HYPERSPECTRAL UNMIXING FROM A CONVEX ANALYSIS AND OPTIMIZATION PERSPECTIVE

Tsung-Han Chan†, Wing-Kin Ma†∗, Chong-Yung Chi†, and A. ArulMurugan†

†Inst. Commn. Eng., National Tsing Hua Univ.
Hsinchu, Taiwan
E-mail: cychi@ee.nthu.edu.tw

∗Dept. Electronic Eng., Chinese Univ. Hong Kong
Shatin, N.T., Hong Kong
E-mail: wkma@ieee.org

ABSTRACT

In hyperspectral remote sensing, unmixing a data cube into spectral signatures and their corresponding abundance fractions plays a crucial role in analyzing the mineralogical composition of a solid surface. This paper describes a convex analysis perspective to (unsupervised) hyperspectral unmixing. Such an endeavor is not only motivated by the recent prevalence of convex optimization in signal processing, but also by the nature of hyperspectral unmixing (specifically, non-negativity and full additivity of abundances) that makes convex analysis a very suitable tool. By the notion of convex analysis, we formulate two optimization problems for solving hyperspectral unmixing, which have the intuitive ideas following the works by Craig and Winter respectively but adopt an optimization treatment different from those previous works. We show the connection of the two hyperspectral unmixing optimization problems, by proving that their optimal solutions become identical when pure pixels exist in the data. We also illustrate how the two problems can be conveniently handled by alternating linear programming. Monte Carlo simulations are presented to demonstrate the efficacy of the two hyperspectral unmixing formulations.

Index Terms— Hyperspectral unmixing, Convex analysis, Simplex geometry, Endmember identifiability, Alternating optimization

1. INTRODUCTION

In exploration of a solid surface using a hyperspectral sensor, hyperspectral unmixing is essential in analyzing the mineralogical composition of the solid surface. It is a signal processing technique that aims to decompose the spectrum of an observed pixel into a set of spectral signatures (or endmembers) and their corresponding proportions (or abundance fractions). Existing hyperspectral unmixing algorithms can be classified into two groups, between which the major distinction lies in whether pure pixels (i.e., pixels that are fully contributed from an endmember) exist per endmember in the given data set or not. Unmixing algorithms without involving pure pixels include minimum volume transform (MVT) [2], minimum volume constrained non-negative matrix factorization (MVC-NMF) [3], and minimum volume simplex analysis (MVSA) [4], to name a few. In [2], Craig reported an unmixing criterion that the endmembers are determined by the vertices of the minimum-volume simplex enclosing all the observed pixels. To find such a simplex, Craig suggested a method (i.e., MVT [2]) that literally moves the faces of a large initial simplex in toward the data cloud. In addition, some algorithms requiring the pure pixel assumption may be computationally less complex, such as pixel purity index (PPI) [5], N-finder (N-FINDR) [6], and vertex component analysis (VCA) [7]. For instance, N-FINDR [6] (proposed by Winter) is based on a criterion that the volume of a simplex formed by the purest pixels (or endmember estimates) is the maximum, and found such purest pixels by inflating the volume of a simplex inside the data set.

In this paper, we provide a convex analysis and optimization perspective to hyperspectral unmixing problems, which have the intuitive ideas from Craig's and Winter's criteria, respectively. The endeavor is not only motivated by the prevalence of convex optimization techniques in signal processing, but also by the fact that some convex analysis concepts, such as affine hull and convex hull, are quite a good match with the nature of hyperspectral unmixing (i.e., the non-negativity and full additivity of abundances). Using convex analysis, we formulate two optimization problems for hyperspectral unmixing using Craig's and Winter's criteria, and prove their optimal solutions to be identical when pure pixels exist. We also demonstrate how to use alternating linear programming to approximate the formulated problems. Finally, some Monte Carlo simulation results are presented, which show a good consistency with our analytical results.

2. PROBLEM STATEMENT

Suppose that a hyperspectral sensor with M spectral bands measures solar electromagnetic radiation reflecting from the N distinct substances. Each pixel of the measured hyperspectral image cube can be described by the following $M \times N$ linear mixing model [1–7]:

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

where $\mathbf{x}[n] = [\, x_1[n], \dots, x_M[n] \,]^T$ is the nth observed pixel vector comprising M spectral bands, $\mathbf{A} = [\, \mathbf{a}_1, \dots, \mathbf{a}_N \,] \in \mathbb{R}^{M \times N}$ denotes the signature matrix whose ith column vector $\mathbf{a}_i$ is the ith endmember, $\mathbf{s}[n] = [\, s_1[n], \dots, s_N[n] \,]^T \in \mathbb{R}^N$ is an abundance vector comprising N fractional abundances, and L is the total number of observed pixel vectors. The goal of hyperspectral unmixing is to estimate $\mathbf{A}$ and $\mathbf{s}[n]$ from the the observed pixels $\mathbf{x}[n]$.

