ROBUST ENDMEMBER EXTRACTION USING WORST-CASE SIMPLEX VOLUME MAXIMIZATION

Tsung-Han Chan[†], *Wing-Kin Ma* [*], *ArulMurugan Ambikapathi*[†], *and Chong-Yung Chi*[†]

[†]Inst. Communications Eng. & Dept. Electrical Eng., National Tsing Hua Univ., Hsinchu, Taiwan
[*]Dept. Electronic Eng., Chinese Univ. of Hong Kong, Shatin, N.T., Hong Kong
E-mail: `{thchan,wkma,aareul}@ieee.org;cychi@ee.nthu.edu.tw`

ABSTRACT

Winter's maximum-volume simplex approach is an efficient and representative endmember extraction approach, as evidenced by the fact that N-FINDR, one of the most widely used class of endmember extraction algorithms, employs simplex volume maximization as its criterion. In this work, we consider a robust generalization of Winter's maximum-volume simplex criterion for the noisy scenario. Our development is based on an observation that the presence of noise would tend to expand the observed data cloud geometrically. The proposed robust Winter criterion is based on a max-min or worst-case approach, where we attempt to counteract the data cloud expansion effects by using a shrunk simplex volume as the metric to maximize. The proposed criterion is implemented by a combination of alternating optimization and projected subgradients. Some simulation results are presented to demonstrate the performance advantages of the proposed robust algorithm.

Index Terms— Endmember Extraction, Simplex Volume Maximization, Worse-case Optimization, Alternating Optimization, Projected Subgradients

1. INTRODUCTION

Hyperspectral endmember extraction has recently received considerable attention due to its various applications such as space object detection and planet exploration, as well as environmental monitoring and military surveillance on Earth [1]. Endmember extraction methods may be classified by the criteria employed. A representative one is Winter's maximum-volume simplex criterion [2], which led to one of the most widely used class of endmember extraction algorithms, namely, N-FINDR [3–5]. The belief in Winter's original work is to search for the set of "purest" pixel vectors from the data cloud through a simplex volume maximization attempt. A key advantage of Winter's criterion is that it is structurally simpler than Craig's minimum-volume enclosing simplex criterion [6, 7] from an optimization viewpoint; this is also a reason behind why N-FINDR algorithms are often simple to implement.

In this paper, we propose a robust generalization of Winter's maximum-volume simplex criterion for the noisy scenario. Recently, there has been growing interest in study of endmember extraction for the noisy scenario; e.g. the optimal Bayesian framework [8] in which noise is explicitly accounted for, simplex identification by split augmented Lagrangian (SISAL) [6] in which soft constraints are used to handle outlier pixels caused by noise under Craig's minimum-volume criterion, and more [9, 10]. We formulate a robust Winter criterion based on a max-min, or worst-case, optimization formulation. The underlying insight is to attempt to counteract the size expansion effects of the data cloud under noise perturbations. To practically realize the robust Winter criterion, we design an optimization algorithm that combines alternating optimization and subgradient techniques. Some simulation results will be provided to demonstrate the merits of the proposed algorithm.

Notations: $\mathbf{1}_N$, $\succeq$, and $\|\cdot\|_2$ represent $N \times 1$ all-one vector, componentwise inequality, and Euclidean norm, respectively.

2. WINTER'S ENDMEMBER EXTRACTION PROBLEM

Assuming that the incident solar radiation gets reflected from the Earth surface through a single bounce and the materials are distinct, each pixel vector of the hyperspectral data cube can be represented by a linear mixing model

