A Simplex Volume Maximization Framework for Hyperspectral Endmember Extraction

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

Abstract—In the late 1990s, Winter proposed an endmember extraction belief that has much impact on endmember extraction techniques in hyperspectral remote sensing. The idea is to find a maximum-volume simplex whose vertices are drawn from the pixel vectors. Winter's belief has stimulated much interest, resulting in many different variations of pixel search algorithms, widely known as N-FINDR, being proposed. In this paper, we take a continuous optimization perspective to revisit Winter's belief, where the aim is to provide an alternative framework of formulating and understanding Winter's belief in a systematic manner. We first prove that, fundamentally, the existence of pure pixels is not only sufficient for the Winter problem to perfectly identify the ground-truth endmembers but also necessary. Then, under the umbrella of the Winter problem, we derive two methods using two different optimization strategies. One is by alternating optimization. The resulting algorithm turns out to be an N-FINDR variant, but, with the proposed formulation, we can pin down some of its convergence characteristics. Another is by successive optimization; interestingly, the resulting algorithm is found to exhibit some similarity to vertex component analysis. Hence, the framework provides linkage and alternative interpretations to these existing algorithms. Furthermore, we propose a robust worst case generalization of the Winter problem for accounting for perturbed pixel effects in the noisy scenario. An algorithm combining alternating optimization and projected subgradients is devised to deal with the problem. We use both simulations and real data experiments to demonstrate the viability and merits of the proposed algorithms.

Index Terms—Alternating optimization, endmember extraction, hyperspectral imaging, projected subgradient method, robust optimization, simplex volume maximization, successive optimization.

I. INTRODUCTION

HYPERSPECTRAL imaging devices with high spectral resolution utilize more than a hundred contiguous spectral bands to produce a set of remotely sensed images, thereby facilitating the identification of the composition of disparate materials over the observed scene [1], [2]. Hyperspectral imaging techniques have been applied to various fields, including space object detection and planet exploration in space [3], [4], as well as environmental monitoring and military surveillance on the Earth [5], [6]. In these hyperspectral images, each pixel physically represents a surface area and could cover more than one material, depending on the spatial resolution of the sensor. Hence, each observed pixel spectrum usually comprises multiple spectra of materials (or endmember signatures). How the endmember signatures can be accurately recovered from the measured data, or, namely, the endmember extraction problem, has been a subject of numerous investigations during the past decade [7]–[9].

A major branch of algorithms for endmember extraction is based on Craig's belief [10], which states that the vertices of a minimum-volume simplex enclosing all the observed pixel vectors may serve as reliable estimates of the endmembers. Algorithms that are based on Craig's belief, either explicitly or implicitly, include iterated constrained endmembers [11], minimum volume constrained nonnegative matrix factorization (NMF) [12], minimum dispersion constrained NMF [13], minimum volume simplex analysis [14], and minimum-volume enclosing simplex (MVES) [15], to name a few. In [15], we have shown that Craig's belief is capable of perfectly identifying the true endmembers when there exist pure pixels for each endmember, i.e., pixels composed of a single endmember. Moreover, by empirical experience, these Craig's belief-based methods work well even when the pure pixel assumption is violated to a certain extent. However, they can be expensive to implement computationally, due mainly to the complexity of Craig's problem.

Another branch of endmember extraction algorithms, which are generally simpler to implement, assumes the existence of pure pixels and attempts to search for those pure pixels as endmember estimates. The pure pixel assumption generally holds for cases where the remote sensing platforms (or aircrafts) fly at a low altitude or perform small-area surveillance [16]. In the late 1990s, Winter proposed an endmember extraction belief which is different from Craig's. Instead of attempting to find the minimum-volume data enclosing simplex, Winter suggested to find a maximum-volume simplex whose vertices are drawn from the hyperspectral data cloud [17]. This resulted in N-FINDR [17]—a class of now widely used endmember extraction algorithms in real hyperspectral data analysis. The principle of N-FINDR is to exhaustively examine a collection of pixel vectors in an attempt to find the maximum-volume

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

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

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

simplex. N-FINDR has drawn much interest in the community, and we have seen many different versions of N-FINDR algorithms based on different pixel search strategies [18]–[21], algorithm structures [22]–[24], and/or computation methods [25], [26] (often with the simplex volume computations). For example, iterative N-FINDR (I-N-FINDR) [18] uses iterative outer and inner loops, with the outer loop relating to pixel vectors and the inner loop relating to endmember estimates; sequential N-FINDR (SQ-N-FINDR) [19] adopts a per-pixel update strategy; successive N-FINDR (SC-N-FINDR) [19] adopts a per-endmember update strategy; the simplex growing algorithm (SGA) [22], [23] identifies one endmember at one time in a successive manner; and random N-FINDR [24] runs a standard N-FINDR multiple times and then fuses the results in an effort to get more accurate endmember estimates. Other than the family of N-FINDR algorithms [17]–[26], there also exist some pure pixel-based methods, such as pixel purity index (PPI) [27], vertex component analysis (VCA) [28], automatic target generation process (ATGP) [29], and automated morphological endmember extraction (AMEE) [30]. PPI [27] projects the data onto a set of randomly generated vectors and estimates the endmembers by extreme projected value search. VCA [28] also estimates an endmember by extreme projected value search and finds another new endmember by using orthogonal complement projection such that the projected data are orthogonal to the endmembers already obtained. ATGP [29] also uses orthogonal complement projection to search for the most distinctive observed pixel vectors as the endmember estimates. AMEE [30], [31] applies the extension of mathematical morphology to the spectral and spatial domain of hyperspectral data for endmember identification.

The main interest of this paper is to revisit Winter's belief using a perspective of optimization. While the intuition in Winter's original work is indeed very insightful, there is little work that places emphasis on establishing more disciplined formulations of Winter's belief, for gaining further insights and for analysis. This paper endeavors to fill this gap. In the first part of this paper, we adopt a continuous optimization formalism in formulating Winter's belief. From the formulated Winter problem, two aspects are explored: endmember identifiability analysis and methods for handling the Winter problem. This paper reveals a number of interesting results. One such result is that, using an alternating optimization strategy, we come up with a method that may be seen as yet another N-FINDR variant, but with desirable provable convergence results. Another result is that, through a successive optimization strategy, we find that VCA [28], as an existing algorithm following an idea different from Winter's, is actually similar to a Winter-based algorithm. Hence, our formulated framework provides linkage and alternative interpretations to N-FINDR and VCA. In the second part of this paper, we consider a generalization of the Winter formulation for the noisy scenario. In particular, we propose a robust Winter formulation for coping with perturbed pixel effects in the noisy scenario, and then, we devise a new algorithm from the formulation. Simulation and experimental results will be provided to demonstrate the advantages and the real applicability of the proposed methods.

A. Contributions and Related Works

The specific contributions of this paper are as follows.
1) From the continuous optimization formulation of the Winter problem, it is shown that under the noiseless model, Winter's belief yields the true endmembers if and only if pure pixels exist. A crucial implication of this is that Winter-based methods would be most suitable for the pure-pixel existent scenario.
2) We introduce two optimization strategies for handling the formulated Winter problem, namely, the alternating optimization and successive optimization strategies. From them, we derive two algorithms, called alternating volume maximization (AVMAX) and successive volume maximization (SVMAX). We also provide some analysis to the endmember identifiability of AVMAX and SVMAX, under the pure pixel condition.
3) We establish connections of the continuous Winter framework to some existing algorithms. As an interesting coincidence, we find out that AVMAX and SVMAX are algorithmically similar to SC-N-FINDR and VCA, respectively. The similarity provides the latter with a new interpretation—SC-N-FINDR and VCA were known to be developed based on pixel search ideas, and now, they may also be regarded as those built from continuous alternating/successive optimization strategies. The similarity also implies that the analysis for AVMAX/SVMAX may be applicable to SC-N-FINDR/VCA.
4) We formulate a robust worst case Winter problem for accounting for the noise effects. The newly formulated problem is in the form of max–min optimization, and it is more difficult to handle than the original Winter problem. An algorithm incorporating alternating optimization and projected subgradients is proposed to tackle the robust Winter problem.

It is worthwhile to discuss some interactions of this paper and some other existing studies. In a previous work [20], the possible connection of VCA and Winter's belief has been alluded to. Here, in order to consolidate the similarity claim, not only the successive optimization concept is needed, but we also need to derive a special determinant decomposition lemma (Lemma 3). Furthermore, we should mention concurrent works on dealing with noise in endmember extraction: In the joint Bayesian algorithm [32], a Bayesian estimation framework accounting for the presence of noise explicitly was employed; in simplex identification by split augmented Lagrangian [33], soft constraints are utilized to mitigate outlier pixel effects caused by noise, and in robust MVES [34], we apply chance constraints on the original MVES constraints. Our worst case optimization for handling the noise effects is nonetheless different from the existing works in terms of the criterion (or belief).

B. Organization of This Paper and Notations

This paper is organized as follows. In Section II, we present the problem statement of endmember extraction and its convex geometry. Section III presents the problem formulation for Winter's belief. Sections IV and V present the AVMAX and SVMAX algorithms, respectively. In Section VI, we present

the robust worst case formulation of Winter's belief and the associated algorithm. Section VII provides some computational complexity analysis. Sections VIII and IX present some simulation and experimental results to demonstrate the viability of the proposed algorithms and their advantages relative to some other existing algorithms. Finally, some conclusions are drawn in Section X.

The notations used throughout this paper are given as follows: $\mathbb{R}$ ($\mathbb{R}_+$), $\mathbb{R}^N$ ($\mathbb{R}_+^N$), and $\mathbb{R}^{M \times N}$ ($\mathbb{R}_+^{M \times N}$) denote the set of real (nonnegative) numbers, $N \times 1$ real (nonnegative) vectors, and $M \times N$ real (nonnegative) matrices, respectively; $\mathbf{1}_N$, $\mathbf{I}_N$, and $\mathbf{e}_i$ represent the $N \times 1$ all-one vector, $N \times N$ identity matrix, and unit vector with the ith entry equal to 1, respectively; $\succeq$ denotes componentwise inequality; $\| \cdot \|_2$ is the Euclidean norm, and $\mathbf{X}^\dagger$ is the pseudoinverse of a matrix $\mathbf{X}$.

II. PROBLEM STATEMENT AND CONVEX GEOMETRY

A. Problem Statement and Assumptions

Consider a scenario where a hyperspectral sensor capable of scanning an object over M spectral bands explores an unknown area of the Earth's land surface with N unknown materials. Assuming that the incident solar radiation reflects from the surface through only a single bounce and the materials are distinct [6], each pixel of the hyperspectral images can be represented by the following $M \times N$ linear mixing model:

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

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

The problem of endmember extraction is to estimate $\mathbf{A}$ from the observed pixel vectors $\mathbf{x}[n]$, assuming knowledge of N. The estimation of N can be regarded as a separate problem and has been studied in [35]–[38]. Some basic assumptions for hyperspectral endmember extraction are as follows [6].

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

(A2) Abundance fractions are proportionally distributed for each $\mathbf{x}[n]$, 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., $\text{rank}(\mathbf{A}) = N$.

There is an additional assumption called the pure pixel assumption, which holds in some practical environments, such as low altitude scanning [16].

(A4) (Pure pixel assumption) There exists an index set $\{\ell_1, \ell_2, \ldots, \ell_N\}$ such that $\mathbf{s}[\ell_i] = \mathbf{e}_i$, yielding $\mathbf{x}[\ell_i] = \mathbf{a}_i$ for $i = 1, \ldots, N$.

B. Convex Geometry Representation

In our previous works [15], [39], [40] (also [41] and [42] for earlier attempts by others), we have developed a convex ge-

ometry framework for analyzing hyperspectral mixing models, using the notion of convex analysis [43]. Here, we follow this previous framework to formulate the problem herein.

In convex analysis, an affine hull of a set of vectors, for example, $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$, is defined as

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

where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^T$. The immediate implication to the hyperspectral mixing model (1) is that every observed pixel vector $\mathbf{x}[n]$ lies in the endmember affine hull $\text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ due to (A2). The affine hull $\text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ can always be represented by

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

for some $(\mathbf{C}, \mathbf{d}) \in \mathbb{R}^{M \times P} \times \mathbb{R}^M$ and $\text{rank}(\mathbf{C}) = P$, where P is the affine dimension. Under (A3), the affine dimension P is $N - 1$. We have shown [39], [40] that $(\mathbf{C}, \mathbf{d})$ in (3) can be identified from the observed pixel vectors $\{\mathbf{x}[1], \ldots, \mathbf{x}[L]\}$, as stated in the following lemma.

Lemma 1 (Affine Set Fitting [39], [40]): Under (A2) and (A3), the affine set fitting parameter $(\mathbf{C}, \mathbf{d})$ in (3) can be obtained from $\{\mathbf{x}[1], \ldots, \mathbf{x}[L]\}$ by

$$\mathbf{d} = \frac{1}{L} \sum_{n=1}^{L} \mathbf{x}[n] \quad (4)$$

$$\mathbf{C} = [\boldsymbol{q}_1(\mathbf{H}\mathbf{H}^T), \boldsymbol{q}_2(\mathbf{H}\mathbf{H}^T), \ldots, \boldsymbol{q}_{N-1}(\mathbf{H}\mathbf{H}^T)] \quad (5)$$

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

Affine set fitting in Lemma 1 provides a way of dimension reduction. Since $\mathbf{x}[n] \in \{\mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \mid \boldsymbol{\alpha} \in \mathbb{R}^{N-1}\}$ and $\mathbf{C}$ is semiunitary, one can affinely transform each observed pixel vector $\mathbf{x}[n]$ by

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^T (\mathbf{x}[n] - \mathbf{d}) \in \mathbb{R}^{N-1}. \quad (6)$$

Substituting (1) into (6), we get

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n]\boldsymbol{\alpha}_i \quad (7)$$

where

$$\boldsymbol{\alpha}_i = \mathbf{C}^T (\mathbf{a}_i - \mathbf{d}) \in \mathbb{R}^{N-1}, \quad i = 1, \ldots, N. \quad (8)$$

One can see that the model (7) is in the same form as (1); however, the dimension of $\tilde{\mathbf{x}}[n]$ is $N - 1$, which is less than that of $\mathbf{x}[n]$. We therefore call $\tilde{\mathbf{x}}[n]$ the dimension-reduced observed pixel vectors of $\mathbf{x}[n]$ and $\boldsymbol{\alpha}_i$ the dimension-reduced endmembers.

