Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images

論文作者：亞魯
ArulMurugan Ambikapathi

指導教授：祁忠勇 博士
Prof. Chong-Yung Chi

國 立 清 華 大 學
通 訊 工 程 研 究 所
博 士 論 文

中 華 民 國 一百 年 八 月

國 立 清 華 大 學

博碩士論文全文電子檔著作權授權書

(提供授權人裝訂於紙本論文書名頁之次頁用)

本授權書所授權之學位論文，為本人於國立清華大學通訊工程研究所_____組， 99 學年度第____學期取得博士學位之論文。

論文題目：Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images
指導教授：祁忠勇

■ 同意

本人茲將本著作，以非專屬、無償授權國立清華大學與台灣聯合大學系統圖書館：基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立清華大學及台灣聯合大學系統圖書館得不限地域、時間與次數，以紙本、光碟或數位化等各種方法收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行線上檢索、閱覽、下載或列印。

論文電子全文上載網路公開之範圍及時間：

本校及台灣聯合大學系統區域網路	■ 中華民國 105 年 8 月 19 日公開
校外網際網路	■ 中華民國 105 年 8 月 19 日公開

■ 全文電子檔送交國家圖書館

授 權 人：亞魯

親筆簽名：_____

中華民國　　年　　月　　日

國 立 清 華 大 學

博碩士紙本論文著作權授權書

(提供授權人裝訂於全文電子檔授權書之次頁用)

本授權書所授權之學位論文，為本人於國立清華大學通訊工程研究所_____組， 99 學年度第____學期取得博士學位之論文。

論文題目：Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images
指導教授：祁忠勇

■ 同意

本人茲將本著作，以非專屬、無償授權國立清華大學，基於推動讀者間「資源共享、互惠合作」之理念，與回饋社會與學術研究之目的，國立清華大學圖書館得以紙本收錄、重製與利用；於著作權法合理使用範圍內，讀者得進行閱覽或列印。

本論文為本人向經濟部智慧局申請專利(未申請者本條款請不予理會)的附件之一，申請文號為：_____，請將論文延至____年____月____日再公開。

授 權 人：亞魯

親筆簽名：_____

中華民國　　年　　月　　日

國家圖書館
博碩士論文電子檔案上網授權書

（提供授權人裝訂於紙本論文本校授權書之後）

ID:GH029664882

本授權書所授權之論文爲授權人在國立清華大學通訊工程研究所 99 學年度第___學期取得博士學位之論文。

論文題目：Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images
指導教授：祁忠勇

茲同意將授權人擁有著作權之上列論文電子全文（含摘要），非專屬、無償授權國家圖書館，不限地域、時間與次數，以微縮、光碟或其他各種數位化方式將上列論文重製，並得將數位化之上列論文及論文電子檔以上載網路方式，提供讀者基於個人非營利性質之線上檢索、閱覽、下載或列印。

☐ 讀者基於非營利性質之線上檢索、閱覽、下載或列印上列論文，應依著作權法相關規定辦理。

☐ 論文電子全文上載網路公開之範圍及時間：中華民國 105 年 8 月 19 日公開。

授權人：亞魯

親筆簽名：_____

民國　　年　　月　　日

國立清華大學博士學位論文

指導教授推薦書

學　系
通訊工程　研究所 ArulMurugan Ambikapathi 君
所提之論文

Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images

經由本人指導撰述，同意提付審查。

指導教授 祁忠勇 (簽章)

中華民國　100　年　8　月　19　日

國立清華大學博士學位論文

考試委員審定書

學　系
通訊工程　研究所 ArulMurugan Ambikapathi 君
所提之論文

Chance Constrained Robust Unmixing Algorithms and Estimation of Number of Endmembers in Hyperspectral Images

經本委員會審查，符合博士資格標準。

學位考試委員會

主持人 薄緯山 (簽章)

委　員 鄭伯昆　任玄

峰陽劉　祁忠勇

中華民國　100　年　8　月　19　日

Abstract

Hyperspectral remote sensing is a powerful technique to identify the materials and their composition in an area by exploiting the spectral diversity of the observed hyperspectral data. The analysis of hyperspectral images obtained for the purpose of mineral identification and quantification is considered in this thesis. The limited spatial resolution of the sensor used for hyperspectral imaging and the presence of noise in the measured hyperspectral data demand an effective hyperspectral unmixing (HU) scheme to extract the underlying endmember signatures and the associated abundance maps distributed over a scene of interest.

Existing HU algorithms are basically devised under either of the two famous unmixing criteria, namely Winter's criterion and Craig's criterion. However, the presence of additive Gaussian noise in the observations expands the actual data cloud and as a consequence, the endmember estimates obtained by applying either Winter's or Craig's criterion based algorithms to the noisy data may no longer be in close proximity to the true endmember signatures. Hence, we propose two robust algorithms, they are Winter's criterion based robust alternating volume maximization (RAVMAX) algorithm and Craig's criterion based robust minimum volume enclosing simplex (RMVES) algorithm. The robust algorithms account for the noise effects in the observations by employing chance constraints, and employ the notion of alternating optimization to handle the resulting non-convex optimization problems. In RAVMAX algorithm, the subproblems involved in each alternating optimization turn out to be convex problems and they can be effectively solved using available con-

vex optimization solvers. On the other hand, the subproblems involved in RMVES algorithm are non-convex and are hence dealt using available sequential quadratic programming solvers.

The HU results can be completely interpretable, only when the number of substances (or endmembers) present in that area is given *a priori*, which however is unknown in practice. Considering the linear mixing model, we propose a hyperspectral data geometry based approach for estimating the number of endmembers by utilizing a successive endmember extraction algorithm (EEA). The approach is fulfilled by two novel algorithms, namely geometry based estimation of number of endmembers - convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm. The GENE-CH and GENE-AH algorithms are based on the fact that all the observed pixel vectors lie in the convex hull and affine hull of the endmember signatures, respectively. The proposed GENE algorithms estimate the number of endmembers by using the Neyman-Pearson hypothesis testing over the endmember estimates provided by a successive EEA until the estimate of the number of endmembers is obtained. Since the estimation accuracies of the proposed GENE algorithms depend on the performance of the EEA used, a reliable, reproducible, and successive EEA, called p-norm based pure pixel identification (TRI-P) algorithm is then proposed. Monte-Carlo simulations and real data experiments on AVIRIS hyperspectral data obtained over the Cuprite mining site, Nevada are performed to demonstrate the efficacy of the proposed RAVMAX, RMVES, GENE, and TRI-P algorithms.

We believe that the proposed chance constrained robust algorithms for hyperspectral unmixing, and data geometry based algorithms for estimating the number of endmembers, will provide a new dimension in analyzing hyperspectral data where noise is always present.

Acknowledgments

Working toward my Ph.D. degree is a hard and challenging experience, and it was with the best support from numerous people I could succeed in that. I feel honored and duty-bound to express my gratitude to them. First and foremost, I would like to place on record, my heart-felt gratitude to Prof. Chong-Yung Chi for being my advisor, mentor, role model, and much more. I will remain grateful for the wonderful opportunities, expert guidance, and support provided by him. I wish to extend my deepest appreciation to Prof. Wing-Kin Ma, Chinese University of Hong Kong, for his dedication and great patience with which he taught me how to deal with the research problems. The oral and written comments of Prof. Chong-Yung Chi and Prof. Wing-Kin Ma are highly perceptive, helpful, and appropriate. They were instrumental in making me understand the core of research.

I wish to express my hearty gratitude to Prof. Jen-Ming Wu, Prof. Yao-Win Peter Hong, and Prof. Meng-Lin Li for serving as expert panel members for my internal oral defense, and Prof. Kun-Shan Chen, Prof. Bor-Chen Kuo, Prof. Hsuan Ren (Shane), and Prof. Yang-Lang Chang for serving as expert panel members for my final oral defense. My sincere gratitude to Prof. Shyang Chang, Prof. Jen-Ming Wu, and Prof. Yao-Win Peter Hong for teaching me the concepts and ideas of communications.

My special thanks to my beloved friend and mentor, Dr. Tsung-Han Chan for all his encouragements, teachings, ideas, and assistances which enhanced my confidence to keep pursuing my Ph.D. degree. I should also thank the members of Wireless Communications and Signal Processing (WCSP) Laboratory, especially Mr. Yi-Lin

Chiou and Mr. Ji-Yuan Liu who helped me in various forms.

I would like to place on record my hearty gratitude to the management and staff of SSN Institutions, India. Especially I would like to thank the president of SSN Institutions Ms. Kala Vijaykumar and the Principal of SSN college of Engineering, Prof. S. Salivahanan for having allowed me to pursue my Ph.D. study. I am honored to be the recipient of Outstanding National Tsing Hua University Scholarship for the academic years 2009-10 and 2010-11. I offer my sincere gratitude for the generous financial supports that I have received throughout my Ph.D. study.

My family members are the core for all my endeavors. I am always grateful to my dear father, Mr. Ambikapathi, loving mother, Dr. Ambika, to whom I would like to dedicate this thesis. My smart younger sister, Mrs. Vanidevi, ever remained my motivator. I must also mention my kind gratitude to my cute little niece Ms. Shree Abirami. I am always love-bound to my beloved soul-mate Mrs. Niraimathi, who defined my inner self. In addition, I am pleased to thank my uncles Mr. Sivanesan and Dr. Murugan, and my dear friends Mr. Siva Shanmugam, Mr. Rajesh Kannan, and Mr. Kannan Keizer, for their affection, motivation and helps. I am blessed to have the above persons as my family members, well wishers, and friends. I firmly believe that getting this Ph.D. degree is a new beginning, and will strive my best to keep going further.

ArulMurugan Ambikapathi

Hsinchu, August 2011.

Table of Contents

Abstract ii

Acknowledgments iv

List of Figures viii

List of Tables x

List of Notations xiii

1 Introduction 1

2 Linear Mixing Model for Hyperspectral Data 10

2.1 Data Preprocessing 13

3 Robust Alternating Volume Maximization Algorithm 18

3.1 Brief review of AVMAX algorithm 19

3.2 Robust AVMAX Formulation and Algorithm 21

3.2.1 Restructuring the AVMAX algorithm 22

3.2.2 Robust AVMAX Algorithm 24

3.3 Simulations 27

3.3.1 Uniform Gaussian noise case 29

3.3.2 Non-uniform Gaussian noise case 30

3.4 Real data experiments 31

3.5 Summary 35

4 Robust Minimum Volume Enclosing Simplex Algorithm **39**

4.1 Brief review of MVES Algorithm 41

4.2 Robust MVES Formulation and Algorithm 43

4.3 Simulations 49

4.3.1 Uniform Gaussian noise case 53

4.3.2 Non-uniform Gaussian noise case 56

4.3.3 Local optimality issues 56

4.4 Real data experiments 59

4.5 Summary 61

5 Estimation of Number of Endmembers Using $p-$norm Based Pure Pixel Identification Algorithm **67**

5.1 Dimension reduction-revisited 68

5.1.1 Convex geometry of the hyperspectral data 70

5.2 Geometry based estimation of number of endmembers (GENE) 71

5.2.1 GENE-Convex Hull (GENE-CH) Algorithm 72

5.2.2 GENE-Affine Hull (GENE-AH) Algorithm 76

5.3 p-norm based Pure Pixel Identification (TRI-P) Algorithm 81

5.4 Simulations 85

5.4.1 Evaluation of TRI-P algorithm 87

5.4.2 Evaluation of GENE-CH and GENE-AH algorithms 87

5.5 Real data experiments 90

5.6 Summary 94

6 Conclusions and Future Works **100**

Bibliography **102**

Publication List of The Author **109**

List of Figures

1.1 Illustration of hyperspectral unmixing. 3

2.1 Illustration of dimension reduction by affine set fitting for N=3, where the geometric center $\mathbf{d}$ of the data cloud $\mathbf{x}[n]$ in M-dimensional space maps to the origin in the $(N-1)$-dimensional space which is also the geometric center of the dimension-reduced data cloud $\tilde{\mathbf{x}}[n]$. 14

3.1 Illustration of Winter's simplex for noisy observations, where $N = 3$. . 21

3.2 Illustration of the effect of η in RAVMAX for $N = 3$. 26

3.3 Endmember signatures taken from library, and the ones estimated by VCA, RAVMAX, and AVMAX algorithms. 34

3.4 Abundance maps obtained by RAVMAX algorithm. 35

3.5 Abundance maps obtained by AVMAX algorithm. 36

3.6 Abundance maps obtained by VCA algorithm. 37

4.1 Illustration of Craig's simplex for noisy observations, where $N = 3$. . 40

4.2 Scatter plot of the dimension-reduced pixels for $N = 3$, illustrating the solutions of RMVES for different values of η. 46

4.3 Average endmember spectral angles obtained by the RMVES algorithm for $\rho = 0.6$ with one VCA initialization, and with ten VCA initializations and choosing the spectral angle associated with the maximum $|\det(\mathbf{H})|$. 58

4.4 Endmember signatures taken from library, and the ones estimated by RMVES, MVES, and VCA algorithms. 62
4.5 Abundance maps obtained by the RMVES algorithm. 63
4.6 Abundance maps obtained by the MVES algorithm. 64
4.7 Abundance maps obtained by the VCA algorithm. 65

5.1 Illustration of (F1) and (F2), for $N = 3$ case. 72
5.2 Working strategy of GENE-CH/GENE-AH algorithm in synchronization with a suitable SEEA. 97
5.3 Illustration of GENE-CH algorithm, when no pure pixel is present in the noise-free hyperspectral data ($N = 3$ case). (a) The endmember estimates are denoted by $\tilde{\mathbf{x}}[l_i], i = 1, ..., N_{\max} = 6$, but $\text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_6]\} \neq \text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ because the true endmembers $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ are not present in the data cloud, whereas $\text{aff}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_k]\} = \text{aff}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, $k = 3, 4, 5, 6$, as shown in (b). 98
5.4 Abundance maps estimated based on endmembers obtained by TRI-P algorithm (with $p = 2$). 99

List of Tables

3.1 Pseudocode for RAVMAX Algorithm. 28

3.2 Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different purity levels (ρ) and SNRs- Uniform Gaussian noise case. 30

3.3 Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different τ values and SNRs, with $\rho = 1$- Non-uniform Gaussian noise case. 31

3.4 Mean-removed spectral angles ϕ (degrees) for VCA, RAVMAX ($\eta = 0.9$) and AVMAX algorithms. 38

4.1 Pseudocode for expanded VCA initialization. 50

4.2 Pseudocode for RMVES Algorithm. 51

4.3 Implementation details for the algorithms under test. 52

4.4 Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different purity levels (ρ) and SNRs- Uniform Gaussian noise case. 54

4.5 Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different τ values and SNRs, with $\rho = 0.7$- Non-uniform Gaussian noise case. 55

4.6 Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), while directly applying SQP to (4.12), for different ρ and SNRs- Uniform Gaussian noise case. 57

4.7 var$\{\phi_{en}\}$ and mean$\{\phi_{en}\}$ of various HU algorithms, when $\rho = 0.6$ and SNR=25 dB, for 50 independent data sets. For each data set the algorithm is applied 10 times, each time with its inherent initialization. 59

4.8 Mean-removed spectral angles ϕ (degrees) between library spectra and endmembers estimated by RMVES, MVES, and VCA algorithms. . . 66

5.1 Pseudocode for GENE-CH and GENE-AH algorithms. 77

5.2 Pseudocode for p-norm based pure pixel (TRI-P) algorithm. 85

5.3 Average ϕ_{en} (degrees) and average computation time T (secs) over the various EEAs for different purity levels (ρ) and SNRs (uniform Gaussian noise case), L =1000, N =8. 88

5.4 Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable) and SNRs- Uniform Gaussian noise case. $N_{\max} = 25$, true $N = 8$, $L = 5000$, $M = 224$, and $\rho = 1$. . 91

5.5 Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable) and SNRs- Non-uniform Gaussian noise case. $N_{\max} = 25$, true $N = 8$, $L = 5000$, $M = 224$, and $\rho = 1$. 92

5.6 Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable), for various purity levels. Uniform Gaussian noise case, true $N = 8$, SNR=30 dB, $N_{\max} = 25$, $L = 5000$, and $M = 224$. 93

5.7 Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable), for various number of endmembers. Uniform Gaussian noise case, $\rho = 1$, SNR=30 dB, $N_{\max} = 25$, $L = 5000$, and $M = 224$. 94

5.8 Number of endmembers estimated by various algorithms, where NA denotes "non-applicable" and $\star$ denotes "out of memory" encountered in Matlab. 95

5.9 Mean-removed spectral angles ϕ (degrees) between library spectra and endmembers estimated by TRI-P ($p = 2$). 95

List of Notations

$\mathbb{R}$, $\mathbb{R}^N$, $\mathbb{R}^{M\times N}$	Set of real numbers, $N\times 1$ vectors, $M\times N$ matrices
$\mathbb{R}_+$, $\mathbb{R}_+^N$, $\mathbb{R}_+^{M\times N}$	Set of non-negative real numbers, $N\times 1$ vectors, $M\times N$ matrices
$\|\mathbf{x}\|_p$	p-norm of a vector $\mathbf{x}$
$\mathbf{1}_N$	$N\times 1$ vector with all the components equal to unity
$\mathbf{0}$	An all-zero vector of proper dimension
$\mathbf{e}_i$	Unit vector of proper dimension with the ith entry equal to 1
$\succeq$	Componentwise inequality
$\mathbf{I}_N$	$N\times N$ identity matrix
$\mathbf{X}^\dagger$	Moore-Penrose inverse or pseudo-inverse of matrix $\mathbf{X}$
range($\mathbf{X}$)	Range space of the matrix $\mathbf{X}$
$\mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma})$	Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$
diag$\{x_1,\ldots,x_N\}$	an $\mathbb{R}^{N\times N}$ Diagonal matrix with diagonal entries $x_1,\ldots,x_N$
det($\mathbf{X}$)	Determinant of a square matrix $\mathbf{X}$
sign($\mathbf{x}$)	A vector whose elements are the signs of the elements in the vector $\mathbf{x}$

$\lvert\mathbf{x}\rvert$	A column vector whose elements are the absolute values of the individual elements in $\mathbf{x}$
$[\mathbf{x}]_i$	ith element of the vector $\mathbf{x}$
$[\mathbf{X}]_{i,:}$	ith row vector of $\mathbf{X}$.
X_i	(i,i)th diagonal element of the matrix $\mathbf{X}$

Chapter 1

Introduction

Hyperspectral Imaging (HI) is a crucial technique to identify the materials and their composition in an area by exploiting the spectral diversity of the observed hyperspectral data. Areas in which HI is employed are diverse and they include mineral identification [1], space object identification [2], analytical chemistry [3], retinal analysis [4], and many others. In this thesis the focus is on the analysis of HI taken with the purpose of mineral identification and quantification, wherein a hyperspectral sensor (usually located on an aircraft or satellite) records the electromagnetic scattering patterns of materials present in an area, over hundreds of spectral bands that range from visible to near-infrared wavelength region. The limited spatial resolution of the sensor used for hyperspectral imaging and the presence of noise in the measured hyperspectral data demand an effective hyperspectral unmixing (HU) scheme to extract the underlying endmember signatures (or simply endmembers) and the associated abundance maps (or abundance fractions) distributed over a scene of interest. The endmember signature corresponds to the reflection pattern of a mineral in different wavelengths and the abundance fraction is the fractional distribution of a mineral over the given scene. Assuming a prior knowledge of the number of endmembers, the design of HU algorithms often (but not always) involves dimension reduction as the

preprocessing step. The dimension reduction algorithms are intended to reduce the complexity of the unmixing algorithm that will be used in the sequel and to some extent they also aid in mitigating the noise effects in the data cloud. Conventional dimension reduction algorithms for hyperspectral data typically include principal component analysis (PCA) [5] and maximum noise fraction (MNF) [6]. A detailed survey of various dimension reduction algorithms can be found in [7].

Conventional HU algorithms are based on linear mixing model (to be explained later in Chapter 2) and can be classified into two main categories. Algorithms in the first category are based on the existence of pure pixels (pixels that are fully contributed by a single endmember) in the given hyperspectral observations, and those in the second category may not require the existence of pure pixels. Figure 1.1 illustrates the notion of pure pixels and mixed pixels in a hyperspectral imaging scenario. The red pixel corresponds to a mixed pixel (contributed by land, vegetation and water) and the blue pixel corresponds to a pure pixel (contributed by only water). Pure-pixel based algorithms include pixel purity index (PPI) [8], N-finder (N-FINDR) [9], simplex growing algorithm (SGA) [10] [11], vertex component analysis (VCA) [12], and alternating volume maximization (AVMAX) [13] [14], to name a few. Among which N-FINDR, SGA and AVMAX algorithms are based on Winter's belief. Winter's belief states that the vertices of the maximum volume simplex inside the observation constructed data cloud will yield high fidelity estimates of the endmember signatures. All the pure-pixel based algorithms aim to find the pure pixels in the given observations, either directly or indirectly, and can identify only the endmember signatures. Hence, the pure-pixel based algorithms are also called as endmember extraction algorithms (EEAs). However, the fully constrained least squares (FCLS) algorithm [15] can be used to estimate the abundance maps from the endmember estimates and the observations, and thereby enabling HU via EEAs. Algorithms in the second category include minimum volume transform (MVT) [16], iterated constrained endmembers

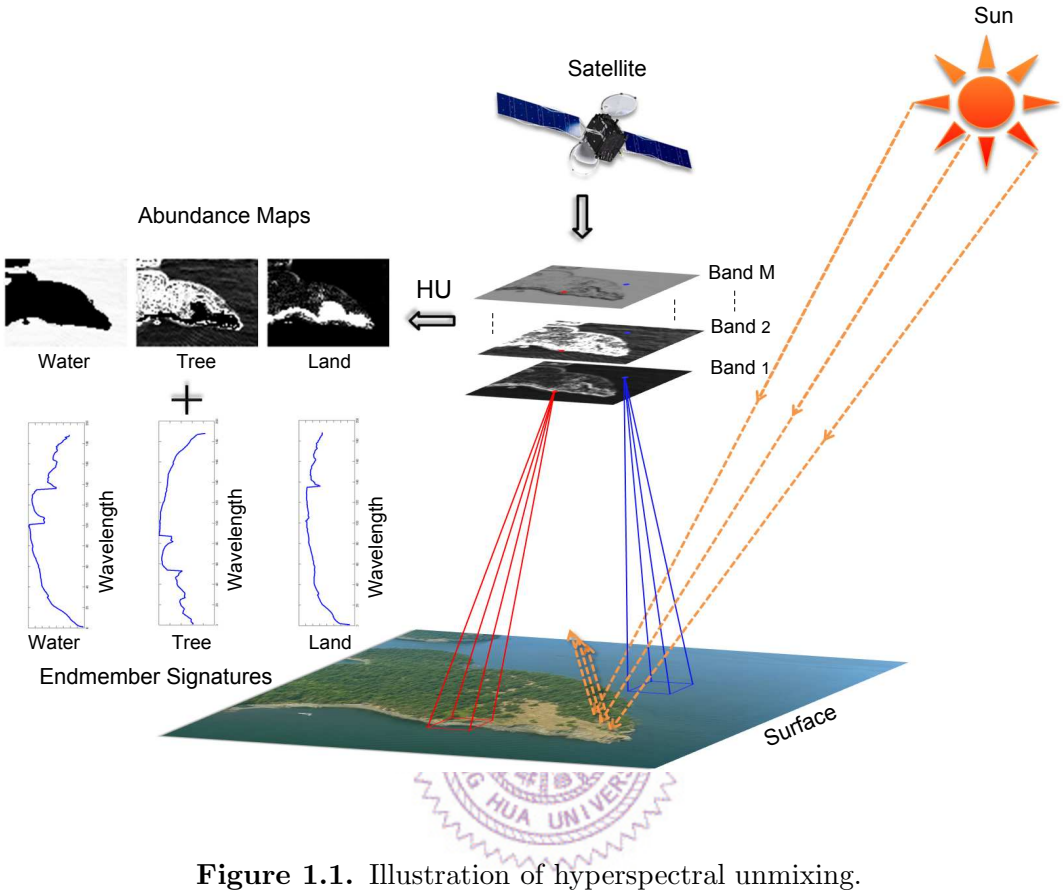


Figure 1.1. Illustration of hyperspectral unmixing.

(ICE) [17], piecewise convex endmember (PCE) [18], alternating projected subgradients (APS) [19], minimum volume constrained non-negative matrix factorization (MVC-NMF) [20], minimum dispersion constrained non-negative matrix factorization (MDC-NMF) [21], minimum volume simplex analysis (MVSA) [22] and minimum volume enclosing simplex (MVES) [23], to name a few. The MVSA and MVES algorithms are directly based on Craig's criterion [16], which states that the vertices of the minimum volume simplex enclosing the data cloud should yield high fidelity estimates of its endmember signatures. Apart from the above mentioned algorithms, there are certain other categories of algorithms that are based on spatial processing and morphological operators [24] [25]. For a recent survey on the various existing

unmixing algorithms, please refer to [26].

Generally speaking, HU algorithms that do not require the existence of pure pixels in the data set are more computationally complicated than pure-pixel based algorithms. On the other hand, the presence of pure pixels cannot be guaranteed for real hyperspectral data that are acquired under poor spatial resolutions. Hence there exists a tradeoff between the two categories of algorithms, in terms of computational tractability and accuracy. Nevertheless, the performance of all the HU algorithms degrades when the observations are noisy.

In recent years, several concurrent works on linear unmixing of noisy observations have been reported. Those works include joint Bayesian algorithm (JBA) [27] and simplex identification via split augmented Lagrangian (SISAL) [28]. The JBA [27] is a statistical HU algorithm that assumes conjugate prior distributions for the abundances and endmember parameters. The endmember signatures are estimated by using a hierarchical Bayesian model, which generates the posterior distributions of both the abundances and the endmember parameters. The SISAL algorithm [28] is based on Craig's criterion, and it employs variable splitting and augmented Lagrangian approach to estimate the minimum volume simplex. Further, it uses soft constraints to mitigate the effects of the outlier pixels.

In this thesis we consider a noisy linear mixing model for the hyperspectral data, where the noise is assumed to be uniform/non-uniform additive Gaussian. An existing dimension reduction method, called affine set fitting [23], has been suitable modified so as to account for the presence of uniform/non-uniform additive Gaussian noise. The linear mixing model and the dimension reduction technique are presented in Chapter 2.

In Chapter 3, we develop a robust version of the Winter's criterion based AVMAX algorithm [13] [14], namely robust AVMAX (RAVMAX) algorithm. The RAVMAX algorithm accounts for the noise effects by employing chance constraints or probability

constraints [29], with a design parameter η that controls the volume of the resulting simplex. We first reformulate the AVMAX subproblems into equivalent problems to which the chance constraints can be suitably applied, and then due to the appropriate range of η, the problems become second-order cone programming (SOCP) problems. Hence, the chance constrained problem can be efficiently solved by any available convex optimization solvers such as `SeDuMi` [30] and `CVX` [31], in an alternating fashion. Some Monte-Carlo simulations and real data experiments are shown to demonstrate the efficacy of the proposed RAVMAX algorithm, in comparison with the conventional pure-pixel based algorithms, including its predecessor AVMAX algorithm.

For hyperspectral data set, where pure pixels cannot be guaranteed for all endmembers, robust algorithm based on Craig's criterion are preferred. Hence, in Chapter 4, we propose a robust minimum volume enclosing simplex (RMVES) algorithm derived from the Craig's criterion based MVES algorithm [23]. Specifically, we formulate the RMVES problem by incorporating chance constraints [29] into Craig's criterion based MVES problem [23], so as to deal with the effect of random noise. Under the Gaussian noise assumption, the chance constrained RMVES problem can be conveniently formulated into a deterministic non-linear program, where the pre-assigned probability given in the chance constraints when coupled with the noise covariance, controls the volume of Craig's simplex. We then handle the resulting problem by alternating optimization, where each subproblem involved therein is specifically handled by readily available sequential quadratic programming (SQP) solver, namely *fmincon* [32]. Some existing benchmark HU methods, including MVES algorithm, are compared with the proposed RMVES algorithm through Monte-Carlo simulations, so as to demonstrate its efficacy. The RMVES algorithm along with MVES and VCA is applied to real hyperspectral data and some inferences regarding the obtained endmember signatures and abundance maps are also discussed.

It should be mentioned that all the above mentioned HU algorithms including

the proposed robust algorithms assume that the number of substances (or endmembers) present in that area is given *a priori*, which however is unknown in practice. Although a number of early efforts have been made in developing algorithms to estimate the number of endmembers (also known as rank estimation [33] [34] or model order selection [35] [36]), the problem of estimating the number of endmembers remains one of the greatest challenges. The vast majority of the existing methods for estimating the number of endmembers can be classified into two categories: information theoretic criteria based methods and eigenvalue thresholding methods. Methods falling into the group of information theoretic criteria includes Akaike's information criterion (AIC) [37], minimum description length (MDL) [38], and Bayesian information criterion (BIC) [39] [40], to name a few. These criteria generally consist of two additive terms: a negative data log-likelihood term and a penalty term. As the number of endmembers increases, the value of the negative data log-likelihood term decreases, whereas the value of the penalty term increases. The best estimate of the number of endmembers is the one that yields the minimum value of the criteria. In AIC, MDL and BIC, the data log-likelihoods in the criteria are identical, but how each method penalizes the overestimation of the number of endmembers makes the difference. Since the criteria require the prior knowledge of the mixture model or likelihood function, the estimation results may suffer from model mismatch errors resulting from incorrect prior information. It has been shown in [41] that the results of AIC and MDL when applied to hyperspectral data are seriously overestimated due to the invalid Gaussian distribution assumption made on the abundances [42].

Methods belonging to the group of eigenvalue thresholding (either explicitly or implicitly) include principal component analysis (PCA) based approaches [43], Neyman-Pearson detection theory based method [44] (also referred to as virtual dimensionality (VD) in [41]), and hyperspectral signal subspace identification by minimum error (HySiMe) [45], to name a few. PCA-based approaches aim to determine the cut-

off threshold between the eigenvalues caused by signals and noise, but the variation between the two adjacent eigenvalues may not be significant in a practical scenario, thereby easily leading to estimation errors in number of endmembers estimation, as demonstrated in [45]. The Neyman-Pearson detection theory based method was first proposed by Harsanyi, Farrand, and Chang (HFC) in 1993, and is termed as HFC method. The HFC method uses Neyman-Pearson detector for a binary hypothesis testing problem, built on the differences in eigenvalues of the sample correlation and sample covariance matrices. The HFC method was later revisited by incorporating the concepts of VD and noise-prewhitening step [41]. HySiMe [45] utilizes a minimum mean square error criterion to estimate the signal subspace in hyperspectral images. It starts with estimating the signal and the noise correlation matrices, and then selects the subset of eigenvectors that best represent the signal subspace in the least squared error sense. In addition to the two aforementioned categories, a Markov chain Monte Carlo (MCMC) based approach was also proposed to estimate the number of endmembers [46]. While the applicability of the algorithm is restricted to the data with a smaller number of endmembers/pixels due to high computational demand, it provides Bayesian estimation of the number of endmembers, with theoretical basis.

In Chapter 5, we propose two hyperspectral data geometry based algorithms for estimating the number of endmembers, namely geometry based estimation of number of endmembers - convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm. The proposed GENE algorithms (GENE-CH and GENE-AH) exploit successive estimation property of a pure-pixel based EEA, and aim to decide when the EEA should stop estimating the next endmember signature. The GENE-CH and GENE-AH algorithms are devised based on the data geometry fact that all the observed pixel vectors should lie in the convex hull (CH) and affine hull (AH) of the endmember signatures, respectively. Since the EEAs identify endmember estimates from the set of observed pixel vectors, the fact pertaining to the data geometry also implies that the

current endmember estimate should lie in the CH/AH of the previously found endmembers when the current endmember estimate is obtained for an overly estimated number of endmembers. In the noisy scenario, the decision of whether the current endmember estimate is in the CH/AH of the previously found endmembers can be formulated as a binary hypothesis testing problem, which we solve by Neyman-Pearson detection theory. The performances of the proposed GENE algorithms depend on the accuracy of pure pixel indices identification of an EEA. However, EEAs that can be used in conjunction with the GENE algorithms are preferred to have the following properties for better estimation accuracy and efficiency:

- Reliability- The EEA can reliably find a set of true endmembers provided that pure pixels exist in the hyperspectral data. Specifically, for the noiseless case, its endmember identifiability can be guaranteed.
- Reproducibility- The EEA provides reproducible endmember estimates for a given hyperspectral data set without need of any random initialization.
- Successive estimation- The EEA estimates the endmembers successively.
- Computational efficiency- The EEA has low computational complexity (as the overall complexity of the GENE algorithms also depends on the complexity of the EEA employed).