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

$$= \sum_{i=1}^{N} s_i[n]\mathbf{a}_i + \mathbf{w}[n],\ n = 1, \ldots, L, \tag{2}$$

where $\mathbf{x}[n] = [\ x_1[n], \ldots, x_M[n]\ ]^T \in \mathbb{R}^M$ denotes the nth observed pixel vector which is constituted by M spectral bands, $\mathbf{w}[n] \in \mathbb{R}^M$ represents noise, $\mathbf{A} = [\ \mathbf{a}_1, \ldots, \mathbf{a}_N\ ] \in \mathbb{R}^{M\times N}$ denotes the signature matrix whose ith column vector $\mathbf{a}_i$ is the ith endmember, N is the number of endmembers, $\mathbf{s}[n] = [\ s_1[n], \ldots, s_N[n]\ ]^T \in \mathbb{R}^N$ denotes the abundance vector associated with the pixel $\mathbf{x}[n]$, and L is the total number of observed pixel vectors. Some general assumptions are as follows [1]:

- (A1) (Non-negativity) $s_i[n] \geq 0$ for all i and n.
- (A2) (Full-additivity) $\sum_{i=1}^{N} s_i[n] = 1$ for all n.
- (A3) $\min\{L, M\} \geq N$ and $\mathbf{a}_1, \ldots, \mathbf{a}_N$ are linearly independent.
- (A4) There exist pure pixels, i.e., $\mathbf{x}[\ell_i] = \mathbf{a}_i$, $i = 1, \ldots, N$ for some index set $\{\ell_1, \ldots, \ell_N\}$.

We consider dimension reduction of the observed pixel vectors by the following procedure [7]:

$$\tilde{\mathbf{x}}[n] \triangleq \mathbf{C}^T(\mathbf{x}[n] - \mathbf{d}),\ n = 1, \ldots, L. \tag{3}$$

Here, $\tilde{\mathbf{x}}[n] \in \mathbb{R}^{N-1}$ is the nth dimension-reduced pixel vector, $\mathbf{d} = \frac{1}{L}\sum_{n=1}^{L} \mathbf{x}[n]$, and $\mathbf{C} = [\ \boldsymbol{q}_1(\mathbf{H}\mathbf{H}^T), \ldots, \boldsymbol{q}_{N-1}(\mathbf{H}\mathbf{H}^T)\ ]$ where $\mathbf{H} = [\ \mathbf{x}[1] - \mathbf{d}, \ldots, \mathbf{x}[L] - \mathbf{d}\ ] \in \mathbb{R}^{M\times L}$ and $\boldsymbol{q}_i(\mathbf{H}\mathbf{H}^T)$ is the ith principal unit-norm eigenvector of $\mathbf{H}\mathbf{H}^T$. Note that the dimension reduction in (3) is lossless in the absence of noise, either from a principal component analysis perspective or from a convex geometry perspective [7]. From (2)-(3), we obtain the following dimension-reduced signal model:

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n]\boldsymbol{\alpha}_i + \tilde{\mathbf{w}}[n],\ n = 1, \ldots, L, \tag{4}$$

This work was supported by the National Science Council (R.O.C.) under Grants NSC 99-2221-E-007-003-MY3, and by a General Research Fund of Hong Kong Research Grant Council (Project No. CUHK415509).

where $\boldsymbol{\alpha}_i = \mathbf{C}^T(\mathbf{a}_i - \mathbf{d}) \in \mathbb{R}^{N-1}$ is the dimension-reduced endmember corresponding to $\mathbf{a}_i$, and $\tilde{\mathbf{w}}[n] = \mathbf{C}^T\mathbf{w}[n]$ denotes a dimension-reduced noise vector.

The goal of endmember extraction is to estimate $\mathbf{a}_1, \ldots, \mathbf{a}_N$ from the observed pixel vectors $\mathbf{x}[n],\ n = 1, \ldots, L$, which is also equivalent to estimating $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$ from the dimension-reduced pixel vectors $\tilde{\mathbf{x}}[n],\ n = 1, \ldots, L$. We consider Winter's endmember extraction belief [2]. Based on the dimension-reduced model (4), we formulate Winter's belief as an optimization criterion:

$$\boxed{\max_{\substack{\boldsymbol{\nu}_i \in \mathcal{F},\\ i=1,\ldots,N}} \quad |\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))|,} \tag{5}$$