A key concept in the geometric formulation is that of convex hulls. The convex hull of a set of vectors, for example, $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$, is defined as

$$\text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \boldsymbol{\alpha}_i \;\middle|\; \boldsymbol{\theta} \in \mathbb{R}_+^N, \mathbf{1}_N^T \boldsymbol{\theta} = 1 \right\}.$$
$$(9)$$
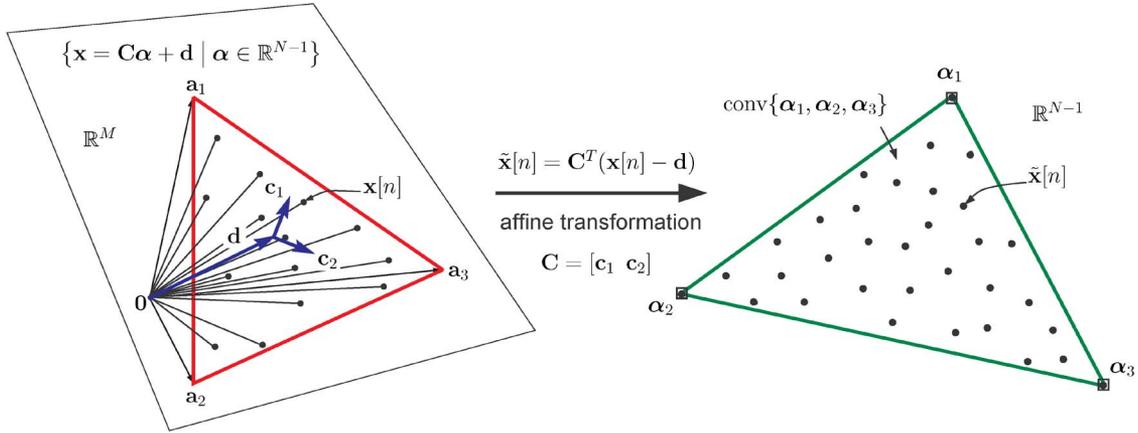
Fig. 1.   Illustration of how data geometry varies via affine transformation for $N = 3$.

Apparently, we have $\tilde{\mathbf{x}}[n] \in \text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ for all n. It has been shown [15] that the convex hull of the (dimension-reduced) true endmembers, $\text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$, is a special kind of convex hulls, namely, a *simplex*. This means that the set of all extreme points, or vertices, is $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ itself. Moreover, we can define its volume given by [44]

$$\text{vol}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N) = |\det(\boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N))| / (N-1)! \quad (10)$$

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

$$\boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N) = \begin{bmatrix} \boldsymbol{\alpha}_1 & \cdots & \boldsymbol{\alpha}_N \\ 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{N \times N}. \quad (11)$$

An illustration of the affine transformation and simplex geometry is shown in Fig. 1. The hyperspectral endmember extraction problem now becomes the estimation of $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$ from the (dimension-reduced) observed pixel vectors $\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$. Once $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$ are found, the endmembers $\mathbf{a}_1, \ldots, \mathbf{a}_N$ can be recovered by the following affine transformation:

$$\mathbf{a}_i = \mathbf{C}\boldsymbol{\alpha}_i + \mathbf{d}, \quad i = 1, \ldots, N. \quad (12)$$

III. FORMULATION OF WINTER'S ENDMEMBER EXTRACTION BELIEF

In 1999, Winter [17] proposed an endmember extraction belief that has now attracted much interest. In that work, it was believed that the ground-truth endmembers can be located by finding a collection of pixel vectors whose simplex volume is the largest. Following the hyperspectral signal geometry derivations in the last section, we herein formulate Winter's belief in the form of continuous optimization. As will be seen, this endeavor will provide many implications. Our continuous optimization formulation of Winter's belief is as follows:

$$\begin{aligned} \max_{\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} \quad & \text{vol}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) \\ \text{s.t.} \quad & \boldsymbol{\nu}_i \in \text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}, \quad i = 1, \ldots, N. \end{aligned} \quad (13)$$
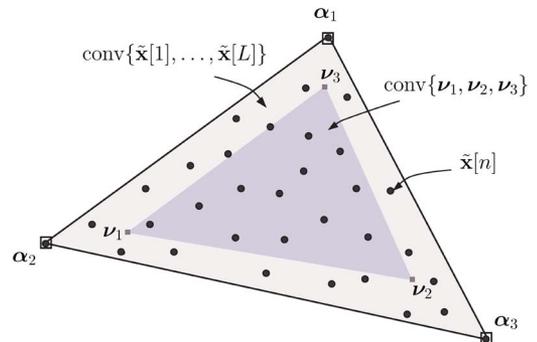
Fig. 2.   Illustration of signal geometry of Winter's belief for $N = 3$.

Problem (13), which we will name the *Winter problem* in the sequel, is to find an N-tuple $(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ from the pixel-constructed convex hull such that the associated simplex volume is maximized. A picture is used to illustrate this in Fig. 2. It should be noted that, in Winter's original work, each endmember estimate $\boldsymbol{\nu}_j$ is restricted to be any vector in $\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$.

With the Winter criterion formulated in (13), the first question that we wish to pin down is *endmember identifiability*: Under the noiseless model, is the Winter problem (13) able to perfectly identify the ground truth $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ and, if yes, what are the underlying conditions? This is addressed in the following theorem.

Theorem 1 (Endmember Identifiability for Winter's Belief): Suppose that noise is absent and that (A1) − (A3) hold. Then, the optimal solution of (13), denoted by $\{\boldsymbol{\nu}_1^\star, \ldots, \boldsymbol{\nu}_N^\star\}$, is uniquely given by $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ if and only if the pure pixel assumption (A4) is true.

Proof: The proof of Theorem 1 is given in Appendix A. ∎

The implications of Theorem 1 are twofold. First, for the pure-pixel case, solving the Winter problem can lead to perfect identification, or error-free extraction, of all the ground-truth endmembers. Second, such perfect identification would not be possible in the absence of pure pixels. The latter further implies that, in the absence of pure pixels, a Winter-based endmember extraction algorithm would be subject to estimation errors even without noise effects. An additional discussion interfacing with practical consideration is as follows.

Remark 1: Our endmember identifiability analysis in Theorem 1 is based on a perspective of *perfect* identification. Although Theorem 1 stipulates that perfect identification by Winter's belief is impossible in the absence of pure pixels, one might find that the estimation errors caused by the lack of pure pixels are acceptable in practice. This is expected to be so when the pure pixel assumption is not too seriously violated. In addition, the pure pixel assumption is satisfied in some practical environments [16].

We have examined the fundamental identifiability characteristics of the Winter problem. Our second question is on optimization methods for the Winter problem. Problem (13) is nonconvex in essence, and securing a globally optimal solution of it can be a significant challenge. However, there are simple efficient ways of handling problem (13) in a suboptimal sense. In the two subsequent sections, we will introduce two different optimization strategies for handling problem (13). The study therein will reveal more properties and connections brought about by Winter's belief.

IV. Alternating Volume Maximization

The first optimization strategy that we consider for the Winter problem is that of alternating optimization. Section IV-A will describe the alternating optimization concept, the resulting algorithm, and its connection to some existing algorithms. This will then be followed by an analysis.

A. Alternating Optimization for the Winter Problem

Alternating optimization, also known as block coordinate descent/ascent and nonlinear Gauss–Seidel, is a pragmatic approach to handling certain classes of difficult optimization problems [45]. Before proceeding to describing the application of alternating optimization to the Winter problem, we consider a simplification of the problem formulation. By (9) and (10), the Winter problem in (13) can be explicitly expressed as

$$\max_{\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} |\det\left(\boldsymbol{\Delta}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N)\right)|$$
$$\text{s.t.} \quad \boldsymbol{\nu}_i \in \mathcal{F}, \quad i = 1,\ldots,N, \quad (14)$$

where

$$\mathcal{F} = \left\{ \boldsymbol{\nu} \in \mathbb{R}^{N-1} \; \middle| \; \boldsymbol{\nu} = \widetilde{\mathbf{X}}\boldsymbol{\theta}, \boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta} = 1 \right\} \quad (15)$$

is the convex hull of $\{\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\}$ and $\widetilde{\mathbf{X}} = [\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]]$. Problem (14) can be simplified to

$$\boxed{\begin{aligned} \max_{\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} \quad & \det\left(\boldsymbol{\Delta}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N)\right) \\ \text{s.t.} \quad & \boldsymbol{\nu}_i \in \mathcal{F}, \quad i = 1,\ldots,N, \end{aligned}} \quad (16)$$

where we take out the absolute value operator from the objective function. The argument for the equivalence of problems (14) and (16) is as follows: The optimal solutions of (14) are order invariant, i.e., if $(\boldsymbol{\nu}_1^\star,\ldots,\boldsymbol{\nu}_N^\star)$ is an optimal solution of (14), then $(\boldsymbol{\nu}_{\pi_1}^\star,\ldots,\boldsymbol{\nu}_{\pi_N}^\star)$, for any permutation index set $\{\pi_1,\ldots,\pi_N\}$, is also an optimal solution to (14). Moreover,

by the basic matrix result that $\det(\mathbf{E}) = -\det(\mathbf{B})$ if $\mathbf{B}$ results from $\mathbf{E}$ by interchanging any two columns, an optimal solution of (14) in which $\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1^\star,\ldots,\boldsymbol{\nu}_N^\star)) \geq 0$ always exists. Hence, an optimal solution of problem (16) is also that of problem (14).

Now, we apply alternating optimization to the simplified Winter problem (16). The following procedure describes how alternating optimization operates:

given a starting point $(\hat{\boldsymbol{\nu}}_1,\ldots,\hat{\boldsymbol{\nu}}_N)$.
repeat the following alternating cycle:
    for $j = 1,\ldots,N$
        solve the jth partial maximization problem of (16):

$$\max_{\boldsymbol{\nu}_j \in \mathcal{F}} \det\left(\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1,\ldots,\hat{\boldsymbol{\nu}}_{j-1},\boldsymbol{\nu}_j,\hat{\boldsymbol{\nu}}_{j+1},\ldots,\hat{\boldsymbol{\nu}}_N)\right) \quad (17)$$

        and update $\hat{\boldsymbol{\nu}}_j$ as a solution of problem (17).
    end
until a stopping rule is satisfied.
output $(\hat{\boldsymbol{\nu}}_1,\ldots,\hat{\boldsymbol{\nu}}_N)$ as an approximate solution to (16).

As seen in the aforementioned procedure, alternating optimization works by maximizing the objective function of (16) with respect to (w.r.t.) only one endmember $\boldsymbol{\nu}_j$ at one time, while holding the other endmember iterates $\hat{\boldsymbol{\nu}}_i$, $i \neq j$, fixed; such partial maximizations are done in a cyclic fashion. The motivation is that the partial maximizations are much easier to solve than its full counterpart (i.e., maximizing the objective function of (16) w.r.t. $\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N$ simultaneously), as will be illustrated next.

The partial maximizations in (17) have tractable solutions. By applying a cofactor expansion to $\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N))$ [44] to the objective function of each partial maximization problem in (17), problem (17) can be reformulated as

$$\max_{\boldsymbol{\nu}_j \in \mathcal{F}} \det\left(\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1,\ldots,\hat{\boldsymbol{\nu}}_{j-1},\boldsymbol{\nu}_j,\hat{\boldsymbol{\nu}}_{j+1},\ldots,\hat{\boldsymbol{\nu}}_N)\right)$$
$$= \max_{\boldsymbol{\nu}_j \in \mathcal{F}} \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}) \quad (18)$$

where

$$\mathbf{b}_j = \left[(-1)^{i+j} \det(\boldsymbol{\mathcal{V}}_{ij}) \right]_{i=1}^{N-1} \in \mathbb{R}^{N-1} \quad (19)$$

and $\boldsymbol{\mathcal{V}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ is a submatrix of $\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1,\ldots,\hat{\boldsymbol{\nu}}_N)$ with the ith row and jth column removed [44]. We see that problem (18), for any j, is a linear program. Even better, the solution to (18) has a closed form, as shown in the following lemma.