Therefore, we also propose a reliable, reproducible, and computationally efficient, successive EEA, called p-norm based pure pixel identification (TRI-P, abbreviated for Triple-P) algorithm. The TRI-P algorithm basically consists of two process: Firstly, the data are projected onto a subspace orthogonal to already found endmember signatures (affine transformation), and secondly, maximum p-norm of the observed pixel vectors is used to identify a new pure pixel (one-dimensional pixel search). The performances of the proposed TRI-P and GENE algorithms are demonstrated through

Monte Carlo simulations for various scenarios and by real data experiments.

The dissertation is organized as follows. The linear mixing model and the dimension reduction technique are detailed in Chapter 2. The first of the proposed robust algorithms, namely RAVMAX is introduced in Chapter 3. Then, in Chapter 4, the craig's criterion based RMVES algorithm is presented. The proposed hyperspectral data geometry based approach for estimating the number of endmembers, and the associated successive EEA, namely TRI-P algorithm are presented in Chapter 5. Finally, some conclusions and future directions are provided in Chapter 6.

Chapter 2

Linear Mixing Model for Hyperspectral Data

The linear mixing model is the one in which each pixel vectors (or simply pixels, for convenience) in the hyperspectral observations is obtained via single reflection of the endmembers present in that location. Such a linear mixing model is conventionally used in HU [1,9,12,13,47,48]. We consider a scenario in which a hyperspectral sensor measures solar electromagnetic radiation from N distinct substances. Owing to low spatial resolution, each observed pixel vector represents a mixture of multiple distinct substances. Hence, each pixel vector of the hyperspectral images measured over M spectral bands can be represented by the following $M \times N$ linear mixing model

$$\mathbf{y}[n] = \mathbf{x}[n] + \mathbf{w}[n], \tag{2.1}$$

$$\mathbf{x}[n] = \mathbf{A}\mathbf{s}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i, \ \forall n = 1, \ldots, L. \tag{2.2}$$

In (2.1), $\mathbf{y}[n] = [\ y_1[n], \ldots, y_M[n]\ ]^T$ represents the nth observed pixel vector comprising M spectral bands, $\mathbf{x}[n] = [\ x_1[n], \ldots, x_M[n]\ ]^T$ corresponds to its noise-free counterpart, and $\mathbf{w}[n] = [\ w_1[n], \ldots, w_M[n]\ ]^T$ is the noise vector. In (2.2),

$\mathbf{A} = [\ \mathbf{a}_1, \ldots, \mathbf{a}_N\ ] \in \mathbb{R}^{M \times N}$ denotes the endmember signature matrix with the ith column vector $\mathbf{a}_i$ being the ith endmember signature, $\mathbf{s}[n] = [\ s_1[n], \ldots, s_N[n]\ ]^T \in \mathbb{R}^N$ is the nth abundance vector comprising N fractional abundances from which the ith abundance map can be defined as $\boldsymbol{s}_i = [\ s_i[1], \ldots, s_i[L]\ ]^T$, and L is the total number of observed pixels. The noise vector $\mathbf{w}[n]$ considered in the signal model (2.1) is zero-mean, uniform/non-uniform additive Gaussian noise vector i.e., $\mathcal{N}(\mathbf{0}, \mathbf{D})$, where $\mathbf{D} = \text{diag}(\sigma_1^2, \ldots, \sigma_M^2) \in \mathbb{R}_+^{M \times M}$ and σ_i^2 denotes the noise variance in the ith spectral band. If $\sigma_i^2 = \sigma_j^2,\ \forall i \neq j$, then it is called uniform Gaussian noise, else it is called non-uniform Gaussian noise. The additive Gaussian noise is a reasonable assumption and is widely used in designing HU algorithms [1, 12, 27].

Assuming prior knowledge about the number of endmembers N, we aim to estimate the endmember signature matrix $\mathbf{A}$ and the abundances $\mathbf{s}[1], \ldots, \mathbf{s}[L]$ from the noisy pixels $\mathbf{y}[1], \ldots, \mathbf{y}[L]$, under the following general assumptions [14] [48] associated with (2.2):

(A1) (Non-negativity condition) $s_i[n] \geq 0\ \forall i, n$.

(A2) (Full additivity condition) $\sum_{i=1}^{N} s_i[n] = 1\ \forall n$.

(A3) $\min\{L, M\} \geq N$ and $\mathbf{A}$ is of full column rank.

(A4) (Pure pixel assumption) There exists at least an index set $\{l_1, \ldots, l_N\}$ such that $\mathbf{x}[l_i] = \mathbf{a}_i$, for $i = 1, \ldots, N$.

(A1) and (A2) hold true in hyperspectral imaging because of the fact that the abundances are nothing but fractional proportions of the endmembers [1], and hence must be non-negative, and the proportions at every pixel location should sum to one. In addition, for hyperspectral images, the number of pixels and that of spectral bands involved are larger than the number of endmembers and the endmember signature of each material is unique, which justifies that (A3) is naturally satisfied due to inherent

characteristics of hyperspectral images and endmembers. Assumption (A4) implies that there exists at least one (location unknown) pixel index for each endmember such that the associated observed pixel vector will be fully contributed by that single endmember, and it usually holds true for hyperspectral images taken with a reasonable spatial resolution [49].

In the ensuing development, we employ two convex analysis concepts, namely *affine hull* and *convex hull* [29]. While the affine hull is employed for dimension reduction, the convex hull is used to infer the geometry of the observations. For ease of later use, they are defined below:

- The *affine hull* of $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is defined as

$$\text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \middle| \mathbf{1}_N^T \boldsymbol{\theta} = 1, \boldsymbol{\theta} \in \mathbb{R}^N \right\}, \tag{2.3}$$

  where $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^T$. An affine hull can be represented as:

$$\text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \mathcal{A}(\mathbf{C}, \mathbf{d}) = \left\{ \mathbf{x} = \mathbf{C}\boldsymbol{\alpha} + \mathbf{d} \middle| \boldsymbol{\alpha} \in \mathbb{R}^P \right\} \tag{2.4}$$

  for some (non-unique) $\mathbf{d} \in \mathbb{R}^M$ and $\mathbf{C} \in \mathbb{R}^{M \times P}$, where $P \leq N - 1$ is the affine dimension. If $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ is affinely independent (i.e., the vectors $\mathbf{a}_1 - \mathbf{a}_N, \ldots, \mathbf{a}_{N-1} - \mathbf{a}_N$ are linearly independent), then the affine dimension $P = N - 1$.

- The *convex hull* of the vectors $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is defined as

$$\text{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \left\{ \mathbf{x} = \sum_{i=1}^{N} \theta_i \mathbf{a}_i \middle| \mathbf{1}_N^T \boldsymbol{\theta} = 1, \boldsymbol{\theta} \in \mathbb{R}_+^N \right\}. \tag{2.5}$$

  A convex hull, $\text{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ is called an $N - 1$ dimensional simplex in $\mathbb{R}^M$ if $\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is affinely independent and is called a *simplest simplex*

when $M = N - 1$. The simplex $\text{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$ has only N extreme points, and they are $\mathbf{a}_1, \ldots, \mathbf{a}_N$.

2.1 Data Preprocessing

In this section, the dimension reduction technique employed for the observations corrupted by uniform and non-uniform Gaussian noise is discussed. For the noisy scenario, we employ the dimension reduction by using a variant of affine set fitting procedure introduced in [23]. To begin with, the affine set fitting procedure for noise-free data is reviewed below for convenience.

It can be readily inferred from the signal model (2.2), (A2), and (A3) that

$$\mathbf{x}[n] \in \text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}. \tag{2.6}$$

Further, it is shown in [23] that the affine hull of the observations is the same as the affine hull of the endmember signatures, i.e.,

$$\text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \text{aff}\{\mathbf{x}[1], \ldots, \mathbf{x}[L]\}, \tag{2.7}$$

which in turn can be represented by the affine set fitting parameters $\mathbf{C}$ and $\mathbf{d}$ in (2.4). Such representation aids in dimension reduction of the observed data. The dimension-reduced pixel vector $\tilde{\mathbf{x}}[n]$ can be obtained by the following affine transformation of $\mathbf{x}[n]$:

$$\tilde{\mathbf{x}}[n] = \mathbf{C}^T(\mathbf{x}[n] - \mathbf{d}) \in \mathbb{R}^{N-1}, \ \forall n = 1, \ldots, L, \tag{2.8}$$

where $(\mathbf{C}, \mathbf{d})$ is an affine set fitting solution given by [23]

$$\mathbf{d} = \frac{1}{L}\sum_{n=1}^{L}\mathbf{x}[n], \tag{2.9}$$

$$\mathbf{C} = [\ \boldsymbol{q}_1(\mathbf{U}_x\mathbf{U}_x^T), \ldots, \boldsymbol{q}_{N-1}(\mathbf{U}_x\mathbf{U}_x^T)\ ], \tag{2.10}$$

where

$$\mathbf{U}_x = [\ \mathbf{x}[1] - \mathbf{d}, \ldots, \mathbf{x}[L] - \mathbf{d}\ ] \in \mathbb{R}^{M\times L}, \tag{2.11}$$

is the mean removed data matrix and $\boldsymbol{q}_i(\mathbf{R})$ denotes the unit-norm eigenvector associated with the ith principal eigenvalue of the matrix $\mathbf{R}$.

It should be mentioned that under (A2) and (A3), the above dimension reduction procedure is a lossless transformation in the absence of noise, and the dimension of the data is reduced from M to $N-1$. An illustration of dimension reduction via the affine set fitting procedure is given in Figure 2.1, where $N = 3$.

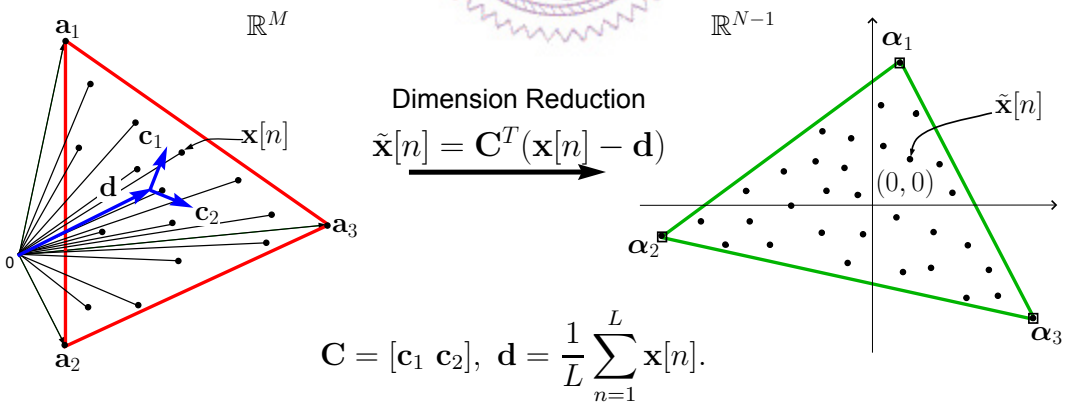


Figure 2.1. Illustration of dimension reduction by affine set fitting for N=3, where the geometric center $\mathbf{d}$ of the data cloud $\mathbf{x}[n]$ in M-dimensional space maps to the origin in the $(N-1)$-dimensional space which is also the geometric center of the dimension-reduced data cloud $\tilde{\mathbf{x}}[n]$.

Since $\sum_{j=1}^{N} s_j[n] = 1$ [(A2)], it follows by substituting the noise-free signal model

(2.2) into (2.8) that

$$\tilde{\mathbf{x}}[n] = \sum_{j=1}^{N} s_j[n]\boldsymbol{\alpha}_j, \ \forall n = 1, \ldots, L, \tag{2.12}$$

where

$$\boldsymbol{\alpha}_j = \mathbf{C}^T(\mathbf{a}_j - \mathbf{d}) \in \mathbb{R}^{N-1} \tag{2.13}$$

is the jth dimension-reduced endmember signature. Moreover, due to $s_i[n] \geq 0$ [(A1)], it can be seen that

$$\tilde{\mathbf{x}}[n] \in \text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} \subset \mathbb{R}^{N-1}, \ \forall n \tag{2.14}$$

and $\text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ is a simplest simplex. This can be regarded as the outcome of the fact that the affine transformation of a simplex is also a simplex, as one can infer from (2.2) that for all n, $\mathbf{x}[n] \in \text{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\}$, and $\text{conv}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} \subset \mathbb{R}^M$ is itself a simplex [by (A3)]. The relation between the dimension-reduced endmember $\boldsymbol{\alpha}_i$ and the true endmember $\mathbf{a}_i$ is given by

$$\mathbf{a}_i = \mathbf{C}\boldsymbol{\alpha}_i + \mathbf{d}, \ i = 1, \ldots, N. \tag{2.15}$$

It is worth mentioning that from (2.12) and under (A4), we have

$$\tilde{\mathbf{x}}[l_i] = \boldsymbol{\alpha}_i, \ \forall i = 1, \ldots, N. \tag{2.16}$$

However, when the data are corrupted by noise, the observations can no longer lie in a single specific affine set. Accordingly, for the noisy scenario, the approximate affine set fitting parameters ($\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$) can be obtained as explained below.

Similar to (2.11), by (2.1) the mean removed noisy data matrix is

$$\mathbf{U}_y = [\ \mathbf{y}[1] - \mathbf{d}, \ldots, \mathbf{y}[L] - \mathbf{d}\ ] = \mathbf{U}_x + \mathbf{W} \in \mathbb{R}^{M \times L}, \tag{2.17}$$

where $\mathbf{W} = [\ \mathbf{w}[1], \ldots, \mathbf{w}[L]\ ]$. Then,

$$\begin{aligned}\mathbf{U}_y\mathbf{U}_y^T &= \mathbf{U}_x\mathbf{U}_x^T + \mathbf{U}_x\mathbf{W}^T + \mathbf{W}^T\mathbf{U}_x + \mathbf{W}\mathbf{W}^T \\ &\cong \mathbf{U}_x\mathbf{U}_x^T + \mathbf{W}\mathbf{W}^T \ = \mathbf{U}_x\mathbf{U}_x^T + L\widehat{\mathbf{D}},\end{aligned} \tag{2.18}$$

which is because the matrices $\mathbf{U}_x\mathbf{W}^T$ and $\mathbf{W}\mathbf{U}_x^T$ asymptotically approach zero matrix for large L (since the noise is zero-mean and independent of the noise-free observations), and $\widehat{\mathbf{D}}$ is defined as

$$\widehat{\mathbf{D}} = \frac{1}{L}\mathbf{W}\mathbf{W}^T, \tag{2.19}$$

which is actually an estimate of the noise covariance matrix $\mathbf{D}$. Replacing $\mathbf{x}[n]$ with $\mathbf{y}[n]$ in (2.9), and by replacing $\mathbf{U}_x\mathbf{U}_x^T$ with $\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}}$ (from (2.18)) in (2.10), respectively, we have the following equations for $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$ as approximations for $\mathbf{C}$ and $\mathbf{d}$, respectively:

$$\widehat{\mathbf{d}} = \frac{1}{L}\sum_{n=1}^{L}\mathbf{y}[n] = \frac{1}{L}\sum_{n=1}^{L}\mathbf{x}[n] + \frac{1}{L}\sum_{n=1}^{L}\mathbf{w}[n] \cong \mathbf{d}, \tag{2.20}$$

$$\widehat{\mathbf{C}} = [\ \boldsymbol{q}_1(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}}), \ldots, \boldsymbol{q}_{N-1}(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}})\ ] \cong \mathbf{C}. \tag{2.21}$$

Note that, (2.20) and (2.21) hold since the noise is of zero mean such that $\frac{1}{L}\sum_{n=1}^{L}\mathbf{w}[n]$ asymptotically approaches zero vector for large L, and $\widehat{\mathbf{D}}$ approaches the true $\mathbf{D}$ for large L, respectively. In this work, the multiple regression analysis based noise covariance estimation method reported in HySiMe [45] is employed to obtain $\widehat{\mathbf{D}}$. With the given $\widehat{\mathbf{D}}$, the affine set fitting solution $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ of the noisy data (given by (2.20) and (2.21)) serves as an approximation to the true $(\mathbf{C}, \mathbf{d})$ and it asymptotically approaches the true $(\mathbf{C}, \mathbf{d})$ for large L.

The obtained $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{d}}$ are then used to obtain the dimension reduced noisy

observations as given by

$$\begin{aligned}\tilde{\mathbf{y}}[n] = \widehat{\mathbf{C}}^T(\mathbf{y}[n] - \widehat{\mathbf{d}}) &= \widehat{\mathbf{C}}^T(\mathbf{x}[n] - \widehat{\mathbf{d}}) + \widehat{\mathbf{C}}^T\mathbf{w}[n] \\ &\cong \tilde{\mathbf{x}}[n] + \tilde{\mathbf{w}}[n], \end{aligned} \tag{2.22}$$

where

$$\tilde{\mathbf{w}}[n] = \widehat{\mathbf{C}}^T\mathbf{w}[n], \tag{2.23}$$

and the approximation in (2.22) is because the affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ for the noisy data serve as a good approximation to the true $(\mathbf{C}, \mathbf{d})$ (see (2.20) and (2.21)). The dimension reduction technique introduced in this chapter will be used in the ensuing developments.

Chapter 3

Robust Alternating Volume Maximization Algorithm

In this chapter, we propose a robust alternating volume maximization (RAVMAX) algorithm based on our previously proposed AVMAX algorithm developed in [13]. The AVMAX algorithm is based on the Winter's criterion which states that the vertices of the maximum volume simplex inscribed within the data cloud should yield high fidelity estimates of the endmembers. However, in the presence of noise, the vertices of the maximum volume simplex could be far away from true endmembers. Hence, application of Winter's criterion based algorithms to noisy hyperspectral data may result in inaccurate estimation of the endmembers. The proposed RAVMAX algorithm accounts for the noise effects in the hyperspectral data by employing chance constraints in its problem formulation. We first reformulate the AVMAX subproblems into equivalent problems to which the chance constraint can be suitably applied, and then reformulate them as second-order cone programming (SOCP) problems. Hence, the chance constrained problem can be efficiently solved by any convex optimization solvers in an alternating fashion. For ease of understanding the proposed RAVMAX algorithm, let us begin by briefly reviewing the AVMAX algorithm.

3.1 Brief review of AVMAX algorithm

The AVMAX algorithm is derived for the noise-free case and is based on the Winter's unmixing criterion [9], which states that under (A4) the vertices of the maximum volume simplex inside the data cloud (observations) yield high fidelity estimates of the endmember signatures. Based on Winter's criterion, the unmixing problem [13] can be written as:

$$\boxed{\begin{array}{ll} \max\limits_{\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N \in \mathbb{R}^{N-1}} & V(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N) \\ \text{s.t.} & \boldsymbol{\nu}_i \in \text{conv}\{\tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\},\ \forall\ i, \end{array}} \tag{3.1}$$

where $\tilde{\mathbf{x}}[n],\ \forall n = 1,\ldots,L$, is defined in (2.8), and $V(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N)$ is the volume of the $(N-1)$-dimensional simplex $\text{conv}\{\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N\}$ in $\mathbb{R}^{N-1}$ and is given by [50],

$$V(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N) = \frac{|\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N))|}{(N-1)!}, \tag{3.2}$$

where

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

By letting $\tilde{\mathbf{X}} = [\ \tilde{\mathbf{x}}[1],\ldots,\tilde{\mathbf{x}}[L]\ ] \in \mathbb{R}^{(N-1)\times L}$ and by (2.5), problem (3.1) can be expressed as

$$\begin{array}{ll} \max\limits_{\substack{\boldsymbol{\nu}_i \in \mathbb{R}^{N-1} \\ \boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N \in \mathbb{R}^L}} & |\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1,\ldots,\boldsymbol{\nu}_N))| \\ \text{s.t.} & \boldsymbol{\nu}_i = \tilde{\mathbf{X}}\boldsymbol{\theta}_i,\quad \boldsymbol{\theta}_i \succeq \mathbf{0},\quad \mathbf{1}_L^T\boldsymbol{\theta}_i = 1\ \forall\ i. \end{array} \tag{3.3}$$

Though the constraints of (3.3) are convex, the non-convexity of the objective function makes the problem difficult to solve. The problem may be handled in a convenient manner by the idea of cofactor expansion and alternating optimization. The cofactor

expansion of the objective function in (3.3) along the jth column is given by

$$\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)) = \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}), \tag{3.4}$$

where $\mathbf{b}_j = [(-1)^{i+j} \det(\boldsymbol{\mathcal{V}}_{ij})]_{i=1}^{N-1} \in \mathbb{R}^{N-1}$ and the term $\boldsymbol{\mathcal{V}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ is a submatrix of $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ with the ith row and jth column removed. We then consider the partial maximization of (3.3) with respect to $\boldsymbol{\nu}_j$ and $\boldsymbol{\theta}_j$, while fixing $\boldsymbol{\nu}_i$ and $\boldsymbol{\theta}_i$ for all $i \neq j$. The problem (3.3) then becomes

$$\begin{aligned} \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \left| \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}) \right| \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \quad \boldsymbol{\theta}_j \succeq \mathbf{0}, \quad \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \tag{3.5}$$

The partial maximization problem (3.5) can be decomposed into the following two linear programs:

$$\begin{aligned} p^\star = \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \quad \boldsymbol{\theta}_j \succeq \mathbf{0}, \quad \mathbf{1}_L^T \boldsymbol{\theta}_j = 1, \end{aligned} \tag{3.6}$$

$$\begin{aligned} q^\star = \min_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \boldsymbol{\nu}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\nu}_j = \tilde{\mathbf{X}} \boldsymbol{\theta}_j, \quad \boldsymbol{\theta}_j \succeq \mathbf{0}, \quad \mathbf{1}_L^T \boldsymbol{\theta}_j = 1. \end{aligned} \tag{3.7}$$

The optimal solution of (3.5) is that of (3.6) if $|p^\star| > |q^\star|$, and that of (3.7) if $|q^\star| > |p^\star|$. This procedure of alternating optimization is performed for all the N columns (one iteration) and the relative change in the volume of the updated $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ is compared with a given threshold. If it exceeds the threshold, we continue with the next updating iteration, else we conclude that the current updated $\boldsymbol{\nu}_j$s are optimum.

Once the optimal solution of (3.3), denoted by $\boldsymbol{\nu}_1^\star, \ldots, \boldsymbol{\nu}_N^\star$ is obtained, the endmember estimates can be recovered by using, $\widehat{\mathbf{a}}_i = \mathbf{C}\boldsymbol{\nu}_i^\star + \mathbf{d}$ (by (2.15)), for all i. It should be recalled that the Winter's criterion based AVMAX algorithm is based on noise-free data. However, when the data is corrupted by noise then the endmembers estimated as extreme points of the Winter's simplex, may be far away from the true endmembers, as illustrated in Figure 3.1.

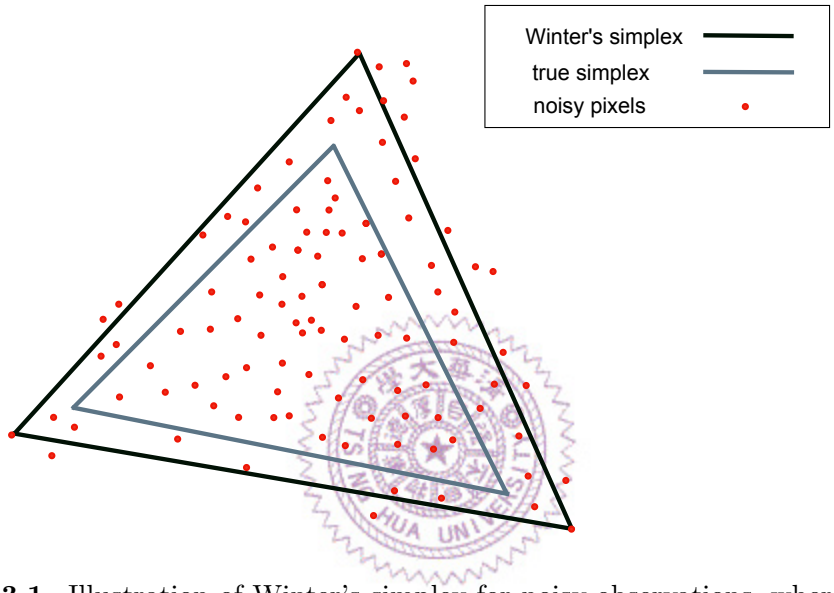


Figure 3.1. Illustration of Winter's simplex for noisy observations, where $N = 3$.

Therefore, we aim to make AVMAX more robust against noise effects by accounting for the randomness in the data, by using chance constraints.

3.2 Robust AVMAX Formulation and Algorithm

We first do some reformulation to (3.6) and (3.7) so that chance constraints can be incorporated into the unmixing problem. Then, we move on to develop a robust version of the AVMAX algorithm.

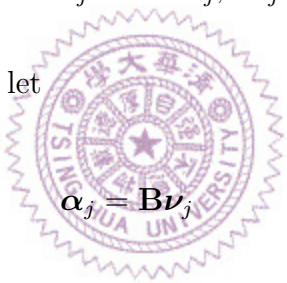
3.2.1 Restructuring the AVMAX algorithm

Let $\mathbf{B} = \text{diag}(\text{sign}(\mathbf{b}_j))$ and $\mathbf{G} = -\mathbf{B}$. Then, we can have $\mathbf{GG} = \mathbf{BB} = \mathbf{I}_{N-1}$, $\mathbf{b}_j^T\mathbf{B} = |\mathbf{b}_j|^T$ and $\mathbf{b}_j^T\mathbf{G} = -|\mathbf{b}_j|^T$. The subproblems (3.6) and (3.7) can then be equivalently written as:

$$\begin{aligned} p^\star = \max_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \mathbf{B}\mathbf{B}\boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \mathbf{B}\boldsymbol{\nu}_j = \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \;\; \boldsymbol{\theta}_j \succeq \mathbf{0}, \;\; \mathbf{1}_L^T\boldsymbol{\theta}_j = 1, \end{aligned} \tag{3.8}$$

$$\begin{aligned} q^\star = \min_{\boldsymbol{\nu}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & \mathbf{b}_j^T \mathbf{G}\mathbf{G}\boldsymbol{\nu}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \mathbf{G}\boldsymbol{\nu}_j = \mathbf{G}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \;\; \boldsymbol{\theta}_j \succeq \mathbf{0}, \;\; \mathbf{1}_L^T\boldsymbol{\theta}_j = 1. \end{aligned} \tag{3.9}$$

Then, by change of variables, we let



$$\boldsymbol{\alpha}_j = \mathbf{B}\boldsymbol{\nu}_j \tag{3.10}$$

and

$$\boldsymbol{\beta}_j = \mathbf{G}\boldsymbol{\nu}_j. \tag{3.11}$$

To facilitate the application of chance constraints to the AVMAX problem, we relax the first equality constraints of (3.8) and (3.9) and thus the corresponding subproblems are given as below:

$$\begin{aligned} p^\star = \max_{\boldsymbol{\alpha}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & |\mathbf{b}_j|^T \boldsymbol{\alpha}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \boldsymbol{\alpha}_j \preceq \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \;\; \boldsymbol{\theta}_j \succeq \mathbf{0}, \;\; \mathbf{1}_L^T\boldsymbol{\theta}_j = 1, \end{aligned} \tag{3.12}$$

$$
\begin{aligned}
q^\star = \quad & \min_{\boldsymbol{\beta}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad -|\mathbf{b}_j|^T \boldsymbol{\beta}_j + (-1)^{N+j} \det(\boldsymbol{\mathcal{V}}_{Nj}) \\
& \text{s.t.} \qquad \boldsymbol{\beta}_j \preceq \mathbf{G}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \quad \boldsymbol{\theta}_j \succeq \mathbf{0}, \quad \mathbf{1}_L^T \boldsymbol{\theta}_j = 1.
\end{aligned} \tag{3.13}
$$

Although we relax the first constraints of the subproblems, we still can show that the optimal solutions of (3.12) and (3.13) are equivalent to that of (3.6) and (3.7), respectively, as proved in the following lemma:

Lemma 1. (Equivalence of subproblems) *The subproblems (3.12) and (3.13) are equivalent to (3.6) and (3.7), respectively.*

Proof: Firstly, it is trivial to show that the objective function of (3.12) is equivalent to that of (3.6) as $\boldsymbol{\alpha}_j = \mathbf{B}\boldsymbol{\nu}_j$, $\mathbf{b}_j^T \mathbf{B} = |\mathbf{b}_j|^T$ and $\mathbf{B}\mathbf{B} = \mathbf{I}$. Next, consider the subproblem (3.12) (after ignoring the constant term in the objective function) and let S denote the constraint set of (3.12). Then we have

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

where $\mathbf{k}_j = [\max_{S_i} [\boldsymbol{\alpha}_j]_i]_{i=1}^{N-1}$ in which $S_i = \{[\boldsymbol{\alpha}_j]_i \leq [\mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j]_i\}$, implying that an optimal solution, denoted by $(\boldsymbol{\alpha}_j^\star, \boldsymbol{\theta}_j^\star)$ will make the equality in $\boldsymbol{\alpha}_j \preceq \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j$ hold (i.e., the constraint will be active). In other words, the optimal solution $(\boldsymbol{\alpha}_j^\star, \boldsymbol{\theta}_j^\star)$ belongs to the set $\{(\boldsymbol{\alpha}_j, \boldsymbol{\theta}_j) \mid \boldsymbol{\alpha}_j = \mathbf{B}\tilde{\mathbf{X}}\boldsymbol{\theta}_j, \boldsymbol{\theta}_j \succeq \mathbf{0}, \mathbf{1}_L^T \boldsymbol{\theta}_j = 1\}$, which is equivalent to the constraint set of (3.6). Hence we can conclude that the subproblems (3.12) and (3.6) are equivalent. By a similar argument the equivalence of (3.13) and (3.7) can be proved. ■

3.2.2 Robust AVMAX Algorithm

Now, we move on to consider the unmixing problem with noisy observations given by (2.1). The corresponding dimension reduced observations are given by (2.22) as

$$\tilde{\mathbf{y}}[n] \cong \tilde{\mathbf{x}}[n] + \widehat{\mathbf{C}}^T \mathbf{w}[n], \tag{3.14}$$

where $\widehat{\mathbf{C}}^T \mathbf{w}[n]$ is a random vector following $\mathcal{N}(\mathbf{0}, \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}})$. In matrix form (considering all the pixels, $n = 1, \ldots, L$), we can write the above equation as: $\tilde{\mathbf{Y}} = \tilde{\mathbf{X}} + \widehat{\mathbf{C}}^T \mathbf{W}$, where $\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}[1], \ldots, \tilde{\mathbf{y}}[L]]$, $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]]$, and $\mathbf{W} = [\mathbf{w}[1], \ldots, \mathbf{w}[L]]$. The first inequality constraint in (3.12) now becomes:

$$\boldsymbol{\alpha}_j \preceq \mathbf{B}(\tilde{\mathbf{Y}} - \widehat{\mathbf{C}}^T \mathbf{W}) \boldsymbol{\theta}_j = \mathbf{B} \tilde{\mathbf{Y}} \boldsymbol{\theta}_j + \mathbf{z}_j, \tag{3.15}$$

where $\mathbf{z}_j \triangleq -\mathbf{B} \widehat{\mathbf{C}}^T \mathbf{W} \boldsymbol{\theta}_j \in \mathbb{R}^{N-1}$ with the distribution $\mathcal{N}(\mathbf{0}, \mathbf{B} \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}} \mathbf{B}^T \|\boldsymbol{\theta}_j\|_2^2)$.

Since the noise-induced vector $\mathbf{z}_j$ is random and unknown, our approach is to consider using chance constraints for (3.15), as shown below

$$\Pr\{[\boldsymbol{\alpha}_j]_i - [\mathbf{B}]_{i,:} \tilde{\mathbf{Y}} \boldsymbol{\theta}_j \leq [\mathbf{z}_j]_i\} \geq \eta, \forall\ i = 1, \ldots, N-1, \tag{3.16}$$

where $0 < \eta < 1$ is a design parameter. A similar equation can be written for the first inequality constraint of (3.13), i.e.,

$$\Pr\{[\boldsymbol{\beta}_j]_i - [\mathbf{G}]_{i,:} \tilde{\mathbf{Y}} \boldsymbol{\theta}_j \leq [\mathbf{z}_j]_i\} \geq \eta, \forall\ i = 1, \ldots, N-1. \tag{3.17}$$

The second-order cone equivalence of a chance constraint has been discussed in [29]. Specifically, for a random variable $\varepsilon \sim \mathcal{N}(\mu, \delta^2)$ and $t \in \mathbb{R}$, one can show that $\Pr(\varepsilon \leq t) \geq \eta$ is true as $t \geq \delta \Phi^{-1}(\eta) + \mu$, where Φ^{-1} is the inverse of the cumulative distribution function (cdf) of the standard normal random variable.

