

FAST ALGORITHMS FOR ROBUST HYPERSPECTRAL ENDMEMBER EXTRACTION BASED ON WORST-CASE SIMPLEX VOLUME MAXIMIZATION

Tsung-Han Chan[†], Ji-Yuan Liou[†], ArulMurugan Ambikapathi[†], Wing-Kin Ma^{*} and Chong-Yung Chi[†]

[†]Inst. Commun. Eng., National Tsinghua Univ.
Hsinchu, Taiwan

E-mail: {thchan, aareul}@ieee.org

^{*}Dept. Electronic Eng., Chinese Univ. Hong Kong
Shatin, N.T., Hong Kong

E-mail: wkma@ieee.org

ABSTRACT

Hyperspectral endmember extraction (EE) is to estimate endmember signatures (or material spectra) from the hyperspectral data of an unexplored area for analyzing the materials and their composition therein. However, the presence of noise in the data posts a serious problem for EE. Recently, robustness against noise has been taken into account in the design of EE algorithms. The robust maximum-volume simplex criterion [1] has been shown to yield performance improvement in the noisy scenario, but its real applicability is limited by its high implementation complexity. In this paper, we propose two fast algorithms to approximate this robust criterion [1], which turns out to deal with a set of partial max-min optimization problems in alternating manner and successive manner, respectively. Some Monte Carlo simulations demonstrate the superior computational efficiency and efficacy of the proposed robust algorithms in the noisy scenario over the robust algorithm in [1] and some benchmark EE algorithms.

Index Terms— Hyperspectral images, Robust endmember extraction, Simplex volume maximization, Fast algorithms

1. INTRODUCTION

Hyperspectral endmember extraction (EE) has been applied in many fields, such as space object detection, environmental monitoring and military surveillance [2]. However, the presence of noise in hyperspectral data is inevitable, and may seriously degrade the performance of EE algorithms. Existing efforts that account for noise effects include joint Bayesian algorithm (JBA) [3], split augmented Lagrangian (SISAL) [4], robust minimum volume enclosing algorithm (RMVES) [5], and others [6], but none of them are based on popular Winter's maximum-volume simplex criterion [7–9].

Very recently, we have proposed a robust generalization of Winter's criterion in the noisy scenario [1], and formulated the robust Winter criterion as a worst-case simplex volume maximization problem. An algorithm called worst-case alternating volume maximization (WAVMAX) that practically realizes the robust Winter criterion has also been proposed [1], but it is quite computationally expensive for large data sizes. In this work, we develop two computationally efficient algorithms to implement the robust Winter criterion reported in [1]. The proposed algorithms, named alternating decoupled volume max-min (ADVMM) and successive decoupled volume max-min (SDVMM), deal with the worst-case simplex volume maximization problem by alternating optimization and by successive optimization, respectively. These optimization principles have been exploited by predecessors of ADVMM and SDVMM; i.e., alternating

volume maximization (AVMAX) and successive volume maximization (SVMAX) [1] that fulfill the original Winter's criterion. Some simulations are presented to demonstrate the efficiency and efficacy of the proposed fast robust algorithms.

Notations: \mathbb{R}^N and $\mathbb{R}^{M \times N}$ denote set of real $N \times 1$ vectors and set of real $M \times N$ matrices, respectively; $\mathbf{1}_N$, \mathbf{I}_N , and \mathbf{e}_i represent $N \times 1$ all-one vector, $N \times N$ identity matrix, and unit column vector with the i th entry equal to 1, respectively; " \succeq ", " $\|\cdot\|$ ", and " \setminus " stand for componentwise inequality, Euclidean norm, and set difference, respectively; $\det(\mathbf{X})$ and \mathbf{X}^\dagger denote the determinant and pseudo-inverse of the matrix \mathbf{X} , respectively; $[\mathbf{x}]_{1:i}$ is an $i \times 1$ column vector formed by the first i elements in \mathbf{x} .