Our convex analysis formulation for hyperspectral unmixing is based on the following assumptions [1]:

(A1) Intensities of all the abundance vectors are non-negative, i.e., $s_i[n] \geq 0$ for all i and n.

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

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

3. A REVIEW OF CONVEX ANALYSIS CONCEPTS

We review some convex analysis concepts, namely affine hull and convex hull [8], for ease of the ensuing development. Given a set of vectors $\{\mathbf{a}_1, \dots, \mathbf{a}_N\} \subset \mathbb{R}^M$ (a set of real M-vectors), the *affine hull* of $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ is defined as

$$\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \;\middle|\; \mathbf{1}_N^T \boldsymbol{\theta} = 1, \; \boldsymbol{\theta} \in \mathbb{R}^N \right\}, \quad (2)$$

where $\boldsymbol{\theta} = [\,\theta_1, \dots, \theta_N\,]^T$ and $\mathbf{1}_N$ is an $N \times 1$ all-one vector. An affine hull can always be represented by an affine set:

$$\text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \{ \mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \mid \boldsymbol{\alpha} \in \mathbb{R}^P \} \triangleq \mathcal{A}(\mathbf{C}, \mathbf{d}), \quad (3)$$

where $\mathcal{A}(\cdot, \cdot)$ is an affine set parameterized by some (non-unique) full column rank $\mathbf{C} \in \mathbb{R}^{M \times P}$ and $\mathbf{d} \in \mathbb{R}^M$, and $P \leq N - 1$ is the affine dimension. If $\mathbf{a}_1, \dots, \mathbf{a}_N$ are affinely independent (or $\mathbf{a}_1 - \mathbf{a}_N, \dots, \mathbf{a}_{N-1} - \mathbf{a}_N$ are linearly independent), then $P = N - 1$.

Given a set of vectors $\{\mathbf{a}_1, \dots, \mathbf{a}_N\} \subset \mathbb{R}^M$, the *convex hull* of $\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ is defined as

$$\text{conv}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \;\middle|\; \mathbf{1}_N^T \boldsymbol{\theta} = 1, \; \boldsymbol{\theta} \succeq \mathbf{0} \right\}, \quad (4)$$

where $\succeq$ denotes componentwise inequality, and $\mathbf{0}$ is an all-zero vector of proper dimension. A convex hull $\text{conv}\{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ is called a *simplex* if $M = N - 1$ and $\mathbf{a}_1, \dots, \mathbf{a}_N$ are affinely independent.

4. CONVEX ANALYSIS TO HYPERSPECTRAL UNMIXING PROBLEMS

Now, let us present how we use convex analysis to formulate problem (1) in a simplex representation. Under (A2) and (A3), one can infer from (1) that

$$\mathbf{x}[n] \in \text{aff}\{\mathbf{a}_1, \dots, \mathbf{a}_N\} = \mathcal{A}(\mathbf{C}, \mathbf{d}), \; \forall\, n \quad (5)$$

for some (unknown) $(\mathbf{C}, \mathbf{d}) \in \mathbb{R}^{M \times (N-1)} \times \mathbb{R}^M$ and $\text{rank}(\mathbf{C}) = N - 1$. The endmember affine hull (or affine set parameters $\mathbf{C}$ and $\mathbf{d}$) can be readily recovered from the observed pixels $\mathbf{x}[1], \dots, \mathbf{x}[L]$, as stated in the following lemma:

Lemma 1. (Endmember affine set construction [9]) *Under* (A2) *and* (A3)*, the endmember affine hull is identical to the observed pixel affine hull:*

$$\mathcal{A}(\mathbf{C}, \mathbf{d}) = \text{aff}\{\mathbf{x}[1], \dots, \mathbf{x}[L]\}. \quad (6)$$

Moreover, $(\mathbf{C}, \mathbf{d})$ can be obtained from $\{\mathbf{x}[1], \dots, \mathbf{x}[L]\}$ by the following closed-form solution

$$\mathbf{d} = \frac{1}{L} \sum_{n=1}^{L} \mathbf{x}[n], \quad (7)$$

$$\mathbf{C} = [\, \boldsymbol{q}_1(\mathbf{U}\mathbf{U}^T), \boldsymbol{q}_2(\mathbf{U}\mathbf{U}^T), \dots, \boldsymbol{q}_{N-1}(\mathbf{U}\mathbf{U}^T)\,], \quad (8)$$

