Spectral variability in hyperspectral unmixing : Multiscale, tensor, and neural network-based approaches

Ricardo Augusto Borsoi

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES FONDAMENTALES ET APPLIQUÉES

UFSC

Variabilité Spectrale en Démélange de données Hyperspectrales: Stratégies Multi-échelles, Tensorielles et Basées sur des Réseaux Neuronaux

Ricardo Augusto Borsoi

Laboratoire J.-L. Lagrange, Observatoire de la Côte d'Azur

Présentée en vue de l'obtention du grade de docteur science pour l'ingénieur
de l'Université Côte d'Azur
en cadre de cotutelle avec l'Université Federale de Santa Catarina
Dirigée par : Cédric Richard / José Carlos Moreira Bermudez
Soutenue le : 18 / 03 / 2021

Devant le jury, composé de :
Florence Tupin, PU, Télécom Paris
Frédéric Pascal, PU, CentraleSupélec
André Ferrari, PU, Université Côte d'Azur
Márcio Costa, PU, UFSC
Aldo von Wangenheim, Prof. Dr. rer.nat., UFSC
Eduardo Camponogara, PU, UFSC
Nikolaos Sidiropoulos, PU, University of Virginia

Spectral Variability in Hyperspectral Unmixing: Multiscale, Tensor, and Neural Network-based Approaches

Jury :

Rapporteurs
Florence Tupin, PU, Télécom Paris, France
Frédéric Pascal, PU, CentraleSupélec, France

Examinateurs
André Ferrari, PU, Université Côte d'Azur, France
Márcio Costa, PU, UFSC, Brazil
Aldo von Wangenheim, Prof. Dr. rer.nat., UFSC, Brazil
Eduardo Camponogara, PU, UFSC, Brazil
Nikolaos Sidiropoulos, PU, University of Virginia, USA

Encadrants
Cédric Richard, PU, Université Côte d'Azur, France
José Carlos Moreira Bermudez, PU, UFSC, Brazil

Spectral Variability in Hyperspectral Unmixing: Multiscale, Tensor, and Neural Network-based Approaches

ABSTRACT

The spectral signatures of the materials contained in hyperspectral images, also called endmembers (EMs), can be significantly affected by variations in atmospheric, illumination or environmental conditions typically occurring within an image. Traditional spectral unmixing (SU) algorithms neglect the spectral variability of the endmembers, what propagates significant mismodeling errors throughout the whole unmixing process and compromises the quality of the estimated abundances. Therefore, significant effort have been recently dedicated to mitigate the effects of spectral variability in SU. However, many challenges still remain in how to best explore *a priori* information about the problem in order to improve the quality, the robustness and the efficiency of SU algorithms that account for spectral variability. In this thesis, new strategies are developed to address spectral variability in SU. First, an (over)-segmentation-based multiscale regularization strategy is proposed to explore spatial information about the abundance maps more effectively. New algorithms are then proposed for both semi-supervised and blind SU, leading to improved abundance reconstruction performance at a small computational complexity. Afterwards, three new models are proposed to represent spectral variability of the EMs in SU, using parametric, tensor, and neural network-based representations for EM spectra at each image pixel. The parametric model introduces pixel-dependent scaling factors over a reference EM matrix to model arbitrary spectral variability, while the tensor-based representation allows one to exploit the high-dimensional nature of the data by means of its underlying low-rank structure. Generative neural networks (such as variational autoencoders or generative adversarial networks) finally allow one to model the low-dimensional manifold of the spectral signatures of the materials more effectively. The proposed models are used to devise three new blind SU algorithms, and to perform data augmentation in library-based SU. Finally, we provide a brief overview of work which extends the proposed strategies to new problems in SU and in hyperspectral image analysis. This includes the use of the multiscale abundance regularization in nonlinear SU, modeling spectral variability and accounting for sudden changes when performing SU and change detection of multitemporal hyperspectral images, and also accounting for spectral variability and changes in the multimodal (i.e., hyperspectral and multispectral) image fusion problem.

Keywords: Hyperspectral images. Endmember variability. Spectral unmixing. Multiscale. Tensors. Neural networks.

Variabilité Spectrale en Démélange de données Hyperspectrales: Stratégies Multi-échelles, Tensorielles et Basées sur des Réseaux Neuronaux

RÉSUMÉ

Les signatures spectrales des composants constitutifs présents dans les images hyperspectrales peuvent être significativement affectées par les variations des conditions atmosphériques, d'illumination ou d'environnement se produisant typiquement dans une image. Les algorithmes traditionnels de démélange spectral (Spectral Unmixing – SU) négligent la variabilité spectrale des composants constitutifs, ce qui propage des erreurs importantes tout au long du processus de démélange et compromet la qualité des abondances estimées. Par conséquent, des efforts importants ont été récemment consacrés à atténuer les effets de la variabilité spectrale dans les procédures de démélange. Cependant, de nombreux défis restent à relever pour savoir comment exploiter au mieux les informations *a priori* sur le problème afin d'améliorer à la fois la qualité et la robustesse des algorithmes de SU qui tiennent compte de la variabilité spectrale des composants. Dans cette thèse, de nouvelles stratégies sont développées pour aborder cette variabilité spectrale. Premièrement, une stratégie de régularisation multi-échelles basée sur la (sur)-segmentation des images est proposée pour explorer plus efficacement les informations spatiales sur les abondances. De nouveaux algorithmes sont ensuite proposés pour le démélange spéctral semi-supervisé et non-supervisé, ce qui se traduit par une amélioration des performances de reconstruction des abondances avec une complexité de calcul réduite. Ensuite, trois nouveaux modèles sont proposés pour représenter la variabilité spectrale des composants constitutifs, en utilisant des représentations paramétriques, tensorielles et basées sur des réseaux neuronaux pour les spectres de ces composants en chaque pixel de l'image. Le modèle paramétrique introduit des facteurs multiplicatifs dépendant des pixels dans une matrice des composants de référence pour modéliser une variabilité spectrale arbitraire, tandis que la représentation basée sur un tenseur permet d'exploiter la grande dimension des données en exploitant sa structure de rang faible sous-jacente. Les réseaux de neurones génératifs (tels que les *variational autoencoders* ou les *generative adversarial networks*) permettent enfin de modéliser la variété de faible dimension des signatures spectrales des matériaux, directement à partir des données observées. Les modèles proposés sont utilisés dans la conception de quatre nouveaux algorithmes de démélange non-supervisés et semi-supervisés. Enfin, nous donnons un bref aperçu des travaux qui étendent les stratégies proposées dans la thèse à de nouveaux problèmes en démélange et en dans l'analyse d'images hyperspectrales. Cela comprend l'utilisation de la régularisation d'abondance multi-échelles en démélange spectral non-linéaire, la modélisation de la variabilité spectrale, la prise en compte des changements soudains lors du démélange et la détection des changements dans les images hyperspectrales multitemporelles, ainsi que la prise en compte de la variabilité spectrale et des changements dans le problème de fusion d'images hyperspectrales et multispectrales.