2. PROBLEM STATEMENT AND ASSUMPTIONS

Consider a $M \times N$ linear spectral mixing model [2]:

$$\mathbf{y}[n] = \mathbf{A}\mathbf{s}[n] + \mathbf{w}[n], \quad n = 1, \dots, L, \quad (1)$$

where $\mathbf{y}[n] = [y_1[n], \dots, y_M[n]]^T \in \mathbb{R}^M$ is the n th observed noisy pixel vector comprising M spectral bands, $\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_N] \in \mathbb{R}^{M \times N}$ denotes the signature matrix whose i th column vector \mathbf{a}_i is the i th endmember signature, $\mathbf{s}[n] = [s_1[n], \dots, s_N[n]]^T \in \mathbb{R}^N$ is the n th abundance vector comprising N fractional abundances, L is the total number of pixels, and $\mathbf{w}[n] = [w_1[n], \dots, w_M[n]]^T \in \mathbb{R}^M$ is the zero-mean random isotropic noise vector with covariance matrix $\sigma^2 \mathbf{I}_M$ where σ^2 is the noise variance.

Endmember extraction problem is to estimate $\mathbf{a}_1, \dots, \mathbf{a}_N$ from the given observed pixel vectors $\mathbf{y}[1], \dots, \mathbf{y}[L]$ with prior knowledge of the number of endmembers N , under the following general assumptions [2]: (A1) $s_i[n] \geq 0$ for all i and n ; (A2) $\sum_{i=1}^N s_i[n] = 1$ for all n ; (A3) $\min\{L, M\} \geq N$ and \mathbf{A} is of full column rank; (A4) (Pure pixel assumption) there exists at least a set of indices $\{\ell_1, \ell_2, \dots, \ell_N\}$ such that $\mathbf{x}[\ell_i] = \mathbf{a}_i$ for $i = 1, \dots, N$.

As a common preprocessing step in hyperspectral image analysis [2], we obtain the dimension reduced observed pixel vectors $\tilde{\mathbf{y}}[n] \in \mathbb{R}^{N-1}$ by the following affine transformation [1]:

$$\tilde{\mathbf{y}}[n] \triangleq \mathbf{C}^T (\mathbf{y}[n] - \mathbf{d}), \quad n = 1, \dots, L, \quad (2)$$

where $\mathbf{d} = \frac{1}{L} \sum_{n=1}^L \mathbf{y}[n]$, $\mathbf{C} = [\mathbf{q}_1(\mathbf{U}\mathbf{U}^T), \dots, \mathbf{q}_{N-1}(\mathbf{U}\mathbf{U}^T)]$, $\mathbf{U} = [\mathbf{y}[1] - \mathbf{d}, \dots, \mathbf{y}[L] - \mathbf{d}] \in \mathbb{R}^{M \times L}$, and $\mathbf{q}_i(\mathbf{U}\mathbf{U}^T)$ denotes the unit-norm eigenvector of $\mathbf{U}\mathbf{U}^T$ associated with the i th principal eigenvalue. Substituting (1) into (2) yields

$$\tilde{\mathbf{y}}[n] = \sum_{i=1}^N s_i[n] \boldsymbol{\alpha}_i + \tilde{\mathbf{w}}[n], \quad n = 1, \dots, L. \quad (3)$$

where $\boldsymbol{\alpha}_i = \mathbf{C}^T (\mathbf{a}_i - \mathbf{d}) \in \mathbb{R}^{N-1}$ is the i th dimension-reduced endmember and $\tilde{\mathbf{w}}[n] \triangleq \mathbf{C}^T \mathbf{w}[n]$ is still the random isotropic noise

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

vector due to $\mathbf{C}^T \mathbf{C} = \mathbf{I}_{N-1}$. After dimension reduction, the aim now is to estimate $\alpha_1, \dots, \alpha_N$ from the dimension reduced observed pixel vectors $\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]$. Once $\alpha_1, \dots, \alpha_N$ are obtained, one can simply recover the endmember estimates by the affine transformation $\mathbf{a}_i = \mathbf{C}\alpha_i + \mathbf{d}$, $i = 1, \dots, N$.

3. BRIEF REVIEW OF WORST-CASE WINTER'S ENDMEMBER EXTRACTION PROBLEM

Winter proposed an EE criterion which states that in the presence of pure pixels, the true endmembers can be determined by finding the vertices of the maximum-volume simplex inside the data cloud $\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]$ [9]. However, a fact is mentioned in [1] that in the presence of additive noise, the simplex volume yielded by Winter's criterion may be larger than that of the true simplex. In other words, the endmember estimates obtained by Winter's criterion may be away from the true endmembers when the observed data are corrupted by noise. To mitigate such effects, we have proposed an idea [1] to pull back the estimates obtained by Winter's criterion by a suitable margin such that (ν_1, \dots, ν_N) are closer to the true endmembers $(\alpha_1, \dots, \alpha_N)$. This idea, as illustrated in Figure 1, can be formulated as the following problem [1]:

$$\begin{aligned} \max_{\substack{\mathbf{v}_i \in \mathbb{R}^{N-1}, \\ i=1, \dots, N}} \left\{ \min_{\substack{\|\mathbf{u}_i\| \leq r, \\ i=1, \dots, N}} \left| \det(\Delta(\mathbf{v}_1 - \mathbf{u}_1, \dots, \mathbf{v}_N - \mathbf{u}_N)) \right| \right\} \\ \text{s.t. } \mathbf{v}_i \in \text{conv}\{\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]\}, i = 1, \dots, N, \end{aligned} \quad (4)$$

where each \mathbf{u}_i lying in a norm ball $\{\mathbf{u} \in \mathbb{R}^{N-1} \mid \|\mathbf{u}\| \leq r\}$ is the pull-back vector, r is the maximum back-off distance,

$$\Delta(\mathbf{t}_1, \dots, \mathbf{t}_N) = \begin{bmatrix} \mathbf{t}_1 & \dots & \mathbf{t}_N \\ 1 & \dots & 1 \end{bmatrix} \in \mathbb{R}^{N \times N}$$

for any $\mathbf{t}_i \in \mathbb{R}^{N-1}$, and $\text{conv}\{\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]\}$ is defined as

$$\text{conv}\{\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]\} = \left\{ \mathbf{y} = \sum_{n=1}^L \theta_n \tilde{\mathbf{y}}[n] \mid \theta \succeq \mathbf{0}, \mathbf{1}_L^T \theta = 1 \right\}, \quad (5)$$

where $\theta = [\theta_1, \dots, \theta_L]^T$. Denoting the optimal solution of problem (4) by $(\hat{\mathbf{v}}_1, \dots, \hat{\mathbf{v}}_N, \hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_N)$, the *robust endmember estimates* are obtained by

$$\hat{\nu}_i = \hat{\mathbf{v}}_i - \hat{\mathbf{u}}_i, i = 1, \dots, N. \quad (6)$$

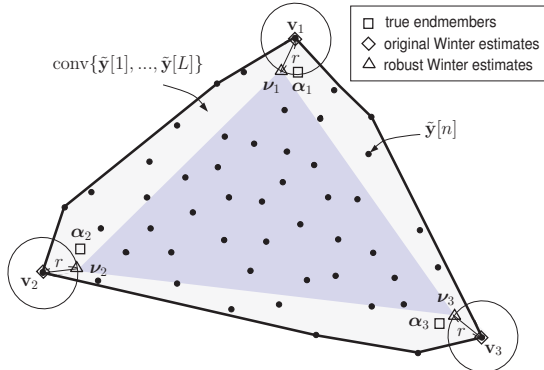


Fig. 1. Illustration of robust Winter's EE problem for $N = 3$.

In [1], we have proposed an algorithm for handling problem (4). Called WAVMAX, the algorithm demonstrates performance improvement in the noisy scenario. However, WAVMAX is expensive to implement.

4. FAST ALGORITHMS FOR WORST-CASE WINTER'S ENDMEMBER EXTRACTION PROBLEM

In this section, we propose two fast algorithms for dealing with the worst-case Winter's problem (4). We utilize alternating optimization and successive optimization to approximate the max-min problem (4) by a sequence of max-min subproblems, leading to alternating decoupled volume max-min (ADVMM) and successive decoupled volume max-min (SDVMM) algorithms, respectively.

4.1. ADVMM Algorithm

By letting $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]] \in \mathbb{R}^{(N-1) \times L}$, $\mathbf{v}_i = \tilde{\mathbf{Y}}\theta_i$, and the property $\det(\mathbf{P}\Delta) = \pm \det(\Delta)$ for any permutation matrix \mathbf{P} , problem (4) can be expressed as

$$\max_{\substack{\theta_i \in \mathcal{S}, \\ i=1, \dots, N}} \left\{ \min_{\substack{\|\mathbf{u}_i\| \leq r, \\ i=1, \dots, N}} \det(\Delta(\tilde{\mathbf{Y}}\theta_1 - \mathbf{u}_1, \dots, \tilde{\mathbf{Y}}\theta_N - \mathbf{u}_N)) \right\} \quad (7)$$