where

$$\mathcal{F} = \{\boldsymbol{\nu} \in \mathbb{R}^{N-1} \mid \boldsymbol{\nu} = \widetilde{\mathbf{X}}\boldsymbol{\theta},\ \boldsymbol{\theta} \succeq \mathbf{0},\ \mathbf{1}_L^T\boldsymbol{\theta} = 1\} \tag{6}$$

is the convex hull of $\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]$, $\widetilde{\mathbf{X}} = [\ \tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\ ]$, and

$$\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N) = \begin{bmatrix} \boldsymbol{\nu}_1 & \cdots & \boldsymbol{\nu}_N \\ 1 & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{N \times N}. \tag{7}$$

The Winter problem (5) seeks to find an estimate of the endmembers, denoted by the N-tuple $(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$, from the pixel-constructed convex hull $\mathcal{F}$ such that the respective simplex volume is the largest.

In [11], we will study the Winter problem in (5) using an optimization perspective. In particular, its connections to N-FINDR and vertex component analysis (VCA) [12] will be shown. In this paper, our interest is in the noisy scenario. Assuming existence of pure pixels [cf., (A4)], one would anticipate that the Winter problem can lead to perfect recovery of the ground-truth endmembers in the absence of noise; this in fact can be shown to be true even by analysis [11]. However, the Winter problem may be sensitive to noise. An illustration is given in Figure 1 to get an idea of how noise may affect the data geometry (for $N = 3$) and the subsequent Winter's endmember estimates. We assume that the noise perturbations are random and isotropic; e.g., i.i.d. Gaussian noise. The noise effects tend to expand the size of the pixel-constructed convex hull $\mathcal{F}$, and subsequently the endmember estimates $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3$ are away from the ground truth $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$.

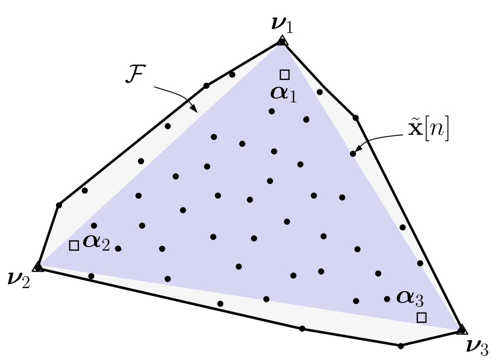

Fig. 1. Data geometry of the Winter problem for $N = 3$.

3. ROBUST GENERALIZATION OF WINTER'S PROBLEM

The noise robust generalization of the Winter problem is developed in this section.

3.1. Robust Winter Formulation by Worst-Case Optimization

As discussed above, noise perturbations may result in expansion of the size of $\mathcal{F}$, as compared to the simplex volume of the ground truth endmembers. Our intuition is to reduce the simplex volume of the endmembers estimate $(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ to a certain extent, thereby attempting to neutralize the size expansion of $\mathcal{F}$. The idea is illustrated in Figure 2. We see that by introducing some reduction of the simplex volume, which physically results in some backoff of the endmember estimates from the boundary of $\mathcal{F}$, the estimate $(\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \boldsymbol{\nu}_3)$ may be brought closer to the ground truth $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$.

Let us formulate the above described robust Winter belief properly. From the original Winter problem (5), a noise robust generalization of the Winter problem is formulated as follows:

$$\boxed{\max_{\substack{\mathbf{v}_i \in \mathcal{F},\\ i=1,\ldots,N}} \quad \left\{ \min_{\substack{\|\mathbf{u}_i\| \le r,\\ i=1,\ldots,N}} \left|\det(\mathbf{v}_1 - \mathbf{u}_1, \ldots, \mathbf{v}_N - \mathbf{u}_N)\right| \right\},} \tag{8}$$

where $\mathbf{u}_i \in \mathbb{R}^{N-1},\ i = 1, \ldots, N$, are error vectors lying in a norm ball $\{\mathbf{u} \in \mathbb{R}^{N-1} \mid \|\mathbf{u}\|_2 \le r\}$, with a given radius r. By letting $(\mathbf{v}_1^\star, \ldots, \mathbf{v}_N^\star, \mathbf{u}_1^\star, \ldots, \mathbf{u}_N^\star)$ be the outer-inner solution of (8), the robust Winter endmember estimates are $\boldsymbol{\nu}_i^\star = \mathbf{v}_i^\star - \mathbf{u}_i^\star$, $i = 1, \ldots, N$.