Lemma 2: Consider the jth partial maximization problem in (18). Let

$$\mathcal{I}_j = \left\{ \ell \in \{1,\ldots,L\} \; \middle| \; \mathbf{b}_j^T \tilde{\mathbf{x}}[\ell] = \max_{n=1,\ldots,L} \mathbf{b}_j^T \tilde{\mathbf{x}}[n] \right\}. \quad (20)$$

A point $\hat{\boldsymbol{\nu}}_j$ is optimal to (18) if and only if it is a convex combination of $\{\tilde{\mathbf{x}}[n]\}_{n \in \mathcal{I}_j}$, i.e., any point

$$\hat{\boldsymbol{\nu}}_j = \sum_{\ell \in \mathcal{I}_j} \beta_\ell \tilde{\mathbf{x}}[\ell] \quad (21)$$

TABLE I
AVMAX ALGORITHM

Given	a convergence tolerance $\varepsilon > 0$, the dimension reduced data $\{\tilde{\mathbf{x}}[n]\}_{n=1}^{L}$, and the number of endmembers N.
Step 1.	initialize $(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ by randomly selecting each $\hat{\boldsymbol{\nu}}_i$ from $\{\tilde{\mathbf{x}}[n]\}_{n=1}^{L}$.
Step 2.	set $j := 1$, $\varrho := \det(\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N))$, and $\zeta = 0$.
Step 3.	calculate $\mathbf{b}_j = [(-1)^{i+j}\det(\boldsymbol{\mathcal{V}}_{ij})]_{i=1}^{N-1}$, where $\boldsymbol{\mathcal{V}}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ with the ith row and jth column removed.
Step 4.	update $\hat{\boldsymbol{\nu}}_j := \tilde{\mathbf{x}}[\ell]$ for any $\ell \in \arg_{n=1,\ldots,L} \max \ \mathbf{b}_j^T \tilde{\mathbf{x}}[n]$.
Step 5.	if $(j \bmod N) \neq 0$, then $j := j + 1$ and go to **Step 3**, else update $\zeta := \zeta + 1$, and compute $\bar{\varrho} = \det(\boldsymbol{\Delta}(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N))$.
Step 6.	if $|\bar{\varrho} - \varrho|/\varrho > \varepsilon$, then set $\varrho := \bar{\varrho}$, $j := 1$, and go to **Step 3**, else output $(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ as an approximate solution to (14).

where $\sum_{\ell \in \mathcal{I}_j} \beta_\ell = 1$, $\beta_\ell \geq 0$ for all $\ell \in \mathcal{I}_j$, is an optimal solution to (18) and vice versa.

Proof: The proof of Lemma 2 is given in Appendix B. ∎

We should add that, for instances where $\mathcal{I}_j$ contains one index only, the optimal solution of (18) is unique and is reduced to

$$\hat{\boldsymbol{\nu}}_j = \tilde{\mathbf{x}}[\ell] \tag{22}$$

$$\ell = \arg\max_n \mathbf{b}_j^T \tilde{\mathbf{x}}[n]. \tag{23}$$

Curiously, we found in our simulation results that the optimal solution of (18) was usually unique.

We name the previously developed alternating optimization method the *AVMAX method*. Table I provides the pseudocode of how we practically implement the AVMAX method. An important observation is as follows.

Remark 2: As it turns out, AVMAX is similar to the N-FINDR algorithms [17], [19], [20] in terms of the algorithmic structures—the former and latter both attempt to maximize the simplex volume by some forms of one-at-a-time pixel search. Among the various N-FINDR algorithms, AVMAX is particularly similar to the SC-N-FINDR algorithm [19]. To be specific, the pseudocode of AVMAX in Table I becomes that of SC-N-FINDR if we replace Step 3 and Step 4 by $\hat{\boldsymbol{\nu}}_j = \tilde{\mathbf{x}}[\ell]$ for any $\ell = \arg\max_{n=1,\ldots,L} |\det(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_{j-1}, \tilde{\mathbf{x}}[n], \hat{\boldsymbol{\nu}}_{j+1}, \ldots, \hat{\boldsymbol{\nu}}_N)|$ and restrict the number of alternating cycles to one (i.e., enforces termination at $\zeta = 1$). From a different perspective, one may see this similarity result as a new alternative interpretation to SC-N-FINDR—SC-N-FINDR may also be regarded as an alternating optimization algorithm under the continuous Winter formalism.

B. Analysis

Having derived an alternating optimization method for the Winter problem (AVMAX), we turn our attention to analysis. From an optimization point of view, a question of interest is whether, for a given problem, an alternating optimization algorithm would converge to a stationary point, e.g., a locally optimal solution. For AVMAX, we have the following result.

Property 1: Suppose that each partial maximization problem (18) has a unique solution.[1] Then, the AVMAX iterate

$(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ converges to a stationary point of the Winter problem (16) as the number of alternating cycles approaches infinity.

Proof: The claim of Property 1 is based on an alternating optimization convergence result for a particular class of optimization problems , Proposition 2.7.1[45]. For the Winter problem, that convergence result states that the alternating optimization iterate $(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ is guaranteed to converge to a stationary point, under the following three premises: 1) Each AVMAX partial maximization yields a unique solution; 2) the objective function of the Winter problem (16) is continuously differentiable; and 3) the feasible set $\mathcal{F}$ is closed convex. It can be verified that the latter two are satisfied by the (reformulated) Winter problem (16). ∎

Another analysis problem that we are interested in is the following: Supposing that pure pixels exist and noise is absent, can AVMAX converge to the ground-truth endmembers and, if yes, how many alternating cycles would be required? This is answered in the following property.

Property 2 (Endmember Identifiability of AVMAX): Suppose that (A1)–(A4) hold, that noise is absent, and that each partial maximization problem (18) has a unique solution. Then, for the first alternating cycle and onward, the AVMAX iterate $\{\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N\}$ is equal to $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$.

Proof: The proof of Property 2 is given in Appendix C. ∎

Property 2 implies that, for the noiseless pure-pixel case, AVMAX will perfectly identify the ground-truth endmembers in its first alternating cycle. Hence, we may intuitively expect that the convergence of AVMAX should be fast for the high signal-to-noise ratio (SNR) regime; this expectation will be shown to be true by simulations.

Remark 3: In Remark 2, we have discussed the similarity between SC-N-FINDR and AVMAX. In particular, SC-N-FINDR may loosely be viewed as a one-alternating-cycle AVMAX algorithm. From Property 2, what we realize is that one alternating cycle is in fact sufficient for the noiseless pure-pixel case. Hence, SC-N-FINDR can be seen as a "good" endmember extraction algorithm, capable of achieving perfect endmember identifiability under the ideal data condition.

V. SUCCESSIVE VOLUME MAXIMIZATION

The second optimization strategy that we propose for the Winter problem is successive optimization.

Successive optimization is an approach that requires a specific decomposition structure of the objective function. To put into context, the Winter problem needs to be cast to a suitable form. Let $\mathbf{w}_i = [\boldsymbol{\nu}_i^T \ 1]^T$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_N]$. The Winter problem in (14) can be equivalently written as

$$\max_{\mathbf{w}_1,\ldots,\mathbf{w}_N \in \mathbb{R}^N} \ |\det(\mathbf{W})|$$
$$\text{s.t.} \quad \mathbf{w}_i \in \overline{\mathcal{F}}, \quad i = 1, \ldots, N \tag{24}$$

where $\overline{\mathcal{F}} = \{\mathbf{w} \in \mathbb{R}^N | \mathbf{w} = [\boldsymbol{\nu}^T \ 1]^T, \boldsymbol{\nu} \in \mathcal{F}\}$, with $\mathcal{F}$ having been defined in (15). The set $\overline{\mathcal{F}}$ can be alternatively represented by

$$\overline{\mathcal{F}} = \{\mathbf{w} \in \mathbb{R}^N \mid \mathbf{w} = \overline{\mathbf{X}}\boldsymbol{\theta}, \boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta} = 1\} \tag{25}$$

---

[1]As mentioned, in practice, we usually found that the AVMAX partial maximization problem (18) yields a unique solution.

where $\bar{\mathbf{x}}[n] = [\tilde{\mathbf{x}}[n]^T \ 1]^T$ and $\overline{\mathbf{X}} = [\bar{\mathbf{x}}[1], \ldots, \bar{\mathbf{x}}[L]]$. Moreover, consider the following matrix lemma which plays an indispensable role in this successive optimization development.

Lemma 3: Let $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_N] \in \mathbb{R}^{M \times N}$. It holds true that

$$\sqrt{\det(\mathbf{Y}^T \mathbf{Y})} = \|\mathbf{y}_1\|_2 \left\| \mathbf{P}_{\mathbf{Y}_{1:1}}^\perp \mathbf{y}_2 \right\|_2 \cdots \left\| \mathbf{P}_{\mathbf{Y}_{1:(N-1)}}^\perp \mathbf{y}_N \right\|_2 \quad (26)$$

where $\mathbf{Y}_{1:j} = [\mathbf{y}_1, \ldots, \mathbf{y}_j]$ and $\mathbf{P}_{\mathbf{Y}_{1:j}}^\perp = \mathbf{I}_M - \mathbf{Y}_{1:j}(\mathbf{Y}_{1:j}^T \mathbf{Y}_{1:j})^\dagger \mathbf{Y}_{1:j}^T$ is the orthogonal complement projector of $\mathbf{Y}_{1:j}$.

Proof: The proof of Lemma 3 is given in Appendix D. ∎

Note that the objective function in (24) is equivalent to $|\det(\mathbf{W})| = \sqrt{\det(\mathbf{W}^T \mathbf{W})}$ since $\mathbf{W}$ is a square matrix. As a result, we can apply Lemma 3 to the objective function of problem (24), thereby obtaining the following equivalent form of the Winter problem:

$$\max_{\mathbf{w}_1, \ldots, \mathbf{w}_N \in \mathbb{R}^N} \quad f_1(\mathbf{w}_1) f_2(\mathbf{w}_1, \mathbf{w}_2) \cdots f_N(\mathbf{w}_1, \ldots, \mathbf{w}_N)$$
$$\text{s.t.} \quad \mathbf{w}_i \in \overline{\mathcal{F}}, \quad i = 1, \ldots, N. \quad (27)$$

where

$$f_1(\mathbf{w}_1) = \|\mathbf{w}_1\|_2 \quad (28)$$

$$f_j(\mathbf{w}_1, \ldots, \mathbf{w}_j) = \left\| \mathbf{P}_{\mathbf{W}_{1:(j-1)}}^\perp \mathbf{w}_j \right\|_2, \quad j = 2, \ldots, N. \quad (29)$$

With the equivalent Winter problem formulation in (27), we are now ready to consider the successive optimization method. The procedure is as follows:

for $j = 1, \ldots, N$

$$\hat{\mathbf{w}}_j = \arg \max_{\mathbf{w}_j \in \overline{\mathcal{F}}} f_j(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{j-1}, \mathbf{w}_j) \quad (30)$$

end
output $(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_N)$ as an approximate solution to (27).

As one may see, successive optimization utilizes the special decomposition structure in the objective function of (27) to perform a recursive determination of the endmembers: The jth endmember $\mathbf{w}_j$ is determined by finding a maximizer of the decomposed subobjective function $f_j(\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{j-1}, \mathbf{w}_j)$, fixing the previously determined endmembers $\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{j-1}$. The obtained partial maximizer $\hat{\mathbf{w}}_j$, together with the previous partial maximizers $\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{j-1}$, is then used to help determine the next endmember. Notice that, unlike alternating optimization, successive optimization is not an iterative method and requires no initial point to start with. Moreover, the partial maximization problems in (30) have simple solutions.

Lemma 4: For each j, the partial maximizer in (30) is given by $\hat{\mathbf{w}}_j = \bar{\mathbf{x}}[\ell]$, where

$$\ell \in \begin{cases} \arg \max_{n=1,\ldots,L} \|\bar{\mathbf{x}}[n]\|_2, & j = 1 \\ \arg \max_{n=1,\ldots,L} \left\| \mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^\perp \bar{\mathbf{x}}[n] \right\|_2, & j > 1 \end{cases} \quad (31)$$

TABLE II
SVMAX ALGORITHM

Given	the dimension reduced data $\{\tilde{\mathbf{x}}[n]\}_{n=1}^L$ and the number of endmembers N.
Step 1.	construct $\bar{\mathbf{x}}[n] = [\tilde{\mathbf{x}}[n]^T \ 1]^T$, $n = 1, \ldots, L$, and set $j = 1$.
Step 2.	obtain $\hat{\boldsymbol{\nu}}_1 = \tilde{\mathbf{x}}[\ell]$ for any $\ell \in \arg \max_n \|\bar{\mathbf{x}}[n]\|_2$, and set $\hat{\mathbf{W}} = \bar{\mathbf{x}}[\ell]$.
Step 3.	update $j := j + 1$ and obtain $\hat{\boldsymbol{\nu}}_j = \tilde{\mathbf{x}}[\ell]$ for any $\ell \in \arg \max_n \|\mathbf{P}_{\hat{\mathbf{W}}}^\perp \bar{\mathbf{x}}[n]\|_2$.
Step 4.	update $\hat{\mathbf{W}} := [\hat{\mathbf{W}} \ \bar{\mathbf{x}}[\ell]] \in \mathbb{R}^{N \times j}$ and go to **Step 3** until $j = N$.
Step 5.	output $(\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N)$ as an approximation solution to (14).

Proof: The proof of Lemma 4 is given in Appendix E. ∎

We name the successive optimization procedure developed earlier the *SVMAX* method. Table II provides a summary of the SVMAX method in pseudocode form. An important remark is as follows.

Remark 4: As another interesting coincidence, SVMAX appears to be similar to the VCA algorithm [28] in algorithmic structures. The notably similar part lies in the result in Lemma 4: If we replace the index selection in (31) by $\ell \in \arg \max_{n=1,\ldots,L} |\mathbf{r}_j^T \bar{\mathbf{x}}[n]|$, where $\mathbf{r}_j = \mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^\perp \boldsymbol{\xi} / \|\mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^\perp \boldsymbol{\xi}\|$ with $\boldsymbol{\xi}$ being randomly generated, then the resulting algorithm would be very similar to VCA. As one can see, both VCA and SVMAX employ some form of orthogonal complement projection onto the previously determined $\hat{\mathbf{w}}_1, \ldots, \hat{\mathbf{w}}_{j-1}$ at each stage, and they both do so in a successive manner.