where $\mathcal{S} = \{\theta \in \mathbb{R}^L \mid \theta \succeq \mathbf{0}, \mathbf{1}_L^T \theta = 1\}$. Optimizing $\theta_1, \dots, \theta_N$ and $\mathbf{u}_1, \dots, \mathbf{u}_N$ jointly in (7) is quite challenging. By alternating optimization, let us consider the partial max-min problem of (7) with respect to (w.r.t.) the pair (θ_j, \mathbf{u}_j) while fixing the other pairs (θ_i, \mathbf{u}_i) for $i \neq j$. The j th partial max-min problem is represented by

$$\max_{\theta_j \in \mathcal{S}} \left\{ \min_{\|\mathbf{u}_j\| \leq r} \det(\Delta(\tilde{\mathbf{Y}}\hat{\theta}_1 - \hat{\mathbf{u}}_1, \dots, \tilde{\mathbf{Y}}\hat{\theta}_N - \hat{\mathbf{u}}_N)) \right\}. \quad (8)$$

The partial max-min problems (8) for $j = 1, \dots, N$ are conducted cyclically until some stopping criterion is satisfied.

Next, we will present how to solve the partial max-min problem (8). By applying a cofactor expansion of $\det(\Delta(\tilde{\mathbf{Y}}\hat{\theta}_1 - \hat{\mathbf{u}}_1, \dots, \tilde{\mathbf{Y}}\hat{\theta}_N - \hat{\mathbf{u}}_N))$ along j th column, we have

$$\mathbf{k}_j^T (\tilde{\mathbf{Y}}\theta_j - \mathbf{u}_j) + (-1)^{N+j} \det(\mathcal{Q}_{Nj}), \quad (9)$$

where

$$\mathbf{k}_j = [(-1)^{i+j} \det(\mathcal{Q}_{ij})]_{i=1}^{N-1} \in \mathbb{R}^{N-1} \quad (10)$$

is a vector with i th element equal to $(-1)^{i+j} \det(\mathcal{Q}_{ij})$ and $\mathcal{Q}_{ij} \in \mathbb{R}^{(N-1) \times (N-1)}$ is a submatrix of $\Delta(\tilde{\mathbf{Y}}\hat{\theta}_1 - \hat{\mathbf{u}}_1, \dots, \tilde{\mathbf{Y}}\hat{\theta}_N - \hat{\mathbf{u}}_N)$ with the i th row and the j th column removed. Then, problem (8) is equivalent to

$$\max_{\theta_j \in \mathcal{S}} \left\{ \min_{\|\mathbf{u}_j\| \leq r} \mathbf{k}_j^T (\tilde{\mathbf{Y}}\theta_j - \mathbf{u}_j) \right\}, \quad (11)$$

where the term $(-1)^{N+j} \det(\mathcal{Q}_{Nj})$ independent of (θ_j, \mathbf{u}_j) is removed without loss of optimality. In addition, since θ_j and \mathbf{u}_j are decoupled, the above problem can be handled by solving the following problems separately:

$$\hat{\mathbf{u}}_j = \arg \max_{\|\mathbf{u}_j\| \leq r} \mathbf{k}_j^T \mathbf{u}_j = r \mathbf{k}_j / \|\mathbf{k}_j\|, \quad (12)$$

$$\hat{\theta}_j = \arg \max_{\theta_j \in \mathcal{S}} \mathbf{k}_j^T \tilde{\mathbf{Y}}\theta_j = \mathbf{e}_\ell, \ell = \arg \max_{n=1, \dots, L} \mathbf{k}_j^T \tilde{\mathbf{y}}[n], \quad (13)$$

where $\hat{\mathbf{u}}_j$ in (12) is obtained by Cauchy-Schwarz inequality and $\hat{\theta}_j$ in (13) can be obtained by [1, Lemma 2]. The pseudo-codes of the proposed ADVMM are given in Table 1 (left part).

Table 1. The pseudo-codes of the proposed ADVMM and SDVMM algorithms for problem (4).