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

Since $\mathbf{x}[n] \in \mathcal{A}(\mathbf{C}, \mathbf{d})$, we can write its affine representation as

$$\mathbf{x}[n] = \mathbf{C}\,\tilde{\mathbf{x}}[n] + \mathbf{d}, \quad (9)$$

where $\tilde{\mathbf{x}}[n]$ is the inverse image of $\mathbf{x}[n]$ under (9), i.e.,

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^\dagger(\mathbf{x}[n] - \mathbf{d}) \in \mathbb{R}^{N-1}, \quad (10)$$

where $\mathbf{C}^\dagger = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T$. The affinely transformed data $\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]$ can be viewed as the dimension-reduced pixels, and have a special data structure to which the simplex geometry can be applied, as stated in the following lemma:

Lemma 2. (Simplex geometry [10]) *Under (A1) to (A3), all the $\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]$ are confined by a simplex $\text{conv}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$:*

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n] \boldsymbol{\alpha}_i \in \text{conv}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\} \subset \mathbb{R}^{N-1}, \; \forall\, n \quad (11)$$

where $\boldsymbol{\alpha}_i = \mathbf{C}^\dagger(\mathbf{a}_i - \mathbf{d}) \in \mathbb{R}^{N-1}$ is the ith dimension-reduced endmember.

Compared to $\mathbf{x}[n]$ in (1) that has M dimensions, the formulation (11) has only $N - 1$ dimensions, and thus will reduce the computational complexity of the subsequent processing. Now, an interesting question is that how we can exploit simplex geometry to estimate $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N$ from $\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]$. The idea from Craig's work [2] (or Winter's work [6]) for hyperspectral unmixing that uses minimum (or maximum) volume simplex fitting approach could be adopted here. Figure 1 illustrates simplex geometry for the case of $N = 3$, where the data cloud is confined by the true simplex (solid-line triangle), and the short dashed-line triangle and long dashed-line triangle are the possible solutions for Craig's and Winter's unmixing criteria, respectively. In the ensuing development, we will formulate two optimization problems for hyperspectral unmixing using the two unmixing criteria.
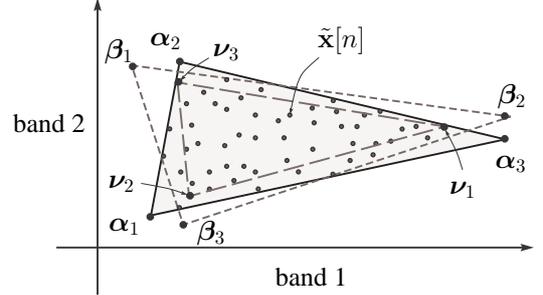
Fig. 1. Scatter plot of the dimension-reduced pixels for $N = 3$, illustrating the Craig's and Winter's criteria for hyperspectral unmixing.

4.1. Minimum volume simplex fitting approach

According to Craig's criterion [2], the unmixing problem of finding a minimum volume simplex enclosing all the dimension-reduced pixels can be written as the following optimization problem [10]:

$$\begin{aligned} \min_{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N \in \mathbb{R}^{N-1}} \quad & V(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N) \\ \text{s.t.} \quad & \tilde{\mathbf{x}}[n] \in \text{conv}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N\}, \; \forall\, n, \end{aligned} \quad (12)$$

where $V(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N)$ is the volume of the simplex $\text{conv}\{\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N\}$ given by [11]

$$V(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N) = \frac{|\det(\boldsymbol{\Delta}(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N))|}{(N-1)!}, \quad (13)$$

where

$$\boldsymbol{\Delta}(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_N) = \begin{bmatrix} \boldsymbol{\beta}_1 & \cdots & \boldsymbol{\beta}_N \\ 1 & \cdots & 1 \end{bmatrix}. \quad (14)$$

Let us consider the endmember identifiability of Craig's criterion, that is, a condition under which the optimal solution of (12) is identical to $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$. Consider the following assumption

(A4) (Pure-pixel assumption) There exist at least one index set $\{\ell_1, \ell_2, \ldots, \ell_N\}$ such that $\tilde{\mathbf{x}}[\ell_i] = \boldsymbol{\alpha}_i$ for $i = 1, \ldots, N$.