In the previous section, we have shown that, for the noiseless pure-pixel case, AVMAX can perfectly identify the true endmembers. In fact, we can show that SVMAX possesses the same desirable property.

Property 3 (Endmember Identifiability of SVMAX): Suppose that (A1)–(A4) hold and that noise is absent. Then, SVMAX identifies the true dimension-reduced endmembers, i.e., $\{\hat{\boldsymbol{\nu}}_1, \ldots, \hat{\boldsymbol{\nu}}_N\} = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$.

Proof: The proof of Property 3 is given in Appendix F. ∎

VI. WORST CASE WINTER'S ENDMEMBER EXTRACTION PROBLEM

In the preceding sections, we have revisited Winter's endmember extraction belief, using the convex geometry notion to formulate the belief in the form of an optimization problem. Then, under the umbrella of the formulated Winter problem, we have established two endmember extraction methods, AVMAX and SVMAX, that exhibit connections to some existing methods. We have also proven that the two proposed methods can perfectly identify the ground-truth endmembers, under the pure pixel assumption and under the noiseless model. In this section, the noisy scenario is considered. A robust generalization of the Winter problem will be proposed, where we intend to mitigate the noise uncertainty effects by adopting a worst case simplex volume strategy. The proposed robust Winter belief and formulation will be described in Section VI-A, while the optimization method for dealing with the newly formulated problem will be developed in Section VI-B.
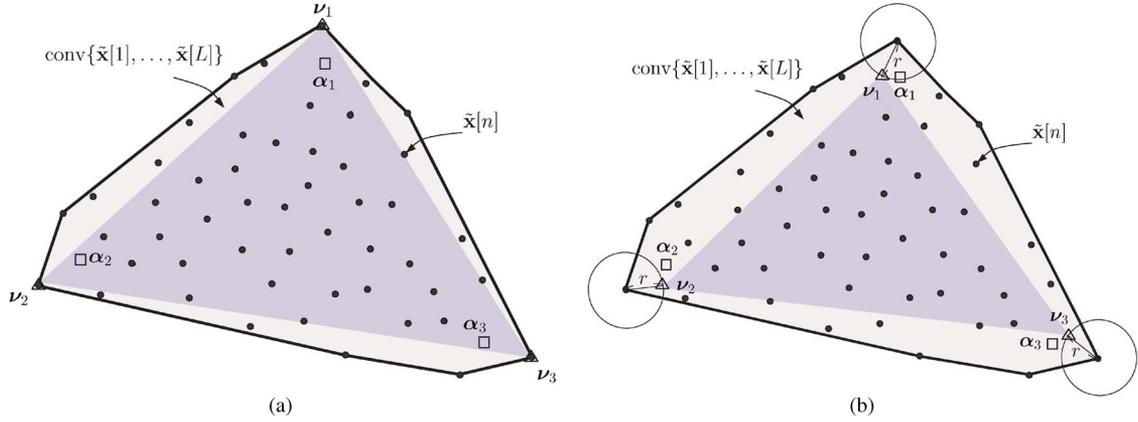
Fig. 3. Illustration of data geometry in the presence of noise for $N = 3$. (a) Solution of original Winter's belief. (b) Solution of the worst case Winter's belief.

A. Worst Case Winter's Belief

In the scenario where the pixel vectors $\tilde{\mathbf{x}}[n]$ are corrupted by additive noise, the simplex volume yielded by Winter's criterion may be larger than that of the ground truth; a picture is shown in Fig. 3(a) to illustrate such a possible instance. Our idea of mitigating such effects is to keep the endmember estimates $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ away from the boundary of the pixel-constructed convex hull $\text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$ by some distance, thereby attempting to bring $(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ closer to the ground truth $(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N)$. A picture illustrating this idea is shown in Fig. 3(b); in the figure, the circles located at the corners of $\text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$ represent the maximum backoff regions of $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$.

We formulate the aforementioned proposed belief as an optimization problem, given as follows:

$$\max_{\substack{\mathbf{v}_i \in \mathbb{R}^{N-1}, \\ i=1,\ldots,N}} \left\{ \min_{\substack{\|\mathbf{u}_i\| \leq r, \\ i=1,\ldots,N}} \text{vol}(\mathbf{v}_1 - \mathbf{u}_1, \ldots, \mathbf{v}_N - \mathbf{u}_N) \right\}$$
$$\text{s.t.} \quad \mathbf{v}_i \in \text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}, \quad i = 1, \ldots, N, \tag{32}$$

where each $\mathbf{u}_i \in \mathbb{R}^{N-1}$, $i = 1, \ldots, N$, is an error vector that lies in a norm ball $\{\mathbf{u} \in \mathbb{R}^{N-1} | \|\mathbf{u}\|_2 \leq r\}$, with the given radius r. By letting $(\mathbf{v}_1^\star, \ldots, \mathbf{v}_N^\star, \mathbf{u}_1^\star, \ldots, \mathbf{u}_N^\star)$ denote the outer–inner solution of (32), the robust Winter endmember estimates are $\boldsymbol{\nu}_i^\star = \mathbf{v}_i^\star - \mathbf{u}_i^\star$, $i = 1, \ldots, N$.

Physically speaking, our formulated problem in (32) takes a worst case robust strategy. The newly incorporated vectors $\mathbf{u}_1, \ldots, \mathbf{u}_N$ are used to enforce the backoffs of $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ in a worst case simplex volume sense, while the radius r represents how much backoff of $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ is allowed. In other words, the value r is also a sort of measure that quantifies the desired robustness against noise. As a rule of thumb, one should increase r when the noise variance, or the magnitude of the noise perturbations, increases.

B. Worst Case AVMAX

Apparently, the worst case robust Winter problem (32) is even more challenging to solve than the original Winter problem (13). In this section, we explain how we handle (32).

We first reformulate problem (32) by using the convex hull representation in (9)

$$\max_{\substack{\boldsymbol{\theta}_i \in \mathcal{S} \\ i=1,\ldots,N}} \left\{ \min_{\substack{\|\mathbf{u}_i\| \leq r, \\ i=1,\ldots,N}} \left| \det\left(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\boldsymbol{\theta}_N - \mathbf{u}_N)\right) \right| \right\} \tag{33}$$

where $\mathcal{S}$ is a simplex given by

$$\mathcal{S} = \left\{ \boldsymbol{\theta} \in \mathbb{R}^L | \boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta} = 1 \right\}. \tag{34}$$

Again, by exploiting the order invariance property of the simplex volume function (as discussed in Section IV-A), problem (33) can be simplified to

$$\max_{\substack{\boldsymbol{\theta}_i \in \mathcal{S} \\ i=1,\ldots,N}} \left\{ \min_{\substack{\|\mathbf{u}_i\|_2 \leq r, \\ i=1,\ldots,N}} \det\left(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\boldsymbol{\theta}_N - \mathbf{u}_N)\right) \right\}. \tag{35}$$

In the ensuing development, the aforementioned problem will be handled in the following manner: We begin by applying alternating optimization to (35) over $\boldsymbol{\theta}_j$'s, where each partial maximization problem w.r.t. each $\boldsymbol{\theta}_j$ is handled by the projected subgradient method [45]. In each of its subgradient updates, the optimal solution of the inner problem of (35) is approximated by alternating optimization over $\mathbf{u}_1, \ldots, \mathbf{u}_N$. The resultant method is named the *worst case AVMAX (WAVMAX)*.

1) Alternating Optimization for (35) over $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$: To derive the alternating optimization of (35) over $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$, we first express the inner problem of (35) as an objective function

$$\varphi(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N) = \min_{\substack{\|\mathbf{u}_i\| \leq r, \\ i=1,\ldots,N}} \det\left(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\boldsymbol{\theta}_N - \mathbf{u}_N)\right). \tag{36}$$

Then, by applying a cofactor expansion to $\det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\boldsymbol{\theta}_N - \mathbf{u}_N))$, the partial maximization problem of (35) w.r.t. $\boldsymbol{\theta}_j$ is formulated as

$$\max_{\boldsymbol{\theta}_j \in \mathcal{S}} \varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j) = \max_{\boldsymbol{\theta}_j \in \mathcal{S}} \left\{ \min_{\substack{\|\mathbf{u}_i\|_2 \leq r, \\ i=1,\ldots,N}} \mathbf{b}_j^T(\mathbf{U})(\widetilde{\mathbf{X}}\boldsymbol{\theta}_j - \mathbf{u}_j) + h_j(\mathbf{U}) \right\} \tag{37}$$

where $\hat{\boldsymbol{\Theta}}_j = [\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_{j-1}, \hat{\boldsymbol{\theta}}_{j+1}, \ldots, \hat{\boldsymbol{\theta}}_N]$, $\mathbf{U} = [\mathbf{u}_1, \ldots, \mathbf{u}_N]$, $\mathbf{b}_j(\mathbf{U}) = [(-1)^{i+j} \det(\mathcal{Q}_{ij})]_{i=1}^{N-1}$ in which $\mathcal{Q}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \mathbf{u}_N)$ with the ith row and the jth column removed, and $h_j(\mathbf{U}) = (-1)^{N+j} \det(\mathcal{Q}_{Nj})$. Note that (37) is a convex problem since it can be easily verified that $\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$ is concave w.r.t. $\boldsymbol{\theta}_j$ [43] due to the pointwise infimum property.[2] However, problem (37) is still difficult to handle due to the fact that $\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$ does not have a closed-form expression and is nondifferentiable. Next, we will concentrate on dealing with problem (37).

2) Projected Subgradient Method for (37): In the optimization literature, there are methods available for solving nondifferentiable convex problems, and those include the subgradient methods, cutting plane methods, and ellipsoid methods [45]. Here, we employ the projected subgradient method to deal with (37) due to its relative implementation simplicity. The projected subgradient method for (37) is as follows.

The basic idea of the projected subgradient method is to generate a sequence of points according to the following iteration:

$$\boldsymbol{\theta}_j^{(k+1)} = \left\{ \boldsymbol{\theta}_j^{(k)} - \gamma_k \mathbf{g}^{(k)} \right\}_{\mathcal{S}} \tag{38}$$

where $\mathbf{g}^{(k)}$ is a subgradient of $-\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$ at $\boldsymbol{\theta}_j^{(k)}$, γ_k is the step size, k is the current iteration number, and $\{\boldsymbol{x}\}_{\mathcal{S}} = \arg\min_{\boldsymbol{\theta} \in \mathcal{S}} \|\boldsymbol{x} - \boldsymbol{\theta}\|_2$ denotes the projection of $\boldsymbol{x}$ onto $\mathcal{S}$. The projected subgradient method keeps track of the best solution found, i.e., at each iteration, we update

$$\varphi_{best}^{(k+1)} = \max\left\{ \varphi_{best}^{(k)}, \varphi\left(\boldsymbol{\theta}_j^{(k+1)}, \hat{\boldsymbol{\Theta}}_j\right) \right\} \tag{39}$$

and update $\hat{\boldsymbol{\theta}}_j = \boldsymbol{\theta}_j^{(k+1)}$ if $\varphi_{best}^{(k+1)} = \varphi(\boldsymbol{\theta}_j^{(k+1)}, \hat{\boldsymbol{\Theta}}_j)$. As a key property, the projected subgradient method can converge to the optimal objective value for certain kinds of step size sequences, e.g., the diminishing step size sequence $\gamma_k = \gamma/\sqrt{k}$ for some $\gamma > 0$ [45].

The projection onto the simplex $\{\boldsymbol{x}\}_{\mathcal{S}}$ can be efficiently implemented by a waterfilling-type algorithm [46]. The subgradients for $-\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$ can be computed as follows. By Danskin's theorem [45], a subgradient of $-\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$ at $\boldsymbol{\theta}_j$ is given by

$$\mathbf{g} = -\widetilde{\mathbf{X}}^T \mathbf{b}_j(\mathbf{U}^\star) \tag{40}$$

where $\mathbf{U}^\star = [\mathbf{u}_1^\star, \ldots, \mathbf{u}_N^\star]$ is the optimal solution of the inner problem of (35) given by

$$(\mathbf{u}_1^\star, \ldots, \mathbf{u}_N^\star)$$
$$= \arg\min_{\substack{\|\mathbf{u}_i\|_2 \le r, \\ i=1,\ldots,N}} \det\left(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \mathbf{u}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \mathbf{u}_N)\right). \tag{41}$$

The aforementioned problem is nonconvex, but it can be approximated by alternating optimization in the same spirit as

---

TABLE III
SUMMARY OF ALTERNATING OPTIMIZATION FOR HANDLING (41)

Given	a convergence tolerance $\varepsilon > 0$, an error tolerance r, the dimension reduced data matrix $\widetilde{\mathbf{X}}$, the parameters $\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_N$, and the number of endmembers N.
Step 1.	initialize $(\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_N)$ with all zero vectors, set $j = 1$ and $\zeta_u = 0$, and calculate $\varrho = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N))$.
Step 2.	compute $\mathbf{k}_j = [(-1)^{i+j}\det(\mathcal{U}_{ij})]_{i=1}^{N-1}$ in which $\mathcal{U}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N)$ with the ith row and the jth column removed.
Step 3.	obtain $\hat{\mathbf{u}}_j = r\mathbf{k}_j/\|\mathbf{k}_j\|_2$.
Step 4.	if $(j \text{ modulo } N) \neq 0$, then $j := j + 1$ and go to **Step 2**, else update $\zeta_u := \zeta_u + 1$ and compute $\bar{\varrho} = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N))$.
Step 5.	if $|\bar{\varrho} - \varrho|/\varrho > \varepsilon$, then set $\varrho := \bar{\varrho}$, $j := 1$, and go to **Step 2**, else output $(\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_N)$ as an approximate solution of (41).