Our robust Winter formulation (8) follows an approach that is generally called max-min or worst-case optimization in the context of robust optimization. The error vectors $\mathbf{u}_1, \ldots, \mathbf{u}_N$ are incorporated to introduce backoff of the estimates $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$, and they are controlled by the inner minimization of (8), which seeks to yield the smallest possible (or worst-case) simplex volume. In that inner minimization, the prescribed parameter r governs how much backoff of $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ is allowed. In other words, the value r 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.

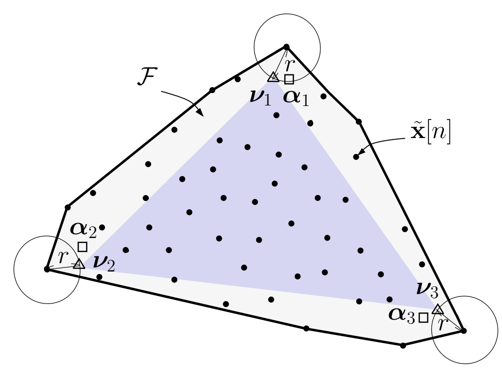

Fig. 2. Data geometry of the worst-case Winter problem.

3.2. Worst-Case Alternating Volume Maximization

Our next endeavor is to develop an optimization algorithm for the robust Winter problem (8). For this purpose, we use the change of variables $\boldsymbol{\nu}_i = \widetilde{\mathbf{X}}\boldsymbol{\theta}_i$, $i = 1, \ldots, N$, to reexpress Problem (8) as

$$\max_{\substack{\boldsymbol{\theta}_i \in \mathcal{S},\\ i=1,\ldots,N}} \quad \varphi(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N), \tag{9}$$

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

$$\mathcal{S} = \{\boldsymbol{\theta} \in \mathbb{R}^L \mid \boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_L^T\boldsymbol{\theta} = 1\}, \tag{10}$$

and $\varphi(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N)$ captures the inner minimization of the problem:

$$\varphi(\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N)=\min_{\substack{\|\mathbf{u}_i\|\leq r,\\ i=1,\ldots,N}}\ \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1-\mathbf{u}_1,\ldots,\widetilde{\mathbf{X}}\boldsymbol{\theta}_N-\mathbf{u}_N)). \tag{11}$$

Note that in (11), we have removed the absolute term [cf., (8)]; this is without loss of generality, as we have shown in [11].

There are two obstacles that make Problem (9) a difficult optimization problem. First, the determinant in the objective function is nonconcave. Second, dealing with max-min optimization may not be easy. We handle the first issue in a practical way by employing alternating optimization: Let $\widehat{\boldsymbol{\theta}}_1,\ldots,\widehat{\boldsymbol{\theta}}_N$ denote the endmember estimates. In alternating optimization, we update $\widehat{\boldsymbol{\theta}}_i$'s by performing the following partial maximization of Problem (9)

$$\widehat{\boldsymbol{\theta}}_j=\arg\max_{\boldsymbol{\theta}_i\in\mathcal{S}}\ \varphi(\boldsymbol{\theta}_j,\widehat{\boldsymbol{\Theta}}_j),\qquad j=1,\ldots,N, \tag{12}$$

where $\widehat{\boldsymbol{\Theta}}_j=[\widehat{\boldsymbol{\theta}}_1,\ldots,\widehat{\boldsymbol{\theta}}_{j-1},\widehat{\boldsymbol{\theta}}_{j+1},\ldots,\widehat{\boldsymbol{\theta}}_N]$. The update process (12) is conducted cyclically until some stopping rule is satisfied. We will call this method *worse-case alternating volume maximization (WAVMAX)* in the sequel.