Mots clés: Images hyperspectrales. Variabilité des composants constitutifs. Démélange spectral. Multi-échelle. Tenseurs. Réseaux de neurones.

Variabilidade dos Endmembers na Separação de dados Hiperespectrais: Estratégias Multi-escala, Tensoriais e Baseadas em Redes Neurais

RESUMO ESTENDIDO

As assinaturas espectrais dos materiais contidos em imagens hiperespectrais, também chamadas de *endmembers* (EMs), podem ser consideravelmente afetadas por variações nas condições atmosféricas, de iluminação, ou ambientais que podem ocorrer dentro de uma imagem. Métodos tradicionais de separação espectral (SU – *spectral unmixing*) não consideram a variabilidade espectral dos endmembers, o que acaba propagando erros por todo o processo de SU e compromete a qualidade das abundâncias estimadas. Consequentemente, muitos trabalhos têm se dedicado à mitigar os efeitos adversos da variabilidade dos endmembers na separação espectral. Não obstante, ainda existem muitos desafios relacionados à exploração de informação *a priori* referente à este problema para melhorar a qualidade, a robustez e a eficiência de algoritmos de SU que consideram a variabilidade dos EMs. Nesta tese, novas estratégias são desenvolvidas para mitigar a variabilidade espectral na separação espectral.

Primeiramente uma estratégia de regularização espacial multiescala baseada em algoritmos de segmentação e sobre-segmentação de imagens é proposta no Capítulo 3 para separação espectral semi-supervisionada esparsa. O problema de unmixing é dividido em dois problemas distintos, em escalas espaciais diferentes, os quais são resolvidos em sequência. A primeira das escalas espaciais contém apenas uma versão aproximada da imagem hiperespectral, enquanto a segunda escala representa a imagem em seu domínio original. Realizando a separação espectral primeiramente no domínio aproximado, obtém-se uma versão aproximada dos mapas de abundância que caracterizam a relação espacial entre os pixels vizinhos. Essa informação pode ser então utilizada na forma de uma regularização eficiente em um segundo problema de separação espectral no domínio original. Ao contrário de outras técnicas de regularização espacial costumeiramente empregadas na literatura (tomo o *Total Variation*, por exemplo), a regularização multiescala proposta conduz a problemas de otimização que não introduzem uma dependência explícita entre as abundâncias em diferentes pixels da imagem, o que permite uma solução muito mais eficiente.

A estratégia de regularização multiescala do Capítulo 3 é então estendida no Capítulo 4 para considerar o problema de SU não-supervisionado, usando um modelo paramétrico para representar variabilidade espectral dos endmembers ao longo da imagem. Ao invés de dividir o problema de separação espectral em dois problemas distintos (em domínios espaciais diferentes), propõe-se a solução de um problema único. Em particular, uma decomposição multiescala diferente é utilizada para representar a imagem em dois domínios espaciais, um sendo a escala de aproximação mencionada anteriormente, e outro contendo apenas os detalhes da imagem hiperespectral e dos mapas de abundância. A regularidade espacial dos mapas de abundância é então favorecida controlando-se sua energia na escala de aproximação e na escala de detalhes individualmente no processo de separação espectral, o qual é formulado como um problema de otimização não-convexo. Utilizando uma abordagem de *Least Squares* alternados e algumas hipóteses simplificadoras, uma solução computacionalmente eficiente é apresentada, levando a uma melhor estimação dos mapas de abundância com um custo computacional inferior.

Posteriormente, nos Capítulos 5–8, novos modelos são propostos para representar as assinaturas espectrais dos endmembers em cada pixel da imagem no problema de separação espectral, utilizando representações paramétricas, tensoriais, e baseadas em redes neurais. Primeiro, um modelo paramétrico é proposto no Capítulo 5 para representar as assinaturas espectrais dos endmembers como uma versão escalada (para cada pixel, material e comprimento de onda) de um conjunto de assinaturas espectrais de referência. O modelo proposto generaliza o modelo de misturas linear, permitindo variações arbitrárias no espectro e conectando a quantidade de variabilidade à amplitude do espectro de referência. As abundâncias e os fatores de escalamento para cada pixel são estimados através da solução de um problema de otimização, o qual incorpora regularizações que promovem a suavidade espacial dessas variáveis para introduzir informação *a priori* ao problema. Simulações mostram que o modelo proposto pode melhorar o desempenho de separação espectral.

Já no Capítulo 6, a propriedade multidimensional dos mapas de abundância e das assinaturas espectrais dos endmembers para cada pixel da imagem é explorada no desenvolvimento de um método de separação espectral que não necessita de um modelo explícito para a assinatura espectral dos endmembers. Para tal, é explorada a representação natural dos dados na forma de tensores, bem como suas estruturas de rank baixo subjacentes, as quais são frequentemente observadas na prática. O problema de separação espectral é formulado como uma decomposição de tensores aproximada, onde os mapas de abundância e endmembers são restringidos à vizinhança de tensores de rank baixo. Uma solução aproximada e eficiente para este problema é proposta utilizando o algoritmo dos *Least Squares* alternados. Uma estratégia simples para a seleção dos ranks dos tensores também foi desenvolvida, reduzindo assim a quantidade de parâmetros a serem ajustados pelo usuário. Resultados indicam que a estratégia proposta conduz a bons resultados de separação espectral com imagens tanto sintéticas quanto reais, com um custo computacional competitivo para imagens de dimensões pequenas.