By letting $\mathbf{Q} = \mathbf{B}\widehat{\mathbf{C}}^T\mathbf{D}\widehat{\mathbf{C}}\mathbf{B}^T \in \mathbb{R}^{(N-1)\times(N-1)}$ and applying the above mentioned chance constraint procedure to (3.16) and (3.17), we have

$$\sqrt{Q_i}\|\boldsymbol{\theta}_j\|_2\Phi^{-1}(1-\eta) \geq [\boldsymbol{\alpha}_j]_i - [\mathbf{B}]_{i,:}\tilde{\mathbf{Y}}\boldsymbol{\theta}_j, \tag{3.18}$$

and

$$\sqrt{Q_i}\|\boldsymbol{\theta}_j\|_2\Phi^{-1}(1-\eta) \geq [\boldsymbol{\beta}_j]_i - [\mathbf{G}]_{i,:}\tilde{\mathbf{Y}}\boldsymbol{\theta}_j. \tag{3.19}$$

for all $i = 1, \ldots, N-1$. By replacing the first constraints of (3.12) and (3.13) with (3.18) and (3.19), respectively, the robust AVMAX problem can then be written as:

$$\begin{aligned} p^\star = \max_{\boldsymbol{\alpha}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & |\mathbf{b}_j^T|\boldsymbol{\alpha}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \sqrt{Q_i}\|\boldsymbol{\theta}_j\|_2\Phi^{-1}(1-\eta) \geq [\boldsymbol{\alpha}_j]_i - [\mathbf{B}]_{i,:}\tilde{\mathbf{Y}}\boldsymbol{\theta}_j, \\ & \boldsymbol{\theta}_j \succeq \mathbf{0}, \ \mathbf{1}_L^T\boldsymbol{\theta}_j = 1, \ \forall\, i = 1, 2, \ldots, N-1. \end{aligned} \tag{3.20}$$

$$\begin{aligned} q^\star = \min_{\boldsymbol{\beta}_j \in \mathbb{R}^{N-1}, \boldsymbol{\theta}_j \in \mathbb{R}^L} \quad & -|\mathbf{b}_j^T|\boldsymbol{\beta}_j + (-1)^{N+j}\det(\boldsymbol{\mathcal{V}}_{Nj}) \\ \text{s.t.} \quad & \sqrt{Q_i}\|\boldsymbol{\theta}_j\|_2\Phi^{-1}(1-\eta) \geq [\boldsymbol{\beta}_j]_i - [\mathbf{G}]_{i,:}\tilde{\mathbf{Y}}\boldsymbol{\theta}_j, \\ & \boldsymbol{\theta}_j \succeq \mathbf{0}, \quad \mathbf{1}_L^T\boldsymbol{\theta}_j = 1, \ \forall\, i = 1, 2, \ldots, N-1. \end{aligned} \tag{3.21}$$

The values of η affect the feasible sets of (3.20) and (3.21); specifically, their convexity. The following are the three possible cases: When $\eta > 0.5$ (i.e., $\Phi^{-1}(1-\eta) < 0$), the first constraints of both subproblems are second-order cone constraints and hence subproblems (3.20) and (3.21) are convex. If $\eta = 0.5$ (i.e., $\Phi^{-1}(1-\eta) = 0$), the first constraints of both subproblems reduce to those of the original AVMAX problem (as in (3.6) and (3.7)), i.e., the constraints become linear (convex). Finally, if $\eta < 0.5$ (i.e., $\Phi^{-1}(1-\eta) > 0$), the constraints become non-convex.

To infer the appropriate range of η, let $\mathcal{F}_1(\eta)$ and $\mathcal{F}_2(\eta)$ be the constraint sets of (3.20) and (3.21), respectively. Since $\Phi^{-1}(1-\eta)$ is a monotonically decreasing function of η, it can be easily seen that $\mathcal{F}_i(\eta_1) \subset \mathcal{F}_i(\eta_2),\ i = 1, 2$, where $\eta_1 > \eta_2$, and therefore $|\det(\mathbf{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))| \propto 1/\eta$. This along with the discussion above for $\eta = 0.5$ confirms that to get a smaller simplex than the simplex estimated by Winter's criterion, the appropriate η should be greater than 0.5. The effect of η is illustrated in Figure 3.2. From our extensive numerical experiences we found that for satisfactory performance, the η value should lie between 0.9 and 1, in which case the subproblems (3.20) and (3.21) are convex. Some more technical aspects are discussed as follows.

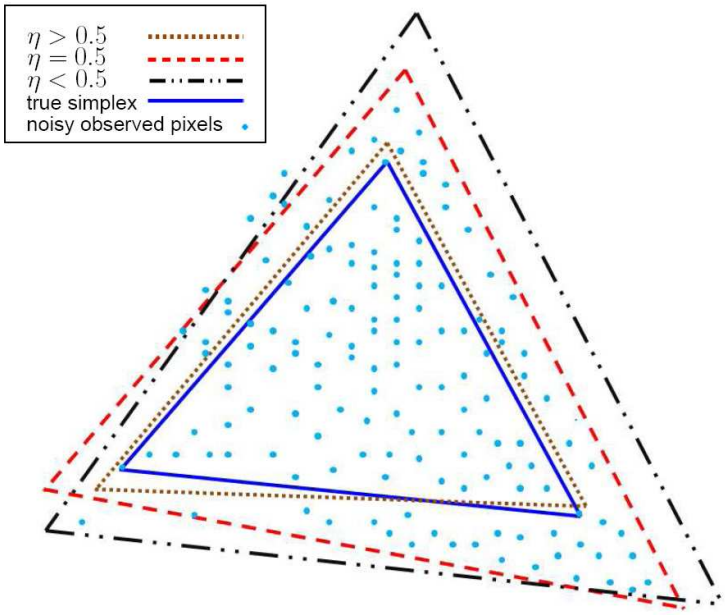


Figure 3.2. Illustration of the effect of η in RAVMAX for $N = 3$.

The subproblems (3.20) and (3.21) are solved in an alternating fashion, similar to the original AVMAX explained in Section 3.1, except for the following difference: after each execution of the subproblems, the corresponding $\boldsymbol{\nu}_j = \mathbf{B}\boldsymbol{\alpha}_j$ (by (3.10)) is obtained if $|p^\star| > |q^\star|$, else $\boldsymbol{\nu}_j = \mathbf{G}\boldsymbol{\beta}_j$ (by (3.11)) is obtained. The proposed RAVMAX algorithm uses any N pixel vectors in $\tilde{\mathbf{X}}$ for its initialization. As the subproblems are convex (for our desired choice, $\eta > 0.5$), they can be solved effectively by using

available convex optimization solvers [31] and [30]. The pseudocode for the RAVMAX algorithm is given in Table 3.1.

3.3 Simulations

This section demonstrates the efficacy of the proposed RAVMAX through comparison with other pure-pixel based HU algorithms. The algorithms considered are N-FINDR [9], VCA [12], and AVMAX [13]. In all these algorithms, FCLS [15] is used to get the abundance maps. The performance of the algorithms under test is evaluated by performing 50 Monte Carlo runs for various purity levels (ρ) and SNRs [23]. A data set with purity level ρ implies the purity factor $1/\sqrt{N} \leq \rho_n = \|\mathbf{s}[n]\|_2 \leq \rho$ [23] for each of its abundance vectors $\mathbf{s}[n]$. The value of ρ_n defines the quantitative dominance of an endmember $\mathbf{a}_i$ in the observed pixel vector $\mathbf{x}[n] = \sum_{i=1}^{N} s_i[n]\mathbf{a}_i, \ \forall n$. The SNR of a data set is defined as

$$\text{SNR} = \sum_{n=1}^{L} \|\mathbf{x}[n]\|_2^2 / ML\sigma^2, \tag{3.22}$$

where σ^2 is the noise variance. The simulation settings are $L = 1000$ (number of pixels), $N = 6$ (number of endmembers) and $M = 224$ (number of observations). In each run, 1000 noise-free observed pixel vectors were synthetically generated following the signal model in (2.2), and the 6 endmembers (i.e., Alunite, Buddingtonite, Calcite, Goethite, Kaolinite, and Muscovite) with 224 bands are selected from USGS library [51], and the abundance vectors $\mathbf{s}[n]$ were generated following Dirichlet distribution $D(\mathbf{s}[n], \boldsymbol{\mu})$ with $\boldsymbol{\mu} = \mathbf{1}_N/N$ [12], for purity levels $\rho = 0.7, 0.85, 1$. In our simulations, the noise covariance matrix is estimated from the observations, using the multiple regression procedure elaborated in HySiMe [45]. The root-mean-square (rms) spectral angles, denoted as ϕ_{en} (ϕ_{ab}) between the true endmembers (abundance maps) and

Table 3.1. Pseudocode for RAVMAX Algorithm.

Given The noisy observed data $\mathbf{y}[n]$, the number of endmembers N, a design parameter $\eta > 0.5$, and a convergence tolerance $\varepsilon > 0$.

Step 1. Estimate the noise covariance matrix (denoted by $\widehat{\mathbf{D}}$) using multiple regression analysis [45].

Step 2. **Dimension reduction**: Obtain the dimension-reduced pixels: $\tilde{\mathbf{y}}[n] = \widehat{\mathbf{C}}^T(\mathbf{y}[n] - \widehat{\mathbf{d}})$ for all n, with the affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ given by (2.21) and (2.20).

Step 3. **Initialization**: Randomly choose any N dimension reduced pixel vectors as $\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N$ and frame the initial matrix

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

Set $j := 1$ and $\varrho := |\det(\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N))|$.

Step 4. Obtain $\mathbf{b}_j = [(-1)^{i+j}\det(\boldsymbol{\mathcal{V}}_{ij})]_{i=1}^{N-1} \in \mathbb{R}^{N-1}$, where $\boldsymbol{\mathcal{V}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ is a submatrix of $\boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$ with the ith row and jth column removed.

Step 5. Find $\mathbf{B} = \mathrm{diag}(\mathrm{sign}(\mathbf{b}_j))$ and let $\mathbf{G} = -\mathbf{B}$. Define $\mathbf{Q} = \mathbf{B}\widehat{\mathbf{C}}^T\mathbf{D}\widehat{\mathbf{C}}\mathbf{B}^T = \mathbf{G}\widehat{\mathbf{C}}^T\mathbf{D}\widehat{\mathbf{C}}\mathbf{G}^T$.

Step 6. Solve the SOCPs (3.20) and (3.21) and obtain their optimal solutions, denoted by $(\boldsymbol{\alpha}_j, \bar{\boldsymbol{\theta}}_j)$ and $(\boldsymbol{\beta}_j, \underline{\boldsymbol{\theta}}_j)$, respectively. Let $p^\star$ and $q^\star$ be the optimal values of (3.20) and (3.21), respectively.

Step 7. If $|p^\star| > |q^\star|$, then update $\boldsymbol{\nu}_j = \mathbf{B}\boldsymbol{\alpha}_j$. Otherwise, update $\boldsymbol{\nu}_j = \mathbf{G}\boldsymbol{\beta}_j$.

Step 8. If (i modulo (N)) $\neq 0$, then $j := j + 1$, and go to **Step 4**,
else

If $|\max\{|p^\star|, |q^\star|\} - \varrho|/\varrho < \varepsilon$, then $\boldsymbol{\Delta}(\widehat{\boldsymbol{\nu}}_1, \ldots, \widehat{\boldsymbol{\nu}}_N) = \boldsymbol{\Delta}(\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_N)$.

Otherwise, set $\varrho := \max\{|p^\star|, |q^\star|\}$, $j := 1$, and go to **Step 4**.

Step 9. Calculate the actual endmember estimates: $\widehat{\mathbf{a}}_i = \widehat{\mathbf{C}}\widehat{\boldsymbol{\nu}}_i + \widehat{\mathbf{d}}$ for $i = 1, ..., N$.

Step 10. Estimate the abundance vectors $\widehat{\mathbf{s}}[1], \ldots, \widehat{\mathbf{s}}[L]$ using FCLS [15].

estimated endmembers (abundance maps) (which have been widely used in HU [1, 12, 23]) are used as the performance indices, which are defined as follows:.

$$\phi_{en} = \min_{\boldsymbol{\pi}\in\Pi_N} \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\arccos\left(\frac{\mathbf{a}_i^T\widehat{\mathbf{a}}_{\pi_i}}{\|\mathbf{a}_i\|\|\widehat{\mathbf{a}}_{\pi_i}\|}\right)\right]^2} \tag{3.23}$$

$$\phi_{ab} = \min_{\boldsymbol{\pi}\in\Pi_N} \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left[\arccos\left(\frac{\boldsymbol{s}_i^T\widehat{\boldsymbol{s}}_{\pi_i}}{\|\boldsymbol{s}_i\|\|\widehat{\boldsymbol{s}}_{\pi_i}\|}\right)\right]^2}, \tag{3.24}$$

where $\widehat{\mathbf{a}}_i$ denotes the estimated endmember signature, $\widehat{\boldsymbol{s}}_i = [\,\widehat{s}_i[1], \ldots, \widehat{s}_i[L]\,]^T$ denotes the estimated ith abundance map, $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_N)^T$, and $\Pi_N = \{\boldsymbol{\pi} \in \mathbb{R}^N \mid \pi_i \in \{1, 2, \ldots, N\},\ \pi_i \neq \pi_j \text{ for } i \neq j\}$ is the set of all the permutations of $\{1, 2, ..., N\}$. In addition, the overall average computation time of an algorithm (averaged over all the scenarios) T (in secs) when implemented in Matlab R2008a and running in a desktop computer equipped with Core i7 – 930 CPU with speed 2.80 GHz, and 12 GB memory, is considered as the computational complexity measure.

3.3.1 Uniform Gaussian noise case

The synthetic data for different SNRs were obtained by adding independent and identically distributed zero-mean Gaussian noise to the noise-free data generated, as per (2.1). The average ϕ_{en}, ϕ_{ab}, and average computation time T, of the unmixing algorithms over SNR = 20, 25, ..., 40 dB and $\rho = 0.7, 0.85, 1$ are shown in Table 3.2, where each bold-faced number denotes the minimum rms spectral angle associated with a specific pair of (ρ, SNR) over all the algorithms. One can readily infer from Table 3.2 that the proposed RAVMAX algorithm generally yields the best performance for all the values of ρ and SNRs.

Table 3.2. Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different purity levels (ρ) and SNRs- Uniform Gaussian noise case.

Methods	ρ	ϕ_{en} SNR (dB)					ϕ_{ab} SNR (dB)					
		20	25	30	35	40	20	25	30	35	40	T
N-FINDR	0.7	5.45	5.31	5.24	5.11	5.16	22.54	21.86	21.63	19.76	19.82	1.28
	0.85	2.65	2.67	2.66	2.65	2.61	9.60	8.37	8.03	7.93	7.77	
	1	1.15	0.58	0.33	0.18	0.10	6.14	3.59	2.13	1.24	0.72	
VCA	0.7	5.77	5.56	5.64	5.56	5.50	31.57	29.97	29.71	28.54	28.38	1.44
	0.85	2.79	2.70	2.67	2.71	2.61	10.83	9.45	9.00	8.89	8.82	
	1	1.12	0.61	0.32	0.18	0.11	6.00	3.45	2.05	1.23	0.76	
AVMAX	0.7	5.50	5.36	5.39	5.13	5.10	24.60	21.94	20.95	18.77	**16.48**	3.56
	0.85	2.77	2.64	2.65	2.69	2.65	9.15	7.96	**7.10**	**6.70**	**6.48**	
	1	1.14	0.61	0.33	0.18	0.10	6.39	3.66	2.13	1.22	0.70	
RAVMAX ($0.9 < \eta < 1$)	0.7	**4.87**	**4.87**	**4.88**	**4.83**	**4.90**	**18.95**	**18.15**	**18.13**	**17.83**	17.94	27.97
	0.85	**2.54**	**2.48**	**2.56**	**2.52**	**2.51**	**8.56**	**7.68**	7.44	7.39	7.34	
	1	**0.79**	**0.43**	**0.24**	**0.14**	**0.08**	**4.34**	**2.60**	**1.56**	**0.98**	**0.59**	

3.3.2 Non-uniform Gaussian noise case

For the non-uniform Gaussian noise case, the noise variances σ_i^2 of the M spectral bands, following a Gaussian shape centered at the $(M/2)$th band, as used in [12] [23], are given by

$$\sigma_i^2 = \sigma^2 \frac{\exp(-\frac{(i-M/2)^2}{2\tau^2})}{\sum_{j=1}^{M} \exp(-\frac{(j-M/2)^2}{2\tau^2})}, \ \forall i = 1, \ldots, M, \tag{3.25}$$

where τ controls the variance of the Gaussian shape among $\sigma_1^2, \ldots, \sigma_M^2$. It corresponds to uniform Gaussian noise for $\tau = \infty$, and one-band noise for $\tau = 0$. Table 3.3 shows the average rms spectral angles for SNR from 15 dB to 40 dB and the average computation time of all the algorithms under test, and the purity level in this case is fixed as 1. Again, it can be inferred that the the proposed RAVMAX algorithm outperforms all the algorithm under test, including its predecessor AVMAX algorithm. In both the noisy scenarios, the increment in the computational complexity of RAVMAX algorithm is the price to be paid for the betterment in the estimation accuracy of the robust algorithm.

Table 3.3. Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different τ values and SNRs, with ρ = 1- Non-uniform Gaussian noise case.

Methods	τ	ϕ_{en} SNR (dB)					ϕ_{ab} SNR (dB)					
		20	25	30	35	40	20	25	30	35	40	T
N-FINDR	∞	1.13	0.60	0.33	0.18	0.09	5.12	3.12	1.89	1.33	0.67	1.21
	18	1.34	0.63	0.32	0.17	0.09	5.14	2.77	1.62	0.96	0.56	
	9	2.11	0.65	0.31	0.17	0.09	7.85	2.82	1.56	0.91	0.52	
VCA	∞	1.15	0.67	0.36	0.21	0.13	5.12	2.94	1.82	1.13	0.71	1.51
	18	1.26	0.66	0.37	0.19	0.11	5.23	2.70	1.59	0.99	0.57	
	9	2.13	0.71	0.35	0.17	0.09	9.43	2.83	1.60	0.89	0.53	
AVMAX	∞	1.14	0.60	0.33	0.18	0.10	5.21	3.14	1.90	1.13	0.67	2.84
	18	1.32	0.63	0.34	0.18	0.10	5.13	2.82	1.66	0.98	0.58	
	9	2.03	0.64	0.32	0.16	0.09	7.74	2.81	1.56	0.91	0.53	
RAVMAX ($0.9 < \eta < 1$)	∞	**0.78**	**0.45**	**0.24**	**0.15**	**0.08**	**3.24**	**2.11**	**1.37**	**0.89**	**0.57**	18.21
	18	**0.95**	**0.49**	**0.27**	**0.15**	**0.09**	**3.08**	**1.94**	**1.23**	**0.78**	**0.50**	
	9	**1.19**	**0.49**	**0.26**	**0.15**	**0.08**	**3.42**	**1.93**	**1.19**	**0.77**	**0.48**	

3.4 Real data experiments

In this section, the RAVMAX, AVMAX and VCA algorithms are applied to AVIRIS real hyperspectral data obtained over the Cuprite Nevada site [52], and the results are compared. The AVIRIS data is well studied in the recent years [12, 23] and the availability of a structured library of endmember signature [51, 53] aids in identification of the mineral maps. The static nature of the Cuprite Nevada site over the recent years, together with the availability of a standard library of minerals makes the data appropriate for conducting real data experiments, so as to validate the algorithms under test. The hyperspectral data over the Cuprite Nevada contains 224 bands with better SNRs in most of the bands [54]. Among the 224 bands, bands 1-2, 104-113, 148-167, and 221-224 were removed due to strong noise or dense water-vapor content in those bands. In our experiment, we considered a 200×200 sub-image (region of interest) of the hyperspectral data, with 188 bands (after removing the bands with poor information). Estimation of the number of endmembers present in a given scene of interest is an important issue (it will be addressed later in Chapter 5), because both the dimension reduction algorithm and the HU algorithm are in need of this

number. To begin with, the eigenvalue distribution (signal energies distribution) of the ROI's data covariance matrix shows that the number of sources according to the principal eigenvalues is approximately 4 ($N = 4$). Obviously, it is an under estimate of N, as the ground truth [51] reports more than 4 endmembers in the ROI. Applying HySiMe [45] to the ROI data yields $N = 18$. On the other hand, for a very similar ROI, it is reported in [12] and [23] that Virtual Dimensionality (VD) [41] estimates the number of endmembers as 14. Also Miao *et al.* reported in their work [20] that for a similar ROI, $N = 9$ and they further claimed that the over estimation of N by other methods can be attributed to the non-linear mixing of the observations in a real scenario. However, for the sake of consistency with our previous work [23], here we consider $N = 14$.

In order to remove some pixels corresponding to inactive constraints and thereby speed up the algorithms, data subsampling is needed. Since the pixels at the periphery (the pixels associated with active constraints) of the dimension-reduced data cloud determines the simplex volume, we first perform some subsampling of the dimension-reduced hyperspectral data $\tilde{\mathbf{y}}[i]$ by using the following convex projection procedure: We consider solving the following optimization problem for all $i = 1, \ldots, L$,

$$\begin{aligned} \min_{\boldsymbol{\theta} \in \mathbb{R}^{L-1}} \quad & \|\tilde{\mathbf{y}}[i] - \tilde{\mathbf{Y}}_i \boldsymbol{\theta}\|^2 \\ \text{s.t.} \quad & \mathbf{1}_{L-1}^T \boldsymbol{\theta} = 1, \ \boldsymbol{\theta} \succeq \mathbf{0}, \end{aligned} \tag{3.26}$$

where $\tilde{\mathbf{Y}}_i = [\ \tilde{\mathbf{y}}[1], \tilde{\mathbf{y}}[2], \ldots, \tilde{\mathbf{y}}[i-1], \tilde{\mathbf{y}}[i+1], \ldots, \tilde{\mathbf{y}}[L]\ ]$. Problem (3.26) is a convex problem and can be solved using available convex optimization solvers, such as `SeDuMi` [30] and `CVX` [31]. The idea is to check each and every pixel in the dimension-reduced data cloud for the case whether it belongs to the convex hull of all the other remaining pixels (i.e., if the optimal value of (3.26) is zero), or not. If yes, the pixel is discarded, if not it is retained (as they correspond to a peripheral pixel). In spite of its initial cost,

this procedure of data subsampling significantly aids in speeding up the algorithm under test. By doing so, we were able to identify 17,965 peripheral pixels out of the 40,000 pixels in the ROI, and we have used them as input for RAVMAX algorithm. Finally, it should be mentioned that the complete hyperspectral data set (without subsampling) is used for the estimation of the abundance maps. When applied to real hyperspectral data, the computational time for RAVMAX, AVMAX (both with subsampled data) and VCA (with full data) are around 70 minutes, 42 minutes, and 4 minutes, respectively. The endmember signatures obtained via RAVMAX (with $\eta = 0.9$), AVMAX, and VCA algorithms, along with the corresponding library signatures are shown in Figure 3.3. The minerals were identified by the visual comparison of the obtained abundance maps with the ones available in [20] [23], [51], [55], and [56]. The abundance maps obtained by the RAVMAX, AVMAX, and VCA algorithms are shown in Figure 3.4, Figure 3.4, and Figure 3.4, respectively. The materials identified are arranged in alphabetical order for ease of visual comparison. It should be mentioned that the difference in the materials identified by the three algorithms could possibly be due to the working nature of the respective algorithms and their sensitivity to initializations. The mean removed spectral angle between the estimated signature $\mathbf{a}_{est}$ and the corresponding library signature $\mathbf{a}_{lib}$ [23] [55] defined as

$$\phi = \arccos\left(\frac{(\mathbf{a}_{est} - \mathbf{m}(\mathbf{a}_{est}))^T(\mathbf{a}_{lib} - \mathbf{m}(\mathbf{a}_{lib}))}{\|\mathbf{a}_{est} - \mathbf{m}(\mathbf{a}_{est})\| \cdot \|\mathbf{a}_{lib} - \mathbf{m}(\mathbf{a}_{lib})\|}\right), \tag{3.27}$$

was used as the performance measure, where $\mathbf{m}(\mathbf{a}) = (\mathbf{1}_M^T\mathbf{a})\mathbf{1}_M(1/M)$ for any vector $\mathbf{a} \in \mathbb{R}^M$. The value of ϕ for the various minerals identified by the algorithms under test are given in Table 3.4. Again, the least ϕ value for an endmember is highlighted as a bold-faced number and the number in parenthesis is the ϕ value for the repeatedly identified endmember. It can be seen from Table 3.4 that for the materials identified, the proposed RAVMAX algorithm performs best (with the minimum average ϕ), and

further, it mostly yields better endmember estimates (minimum ϕ) than the AVMAX and VCA algorithms.

Since the available AVIRIS real data have a relatively good SNR (roughly around 30 dB), we have also tried the HU algorithms for real data with artificially added noise. However, the majority of the so obtained abundance maps were inconclusive when compared with available groundtruths [20] [23], [51], [55], [56], and hence the mineral identification and quantification could not be performed for the endmember and abundance estimates of the noise added AVIRIS real data.

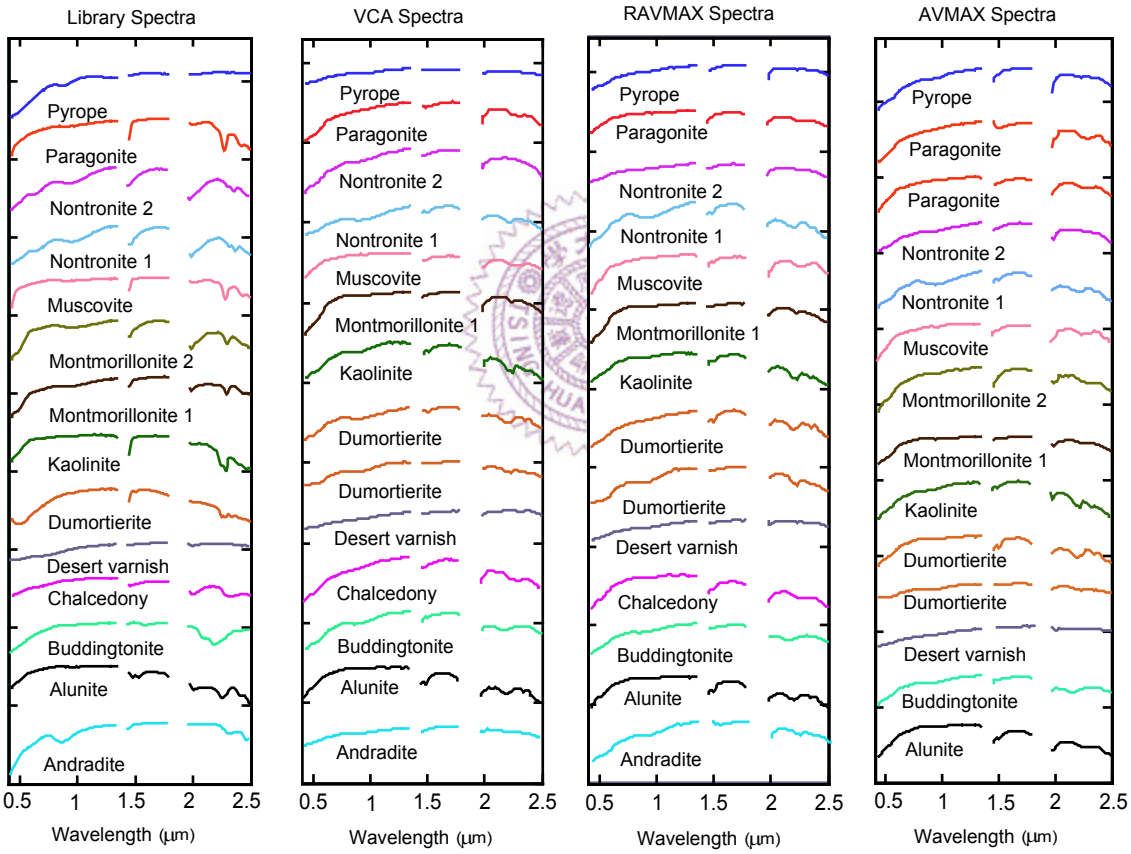


Figure 3.3. Endmember signatures taken from library, and the ones estimated by VCA, RAVMAX, and AVMAX algorithms.

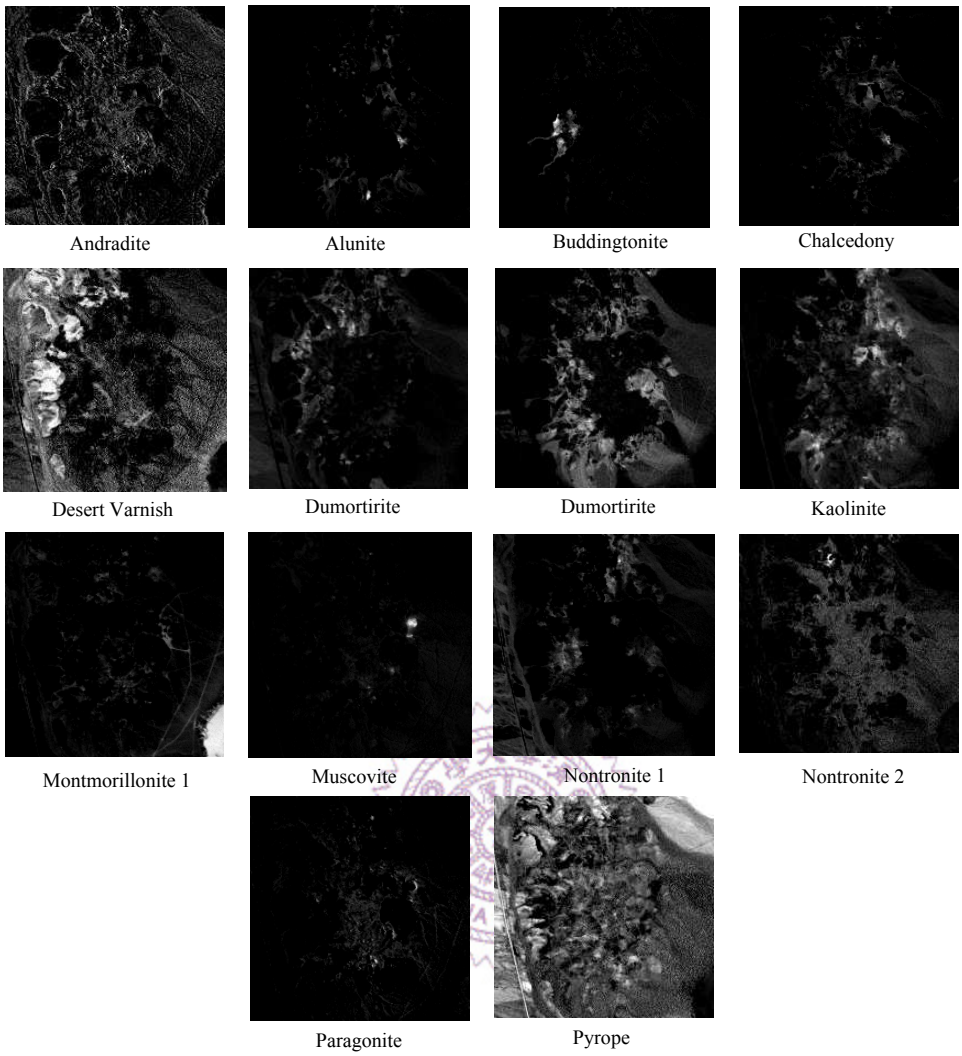


Figure 3.4. Abundance maps obtained by RAVMAX algorithm.