The key issue of WAVMAX lies in handling (12). By applying a cofactor expansion to $\det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1-\mathbf{u}_1,\ldots,\widetilde{\mathbf{X}}\boldsymbol{\theta}_N-\mathbf{u}_N))$, we rewrite the objective function of (12) as

$$\varphi(\boldsymbol{\theta}_j,\boldsymbol{\Theta}_j)=\min_{\substack{\|\mathbf{u}_i\|\leq r,\\ i=1,\ldots,N}}\ \mathbf{b}_j^T(\mathbf{U})(\widetilde{\mathbf{X}}\boldsymbol{\theta}_j-\mathbf{u}_j)+h_j(\mathbf{U}). \tag{13}$$

Here, $\boldsymbol{\Theta}_j$ is defined in the same way as $\widehat{\boldsymbol{\Theta}}_j$, $\mathbf{U}=[\mathbf{u}_1,\ldots,\mathbf{u}_N]$, $\mathbf{b}_j(\mathbf{U})=[(-1)^{i+j}\det(\boldsymbol{\mathcal{Q}}_{ij})]_{i=1}^{N-1}$ where $\boldsymbol{\mathcal{Q}}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\boldsymbol{\theta}_1-\mathbf{u}_1,\ldots,\widetilde{\mathbf{X}}\boldsymbol{\theta}_N-\mathbf{u}_N)$ with the ith row and the jth column removed, and $h_j(\mathbf{U})=(-1)^{N+j}\det(\boldsymbol{\mathcal{Q}}_{Nj})$. It can be shown that (13) is a concave function, and it follows that Problem (12) is a convex optimization problem, a merit of considering WAVMAX.

3.3. Projected Subgradient Method for WAVMAX

While the partial maximization problems of WAVMAX in (12) are structurally simpler than the original robust Winter problem in (9), the max-min problem structure of (12) remains an obstacle. Fortunately, the fact that (12) is a convex problem enables us to deal with it using the subgradient approach [13]. The latter is a key approach to solving non-differentiable convex problems, and Problem (12) fits into that scope.

To be precise, we employ the projected subgradient method to deal with Problem (12). The basic idea is to generate a sequence of points according to the following iteration

$$\boldsymbol{\theta}_j^{(k+1)}=\left\{\boldsymbol{\theta}_j^{(k)}-\gamma_k\mathbf{g}^{(k)}\right\}_{\mathcal{S}}, \tag{14}$$

where $\mathbf{g}^{(k)}$ is a subgradient of $-\varphi(\boldsymbol{\theta}_j,\widehat{\boldsymbol{\Theta}}_j)$ 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

$$\varphi_{best}^{(k+1)}=\max\left\{\varphi_{best}^{(k)},\varphi(\boldsymbol{\theta}_j^{(k+1)},\widehat{\boldsymbol{\Theta}}_j)\right\}, \tag{15}$$

and update $\widehat{\boldsymbol{\theta}}_j=\boldsymbol{\theta}_j^{(k+1)}$ if $\varphi_{best}^{(k+1)}=\varphi(\boldsymbol{\theta}_j^{(k+1)},\widehat{\boldsymbol{\Theta}}_j)$. 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 $\gamma>0$ [13].