ADVMM Algorithm	SDVMM Algorithm
<p>Given tolerance $\varepsilon > 0$, back-off distance r, $\tilde{\mathbf{Y}}$ and N.</p> <p>S1. randomly select $(\hat{\theta}_1, \dots, \hat{\theta}_N)$ from $\{\mathbf{e}_i\}_{i=1}^L$ and set $\hat{\mathbf{u}}_1 = \dots = \hat{\mathbf{u}}_N = \mathbf{0}$.</p> <p>S2. set $j := 1$, $\varrho := \det(\Delta(\tilde{\mathbf{Y}}\hat{\theta}_1 - \hat{\mathbf{u}}_1, \dots, \tilde{\mathbf{Y}}\hat{\theta}_N - \hat{\mathbf{u}}_N))$.</p> <p>S3. compute \mathbf{k}_j by (10), and update $\hat{\mathbf{u}}_j$ by (12) and $\hat{\theta}_j$ by (13).</p> <p>S4. if $(j \bmod N) \neq 0$, then $j := j + 1$ and go to S3, else compute $\bar{\varrho} = \det(\Delta(\tilde{\mathbf{Y}}\hat{\theta}_1 - \hat{\mathbf{u}}_1, \dots, \tilde{\mathbf{Y}}\hat{\theta}_N - \hat{\mathbf{u}}_N))$.</p> <p>S5. if $\bar{\varrho} - \varrho /\varrho > \varepsilon$, then set $\varrho := \bar{\varrho}$, $j := 1$, and go to S3, else output $\hat{\mathbf{v}}_j = \tilde{\mathbf{Y}}\hat{\theta}_j - \hat{\mathbf{u}}_j$, $\forall j$ as an approximate solution to (4).</p>	<p>Given back-off distance r, $\{\tilde{\mathbf{y}}[n]\}_{n=1}^L$ and N.</p> <p>S1. construct $\tilde{\mathbf{y}}[n] = [\tilde{\mathbf{y}}[n]^T \ 1]^T$, $\forall n$ and set $\hat{\mathbf{H}}_{1:0} = \mathbf{I}_N$ and $j = 0$.</p> <p>S2. update $j := j + 1$ and obtain $\hat{\mathbf{w}}_j$ by (23), and $\hat{\mathbf{z}}_j$ by (20).</p> <p>S3. set $[\hat{\mathbf{z}}_j]_N = 0$, update $\hat{\mathbf{H}}_{1:j} := [\hat{\mathbf{H}}_{1:(j-1)} \ \hat{\mathbf{w}}_j - \hat{\mathbf{z}}_j]$ and go to S2 until $j = N$.</p> <p>S4. output $\hat{\mathbf{v}}_j = [\hat{\mathbf{w}}_j]_{1:N-1} - [\hat{\mathbf{z}}_j]_{1:N-1}$, $\forall j$ as an approximate solution to (4).</p>

4.2. SDVMM Algorithm

By letting $\mathbf{w}_i = [\mathbf{v}_i^T \ 1]^T$, $\mathbf{z}_i = [\mathbf{u}_i^T \ 0]^T$ and $\tilde{\mathbf{y}}[n] = [\tilde{\mathbf{y}}[n]^T \ 1]^T$, problem (4) can be rewritten as

$$\max_{\substack{\mathbf{w}_i \in \mathcal{F}, \\ i=1, \dots, N}} \left\{ \min_{\substack{\|\mathbf{z}_i\| \leq r, \\ \mathbf{e}_N^T \mathbf{z}_i = 0, \forall i}} \left| \det([\mathbf{w}_1 - \mathbf{z}_1, \dots, \mathbf{w}_N - \mathbf{z}_N]) \right| \right\} \quad (14)$$

where $\mathcal{F} = \text{conv}\{\tilde{\mathbf{y}}[1], \dots, \tilde{\mathbf{y}}[L]\}$. It has been shown in [1, Lemma 3] that problem (14) can be equivalently represented by

$$\max_{\substack{\mathbf{w}_i \in \mathcal{F}, \\ i=1, \dots, N}} \min_{\substack{\|\mathbf{z}_i\| \leq r, \\ \mathbf{e}_N^T \mathbf{z}_i = 0, \forall i}} \prod_{j=1}^N f((\mathbf{w}_1, \mathbf{z}_1), \dots, (\mathbf{w}_j, \mathbf{z}_j)) \quad (15)$$

where

$$f((\mathbf{w}_1, \mathbf{z}_1), \dots, (\mathbf{w}_j, \mathbf{z}_j)) = \|\mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp (\mathbf{w}_j - \mathbf{z}_j)\|, \quad (16)$$