TABLE IV
SUMMARY OF WAVMAX ALGORITHM

Given	a convergence tolerance $\varepsilon > 0$, an error tolerance r, the dimension reduced data matrix $\widetilde{\mathbf{X}}$, the number of endmembers N, the subgradient step size γ, and the maximum number of subgradient iterations K.
Step 1.	initialize $(\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_N)$ by AVMAX and obtain $\hat{\mathbf{U}}$ by Table III.
Step 2.	*Alternating optimization over $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$:* set $j := 1$, $\zeta = 0$, and compute $\varrho = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N))$.
Step 3.	*Projected subgradient iterations for $\boldsymbol{\theta}_j$:* 3.1. set $k = 1$ and $\varphi_{best} = 0$. 3.2. calculate $\mathbf{b}(\hat{\mathbf{U}}) = [(-1)^{i+j}\det(\mathcal{Q}_{ij})]_{i=1}^{N-1}$ where $\mathcal{Q}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N)$ with the ith row and the jth column removed. 3.3. update $$\boldsymbol{\theta}_j := \left\{ \boldsymbol{\theta}_j + \gamma_k \widetilde{\mathbf{X}}^T \mathbf{b}(\hat{\mathbf{U}}) \right\}_{\mathcal{S}},$$ where $\gamma_k = \gamma/\sqrt{k}$ and $\{\boldsymbol{x}\}_{\mathcal{S}}$ is the projection $\boldsymbol{x}$ onto the simplex using a water-filling method [46]. 3.4. update $\hat{\mathbf{U}}$ by Table III with the given $(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j)$. 3.5. update $\hat{\boldsymbol{\theta}}_j := \boldsymbol{\theta}_j$ if $\varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j) > \varphi_{best}$ and $$\varphi_{best} := \max\left\{ \varphi_{best}, \varphi(\boldsymbol{\theta}_j, \hat{\boldsymbol{\Theta}}_j) \right\}.$$ 3.6. update $k := k + 1$ and go to 3.2 until $k > K$.
Step 4.	if $(j \text{ modulo } N) \neq 0$, then $j := j + 1$ and go to **Step 3**, else update $\zeta := \zeta + 1$, and compute $\hat{\mathbf{U}}$ by Table III and $\bar{\varrho} = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N))$.
Step 5.	if $|\bar{\varrho} - \varrho|/\varrho > \varepsilon$, then set $\varrho := \bar{\varrho}$, $j := 1$, and go to **Step 3**, else output $\hat{\boldsymbol{\nu}}_j = \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_j - \hat{\mathbf{u}}_j$ for all j as the robust Winter's estimates.

AVMAX. The partial minimization problem of (41) w.r.t. $\mathbf{u}_j$ can be expressed as

$$\min_{\|\mathbf{u}_j\|_2 \le r} \quad \mathbf{k}_j^T(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_j - \mathbf{u}_j) + (-1)^{N+j}\det(\mathcal{U}_{Nj}) \tag{42}$$

where $\mathbf{k}_j = [(-1)^{i+j}\det(\mathcal{U}_{ij})]_{i=1}^{N-1}$ in which $\mathcal{U}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_1 - \hat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\hat{\boldsymbol{\theta}}_N - \hat{\mathbf{u}}_N)$ with the ith row and the jth column removed. By Cauchy–Schwarz inequality, it can be easily shown that the solution to (42) is uniquely given by $\hat{\mathbf{u}}_j = r\mathbf{k}_j/\|\mathbf{k}_j\|_2$. The pseudocode of alternating optimization for (41) is described in Table III.

In summary, the resultant pseudocode of the proposed WAVMAX method is given in Table IV.

VII. COMPUTATIONAL COMPLEXITY

The computational complexity of the proposed methods and some existing benchmark methods is briefly discussed in this

---

[2]Let $\mathcal{Y}$ be a nonempty continuous set. The pointwise infimum property states that, if $f(\boldsymbol{x}, \boldsymbol{y})$ is concave in $\boldsymbol{x}$ for any $\boldsymbol{y} \in \mathcal{Y}$, then $g(\boldsymbol{x}) = \inf_{\boldsymbol{y} \in \mathcal{Y}} f(\boldsymbol{x}, \boldsymbol{y})$ is concave in $\boldsymbol{x}$.

TABLE V
COMPUTATIONAL COMPLEXITY ORDER OF THE VARIOUS ENDMEMBER EXTRACTION METHODS, WHERE $\eta \in (2.3, 2.9)$ [28], N IS THE NUMBER OF
ENDMEMBERS, L IS THE NUMBER OF PIXELS, ζ IS THE NUMBEROF ITERATIONS, K IS THE MAXIMUM NUMBER OF SUBGRADIENT
ITERATIONS, ζ_u IS THE NUMBER OF ALTERNATING CYCLES FOR HANDLING (41), AND ζ_w IS THE NUMBER OF
ITERATIONS REQUIRED FOR WATERFILLING ALGORITHM [46] TO IMPLEMENT $\{\boldsymbol{x}\}_{\mathcal{S}}$

Methods	VCA	SGA	I-N-FINDR	SQ-N-FINDR		
Complexity order	$\mathcal{O}(N^2 L)$	$\mathcal{O}(\sum_{i=2}^{N} i^{\eta})$	$\mathcal{O}(\zeta N^2 L)$	$\mathcal{O}(\zeta N^2 L)$		
Methods	SC-N-FINDR	AVMAX	SVMAX	WAVMAX		
Complexity order	$\mathcal{O}(N^2 L)$	$\mathcal{O}(\zeta N^2 L)$	$\mathcal{O}(N^2 L)$	$\mathcal{O}\left(N\zeta K\left((N-1)L + \zeta_w L + \zeta_u N^2 (N-1)^{\eta}\right)\right)$		

section. Since the AVMAX and SVMAX algorithms (as shown in Tables I and II) involve simple matrix/vector additions and multiplications, we can easily verify that the complexities of the AVMAX and SVMAX algorithms are $\mathcal{O}(N^2 L\zeta)$ and $\mathcal{O}(N^2 L)$, respectively, where ζ is the total number of alternating cycles used. Moreover, the computational complexity of the WAVMAX algorithm (as shown in Tables III and IV) is characterized as follows. Denote the number of alternating cycles over $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$, the maximum number of subgradient iterations, the number of alternating cycles for handling (41), and the number of iterations required for the waterfilling algorithm [46] to implement simplex projection $\{\boldsymbol{x}\}_{\mathcal{S}}$ by ζ, K, ζ_u, and ζ_w, respectively. The computational complexity of the WAVMAX method can be shown to be

$$\mathcal{O}\left(N\zeta K\left((N-1)L + \zeta_w L + \zeta_u N^2 (N-1)^{\eta}\right)\right) \quad (43)$$

where $\eta \in (2.3, 2.9)$ [28].

We also provide the computational complexities of some existing benchmark methods. The computational complexities of VCA and SGA have been reported in [22] and [28] as $\mathcal{O}(N^2 L)$ and $\mathcal{O}(\sum_{i=2}^{N} i^{\eta})$, respectively. For N-FINDR variants, such as I-N-FINDR [18], SQ-N-FINDR [19], and SC-N-FINDR [19], we consider the application of the fast simplex volume computation method using the Woodbury matrix identity [25] to them, and hence, the complexity of each simplex volume computations involved reduces from $\mathcal{O}(N^{\eta})$ to $\mathcal{O}(N)$. Therefore, it is easy to verify that the computational complexity orders of I-N-FINDR, SQ-N-FINDR, and SC-N-FINDR are $\mathcal{O}(\zeta N^2 L)$, $\mathcal{O}(\zeta N^2 L)$, and $\mathcal{O}(N^2 L)$, respectively. We summarized the aforementioned complexity analysis in Table V. One can see that VCA, SC-N-FINDR, and SVMAX have the same complexity order. Since ζ, ζ_w, and ζ_u are theoretically unknown, we will use the average computation time of the methods to illustrate their respective computational complexity in Section VIII.

VIII. COMPUTER SIMULATIONS

In this section, three Monte Carlo simulations are presented to demonstrate the performance of the proposed AVMAX, SVMAX, and WAVMAX algorithms. In each Monte Carlo simulation, 100 independent runs were performed. Section VIII-A presents the results for different SNRs. Section VIII-B presents the results for different numbers of pixels. Section VIII-C presents the results for different numbers of endmembers. For comparison, we also tested five existing endmember extraction algorithms: I-N-FINDR [18], SQ-N-FINDR [19],

SC-N-FINDR [19], SGA [22], and VCA [28]. Note that the fast simplex volume computation method using the Woodbury matrix identity [24] was used in I-N-FINDR, SQ-N-FINDR, and SC-N-FINDR, and the affine set fitting [39] (as presented in Lemma 1) was used for dimension reduction in I-N-FINDR, SQ-N-FINDR, SC-N-FINDR, and SGA for fair comparison.

Let $\hat{\mathbf{a}}_1, \ldots, \hat{\mathbf{a}}_N$ denote a set of the endmember estimates. The root-mean-square spectral angle distance between endmembers and their estimates was used as a performance measure [28]

$$\phi = \min_{\boldsymbol{\pi} \in \Pi_N} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left[\arccos\left(\frac{\mathbf{a}_i^T \hat{\mathbf{a}}_{\pi_i}}{\|\mathbf{a}_i\| \|\hat{\mathbf{a}}_{\pi_i}\|}\right)\right]^2} \quad (44)$$

where $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)$ and $\Pi_N = \{\boldsymbol{\pi} \in \mathbb{R}^N | \pi_i \in \{1, 2, \ldots, N\}, \pi_i \neq \pi_j$ for $i \neq j\}$ is the set of all the permutations of $\{1, 2, \ldots, N\}$. The performance measures defined in (44) with $N!$ permutations $\boldsymbol{\pi}$ can be efficiently solved by the Hungarian algorithm [47].

Some simulation settings are as follows. The convergence tolerance for the AVMAX algorithm and the WAVMAX algorithm and its associated subalgorithm in Table III is set to $\varepsilon = 5 \times 10^{-5}$. In addition, we set the maximum number of subgradient iterations $K = 5$ and step size $\gamma = 1$ in the WAVMAX algorithm. For comparison of computational complexity, the computation time T (in seconds) of each algorithm (implemented in Mathworks Matlab R2008a) running on a desktop computer equipped with 2.80-GHz Core i7-930 CPU and 12-GB memory is used as our computational complexity measure.

A. Monte Carlo Simulations for Various SNRs

Eight endmembers (i.e., Carnallite, Biotite, Actinolite, Andradite, Clintonite, Diaspore, Goethite, and Halloysite) with 224 bands selected from the U.S. Geological Survey (USGS) library [48] were used to produce 1000 noise-free observed pixel vectors (i.e., $N = 8$, $M = 224$, and $L = 1000$). The corresponding abundances were generated following a Dirichlet distribution with $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_N)^T = (1/N)\mathbf{1}_N$[28] to automatically enforce (A1) and (A2). To ensure (A4), the $\mathbf{s}[\ell_i] = \mathbf{e}_i$ for $i = 1, \ldots, N$ were randomly added in the generated abundances. The noisy data were eventually obtained by adding independent and identically distributed zero-mean Gaussian noise to the noise-free data for different SNRs, where SNR $= \sum_{n=1}^{L} \|\mathbf{x}[n]\|^2 / (\sigma^2 M L)$ in which σ^2 is the noise variance. The average ϕ and T of all the endmember extraction algorithms over SNR $= 5, 15, \ldots, 45, \infty$ (dB) are shown in

TABLE VI
PERFORMANCE COMPARISON OF AVERAGE ϕ (DEGREES), AVERAGE T
(SECONDS), AND AVERAGE ζ OVER DIFFERENT ENDMEMBER
EXTRACTION METHODS FOR $N = 8$, $L = 1000$, AND VARIOUS SNRs

Algorithms		SNR (dB)					
		5	15	25	35	45	∞
VCA	ϕ	15.34	3.79	1.26	0.44	0.13	**0**
	T	0.12	0.08	0.08	0.06	0.05	0.03
SGA	ϕ	13.92	3.34	0.96	0.29	**0.09**	**0**
	T	0.11	0.09	0.08	0.08	0.08	0.07
I-N-FINDR	ϕ	14.43	3.50	1.07	0.32	0.10	**0**
	T	0.21	0.20	0.27	0.17	0.21	0.29
	ζ	3.30	3.75	2.66	2.72	2.86	2.93
SQ-N-FINDR	ϕ	14.17	3.49	1.08	0.32	0.10	**0**
	T	0.19	0.15	0.13	0.12	0.12	0.12
	ζ	3.28	2.61	2.36	2.07	2.10	2.10
SC-N-FINDR	ϕ	14.59	3.74	1.18	0.32	0.11	**0**
	T	0.08	0.07	0.06	0.06	0.06	0.06
AVMAX	ϕ	15.00	3.55	1.07	0.32	0.10	**0**
	T	0.05	0.04	**0.03**	**0.03**	**0.03**	**0.02**
	ζ	3.97	3.61	3.12	2.95	2.73	2.00
SVMAX	ϕ	14.23	3.33	0.94	0.28	0.09	**0**
	T	**0.04**	**0.03**	**0.03**	**0.03**	**0.03**	**0.02**
WAVMAX	ϕ	**12.84**	**3.08**	**0.92**	0.30	**0.09**	**0**
	T	134.88	72.21	69.23	55.04	53.48	54.51
	ζ	5.05	3.32	3.58	3.08	3.15	3.28