The projection onto the simplex $\{\boldsymbol{x}\}_{\mathcal{S}}$ can be efficiently implemented by a waterfilling-type algorithm [14]. The subgradients of $-\varphi(\boldsymbol{\theta}_j,\widehat{\boldsymbol{\Theta}}_j)$ can be computed as follows. By Danskin's theorem [13], a subgradient of $-\varphi(\boldsymbol{\theta}_j,\widehat{\boldsymbol{\Theta}}_j)$ at $\boldsymbol{\theta}_j$ is given by

$$\mathbf{g}=-\widetilde{\mathbf{X}}^T\mathbf{b}_j(\mathbf{U}^\star), \tag{16}$$

where $\mathbf{U}^\star=[\mathbf{u}_1^\star,\ldots,\mathbf{u}_N^\star]$ is the optimal solution of (12); i.e.,

$$(\mathbf{u}_1^\star,\ldots,\mathbf{u}_N^\star)=\arg\min_{\substack{\|\mathbf{u}_i\|_2\leq r,\\ i=1,\ldots,N}}\ \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1-\mathbf{u}_1,\ldots,\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N-\mathbf{u}_N)). \tag{17}$$

The above problem is nonconvex, but can be approximated by alternating optimization. To be concise, here we only describe the key steps. The partial minimization problem of (17) w.r.t. $\mathbf{u}_j$ can be expressed as

$$\min_{\|\mathbf{u}_j\|_2\leq r}\ \mathbf{k}_j^T(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_j-\mathbf{u}_j)+(-1)^{N+j}\det(\boldsymbol{\mathcal{U}}_{Nj}), \tag{18}$$

where $\mathbf{k}_j=[(-1)^{i+j}\det(\boldsymbol{\mathcal{U}}_{ij})]_{i=1}^{N-1}$ and $\boldsymbol{\mathcal{U}}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1-\widehat{\mathbf{u}}_1,\ldots,\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N-\widehat{\mathbf{u}}_N)$ with the ith row and the jth column removed. By Cauchy-Schwarz inequality, it can be easily shown that the solution to (18) is uniquely given by $\widehat{\mathbf{u}}_j=r\mathbf{k}_j/\|\mathbf{k}_j\|_2$. The pseudo code of alternating optimization for (17) is given in Table 1.

The development of the WAVMAX method is complete. A summary of WAVMAX in pseudo code form is given in Table 2.

Table 1. A Summary of Alternating Optimization for Handling (17)

Given	a convergence tolerance $\varepsilon>0$, an error tolerance r, the dimension reduced data matrix $\widetilde{\mathbf{X}}$, the parameters $\widehat{\boldsymbol{\theta}}_1,\ldots,\widehat{\boldsymbol{\theta}}_N$, and the number of endmembers N.
Step 1.	initialize $(\widehat{\mathbf{u}}_1,\ldots,\widehat{\mathbf{u}}_N)$ with all zero vectors, set $j=1$, and calculate $\varrho=\det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1-\widehat{\mathbf{u}}_1,\ldots,\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N-\widehat{\mathbf{u}}_N))$.
Step 2.	compute $\mathbf{k}_j=[(-1)^{i+j}\det(\boldsymbol{\mathcal{U}}_{ij})]_{i=1}^{N-1}$ in which $\boldsymbol{\mathcal{U}}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1-\widehat{\mathbf{u}}_1,\ldots,\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N-\widehat{\mathbf{u}}_N)$ with the ith row and the jth column removed.
Step 3.	obtain $\widehat{\mathbf{u}}_j=r\mathbf{k}_j/\|\mathbf{k}_j\|_2$.
Step 4.	if (j modulo N) $\neq 0$, then $j:=j+1$ and go to **Step 2**, else compute $\bar{\varrho}=\det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1-\widehat{\mathbf{u}}_1,\ldots,\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N-\widehat{\mathbf{u}}_N))$,
Step 5.	if $\lvert\bar{\varrho}-\varrho\rvert/\varrho>\varepsilon$, then set $\varrho:=\bar{\varrho}$, $j:=1$, and go to **Step 2**, else output $(\widehat{\mathbf{u}}_1,\ldots,\widehat{\mathbf{u}}_N)$ as an approximate solution of (17).