in which $\mathbf{H}_{1:j} = [\mathbf{w}_1 - \mathbf{z}_1, \dots, \mathbf{w}_j - \mathbf{z}_j]$, $\mathbf{P}_{\mathbf{H}_{1:j}}^\perp = \mathbf{I}_N - \mathbf{H}_{1:j}(\mathbf{H}_{1:j}^T \mathbf{H}_{1:j})^{-1} \mathbf{H}_{1:j}^T$ is the orthogonal complement projector of $\mathbf{H}_{1:j}$, and $\mathbf{P}_{\mathbf{H}_{1:0}}^\perp = \mathbf{I}_N$. Solving problem (15) w.r.t. $2N$ -tuple $(\mathbf{w}_1, \dots, \mathbf{w}_N, \mathbf{z}_1, \dots, \mathbf{z}_N)$ is difficult. We approximate problem (15) by successive optimization as follows:

$$(\hat{\mathbf{w}}_j, \hat{\mathbf{z}}_j) = \arg \max_{\mathbf{w}_j \in \mathcal{F}} \min_{\substack{\|\mathbf{z}_j\| \leq r, \\ \mathbf{e}_N^T \mathbf{z}_j = 0}} f((\hat{\mathbf{w}}_1, \hat{\mathbf{z}}_1), \dots, (\hat{\mathbf{w}}_{j-1}, \hat{\mathbf{z}}_{j-1}), (\mathbf{w}_j, \mathbf{z}_j)) \quad (17)$$

from $j = 1$ to N . The solution $(\hat{\mathbf{w}}_j, \hat{\mathbf{z}}_j)$ can be obtained by handling the j th max-min subproblem with the previous $(j-1)$ max-min subproblem solutions $\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_{j-1}, \hat{\mathbf{z}}_1, \dots, \hat{\mathbf{z}}_{j-1}$ given. Unlike alternating optimization, the methodology presented here is initialization free and only needs to solve (17) successively for $j = 1, \dots, N$.

The issue that remains is how we handle each difficult (non-convex) max-min subproblem (17). By relaxing $\mathbf{e}_N^T \mathbf{z}_j = 0$, it can be shown that a closed-form solution to (17) exists. To see this, by (16), problems (17) with $\mathbf{e}_N^T \mathbf{z}_j = 0$ relaxed is

$$\max_{\mathbf{w}_j \in \mathcal{F}} \min_{\|\mathbf{z}_j\| \leq r} \left\| \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp (\mathbf{w}_j - \mathbf{z}_j) \right\|, \quad j = 1, \dots, N. \quad (18)$$

The inner problem of (18) for any $\mathbf{w}_j \in \mathcal{F}$ is given by

$$\hat{\mathbf{z}}_j = \arg \min_{\|\mathbf{z}_j\| \leq r} \left\| \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp (\mathbf{w}_j - \mathbf{z}_j) \right\|. \quad (19)$$

Problem (19) is convex and Slater's condition holds. The optimal solution of problem (19) can be derived by Karush–Kuhn–Tucker (KKT) conditions, as stated in the following lemma:

Lemma 1. For any $\mathbf{w}_j \in \mathcal{F}$, problem (19) has an analytical solution given by

$$\hat{\mathbf{z}}_j = r \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \mathbf{w}_j / \left\| \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \mathbf{w}_j \right\|, \quad \mathbf{w}_j \in \mathcal{W}(r), \quad (20)$$

$$\hat{\mathbf{z}}_j \in \{ \mathbf{z}_j \mid \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp (\mathbf{w}_j - \mathbf{z}_j) = \mathbf{0} \}, \quad \mathbf{w}_j \in \mathbb{R}^N \setminus \mathcal{W}(r), \quad (21)$$

where $\mathcal{W}(r) = \{ \mathbf{w} \in \mathbb{R}^N \mid \|\mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \mathbf{w}\| > r \}$.

Proof: The proof of Lemma 1 is given in Appendix. ■

It is trivial to see that the solution (21) always yields zero objective value in (18), and hence the optimal solution (20) is considered. Substituting (20) into (18) yields

$$\max_{\mathbf{w}_j \in \mathcal{F} \cap \mathcal{W}(r)} \left\| \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \mathbf{w}_j \right\|. \quad (22)$$

The optimal solution of (22) can be easily obtained by following the proof in [1, Lemma 4]; it is given by

$$\hat{\mathbf{w}}_j = \tilde{\mathbf{y}}[\ell], \quad \ell = \arg \max_{n \in \mathcal{N}_j} \left\| \mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \tilde{\mathbf{y}}[n] \right\|, \quad (23)$$

where $\mathcal{N}_j = \{ n \mid \|\mathbf{P}_{\mathbf{H}_{1:(j-1)}}^\perp \tilde{\mathbf{y}}[n]\| > r, n = 1, \dots, L \}$.