TABLE VII
PERFORMANCE COMPARISON OF AVERAGE ϕ (DEGREES), AVERAGE T
(SECONDS), AND AVERAGE ζ OVER DIFFERENT ENDMEMBER
EXTRACTION METHODS FOR $N = 8$, SNR $= 15$ (dB),
AND VARIOUS NUMBERS OF PIXELS L

Algorithms		The number of pixels (L)					
		500	1000	2000	4000	8000	16000
VCA	ϕ	4.32	3.80	3.45	3.43	3.66	3.49
	T	0.05	0.05	0.05	0.07	0.11	0.32
SGA	ϕ	3.95	3.34	3.01	2.97	2.94	3.01
	T	0.04	0.06	0.12	0.23	0.49	1.41
I-N-FINDR	ϕ	4.04	3.50	3.19	3.10	3.10	3.12
	T	0.08	0.12	0.24	0.47	1.68	2.46
	ζ	2.77	2.75	2.84	2.78	2.87	2.74
SQ-N-FINDR	ϕ	4.07	3.50	3.19	3.11	3.11	3.12
	T	0.07	0.12	0.25	0.50	1.01	2.70
	ζ	2.63	2.59	2.73	2.75	2.78	2.72
SC-N-FINDR	ϕ	4.49	3.87	3.49	3.41	3.56	3.50
	T	0.03	0.05	0.10	0.19	0.38	1.01
AVMAX	ϕ	4.24	3.60	3.21	3.09	3.09	3.13
	T	0.02	**0.02**	0.03	**0.04**	**0.07**	**0.20**
	ζ	3.50	3.64	3.64	3.81	3.71	3.77
SVMAX	ϕ	3.96	3.33	3.07	2.94	2.95	3.03
	T	**0.01**	**0.02**	**0.02**	**0.04**	**0.07**	**0.20**
WAVMAX	ϕ	**3.87**	**3.27**	**2.77**	**2.60**	**2.55**	**2.52**
	T	21.90	61.12	185.61	729.88	2991.57	14321.55
	ζ	3.53	3.38	2.87	3.10	3.31	3.30

TABLE VIII
PERFORMANCE COMPARISON OF AVERAGE ϕ (DEGREES), AVERAGE T
(SECONDS), AND AVERAGE ζ OVER DIFFERENT ENDMEMBER
EXTRACTION METHODS FOR $L = 1000$, SNR $= 15$ (dB),
AND VARIOUS NUMBERS OF ENDMEMBERS N

Algorithms		The number of endmembers (N)					
		4	6	8	10	12	14
VCA	ϕ	2.93	4.07	5.71	7.59	9.25	10.93
	T	0.03	0.03	0.6	0.15	0.20	0.19
SGA	ϕ	2.62	3.36	5.48	7.50	9.97	11.54
	T	0.03	0.05	0.06	0.08	0.10	0.12
I-N-FINDR	ϕ	2.83	3.68	5.13	6.83	8.46	9.91
	T	0.13	0.20	0.20	0.33	0.36	0.33
	ζ	2.23	2.47	3.00	3.35	3.52	3.83
SQ-N-FINDR	ϕ	2.83	3.66	5.21	6.52	8.74	10.07
	T	0.06	0.09	0.13	0.18	0.24	0.31
	ζ	2.18	2.34	2.77	3.09	3.44	3.88
SC-N-FINDR	ϕ	2.88	3.99	5.73	7.20	9.50	10.18
	T	0.03	0.04	0.06	0.07	0.09	0.10
AVMAX	ϕ	2.83	3.68	5.34	6.93	8.90	10.28
	T	**0.01**	**0.01**	0.02	0.03	**0.03**	0.05
	ζ	2.88	3.26	3.64	4.07	4.42	4.64
SVMAX	ϕ	2.55	3.60	5.46	6.93	9.50	11.17
	T	**0.01**	**0.01**	**0.02**	**0.02**	**0.03**	**0.03**
WAVMAX	ϕ	**2.11**	**3.18**	**4.72**	**6.23**	**7.94**	**9.78**
	T	21.08	43.23	67.02	108.74	193.24	286.21
	ζ	3.04	3.53	3.86	5.01	7.89	10.64

and VCA. For most cases of the SNRs under test, WAVMAX outperforms all the other algorithms, and the computation time of AVMAX and SVMAX is less than that of all the other algorithms. The average number of iterations ζ spent by I-N-FINDR, SQ-N-FINDR, AVMAX, and WAVMAX algorithms are also displayed in Table VI. It can be inferred from Table VI that, as the SNR increases, the ζ of SQ-N-FINDR and AVMAX algorithms decreases. This meets our intuitive expectation that the AVMAX algorithm can quickly converge under the high SNR regime. In particular, in the absence of noise (SNR $= \infty$), the number of iterations ζ for the AVMAX algorithm to terminate is equal to 2; this also further validates our claim in Property 2 that AVMAX iterate equals the true endmembers after the first alternating cycle.

B. Monte Carlo Simulations for Various Numbers of Pixels

The synthetic data were generated in the same manner as in Section VIII-A, where the SNR is fixed at 15 dB and the number of pixels varies from 500 to 16 000. The average ϕ and T of the endmember extraction methods for the synthetic data sets with different L's are shown in Table VII. The error tolerance r for WAVMAX is the same as that in Section VIII-A for SNR $= 15$ dB. One can see that, for all the values of L under test, WAVMAX outperforms all the other algorithms, and the AVMAX and SVMAX spent less computation time T than the others. One can also see that the computation time of WAVMAX is the price for its better accuracy. As a future work, we may consider redundant constraint removal as a preprocessing to reduce the complexity of WAVMAX. Moreover, the number of iterations ζ of AVMAX slightly increases as L becomes larger while that of I-FINDR, SQ-N-FINDR, and WAVMAX seems independent of L.

Tables VI, where each bold-faced number denotes the minimum ϕ and minimum T for a specific SNR over all the algorithms. The error tolerance r for the proposed WAVMAX algorithm was set to $r = \lambda\sigma$, where λ is a tunable constant and was set to 1.3 and the noise standard deviation σ is assumed to be perfectly known in the simulations. One can see from Table VI that the average ϕ of all the algorithms gradually decreases as the SNR goes up, and they are equal to zero in the absence of noise. This directly validates our analytical results that AVMAX and SVMAX can yield perfect endmember identification in the noise-free case. The simulation results also show that the same perfect endmember identification result applies to their similar counterparts, SC-N-FINDR

Fig. 4. (a) Endmember signatures provided by the USGS library and endmember estimates obtained by (b) SGA and the proposed (c) AVMAX, (d) SVMAX, and (e) WAVMAX algorithms.

C. Monte Carlo Simulations for Various Numbers of Endmembers

Again, the synthetic data were generated in the same manner as in Section VIII-A, where the SNR is set to 15 dB and the N endmembers varying from 4 to 14 are randomly picked from the USGS library [48]. The average ϕ and T of the endmember extraction methods for the synthetic data sets with different N's are shown in Table VIII. The error tolerance r for WAVMAX is the same as that in Section VIII-A for SNR = 15 dB. One can observe that the performance of all the algorithms under test degrades as the number of endmembers increases. Moreover, WAVMAX outperforms the other methods, and the computation time of AVMAX and SVMAX is less than that of the other algorithms for all the values of N under test. The average ζ increases as N goes up for I-N-FINDR, SQ-N-FINDR, AVMAX, and WAVMAX.

IX. REAL HYPERSPECTRAL IMAGE EXPERIMENTS

In this section, the SGA algorithm [22] and the proposed AVMAX, SVMAX, and WAVMAX algorithms were tested on the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) hyperspectral data taken over the Cuprite mining site, Nevada, in 1997 [49]. After removing the spectral bands with low SNRs (owing to atmospheric effects), a total of 188 bands out of the 224 bands (excluding the bands 1–2, 104–113, 148–167, and 221–224) are chosen for the experiment. As has been done in [12], [15], and [28], subimages of 200 by 200 pixels for these bands are considered. Hyperspectral signal identification by minimum error [38] was used to estimate the number of endmembers in this region of interest; the result is $N = 18$. The four algorithms, SGA, AVMAX, SVMAX, and WAVMAX (with $r = 30$), were applied to the cropped data, and the fully constrained least squares method [50] with
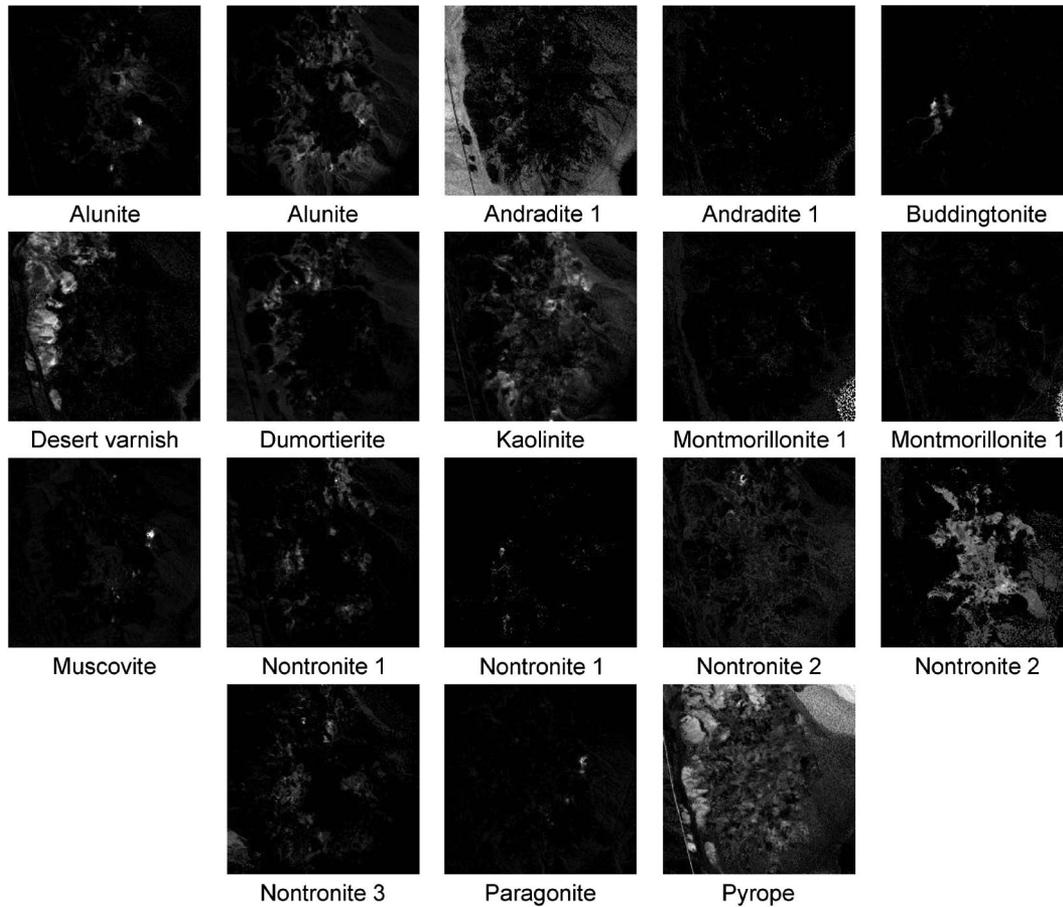
Fig. 5. Eighteen respective estimated abundances obtained by WAVMAX algorithm.

TABLE IX
MEAN-REMOVED SPECTRAL ANGLES ϕ (DEGREES) BETWEEN LIBRARY SPECTRA AND ENDMEMBERS ESTIMATED BY SGA, AVMAX, SVMAX, AND WAVMAX

	SGA	AVMAX	SVMAX	WAVMAX
Alunite	**17.83** (21.83)	25.07 (32.28)	**17.83**	25.07 (32.26)
Andradite#1	19.72	**15.65**	-	17.82 (21.62)
Andradite#2	-	**30.72**	-	-
Buddingtonite	26.19	**23.93**	26.19	25.84
Chalcedony	-	-	32.85 (**17.53**)	-
Desert Varnish	-	**11.96**	15.48	**11.96**
Dumortierite	31.48	**22.09**	31.48 (26.13)	31.49
Kaolinite	32.43 (34.22)	**24.30**	31.30	31.30
Montmorillonite#1	17.24 (**15.95**)	18.06	18.83	17.63 (17.15)
Montmorillonite#2	**20.22**	-	**20.22**	-
Muscovite	37.87	33.71	37.87	**31.48**
Nontronite#1	24.71 (26.97) (27.21)	24.00	24.00 (**23.91**) (23.63)	23.99 (26.37)
Nontronite#2	19.30	15.92 (22.62) (27.16)	22.28 (**15.89**)	15.91 (27.13)
Nontronite#3	**25.61**	30.31 (29.22)	-	26.21
Paragonite	**33.99**	36.85	36.85	35.57
Pyrope	12.02	15.67	**12.02**	14.32
Average ϕ	24.71	24.42	24.13	**24.05**

the estimated endmembers, as shown in Fig. 4, was used to estimate the corresponding abundance maps. The minerals were then identified based on the available ground-truth abundance maps [51], [52] and the results reported in [12], [15], [28], and [46]. Due to space limit, we only display the abundance maps obtained by the proposed WAVMAX algorithm in Fig. 5.