3.5 Summary

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

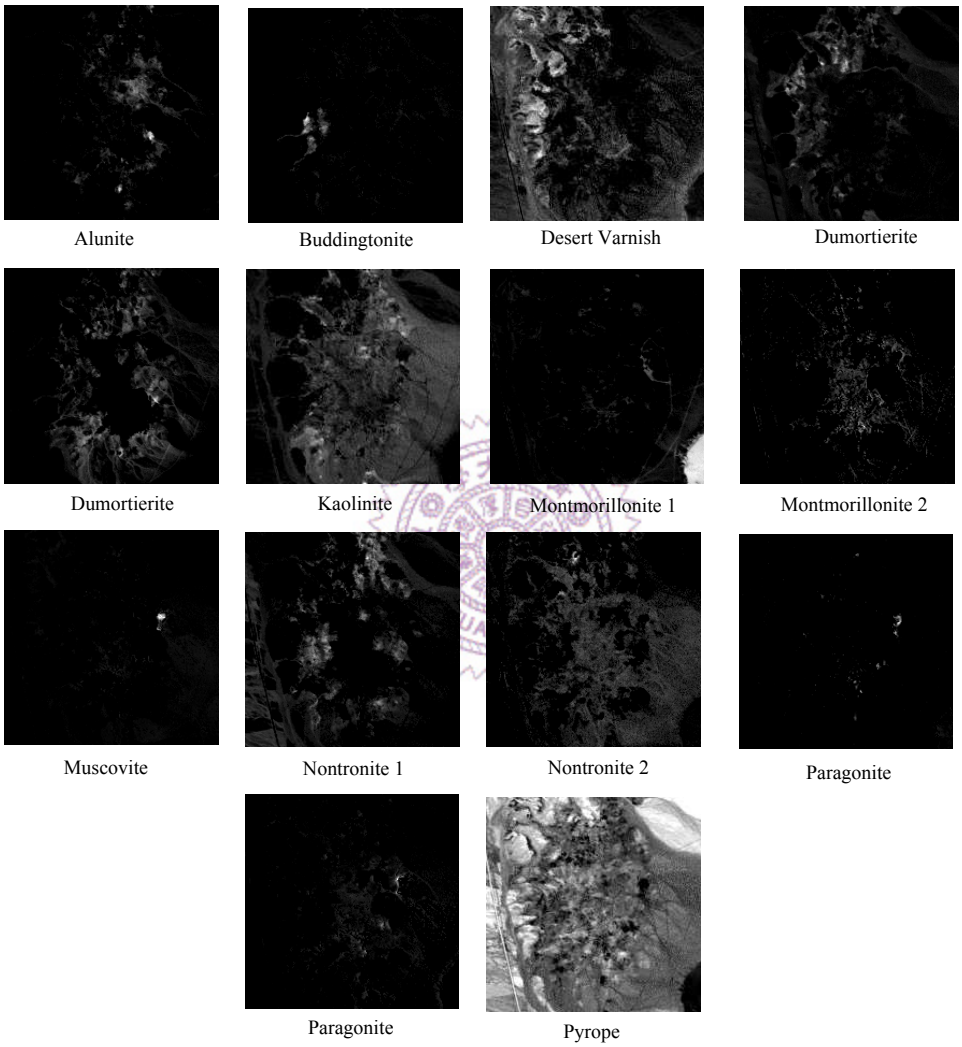


Figure 3.5. Abundance maps obtained by AVMAX algorithm.

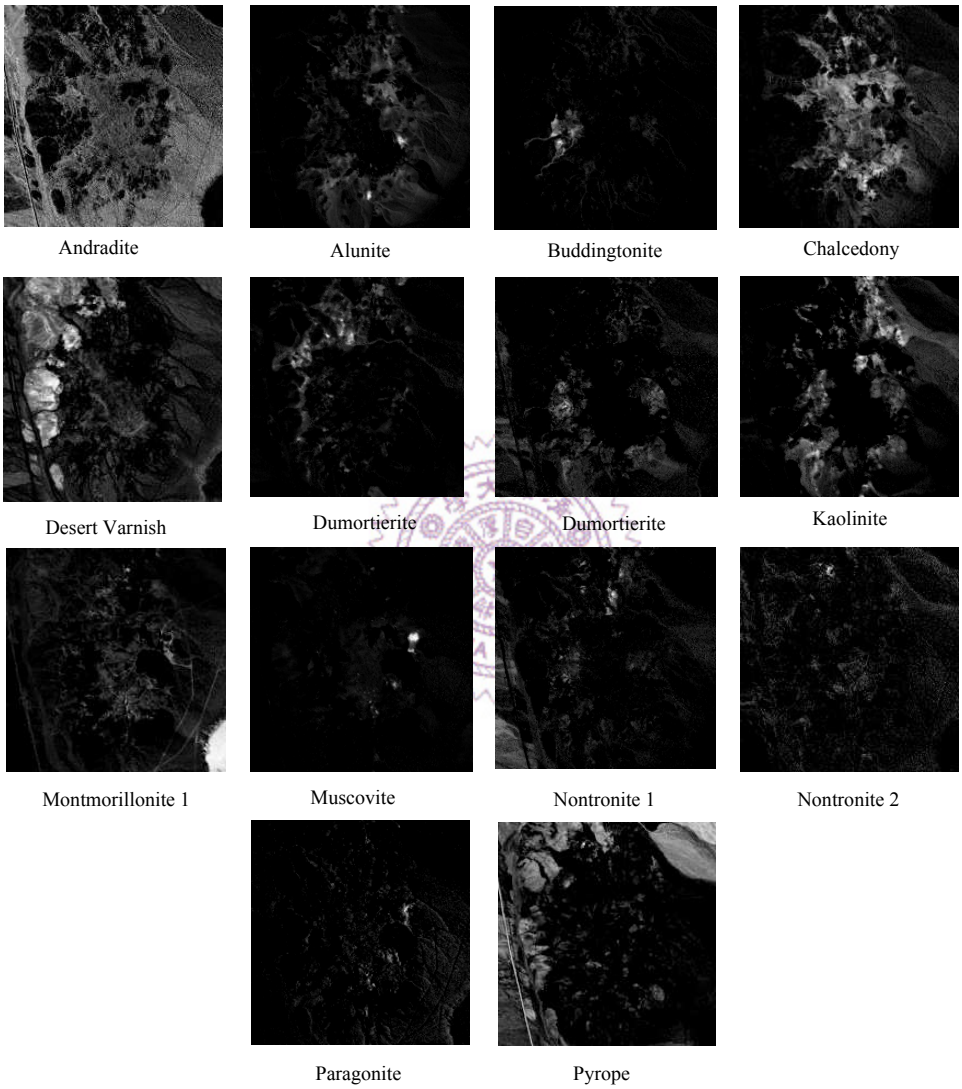


Figure 3.6. Abundance maps obtained by VCA algorithm.

Table 3.4. Mean-removed spectral angles ϕ (degrees) for VCA, RAVMAX ($\eta = 0.9$) and AVMAX algorithms.

	VCA	**RAVMAX**	AVMAX
Andradite	**18.49**	22.06	—
Alunite	**17.74**	21.68	23.70
Buddingtonite	27.25	24.64	**24.22**
Chalcedony	31.9	**19.53**	—
Desert Vanish	12.12	16.09	**12.04**
Dumortierite	31.95 (32.01)	**25.03** (31.40)	27.07 (31.49)
Kaolinite	30.33	**24.16**	34.08
Montmorillonite1	18.06	**17.86**	18.53
Montmorillonite2	—	—	**28.42**
Muscovite	**32.7**	39.05	39.82
Nontronite 1	24.66	23.45	**23.32**
Nontronite 2	21.51	**17.16**	18.30
Paragonite	35.91	**31.34**	33.88 (36.60)
Pyrope	25.59	**23.04**	27.24
Average ϕ	25.73	**24.03**	28.40

Chapter 4

Robust Minimum Volume Enclosing Simplex Algorithm

For hyperspectral data set, where pure pixels cannot be guaranteed for all endmembers (i.e., when (A4) does not hold true), robust algorithm based on Craig's criterion are preferred. The Craig's criterion states that the vertices of the minimum volume simplex enclosing the data cloud serve as high fidelity estimates of the endmember signatures. In this chapter, we propose a robust minimum volume enclosing simplex (RMVES) algorithm, which is derived from the Craig's criterion based minimum volume enclosing simplex (MVES) algorithm. A toy example shown in Figure 4.1 demonstrates how Craig's criterion may yield inaccurate endmember estimates for the noisy observations (shown as dots).

Clearly, the random noise expands the data cloud, and consequently the volume (vertices) of the simplex formed by Craig's criterion (shown in dashed line) is larger than (far from) that of the true simplex (shown in solid line). Therefore, for the hyperspectral data with noise, the aim is to fit a simplex for the observed data, whose volume is less than that of the simplex volume obtained by Craig's criterion, so that the resultant simplex will be close to the true simplex. Motivated by this,

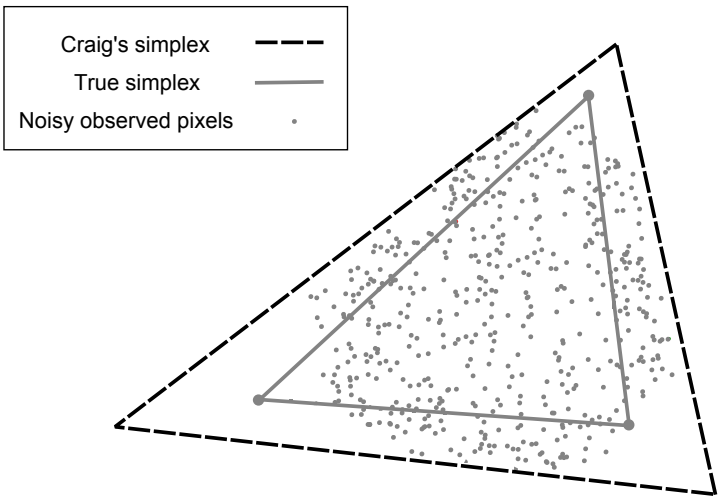


Figure 4.1. Illustration of Craig's simplex for noisy observations, where $N = 3$.

we formulate the RMVES problem by incorporating probability constraints or chance constraints [29] into Craig's criterion based MVES problem [23], so as to deal with the effect of random noise. Under the Gaussian noise assumption, the chance constrained RMVES problem can be conveniently formulated into a deterministic non-linear program, where the pre-assigned probability given in the chance constraints when coupled with the noise covariance, controls the volume of Craig's simplex. We then handle the resulting problem by alternating optimization, where each subproblem involved therein is specifically handled by readily available sequential quadratic programming solvers. First, let us present a brief review of the existing MVES algorithm and then move on to propose the robust version.

4.1 Brief review of MVES Algorithm

The MVES algorithm is derived for noise-free case and it aims to solve the following optimization problem [13] [23]:

$$\boxed{\begin{array}{ll} \min\limits_{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\in\mathbb{R}^{N-1}} & V(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N) \\ \text{s.t.} & \tilde{\mathbf{x}}[n]\in \text{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\},\ \forall n. \end{array}} \tag{4.1}$$

Here, $\tilde{\mathbf{x}}[n]$ is the dimension-reduced noise-free observation (given by (2.8)), $\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N$ correspond to the dimension reduced endmember signatures, and $V(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N)$ is the volume of the simplex $\text{conv}\{\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N\}$, defined as

$$V(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_N)=\frac{|\det(\mathbf{B})|}{(N-1)!}, \tag{4.2}$$

where $(N-1)!$ represents factorial of $N-1$ and

$$\mathbf{B}=[\ \boldsymbol{\beta}_1-\boldsymbol{\beta}_N,\ldots,\boldsymbol{\beta}_{N-1}-\boldsymbol{\beta}_N\ ]\in\mathbb{R}^{(N-1)\times(N-1)}. \tag{4.3}$$

By (4.3) and (2.5), the constraints of (4.1) can be expressed as

$$\tilde{\mathbf{x}}[n]=\sum_{i=1}^{N}s_i[n]\boldsymbol{\beta}_i=\mathbf{B}\mathbf{s}'[n]+\boldsymbol{\beta}_N,\ \forall n, \tag{4.4}$$

where

$$\mathbf{s}'[n]=[s_1[n],\ldots,s_{N-1}[n]]^T\in\mathbb{R}_+^{N-1}, \tag{4.5a}$$

$$s_N[n]=1-\mathbf{1}_{N-1}^T\mathbf{s}'[n]\geq 0, \tag{4.5b}$$

are the abundance fractions of the nth pixel in the observations. Note that (4.5a) and (4.5b) jointly enforce the non-negativity and full-additivity constraints on the

abundances. Equation (4.4) can be rewritten as

$$\mathbf{s}'[n] = \mathbf{B}^{-1}(\tilde{\mathbf{x}}[n] - \boldsymbol{\beta}_N) = \mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g},\ \forall n, \tag{4.6}$$

where

$$\mathbf{H} = \mathbf{B}^{-1}, \tag{4.7a}$$

$$\mathbf{g} = \mathbf{B}^{-1}\boldsymbol{\beta}_N. \tag{4.7b}$$

Substituting (4.6) into (4.5) gives rise to the following equivalent constraints

$$s_i[n] = \mathbf{h}_i^T\tilde{\mathbf{x}}[n] - g_i \geq 0,\ \forall i = 1, \ldots, N-1, n = 1, \ldots, L, \tag{4.8a}$$

$$s_N[n] = 1 - \mathbf{1}_{N-1}^T(\mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g}) \geq 0,\ \forall n = 1, \ldots, L, \tag{4.8b}$$

where $\mathbf{h}_i^T$ is the ith row vector of $\mathbf{H}$, g_i is the ith element of $\mathbf{g}$. Hence, the minimization problem in (4.1) is equivalent to the following maximization problem

$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} \quad |\det(\mathbf{H})| \tag{4.9a}$$

$$\text{s.t.} \quad \mathbf{h}_i^T\tilde{\mathbf{x}}[n] - g_i \geq 0,\ \forall i, n, \tag{4.9b}$$

$$1 - \mathbf{1}_{N-1}^T(\mathbf{H}\tilde{\mathbf{x}}[n] - \mathbf{g}) \geq 0,\ \forall n. \tag{4.9c}$$

After some simplifications, the non-convex optimization problem (4.9) is then handled via alternating optimization, where each and every subproblem is convex (in fact, linear program) and can be solved by using available convex optimization solvers such as `SeDuMi` [30] and `CVX` [31]. Let us emphasize that the MVES algorithm in [23] considers a noise-free scenario, and hence for noisy hyperspectral data, the solution, i.e., the simplex $\text{conv}\{\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_N\}$, obtained by MVES may be far from the true one.

4.2 Robust MVES Formulation and Algorithm

In this section, we develop a robust HU algorithm that aims to accounts for the presence of noise in the observations. The presence of noise in the observations expands the data cloud (even after dimension reduction). Hence, for noisy data, the minimum volume simplex estimated by (4.9) will be larger than the true simplex and therefore its vertices (endmember signatures) will be away from those of the true simplex (as illustrated in Figure 4.1). Given the noisy observations, our aim now is to find a minimum volume simplex that can be closer to the true simplex. This means that we allow some pixels in the noisy data cloud to be outside the estimated simplex.

From (2.22), we have $\tilde{\mathbf{x}}[n] \cong \tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]$. Substituting this for $\tilde{\mathbf{x}}[n]$ in (4.9), we have

$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} \quad |\det(\mathbf{H})| \tag{4.10a}$$

$$\text{s.t.} \quad \mathbf{h}_i^T(\tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]) - g_i \geq 0,\ \forall i, n, \tag{4.10b}$$

$$1 - \mathbf{1}_{N-1}^T\mathbf{H}(\tilde{\mathbf{y}}[n] - \tilde{\mathbf{w}}[n]) + \mathbf{1}_{N-1}^T\mathbf{g} \geq 0,\ \forall n. \tag{4.10c}$$

However, $\tilde{\mathbf{w}}[n] = \widehat{\mathbf{C}}^T\mathbf{w}[n]$ is an unknown random vector. We therefore need to consider this uncertainty in the constraints of (4.10). In the ensuing development, we will make use of chance constraints or probabilistic constraints [29] that can account for the randomness in the observations. A chance constrained counterpart of problem (4.10) is proposed as follows:

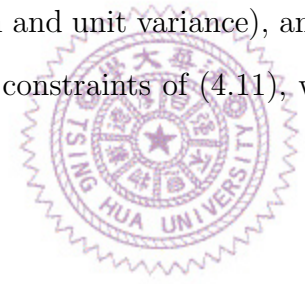
$$\max_{\substack{\mathbf{H}\in\mathbb{R}^{(N-1)\times(N-1)},\\ \mathbf{g}\in\mathbb{R}^{N-1}}} \quad |\det(\mathbf{H})| \tag{4.11a}$$

$$\text{s.t.} \quad \Pr(u_i[n] \leq \mathbf{h}_i^T\tilde{\mathbf{y}}[n] - g_i) \geq \eta,\ \forall i, n, \tag{4.11b}$$

$$\Pr(z[n] \leq 1 - \mathbf{1}_{N-1}^T\mathbf{H}\tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T\mathbf{g}) \geq \eta,\ \forall n, \tag{4.11c}$$

where $\Pr(\cdot)$ represents probability, $u_i[n] \triangleq \mathbf{h}_i^T \widehat{\mathbf{C}}^T \mathbf{w}[n]$ is a random variable with the distribution $\mathcal{N}(0, \mathbf{h}_i^T \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}} \mathbf{h}_i)$, $i = 1, \ldots, N-1$, $z[n] \triangleq -\mathbf{1}_{N-1}^T \mathbf{H} \widehat{\mathbf{C}}^T \mathbf{w}[n]$ is a random variable with the distribution $\mathcal{N}(0, \mathbf{1}_{N-1}^T \mathbf{H} \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}} \mathbf{H}^T \mathbf{1}_{N-1})$, and $\eta \in (0,1)$ is a design parameter. Note that in (4.11), the hard constraints in (4.10), i.e., (4.10b) and (4.10c) are replaced by the soft constraints (4.11b) and (4.11c), respectively. Specifically, in (4.11), we only require that the constraints (4.10b) and (4.10c) hold true with probability no less than η.

The chance constraints in (4.11) can be further simplified by normalizing the random variables involved [29]. Recall from Section 3.2.2 that for a random variable $\varepsilon \sim \mathcal{N}(\mu, \delta^2)$ and $t \in \mathbb{R}$, $\Pr(\varepsilon \leq t) \geq \eta$ is true as $t \geq \delta \Phi^{-1}(\eta) + \mu$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal random variable (Gaussian random variable with zero mean and unit variance), and $\Phi^{-1}(\cdot)$ is the inverse of $\Phi(\cdot)$. Applying this procedure to the constraints of (4.11), we have the following RMVES problem:



$$\max_{\substack{\mathbf{H} \in \mathbb{R}^{(N-1)\times(N-1)}, \\ \mathbf{g} \in \mathbb{R}^{N-1}}} \quad |\det(\mathbf{H})| \tag{4.12a}$$

$$\text{s.t.} \quad \Phi^{-1}(\eta) \sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i, \ \forall i, n, \tag{4.12b}$$

$$\Phi^{-1}(\eta) \sqrt{\mathbf{1}_{N-1}^T \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^T \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T \mathbf{g}, \ \forall n, \tag{4.12c}$$

where $\mathbf{Q} = \mathbf{H} \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}} \mathbf{H}^T$, and $Q_i = \mathbf{h}_i^T \widehat{\mathbf{C}}^T \mathbf{D} \widehat{\mathbf{C}} \mathbf{h}_i$. Let the constraint set of (4.12) be

$$\begin{aligned} \mathcal{F}(\eta) = \Big\{ & (\mathbf{H}, \mathbf{g}) \mid \Phi^{-1}(\eta) \sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i, \ \forall i, n, \\ & \Phi^{-1}(\eta) \sqrt{\mathbf{1}_{N-1}^T \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^T \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T \mathbf{g}, \ \forall n \Big\}. \end{aligned} \tag{4.13}$$

As $\Phi^{-1}(\eta)$ is a monotone increasing function of η, one can infer that

$$\mathcal{F}(\eta_1) \subseteq \mathcal{F}(\eta_2),\ \forall \eta_1 \geq \eta_2. \tag{4.14}$$

Thus, the optimal $|\det(\mathbf{H})|$ will be larger for smaller η. Also note from (4.7a) that an increased value of $|\det(\mathbf{H})|$ corresponds to a decreased value of $|\det(\mathbf{B})| = 1/|\det(\mathbf{H})|$ (i.e., a smaller simplex). It can be readily seen that, when $\eta = 0.5$, the constraint set $\mathcal{F}(0.5)$ of (4.12) is identical to that of (4.9) (as $\Phi^{-1}(\eta = 0.5) = 0$) with $\tilde{\mathbf{x}}[n]$ replaced by $\tilde{\mathbf{y}}[n]$. Simply speaking, for noisy observations our aim is to find a simplex whose volume is minimized further when compared to that of the simplex obtained by Craig's criterion based MVES algorithm (see Figure 4.1). This, together with (4.14), confirms that the appropriate range of η for robust design should be between 0 and 0.5. It is also worthwhile to mention that when $\eta = 0$, $|\det(\mathbf{H})|$ becomes unbounded as $\mathcal{F}(0) = \mathbb{R}^{(N-1)\times(N-1)} \times \mathbb{R}^{N-1}$ and hence $|\det(\mathbf{B})|$ becomes zero, and for the other extreme case of $\eta = 1$, problem (4.12) becomes infeasible (as could be inferred from (4.13)), since $\mathcal{F}(1)$ is an empty set.

Note that the values of η affect the feasible set of (4.12). When $\eta > 0.5$ (i.e., $\Phi^{-1}(\eta) > 0$), the constraints (4.12b) and (4.12c) are second-order cone constraints (convex). These constraints ensure that the corresponding abundances of each noisy observation are positive, rendering all the observations to be inside the estimated simplex. However, from (4.14), one can infer that the estimated simplex when $\eta > 0.5$ will be larger than the one estimated when $\eta = 0.5$ (simplex estimated by MVES). On the other hand, if $\eta < 0.5$ (i.e., $\Phi^{-1}(\eta) < 0$), the constraints (4.12b) and (4.12c) become non-convex in $(\mathbf{H}, \mathbf{g})$. However, these constraints allow some negativity in the abundance fractions, by means of which some noisy pixels can be left out of the estimated simplex. Recall that for the robust design, some noisy pixels must be left out of the simplex such that the estimated simplex can be close to the true simplex

(as could be inferred from Figure 4.1). Thus we again conclude that the apt range of the design parameter η should be between 0 and 0.5.

Figure 4.2 illustrates a scatter plot (for $N = 3$) of the dimension-reduced noisy observations and optimal simplexes, $\text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ of the RMVES problem (4.12) for different values of η. As elaborately discussed above, one can see from Figure 4.2 that when $\eta < 0.5$, the solution of RMVES problem indeed approaches the true simplex (as some of the noisy pixels are left outside the simplex). When $\eta = 0.5$ the minimum volume simplex tightly encloses all the observations, and when $\eta > 0.5$, the simplex expands, but still encloses all the observations. Note that for $\eta \geq 0.5$, the estimated simplex is away from the true simplex.

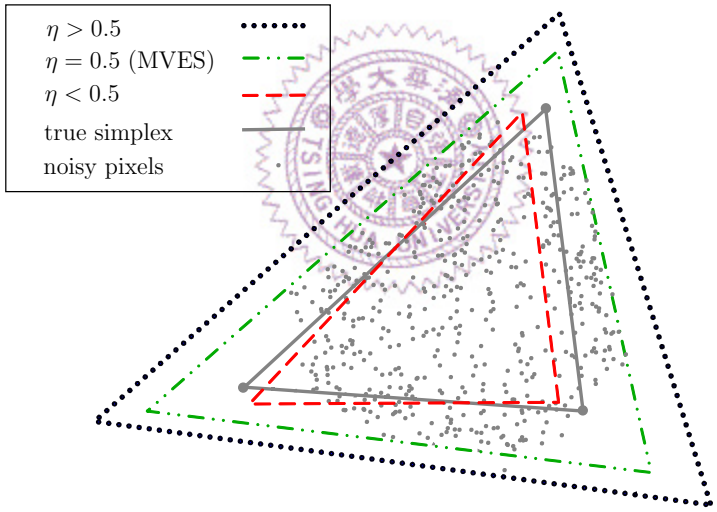


Figure 4.2. Scatter plot of the dimension-reduced pixels for $N = 3$, illustrating the solutions of RMVES for different values of η.

While the feasible set of (4.12) could be convex or non-convex (depending on η), the objective function of (4.12) is always non-convex. Inspired by our previous works [23], [57], [58], and [59], we utilize the alternating optimization methodology, by virtue of which we form subproblems that are hopefully "less non-convex" than (4.12). We consider the cofactor expansion for $\det(\mathbf{H})$ along the ith row of $\mathbf{H}$:

$$\det(\mathbf{H}) = \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}), \tag{4.15}$$

where $\boldsymbol{\mathcal{H}}_{ij}$ is the submatrix of $\mathbf{H}$ with the ith row and jth column removed. One can observe from (4.15) that $\det(\mathbf{H})$ is linear in each $\mathbf{h}_i^T$, which enables us to update $\mathbf{h}_i^T$ and g_i while fixing the other rows of $\mathbf{H}$ and the other entries of $\mathbf{g}$. Then the partial maximization of (4.12) with respect to $\mathbf{h}_i^T$ and g_i can be formulated as

$$\max_{\mathbf{h}_i^T, g_i} \quad \Big| \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}) \Big| \tag{4.16a}$$

$$\text{s.t.} \quad \Phi^{-1}(\eta) \sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i, \ \forall n, \tag{4.16b}$$

$$\Phi^{-1}(\eta) \sqrt{\mathbf{1}_{N-1}^T \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^T \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T \mathbf{g}, \ \forall n. \tag{4.16c}$$

The objective function in (4.16) is still non-convex, but can be handled by breaking it into the following two optimization problems

$$p^\star = \max_{\mathbf{h}_i^T, g_i} \quad \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}) \tag{4.17a}$$

$$\text{s.t.} \quad \Phi^{-1}(\eta) \sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i, \ \forall n,$$

$$\Phi^{-1}(\eta) \sqrt{\mathbf{1}_{N-1}^T \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^T \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T \mathbf{g}, \ \forall n,$$

$$q^\star = \min_{\mathbf{h}_i^T, g_i} \quad \sum_{j=1}^{N-1} (-1)^{i+j} h_{ij} \det(\boldsymbol{\mathcal{H}}_{ij}) \tag{4.17b}$$

$$\text{s.t.} \quad \Phi^{-1}(\eta) \sqrt{Q_i} \leq \mathbf{h}_i^T \tilde{\mathbf{y}}[n] - g_i, \ \forall n,$$

$$\Phi^{-1}(\eta) \sqrt{\mathbf{1}_{N-1}^T \mathbf{Q} \mathbf{1}_{N-1}} \leq 1 - \mathbf{1}_{N-1}^T \mathbf{H} \tilde{\mathbf{y}}[n] + \mathbf{1}_{N-1}^T \mathbf{g}, \ \forall n.$$

As discussed earlier, the apt range of the design parameter is $\eta < 0.5$, and hence unlike in [23], [57], and [59], the subproblems here, (4.17a) and (4.17b) are non-convex, which make problem (4.12) yet difficult to solve. Sequential quadratic programming (SQP) [60] is a well-known method to handle non-convex problems, and here we

employ the available MATLAB based SQP solver, namely *fmincon* [32] to handle the subproblems (4.17a) and (4.17b). It should be noted that though the subproblems (4.17a) and (4.17b) are non-convex, the objective functions are now linear (convex). From our simulation experience, we found that this strategy substantially aids in mitigating the local optimality issue associated with the problem (4.12) (this issue will be addressed later in Section 4.3.3). The optimal solution of (4.16), denoted by $(\mathbf{h}_i^{\star T}, g_i^{\star})$, is chosen as an optimal solution of the maximization problem (4.17a) if $|p^{\star}| > |q^{\star}|$, and that of the minimization problem (4.17b) if $|q^{\star}| > |p^{\star}|$. This row-wise maximization is conducted cyclically (i.e., $i := (i \text{ modulo } (N-1)) + 1$ via each row update of $\mathbf{H}$). We define one iteration as one full update of the matrix $\mathbf{H}$ and the vector $\mathbf{g}$. At each iteration, if the relative change in $|\det(\mathbf{H})|$ between the current and the previous iterations exceeds a given threshold, then we continue with the next updating iteration, else the updated $\mathbf{H}$ and $\mathbf{g}$ are the obtained estimates.

Suppose that a solution $(\mathbf{H}^{\star}, \mathbf{g}^{\star})$ is obtained by the above mentioned alternating maximization method. From (2.15), (4.7a), (4.7b), the endmember signatures can then be recovered by $\widehat{\mathbf{a}}_i = \mathbf{C}\widehat{\boldsymbol{\alpha}}_i + \mathbf{d}$ for $i = 1, \ldots, N$, where

$$\widehat{\boldsymbol{\alpha}}_N = (\mathbf{H}^{\star})^{-1}\mathbf{g}^{\star}, \tag{4.18}$$

$$[\widehat{\boldsymbol{\alpha}}_1, ..., \widehat{\boldsymbol{\alpha}}_{N-1}] = \widehat{\boldsymbol{\alpha}}_N \mathbf{1}_{N-1}^T + (\mathbf{H}^{\star})^{-1}. \tag{4.19}$$

Since some abundance fractions $s_i[n]$ can be negative for those pixels outside the simplex $\text{conv}\{\widehat{\boldsymbol{\alpha}}_1, \ldots, \widehat{\boldsymbol{\alpha}}_N\}$, the abundance estimates $\widehat{\mathbf{s}}[1], \ldots, \widehat{\mathbf{s}}[L]$ are therefore obtained by using the FCLS algorithm [15], which can ensure the non-negativity and full additivity of the estimated abundances. The above illustrated procedure is collectively termed as the robust minimum volume enclosing simplex (RMVES) algorithm.

The proposed RMVES algorithm uses the well known VCA [12] for the initialization of (4.16) [47]. The endmember estimates obtained by VCA are first expanded

until all the dimension-reduced data are well within the simplex formed by the expanded endmember estimates. The expanded endmember estimates are then used to get the initial estimates of $\mathbf{H}$ (by (4.7a)) and $\mathbf{g}$ (by (4.7b)). The pseudocode for the expanded VCA, that is used for RMVES initialization [22] is given in Table 4.1. The pseudocode for the proposed RMVES algorithm is given in Table 4.2.

We conclude this section by summarizing key distinctions between the MVES algorithm [23] and the proposed RMVES algorithm. In contrast to the MVES algorithm that only considers hard constraints assuming the absence of noise, the RMVES algorithm employs soft constraints (chance constraints) to account for the presence of noise in the observed data. In addition, the RMVES algorithm uses the chance constraint parameter η to control the volume of the estimated simplex such that some noisy dimension reduced data $\tilde{\mathbf{y}}[n]$ naturally fall outside the estimated simplex for $\eta < 0.5$, and it yields the same endmember estimates with the MVES algorithm for $\eta = 0.5$. While both algorithms employ the alternating optimization approach, the RMVES algorithm resorts to a series of non-linear and non-convex programs (handled via SQP) to obtain the desired endmember estimates, while the MVES algorithm solves a series of linear programs. Finally, in RMVES algorithm the endmember estimates along with the hyperspectral data are used to estimate the abundances via FCLS algorithm, whereas the MVES algorithm yields endmember estimates and abundance estimates simultaneously.

4.3 Simulations

In this section, the efficacy of the proposed RMVES algorithm is demonstrated. The results shown in this section are averaged over 50 Monte Carlo runs for each considered scenario. To show the robustness of the proposed algorithm, two different noisy scenarios are considered. The first scenario is for the observations corrupted with

Table 4.1. Pseudocode for expanded VCA initialization.

Given	The affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ and the endmember matrix estimated by VCA [12], say $\mathbf{A}_0$.
Step 1.	Obtain the dimension-reduced VCA endmember estimates $\mathbf{E}_0 = \widehat{\mathbf{C}}^T(\mathbf{A}_0 - \widehat{\mathbf{d}}\mathbf{1}_N^T)$.
Step 2.	Subtract the mean vector $\boldsymbol{\mu}_{E_0} = \mathbf{E}_0\mathbf{1}_N(1/N)$ from $\mathbf{E}_0$, i.e., $\bar{\mathbf{E}} = \mathbf{E}_0 - \boldsymbol{\mu}_{E_0}\mathbf{1}_N^T$.
Step 3.	Repeat $\mathbf{E}_0 := \mathbf{E}_0 + k\bar{\mathbf{E}}$ until all the elements of $\mathbf{E}_0^\dagger\mathbf{y}[n]$ are non-negative for all $n = 1, \ldots, L$, where k is a constant (say $k = 5$).
Step 4.	The output $\mathbf{E}_0 \triangleq [\boldsymbol{\beta}_1^{(0)}, \ldots, \boldsymbol{\beta}_N^{(0)}]$ can then be used to obtain the initial $\mathbf{H}^{(0)} = [\boldsymbol{\beta}_1^{(0)} - \boldsymbol{\beta}_N^{(0)}, \ldots, \boldsymbol{\beta}_{N-1}^{(0)} - \boldsymbol{\beta}_N^{(0)}]^{-1}$ and $\mathbf{g}^{(0)} = \mathbf{H}^{(0)}\boldsymbol{\beta}_N^{(0)}$.

uniform Gaussian noise (Section 4.3.1). The performance is studied under different purity levels and for different signal-to-noise ratios (SNRs). The second scenario is that the observations are corrupted by non-uniform Gaussian noise (Section 4.3.2). In this case, we fix the purity level and evaluate the performance of the algorithms under test for different distributions of the non-uniform Gaussian noise variance over various SNR levels. Finally, in Section 4.3.3, we study the local optimality issue associated with handling the non-convex problems (4.12) and (4.16) (via RMVES algorithm), to show the substantial performance improvement of the latter over the former. Several existing algorithms are used to compare the performance of the proposed RMVES algorithm. The pure-pixel based methods used are N-FINDR [9], SGA [10], and VCA [12], and then FCLS [15] is used to find the abundances associated with the obtained endmember estimates. The other methods used are ICE [17], APS [19], MVC-NMF [20], MVSA [22], SISAL [28] and MVES [23]. The parameter settings and algorithmic details of the algorithms under test are summarized in Table 4.3 (for other parameters of the algorithms not listed in the table, their default values mentioned

Table 4.2. Pseudocode for RMVES Algorithm.