A despeito de um bom desempenho prático, estratégias baseadas em modelos paramétricos e tensoriais não exploram uma propriedade muito importante da variabilidade espectral dos endmembers: as assinaturas espectrais são geralmente confinadas a um manifold de dimensão baixa. O Capítulo 7 propõe explorar esse fato utilizando uma representação baseada em redes neurais generativas (como autoencoders variacionais e redes generativas adversárias), as quais representam o manifold onde reside o espectro dos endmembers como a imagem de uma função sobre um espaço Euclideano de baixa dimensão. Baseado nessa representação, uma abordagem de separação espectral é proposta. Primeiramente, pixels ditos *puros*, os quais representam instâncias da assinatura espectral de cada endmember, são extraídos da imagem e utilizados para treinar as redes neurais generativas, uma para cada material na cena. Posteriormente, os mapas de abundância e as representações latentes dos endmembers para cada pixel são obtidas através da solução de um problema de otimização inspirado por métodos de fatoração de matrizes. Resultados experimentais mostraram um excelente desempenho do método para cenas contendo pixels puros. O Capitulo 8 estende essa abordagem para o problema de separação espectral semi-supervisionado com bibliotecas espectrais contendo uma quantidade limitada de assinaturas espectrais. Neste caso, os modelos generativos para cada endmember são aprendidos diretamente a partir da biblioteca espectral disponível, e são posteriormente utilizados para gerar novas assinaturas sintéticas usando uma estratégia de *data augmentation*. Resultados experimentais mostraram uma performance comparável com aquela obtida utilizando assinaturas geradas a partir de modelos físicos adequados à cena, mas sem a necessidade do conhecimento *a priori* de tais processos.

Por fim, o Capítulo 9 apresenta um breve olhar sobre outras contribuições realizadas durante a duração desta tese que estendem os métodos propostos nos Capítulos 3–8 para resolver outros desafios relacionados à separação espectral e, de maneira mais geral, à análise de imagens hiperespectrais. Primeiramente, o método de regularização espacial multiescalas proposto no Capítulo 3 é estendido para resolver o problema de separação espectral não-linear usando kernels. Além de estender a formulação anterior para um modelo de mistura não-paramétrico, uma metodologia também é proposta para determinar os parâmetros do algoritmo automaticamente. Além disso, uma solução eficiente para os problemas de otimização resultantes é desenvolvida fazendo-se uso de sua dualidade forte e do método da bisseção.

Posteriormente, o modelo paramétrico desenvolvido no Capítulo 5 para representar os endmembers dentro de uma mesma imagem é estendido para representar a variabilidade espectral encontrada em imagens adquiridas em diferentes instantes de tempo. Utilizando uma estratégia de filtragem Bayesiana e o algoritmo *Expectation Maximization*, um método de separação espectral é proposto para sequencias de imagens sem variações abruptas nos mapas de abundâncias em instantes de tempo adjacentes. Para mitigar esta última limitação, um método baseado em bibliotecas espectrais é proposto para realizar unmixing e detecção de mudanças em sequências de imagens hiperespectrais baseado em um problema de otimização combinatório. De modo a reduzir a complexidade computacional do método, uma solução aproximada é proposta, e garantias teóricas de desempenho são derivadas.

Finalmente, a variabilidade espectral e espacial é abordada dentro do problema de fusão de imagens hiperespectrais e multiespectrais, a qual visa obter imagens de alta resolução espacial e espectral. Considerando que as assinaturas espectrais subjacentes a cada uma das imagens podem ser diferentes, primeiramente estendemos o modelo paramétrico do Capítulo 5 para representar a variabilidade entre as imagens de diferentes modalidades. Isso levou a uma melhora de desempenho considerável, especialmente em pares de imagens onde a variabilidade espectral espacialmente invariante. Para abordar o problema onde a variabilidade pode ocorrer tanto espectral quanto espacialmente, o problema de fusão de imagens é formulado como a fatoração de tensores acoplada, sujeito a presença de variabilidade aditiva arbitrária entre as imagens. Assumindo que tanto as imagens quanto o termo representando a variabilidade podem ser representados como tensores de rank baixo, dois novos algoritmos são propostos (um puramente algébrico e outro baseado em uma estratégia de otimização), levando a uma melhora no desempenho prático. Resultados teóricos garantem a recuperação exata da imagem de interesse quando certas condições sobre os ranks são atendidas.

Palavras-chave: Imagens hiperespectrais. Variabilidade dos endmembers. Separação espectral. Multiescala. Tensores. Redes neurais.

ACKNOWLEDGEMENTS

I am very thankful to the many people who made these past years very enriching, both academically as well as personally. First, I would like to thank my supervisors, José Carlos Moreira Bermudez and Cédric Richard, for the exceptional guidance and support they gave me throughout this whole period. I was fortunate to learn from their expertise and knowledge and from their detailed feedback. I would also like to thank Profs. Florence Tupin, Frédéric Pascal, André Ferrari, Márcio Costa, Aldo von Wangenheim, Eduardo Camponogara and Nikolaos Sidiropoulos for accepting to by part of my thesis committee, specially Profs. Florence and Frédéric, who took the time to review my thesis manuscript.

I am also grateful to all the people I had the opportunity to work with during these years. For Tales Imbiriba, my "academic brother", goes my deepest gratitude. His willingness to openly explain his work and ideas were immensely helpful, and our countless discussions over these years (accompanied by copious amounts of tea, coffee and biscuit) are an integral part of much of the work in this thesis.

Thanks to Alina Zare, Jean-Yves Tourneret, Jocelyn Chanussot, Lucas Drumetz, and Christian Jutten for joining in on the review paper on spectral variability, for Clémence Prevost, David Brie and Konstantin Usevich, for our fruitful collaboration on tensors and image fusion, and for Pau Closas, André Ferrari and Jie Chen, who also helped to bring the time dimension into play. Thanks also to Mircea Moscu, for all the fun we had investigating the connectivity of graph signals. Finally, thanks to Diego do Carmo and Márcio Costa, from whom I learned a lot about hearing aids and speech signal processing.

A big thanks goes also to all my former labmates in Brazil and in France, especially Alohotsy, Daniel, Diego, Fabio, Filipe, Gustavo, Ikram, Jhonny, Manfred, Mircea, Paulo, Rafael, Renata, Victor, Wemerson, and Wilmer. A very special thanks to Guilherme Costa, for the friendship and for all the wise advice he is always ready to share. Thanks also to Ezequiel, with whom I was lucky to share many moments of this journey.

I am also very grateful to my supervisors and to the people at UFSC (Fábio de Queiroz, Bartolomeu Uchôa-Filho) and at UCA (Elisabeth Taffin de Givenchy) for helping me with all the bureaucratic process involved in enrolling for the dual doctoral degree. Another big thanks goes to Wilson Costa, Marcelo Siqueira and Lis Cunha, who do a wonderful job in taking care of all the bureaucracy at UFSC so efficiently.