The mean-removed spectral angle of the estimated endmember $\widehat{\mathbf{a}}$ and the associated library endmember signature $\mathbf{a}$ [53] is considered as the performance measure and is defined as

$$\bar{\phi} = \arccos\left(\frac{(\widehat{\mathbf{a}} - \mathbf{m}(\widehat{\mathbf{a}}))^T (\mathbf{a} - \mathbf{m}(\mathbf{a}))}{\|\widehat{\mathbf{a}} - \mathbf{m}(\widehat{\mathbf{a}})\| \|\mathbf{a} - \mathbf{m}(\mathbf{a})\|}\right) \quad (45)$$

where $\mathbf{m}(\mathbf{a}) = (\mathbf{1}_M^T \mathbf{a}/M)\mathbf{1}_M$ for any vector $\mathbf{a} \in \mathbb{R}^M$. Table IX shows the mean-removed spectral angle $\bar{\phi}$ for the estimated endmembers obtained by the various methods, where the numbers in parentheses stand for $\bar{\phi}$ of the endmember classified

to the same mineral. As can be seen in Table IX, although WAVMAX yields the least average mean-removed spectral angle, the difference w.r.t. SGA, AVMAX, and SVMAX algorithms is less than $1°$. It should be mentioned that the Cuprite data taken by AVIRIS have a very high SNR [54], and hence, the performance improvement of WAVMAX would not be significant. We anticipate that WAVMAX may yield more promising results when the SNR of the hyperspectral data is low, for instance, the data acquired by the Earth Observing-1 Hyperion sensor [54]. Even so, the results (in Figs. 4 and 5 and Table IX) still validate the real applicability of the proposed algorithms.

X. Conclusion

We have used a continuous optimization perspective to revisit Winter's endmember extraction problem and explore the subsequent implications. In the first part of this paper, we consider a continuous optimization formulation of the Winter problem as an alternative way to study Winter's belief. In the process, we have shown by analysis that the existence of pure pixels is the sufficient and necessary endmember identifiability condition for Winter's belief. We have introduced two optimization strategies for the Winter problem, namely, alternating optimization and successive optimization, and used them to derive two algorithms, namely, AVMAX and SVMAX, respectively. We have explored their connections to some existing algorithms, namely, SC-N-FINDR and VCA. Some analyses for AVMAX and SVMAX have also been conducted. Both Monte Carlo simulation results and real data experiments have been conducted to demonstrate the performance of AVMAX and SVMAX relative to some other existing algorithms. The numerical results have also validated some implications obtained from our analysis, e.g., the potential for AVMAX to have fast convergence under high SNRs.

In the second part of this paper, we have established a worst case Winter endmember extraction problem with the aim of accounting for the presence of noise in hyperspectral data. Specifically, we have proposed a WAVMAX algorithm using a combination of alternating optimization and projected subgradients. The practical applicability of WAVMAX has been demonstrated by Monte Carlo simulations and real data experiments.

Appendix

A) Proof of Theorem 1: We first prove the sufficiency of Theorem 1. Under (A4), it can be easily verified that

$$\text{conv}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\} = \text{conv}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}. \quad (46)$$

Hence, the constraints in (13) become $\boldsymbol{\nu}_i \in \text{conv}\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$ for all i. With (11), this result can be equivalently represented by

$$\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) = \boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N)\mathbf{Q}^T \quad (47)$$

where $\mathbf{Q} \in \mathbb{R}_+^{N \times N}$, $\mathbf{Q}\mathbf{1}_N = \mathbf{1}_N$. In [55, Lemma 1], we have proved that $|\det(\mathbf{Q})| \le 1$ for any $\mathbf{Q} \in \mathbb{R}_+^{N \times N}$, $\mathbf{Q}\mathbf{1}_N = \mathbf{1}_N$,

and that equality holds if and only if $\mathbf{Q}$ is a permutation matrix. Hence, we infer that

$$\text{vol}(\boldsymbol{\nu}_1, \dots, \boldsymbol{\nu}_N) = \text{vol}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N)|\det(\mathbf{Q})|$$
$$\le \text{vol}(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N) \quad (48)$$

and that the aforementioned equality holds if and only if $\mathbf{Q}$ is a permutation matrix. This further implies that the optimum of Winter's problem in (13) is attained only by $(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N)$ or any of its permuted counterparts.

Next, we show the necessity of Theorem 1. Suppose that $\{\boldsymbol{\nu}_1^\star, \dots, \boldsymbol{\nu}_N^\star\} = \{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$. This means that $\boldsymbol{\alpha}_i \in \text{conv}\{\tilde{\mathbf{x}}[1], \dots, \tilde{\mathbf{x}}[L]\}$ for all i. Since $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$ are affinely independent, every $\boldsymbol{\alpha}_i$ cannot be represented by any nontrivial convex combination of $\{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$. Hence, by (7), we must have $\boldsymbol{\alpha}_i = \tilde{\mathbf{x}}[\ell_i]$ for some ℓ_i and for all i; that is, the pure pixel assumption (A4). ∎

B) Proof of Lemma 2: By substituting $\boldsymbol{\nu}_j = \widetilde{\mathbf{X}}\boldsymbol{\theta}_j$ into the objective function, problem (18) can be equivalently written as

$$\max_{\boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1} \mathbf{b}_j^T \widetilde{\mathbf{X}}\boldsymbol{\theta}_j \quad (49)$$

where the term $(-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj})$ in (18) is removed without loss of generality (w.l.o.g.). By letting $\theta_{jn} = [\boldsymbol{\theta}_j]_n$, we have

$$\mathbf{b}_j^T \widetilde{\mathbf{X}}\boldsymbol{\theta}_j = \sum_{n=1}^L \theta_{jn} \mathbf{b}_j^T \tilde{\mathbf{x}}[n] \le \max_{n=1,\dots,L} \mathbf{b}_j^T \tilde{\mathbf{x}}[n] \quad (50)$$

for any $\boldsymbol{\theta}_j \succeq \mathbf{0}$ and $\mathbf{1}_L^T \boldsymbol{\theta}_j = 1$. Moreover, it can be verified that the equality in (50) holds if and only if

$$\sum_{n \in \mathcal{I}_j} \theta_{jn} = 1,$$
$$\mathcal{I}_j = \left\{\ell \in \{1, \dots, L\} \Big| \mathbf{b}_j^T \tilde{\mathbf{x}}[\ell] = \max_{n=1,\dots,L} \mathbf{b}_j^T \tilde{\mathbf{x}}[n]\right\}. \quad (51)$$

Hence, the solution (21) directly follows from (51). ∎

C) Proof of Property 2: Consider Step 4 in Table I for $j = 1$. Under the assumption that the solution of each alternating maximizer is unique, we have

$$\hat{\boldsymbol{\nu}}_1 = \tilde{\mathbf{x}}[\ell], \quad \ell = \arg \max_{n=1,\dots,L} \bar{\mathbf{b}}_1^T \bar{\mathbf{x}}[n] \quad (52)$$

where $\bar{\mathbf{b}}_1 = [\mathbf{b}_1^T \quad (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{N1})]^T$ and $\bar{\mathbf{x}}[n] = [\tilde{\mathbf{x}}[n]^T \quad 1]^T$. Recall that $\bar{\mathbf{x}}[n] = \sum_{i=1}^N s_i[n]\bar{\boldsymbol{\alpha}}_i$ and $\bar{\boldsymbol{\alpha}}_i = [\boldsymbol{\alpha}_i^T \quad 1]^T$. Then, we can infer from (52) and assumptions (A1) and (A2) that

$$\max_n \bar{\mathbf{b}}_1^T \bar{\mathbf{x}}[n] = \max_n \sum_{i=1}^N s_i[n]\bar{\mathbf{b}}_1^T \bar{\boldsymbol{\alpha}}_i \le \max_{i=1,\dots,N} \bar{\mathbf{b}}_1^T \bar{\boldsymbol{\alpha}}_i. \quad (53)$$

Assume w.l.o.g. that $\bar{\mathbf{b}}_1^T \bar{\boldsymbol{\alpha}}_1 = \max_{i=1,\dots,N} \bar{\mathbf{b}}_1^T \bar{\boldsymbol{\alpha}}_i$. By the solution uniqueness in (52), it can be easily verified that the equality in (53) holds if and only if $\hat{\boldsymbol{\nu}}_1 = \tilde{\mathbf{x}}[\ell] = \boldsymbol{\alpha}_1$. Next, consider

$j \geq 2$. Since $\bar{\mathbf{b}}_j^T$ is the jth row vector of the adjoint matrix of $\boldsymbol{\Delta}(\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{j-1}, \hat{\boldsymbol{\nu}}_j, \ldots, \hat{\boldsymbol{\nu}}_N)$[44], it satisfies

$$\bar{\mathbf{b}}_j^T \bar{\boldsymbol{\alpha}}_i = 0, \quad i = 1, \ldots, j-1 \tag{54}$$

due to $\mathrm{adj}(\mathbf{B})\mathbf{B} = \det(\mathbf{B})\mathbf{I}$ where $\mathrm{adj}(\mathbf{B})$ is the adjoint matrix of a square matrix $\mathbf{B}$. Hence, it follows from (54) and the assumptions (A1) and (A2) that

$$\max_n \bar{\mathbf{b}}_j^T \bar{\mathbf{x}}[n] = \max_n \sum_{i=j}^N s_i[n] \bar{\mathbf{b}}_j^T \bar{\boldsymbol{\alpha}}_i \leq \max_{i \in \{j, \ldots, N\}} \bar{\mathbf{b}}_j^T \bar{\boldsymbol{\alpha}}_i. \tag{55}$$

By employing the same aforementioned proof, we obtain, w.l.o.g., $\hat{\boldsymbol{\nu}}_j = \tilde{\mathbf{x}}[\ell] = \boldsymbol{\alpha}_j$. The proof of Property 2 is therefore complete. ∎

D) Proof of Lemma 3: To prove Lemma 3, we first provide the following lemma.

Lemma 5: Let $\mathbf{A} \in \mathbb{R}^{m \times p}$, $\mathbf{B} \in \mathbb{R}^{m \times q}$, and $\mathbf{C} = [\mathbf{A} \ \mathbf{B}]$. The orthogonal complement projector of $\mathbf{C}$, denoted by $\mathbf{P}_{\mathbf{C}}^{\perp}$, is identical to

$$\mathbf{P}_{\mathbf{C}}^{\perp} = \mathbf{P}_{\tilde{\mathbf{B}}}^{\perp} \mathbf{P}_{\mathbf{A}}^{\perp}$$

where $\tilde{\mathbf{B}} = \mathbf{P}_{\mathbf{A}}^{\perp} \mathbf{B}$.

Proof: Let $\mathbf{D} = [\mathbf{A} \ \tilde{\mathbf{B}}]$. As a basic result of matrix analysis, the range spaces of $\mathbf{C}$ and $\mathbf{D}$ are identical. Thus, their orthogonal projectors are identical, i.e., $\mathbf{P}_{\mathbf{C}} = \mathbf{P}_{\mathbf{D}}$. Let us consider $\mathbf{P}_{\mathbf{D}}$

$$\mathbf{P}_{\mathbf{D}} = \mathbf{D}(\mathbf{D}^T\mathbf{D})^{\dagger}\mathbf{D}^T = [\mathbf{A} \ \tilde{\mathbf{B}}] \begin{bmatrix} \mathbf{A}^T\mathbf{A} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{B}}^T\tilde{\mathbf{B}} \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{A}^T \\ \tilde{\mathbf{B}}^T \end{bmatrix}$$

$$= \mathbf{A}(\mathbf{A}^T\mathbf{A})^{\dagger}\mathbf{A}^T + \tilde{\mathbf{B}}(\tilde{\mathbf{B}}^T\tilde{\mathbf{B}})^{\dagger}\tilde{\mathbf{B}}^T = \mathbf{P}_{\mathbf{A}} + \mathbf{P}_{\tilde{\mathbf{B}}}. \tag{56}$$

Moreover, we have

$$\mathbf{P}_{\mathbf{C}}^{\perp} = \mathbf{I}_m - \mathbf{P}_{\mathbf{C}} = \mathbf{I}_m - \mathbf{P}_{\mathbf{D}} = \mathbf{I}_m - \mathbf{P}_{\mathbf{A}} - \mathbf{P}_{\tilde{\mathbf{B}}}$$

$$= \mathbf{I}_m - \mathbf{P}_{\mathbf{A}} - \mathbf{P}_{\tilde{\mathbf{B}}} + \mathbf{P}_{\mathbf{A}}\mathbf{P}_{\tilde{\mathbf{B}}} \tag{57}$$

$$= (\mathbf{I}_m - \mathbf{P}_{\tilde{\mathbf{B}}})(\mathbf{I}_m - \mathbf{P}_{\mathbf{A}}) = \mathbf{P}_{\tilde{\mathbf{B}}}^{\perp}\mathbf{P}_{\mathbf{A}}^{\perp} \tag{58}$$

where (57) is owing to $\mathbf{P}_{\mathbf{A}}\mathbf{P}_{\tilde{\mathbf{B}}} = \mathbf{0}$. ∎

Now, we prove Lemma 3. Consider two cases as follows: (C1) $\mathbf{Y}$ is of full column rank, and (C2) $\mathbf{Y}$ is not of full column rank. For (C1), $\mathbf{Y}^T\mathbf{Y}$ is positive definite, and $\det(\mathbf{Y}^T\mathbf{Y}) > 0$. By considering the partitioned form

$$\mathbf{Y}^T\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1^T\mathbf{y}_1 & \mathbf{y}_1^T\mathbf{Y}_{2:N} \\ \mathbf{Y}_{2:N}^T\mathbf{y}_1 & \mathbf{Y}_{2:N}^T\mathbf{Y}_{2:N} \end{bmatrix} \tag{59}$$

and by using Schur's formula [43], we obtain

$$\det(\mathbf{Y}^T\mathbf{Y})$$

$$= \det(\mathbf{y}_1^T\mathbf{y}_1)\det\left(\mathbf{Y}_{2:N}^T\mathbf{Y}_{2:N} - \mathbf{Y}_{2:N}^T\mathbf{y}_1(\mathbf{y}_1^T\mathbf{y}_1)^{-1}\mathbf{y}_1^T\mathbf{Y}_{2:N}\right)$$

$$= \|\mathbf{y}_1\|_2^2 \det\left(\mathbf{Y}_{2:N}^T\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{2:N}\right). \tag{60}$$