Given The noisy observed data $\mathbf{y}[n]$, the number of endmembers N, a design parameter η, and a convergence tolerance $\varepsilon > 0$.

Step 1. Estimate the noise covariance matrix (denoted by $\widehat{\mathbf{D}}$) using multiple regression analysis [45].

Step 2. **Dimension reduction**: Obtain the dimension-reduced pixels: $\tilde{\mathbf{y}}[n] = \widehat{\mathbf{C}}^T(\mathbf{y}[n] - \widehat{\mathbf{d}})$ for all n, with the affine set fitting parameters $(\widehat{\mathbf{C}}, \widehat{\mathbf{d}})$ given by (2.20) and (2.21).

Step 3. **Initialization**: Let $\mathbf{H} = \mathbf{H}^{(0)}$ and $\mathbf{g} = \mathbf{g}^{(\mathbf{0})}$, where $\mathbf{H}^{(0)}$ and $\mathbf{g}^{(0)}$ are given by the expanded VCA (as explained in Table 4.1). Set $i := 1$ and $\varrho := |\det(\mathbf{H})|$.

Step 4. Define $\mathbf{Q} = \mathbf{H}\widehat{\mathbf{C}}^T\widehat{\mathbf{D}}\widehat{\mathbf{C}}\mathbf{H}^T$ and let $\boldsymbol{\mathcal{H}}_{ij} \in \mathbb{R}^{(N-1)\times(N-1)}$ denote the submatrix of $\mathbf{H}$ with the ith row and jth column removed. Then, handle (4.17a) and (4.17b) using SQP and obtain their solutions, denoted by $(\bar{\mathbf{h}}_i^T, \bar{g}_i)$ and $(\check{\mathbf{h}}_i^T, \check{g}_i)$, respectively. Let $p^\star$ and $q^\star$ be the optimal values of (4.17a) and (4.17b), respectively.

Step 5. If $|p^\star| > |q^\star|$, update $(\mathbf{h}_i^T, g_i) := (\bar{\mathbf{h}}_i^T, \bar{g}_i)$, otherwise $(\mathbf{h}_i^T, g_i) := (\check{\mathbf{h}}_i^T, \check{g}_i)$.

Step 6. If (i modulo $(N-1)$) $\neq 0$, then $i := i + 1$, and go to **Step 4**, else

If $|\max\{|p^\star|, |q^\star|\} - \varrho|/\varrho < \varepsilon$, then $\mathbf{H}^\star = \mathbf{H}$ and $\mathbf{g}^\star = \mathbf{g}$.

Otherwise, set $\varrho := \max\{|p^\star|, |q^\star|\}$, $i := 1$, and go to **Step 4**.

Step 7. Calculate the dimension reduced endmember estimates: $\widehat{\boldsymbol{\alpha}}_N = (\mathbf{H}^\star)^{-1}\mathbf{g}^\star$ and $[\widehat{\boldsymbol{\alpha}}_1, ..., \widehat{\boldsymbol{\alpha}}_{N-1}] = \widehat{\boldsymbol{\alpha}}_N \mathbf{1}_{N-1}^T + (\mathbf{H}^\star)^{-1}$.

Step 8. Calculate the actual endmember estimates: $\widehat{\mathbf{a}}_i = \widehat{\mathbf{C}}\widehat{\boldsymbol{\alpha}}_i + \widehat{\mathbf{d}}$ for $i = 1, ..., N$.

Step 9. Estimate the abundance vectors $\widehat{\mathbf{s}}[1], \ldots, \widehat{\mathbf{s}}[L]$ by using the FCLS algorithm [15].

---

in their respective references are used). The root-mean-square (rms) spectral angles, denoted as ϕ_{en} (ϕ_{ab}) between the true endmembers (abundance maps) and estimated endmembers (abundance maps) (which are defined in (3.23) and (3.24), respectively) are used as the performance indices. The average computation time (averaged over all the scenarios and over all independent runs) of all the algorithms (i.e., the average time required to estimate both the endmember signatures and the abundances), when implemented in Matlab R2008a and running in a desktop computer equipped with Core i7 − 930 CPU with speed 2.80 GHz, and 12 GB memory, is considered as the computational complexity measure.

Table 4.3. Implementation details for the algorithms under test.

Algorithm	**Algorithmic details**
NFINDR	Dimension reduction: PCA; Initialization: Randomly selected data points; Abundance estimation: FCLS.
SGA	Dimension reduction: variable PCA; Initialization: Randomly selected target pixel; Abundance estimation: FCLS.
VCA	Dimension reduction: SVD/PCA; Abundance estimation: FCLS.
ICE	Dimension reduction: None; Initialization: VCA and FCLS; Convergence tolerance: 10^{-5}; Maximum iterations: 1000 Regularization parameter: 0.01.
APS	Dimension reduction: None; Initialization: VCA and FCLS; Convergence tolerance: 2×10^{-4}; Maximum iterations: 500; Number of subgradient updates: 20; Regularization parameter: 0.001.
MVC-NMF	Dimension reduction: SVD; Initialization: VCA and FCLS; Convergence tolerance: 10^{-6}; Regularization parameter: 0.001.
MVSA	Dimension reduction: PCA; Initialization: Expanded VCA; Regularization parameter: 10^{-6}.
SISAL	Dimension reduction: PCA; Initialization: Expanded VCA; Regularization parameter: 0.010 to 0.035 (depending on SNR).
MVES	Dimension reduction: Affine set fitting; Initialization: Solving feasibility problem; Convergence tolerance: 10^{-8}.
RMVES	Dimension reduction: Modified affine set fitting; Initialization: Expanded VCA; Convergence tolerance: 10^{-6}; Design parameter η: 0.001.

In the simulations, 8 endmembers, Alunite, Andradite, Buddingtonite, Calcite, Chalcedony, Chlorite, Desert Varnish, and Halloysite with 224 spectral bands selected from USGS library [51, 53] are used to generate the synthetic observations. The abundance vectors were generated by following the Dirichlet distribution [12] [23],

which can ensure the assumptions on the abundance vectors (i.e., (A1) and (A2)). The purity levels considered in our simulations are $\rho = 0.6, 0.8, 1$.

As previously inferred for the RMVES algorithm, the apt value of η must be less than 0.5, and from our extensive simulation experience we found that η should lie in the range of 0.01 to 0.0001. Therefore, in each simulation and real data experiment (see Section 4.4), the η value for the RMVES algorithm is set to 0.001 (constant) for all SNRs. The reasons behind fixing a η value can be explained by scrutinizing the constraints of problem (4.16), where Q_i and $\mathbf{Q}$ in (4.16b) and (4.16c), respectively are functions of the noise covariance matrix $\mathbf{D}$. This implies, for a fixed η value, the noise covariance matrix indirectly determines the lower bound of the abundances, which in turn controls the volume of the estimated simplex. Further, since (4.16) is non-convex, for each given data set (under a given scenario), the RMVES algorithm is executed 10 times with 10 different expanded VCA initializations (since each time VCA yields different endmember estimates for noisy observations), and the endmember signature and abundance fractions associated with the largest $|\det(\mathbf{H})|$ is chosen as the optimal endmember and abundance estimates for the data under consideration. Such a technique has been applied before in handling non-convex problems, e.g., [61].

4.3.1 Uniform Gaussian noise case

For the uniform Gaussian noise case, the noisy data were obtained by adding independent and identically distributed (i.i.d.) zero-mean Gaussian noise vector to the noise-free data for different SNRs (defined in (3.22)). To maintain non-negativity of the observed pixels, we artificially set the negative values of the noisy pixels to zero. Table 4.4 shows the rms endmember spectral angle ϕ_{en} (in degrees), rms abundance spectral angle ϕ_{ab} (in degrees), and the overall average computation time (over all the scenarios) T (in secs) of all the algorithms under test, for the observations corrupted

Table 4.4. Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different purity levels (ρ) and SNRs- Uniform Gaussian noise case.

Methods	ρ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
N-FINDR	0.6	9.35	8.98	8.83	9.21	9.13	8.93	50.79	48.38	47.02	47.40	47.79	47.40	1.49
	0.8	7.22	6.08	5.45	5.00	4.97	4.87	39.41	29.38	21.81	19.64	19.33	18.28	
	1	5.37	2.31	1.48	1.06	0.51	0.32	31.84	13.38	8.18	5.77	2.93	1.78	
SGA	0.6	8.99	8.62	8.76	8.59	8.70	8.58	51.77	48.05	47.48	47.40	47.23	47.28	1.35
	0.8	6.76	4.73	4.23	4.16	3.95	3.93	41.76	28.36	21.98	20.56	19.09	18.83	
	1	5.28	2.13	**1.22**	0.72	0.52	0.40	32.01	13.93	8.27	5.11	3.42	2.44	
VCA	0.6	9.35	8.97	8.33	8.05	8.06	7.82	51.41	47.64	45.98	44.93	45.10	44.69	1.50
	0.8	7.16	5.86	5.38	4.63	4.30	4.30	41.91	30.53	29.25	24.14	20.30	18.90	
	1	5.29	2.39	3.02	1.17	0.57	0.39	34.28	15.28	16.24	6.81	3.49	1.99	
ICE	0.6	8.70	8.19	8.10	8.41	8.14	8.22	**39.98**	33.57	32.12	30.16	30.11	30.03	420.61
	0.8	**5.33**	3.96	3.83	3.74	3.73	3.71	**28.77**	19.40	16.62	14.48	14.50	12.68	
	1	**4.17**	2.53	2.22	2.02	1.96	1.95	**23.65**	14.71	10.96	9.11	8.47	8.28	
APS	0.6	10.67	8.99	9.26	9.77	9.81	10.28	44.45	37.67	34.61	30.71	28.75	26.50	16.76
	0.8	9.00	5.09	7.21	6.46	5.75	5.65	36.33	23.65	24.00	17.52	15.21	14.69	
	1	7.11	**1.96**	5.08	**0.59**	0.70	0.55	32.49	**13.29**	14.43	**4.60**	3.51	2.31	
MVC-NMF	0.6	11.88	10.69	10.91	10.25	9.89	9.72	45.25	36.07	34.96	33.41	32.11	32.04	72.25
	0.8	11.06	6.04	6.69	4.89	3.69	5.35	37.32	22.93	23.70	14.70	10.41	9.93	
	1	10.27	4.50	4.64	1.12	**0.39**	**0.16**	30.47	14.17	12.12	4.82	**2.58**	**1.50**	
MVSA	0.6	14.35	13.18	10.59	8.01	5.40	3.51	45.63	36.31	24.33	15.98	10.17	5.81	3.04
	0.8	14.22	12.72	10.39	7.78	5.02	3.62	40.85	29.42	21.02	13.79	8.56	5.68	
	1	14.47	12.30	10.63	8.06	5.14	3.44	38.49	27.04	19.51	13.08	8.04	4.82	
SISAL	0.6	**8.10**	7.33	6.11	5.25	3.82	2.67	46.93	38.75	26.38	20.21	13.68	8.49	1.91
	0.8	6.48	**3.53**	**1.54**	**0.80**	**0.58**	**0.41**	36.56	20.06	**9.80**	**5.20**	**3.22**	**2.13**	
	1	5.15	2.41	1.77	1.19	0.72	0.42	29.51	13.54	**7.89**	4.76	2.82	1.68	
MVES	0.6	15.01	12.37	9.70	7.44	5.18	3.53	42.33	33.07	21.90	14.76	9.46	5.76	46.44
	0.8	14.20	12.36	10.25	7.68	5.16	3.68	43.28	32.14	22.41	15.00	9.55	5.72	
	1	14.63	12.59	10.30	7.72	5.28	3.65	44.05	32.90	23.19	15.41	10.15	6.46	
RMVES	0.6	9.14	**5.72**	**3.26**	**2.31**	**1.58**	**1.09**	41.49	**26.82**	**13.13**	**7.90**	**4.73**	**2.84**	118.97
	0.8	8.72	5.13	3.42	2.43	1.68	1.11	35.63	**18.48**	10.30	6.49	4.18	2.66	
	1	8.74	5.63	4.41	3.55	2.59	1.77	32.94	16.83	10.85	7.73	6.10	3.61	

by uniform Gaussian noise, with SNRs between 15 dB and 40 dB, and $\rho = 0.6, 0.8,$ and 1. For each scenario, the best (minimum ϕ_{en} and ϕ_{ab}) out of all the algorithms is highlighted as a bold-faced number. As one can infer from Table 4.4, the proposed RMVES algorithm achieves the best performance for the highly mixed case ($\rho = 0.6$) when SNR$\geq$ 20 dB, and for all the values of ρ the RMVES algorithm performs better than its predecessor, MVES algorithm. For the moderately mixed case ($\rho = 0.8$), SISAL shows best performance for SNR$\geq$ 20 dB, followed by RMVES for SNR$\geq$ 25 dB. As expected, when pure pixels exist in the data ($\rho = 1$), the performances of the pure-pixel based algorithms such as N-FINDR, VCA and SGA are comparable with each other, and in some scenarios ICE and APS perform well. As far as the

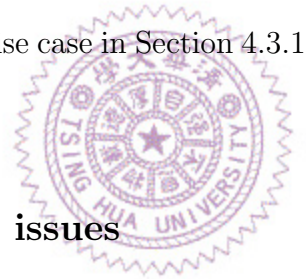
Table 4.5. Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), over the various unmixing algorithms for different τ values and SNRs, with $\rho = 0.7$- Non-uniform Gaussian noise case.

Methods	τ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
N-FINDR	∞	8.14	7.77	7.43	7.23	7.38	7.34	46.89	41.50	37.13	35.12	36.99	34.41	1.34
	18	10.08	8.41	7.19	7.41	7.25	7.28	53.02	47.10	41.89	36.13	33.72	33.62	
	9	11.27	8.96	7.71	7.66	7.60	7.43	56.13	50.63	44.96	40.66	35.64	34.21	
SGA	∞	8.06	6.98	6.50	6.43	6.45	6.45	47.86	40.82	37.96	37.27	37.10	36.73	1.41
	18	10.04	8.02	7.30	6.55	6.47	6.59	53.53	49.58	42.08	37.58	37.26	36.00	
	9	11.37	8.72	7.83	6.85	6.57	6.54	57.33	51.34	46.95	41.95	37.14	35.89	
VCA	∞	8.38	7.68	6.54	6.38	6.62	6.29	47.76	40.46	38.97	36.26	36.61	36.54	1.43
	18	9.89	8.22	6.85	6.49	6.34	6.17	56.48	49.81	41.35	39.59	36.69	35.69	
	9	10.91	**8.68**	7.56	6.78	6.41	6.37	59.34	53.39	46.54	40.46	36.58	36.36	
ICE	∞	**6.02**	4.89	4.93	4.62	4.56	4.76	**31.60**	22.25	19.54	17.26	15.99	17.87	470.17
	18	10.95	**6.65**	4.83	4.71	4.70	4.78	**47.79**	37.31	19.12	17.97	17.71	17.47	
	9	14.05	8.95	**5.22**	4.67	4.92	4.65	**51.20**	**44.73**	26.00	16.75	18.77	17.43	
APS	∞	10.37	8.33	9.08	9.22	9.81	9.57	39.78	32.09	27.98	25.33	25.06	23.47	24.36
	18	17.40	11.71	10.42	9.47	9.15	9.67	50.20	41.77	32.22	26.72	25.01	23.03	
	9	20.64	15.06	12.76	9.82	9.68	9.97	53.21	46.37	37.55	28.35	25.50	24.03	
MVC-NMF	∞	11.94	9.79	9.61	9.31	7.98	7.85	40.63	32.54	32.96	26.58	23.13	21.75	44.62
	18	11.19	11.75	13.05	9.55	9.11	8.20	51.33	40.48	33.32	29.39	25.23	23.91	
	9	13.05	11.44	17.72	10.83	8.58	8.69	54.36	47.72	39.38	29.93	26.40	23.53	
MVSA	∞	14.37	12.87	10.38	7.56	4.90	3.43	44.05	32.82	22.60	14.14	8.72	5.57	2.96
	18	15.54	14.62	13.27	8.71	5.54	3.74	59.47	50.44	33.29	17.00	9.44	5.55	
	9	15.35	14.42	12.93	11.91	6.41	4.20	61.96	56.46	45.67	28.32	11.02	6.38	
SISAL	∞	7.37	5.84	3.87	**2.17**	1.83	1.91	41.23	30.66	17.97	8.84	6.79	5.52	2.41
	18	**9.76**	8.13	5.92	**2.94**	**1.38**	1.59	52.81	43.75	30.08	12.16	6.07	4.49	
	9	**10.68**	8.93	7.28	4.90	**1.50**	**1.39**	58.75	49.39	38.27	23.55	6.27	4.70	
MVES	∞	14.26	12.15	10.17	7.46	5.15	3.42	42.90	32.38	22.26	14.80	9.33	5.66	38.58
	18	14.98	14.31	13.08	8.30	5.30	3.77	52.34	44.75	32.05	16.42	9.43	5.56	
	9	15.92	15.17	12.55	11.17	6.35	4.25	53.70	49.77	40.40	26.81	11.64	6.83	
RMVES	∞	8.52	**4.85**	**3.23**	2.44	**1.48**	**1.17**	37.45	**20.63**	**11.44**	**6.99**	**4.23**	**2.63**	159.44
	18	10.44	7.89	**4.17**	2.97	1.63	**1.08**	52.07	**32.96**	**14.57**	**8.13**	**4.11**	**2.56**	
	9	11.73	9.07	6.57	**3.13**	2.24	1.44	63.06	45.89	**25.71**	**8.97**	**5.33**	**3.23**	

average computation time of the algorithms is concerned, Table 4.4 confirms that the values of the average computation time for pure-pixel based algorithms (N-FINDR, VCA and SGA) are less than those of other algorithms. Among the other algorithms, SISAL costs the least average computation time and the average computation time of RMVES algorithm is less than that of ICE algorithm. As can be seen from Table 4.4, the computational complexity of RMVES algorithm is the price for the accuracy, especially when the hyperspectral data are highly mixed.

4.3.2 Non-uniform Gaussian noise case

For the non-uniform Gaussian noise case, the noise variances σ_i^2 of the M spectral bands, following a Gaussian shape centered at the $(M/2)$th band, as explained in Section 3.3.2. Table 4.5 shows the average rms spectral angles and the average computation time of all the algorithms under test, for SNR from 15 dB to 40 dB, $\rho = 0.7$, and $\tau = \infty, 18$, and 9. For the non-uniform Gaussian noise case, in most of the scenarios under consideration, RMVES performs best followed by SISAL and ICE. It is worth mentioning that for $\tau = 9$ some performance drop of RMVES may be due to the estimation of the noise covariance matrices in such scenarios. Here again, in all the scenarios, the RMVES algorithm performs better than its predecessor, MVES algorithm. The observations on the average computation time of the algorithms under test for the uniform Gaussian noise case in Section 4.3.1 also apply to the non-uniform Gaussian noise case here.



4.3.3 Local optimality issues

In this subsection, let us investigate the local optimality issues while attempting to solve the non-convex problems (4.12) and (4.16) (via RMVES algorithm). Table 4.6 shows the average ϕ_{en} and ϕ_{ab} for the uniform Gaussian noise case, and the avearge computation time T (in seconds), when (4.12) is directly handled using SQP. Here again the η value used is 0.001 and (4.12) is executed 10 times with 10 different VCA initializations (as in Table 4.1). The estimates associated with the maximum value of $|\det(\mathbf{H})|$ is chosen as the solution for (4.12). As one can infer by comparing Table 4.4 and Table 4.6, attempting to solve (4.16) in a cyclic optimization fashion (RMVES algorithm) significantly improves the performance of the algorithm. Numerical results from these tables indicate that cyclic optimization is not as susceptible to local optimality issues as attempting to solve (4.12) directly by SQP.

Table 4.6. Average ϕ_{en} (degrees), ϕ_{ab} (degrees), and average computation time T (secs), while directly applying SQP to (4.12), for different ρ and SNRs- Uniform Gaussian noise case.

Method	ρ	ϕ_{en} SNR (dB)						ϕ_{ab} SNR (dB)						
		15	20	25	30	35	40	15	20	25	30	35	40	T
Applying SQP to (4.12)	0.6	21.02	20.34	20.86	20.48	20.75	20.14	48.69	46.00	45.95	43.48	43.34	41.90	5.64
	0.8	19.41	20.10	20.47	17.89	16.98	17.11	52.88	49.61	48.48	44.49	40.90	41.10	
	1	18.86	18.91	16.56	15.49	15.39	15.44	55.46	52.25	46.09	44.16	44.18	44.23	

Another interesting observation is the variation in the average endmember spectral angle when handling (4.16) (via RMVES algorithm) with one VCA initialization (as in Table 4.1) and with ten different VCA initializations. In Figure 4.3, the asterisks correspond to the average endmember spectral angle (averaged over 50 independent runs) obtained by executing the RMVES algorithm with a single VCA initialization in each of the 50 independent runs, for $\rho = 0.6$ and SNR (uniform Gaussian noise) varying from 15 dB to 40 dB with the step size of 5 dB. The triangles in the figure represent the average endmember spectral angles associated with the RMVES algorithm for the same scenarios, but with ten VCA initializations in each of the 50 independent runs, and choosing the spectral angle associated with the maximum $|\det(\mathbf{H})|$ (i.e., those ϕ_{en}s shown in Table 4.4). As one can infer, the differences between the average endmember spectral angles obtained by RMVES with one VCA initialization and by RMVES with ten VCA initializations, for all the scenarios are within two degrees. This indicates that the local optimality issue associated with (4.16) is not severe, but still exists (even after linearizing the objective function of (4.12) via cofactor expansion), since the constraints sets are non-convex.

Analyzing the local optimality issue of the existing HU algorithms in a fair fashion is non-trivial due to their distinct characteristics and can be a separate research topic. However, to illustrate the existence of local optimality in other HU methods, we proceed with the following simulations. A total of 50 data sets have been independently

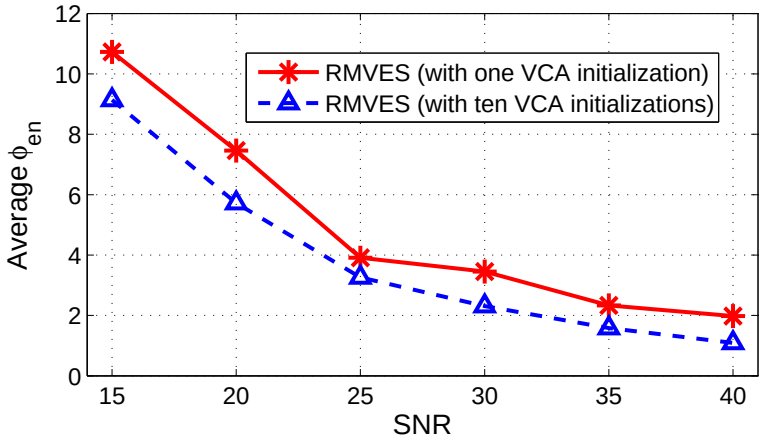


Figure 4.3. Average endmember spectral angles obtained by the RMVES algorithm for $\rho = 0.6$ with one VCA initialization, and with ten VCA initializations and choosing the spectral angle associated with the maximum $|\det(\mathbf{H})|$.

generated with $\rho = 0.6$, under uniform Gaussian noise case with SNR 25 dB, and the HU algorithms were applied 10 times (*each time with a different initialization*) to each of the data set. Then, for the ith data set, the variance and mean of ϕ_{en} (over the 10 executions of each algorithm on the ith data set), denoted by $\text{var}\{\phi_{en}\}_i$ and $\text{mean}\{\phi_{en}\}_i$, respectively are noted and the average variance and mean over 50 independent runs are computed respectively as,

$$\text{var}\{\phi_{en}\} = \frac{1}{50}\sum_{i=1}^{50} \text{var}\{\phi_{en}\}_i,$$

and

$$\text{mean}\{\phi_{en}\} = \frac{1}{50}\sum_{i=1}^{50} \text{mean}\{\phi_{en}\}_i.$$

Table 4.7 shows the $\text{var}\{\phi_{en}\}$ and $\text{mean}\{\phi_{en}\}$ of the various algorithms under test. The $\text{mean}\{\phi_{en}\}$ values in Table 4.7 are comparable with the corresponding values shown in Table 4.4, where the average values of ϕ_{en} are over 50 independent runs without using 10 different initializations for each independent run. As can be

seen from Table 4.7, in addition to SGA [10] and APS [19], the var$\{\phi_{en}\}$ for MVES algorithm is also zero because the MVES algorithm is initialized by solving the feasibility problem [23] (that is inherent in the MVES algorithm), which being convex always yields the same initial estimates for each data set (i.e., the same initialization is yielded for the same data set).

Table 4.7. var$\{\phi_{en}\}$ and mean$\{\phi_{en}\}$ of various HU algorithms, when $\rho = 0.6$ and SNR=25 dB, for 50 independent data sets. For each data set the algorithm is applied 10 times, each time with its inherent initialization.

Algorithm	var$\{\phi_{en}\}$ (in degrees^2)	mean$\{\phi_{en}\}$ (in degrees)
NFINDR	1.16	9.06
SGA	0	8.52
VCA	1.23	8.17
ICE	1.64	8.38
APS	0	9.26
MVC-NMF	5.04	10.69
MVSA	3.02	10.81
SISAL	2.38	6.74
MVES	0	9.71
RMVES	3.77	4.29

4.4 Real data experiments

In this section, we demonstrate the performance of the proposed RMVES algorithm using the AVIRIS data taken over the Cuprite Nevada site [52], the details of which are provided in Chapter 3 Section 3.4. The AVIRIS data are challenging because of two reasons. One reason is that the spatial resolution is not so high (about 17 m per pixel) and the other reason is that the true total number of endmembers and the associated minerals are yet to be accurately identified. Based on our the discussions in Section 3.4, the number of endmembers in this real data experiment is set to be 14.

Two existing algorithms, namely MVES and VCA were also tested with the same real data, and compared with the proposed RMVES algorithm. It should be noted that the non-pure pixel based algorithms are sensitive to initialization and hence their performances vary with the initialization. In order to have a fair comparison with our RMVES algorithm, the MVES algorithm was applied to the hyperspectral data with 10 different expanded VCA initializations (which are also feasible initializations for MVES) and the estimates associated with the least simplex volume is chosen as the results of MVES algorithm.

The subsampled real data discussed in Section 3.4 is used as input for RMVES and MVES algorithms, and the complete hyperspectral data set (without subsampling) is used for the estimation of the abundance maps. When applied to real hyperspectral data, the computational time for RMVES, MVES (both with subsampled data) and VCA (with full data) are around 145 minutes (which is around 8 hours on average for the complete data), 119 minutes and 4 minutes, respectively. The endmember signatures obtained via RMVES, MVES and VCA, along with the corresponding library signatures are shown in Figure 4.4. The minerals were identified by the visual comparison of the obtained abundance maps with the ones available in [20] [23] [51] [55] and [56]. The materials identified are arranged in alphabetical order for ease of visual comparison and the abundance maps obtained by the RMVES, MVES, and VCA algorithms are shown in Figure 4.5, Figure 4.6, and Figure 4.7, respectively. It should be mentioned that the difference in the materials identified by the three algorithms could possibly be due to the working nature of the respective algorithms and their sensitivity to initializations. The mean removed spectral angle between the estimated signature $\mathbf{a}_{est}$ and the corresponding library signature $\mathbf{a}_{lib}$ (defined in (3.27)) is used as the performance measure. The value of ϕ for the various minerals identified by the algorithms under test are given in Table 4.8. Again, the least ϕ value for an endmember is highlighted as a bold-faced number and the number in parenthesis is

the ϕ value for the repeatedly identified endmember. It can be seen from Table 4.8 that for the materials identified, the proposed RMVES algorithm performs best (with the minimum average ϕ), and further, it mostly yields better endmember estimates (minimum ϕ) than its predecessor MVES algorithm and the pure-pixel based VCA algorithm. Note that few of the mean removed spectral angles for endmembers (e.g., Buddingtonite, Kaolinite#1) identified by RMVES algorithm are marginally higher than that of the ones obtained by MVES. This could be attributed to some orientation of the simplex obtained by RMVES, with respect to the simplex of the true endmembers. In our simulations and real data experiments, HySiMe [45] was used to estimate the noise covariance matrix. Better estimation of the noise covariance matrix should further enhance the performance of the RMVES algorithm.

As mentioned in Section 3.4, we have also tried the HU algorithms for real data with artificially added noise. However, the majority of the so obtained abundance maps were inconclusive when compared with available groundtruths [20] [23], [51], [55], [56], and hence the mineral identification and quantification could not be performed for the endmember and abundance estimates obtained for the noise added AVIRIS real data.

4.5 Summary

We have presented a robust HU algorithm, namely RMVES (as shown in Table 4.2), for effective unmixing of mixed hyperspectral data corrupted by uniform or non-uniform Gaussian noise. The dimension reduction via affine set fitting procedure has been suitably modified for noisy hyperspectral observations. The randomness caused by noise has been dealt by incorporating chance constraints in the unmixing problem formulation with a design parameter η. A detailed analysis on the role of η has been presented, and it was concluded that η must be less than 0.5, which along with the

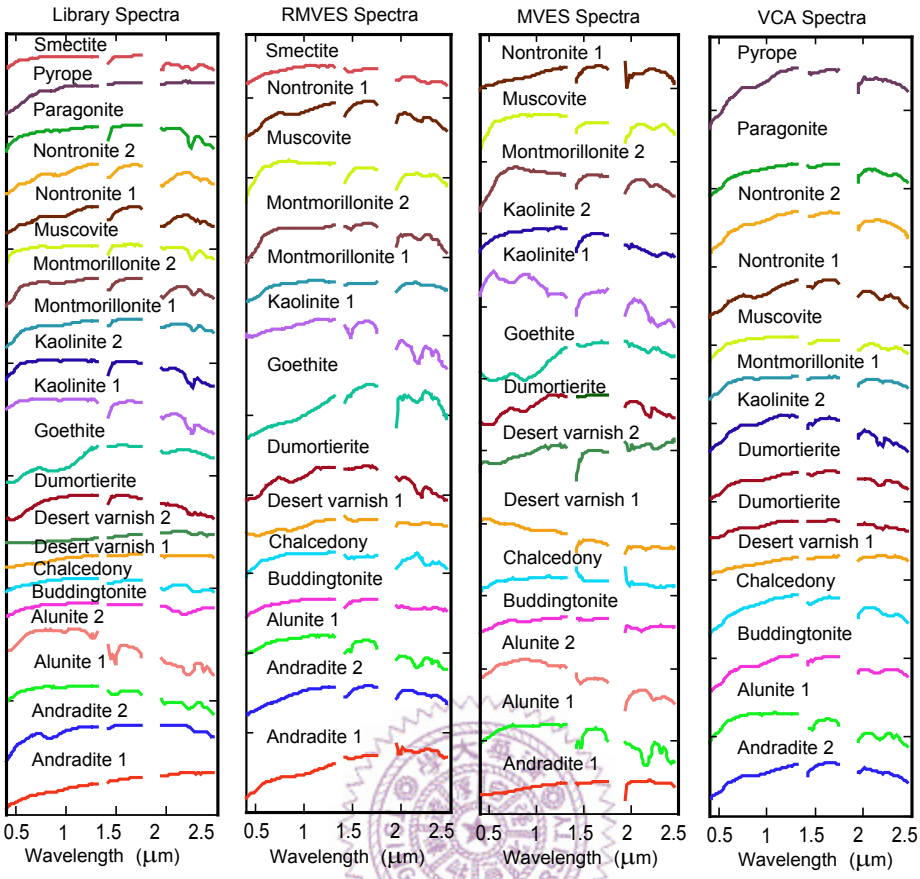


Figure 4.4. Endmember signatures taken from library, and the ones estimated by RMVES, MVES, and VCA algorithms.

objective function results in a non-convex optimization problem. To minimize the effect of non-convexity of the objective function, an alternating optimization concept has been utilized. The partial maximization problems involved are handled by using available sequential quadratic programming (SQP) solvers. Finally, Monte-Carlo simulations and real data experiments presented in Sections 4.3 and 4.4, respectively, demonstrate the superior performance of the proposed RMVES algorithm over several existing benchmark methods, including its predecessor, the MVES algorithm.