I also gratefully acknowledge those who supported my research so far and made this thesis possible. I thank the National Council for Scientific and Technological Development (CNPq), who sponsored this thesis, and my stay in France. Thanks also to the GdR ISIS and to Prof. Audrey Giremus for fostering the collaboration with Clémence, David and Konstantin at the Université de Lorraine through a *bourse de mobilité*.

Last but foremost, I want to thank my parents, Francisco and Rosa Borsoi, for their continued love and support in each moment of this journey.

"Son of man,
You cannot say, or guess, for you know only
A heap of broken images"
(T.S. Eliot, *The Waste Land*, 1922)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ADMM	Alternating Direction Method of Multipliers
AEC	Autoencoder
ALS	Alternating Least Squares
ATREM	Atmospheric Removal
AVIRIS	Airborne Visible/Infrared Imaging Spectrometer
BCCB	Block Circulant matrix with Circulant Blocks
BPT	Binary Partition Tree
CPD	Canonical Polyadic Decomposition
DeepGUn	Deep Generative Unmixing algorithm
DGM	Deep Generative Model
ELM	Empirical Line Method
ELMM	The Extended Linear Mixing Model
EM	Endmember
FCLS	Fully Constrained Least Squares
FLAASH	Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes
GAN	Generative Adversarial Network
GLMM	Generalized Linear Mixing Model
HI	Hyperspectral Image
HR	High Resolution
HS	Hyperspectral
IR	Infrared
LMM	Linear Mixing Model
LS-SVR	Least-Squares Support Vector Regression
MESMA	Multiple Endmember Spectral Mixture Analysis
MI	Multispectral Image
MSE	Mean Squared Error
MTSU	Multitemporal Spectral Unmixing
MUA	Multiscale sparse Unmixing Algorithm
MUA-SV	Multiscale Unmixing Algorithm Accounting for Spectral Variability
NDVI	Normalized Difference Vegetation Index
NN	Neural Network
NRMSE	Normalized Root Mean Squared Error
NTF	Nonnegative Tensor Factorization
PLMM	Perturbed Linear Mixing Model
RMSE	Root Mean Squared Error
RTF	Radiative Transfer Function
SAM	Spectral Angle Mapper
SCLS	Scaled Constrained Least Squares

SLIC	Simple Linear Iterative Clustering
SNR	Signal-to-Noise Ratio
SRE	Signal to Reconstruction Error
SU	Spectral Unmixing
SVM	Support Vector Machine
TV	Total Variation
ULTRA	Unmixing with Low-rank Tensor Regularization Algorithm
ULTRA-V	Unmixing with Low-rank Tensor Regularization Algorithm accounting for endmember Variability
USGS	United States Geological Survey
VAE	Variational Autoencoder
VCA	Vertex Component Analysis
w.r.t.	with respect to

LIST OF SYMBOLS

P	Number of materials (endmembers) in a hyperspectral image
L	Number of spectral bands in a hyperspectral image
N	Number of pixels in a hyperspectral image
N_1, N_2	Number of rows and columns of a hyperspectral image
S	Dimension of the coarse representation of the hyperspectral image and abundance maps
$\widetilde{P}$	Number of spectral signatures contained in a spectral library
R	Order of a tensor
$K_\bullet$	Canonical poliadic rank of a tensor
F_i	Mode-i rank of a tensor
H	Dimensionality of the latent space of a generative endmember model
U_p	Number of pure pixels of the p-th endmember extracted from the image
C_p	Number of spectral signatures of the p-th endmember in the spectral library
Q_s	Number of spectral signatures sampled from the generative model for library augmentation
T	Number of images in a multitemporal hyperspectral image sequence
y_{sensor}	Reflectance of a given pixel observed at the sensor
y_{atm}	Reflectance of the atmosphere
y_{s}	Surface reflectance
y_{avg}	Average surface reflectance in a region around a pixel
T_{g}	Gaseous transmittance
$T_{\uparrow}$	Upward scattering transmittance
$T_{\downarrow}$	Downward scattering transmittance
r	Ratio between diffuse and total transmittance for the ground-to-sensor path
s	Spherical albedo of the atmosphere
$\boldsymbol{y}_n$	$L \times 1$ hyperspectral image pixel indexed by n.
$\boldsymbol{a}_n$	$P \times 1$ vector containing the abundances of every endmember in the n-th pixel
$\boldsymbol{e}_n$	$L \times 1$ vector representing additive noise and modeling errors in the n-th pixel
$\boldsymbol{Y}$	Hyperspectral image (ordered as an $L \times N$ matrix)
$\boldsymbol{A}$	$P \times N$ matrix containing the abundances corresponding to all pixels as its columns
$\boldsymbol{E}$	$L \times N$ matrix containing the additive noise and modeling errors for all image pixels as its columns
$\boldsymbol{M}_0$	$L \times P$ matrix containing the reference spectral signatures of the P endmembers
$\boldsymbol{m}_{0,p}$	$L \times 1$ vector containing the spectrum of the p-th reference endmember
$\boldsymbol{M}_n$	$L \times P$ matrix containing the spectral signatures of the endmembers in the n-th pixel
$\mathscr{M}$	$L \times P \times N$ tensor containing the endmember matrices for all pixels