Let us apply Schur's formula again to the second term of (60)

$$\det\left(\mathbf{Y}_{2:N}^T\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{2:N}\right)$$

$$= \det\left(\left(\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{2:N}\right)^T\left(\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{2:N}\right)\right)$$

$$= \|\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{y}_2\|_2^2 \det\left(\mathbf{Y}_{3:N}^T\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{P}_{\tilde{\mathbf{y}}_2}^{\perp}\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{3:N}\right)$$

$$= \|\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{y}_2\|_2^2 \det\left(\left(\mathbf{P}_{\tilde{\mathbf{y}}_2}^{\perp}\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{3:N}\right)^T\left(\mathbf{P}_{\tilde{\mathbf{y}}_2}^{\perp}\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{Y}_{3:N}\right)\right) \tag{61}$$

where $\tilde{\mathbf{y}}_2 = \mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{y}_2$. By Lemma 5, we have $\mathbf{P}_{\tilde{\mathbf{y}}_2}^{\perp}\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp} = \mathbf{P}_{\mathbf{Y}_{1:2}}^{\perp}$, and hence

$$\det(\mathbf{Y}^T\mathbf{Y}) = \|\mathbf{y}_1\|_2^2 \|\mathbf{P}_{\mathbf{Y}_{1:1}}^{\perp}\mathbf{y}_2\|_2^2$$

$$\times \det\left(\left(\mathbf{P}_{\mathbf{Y}_{1:2}}^{\perp}\mathbf{Y}_{3:N}\right)^T\left(\mathbf{P}_{\mathbf{Y}_{1:2}}^{\perp}\mathbf{Y}_{3:N}\right)\right). \tag{62}$$

By repeating the aforementioned procedure, we obtain the decomposition in (26).

For (C2), we have $\det(\mathbf{Y}^T\mathbf{Y}) = 0$. Schur's formula may no longer apply in this case, but it can be easily verified that the formula on the right-hand side of (26) is also zero. The proof of Lemma 3 is therefore complete. ∎

E) Proof of Lemma 4: Consider (30). By substituting $\mathbf{w}_1 = \overline{\mathbf{X}}\boldsymbol{\theta}$ into the objective function of (30) for $j = 1$ and by triangle inequality [43], we have

$$\max_{\boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T\boldsymbol{\theta}=1} \left\| \sum_{n=1}^L \theta_n \bar{\mathbf{x}}[n] \right\|_2 \leq \max_{\boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T\boldsymbol{\theta}=1} \sum_{n=1}^L \theta_n \|\bar{\mathbf{x}}[n]\|_2$$

$$\leq \max_{n=1,\ldots,L} \|\bar{\mathbf{x}}[n]\|_2. \tag{63}$$

It can be easily verified that the aforementioned equality is achieved if and only if $\boldsymbol{\theta} = \mathbf{e}_\ell$ for any $\ell \in \arg\max_{n=1,\ldots,L} \|\bar{\mathbf{x}}[n]\|_2$. Hence, the solution $\hat{\mathbf{w}}_1 = \bar{\mathbf{x}}[\ell]$ is arrived. The proof for (30) when $j > 1$ is the same as above and hence is omitted for brevity. ∎

F) Proof of Property 3: Let $\bar{\boldsymbol{\alpha}}_i = [\boldsymbol{\alpha}_i^T \ 1]^T$. By assumptions (A1) and (A2) and triangle inequality, we have

$$\|\bar{\mathbf{x}}[n]\|_2 = \left\| \sum_{i=1}^N s_i[n]\bar{\boldsymbol{\alpha}}_i \right\|_2 \leq \sum_{i=1}^N s_i[n]\|\bar{\boldsymbol{\alpha}}_i\|_2 \leq \max_i \|\bar{\boldsymbol{\alpha}}_i\|_2 \tag{64}$$

where the aforementioned equality holds if and only if $n = \ell$ such that $\bar{\mathbf{x}}[\ell] = \bar{\boldsymbol{\alpha}}_i$ for any $i \in \arg\max_{k=1,\ldots,N} \|\bar{\boldsymbol{\alpha}}_k\|_2$. Assume w.l.o.g. that $\|\bar{\boldsymbol{\alpha}}_1\| = \max_{i=1,\ldots,N} \|\bar{\boldsymbol{\alpha}}_i\|_2$. Hence, we can obtain $\hat{\mathbf{w}}_1 = \bar{\boldsymbol{\alpha}}_1$. Next, consider $j \geq 2$. Suppose that $\hat{\mathbf{W}}_{1:(j-1)} = [\bar{\boldsymbol{\alpha}}_1, \ldots, \bar{\boldsymbol{\alpha}}_{j-1}]$ are obtained. By $\|\mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^{\perp}\bar{\boldsymbol{\alpha}}_i\|_2 = 0$ for all $i < j$, it holds that

$$\left\| \mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^{\perp}\bar{\mathbf{x}}[n] \right\|_2 \leq \sum_{i=j}^N s_i[n] \left\| \mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^{\perp}\bar{\boldsymbol{\alpha}}_i \right\|_2$$

$$\leq \max_{i \in \{j,\ldots,N\}} \left\| \mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^{\perp}\bar{\boldsymbol{\alpha}}_i \right\|_2 \tag{65}$$

where the equality holds if and only if $n = \ell$ such that $\bar{\mathbf{x}}[\ell] = \bar{\boldsymbol{\alpha}}_i$ for any $i \in \arg\max_{k=j,\ldots,N} \|\mathbf{P}_{\hat{\mathbf{W}}_{1:(j-1)}}^{\perp}\bar{\boldsymbol{\alpha}}_k\|_2$. Using the same argument as above, we get $\hat{\mathbf{w}}_j = \bar{\boldsymbol{\alpha}}_j$ w.l.o.g. The proof of Property 3 is therefore complete. ∎

REFERENCES

[1] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*, 2nd ed.   New York: Wiley, 2004.

[2] J. A. Richards, "Analysis of remotely sensed data: The formative decades and the future," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 422–432, Mar. 2005.

[3] B. A. Campbell, *Radar Remote Sensing of Planetary Surfaces*. New York: Cambridge Univ. Press, 2002.

[4] R. N. Clark, G. A. Swayze, K. E. Livo, R. F. Kokaly, S. Sutley, J. B. Dalton, R. R. McDougal, and C. A. Gent, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS tetracorder and expert systems," *J. Geophys. Res.*, vol. 108, no. 12, pp. 5–44, Dec. 2003.

[5] G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 12–16, Jan. 2002.

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

[7] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.

[8] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," in *Proc. 2nd IEEE Workshop Hyperspectral Image Signal Process.: Evolution Remote Sens. (WHISPERS)*, Reykjavik, Iceland, Jun. 14–16, 2010, pp. 1–4.

[9] J. M. Bioucas-Dias and A. Plaza, "Hyperspectral unmixing: Geometrical, statistical, and sparse regression-based approaches," in *Proc. SPIE—Image and Signal Processing for Remote Sensing XVI*, Toulouse, France, Sep. 20, 2010, vol. 7830, p. 783 00A.

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

[11] 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.

[12] 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.

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

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

[15] 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.

[16] 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.

[17] 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.

[18] 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. Multispectral, Hyperspectral, Ultraspectral Imagery*, Aug. 2004, vol. 5425, pp. 31–41.

[19] C.-C. Wu, S. Chu, and C.-I. Chang, "Sequential N-FINDR algorithms," *Proc. SPIE*, vol. 7086, p. 708 60C, Aug. 2008.

[20] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "Variants of N-FINDR algorithm for endmember extraction," *Proc. SPIE*, vol. 7109, p. 710 90G, Oct. 2008.

[21] M. Zortea and A. Plaza, "A quantitative and comparative analysis of different implementations of N-FINDR: A fast endmember extraction algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 787–791, Oct. 2009.

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

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

[24] C.-I. Chang, C.-C. Wu, and C.-T. Tsai, "Random N-finder (N-FINDR) endmember extraction algorithms for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 641–656, Mar. 2011.

[25] W. Xiong, C.-I. Chang, and K. Kalpakis, "Fast algorithms to implement N-FINDR for hyperspectral endmember extraction," in *Proc. SPIE Conf. Algorithms Technol. Multispectral, Hyperspectral, Ultraspectral Imagery*, Orlando, FL, Apr. 5–9, 2010, vol. 7695, p. 769 51Q.

[26] W. Xiong, C.-T. Tsai, C.-W. Yang, and C.-I. Chang, "Convex cone-based endmember extraction for hyperspectral imagery," *Proc. SPIE*, vol. 7812, p. 781 20H, Apr. 13, 2010.

[27] 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, vol. 1, pp. 23–26.

[28] 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.

[29] 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.

[30] 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.

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

[32] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero, "Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4355–4368, Nov. 2009.

[33] J. M. B. Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. 1st IEEE Workshop Hyperspectral Image Signal Process.: Evolution Remote Sens.*, Grenoble, France, Aug. 26–28, 2009, pp. 1–4.

[34] 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 Int. Conf. Acoust., Speech, Signal Process.*, Dallas, TX, Mar. 14–19, 2010, pp. 1202–1205.

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

[36] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[37] 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.

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

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

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

[41] J. W. Boardman, "Automating spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. Summ. 4th Annu. JPL Airborne Geosci. Workshop*, Dec. 9–14, 1993, vol. 1, pp. 11–14.

[42] J. W. Boardman, "Geometric mixture analysis of imaging spectrometry data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Pasadena, CA, Aug. 8–12, 1994, vol. 4, pp. 2369–2371.

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

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

[45] D. P. Bertsekas, *Nonlinear Programming*.   Belmont, MA: Athena Scientific, 1999.

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

[47] H. W. Kuhn, "The Hungarian method for the assignment method," *Nav. Res. Logist. Quart.*, vol. 2, pp. 83–97, 1955.

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

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

[50] 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.

[51] G. Swayze, R. Clark, S. Sutley, and A. Gallagher, "Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada," in *Proc. Summ. 3rd Annu. JPL Airborne Geosci. Workshop*, 1992, vol. 2, pp. 47–49.

[52] G. Swayze, "The hydrothermal and structural history of the Cuprite Mining District, southwestern Nevada: An integrated geological and geophysical approach," Ph.D. dissertation, Univ. Colorado, Boulder, CO, 1997.

[53] 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 μm," U.S. Geol. Surv., Denver, CO, Open File Rep. 93-592, 1993.

[54] F. A. Kruse, J. W. Boardman, and J. F. Huntington, "Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1388–1400, Jun. 2003.

[55] 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," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 875–888, May 2010.

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

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

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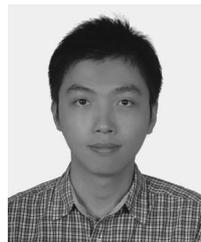
Tsung-Han Chan (S'08–M'09) received the B.S. degree from the Department of Electrical Engineering, Yuan Ze University, Taoyuan, Taiwan, in 2004, and the Ph.D. degree from the Institute of Communications Engineering, National Tsing Hua University (NTHU), Hsinchu, Taiwan, in 2009.

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

Wing-Kin Ma (S'96–M'01) received the B.Eng. (first-class honors) degree in electrical and electronic engineering from the University of Portsmouth, Portsmouth, U.K., in 1995, and the M.Phil. and Ph.D. degrees in electronic engineering from The Chinese University of Hong Kong (CUHK), Shatin, Hong Kong, in 1997 and 2001, respectively.

His Ph.D. dissertation was commended to be "of very high quality and well deserved honorary mentioning" by the Faculty of Engineering, CUHK, in 2001. He is currently an Assistant Professor with the Department of Electronic Engineering, CUHK. From 2005 to 2007, he was also an Assistant Professor with the Institute of Communications Engineering, National Tsing Hua University, Hsinchu, Taiwan, where he is still holding an adjunct position. Prior to becoming a Faculty Member, he held various research positions with McMaster University, Hamilton, ON, Canada; CUHK; and the University of Melbourne, Melbourne, Australia. His research interests are in signal processing and communications, with a recent emphasis on multiple-input–multiple-output communications, blind signal processing, and convex optimization techniques.

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

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

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

Dr. Chi has been a Technical Program Committee member for many IEEE sponsored and cosponsored workshops, symposiums, and conferences on signal processing and wireless communications, including being the Co-organizer and General Cochairman of the 2001 IEEE Workshop on Signal Processing Advances in Wireless Communications, the Cochair of the Signal Processing for Communications (SPC) Symposium, ChinaCOM 2008, and the Lead Cochair of the SPC Symposium, ChinaCOM 2009. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (5/2001–4/2006), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II (1/2006–12/2007), the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I (1/2008–12/2009), and the IEEE SIGNAL PROCESSING LETTERS (6/2006–5/2010), a member of the Editorial Board of the European Association for Signal Processing (EURASIP) Signal Processing Journal (6/2005–5/2008), and an Editor (7/2003–12/2005) as well as a Guest Editor (2006) of the EURASIP Journal on Applied Signal Processing. He was a member of the 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.