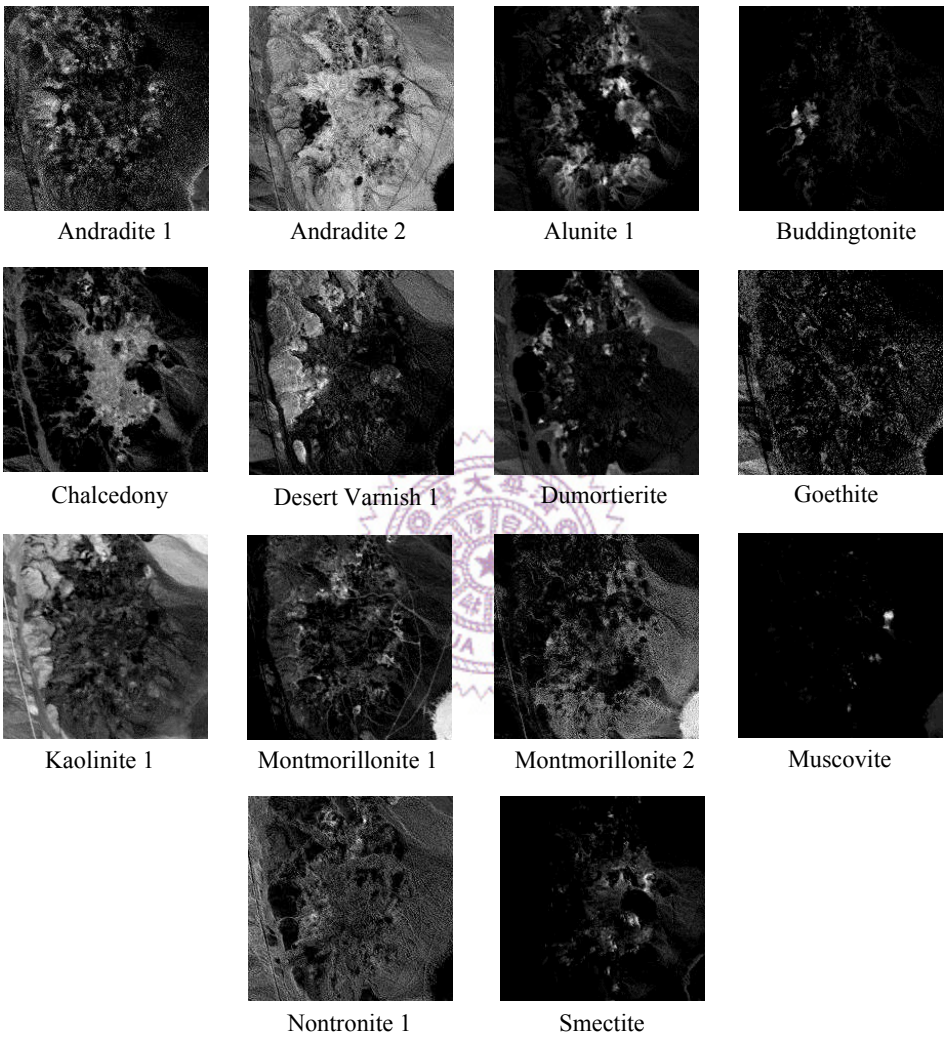


Figure 4.5. Abundance maps obtained by the RMVES algorithm.

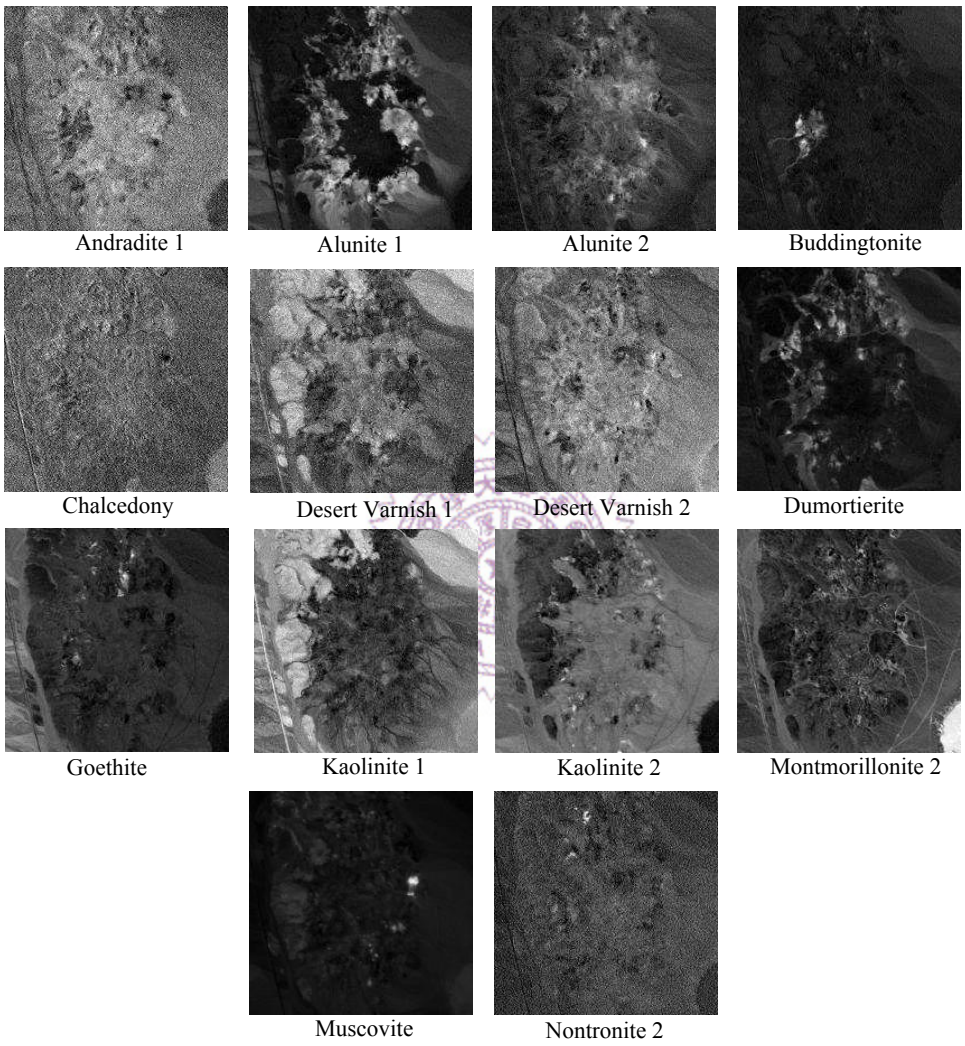


Figure 4.6. Abundance maps obtained by the MVES algorithm.

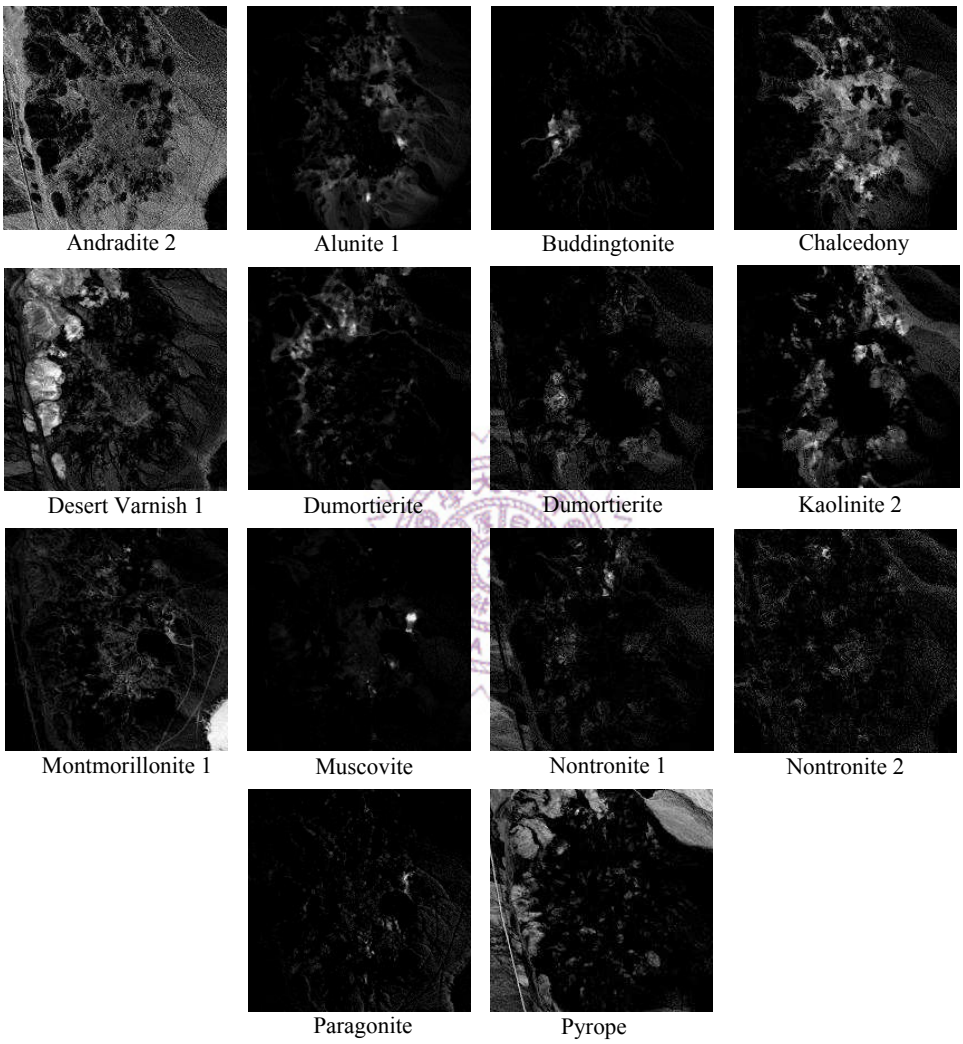


Figure 4.7. Abundance maps obtained by the VCA algorithm.

Table 4.8. Mean-removed spectral angles ϕ (degrees) between library spectra and endmembers estimated by RMVES, MVES, and VCA algorithms.

	RMVES	MVES	VCA
Andradite#1	**9.36**	25.61	-
Andradite#2	24.52	-	**18.49**
Alunite#1	**15.92**	21.85	17.74
Alunite#2	-	**17.72**	-
Buddingtonite	23.54	**22.98**	27.25
Chalcedony	**27.74**	38.25	31.9
Desert Varnish#1	20.99	18.64	**12.12**
Desert Varnish#2	-	**43.04**	-
Dumortierite	**20.77**	29.32	31.95 (32.01)
Goethite	**17.71**	19.05	-
Kaolinite#1	27.25	**26.50**	-
Kaolinite#2	-	**21.09**	32.49
Montmorillonite#1	22.99	-	**18.06**
Montmorillonite#2	**24.34**	26.00	-
Muscovite	39.63	44.19	**32.7**
Nontronite#1	**22.95**	-	24.66
Nontronite#2	-	28.83	**21.51**
Paragonite	-	-	**35.91**
Pyrope	-	-	**25.59**
Smectite	**22.53**	-	-
Average ϕ	**22.87**	27.11	25.73

Chapter 5

Estimation of Number of Endmembers Using $p-$norm Based Pure Pixel Identification Algorithm

The robust HU algorithms presented in Chapter 3 and Chapter 4, and the existing HU algorithms assume that the number of endmembers in the hyperspectral data is known *a priori*, which, however is not possible in practical scenarios. Hence, in this chapter, we propose a data geometry based approach to estimate the number of endmembers. Since, the dimension reduction technique introduced in Chapter 2 assumes that the number of endmembers N is known before hand, we first present the dimension reduction procedure for an unknown N. The geometric properties associated with the dimension-reduced hyperspectral data are then discussed. Based on the geometric properties we then move on propose two novel algorithms, namely geometry based estimation of number of endmembers - convex hull (GENE-CH) algorithm and affine hull (GENE-AH) algorithm. Since the estimation accuracies of the proposed GENE algorithms depend on the performance of the EEA used, a reliable, reproducible, and successive EEA (SEEA), called p-norm based pure pixel identification (TRI-P)

algorithm is then proposed. The performance of the proposed TRI-P algorithm, and the estimation accuracies of the GENE algorithms are demonstrated through Monte Carlo simulations and real data experiments.

5.1 Dimension reduction-revisited

To explain the dimension reduction technique with unknown N, let us first consider the noise-free scenario (i.e., $\mathbf{y}[n] = \mathbf{x}[n],\ \forall n$, and $\mathbf{U}_y = \mathbf{U}_x$). Recall from (2.4), (2.6), and (2.7) that

$$\mathbf{y}[n] \in \text{aff}\{\mathbf{a}_1, \ldots, \mathbf{a}_N\} = \text{aff}\{\mathbf{y}[1], \ldots, \mathbf{y}[L]\} = \mathcal{A}(\mathbf{C}, \mathbf{d}). \tag{5.1}$$

When N is known, an optimal solution for the affine-set fitting parameter $(\mathbf{C}, \mathbf{d}) \in \mathbb{R}^{M\times(N-1)} \times \mathbb{R}^M$ is given by (2.9) and (2.10) (with $\mathbf{U}_x = \mathbf{U}_y$). However, in practice, the number of endmembers N is unknown and ought to be estimated. Therefore, by assuming $N_{\max}$ to be a maximum bound on the number of endmembers, where $N \le N_{\max} \le M$, similar to the one in [23], we can obtain an affine set fitting parameter $(\boldsymbol{\mathcal{C}}, \mathbf{d}) \in \mathbb{R}^{M\times(N_{\max}-1)} \times \mathbb{R}^M$, as below:

$$\boldsymbol{\mathcal{C}} = [\ \boldsymbol{q}_1(\mathbf{U}_y\mathbf{U}_y^T), \ldots, \boldsymbol{q}_{N_{\max}-1}(\mathbf{U}_y\mathbf{U}_y^T)\ ], \tag{5.2}$$

and $\mathbf{d}$ is defined in (2.9). From (2.10) and (5.2), it can be easily verified that

$$\mathbf{y}[n] \in \mathcal{A}(\mathbf{C}, \mathbf{d}) \subseteq \mathcal{A}(\boldsymbol{\mathcal{C}}, \mathbf{d}). \tag{5.3}$$

Then, by virtue of (2.4), and since $\boldsymbol{\mathcal{C}}$ given by (5.2) is semi-unitary, the dimension-reduced pixel vectors $\tilde{\mathbf{y}}[n]$ can be obtained by the following affine transformation of

$\mathbf{y}[n]$

$$\tilde{\mathbf{y}}[n] = \boldsymbol{\mathcal{C}}^T(\mathbf{y}[n] - \mathbf{d}) \in \mathbb{R}^{N_{\max}-1}. \tag{5.4}$$

For the noisy scenario, similar to (2.20) and (2.21), the approximate affine set fitting parameter $(\widehat{\boldsymbol{\mathcal{C}}}, \widehat{\mathbf{d}})$ can be obtained as

$$\widehat{\mathbf{d}} = \frac{1}{L}\sum_{n=1}^{L}\mathbf{y}[n] = \frac{1}{L}\sum_{n=1}^{L}\mathbf{x}[n] + \frac{1}{L}\sum_{n=1}^{L}\mathbf{w}[n] \cong \mathbf{d}, \tag{5.5}$$

$$\widehat{\boldsymbol{\mathcal{C}}} = [\ \boldsymbol{q}_1(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}}), \ldots, \boldsymbol{q}_{N_{\max}-1}(\mathbf{U}_y\mathbf{U}_y^T - L\widehat{\mathbf{D}})\ ] \cong \boldsymbol{\mathcal{C}}, \tag{5.6}$$

where $\widehat{\mathbf{D}}$ is defined as

$$\widehat{\mathbf{D}} = \frac{1}{L}\mathbf{W}\mathbf{W}^T, \tag{5.7}$$

is an estimate of the noise covariance matrix. As shown in [62], for given $\widehat{\mathbf{D}}$, the affine set fitting solution $(\widehat{\boldsymbol{\mathcal{C}}}, \widehat{\mathbf{d}})$ (given by (2.20) and (2.21)) can be shown to be an approximation to the true $(\boldsymbol{\mathcal{C}}, \mathbf{d})$ and it asymptotically approaches the true $(\boldsymbol{\mathcal{C}}, \mathbf{d})$ for large L. In practical situations, the multiple regression analysis based noise covariance estimation method reported in HySiMe [45] can be used to estimate $\widehat{\mathbf{D}}$. Thus, in the noisy scenario, the dimension reduced observations can be represented as

$$\tilde{\mathbf{y}}[n] = \widehat{\boldsymbol{\mathcal{C}}}^T(\mathbf{y}[n] - \widehat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}, \tag{5.8}$$

in which $\widehat{\mathbf{d}}$ and $\widehat{\boldsymbol{\mathcal{C}}}$, are given by (5.5) and (5.6), respectively. Further, due to (2.1), (2.2), and (A2), we have

$$\tilde{\mathbf{y}}[n] = \tilde{\mathbf{x}}[n] + \tilde{\mathbf{w}}[n],\ n = 1, \ldots, L, \tag{5.9}$$

where

$$\tilde{\mathbf{x}}[n] = \sum_{i=1}^{N} s_i[n]\boldsymbol{\alpha}_i,\ n = 1, \ldots, L, \tag{5.10}$$

in which

$$\boldsymbol{\alpha}_i = \widehat{\boldsymbol{\mathcal{C}}}^T(\mathbf{a}_i - \widehat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1},\ i = 1, \ldots, N \tag{5.11}$$

is the ith dimension-reduced endmember, and $\tilde{\mathbf{w}}[n] \triangleq \widehat{\boldsymbol{\mathcal{C}}}^T\mathbf{w}[n] \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$, in which

$$\boldsymbol{\Sigma} = \widehat{\boldsymbol{\mathcal{C}}}^T\mathbf{D}\widehat{\boldsymbol{\mathcal{C}}} \in \mathbb{R}^{(N_{\max}-1)\times(N_{\max}-1)}. \tag{5.12}$$

Note that the dimension-reduced endmember defined in (5.11) is different from the one defined in earlier chapters as the dimension now is $\mathbb{R}^{N_{\max}-1}$ instead of $\mathbb{R}^{N-1}$. The relation between the dimension-reduced endmember $\boldsymbol{\alpha}_i$ and the true endmember $\mathbf{a}_i$ is then given by

$$\mathbf{a}_i = \widehat{\boldsymbol{\mathcal{C}}}\boldsymbol{\alpha}_i + \widehat{\mathbf{d}},\ i = 1, \ldots, N. \tag{5.13}$$

It is worth mentioning that under (A4), from (5.9) and (5.10), we have

$$\tilde{\mathbf{y}}[l_i] = \boldsymbol{\alpha}_i + \tilde{\mathbf{w}}[l_i],\ \forall i = 1, \ldots, N, \tag{5.14}$$

which is essential in the development of the proposed GENE algorithms in Section 5.2.

5.1.1 Convex geometry of the hyperspectral data

In this subsection we will present the convex geometry of the noise-free dimension-reduced hyperspectral data given by (5.10). The convex geometry will lay a solid platform for the ensuing sections, though the presence of noise will be taken into

account therein. Based on (5.10), we have the following facts:

(F1) In the noise-free case, by (A1)-(A4), any dimension-reduced pixel vectors $\tilde{\mathbf{x}}[n]$ lie in the convex hull of the dimension-reduced endmember signatures, and

$$\text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\} = \text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}, \tag{5.15}$$

in which $\text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ is a simplex with N extreme points being $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N$. A more general case of (F1) can be obtained by relaxing (A1) and (A4), as stated next.

(F2) In the noise-free case, by (A2) and (A3), any dimension-reduced pixel vectors $\tilde{\mathbf{x}}[n]$ lie in the affine hull of the dimension-reduced endmember signatures, and

$$\text{aff}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\} = \text{aff}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}, \tag{5.16}$$

with the affine dimension equal to $N - 1$.

A simple illustration of (F1) and (F2), for $N = 3$ case is shown in Figure 5.1. These geometrical properties of the observed hyperspectral data play a significant role in the proposed algorithms for estimating the number of endmembers, which will be presented in the next section.

5.2 Geometry based estimation of number of endmembers (GENE)

In this section, we make use of the key geometric characteristics of the observed dimension-reduced hyperspectral data, i.e (F1) and (F2) to systematically estimate the number of endmembers present in the hyperspectral data. Thus the proposed

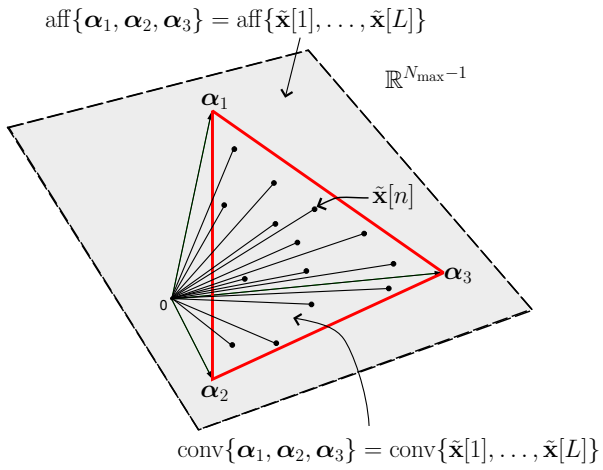


Figure 5.1. Illustration of (F1) and (F2), for $N = 3$ case.

algorithms are aptly named as geometry based estimation of number of endmembers (GENE) algorithms. In the first subsection, we propose the GENE-CH algorithm based on the convex hull geometry (F1), provided that (A4) holds true. However, for data with (A4) violated, the GENE-AH algorithm is proposed in the subsequent subsection. GENE-AH is based on (F2) and it turns out to be more robust than the GENE-CH algorithm against absence of pure pixels in the data (which will be confirmed by the simulations in Section 5.4).

5.2.1 GENE-Convex Hull (GENE-CH) Algorithm

In this subsection we assume that (A4) holds true. Suppose that a reliable, SEEA has found the pixel indices $l_1, \ldots, l_N, l_{N+1}, \ldots, l_{k-1}, l_k$, in which $l_1, \ldots, l_N$ are pure pixel indices. Here, l_k is the current pixel index estimate and $\{l_1, l_2, \ldots, l_{k-1}\}$ are the previously found pixel index estimates, and $k \leq N_{\max}$. Then by (5.9) and (5.14), it can be readily inferred that

$$\tilde{\mathbf{y}}[l_i] = \tilde{\mathbf{x}}[l_i] + \tilde{\mathbf{w}}[l_i], \ i = 1, \ldots, k, \tag{5.17}$$

where

$$\tilde{\mathbf{x}}[l_i] = \begin{cases} \boldsymbol{\alpha}_i, & i = 1, \ldots, N, \\ \sum_{j=1}^{N} s_j[l_i]\boldsymbol{\alpha}_j, & i = N+1, \ldots, k. \end{cases} \tag{5.18}$$

To explain the idea behind GENE-CH, let us first consider the noise-free scenario, i.e., $\tilde{\mathbf{w}}[l_i] = \mathbf{0}$, $\forall i = 1, \ldots, k$, in (5.17). Recall from (F1) that the total number of extreme points in $\text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$ is N. That is to say, if $\tilde{\mathbf{y}}[l_k] = \tilde{\mathbf{x}}[l_k]$ cannot contribute a new extreme point to the $\text{conv}\{\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}], \tilde{\mathbf{y}}[l_k]\}$, i.e.,

$$\text{conv}\{\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}], \tilde{\mathbf{y}}[l_k]\} = \text{conv}\{\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}]\} = \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}, \tag{5.19}$$

or in other words, if $\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, then it can be inferred by (5.18) that all the endmembers are already found, that is $k \geq N+1$. Therefore, the smallest k such that $\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, must take the value of $N+1$, and thus N can be estimated as $k-1$, provided that the smallest k can be reliably estimated. However, in a real scenario, since only noisy $\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_k]$ are available (rather than $\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_k]$), in the process of estimating the number of endmembers, the presence of noise in the $\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_k]$ must be taken into account. To this end, we propose a Neyman-Pearson hypothesis [63] testing based method to determine whether $\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, or not, based on noisy $\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_k]$. The idea is to find the smallest k for which $\tilde{\mathbf{y}}[l_k]$ is closest to $\text{conv}\{\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}]\}$.

Let us consider the following constrained least squares problem:

$$\boldsymbol{\theta}^\star = \arg \min_{\boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_{k-1}^T \boldsymbol{\theta} = 1} \quad \|\tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}\|_2^2, \tag{5.20}$$

where

$$\widehat{\mathbf{A}}_{k-1} = [\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}]] \in \mathbb{R}^{(N_{\max}-1)\times(k-1)}. \tag{5.21}$$

The optimization problem in (5.20) is convex and can be solved by using available convex optimization solvers such as `SeDuMi` [30] and `CVX` [31]. We define the fitting error vector $\boldsymbol{e}$ as below:

$$\begin{aligned} \boldsymbol{e} = & \ \tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}^\star & (5.22)\\ = & \ \boldsymbol{\mu}_k + \left(\tilde{\mathbf{w}}[l_k] - \sum_{i=1}^{k-1} \theta_i^\star \tilde{\mathbf{w}}[l_i]\right) \in \mathbb{R}^{N_{\max}-1}, \quad \text{(by (5.17))} & (5.23) \end{aligned}$$

where

$$\boldsymbol{\mu}_k = \tilde{\mathbf{x}}[l_k] - \sum_{i=1}^{k-1} \theta_i^\star \tilde{\mathbf{x}}[l_i]. \tag{5.24}$$

Then the following can be observed from (5.23):

- If $\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, then it implies that $\tilde{\mathbf{x}}[l_k] - \sum_{i=1}^{k-1} \theta_i' \tilde{\mathbf{x}}[l_i] = \mathbf{0}$, for some $\boldsymbol{\theta}' = [\theta_1', \ldots, \theta_{k-1}']^T \succeq \mathbf{0}$, $\mathbf{1}_{k-1}^T \boldsymbol{\theta}' = 1$. Since $\boldsymbol{\theta}^\star \simeq \boldsymbol{\theta}'$ makes $\boldsymbol{\mu}_k \simeq \mathbf{0}$ (due to $\boldsymbol{\mu}_k = \mathbf{0}$ for the noise-free case), then $\boldsymbol{e}$ can be approximated as a zero-mean Gaussian random vector, i.e., $\boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \xi^\star \boldsymbol{\Sigma})$, where

  $$\xi^\star = 1 + \theta_1^{\star 2} + \theta_2^{\star 2} + \cdots + \theta_{k-1}^{\star 2}, \tag{5.25}$$

  and $\boldsymbol{\Sigma}$ is given by (5.12).

- If $\tilde{\mathbf{x}}[l_k] \notin \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, then $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{\mu}_k, \xi^\star \boldsymbol{\Sigma})$ is a non-zero mean Gaussian random vector.

Now define

$$r = \boldsymbol{e}^T(\xi^\star \boldsymbol{\Sigma})^{-1}\boldsymbol{e}. \tag{5.26}$$

When $\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}$, it is easy to see that r can be approximated as a central Chi-square distributed random variable, and otherwise r is a non-central Chi-square distributed random variable [64]. In both cases, the degrees of freedom is $N_{\max} - 1$. Hence, we consider the following two hypotheses:

$$H_0 \ (\tilde{\mathbf{x}}[l_k] \in \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}) : \ r \sim f_{\chi^2}(x, N_{\max} - 1) \tag{5.27a}$$

$$H_1 \ (\tilde{\mathbf{x}}[l_k] \notin \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}) : \ r \sim f_{N\chi^2}(x, N_{\max} - 1, \|\boldsymbol{\mu}_k\|_2^2). \tag{5.27b}$$

Here, $f_{\chi^2}(x, N_{\max}-1)$ is the pdf of central Chi-square distribution and is given by [64]

$$f_{\chi^2}(x, k) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{(k/2)-1} e^{-x/2}, & x \geq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{5.28}$$

where $\Gamma(k/2)$ denotes the Gamma function. However, the non-central Chi-square pdf $f_{N\chi^2}(x, N_{\max} - 1, \|\boldsymbol{\mu}_k\|_2^2)$ is unknown, as $\boldsymbol{\mu}_k$ is unknown. Therefore, we use Neyman-Pearson classifier rule for the hypothesis testing problem:

$$\text{Decide } H_0 \ \text{ if } \ r < \kappa \tag{5.29a}$$

$$\text{Decide } H_1 \ \text{ if } \ r > \kappa, \tag{5.29b}$$

where κ can be found by minimizing the $P(H_0|H_1)$ subject to $P(H_1|H_0) \leq P_{\text{FA}}$, in which P_{FA} is the preassigned acceptable false alarm rate. Obviously, the optimal

value of κ should satisfy [63]

$$\int_{\kappa}^{\infty} f_{\chi^2}(x, N_{\max} - 1)dx = P_{\text{FA}}.$$

Hence the decision rules in (5.29) can be equivalently written as

$$\text{Decide } H_0 \text{ if } \int_{r}^{\infty} f_{\chi^2}(x, N_{\max} - 1)dx > \int_{\kappa}^{\infty} f_{\chi^2}(x, N_{\max} - 1)dx = P_{\text{FA}}, \quad (5.30a)$$

$$\text{Decide } H_1 \text{ if } \int_{r}^{\infty} f_{\chi^2}(x, N_{\max} - 1)dx < P_{\text{FA}}. \quad (5.30b)$$

The integral in (5.30) can be easily computed as follows:

$$\int_{r}^{\infty} f_{\chi^2}(x, N_{\max} - 1)dx = 1 - \frac{\gamma(r/2, (N_{\max} - 1)/2)}{\Gamma((N_{\max} - 1)/2)}, \quad (5.31)$$

where $\gamma(x/2, (N_{\max} - 1)/2)$ is the lower incomplete Gamma function [65]. Once the integral is evaluated, one of the hypotheses should be true, based on (5.30). The entire procedure for GENE-CH is summarized in Table 5.1. The working strategy of GENE-CH algorithm in synchronization with a suitable SEEA is demonstrated in Figure 5.2.

5.2.2 GENE-Affine Hull (GENE-AH) Algorithm

Recall that the GENE-CH algorithm is based on the assumption that the pure pixels are present in the data (i.e., (A4) holds true). However, for practical hyperspectral data the presence of pure pixels cannot be guaranteed. In this case, the dimension-reduced endmembers estimated by an EEA can be expressed in general as in (5.17),

Table 5.1. Pseudocode for GENE-CH and GENE-AH algorithms.

Given noisy hyperspectral data $\mathbf{y}[n]$, sufficiently large maximum number of endmembers $N \leq N_{\max} \leq M$, false alarm probability P_{FA}, and estimate of noise covariance matrix $\widehat{\mathbf{D}}$.

Step 1. Compute $(\widehat{\boldsymbol{\mathcal{C}}}, \widehat{\mathbf{d}})$ given by (5.5) and (5.6).

Step 2. Obtain the first pixel index l_1 by a SEEA and compute $\tilde{\mathbf{y}}[l_1] = \widehat{\boldsymbol{\mathcal{C}}}^T(\mathbf{y}[l_1] - \widehat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}$. Set $k = 2$.

Step 3. Obtain the kth pixel index l_k using the SEEA and compute $\tilde{\mathbf{y}}[l_k] = \widehat{\boldsymbol{\mathcal{C}}}^T(\mathbf{y}[l_k] - \widehat{\mathbf{d}}) \in \mathbb{R}^{N_{\max}-1}$ and form $\widehat{\mathbf{A}}_{k-1} = [\tilde{\mathbf{y}}[l_1], \ldots, \tilde{\mathbf{y}}[l_{k-1}]] \in \mathbb{R}^{(N_{\max}-1)\times(k-1)}$.

Step 4. Solve the following:

$$\mathrm{GENE-CH}: \boldsymbol{\theta}^\star = \arg \min_{\boldsymbol{\theta} \succeq \mathbf{0}, \mathbf{1}_{k-1}^T \boldsymbol{\theta} = 1} \|\tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}\|_2^2,$$

$$\mathrm{GENE-AH}: \boldsymbol{\theta}^\star = \arg \min_{\mathbf{1}_{k-1}^T \boldsymbol{\theta} = 1} \|\tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}\|_2^2,$$

and calculate $\boldsymbol{e} = \tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}^\star$.

Step 5. Compute $r = \boldsymbol{e}^T(\xi^\star \boldsymbol{\Sigma})^{-1}\boldsymbol{e}$, where $\xi^\star = 1 + \boldsymbol{\theta}^{\star T}\boldsymbol{\theta}^\star$ and $\boldsymbol{\Sigma} = \widehat{\boldsymbol{\mathcal{C}}}^T \mathbf{D} \widehat{\boldsymbol{\mathcal{C}}}$.