4. SIMULATIONS AND CONCLUSION

A Monte Carlo simulation of one hundred independent runs is presented to demonstrate the performance of the proposed WAVMAX algorithm. Five existing endmember extraction algorithms, I-N-FINDR [3], SQ-N-FINDR [4], SC-N-FINDR [4], SGA [5], and VCA [12], were tested for comparison. In each run, 8 endmember signatures with 224 bands selected from the U.S. geological survey (USGS) library, abundances generated following Dirichlet distribution [12], and additive zero-mean white Gaussian noise were used to generate the 1000 noisy observed pixel vectors with various signal-to-noise ratios (SNRs), where $\mathrm{SNR}=\sum_{n=1}^{L}\|\mathbf{As}[n]\|_2^2/(\sigma^2 ML)$ in which σ^2 is the noise variance. The root-mean-square (rms) spectral angle distance, denoted by ϕ, was used as the performance measure [12]. The computation time T (in secs) of each method (implemented on Matlab) running in a computer equipped with Core i7-930 CPU 2.80GHz and 12GB memory is used as the complexity measure.

Table 2. A Summary of WAVMAX Algorithm

Given a convergence tolerance $\varepsilon > 0$, 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 $(\widehat{\boldsymbol{\theta}}_1, \ldots, \widehat{\boldsymbol{\theta}}_N)$ by alternating volume maximization (AVMAX) [11] and obtain $\widehat{\mathbf{U}}$ by Table 1.

Step 2. *Alternating optimization over* $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$*:* set $j := 1$ and compute $\varrho = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1 - \widehat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N - \widehat{\mathbf{u}}_N))$.

Step 3. *Projected subgradient iterations for* $\boldsymbol{\theta}_j$

- **3.1.** set $k = 1$ and $\varphi_{best} = 0$.
- **3.2.** calculate $\mathbf{b}(\widehat{\mathbf{U}}) = [(-1)^{i+j}\det(\boldsymbol{\mathcal{Q}}_{ij})]_{i=1}^{N-1}$ where $\boldsymbol{\mathcal{Q}}_{ij}$ is a submatrix of $\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1 - \widehat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N - \widehat{\mathbf{u}}_N)$ with the ith row and the jth column removed.
- **3.3.** update

$$\boldsymbol{\theta}_j := \left\{ \boldsymbol{\theta}_j + \gamma_k \widetilde{\mathbf{X}}^T \mathbf{b}(\widehat{\mathbf{U}}) \right\}_{\mathcal{S}},$$

  where $\gamma_k = \gamma/\sqrt{k}$ and $\{\boldsymbol{x}\}_{\mathcal{S}}$ is the projection $\boldsymbol{x}$ onto the simplex using a water-filling method [14].
- **3.4.** update $\widehat{\mathbf{U}}$ by Table 1 with the given $(\boldsymbol{\theta}_j, \widehat{\boldsymbol{\Theta}}_j)$.
- **3.5.** update $\widehat{\boldsymbol{\theta}}_j := \boldsymbol{\theta}_j$ if $\varphi(\boldsymbol{\theta}_j, \widehat{\boldsymbol{\Theta}}_j) > \varphi_{best}$ and

$$\varphi_{best} := \max \left\{ \varphi_{best}, \varphi(\boldsymbol{\theta}_j, \widehat{\boldsymbol{\Theta}}_j) \right\}.$$

- **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 compute $\widehat{\mathbf{U}}$ by Table 1 and $\bar{\varrho} = \det(\boldsymbol{\Delta}(\widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_1 - \widehat{\mathbf{u}}_1, \ldots, \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_N - \widehat{\mathbf{u}}_N))$.

Step 5. if $|\bar{\varrho} - \varrho|/\varrho > \varepsilon$, then set $\varrho := \bar{\varrho}$, $j := 1$, and go to **Step 3**, otherwise, $\widehat{\boldsymbol{\nu}}_j = \widetilde{\mathbf{X}}\widehat{\boldsymbol{\theta}}_j - \widehat{\mathbf{u}}_j$ for all j.