$\mathscr{Y}$	$L \times N_1 \times N_2$ tensor containing the hyperspectral image pixels
$\mathscr{A}$	$P \times N_1 \times N_2$ tensor containing the abundance vectors for all pixels
$\mathscr{E}$	$L \times N_1 \times N_2$ tensor containing the observation noise
$\mathscr{P}$	Low-rank approximation of tensor $\mathscr{M}$
$\mathscr{Q}$	Low-rank approximation of tensor $\mathscr{A}$
$\boldsymbol{M}_{\text{Lib}}$	$L \times \widetilde{P}$ spectral library containing endmember spectra known *a priori*
$\boldsymbol{A}^{\text{Lib}}$	$\widetilde{P} \times N$ matrix containing the fractional abundances corresponding to each element of $\boldsymbol{M}_{\text{Lib}}$
$\mathcal{M}$	Set containing all the $L \times P$ endmember matrices that can be constructed using specta in a library known *a priori*
$\mathcal{M}_p$	Set containing spectral signatures of the p-th endmember known *a priori*
$\boldsymbol{W}$	$N \times S$ matrix mapping from the original spatial domain to the coarse approximation domain
$\boldsymbol{W}^*$	$S \times N$ matrix mapping from the coarse approximation domain to the original spatial domain
$\widetilde{\boldsymbol{W}}$	$N \times N + S$ matrix mapping from the original spatial domain to a coarse representation and its complement
$\widetilde{\boldsymbol{W}}^{\dagger}$	Right inverse of $\mathcal{W}$
$W_{i,j}$	Element in the i-th row and j-th column of $\boldsymbol{W}$
$\Omega_s(n)$	Number of pixels contained in the n-th superpixel
$RE_{\mathcal{C}}, RE_{\mathcal{D}}$	Reconstruction errors at the coarse and detail spatial scales
$\boldsymbol{Y}_{\mathcal{C}}$	$L \times S$ representation of the hyperspectral image $\boldsymbol{Y}$ in the coarse spatial domain
$\boldsymbol{y}_{\mathcal{C}_i}$	i-th column of $\boldsymbol{Y}_{\mathcal{C}}$
$\boldsymbol{A}_{\mathcal{C}}$	$P \times S$ representation of the abundances $\boldsymbol{A}$ in the coarse spatial domain
$\boldsymbol{a}_{\mathcal{C}_i}$	i-th column of $\boldsymbol{A}_{\mathcal{C}}$
$\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}$	$\widetilde{P} \times S$ representation of $\boldsymbol{A}^{\text{Lib}}$ in the coarse spatial domain
$\widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}}$	$\widetilde{P} \times S$ matrix containing the estimate of $\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}$
$\widehat{\boldsymbol{A}}_{\mathcal{C}*}^{\text{Lib}}$	$\widetilde{P} \times N$ matrix containing the representation of $\widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}}$ in the original spatial domain
$\boldsymbol{Y}_{\mathcal{D}}$	$L \times N$ matrix containing the spatial details of the hyperspectral image $\boldsymbol{Y}$
$\boldsymbol{y}_{\mathcal{D}_i}$	i-th column of $\boldsymbol{Y}_{\mathcal{D}}$
$\boldsymbol{A}_{\mathcal{D}}$	$P \times N$ matrix containing the spatial details of the abundances $\boldsymbol{A}$
$\boldsymbol{a}_{\mathcal{D}_i}$	i-th column of $\boldsymbol{A}_{\mathcal{D}}$
$\boldsymbol{E}_{\mathcal{C}}$,$\widetilde{\boldsymbol{E}}_{\mathcal{C}}$	$L \times N$ representation of the noise and modeling errors $\boldsymbol{E}$ in the coarse spatial domain
$\boldsymbol{E}_{\mathcal{D}}$,$\widetilde{\boldsymbol{E}}_{\mathcal{D}}$	$L \times N$ matrix containing the spatial details of the noise and modeling errors $\boldsymbol{E}$
$\boldsymbol{M}_{\mathcal{C}_i}$	Average of all $\boldsymbol{M}_j$ inside the i-th superpixel

$\boldsymbol{M}_{\mathcal{C}_i^*}$	Average of all $\boldsymbol{M}_j$ inside the superpixel that contains the i-th pixel
$\boldsymbol{M}_{\mathcal{D}_i}$	Variability of $\boldsymbol{M}_i$ with respect to the average endmember matrix of its neighborhood $\boldsymbol{M}_{\mathcal{C}_i^*}$
$\operatorname{tr}(\cdot)$	Matrix trace operator
$\|\cdot\|_{1,1}$	Matrix $\mathcal{L}_1$ norm
$\|\cdot\|_F$	Frobenius norm
$\|\cdot\|_{2,1}$	Mixed $\mathcal{L}_{2,1}$ norm of a matrix
$[\cdot]_{i,j}$	Element in the i-th row and j-th column of a matrix
$[\cdot]_i$	i-th column of a matrix
$[\cdot]_{:,:,n}$	Matrix obtained by fixing the mode-3 index of a tensor
$[\cdot]_{\ell,k,:}$	Slice obtained by fixing all but the third mode of a tensor
$\iota_+(\cdot)$	Indicator function of the nonnegative orthant
$\iota_{\mathcal{S}^1}(\cdot)$	Indicator function of the unity simplex
$\operatorname{soft}(\cdot,\cdot)$	Soft threshold function
$\operatorname{vec}(\cdot)$	Vectorization operator
$\operatorname{diag}(\cdot)$	Diagonal matrix
$(\cdot)^\dagger$	Pseudoinverse operator
$\mathcal{P}_+$	Operator that projects each element of a matrix onto the nonnegative orthant
$\mathcal{H}_v, \mathcal{H}_h$	Linear operators that compute the vertical and horizontal gradients of an image
$\mathcal{F}, \mathcal{F}^{-1}$	Bi-dimensional discrete Fourier transform and its inverse
$\lvert\cdot\rvert$	Cardinality of a set
$\operatorname{supp}_{\boldsymbol{x}}(g)$	Support of a function g as a function of $\boldsymbol{x}$
$\mathbb{1}_x$	Indicator function containing ones on the support of x
$\odot$	Hadamard (elementwise) product
$\otimes$	Kronecker product
$\mathcal{O}$	Standard big-O notation
$\dim \operatorname{span}$	Dimension of the vector space generated by a set of vectors
$\operatorname{SVD}(\boldsymbol{X})$	Set of singular values of $\boldsymbol{X}$ sorted descending in value
$\operatorname{diff}(\boldsymbol{x})$	Vector containing first order (forward) differences of the elements of $\boldsymbol{x}$
$\mathcal{S}^1$	Unity simplex
$\mathrm{E}_\varsigma\{\cdot\}$	Expected value operator with respect to the distribution ς
$\mathrm{KL}(\cdot\|\cdot)$	Kullback-Leibler divergence between two distributions
TDiag_R	R-dimensional diagonal tensor
$\circ$	Outer product
$\times_k$	Mode-k product
$\times^k$	Contracted mode-k product
$[\![\ldots]\!]$	Full multilinear product