Step 6. Calculate $\psi = \int_r^\infty f_{\chi^2}(x, N_{\max} - 1)dx$ by (5.31).

Step 7. If $\psi > P_{\mathrm{FA}}$, then output $k-1$ as the estimate for number of endmembers, else $k := k + 1$ and if $k \leq N_{\max}$ go to **Step 3**.

---

where

$$\tilde{\mathbf{x}}[l_i] = \sum_{j=1}^{N} s_j[l_i]\boldsymbol{\alpha}_j, \ \forall i = 1, \ldots, k. \tag{5.32}$$

For such hyperspectral data, GENE-CH may not provide an accurate estimate of the number of endmembers. A pictorial illustration is given in Figure 5.3, where $N = 3$ endmembers $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$ are not present in the noise-free hyperspectral data. For this case, the endmember estimates, denoted by $\tilde{\mathbf{x}}[l_i], i = 1, ..., N_{\max} = 6$, obtained by an EEA are shown in Figure 5.3(a) and can be expressed as

$$\tilde{\mathbf{x}}[l_i] = \sum_{j=1}^{3} s_j[l_i]\boldsymbol{\alpha}_j, \ i = 1, \ldots, N_{\max} = 6, \quad \text{(by (5.17) and (5.32))} \tag{5.33}$$

where $l_1, \ldots, l_6$ are the pixel indices provided by the EEA under consideration. Then, as can be inferred from Figure 5.3(a), for the $\text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_6]\}$, there can be more than 3 extreme points which in fact is 6 in this case. In other words,

$$\tilde{\mathbf{x}}[l_k] \notin \text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}, \ k = 2, 3, 4, 5, 6, \tag{5.34}$$

which means that the hypothesis H_1 given by (5.27b) will be true even for $k > N = 3$. Hence, using the fact (F1) will obviously result in an overestimation of the number of endmembers for this case. However, from Figure 5.3(b), it can be readily inferred that

$$\tilde{\mathbf{x}}[l_k] \notin \text{aff}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}, \ k = 2, 3, \text{ and}$$
$$\tilde{\mathbf{x}}[l_k] \in \text{aff}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_{k-1}]\}, \ k = 4, 5, 6.$$

Motivated by the above illustration, we next propose the GENE-AH algorithm.

The GENE-AH algorithm uses the fact (F2), which states that in the noise-free case, the affine dimension of $\text{aff}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$ is $N - 1$. This implies that in the noise-free case, if $\tilde{\mathbf{y}}[l_k] = \tilde{\mathbf{x}}[l_k]$ cannot contribute an increment to the affine dimension

of $\text{aff}\{\tilde{\mathbf{y}}[l_1],\ldots,\tilde{\mathbf{y}}[l_{k-1}],\tilde{\mathbf{y}}[l_k]\}$, i.e.,

$$\text{aff}\{\tilde{\mathbf{y}}[l_1],\ldots,\tilde{\mathbf{y}}[l_{k-1}],\tilde{\mathbf{y}}[l_k]\} = \text{aff}\{\tilde{\mathbf{y}}[l_1],\ldots,\tilde{\mathbf{y}}[l_{k-1}]\} = \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}, \quad (5.35)$$

or in other words, if $\tilde{\mathbf{x}}[l_k] \in \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}$, then $k \geq N+1$. Therefore, the smallest k such that $\tilde{\mathbf{x}}[l_k] \in \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}$, must take the value of $N+1$, and thus N can be estimated as $k-1$. As presented in Section 5.2.1, again we use Neyman-Pearson hypothesis [63] testing to determine whether $\tilde{\mathbf{x}}[l_k] \in \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}$, or not, based on noisy $\tilde{\mathbf{y}}[l_1],\ldots,\tilde{\mathbf{y}}[l_k]$. The idea is to find the smallest k for which $\tilde{\mathbf{y}}[l_k]$ is closest to $\text{aff}\{\tilde{\mathbf{y}}[l_1],\ldots,\tilde{\mathbf{y}}[l_{k-1}]\}$. The details are as follows:

As in (5.20), we consider solving the following constrained least squares problem:

$$\boldsymbol{\theta}^\star = \arg \min_{\mathbf{1}_{k-1}^T\boldsymbol{\theta}=1} \|\tilde{\mathbf{y}}[l_k] - \widehat{\mathbf{A}}_{k-1}\boldsymbol{\theta}\|_2^2, \quad (5.36)$$

where $\widehat{\mathbf{A}}_{k-1}$ is defined in (5.21). Again, since (5.36) is convex, $\boldsymbol{\theta}^\star$ can be obtained by available convex optimization solvers [30, 31]. By defining the fitting error vector $\boldsymbol{e}$ as in (5.22), we have the following inferences:

- if $\tilde{\mathbf{x}}[l_k] \in \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}$, then it can be approximated that $\boldsymbol{e} \sim \mathcal{N}(\mathbf{0}, \xi^\star\boldsymbol{\Sigma})$.
- if $\tilde{\mathbf{x}}[l_k] \notin \text{aff}\{\tilde{\mathbf{x}}[l_1],\ldots,\tilde{\mathbf{x}}[l_{k-1}]\}$, then $\boldsymbol{e} \sim \mathcal{N}(\boldsymbol{\mu}_k, \xi^\star\boldsymbol{\Sigma})$,

where $\boldsymbol{\mu}_k$, $\xi^\star$, and $\boldsymbol{\Sigma}$ are defined in (5.24), (5.25), and (5.12), respectively. Defining the random variable r as in (5.26), a similar Neyman-Pearson hypothesis testing procedure can be devised for GENE-AH to estimate the number of endmembers present in the data. As will be seen from the simulations (see Section 5.4), the GENE-AH algorithm yields better performance demonstrating its robustness against the absence of pure pixels in the data. The procedure for GENE-AH is similar to the one in Table 5.1, except that in Step 4, the optimal $\boldsymbol{\theta}^\star$ is obtained by solving (5.36).

Some conceptual distinctions between the the proposed GENE algorithms and some existing benchmark algorithms are worth discussion here. Most of the available algorithms for the estimation of number of endmembers, such as HySiMe [45] and orthogonal subspace projection (OSP) [66], are directly or indirectly based on the projection power on the most apt range space of the hyperspectral data, which in turn is based only on (A3). In other words those methods consider the following fact:

$$\tilde{\mathbf{x}}[l_i] \in \text{range}(\mathbf{G}), \tag{5.37}$$

where $\mathbf{G} = [\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N]$. However, GENE-CH and GENE-AH involve convex hull and affine hull i.e.,

$$\tilde{\mathbf{x}}[l_i] \in \text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} \subset \text{aff}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} \subset \text{range}(\mathbf{G}), \tag{5.38}$$

as they not only make use of (A3), but also (A1) for GENE-AH algorithm, and (A1) and (A2) for GENE-CH algorithm. The advantage of considering the assumptions on abundances will be more evident in the simulation and real data experiment sections (i.e., Section 5.4 and Section 5.5, respectively).

It should be noted that the estimation accuracies of both the GENE-CH and GENE-AH algorithms depend on the performance of the EEA used. Hence, in the next section, we propose a reliable (with theoretical support for endmember identifiability), reproducible (without any initialization), and SEEA, namely p-norm based pure pixel identification (TRI-P) algorithm.

5.3 p-norm based Pure Pixel Identification (TRI-P) Algorithm

The proposed TRI-P algorithm aims to find the pure pixel indices (and thereby the endmembers) from the hyperspectral observations. Throughout the derivation of the TRI-P algorithm, we focus on a dimension-reduced noise-free signal model (given by (5.4)) by means of which the endmember identifiability of the TRI-P algorithm can be theoretically proved (noisy scenarios will be considered in simulations in Section 5.4.1). We begin by incorporating the assumption (A2) in (5.9) so that we have the following augmented dimension-reduced data:

$$\bar{\mathbf{x}}[n] = \begin{bmatrix} \tilde{\mathbf{x}}[n] \\ 1 \end{bmatrix} = \sum_{i=1}^{N} s_i[n]\bar{\boldsymbol{\alpha}}_i \in \mathbb{R}^{N_{\max}}, \tag{5.39}$$

where

$$\bar{\boldsymbol{\alpha}}_i = [\boldsymbol{\alpha}_i^T 1]^T, \ i = 1, \ldots, N, \tag{5.40}$$

are the augmented dimension-reduced endmembers.

We now find the first dimension-reduced endmember by p-norm maximization procedure as follows: Considering the p-norm of all the pixel vectors in the augmented dimension-reduced data cloud $\bar{\mathbf{X}} = [\ \bar{\mathbf{x}}[1], \ldots, \bar{\mathbf{x}}[L]\ ]$, by the triangle inequality, (A1), and (A2), one can infer from (5.39) that for all n,

$$\|\bar{\mathbf{x}}[n]\|_p \leq \sum_{i=1}^{N} s_i[n] \cdot \|\bar{\boldsymbol{\alpha}}_i\|_p \leq \max_{i=1,\ldots,N}\{\|\bar{\boldsymbol{\alpha}}_i\|_p\}, \tag{5.41}$$

where $p \geq 1$. The equality in (5.41) holds if and only if $n = l_i$ (by (A4)) for any $i \in \arg\max_{k=1,\ldots,N}\{\|\bar{\boldsymbol{\alpha}}_k\|_p\}$. Thus, a dimension-reduced endmember can be identified

by

$$\boldsymbol{\alpha}_1 = \tilde{\mathbf{x}}[l_1] \text{ for any } l_1 \in \arg\max_{n=1,\ldots,L}\{\|\bar{\mathbf{x}}[n]\|_p\}. \tag{5.42}$$

Once the first endmember is found, the other endmembers $\boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_N$ can be obtained successively by the following general procedure: Suppose that the augmented dimension-reduced endmembers (given by (5.40)) $\bar{\boldsymbol{\alpha}}_1, \ldots, \bar{\boldsymbol{\alpha}}_{k-1}$ (where $k-1 < N$) are already identified. Let

$$\mathbf{Q} = [\bar{\boldsymbol{\alpha}}_1, \ldots, \bar{\boldsymbol{\alpha}}_{k-1}] \in \mathbb{R}^{N_{\max}\times(k-1)}. \tag{5.43}$$

To find a new endmember different from $\bar{\boldsymbol{\alpha}}_1, \ldots, \bar{\boldsymbol{\alpha}}_{k-1}$, we consider the following orthogonal complement subspace projection:

$$\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\mathbf{x}}[n] = \sum_{i=k}^{N} s_i[n]\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\boldsymbol{\alpha}}_i, \forall n, \quad \text{(by (5.39))} \tag{5.44}$$

where $\mathbf{P}_{\mathbf{Q}}^{\perp} = \mathbf{I}_{N_{\max}} - \mathbf{Q}(\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T$ is the orthogonal complement projector of $\mathbf{Q}$. Again, by the triangle inequality, (A1), (A2), and (5.44), we have

$$\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\mathbf{x}}[n]\|_p \le \sum_{i=k}^{N} s_i[n] \cdot \|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\boldsymbol{\alpha}}_i\|_p \le \max_{i=k,\ldots,N}\{\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\boldsymbol{\alpha}}_i\|_p\}. \tag{5.45}$$

The equality in (5.45) holds if and only if $n = l_j$ (by (A4)) for any $j \in \arg\ \max_{i=k,\ldots,N}\{\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\boldsymbol{\alpha}}_i\|_p\}$. Therefore, one can find a new dimension-reduced endmember as

$$\boldsymbol{\alpha}_j = \tilde{\mathbf{x}}[l_j] \text{ for any } l_j \in \arg\max_{n=1,\ldots,L}\{\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\mathbf{x}}[n]\|_p\}, \tag{5.46}$$

and $\boldsymbol{\alpha}_j = \tilde{\mathbf{x}}[l_j] \notin \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{k-1}\}$.

The above endmember estimation methodology is called the TRI-P algorithm, and it can identify all the dimension-reduced endmembers, as stated in the following

lemma.

Lemma 2. *Under (A1)-(A4), with N known, and in the absence of noise, TRI-P algorithm yields $\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\}$ such that the simplex* $\text{conv}\{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_N\} = \text{conv}\{\tilde{\mathbf{x}}[1], \ldots, \tilde{\mathbf{x}}[L]\}$.

Once the dimension-reduced endmembers are found, the corresponding endmembers can be obtained by (5.13).

At this juncture, it is worthwhile to point out some characteristics of the proposed TRI-P algorithm.

- Lemma 1 is valid only if (A4) is satisfied and the number of endmembers N is perfectly known. However, if (A4) is satisfied and N is unknown (which is the case in the proposed GENE-CH algorithm), then $N_{\max}$ pixel indices can be obtained by TRI-P algorithm, where the first N pixel indices will be a set of pure pixel indices. In this case, an interesting question is: What will be the pixel index l_k obtained by TRI-P algorithm when $k > N$? It can be shown from (5.9) and (A3) that in the noiseless case $\bar{\mathbf{x}}[n] \in \text{range}(\bar{\mathbf{G}})$, where $\bar{\mathbf{G}} = [\bar{\boldsymbol{\alpha}}_1, \ldots, \bar{\boldsymbol{\alpha}}_N]$. Therefore, under (A1)-(A4), when finding $\boldsymbol{\alpha}_k$ for $k > N$, we have by (5.43) that $\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\mathbf{x}}[n]\|_p = 0, \ \forall n$. Equation (5.46) will therefore yield a pixel index l_k for which

$$\tilde{\mathbf{x}}[l_k] = \sum_{j=1}^{N} s_j[l_k]\boldsymbol{\alpha}_j, \ \forall k > N, \tag{5.47}$$

  can be any pixel in the data cloud.

- On the other hand, if (A4) is not satisfied and N is perfectly known. Then the set of pixel indices $\{l_1, \ldots, l_N\}$ corresponds to a set of so-called "purest pixels" available in the data cloud.

- Finally, if both (A4) is not satisfied and N is unknown (which is the case in the

proposed GENE-AH algorithm), then still $N_{\max}$ pixel indices can be obtained, where the first N pixel indices will be indices corresponding to a set of so-called "purest pixels" in the data cloud, whereas pixel index l_k when $k > N$ can be any pixel in the data cloud. That is,

$$\tilde{\mathbf{x}}[l_k] = \sum_{j=1}^{N} s_j[l_k]\boldsymbol{\alpha}_j, \ \forall k > N. \tag{5.48}$$

- While other existing pure-pixel based EEAs such as pixel purity index (PPI) [8], vertex component analysis (VCA) [12], N-FINDR [9] [67] [68] [69], and simplex growing algorithm (SGA) [10] require initializations (though the outcome of SGA is insensitive to initialization [10]), TRI-P does not require initialization of any form, and hence the solution is unique for a given hyperspectral data. Therefore, the endmember estimates of the TRI-P algorithms are reproducible even in the presence of noise, i.e., they always yield the same endmember estimates for a given hyperspectral data.

- It can be easily shown that both SVMAX algorithm [14] which is developed based on Winter's unmixing criterion, and ATGP algorithm [70] which is based on target detection are the special cases of the pixel search based TRI-P algorithm with $p = 2$.

The entire TRI-P algorithm is summarized in Table 5.2. Though the TRI-P algorithm in Table 5.2 alone can yield a set of pixel indices $\{l_1, \ldots, l_{N_{\max}}\}$, if used in conjunction with the GENE algorithms (presented in Table 5.1), the pixel index l_1 needed in Step 2 of Table 5.1 is provided by Step 1 of Table 5.2, while l_k needed in Step 3 of Table 5.1 for any $1 < k \leq N_{\max}$ is successively provided by Step 2 of Table 5.2, for each $k > 1$. In other words, TRI-P algorithm with the above characteristics serves as a good candidate for successively providing the pixel indices to feed the

proposed GENE algorithms in a synchronous fashion.

Table 5.2. Pseudocode for p-norm based pure pixel (TRI-P) algorithm.

Given dimension-reduced observations $\tilde{\mathbf{x}}[n]$, $\bar{\mathbf{x}}[n]$ given by (5.39), and maximum number of endmembers $N_{\max}$.

Step 1. Set $k = 1$. Obtain $\bar{\boldsymbol{\alpha}}_1 = \bar{\mathbf{x}}[l_1]$ for any $l_1 \in \arg\max_n\{\|\bar{\mathbf{x}}[n]\|_p\}$. Let $\mathbf{Q} = \bar{\boldsymbol{\alpha}}_1$.

Step 2. Update $k := k + 1$ and obtain $\bar{\boldsymbol{\alpha}}_k = \bar{\mathbf{x}}[l_k]$ for any $l_k \in \arg\max_n\{\|\mathbf{P}_{\mathbf{Q}}^{\perp}\bar{\mathbf{x}}[n]\|_p\}$.

Step 3. Update $\mathbf{Q} := [\mathbf{Q}\ \bar{\boldsymbol{\alpha}}_k] \in \mathbb{R}^{N_{\max}\times k}$ and go to **Step 2** until $k = N_{\max} - 1$.

Output the pixel indices $\{l_1, \ldots, l_{N_{\max}}\}$

5.4 Simulations

In this section various Monte Carlo simulations are performed to analyze the performances of the proposed algorithms, namely TRI-P, GENE-CH and GENE-AH algorithms. In the first subsection, the effectiveness of the proposed TRI-P (for $p = 1, 2$, and ∞) algorithm is studied. Algorithms that are considered for comparison with TRI-P algorithm are VCA [12], iterative N-FINDR (I-N-FINDR) [67], successive N-FINDR (SC-N-FINDR), sequential N-FINDR (SQ-N-FINDR) [68], SGA [10], and alternating volume maximization (AVMAX) [14]. The algorithmic details for those EEAs under test are as follows: Affine set fitting [23] is employed for dimension reduction in I-N-FINDR, SC-N-FINDR, SQ-N-FINDR, and AVMAX, while VCA uses either singular value decomposition (SVD) or PCA based on the signal-to-noise ratio (SNR). To have fair complexity comparison with other methods, VCA is supplied

with the SNR value (instead of letting the VCA algorithm estimate the SNR value). The convergence tolerance for I-N-FINDR, SQ-N-FINDR, and AVMAX is set to 10^{-6}.

In the second subsection, simulations are performed to study the effectiveness of GENE-CH and GENE-AH algorithms in various scenarios. Algorithms that are considered for comparison are HySiMe [45], HFC, NWHFC [41], ATGP-NPD [71], and MINMAX-SVD [34]. The GENE algorithms, HFC, NWHFC, and ATGP-NPD are evaluated for the following false alarm probability: $10^{-3}, 10^{-4}$, 10^{-5} and 10^{-6}, and for GENE, NWHFC, and ATGP-NPD algorithms, the true noise covariance matrix is supplied for each simulated data.

In both subsections, for all the scenarios under consideration, 100 Monte Carlo runs are performed. The average root-mean-square (rms) spectral angle given in (3.23) between the true and the estimated endmember signatures is used as the performance index for evaluating the performances of the EEAs under test. Lower spectral angle ϕ corresponds to better performance of the EEA. For performance comparison of the algorithms under test for estimating the number of endmembers, the mean and standard deviation of the estimated number of endmembers are calculated.

In the simulations, the endmembers are chosen from the USGS library [53]. The endmembers considered in our simulations are from the following pool: Alunite, Andradite, Buddingtonite, Chalcedony, Desert varnish, Goethite, Halloysite, Kaolinite, Montmorillonite, Muscovite, Nontronite, Pyrope, Ammonium smectite, Calcite, Dicktite, Dumortierite, Hematite, Jarosite, Opal, and Paragonite, with $M = 224$. The abundance vectors $\mathbf{s}[n],\ n = 1, \ldots, L$ are generated by following the Dirichlet distribution [12] [23], which ensures that the assumptions (i.e., (A1) and (A2)) hold true for the simulated hyperspectral data. In addition to the number of endmembers N, and the number of pixels L, the other parameters that define a particular scenario are purity level ρ of the data set (defined in Section 3.3) and SNR (defined in (3.22)).

5.4.1 Evaluation of TRI-P algorithm

Here, the first 8 endmembers are considered ($N = 8$) from the above mentioned pool and additive white Gaussian noise is added to the noise-free data and the noisy observations are generated as per (2.1). The noise in each band is assumed to be independent and identically distributed Gaussian with zero mean and variance σ^2. The EEAs are tested for different purity levels and for different SNRs. The average rms spectral angles ϕ_{en} (defined in (3.23)) for the EEAs under test for SNRs ranging from 10 dB to 40 dB, and for no noise case (SNR $=\infty$), with $\rho = 0.8, 0.9, 1$, and $L = 1000$ pixels are shown in Table 5.3. Though the EEAs are designed for $\rho = 1$, we consider cases with different purity levels so as to study the performances of the algorithms when the pure pixel assumption is violated. The bold-faced numbers in Table 5.3 correspond to the minimum average rms spectral angle for a specific pair of (ρ, SNR), over all the algorithms under test. It can be observed from Table 5.3 that for $\rho = 1$, and 0.9, TRI-P ($p = 2$) wins in almost all situations, for $\rho = 0.8$ (highly mixed case), TRI-P ($p = 1$) performs well in many cases. The average computation time T (over all the scenarios under consideration) of each algorithm implemented in Matlab R2008a and running in a desktop computer equipped with Core i7 $-$ 930 CPU with speed 2.80 GHz, and 12 GB memory, is also shown Table 5.3. It can be observed from Table 5.3 that, the TRI-P algorithm, besides better performance, also offers the highest computational efficiency.

5.4.2 Evaluation of GENE-CH and GENE-AH algorithms

The GENE-CH and GENE-AH algorithms introduced in Section 5.2 are tested on the simulated hyperspectral data. Since, out of all the EEAs considered in Section 5.4.1, TRI-P (with $p = 2$) offered the best performance, the pure pixel indices required for GENE-CH and GENE-AH (see Table 5.1) are obtained from the TRI-P (with $p = 2$)

Table 5.3. Average ϕ_{en} (degrees) and average computation time T (secs) over the various EEAs for different purity levels (ρ) and SNRs (uniform Gaussian noise case), L =1000, N =8.

Methods	ρ	ϕ_{en} (degrees) SNR (dB)								T (secs)
		10	15	20	25	30	35	40	∞	
VCA	0.8	9.34	6.85	5.28	4.98	4.55	4.26	4.39	4.03	0.0413
	0.9	8.88	5.43	3.65	3.14	2.66	2.60	2.47	2.03	
	1	8.65	4.28	2.15	1.57	0.84	0.53	0.32	0.02	
I-N-FINDR	0.8	9.24	6.40	4.93	4.70	4.54	4.54	4.60	4.50	0.1212
	0.9	8.69	4.88	3.32	2.81	2.66	2.68	2.62	2.49	
	1	8.39	3.87	1.91	1.06	0.60	0.35	0.23	0.12	
SC-N-FINDR	0.8	9.57	7.14	5.94	5.44	5.27	5.04	4.94	4.80	0.0552
	0.9	9.11	5.84	3.76	3.14	2.98	2.86	2.79	2.79	
	1	8.55	4.32	2.07	1.23	0.71	0.38	0.31	0.12	
SQ-N-FINDR	0.8	9.25	6.37	4.88	4.70	4.50	4.51	4.63	4.51	0.1361
	0.9	8.98	4.94	3.32	2.81	2.66	2.68	2.62	2.49	
	1	8.23	3.97	1.91	1.06	0.60	0.35	0.23	0.12	
SGA	0.8	8.71	6.31	5.03	4.59	4.61	4.56	4.43	4.25	0.1886
	0.9	8.51	5.01	3.10	2.66	2.40	2.39	2.37	2.12	
	1	**7.96**	3.76	1.78	0.99	0.57	0.34	0.22	**0.01**	
AVMAX	0.8	9.34	6.52	5.13	4.77	4.60	4.62	4.55	4.17	0.0207
	0.9	8.91	5.26	3.34	2.82	2.67	2.68	2.66	2.36	
	1	8.56	4.00	1.90	1.06	0.60	0.35	0.23	**0.01**	
TRI-P ($p = 1$)	0.8	9.46	6.25	**4.59**	**4.22**	**4.05**	**3.94**	**4.01**	4.02	**0.0166**
	0.9	9.07	4.93	3.16	2.54	2.29	2.18	2.19	**1.97**	
	1	8.64	3.95	1.88	1.00	0.56	0.34	0.22	**0.01**	
TRI-P ($p = 2$)	0.8	9.02	**6.22**	4.63	4.35	4.30	4.25	4.19	**3.95**	0.0170
	0.9	8.59	**4.54**	**2.64**	**2.17**	**2.03**	**2.04**	**2.04**	**1.97**	
	1	8.10	**3.74**	**1.75**	**0.95**	**0.55**	**0.33**	**0.21**	**0.01**	
TRI-P ($p = \infty$)	0.8	8.48	6.28	4.98	4.72	4.63	4.62	4.57	4.30	0.0172
	0.9	**8.36**	4.87	3.11	2.64	2.41	2.39	2.37	2.13	
	1	**8.04**	3.92	1.80	0.99	0.56	**0.33**	0.22	**0.01**	

algorithm. There are totally four scenarios under consideration.

Scenario 1: The endmembers used are the same as used in the previous subsection ($N = 8,\ M = 224$) with 5000 pixels and $N_{\max} = 25$. As in Section 5.4.1, uniform Gaussian noise was added to produce noisy hyperspectral data for SNRs 15, 25, 35 and 45 dB. The mean and the standard deviation of the estimated number of endmembers over 100 independent runs, for the algorithms under test are shown in Table 5.4. From Table 5.4, it can be readily observed that for this scenario, the estimation accuracies

of the proposed algorithms (GENE-CH in particular) are the best for low SNR (15 dB), and for other SNRs, GENE-CH and GENE-AH with $P_{\text{FA}} = 10^{-6}$ perform well and so does HySiMe. It should be noted that for SNR $\geq$ 25 dB, as the P_{FA} decreases, the standard deviations reduce to zero. The performances of the other algorithms under test are below par, and that of HFC (for SNR $\geq$ 25 dB) and NWHFC (for all SNRs) are independent of SNR.

Scenario 2: We study the performances of the algorithms under test in the case when the data are corrupted by non-uniform Gaussian noise, while maintaining the other parameters for data generation used in the first scenario. The noise in each band is considered to be uncorrelated, but with different variances in each band. The noise variances σ_i^2 in each of the M spectral bands follow a Gaussian shape that is centered at the $(M/2)$th band and are given by (3.25). The τ value in the simulations is set to 36 and the purity level ρ is fixed to be 1. Values of mean±standard deviation of the number of endmembers estimated by the algorithms under test are also shown in Table 5.5. Here again, for low SNR (15 dB), the proposed GENE algorithms fares well. For other SNRs, GENE-AH with $P_{\text{FA}} = 10^{-6}$ and HySiMe yield the best performance. It is worthwhile to mention that contrary to the GENE algorithms, the performance of HySiMe algorithm is almost independent of the noise types (uniform or non-uniform Gaussian noise).

Scenario 3: The purity level ρ of the hyperspectral data is allowed to vary while maintaining $N = 8$. The data are corrupted by uniform Gaussian noise with SNR = 30 dB. For the case with $N = 8$, $M = 224$, $L = 5000$, and $N_{\max} = 25$, values of mean±standard deviation of the number of endmembers estimated by the algorithms under test are tabulated in Table 5.6. It can be readily seen from Table 5.6 that when purity level is smaller, GENE-CH overestimates the number of endmembers which is consistent with the discussions in Section 5.2.2 and the illustration in Figure 5.3. On the other hand, GENE-AH with $P_{\text{FA}} = 10^{-6}$ and HySiMe correctly estimates the

number of endmembers.

Scenario 4: Finally, the number of endmember N is allowed to vary as $8, 12, 16$ and 20, while maintaining $\rho = 1, M = 224$, and $L = 5000$. Here again, the data is corrupted by uniform Gaussian noise with SNR = 30 dB. Values of mean±standard deviation of the number of endmembers estimated by the algorithms under test, are also tabulated in Table 5.7. It can be observed from Table 5.7 that for higher number of endmember $N = 16, 20$ GENE-CH yields the best performance followed by GENE-AH. For $N = 8, 12$ both GENE-AH with $P_{\text{FA}} = 10^{-6}$ and HySiMe yield the best performance.

Thus, it can be concluded from the above simulation results that the GENE-CH algorithm is more suitable for data with pure pixels (i.e., for data with (A4) satisfied) and larger number of endmembers, while the GENE-AH algorithm is the better choice for a general hyperspectral data without any such prior information.

5.5 Real data experiments

In this section, the proposed GENE-CH and GENE-AH algorithms using TRI-P ($p = 2$), and some other algorithms for estimating the number of endmembers, are tested with AVIRIS real hyperspectral data obtained over the Cuprite Nevada site [52]. The details of the AVIRIS data used in this experiment is presented in Section 3.4. The algorithms under test are GENE-CH, GENE-AH, HySiMe, HFC, NWHFC, ATGP-NPD and MINMAX-SVD. The estimated number of endmembers obtained by the algorithms under test are given in Table 5.8. For the GENE algorithms $N_{\max}$ is set to 100 and for all the algorithms P_{FA} is set to 10^{-8}, wherever applicable, in order to get a reliable estimate. The noise covariance matrix for this real data is estimated by multiple regression analysis [45] and is supplied for GENE, NWHFC, and ATGP-NPD algorithms. As can be seen from Table 5.8, the estimated numbers

Table 5.4. Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable) and SNRs- Uniform Gaussian noise case. $N_{\max} = 25$, true $N = 8$, $L = 5000$, $M = 224$, and $\rho = 1$.

Methods	P_{FA}	Uniform Gaussian noise			
		SNR (dB)			
		15	25	35	45
GENE-CH (TRI-P, $p = 2$)	10^{-3}	8.26±1.07	8.27±0.56	8.16±0.36	8.16±0.41
	10^{-4}	7.56±0.67	8.01±0.09	8.02±0.17	8.05±0.21
	10^{-5}	7.21±0.51	8.00±0	8.00±0	8.01±0.09
	10^{-6}	6.99±0.41	8.00±0	8.00±0	8.00±0
GENE-AH (TRI-P, $p = 2$)	10^{-3}	7.78±0.75	8.07±0.29	8.06±0.23	8.03±0.19
	10^{-4}	7.27±0.67	8.00±0	8.02±0.14	8.00±0
	10^{-5}	6.90±0.65	8.00±0	8.00±0	8.00±0
	10^{-6}	6.52±0.70	8.00±0	8.00±0	8.00±0
HYSIME	–	6.10±0.30	8.00±0	8.00±0	8.00±0
HFC	10^{-3}	4.05±0.21	5.00±0	5.00±0	5.00±0
	10^{-4}	4.00±0	5.00±0	5.00±0	5.00±0
	10^{-5}	3.99±0.09	5.00±0	5.00±0	5.00±0
	10^{-6}	3.99±0.09	5.00±0	5.00±0	5.00±0
NW-HFC	10^{-3}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-4}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-5}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-6}	5.00±0	5.00±0	5.00±0	5.00±0
ATGP-NPD	10^{-3}	18.14±1.52	30.86±1.92	34.95±1.70	35.89±1.91
	10^{-4}	14.78±1.18	25.91±1.68	29.92±1.53	30.68±1.62
	10^{-5}	12.60±0.97	22.41±1.30	26.29±1.27	27.02±1.50
	10^{-6}	10.94±0.87	19.91±1.21	24.25±1.19	24.93±0.89
MINMAX-SVD	–	3.31±1.81	2.66±1.48	2.85±1.31	3.47±1.12

of endmembers are different for each algorithm. For the Cuprite data set, initially it was concluded that there are about 13 minerals (endmembers) in the site, and later it was increased to nearly 70 mineral compounds (endmembers) [54]. Hence, it is difficult to comment on the estimation accuracies of the algorithms under test. However, the good estimation accuracy of GENE-AH algorithm (as inferred from Table 5.4 to Table 5.7), makes 27 (see Table 5.8), a reasonable estimate for this data set. On the other hand, GENE-CH overestimates the number of endmembers. This may be attributed to the fact that in the absence of pure pixels, GENE-CH indeed

Table 5.5. Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable) and SNRs- Non-uniform Gaussian noise case. $N_{\max} = 25$, true $N = 8$, $L = 5000$, $M = 224$, and $\rho = 1$.