We should mention that the constraint $\mathbf{w}_j \in \mathcal{W}(r)$ is to ensure the meaningful solution of problem (18). In fact, one can properly choose an r such that $\mathbf{w}_j \in \mathcal{W}(r)$, $j = 1, \dots, N$ are all satisfied. Also, if the $(\hat{\mathbf{w}}_j, \hat{\mathbf{z}}_j)$ is obtained, we can artificially set $[\hat{\mathbf{z}}_j]_N = 0$ to ensure the feasibility of $(\hat{\mathbf{w}}_j, \hat{\mathbf{z}}_j)$ to problem (17). The pseudo-codes of the SDVMM algorithm are given in Table 1 (right part).

5. SIMULATION AND CONCLUSION

Monte Carlo simulations of 100 independent runs are performed to demonstrate the performance of the proposed ADVMM and SDVMM algorithms¹, compared to the four existing methods, SQ-N-FINDR [7], SC-N-FINDR [7], SGA [8], and WAVMAX [1]. The root-mean-square (rms) spectral angle distance, denoted as ϕ (in degrees), was used as the error performance measure [1]. The computation time T (in secs) of each algorithm (implemented in Mathworks Matlab R2008a) running in a desktop computer equipped with Core i7-930 CPU 2.80 GHz, 12GB memory is used as our computational complexity measure. In each run, the observed data were synthetically generated following (1) where $N = 8$ endmember signatures with $M = 224$ bands were selected from the U.S. geological survey (USGS) library [10], the abundance vectors were generated following Dirichlet distribution [1], and zero-mean white Gaussian noise vectors were added for different signal-to-noise ratios (SNRs),

¹A Matlab implementation is provided at <http://mx.nthu.edu.tw/~tsunghan>

Table 2. Performance comparison of average ϕ (degrees) and average T (secs) over some existing EE methods.

Algorithms		Case I: $M = 224, N = 8, L = 1000$						Case II: $M = 224, N = 8, \text{SNR} = 15 \text{ dB}$					
		SNR (dB)						Number of pixels (L)					
		5	10	15	20	25	∞	250	500	1000	2000	4000	8000
SGA [8]	ϕ	13.92	8.06	3.34	1.73	0.96	0.00	5.32	3.95	3.34	3.01	2.97	2.94
	T	0.126	0.124	0.116	0.113	0.111	0.111	0.118	0.192	0.255	0.577	1.054	2.231
SQ-N-FINDR [7]	ϕ	14.13	8.06	3.50	1.86	1.07	0.00	5.23	4.08	3.50	3.19	3.11	3.11
	T	0.179	0.164	0.138	0.130	0.120	0.110	0.109	0.179	0.247	0.556	0.907	1.664
SC-N-FINDR [7]	ϕ	14.90	8.50	3.92	2.00	1.08	0.00	5.99	4.66	3.92	3.52	3.54	3.34
	T	0.068	0.066	0.059	0.056	0.055	0.054	0.066	0.096	0.120	0.244	0.384	0.649
WAVMAX [1]	ϕ	13.30	8.10	3.15	1.75	1.03	0.00	5.24	4.01	3.15	2.92	2.73	2.58
	T	42.879	55.778	43.853	45.579	46.696	39.285	11.518	30.861	65.850	233.859	778.585	2997.140
ADVMM	ϕ	12.95	7.27	3.15	1.75	1.03	0.00	5.16	3.87	3.15	2.85	2.70	2.61
	T	0.042	0.038	0.027	0.020	0.017	0.015	0.066	0.073	0.072	0.111	0.166	0.218
SDVMM	ϕ	13.50	7.45	3.00	1.59	0.90	0.00	5.04	3.66	3.00	2.69	2.49	2.42
	T	0.028	0.026	0.019	0.016	0.014	0.014	0.049	0.056	0.060	0.099	0.149	0.184

where $\text{SNR} = \sum_{n=1}^L \|\mathbf{x}[n]\|^2 / \sigma^2 ML$. For the proposed methods, the convergence tolerance $\varepsilon = 5 \times 10^{-5}$ and $r = 1.3\sigma$, where noise power σ is assumed to be known.

Table 2 shows the average ϕ and T per realization over various algorithms. The minimum ϕ and T for a specific SNR or L are highlighted by bold-faced numbers. In Case I, ADVMM outperforms the other algorithms for SNR = 5, 10 dB, and SDVMM performs best for SNR ≥ 15 dB. In Case II, the performance of SDVMM is better than the other algorithms for various values of L under test. Besides, the proposed ADVMM and SDVMM algorithms are faster than all the other existing methods, and are more than 1000 times faster than WAVMAX.