$\mathrm{mat}_i(\cdot)$	Matricization of a tensor obtained by stacking all fibers along its i-th dimension.
$f(\cdot,\cdot)$	Parametric function representing the spectra of the endmembers
$\mathcal{R}(\cdot)$	Regularization functional
$\mathcal{J}(\cdot)$	Cost function
$\mathcal{N}$	Set of pixels comprising a compact spatial neighborhood
$\boldsymbol{\Upsilon}_{\mathcal{P},p}$	Set of pure pixels of the p-th endmember
β	Regularization parameter
$\lambda_\bullet$	Regularization parameter
ρ	Regularization parameter
$N_{\boldsymbol{x}}$	Number of elements in vector $\boldsymbol{x}$
$\epsilon_\bullet$	Stopping criterion for optimization problems
$\boldsymbol{\mu}$	Vector of Lagrange multipliers
$\mathcal{G}_\eta$	Generator (decoder) neural network
$\mathcal{D}_\tau$	Encoder neural network
$\mathcal{C}_\tau$	Discriminator in a generative adversarial network
τ, η	Parameters of the neural networks
$\widetilde{\mathcal{G}}(\cdot)$	Concatenation of the generative models for each endmember
$\boldsymbol{Z}_n$	$H \times P$ matrix containing the latent endmember representations
$\boldsymbol{z}_{p,n}$	p-th column of $\boldsymbol{Z}_n$
$\mathcal{Z}$	$H \times P \times N$ tensor containing the latent endmember representations for all pixels
$\boldsymbol{Z}_0$	$H \times P$ matrix containing the latent representations of $\boldsymbol{M}_0$
$\boldsymbol{z}_{p,0}$	p-th column of $\boldsymbol{Z}_0$
X, Z	Random variables corresponding to observations $\boldsymbol{x}$ and $\boldsymbol{z}$
$p(X)$	Probability density function of X
$q_\tau(Z\|X)$	Variational approximation to the intractable posterior $p(Z\|X)$
$\boldsymbol{\Theta}$	Matrix containing the parameter vectors of an endmember variability model for all pixels
$\boldsymbol{\theta}_n$	n-th column of $\boldsymbol{\Theta}$
$\boldsymbol{dM}_n$	$L \times P$ matrix representing additive endmember variability in the n-th pixel
$\boldsymbol{\Phi}$	$P \times N$ matrix containing endmember-dependent scaling factors for all pixels
$\boldsymbol{\phi}_n$	n-th column of $\boldsymbol{\Phi}$
ϕ_n	Scalar pixel-dependent scaling factor
$\boldsymbol{\Psi}_n$	$L \times P$ matrix containing endmember and band-dependent scaling factors for the n-th pixel
Ψ	$L \times P \times N$ tensor containing the scaling matrices $\boldsymbol{\Psi}_n$ for all pixels
$\boldsymbol{\zeta}_n$	Function representing nonlinear mixing effects at the n-th pixel
$\boldsymbol{\zeta}_{\mathcal{C}_i}$	Function representing nonlinear mixing effects at the i-th superpixel

$\boldsymbol{\zeta}_{\mathcal{C}_n^*}$	Function representing nonlinear mixing effects at the superpixel which contains the n-th pixel
$\boldsymbol{\xi}_n$	$L \times 1$ vector containing the bandwise reconstruction error for the nonlinear mixing mode for the n-th pixel
$\boldsymbol{\xi}_{\mathcal{C}_i}$	$L \times 1$ vector containing the bandwise reconstruction error for the nonlinear mixing mode for the i-th superpixel
$\boldsymbol{\xi}_{\zeta,n}$	$L \times 1$ vector containing the inter-scale variations of the nonlinear mixing effects for the n-th pixel
$\boldsymbol{\Sigma}_{\boldsymbol{e}}$	$L \times L$ covariance matrix of the observation noise $\boldsymbol{e}_n$
ν_{err}	Modeling errors
V_0, V_1	Constants related to the coarse and fine scale reconstruction errors for a nonlinear mixing model
$\boldsymbol{Y}^{(t)}$	$L \times N$ representation of a hyperspectral image indexed at time t
$\boldsymbol{A}^{(t)}$	$P \times N$ representation of the abundance maps indexed at time t
$\boldsymbol{M}^{(t)}$	$L \times P$ endmember matrix indexed at time t
$\boldsymbol{E}^{(t)}$	$L \times N$ representation of the observation noise indexed at time t
$\boldsymbol{\theta}^{(t)}$	Vector containing the parameters of an endmember variability model indexed at time t
$\boldsymbol{c}^{(t)}$	Temporal innovation vector of $\boldsymbol{\theta}^{(t-t)}$
$\boldsymbol{r}^{(t)}$	$LN \times 1$ vector containing observation noise and the effect of abundance variations indexed at time t
$\boldsymbol{C}$	Covariance matrix of $\boldsymbol{c}^{(t)}$
$\boldsymbol{\Psi}^{(t)}$	$L \times P$ matrix containing the endmember scaling factors indexed at time t
$\boldsymbol{\Delta A}^{(t)}$	$P \times N$ matrix containing the fluctuation of the abundances about its mean at time t
$\boldsymbol{A}_{\text{avg}}$	$P \times N$ matrix containing the average abundance values
$\boldsymbol{\delta}^{(t)}_{\text{small},n}$	$P \times 1$ vector containing small abundance changes at pixel n and time t
$\boldsymbol{\delta}^{(t)}_{\text{large},n}$	$P \times 1$ vector containing large abundance changes at pixel n and time t
$\boldsymbol{M}^{(t)}_n$	$L \times P$ endmember matrix indexed at pixel n and time t
$RE^{(t)}_{M,n}$	Reconstruction error of the endmember selection problem at pixel n and time t
RE_0	Threshold for detecting abundance changes during endmember selection
$\boldsymbol{Y}_m$	$L \times \widetilde{N}$ representation of a multispectral image
$\widetilde{N}$	Number of pixels in a multispectral image
$\widetilde{N}_1, \widetilde{N}_2$	Number of rows and columns of a multispectral image
L_m	Number of spectral bands in a multispectral image
$\boldsymbol{E}_m$	$L \times \widetilde{N}$ matrix containing the additive noise in a multispectral image
$\mathcal{Y}_m$	$L \times \widetilde{N}_1 \times \widetilde{N}_2$ tensor containing the multispectral image pixels
$\mathcal{E}_m$	$L \times \widetilde{N}_1 \times \widetilde{N}_2$ tensor containing the observation noise in the multispectral image

$\boldsymbol{D}_{\text{spat}}$	$\widetilde{N} \times N$ matrix representing optical blurring due to the sensor point spread function and spatial downsampling
$\boldsymbol{D}_{\text{spec}}$	$L_m \times L$ matrix containing the spectral response functions of the multispectral instrument
$\boldsymbol{D}_{k,\text{spat}}$	$\widetilde{N} \times N$ matrix representing spatial blurring and downsampling in the hyperspectral sensor for the k-th spatial dimension
$\boldsymbol{S}$	$L \times \widetilde{N}$ high resolution image in the variability-free model
$\boldsymbol{S}_h, \boldsymbol{S}_m$	$L \times \widetilde{N}$ high resolution images underlying the hyperspectral and multispectral observations
$\mathcal{S}_h, \mathcal{S}_m$	$L \times \widetilde{N}_1 \times \widetilde{N}_2$ tensors representations of $\boldsymbol{S}_h$ and $\boldsymbol{S}_m$
$\mathbb{\Delta}$	$L \times \widetilde{N}_1 \times \widetilde{N}_2$ additive variability tensor representing changes between the hyperspectral and multispectral images
$\boldsymbol{M}_m$	$L \times P$ matrix containing the endmembers in the multispectral image
$\boldsymbol{A}_{\text{HR}}$	$P \times \widetilde{N}$ matrix containing high resolution abundance maps
Ω_S, Ω_Δ	Sets containing prior information about the high resolution image and the variability factor