Step 6. output $(\widehat{\boldsymbol{\nu}}_1, \ldots, \widehat{\boldsymbol{\nu}}_N)$ as the robust Winter's estimates.

Some parameter settings are as follows: The convergence tolerance for the WAVMAX algorithm and its associated sub-algorithm in Table I was set to $\varepsilon = 5 \times 10^{-5}$. We also set the maximum number of subgradient iterations $K = 5$ and step size $\gamma = 1$ in the WAVMAX algorithm. The error tolerance r for the proposed WAVMAX algorithm was set to $r = 1.3\sigma$.

Table 3 shows the average ϕ and T of all the endmember extraction algorithms over $\mathrm{SNR} = 5, 15, ..., 45$ (dB). Herein, each boldface value denotes the best performance among the tested algorithms for a specific SNR. One can see from Table 3 that WAVMAX yields better endmember extraction performance than the other algorithms for $\mathrm{SNR} \leq 25$dB. The performance gap is particularly noticeable for low SNRs. For high SNRs such as $\mathrm{SNR} \geq 35$dB, all the algorithms are quite on par in performance, with WAVMAX and SGA being the best two. Table 3 also shows that the computation time required by WAVMAX is higher. Future work should consider complexity reduction of WAVMAX. Also, the performance of WAVMAX on real data, such as the Cuprite scene dataset, should be investigated.

In conclusion, we have established a robust Winter endmember extraction approach for the noisy scenario. The approach is based on max-min or worst-case robust optimization. The resultant algorithm, WAVMAX, has been numerically demonstrated to yield robustness against noise and improved endmember extraction performance compared to some other algorithms.

Table 3. Performance comparison of average ϕ (degrees) and average T (secs) 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
	T	0.12	0.08	0.08	**0.06**	**0.05**
SGA	ϕ	13.92	3.34	0.96	**0.29**	**0.09**
	T	0.11	0.09	0.08	0.08	0.08
I-N-FINDR	ϕ	14.43	3.50	1.07	0.32	0.10
	T	0.21	0.20	0.27	0.17	0.21
SQ-N-FINDR	ϕ	14.17	3.49	1.08	0.32	0.10
	T	0.19	0.15	0.13	0.12	0.12
SC-N-FINDR	ϕ	14.59	3.74	1.18	0.32	0.11
	T	**0.08**	**0.07**	**0.06**	**0.06**	0.06
WAVMAX	ϕ	**12.84**	**3.08**	**0.92**	0.30	**0.09**
	T	134.88	72.21	69.23	55.04	53.48

5. REFERENCES

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

[2] 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.

[3] 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 and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, Aug. 2004, vol. 5425, pp. 31–41.

[4] Chao-Cheng Wu, Shihyu Chu, and Chein-I Chang, "Sequential N-FINDR algorithms," *Proc. of SPIE*, vol. 7086, Aug. 2008.

[5] 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.

[6] J. M. B. Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. First IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Grenoble, France, Aug. 26-28, 2009.

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

[8] 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 Processing*, vol. 57, no. 11, pp. 4355–4368.

[9] 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 International Conference on Acoustics, Speech, and Signal Processing*, Dallas, Texas, USA, Mar. 14-19, 2010, pp. 1202–1205.

[10] Alina Zare and Paul Gader, "L1-endmembers: A robust endmember detection and spectral unmixing algorithm," in *Proc. of SPIE*, Dec. 9-14, 2010, vol. 7695, p. 76951L.

[11] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "An optimization perspective on Winter's endmember extraction belief," to appear in *IEEE International Geoscience and Remote Sensing Symposium*, Sendai, Japan, Aug. 1-5, 2011.

[12] 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.

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

[14] A. Zymnis, S.-J. Kim, J. Skaf, M. Parente, and S. Boyd, "Hyperspectral image unmixing via alternating projected subgradients," in *Proc. 41st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 4-7, 2007.