HySime [12]. The performance index employed is the root-mean-square (rms) spectral angle between the true one and estimated one [1, 4, 8]. The rms spectral angles between endmembers and their estimates are denoted as  $\phi_{en}$ , and those between abundance maps and their estimates are denoted as  $\phi_{ab}$ . The average  $\phi_{en}$  and  $\phi_{ab}$  of the unmixing algorithms over SNR = 20, 25, ..., 40 dB and  $\rho = 0.7, 0.85, 1$  are shown in Table 1, where each bold-faced number denotes the minimum rms spectral angle associated with a specific pair of ( $\rho$ , SNR) over all the algorithms. One can readily infer from Table 1 that the proposed RAVMAX algorithm generally yields the best performance for all the values of  $\rho$  and SNRs.

## 6. CONCLUSION

To account for the noise effects in an HU framework, we have presented a robust HU algorithm, i.e., RAVMAX. Here, we reformulated the original AVMAX problem with deterministic constraints into the one with chance constraints. The RAVMAX problem can be efficiently solved by using available second-order cone program solvers. The simulation results demonstrate the superior performance of RAVMAX over some existing benchmark HU algorithms including the original AVMAX. The performance of RAVMAX with real hyperspectral data is currently under investigation.

## 7. APPENDIX

**Proof of Lemma 2:** Firstly, it is trivial to show that the objective function of (19) is equivalent to that of (13) as  $\boldsymbol{\alpha}_j = \mathbf{B}\boldsymbol{\nu}_j$ ,  $\mathbf{b}_j^T \mathbf{B} = |\mathbf{b}_j|^T$  and  $\mathbf{B}\mathbf{B} = \mathbf{I}$ . Next, consider the subproblem (19) (after ignoring the constant term in the objective function) and let  $S$  denote

the constraint set of (19). Then we have

$$\max_{\boldsymbol{\alpha}_j \in S} |\mathbf{b}_j|^T \boldsymbol{\alpha}_j = |\mathbf{b}_j|^T \mathbf{k}_j,$$

where  $\mathbf{k}_j = [\max_{S_i} [\boldsymbol{\alpha}_j]_i]_{i=1}^{N-1}$  in which  $S_i = \{[\boldsymbol{\alpha}_j]_i \leq [\mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j]_i\}$ , implying that an optimal solution, denoted by  $(\boldsymbol{\alpha}_j^*, \boldsymbol{\theta}_j^*)$  will make the equality in  $\boldsymbol{\alpha}_j \preceq \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j$  hold (i.e., the constraint will be active). In other words, the optimal solution  $(\boldsymbol{\alpha}_j^*, \boldsymbol{\theta}_j^*)$  belongs to the set  $\{(\boldsymbol{\alpha}_j, \boldsymbol{\theta}_j) \mid \boldsymbol{\alpha}_j = \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_1^T \boldsymbol{\theta}_j = 1\}$ , which is equivalent to the constraint set of (13). Hence we can conclude that the subproblems (19) and (13) are equivalent. By a similar argument the equivalence of (20) and (14) can be proved.

## 8. REFERENCES

- [1] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44-57, Jan. 2002.
- [2] 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.
- [3] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE Conf. Imaging Spectrometry, Pasadena, CA*, Oct. 1999, pp. 266-275.
- [4] 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.
- [5] T. H. Chan, W.-K. Ma, C.-Y. Chi and A. Ambikapathi, "Hyperspectral unmixing from a convex analysis and optimization perspective," in *Proc. First IEEE WHISPERS*, Grenoble, France, August 26-28, 2009.
- [6] A. Ambikapathi, T.-H. Chan, W.-K. Ma and C.-Y. Chi, "A robust minimum volume enclosing simplex algorithm for hyperspectral unmixing," in *Proc. IEEE ICASSP*, Dallas, Mar. 14-19, 2010, pp. 1202-1205.
- [7] S. Boyd and L. Vandenberghe, *Convex Optimization*, UK: Cambridge Univ. Press, 2004.
- [8] 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 Processing*, vol. 57, no. 11, pp. 4418-4432, Nov. 2009.
- [9] G. Strang, *Linear Algebra and Its Applications*, CA: Thomson, 4th edition, 2006.
- [10] 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.
- [11] Tech. Rep., Available online: <http://speclab.cr.usgs.gov/cuprite.html>.
- [12] 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.