The above assumption is frequently employed in pure-pixel based unmixing methods [5–7], and can be proven to be a sufficient endmember identifiability condition of Craig's unmixing criterion. This is stated in the following theorem:

Theorem 1. (Endmember identifiability of Craig's criterion)
Under (A4), the optimal solution of (12) is uniquely given by $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$.

The proof of Theorem 1 is given in Appendix.

For ease of algorithm development for problem (12) (to be presented in Section 5), an explicit form of (12) is derived as follows. An alternative expression of the cost function in (12) is given by [11]

$$V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N) = \frac{|\det(\mathbf{B})|}{(N-1)!}, \tag{15}$$

where $\mathbf{B} = [\, \boldsymbol{\beta}_1 - \boldsymbol{\beta}_N, \ldots, \boldsymbol{\beta}_{N-1} - \boldsymbol{\beta}_N \,] \in \mathbb{R}^{(N-1)\times(N-1)}$. In addition, by (4), we rewrite the constraint of (12) in terms of $\mathbf{B}$ as

$$\tilde{\mathbf{x}}[n] = \boldsymbol{\beta}_N + \mathbf{B}\boldsymbol{\theta}_n, \tag{16}$$

where $\boldsymbol{\theta}_n \succeq \mathbf{0}$ and $\mathbf{1}_{N-1}^T \boldsymbol{\theta}_n \leq 1$. Hence, problem (12) can be equivalently written as

$$\begin{aligned} \min_{\substack{\mathbf{B}\in\mathbb{R}^{(N-1)\times(N-1)}, \\ \boldsymbol{\beta}_N, \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_L \in \mathbb{R}^{N-1}}} \quad & |\det(\mathbf{B})| \\ \text{s.t.} \quad & \boldsymbol{\theta}_n \succeq \mathbf{0}, \; \mathbf{1}_{N-1}^T \boldsymbol{\theta}_n \leq 1, \\ & \tilde{\mathbf{x}}[n] = \boldsymbol{\beta}_N + \mathbf{B}\boldsymbol{\theta}_n, \; \forall\, n. \end{aligned} \tag{17}$$

By letting $\boldsymbol{\theta}_n = \mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g}$ for all n where $\mathbf{H} = \mathbf{B}^{-1}$ and $\mathbf{g} = \mathbf{B}^{-1}\boldsymbol{\beta}_N$, one can eliminate the variables $\boldsymbol{\theta}_n$ for all n in (17) and come up with

$$\begin{aligned} \max_{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)}, \; \mathbf{g}\in\mathbb{R}^{N-1}} \quad & |\det(\mathbf{H})| \\ \text{s.t.} \quad & \mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g} \succeq \mathbf{0}, \\ & \mathbf{1}_{N-1}^T(\mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g}) \leq 1, \; \forall\, n, \end{aligned} \tag{18}$$

where the objective function is nonconvex but the constraints are affine (or convex). Problem (18) is a nonconvex problem.

4.2. Maximum volume simplex fitting approach

We now turn our attention to Winter's criterion [6]. Based on that, the unmixing problem of finding a maximum volume simplex within the set of the dimension-reduced pixels can be formulated as an optimization problem below:

$$\begin{aligned} \max_{\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} \quad & V(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) \\ \text{s.t.} \quad & \boldsymbol{\nu}_i \in \text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}, \; \forall\, i. \end{aligned} \tag{19}$$

Following the proof in Theorem 1, the endmember identifiability of Winter's criterion can also be proven under the pure-pixel assumption (see Appendix for the proof):

Theorem 2. (Endmember identifiability of Winter's criterion)
Under (A4), the optimal solution of (19) is uniquely given by $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$.

Likewise, we here formulate an explicit form of (19) for ease of algorithm development. By (4), (13), and letting $\overline{\mathbf{X}} = [\, \tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L] \,] \in \mathbb{R}^{(N-1)\times L}$, problem (19) can be expressed as

$$\begin{aligned} \max_{\substack{\boldsymbol{\nu}_i \in \mathbb{R}^{N-1} \\ \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N \in \mathbb{R}^L}} \quad & |\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))| \\ \text{s.t.} \quad & \boldsymbol{\nu}_i = \overline{\mathbf{X}}\boldsymbol{\theta}_i, \; \boldsymbol{\theta}_i \succeq \mathbf{0}, \; \mathbf{1}_L^T \boldsymbol{\theta}_i = 1 \; \forall\, i. \end{aligned} \tag{20}$$

Again, problem (20) is nonconvex since its objective function is nonconvex. Nevertheless, the constraints are affine (or convex).