Methods	P_{FA}	Non-uniform Gaussian noise ($\tau = 36$) SNR (dB)			
		15	25	35	45
GENE-CH (TRI-P, $p = 2$)	10^{-3}	9.07±1.11	8.44±0.68	8.51±0.84	8.26±0.52
	10^{-4}	8.51±0.74	8.19±0.42	8.15±0.41	8.10±0.34
	10^{-5}	8.27±0.60	8.07±0.25	8.03±0.19	8.02±0.20
	10^{-6}	8.13±0.44	8.05±0.21	8.02±0.17	8.00±0
GENE-AH (TRI-P, $p = 2$)	10^{-3}	8.26±0.48	8.10±0.30	8.08±0.27	8.09±0.28
	10^{-4}	8.11±0.35	8.03±0.19	8.00±0	8.01±0.09
	10^{-5}	7.95±0.50	8.01±0.09	8.00±0	8.00±0
	10^{-6}	7.91±0.47	8.00±0	8.00±0	8.00±0
HYSIME	–	6.09±0.28	8.00±0	8.00±0	8.00±0
HFC	10^{-3}	3.00±0	5.00±0	5.00±0	5.00±0
	10^{-4}	3.00±0	5.00±0	5.00±0	5.00±0
	10^{-5}	3.00±0	5.00±0	5.00±0	5.00±0
	10^{-6}	2.99±0.09	5.00±0	5.00±0	5.00±0
NW-HFC	10^{-3}	6.04±0.19	6.14±0.34	6.13±0.37	6.13±0.33
	10^{-4}	6.03±0.17	6.04±0.19	6.04±0.19	6.04±0.19
	10^{-5}	6.01±0.10	6.03±0.17	6.03±0.17	6.03±0.17
	10^{-6}	5.99±0.09	6.03±0.17	6.03±0.17	6.03±0.17
ATGP-NPD	10^{-3}	24.60±1.82	33.30±1.83	35.41±1.65	35.75±1.75
	10^{-4}	20.40±1.39	28.29±1.49	30.71±1.69	30.93±1.74
	10^{-5}	17.52±1.15	24.90±1.25	26.77±1.28	26.97±1.43
	10^{-6}	15.41±1.08	22.54±1.22	24.59±1.15	24.81±0.98
MINMAX-SVD	–	3.65±1.68	3.10±1.70	3.15±1.62	3.30±1.11

overestimates the number of endmembers (see Table 5.6).

To show the applicability of TRI-P ($p = 2$) algorithm in real data, a quantitative measurement of the endmembers estimated by TRI-P ($p = 2$) algorithm for the considered real data with $N = 27$, namely the mean removed spectral angle ϕ between the estimated signature and the corresponding library signature, defined in (3.27) is considered. The value of ϕ for the various minerals identified by the TRI-P ($p = 2$), is given in Table 5.9, and the numbers in the parentheses correspond to the values of ϕ for repeatedly identified materials. The abundance maps corresponding to the

Table 5.6. Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable), for various purity levels. Uniform Gaussian noise case, true $N = 8$, SNR=30 dB, $N_{\max} = 25$, $L = 5000$, and $M = 224$.

Methods	P_{FA}	$N = 8$, SNR=30 dB			
		Purity Level ρ			
		0.8	0.85	0.9	0.95
GENE-CH (TRI-P, $p = 2$)	10^{-3}	14.06±3.68	12.79±2.80	9.98±1.44	8.80±0.93
	10^{-4}	12.82±2.94	10.97±2.38	9.28±1.28	8.33±0.58
	10^{-5}	12.13±2.70	10.23±1.92	8.85±0.98	8.17±0.45
	10^{-6}	11.65±2.69	9.85±1.71	8.61±0.92	8.11±0.38
GENE-AH (TRI-P, $p = 2$)	10^{-3}	8.10±0.31	8.05±0.21	8.14±0.34	8.09±0.28
	10^{-4}	8.02±0.14	8.01±0.09	8.00±0	8.01±0.09
	10^{-5}	8.00±0	8.01±0.09	8.00±0	8.00±0
	10^{-6}	8.00±0	8.00±0	8.00±0	8.00±0
HYSIME	–	8.00±0	8.00±0	8.00±0	8.00±0
HFC	10^{-3}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-4}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-5}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-6}	5.00±0	5.00±0	5.00±0	5.00±0
NW-HFC	10^{-3}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-4}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-5}	5.00±0	5.00±0	5.00±0	5.00±0
	10^{-6}	5.00±0	5.00±0	5.00±0	5.00±0
ATGP-NPD	10^{-3}	35.34±1.97	35.32±2.06	34.73±1.80	34.10±1.85
	10^{-4}	29.75±1.63	29.50±1.62	29.43±1.57	29.07±1.45
	10^{-5}	25.84±1.45	25.51±1.39	25.55±1.27	25.38±1.22
	10^{-6}	22.98±1.42	22.96±1.15	23.17±1.23	22.90±1.20
MINMAX-SVD	–	3.20±1.88	2.90±1.81	2.91±1.67	3.15±1.81

endmembers estimated by TRI-P ($p = 2$) are obtained by fully constrained least squares (FCLS) algorithm [15] and are shown in Figure 5.4. Note that the minerals in Table 5.9 and Figure 5.4 are arranged in alphabetical order and the minerals were identified by the visual comparison of the obtained abundance maps with the ones available in [20, 23, 51, 55, 56].

Table 5.7. Mean±standard deviation of the estimated number of endmembers for various algorithms over 100 independent runs, with different false alarm probabilities P_{FA} (whenever applicable), for various number of endmembers. Uniform Gaussian noise case, $\rho = 1$, SNR=30 dB, $N_{\max} = 25$, $L = 5000$, and $M = 224$.

Methods	P_{FA}	$\rho = 1$, SNR=30 dB			
		Number of Endmembers N			
		8	12	16	20
GENE-CH (TRI-P, $p = 2$)	10^{-3}	8.19±0.40	12.28±0.62	15.86±0.51	19.82±0.55
	10^{-4}	8.02±0.17	12.04±0.19	15.77±0.46	19.78±0.50
	10^{-5}	8.00±0	12.03±0.17	15.77±0.46	19.74±0.52
	10^{-6}	8.00±0	12.02±0.14	15.72±0.47	19.70±0.52
GENE-AH (TRI-P, $p = 2$)	10^{-3}	8.06±0.23	12.02±0.14	14.99±0.38	18.01±0.46
	10^{-4}	8.00±0	12.00±0	14.76±0.42	17.75±0.50
	10^{-5}	8.00±0	12.00±0	14.57±0.49	17.51±0.54
	10^{-6}	8.00±0	12.00±0	14.32±0.46	17.17±0.66
HYSIME	–	8.00±0	12.00±0	14.00±0	16.15±0.35
HFC	10^{-3}	5.00±0	7.81±0.44	8.35±0.49	5.01±0.61
	10^{-4}	5.00±0	7.14±0.68	8.66±0.27	4.19±0.63
	10^{-5}	5.00±0	6.44±0.53	7.93±0.25	3.67±0.60
	10^{-6}	5.00±0	6.10±0.46	7.76±0.47	3.23±0.52
NW-HFC	10^{-3}	5.00±0	7.79±0.53	9.39±0.54	7.01±0.74
	10^{-4}	5.00±0	7.18±0.70	9.15±0.35	6.23±0.69
	10^{-5}	5.00±0	6.46±0.62	8.97±0.30	5.46±0.77
	10^{-6}	5.00±0	5.96±0.58	8.80±0.42	4.78±0.70
ATGP-NPD	10^{-3}	33.88±1.73	44.22±2.09	49.30±2.40	55.59±2.65
	10^{-4}	28.68±1.56	37.28±1.55	41.69±2.15	47.16±1.80
	10^{-5}	25.25±1.25	32.83±1.60	36.33±2.05	40.92±1.89
	10^{-6}	22.93±1.14	28.64±1.61	31.90±1.74	36.01±1.76
MINMAX-SVD	–	2.73±1.64	3.57±2.01	4.33±2.01	3.87±1.98

5.6 Summary

In this work, we have considered the estimation of number of endmembers in hyperspectral images, which has been a challenging problem prevailing in the field of hyperspectral image analysis. To this end, we have presented two convex geometry based algorithms, namely GENE-CH and GENE-AH algorithms, which make use of the fact that the observed dimension-reduced observations lie in the convex hull and affine hull of the endmember signatures, respectively. The GENE algorithms em-

Table 5.8. Number of endmembers estimated by various algorithms, where NA denotes "non-applicable" and $\star$ denotes "out of memory" encountered in Matlab.

Algorithms	P_{FA}	Estimated N
GENE-CH	10^{-8}	76
GENE-AH	10^{-8}	27
HYSIME	NA	21
HFC	10^{-8}	11
NWHFC	10^{-8}	11
ATGP-NPD-Gaussian Noise	10^{-8}	61
ATGP-NPD-Laplacian Noise	10^{-8}	36
ATGP-NPD-Modified Gaussian Noise	10^{-8}	44
MINMAX-SVD	NA	$\star$

Table 5.9. Mean-removed spectral angles ϕ (degrees) between library spectra and endmembers estimated by TRI-P ($p = 2$).

Minerals	ϕ (degrees)
Alunite	17.91
Andradite	18.03 (19.16)
Buddingtonite	26.33
Chalcedony	19.85 (30.29)
Desert Varnish	11.92
Dumortierite	26.13 (31.43)
Kaolinite	24.34 (32.85)
Montmorillonite#1	18.79
Montmorillonite#2	20.34 (20.61)
Muscovite	34.01 (37.71)
Nontronite#1	23.89 (24.07) (25.38) (25.12)
Nontronite#2	15.35 (23.07) (26.02) (27.01)
Nontronite#3	22.76
Paragonite	36.73
Pyrope	12.20
Average ϕ	24.12

ploy a Neyman-Pearson hypothesis testing strategy to estimate the true number of endmembers. Any SEEA can be employed with the GENE algorithms, however, a reliable, reproducible and computationally efficient EEA will result in an improved estimation accuracy. Furthermore, we have also presented a simple, effective, and

computationally efficient EEA, namely TRI-P algorithm with its endmember identifiability proven. Simulation results confirm the superior efficacy of TRI-P ($p = 2$) and TRI-P ($p = 1$), for pure and mixed hyperspectral data, respectively. It is also shown via simulations that to estimate the number of endmembers, GENE-CH is preferred when pure pixels are present and the number of endmembers is large. In general, when no such information is available for the hyperspectral data, GENE-AH should be the advisable choice for estimating the number of endmembers. In addition, the real hyperspectral data experiment also exhibits the practical applicability of the proposed GENE-AH algorithm and TRI-P algorithm.

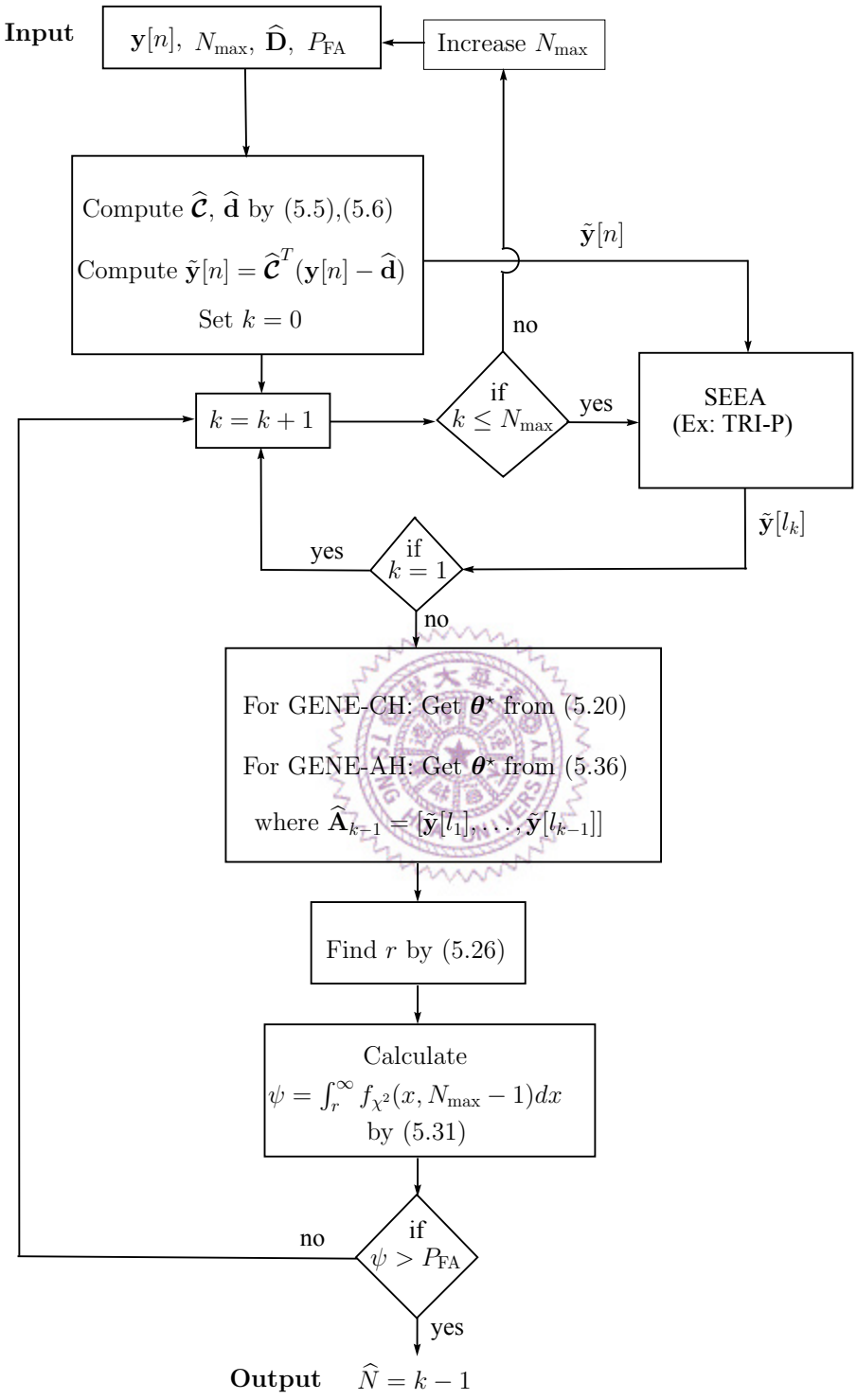


Figure 5.2. Working strategy of GENE-CH/GENE-AH algorithm in synchronization with a suitable SEEA.

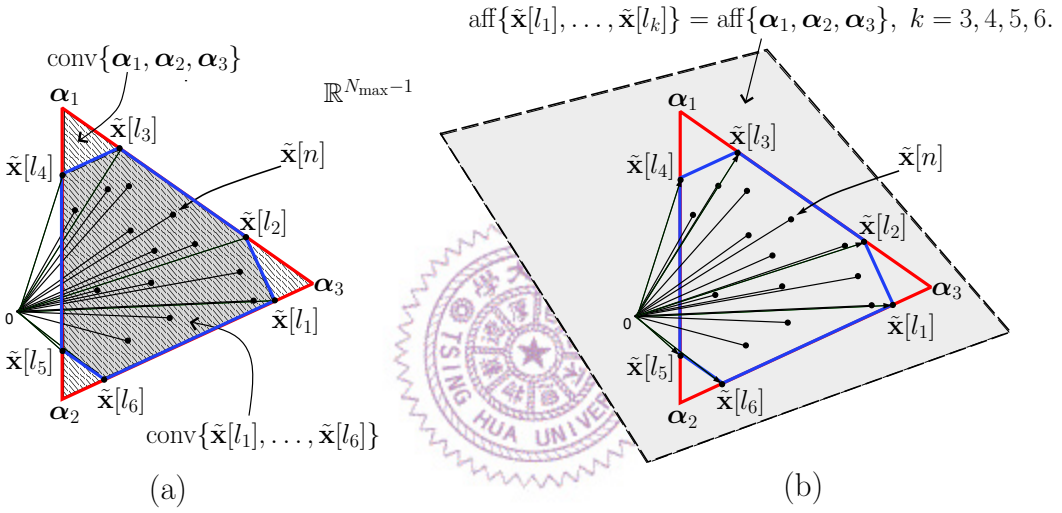


Figure 5.3. Illustration of GENE-CH algorithm, when no pure pixel is present in the noise-free hyperspectral data ($N = 3$ case). (a) The endmember estimates are denoted by $\tilde{\mathbf{x}}[l_i], i = 1, ..., N_{\max} = 6$, but $\text{conv}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_6]\} \neq \text{conv}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$ because the true endmembers $\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3$ are not present in the data cloud, whereas $\text{aff}\{\tilde{\mathbf{x}}[l_1], \ldots, \tilde{\mathbf{x}}[l_k]\} = \text{aff}\{\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3\}$, $k = 3, 4, 5, 6$, as shown in (b).

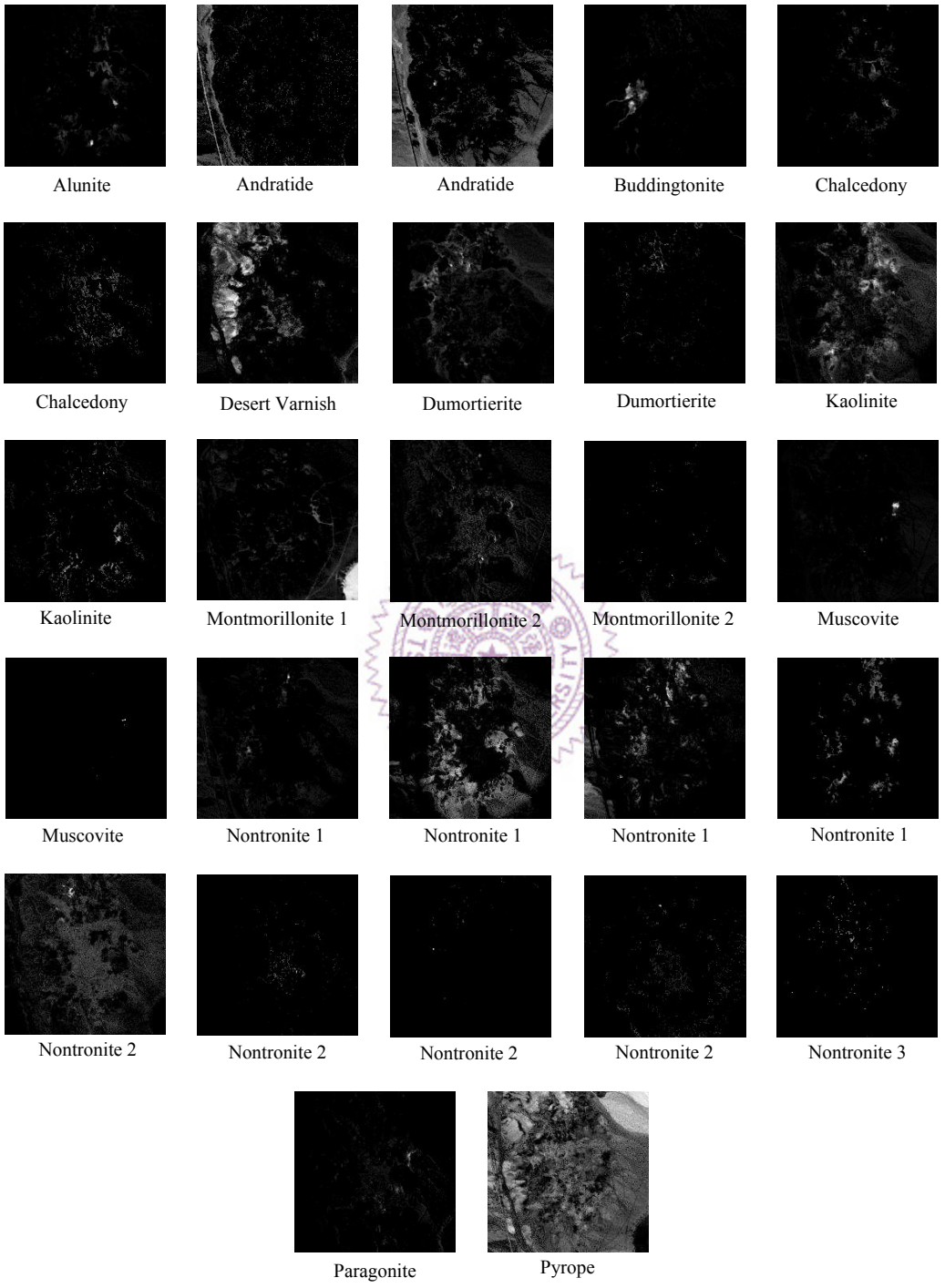


Figure 5.4. Abundance maps estimated based on endmembers obtained by TRI-P algorithm (with $p = 2$).

Chapter 6

Conclusions and Future Works

In this dissertation, for effectively mitigating the noise effect in the hyperspectral image data on the HU process, we have presented two robust HU algorithms, namely RAVMAX and RMVES algorithms for robustly unmixing noisy hyperspectral data with and without pure pixels, respectively. In fact, to the best of our knowledge, it is the first time in HU that the presence of noise in the data is systematically and quantitatively considered in the problem formulation of the HU algorithms. In both algorithms, the randomness caused by noise has been dealt by incorporating chance constraints in the unmixing problem formulation with a design parameter η. The chance constraints are used to control the constraint set which in turn controls the resultant estimated simplex volume. A detailed analysis on the role of η has been presented, and it was concluded that η must be less than 0.5, in RMVES algorithm and must be greater than 0.5 for RAVMAX algorithm. The value of η affects the convexity of the subproblems involved in the robust algorithms. As a consequence of it, the subproblems in RAVMAX are convex and hence solved by available convex optimization solvers, whereas those in RMVES are non-convex and SQP solvers are used to handle them. The efficacy of the robust algorithms are demonstrated through extensive simulations and real data experiments using the standard AVIRIS data.

One of the critical challenges in hyperspectral analysis, namely estimation of number of endmembers, has also been addressed in this dissertation. The data geometry based approach for estimating the number of endmembers resulted in two novel algorithms namely GENE-CH and GENE-AH algorithms. A SEEA, namely TRI-P algorithm has been presented as a suitable candidate to work in synchronization with the proposed GENE algorithms. Both simulations and real data experiments confirmed the superior efficacy of the proposed GENE and TRI-P algorithms.

The algorithms presented in this thesis leave some worthwhile future research directions as follows. In addition to the presence of noise, another challenging problem faced by the researchers in hyperspectral analysis is the presence of outliers in the data. An algorithm that is not only robust against noise, but also robust against outliers in the hyperspectral data will be an interesting, but challenging future direction. The design parameter η plays a very important in the performance of robust algorithms. The appropriate values for η ($(0.9, 1)$ for RAVMAX and 0.001 for RMVES) were obtained through empirical experience. However, other more appropriate choices for η that can still enhance the performance of the RMVES and RAVMAX algorithms, for different scenarios, are worth investigations. An robust algorithm that can account for the orientation of the estimated endmember simplex in addition to the volume shrinkage of the estimated simplex should further improve the performance of the robust algorithms.

Algorithms for estimation of number of endmembers, which are robust against the presence of outliers and against model mismatch errors (which may occur when multiple reflections are recorded by the hyperspectral sensor), will also be an interesting future research. It will also be interesting to study how the proposed algorithms perform in other potential applications, such as analytical chemistry, deconvolution of genomic signals, biomedical images (dynamic contrast enhanced magnetic resonance images, DCE-MRI), to name a few, where the linear mixing model is valid.

Bibliography

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

[2] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Elsevier Journal of Linear Algebra and Its Applications*, vol. 416, no. 1, pp. 29-47, July 2006.

[3] M. B. Lopes, J. C. Wolff, J. B. Dias, and M. Figueiredo, "NIR hyperspectral unmixing based on a minimum volume criterion for fast and accurate chemical characterisation of counterfeit tablets," *Analytical Chemistry*, vol. 82, no. 4, pp. 1462-1469, Jan. 2010.

[4] W. R. Johnson, D. W. Wilson, and W. Fink, "Snapshot hyperspectral imaging in opthomology," *Journal of Biomedical Optics*, published online on Feb. 2007.

[5] M. O. Smith, P. E. Johnson, and J. B. Adams, "Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis," *Journal Geophys. Res.*, vol. 90, no. 2, pp. C797-C804, Feb. 1985.

[6] A. A. Green, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 1, pp. 65-74, May 1988.

[7] I. K. Fodor, "A survey of dimension reduction techniques," *Techical Report,* Lawrence Livermore National Laboratory, May 2002.

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

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

[11] 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.

[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] T.-H. Chan, C.-Y. Chi, W.-K. Ma, and A. Ambikapathi "Hyperspectral unmixing from a convex analysis and optimization perspective," in *Proc. First IEEE WHISPERS*, Grenoble, France, Aug. 26-28, 2009.

[14] 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. Geosci. Remote Sens.- Special Issue on Spectral Unmixing of Remotely Sensed Data*, 2011.

[15] 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.

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

[17] 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.

[18] A. Zare and P. Gader, "PCE: Piecewise convex endmember detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 6, pp. 2620-2632, June 2010.

[19] 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.,* Pacific Grove, CA, Nov. 4-7, 2007, pp. 1164-1168.

[20] 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.

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

[22] 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.

[23] 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.

[24] 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.

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

[26] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," in *Proc. Second IEEE WHISPERS*, Reykjavik, Iceland, June 14-16, 2010.

[27] 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.

[28] J. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *Proc. First IEEE WHISPERS*, Grenoble, France, August 26-28, 2009.

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

[30] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones", *Optimization Methods and Software* vol. 11-12, pp. 625-653, 1999.

[31] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 1.21. http://cvxr.com/cvx, Oct. 2010.

[32] MATLAB Optimization Toolbox - Version 7.6 (R2008a), The MathWorks, Inc., Natick, MA, USA.

[33] P. Stoica, "Maximum likelihood parameter and rank estimation in reduced-rank multivariate linear regressions," *IEEE Trans. Signal Processing*, vol. 44, no. 12, pp. 3069-3078, Dec. 1996.

[34] C. Kuybeda, D. Malah, and M. Barzohar, " Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors," *IEEE Trans. Signal Processing*, vol. 55, no. 12, pp. 5579-5592, Dec. 2007.

[35] P. Stoica and Y. Selen, "Model-order selection: A review of information criterion rules," *IEEE Signal Process. Mag.*, vol. 21, no. 4, pp. 36-47, July 2004.

[36] K. P. Burnham, *Model selection and multimodel inference: A practical information theoretic approach*, Springer Publications, 2002.

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

[38] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465-471, Sep. 1978.

[39] G. Schwarz, "Estimating the dimension of a model," *The Annals of Statistics*, vol. 6, no. 2, pp. 461-464, Mar. 1978.

[40] M. W. Graham and D. J. Miller, "Unsupervised learning of parsimonious mixtures on large spaces with integrated feature and component selection," *IEEE Trans. Signal Processing*, vol. 54, no. 4, pp. 1289-1303, April 2006.

[41] 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.

[42] M. Wax and T. Kailath, "Detection of signals by information criteria," *IEEE Trans. Acoustics, Speech, and Signal Process.,* vol. ASSP-33, pp. 387-392, April 1985.

[43] P. R. Peres-Neto, D. A. Jackson, and K. M. Somers, "How many principal components? stopping rules for determining the number of non-trivial axes revisited," *Elsevier Journal of Comput. Stat. and Data Analysis*, vol. 49, pp. 974-997, July 2004.

[44] J. Harsanyi, W. Farrand, and C.-I Chang, "Determining the number and identity of spectral endmembers: An integrated approach using Neyman-Pearson eigenthresholding and iterative constrained RMS error minimization," in *Proc. 9th Thematic Conf. Geologic Remote Sensing*, Feb. 1993.

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

[46] O. Eches, N. Dobigeon, and J.-Y. Tourneret, "Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm," *IEEE J. Sel. Topics Signal Processing*, vol. 3, no. 3, pp. 582-591, June 2010.

[47] A. Ambikapathi, T.-H. Chan, W.-K.Ma, and C.-Y. Chi, "A robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," in *Proc. IEEE ICASSP*, Dallas, Texas, Mar. 14-19, 2010, pp. 1202-1205.

[48] A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "A robust alternating volume maximization algorithm for endmember extraction in hyperspectral images," in *Proc. IEEE WHISPERS*, Reykjavik, Iceland, June 14-16, 2010.

[49] 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.

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

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

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

[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," *U. S. Geol. Surv., Open File Rep.* pp. 93-592, 1993.

[54] R. N. Clarke, and G. A. Swayze, "Evolution in imaging spectroscopy analysis and sensor signal-to-noise: An examination of how far we have come," in *Proc. 6th Annual JPL Airborne Earth Science Workshop*, Mar. 4-8, 1996.

[55] T. H. Chan, "Convex analysis based non-negative blind source separation for biomedical and hyperspectral image analysis," Ph.D. dissertation, National Tsing Hua Univ., Taiwan, 2009.

[56] J. M. P. Nascimento, "Unsupervised hyperspectral unmxing," Ph.D. dissertation, Technical University of Lisbon, Lisbon, 2006.

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

[58] 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 Chapter 7, *Convex Optimization in Signal Processing and Communications*, Editors: D. P. Palomar and Y. C. Eldar, UK: Cambridge University Press, 2010.

[59] F.-Y. Wang, C.-Y. Chi, T.-H. Chan, and Y. Wang, "Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 875-888, May 2010.

[60] P. T. Boggs and J. W. Tolle, "Sequential quadratic programming," *Acta Numerica*, pp. 1-51, 1996.

[61] S. Jia and Y. Qian "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 161-173, Jan. 2009.

[62] 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. Geosci. Remote Sens. - Special Issue on Spectral Unmixing of Remotely Sensed Data*, 2011.

[63] L. C. Ludeman, *Random Processes Filtering, Estimation, and Detection*, Wiley-Interscience Publication, 2003.

[64] M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions*, 2nd ed., Wiley-Interscience Publication, 1993.

[65] G. Arfken and H. Weber, *Mathematical Methods for Physicists*, Harcourt Academic Press, 2000.

[66] C.-I. Chang, W. Xiong, W. Liu, M.-L. Chang, C.-C. Wu, and C.-C. Chen, "Linear spectral mixture analysis based approaches to estimation of virtual dimensionality in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 3960-3979, Nov. 2010.

[67] 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*, vol. 5425, Aug. 2004, pp. 31-41.

[68] C.-C. Wu, S. Chu, and C.-I. Chang, "Sequential N-FINDR algorithms," *Proc. of SPIE*, vol. 7086, Aug. 2008, pp. 70860C1-70860C12.

[69] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "Variants of N-FINDR algorithm for endmember extraction," *Proc. of SPIE*, vol. 7109, Oct. 2008, pp. 71090G1-71090G8

[70] H. Ren and C.-I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aer. Elec. Sys.*, Vol. 39, no. 4, pp. 1232-1249, Oct. 2003.

[71] C.-I. Chang, W. Xiong, H.-M. Chen, and J.-W. Chai, "Maximum orthogonal subspace projection appraoch to estimating the number of spectral signal sources in hyperspectral imagery," *IEEE J. Sel. Topics Signal Processing*, vol. 5, no. 3, pp. 504-520, June 2011.

Publication List of The Author

Journal papers:

1. 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.

2. 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.

Journal paper under review:

1. A. Ambikapathi, T.-H. Chan, C.-Y. Chi, and K. Keizer, "Hyperspectral data geometry based estimation of number of endmembers using p-norm based pure pixel identification algorithm," submitted to *IEEE Trans. Geoscience and Remote Sensing* (under the first round review).

International conference papers:

1. T.-H. Chan, W.-K. Ma, C.-Y. Chi, and A. Ambikapathi, "Hyperspectral unmixing from a convex analysis and optimization perspective," in *Proc. First IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Grenoble, France, August 26-28, 2009. (Invited paper)

2. 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 (ICASSP)*, Dallas, Texas, Mar. 14-19, 2010, pp. 1202-1205.

3. A. Ambikapathi, T.-H. Chan, W.-K. Ma, and C.-Y. Chi, "A robust alternating volume maximization algorithm for endmember extraction in hyperspectral images," in *Proc. Second IEEE WHISPERS*, Reykjavik, Iceland, June 14-16, 2010.

4. A. Ambikapathi, T.-H. Chan, C.-Y. Chi, and K. Keizer, "Two effective and computationally efficient pure-pixel based algorithms for hyperspectral endmember extraction," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 22-27, 2011, pp. 1369-1372.

5. T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "Robust endmember extraction using worst-case simplex volume maximization," in *Proc. Third IEEE WHISPERS*, Lisbon, Portugal, June 6-9, 2011. (This paper was awarded the **Best Paper Award**)

6. T.-H. Chan, W.-K. Ma, A. Ambikapathi, and C.-Y. Chi, "An optimization perspective on Winter's endmember extraction belief," to appear in *Proc. IEEE International Geoscience and Remote Sensing Symposium (IGARSS'11)*, Vancouver, Canada, July 24-29, 2011. (Invited paper)

7. T.-H. Chan, C.-J. Song, A. Ambikapathi, C.-Y. Chi, and W.-K. Ma, "Fast alternating volume maximization algorithm for blind separation of non-negative sources," to appear in *Proc. IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Beijing, China, Sept. 18-21, 2011.