In conclusion, we have developed two fast, noise-robust Winter criterion based hyperspectral EE algorithms, namely ADVMM and SDVMM, by using alternative and successive optimization strategies, respectively. Simulation results have shown superior efficacy and computational efficiency of the proposed methods over some existing benchmark EE algorithms.

6. APPENDIX

The KKT conditions of Problem (19) are as below:

$$\left(\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp + \hat{\lambda} \mathbf{I}_N \right) \hat{\mathbf{z}}_j = \mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j, \quad (24a)$$

$$\hat{\lambda} (\|\hat{\mathbf{z}}_j\|^2 - r^2) = 0, \quad (24b)$$

$$\|\hat{\mathbf{z}}_j\|^2 - r^2 \leq 0, \quad \hat{\lambda} \geq 0, \quad (24c)$$

where $\hat{\mathbf{z}}_j$ and $\hat{\lambda}$ are primal and dual optimal points of (19). By (24a), (24c), we have

$$\|\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j\| \leq \|\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp + \hat{\lambda} \mathbf{I}_N\| \|\hat{\mathbf{z}}_j\| \leq (1 + \hat{\lambda})r, \quad (25)$$

where the first inequality is due to the inequality of the operator norm, and the second inequality is due to the eigenvalues of a projection matrix equal to either zero or one.

Two cases on $\mathbf{w}_j \in \mathcal{F}$ are considered: (C1) $\mathbf{w}_j \in \mathcal{W}(r)$ and (C2) $\mathbf{w}_j \in \mathbb{R}^N \setminus \mathcal{W}(r)$. We first consider (C1). By (C1) and (25), we can have $\hat{\lambda} > 0$, which implies $\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp + \hat{\lambda} \mathbf{I}_N$ is of full rank. Then, (24a) becomes

$$\hat{\mathbf{z}}_j = (\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp + \hat{\lambda} \mathbf{I}_N)^{-1} \mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j. \quad (26)$$

By idempotence property and eigenvalue decomposition of $\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp$ and through some mathematical derivations, (26) can be

shown to be

$$\hat{\mathbf{z}}_j = \frac{1}{1 + \hat{\lambda}} \mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j. \quad (27)$$

By (27) and $\hat{\lambda} > 0$, (24b) becomes $\|\hat{\mathbf{z}}_j\| = \|\frac{1}{1 + \hat{\lambda}} \mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j\| = r$, which yields

$$\hat{\lambda} = r^{-1} \|\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j\| - 1. \quad (28)$$

Therefore, by (27) and (28), the solution (20) can be obtained.

We next consider (C2); i.e., $\mathbf{w}_j \in \mathbb{R}^N \setminus \mathcal{W}(r)$. As it can be shown by (28) that $\hat{\lambda} > 0$ implies (C1), then by contradiction, (C2) implies $\hat{\lambda} = 0$. Hence (24a) reduces to

$$\mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \hat{\mathbf{z}}_j = \mathbf{P}_{\hat{\mathbf{H}}_{1:(j-1)}}^\perp \mathbf{w}_j, \quad (29)$$

which leads to (21). ■

7. REFERENCES

- [1] T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," to appear in *IEEE Trans. Geoscience and Remote Sensing - Special Issue on Spectral Unmixing of Remotely Sensed Data*, 2011.
- [2] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44-57, Jan. 2002.
- [3] 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, Nov. 2009.
- [4] 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.
- [5] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "Chance constrained robust minimum volume enclosing simplex algorithm for hyperspectral unmixing," to appear in *IEEE Trans. Geoscience and Remote Sensing - Special Issue on Spectral Unmixing of Remotely Sensed Data*, 2011.
- [6] A. Zare and P. Gader, "L1-endmembers: A robust endmember detection and spectral unmixing algorithm," in *Proc. of SPIE*, vol. 7695, p. 76951L, Dec. 9-14, 2010.
- [7] C.-C. Wu, S. Chu, and C.-I Chang, "Sequential N-FINDR algorithms," in *Proc. of SPIE*, vol. 7086, Aug. 2008.
- [8] 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, April 2010.
- [9] 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.
- [10] Tech. Rep., Available online: <http://speclab.cr.usgs.gov/cuprite.html>.