1 INTRODUCTION

"The sky above the port was the color of television, tuned to a dead channel."
William Gibson, *Neuromancer*

The introduction of the first multispectral scanners in the 1960s allowed reflectance spectra to be sampled with higher flexibility and resolution than ordinary monochromatic or RGB cameras. The spectral resolution of these instruments was soon further enhanced with the development of hyperspectral (HS) sensors, which are able to sample reflectance spectra at a large number of contiguous wavelength intervals. Modern HS instruments are able to acquire hundreds or even thousands of spectral bands from the visible up to the infrared spectra [1]. An illustrative depiction of a hyperspectral image (HI) and its acquisition in a remote sensing context can be seen in Figs. 1 and 2.

The high spectral resolution of HIs makes them an important tool for the precise identification and discrimination of different materials in a scene. HIs contribute significantly to different fields and are now at the core of a vast number of applications such as space exploration [3], land-use analysis, mineral detection, environment monitoring, field surveillance [1, 4], disease diagnosis and image-guided surgery [5].

Notwithstanding the advantages brought forth by their high spectral resolution, hyperspectral cameras operate on a delicate trade-off between spatial resolution and signal-to-noise ratio (SNR). This happens since the light observed at the sensor is decomposed into several spectral bands, which in turn demands the pixel size to be large enough to attain an acceptable SNR. When combined with a large target-to-sensor distance, which is common in many applications, this leads to images with low spatial resolution [6].

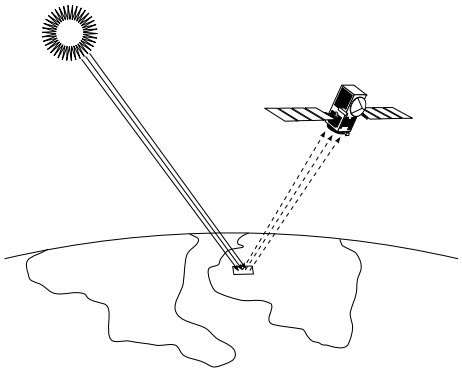


Figure 1 – The solar radiation is reflected at the surface of the Earth and captured by a sensor onboard a satellite [2].

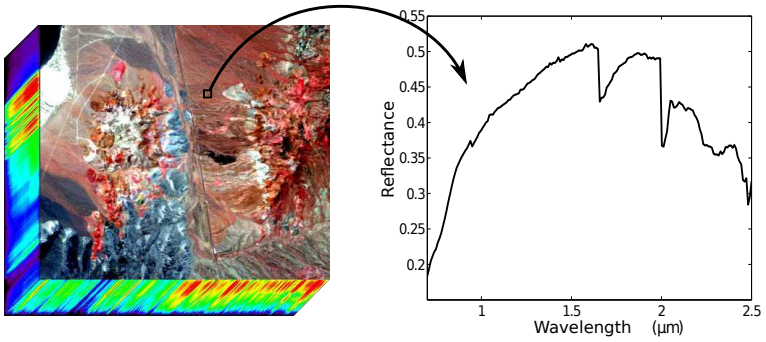


Figure 2 – Illustration of a pixel from a hyperspectral image [2].

1.1 THE SPECTRAL UNMIXING PROBLEM

The limited spatial resolution of HIs means that each image pixel is actually a mixture of P different pure materials, whose spectra are termed *endmembers* (EM), present in the scene [7]. This mixing process conceals important information about the pure materials and their distribution in an HI. Spectral unmixing (SU) aims to solve this problem by decomposing a hyperspectral image into the spectral signatures of the endmembers and their fractional *abundance* proportions for each pixel [8].

The simplest and most widely used model to represent the interaction between light and the EMs in the scene is the Linear Mixing Model (LMM) [7], which represents a given pixel $\boldsymbol{y}_n$ indexed by n with L spectral bands as:

$$\boldsymbol{y}_n = \boldsymbol{M}_0\,\boldsymbol{a}_n + \boldsymbol{e}_n, \text{ subject to } \mathbf{1}^\top \boldsymbol{a}_n = 1 \text{ and } \boldsymbol{a}_n \geqslant \mathbf{0}\,, \tag{1}$$

where $\boldsymbol{M}_0 = [\boldsymbol{m}_{0,1}, \ldots, \boldsymbol{m}_{0,P}]$ is an $L \times P$ matrix whose columns are the P endmembers, $\boldsymbol{a}_n$ is a vector containing the abundances of every endmember in the pixel $\boldsymbol{y}_n$ and $\boldsymbol{e}_n$ is an additive noise vector. The constraints in (1) specify that all abundances should be nonnegative and sum to one, which allows them to be interpreted as the proportion of the pixel $\boldsymbol{y}_n$ corresponding to each EM.

Under the LMM, the SU problem consists of recovering the endmember spectra $\boldsymbol{M}_0$ and the abundances $\boldsymbol{a}_n$ given an observed HI $\boldsymbol{Y} = \left[\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\right]$ with N pixels. This problem, which is generally non-convex, is equivalent to the blind source separation of $\boldsymbol{Y}$. Nevertheless, it attracted a great amount of interest from the community and multiple strategies, not always originated from existing blind source separation solutions, have been proposed to solve it [9].