Summarizing all the above results (Theorem 1 and Theorem 2), we can readily conclude that

Corollary 1. (Equivalence of Craig's and Winter's criteria) *Suppose that there exist at least one pure pixel per endmember in the data set [(A4)]. Then, problems (12) and (19) can identically yield the true dimension-reduced endmembers* $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$.

5. ALGORITHMS

In this section, we provide an alternating optimization approach to the non-convex optimization problems (18) and (20). The idea is motivated by the cofactor expansion [11]. Since such an approach to problem (18) has been reported in [10], we here only demonstrate how we use the similar approach to tackle problem (20). Now, consider the cofactor expansion for $\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))$ as follows:

$$\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)) = \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}), \tag{21}$$

where $\mathbf{b}_j = [(-1)^{i+j}\det(\boldsymbol{\mathcal{V}}_{ij})]_{i=1}^{N-1} \in \mathbb{R}^{N-1}$ and the matrix $\boldsymbol{\mathcal{V}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ is a submatrix of $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ with the ith row and jth column removed [11]. It is apparent from (21) that $\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))$ is affine in each $\boldsymbol{\nu}_j$. Hence, we consider the partial maximization of (20) with respect to $\boldsymbol{\nu}_j$ and $\boldsymbol{\theta}_j$, while fixing $\boldsymbol{\nu}_i$ and $\boldsymbol{\theta}_i$ for $i \neq j$; that is,

$$\begin{aligned} \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \overline{\mathbf{X}}\boldsymbol{\theta}_j, \; \boldsymbol{\theta}_j \succeq \mathbf{0}, \; \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \tag{22}$$

Problem (22) is still nonconvex due to the nonconvex objective function, but it can be solved in a globally optimal manner by breaking it into two linear programs (LPs):

$$\begin{aligned} p^\star = \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \overline{\mathbf{X}}\boldsymbol{\theta}_j, \; \boldsymbol{\theta}_j \succeq \mathbf{0}, \; \mathbf{1}_L^T \boldsymbol{\theta}_j = 1, \end{aligned} \tag{23}$$

$$\begin{aligned} q^\star = \min_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \overline{\mathbf{X}}\boldsymbol{\theta}_j, \; \boldsymbol{\theta}_j \succeq \mathbf{0}, \; \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \tag{24}$$

The optimal solution of (22) is chosen as that of (23) if $|p^\star| > |q^\star|$, and that of (24) if $|q^\star| > |p^\star|$. The partial maximization is conducted cyclically until some stopping rule is satisfied. Once the optimal solutions of (20), denoted by $\boldsymbol{\nu}_1^\star, \ldots, \boldsymbol{\nu}_N^\star$, are obtained, one can simply recover the endmember estimates by $\hat{\mathbf{a}}_i = \mathbf{C}\hat{\boldsymbol{\nu}}_i + \mathbf{d}$ for all i. To initialize the above approach, we can find some feasible $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ by randomly selecting the N dimension-reduced pixels from $\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]$. The proposed alternating linear programming for (20) is then termed as alternating volume maximization (AVMAX). Similarly for problem (18), we call the approach reported in [10] as alternating volume minimization (AVMIN) in this paper for ease of comparison with AVMAX for (20).

6. SIMULATIONS AND CONCLUSION

We performed one hundred independent runs of our proposed AVMIN and AVMAX for performance evaluation. In each run, 5000 observed pixels were synthetically generated following the signal model in (1), where the 5 endmembers with 224 spectral bands were randomly selected from U.S. geological survey (USGS) library [12], and the abundance vectors $\mathbf{s}[n]$ were generated following Dirichlet distribution $D(\mathbf{s}[n], \boldsymbol{\mu})$ with $\boldsymbol{\mu} = \frac{1}{N}\mathbf{1}_N$. We also tested MVSA [4] and N-FINDR [6] for performance comparison.

Two simulation scenarios are considered. One is for the data containing pure pixels, while the other for the data set without them. For the former case, we added the pure pixels in the data set. For the latter case, we discarded the observed pixels with the 2-norm of abundance vectors larger than 0.7. The sum square error (SSE) between the true endmembers and estimated ones is used as the performance index. The results are shown in Table 1. One can see that AVMIN and AVMAX (and MVSA and N-FINDR) achieve perfect endmember separation (with zero SSEs) when pure pixels exist. This observation directly supports our analytical results [Corollary 1]. For the data set without pure pixels, the performance of AVMIN and MVSA (and AVMAX and N-FINDR) are very competitive.

In conclusion, we have presented two convex analysis based formulations with intuitive ideas from Craig's and Winter's works. We have proven that these two formulated problems achieve endmember identifiability and lead to identical optimal solutions under the pure-pixel assumption. The AVMIN and AVMAX for approximating the two problems (18) and (20) respectively were also presented. Simulation results provide a validation of our analytical results.

Table 1. Performance evaluation (SSE) for the two scenarios.

Methods	AVMIN	MVSA	AVMAX	N-FINDR
Pure pixels	0	0	0	0
No pure pixels	0.0087	0.0565	10.3687	10.7585

7. APPENDIX

Proof of Theorem 1. The constraint of (12) is equivalent to

$$\mathrm{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\} \subseteq \mathrm{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}. \quad (25)$$

Under (A4), we can have

$$\mathrm{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\} = \mathrm{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}. \quad (26)$$

Hence, (25) becomes $\mathrm{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} \subseteq \mathrm{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$, which means $\boldsymbol{\alpha}_i \in \mathrm{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$, i.e.,

$$\boldsymbol{\alpha}_i = \sum_{j=1}^{N} \theta_{ij}\boldsymbol{\beta}_j \quad (27)$$

where $\sum_{j=1}^{N} \theta_{ij} = 1$ and $\theta_{ij} \geq 0$ for $i = 1, \ldots, N$. Then, from (14) and (27), one can easily infer that

$$\boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N) = \boldsymbol{\Delta}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N)\boldsymbol{\Theta}^T, \quad (28)$$

where $\boldsymbol{\Theta} = [\theta_{ij}] \in \mathbb{R}_+^{N \times N}$ and $\boldsymbol{\Theta}\mathbf{1}_N = \mathbf{1}_N$. By (28) and (13), we can have

$$V(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N) = \det \boldsymbol{\Delta}(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N)\boldsymbol{\Theta}^T \ /(N-1)! \quad (29)$$

$$= V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N)|\det(\boldsymbol{\Theta})|. \quad (30)$$

According to Lemma 1 reported in [13] that $|\det(\boldsymbol{\Theta})| \leq 1$ for which the equality holds if and only if $\boldsymbol{\Theta}$ is a permutation matrix, one can easily see that

$$V(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N) \leq V(\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N), \quad (31)$$

and the equality holds (or the the optimality of (12) is achieved) if and only if $\boldsymbol{\Theta}$ is a permutation matrix, implying that the optimum solution for $\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$ is exactly $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ by (27).

Proof of Theorem 2. The approach to proving Theorem 2 is similar to that in Theorem 1. Under (A4), the constraint of (19) becomes $\boldsymbol{\nu}_i \in \mathrm{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ for all i, and can be written as

$$\boldsymbol{\nu}_i = \sum_{j=1}^{N} \theta_{ij}\boldsymbol{\alpha}_j, \quad (32)$$

where $\sum_{j=1}^{N} \theta_{ij} = 1$ and $\theta_{ij} \geq 0$ for all i. By the definition of $\boldsymbol{\Theta}$ in (28), we can write an equivalent form of (32) given by

$$\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) = \boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N)\boldsymbol{\Theta}^T. \quad (33)$$

By (13) and Lemma 1 in [13], we can easily infer from (33) that

$$V(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) = V(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N)|\det(\boldsymbol{\Theta})|, \quad (34)$$

$$\leq V(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N), \quad (35)$$

and the equality holds (or the the optimality of (19) is achieved) if and only if $\boldsymbol{\Theta}$ is a permutation matrix. This implies that the optimum solution for $\{\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N\}$ is exactly $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ by (32).

8. REFERENCES

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

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

[3] 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.

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

[5] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVRIS data," in *Proc. Summ. JPL Airborne Earth Sci. Workshop*, Pasadena, CA, Dec. 9-14, 1995, vol. 1, pp. 23–26.

[6] 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.

[7] 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.

[8] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, 2004.

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

[10] 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," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Taipei, Taiwan, April 19-24, 2009, pp. 1089-1092.

[11] G. Strang, *Linear Algebra and Its Applications*, CA: Thomson, 4th edition, 2006.

[12] Tech. Rep., Available online: http://speclab.cr.usgs.gov/cuprite.html.

[13] F.-Y. Wang, C.-Y. Chi, T.-H. Chan, and Y. Wang, "Non-negative least-correlated component analysis for separation of dependent sources by volume maximization," to appear in *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2009.