1.1.1 Spectral variability in SU

Traditionally, the LMM assumes that the signatures $\boldsymbol{M}_0$ of the pure materials are the same for all pixels $\boldsymbol{y}_n$, $n = 1, \ldots, N$ in the HI. Although this assumption leads to a well-posed and computationally simpler framework, it limits the applicability of the LMM

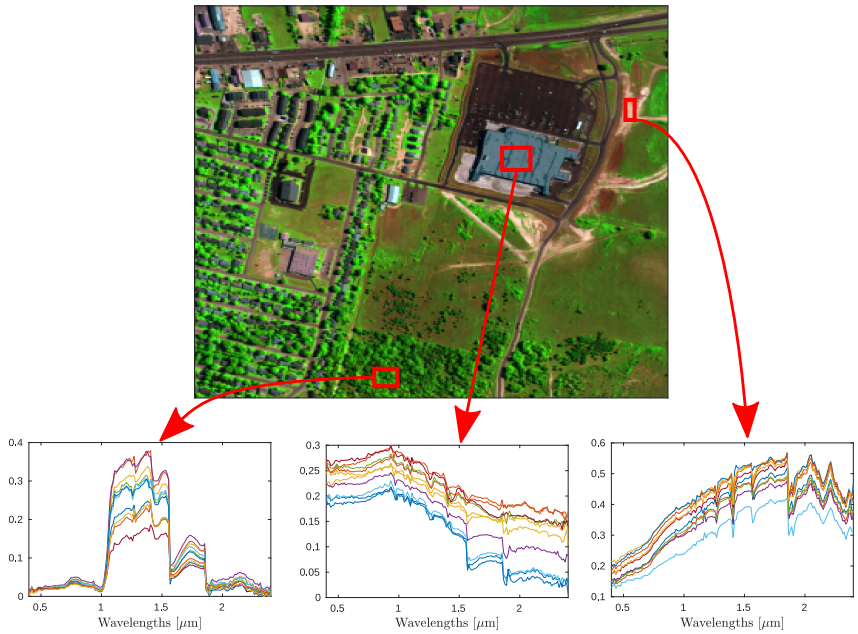


Figure 3 – Spectral variability is ubiquitous in hyperspectral images: the pixels in regions composed of a single material (e.g., tree, roof and soil in the image above) can contain very different spectral signatures.

since it can jeopardize the accuracy of estimated abundances in many circumstances due to the spectral variability of the endmembers.

Spectral variability is an effect commonly observed in many scenes in which the spectral signatures of the pure constituent materials vary across the observed HI, as illustrated in Fig. 3. It can be caused, for instance, by variable illumination and atmospheric conditions. Variability can also be intrinsic to the very definition of a pure material, such as signatures of a single vegetation species varying significantly due to different growing and environmental conditions [10, 11].

In this context, the use of a single matrix $\boldsymbol{M}_0$ for all pixels in the LMM (1) leads to problems such as *proportion indeterminacy*, where errors in the estimation of the endmember spectra at each pixel propagate to the estimated abundances. This results in erroneous abundance estimation and in the selection of too many endmembers to represent the spectrum of each pixel $\boldsymbol{y}_n$ [10, 12, 11]. Due to the significant impact of endmember variability on abundance estimation quality, a lot of effort has recently been dedicated to develop algorithms that are able to obtain better abundance estimates in this scenario.

The most general form of the LMM considering spectral variability generalizes (1) to allow for a different endmember matrix for each pixel, resulting in:

$$\boldsymbol{y}_n = \boldsymbol{M}_n \boldsymbol{a}_n + \boldsymbol{e}_n, \text{ subject to } \mathbf{1}^\top \boldsymbol{a}_n = 1 \text{ and } \boldsymbol{a}_n \geqslant \mathbf{0}\,, \tag{2}$$

for $n = 1, \ldots, N$, where $\boldsymbol{M}_n \in \mathbb{R}^{L \times P}$ is the n-th pixel endmember matrix.

SU considering spectral variability can be generally defined as two complementary problems related to: 1) the recovery of the abundances; and 2) the recovery of the

endmembers. These can be defined as:

P_1 : To mitigate the adverse effects of spectral variability in the abundance estimation;

P_2 : To estimate the spectral signatures of the endmembers present in each pixel of the image.

Substantial interest has been recently raised by both of these problems. While all SU methods must deal with P_1 while accounting for spectral variability, not all of them take P_2 into consideration due to the additional difficulty it entails.

Several review papers exist on the topic and provide an extensive discussion of the algorithms devised to solve these problems [11, 10, 13, 14]. In particular, the recent review [15] (which was written as part of this thesis work) provides a comprehensive categorization of existing methods.

1.2 CHALLENGES AND MOTIVATION

Although many works proposed to address the SU problem with spectral variability, the challenging nature of this problem, both from a theoretical as well as from a practical standpoint, still make it a very active research topic. In particular, devising improved models to represent the endmembers and the abundances is critical to obtain high-quality unmixing results. Moreover, the need to process large amounts of information also calls for computationally efficient solutions.

1.2.1 Modeling the endmembers and the abundances

Due to the ill-posedness of the SU problem when spectral variability is considered in (2), incorporating adequate *a priori* information is critical to obtain good results. Therefore, a considerable amount of work has been dedicated to devise appropriate models to represent variability of EM spectra.

Early approaches considered modeling EMs as sets of spectral signatures called *spectral libraries*, whose elements encompass several variants of the spectrum of the material in question [16]. SU then involves finding which spectral signatures from the library are present in each pixel. Despite achieving significant popularity in, e.g., remote sensing applications [11], these approaches carry a significant downside in that the spectral libraries must be known a priori. This usually entails resource-consuming laboratory or *in situ* measurements to obtain spectra that are good representatives of the materials in the scene. Moreover, the computational complexity of many spectral library-based SU method increases very quickly with the number of spectra in the library.

A different line of work proposed to model the EMs using statistical distributions, with the spectral signatures of each element in each pixel being realizations of vector random variables [17, 18]. A significant advantage of this approach is that spectral libraries

are no longer required in order to perform SU, what allows for fully unsupervised approaches [10]. Moreover, the statistical formalism allows one to adopt principled methods to estimate the abundances and the parameters of the EM distributions from the data. However, the lack of mathematical tractability in these approaches imposes the use of simple distributions to represent the EMs in the scene (such as, e.g., a Gaussian), which may not be physically accurate. Moreover, the computational cost associated with these techniques can be very high.

Recently, physically motivated parametric models have been employed to represent EM spectra [19, 20]. The SU problem then consists of recovering both the abundances and a set of parameters which are low-dimensional representatives of the spectral signatures in the scene. This approach has demonstrated a good trade-off between the accuracy of SU results and the computation times of the algorithms. Moreover, there exists significant freedom in the definition of the parametric models. However, great care is necessary when constructing both the parametric model and the SU algorithm in order to obtain good results.

Besides modeling endmember spectra, another important step lies in choosing an appropriate model to represent the spatial arrangement of the abundances. The use of spatial contextual information has proven very successful in SU. One popular way to exploit this kind of information is to promote abundances of spatially close pixels to be similar to each other, which is the subject of, e.g., the Total Variation (TV) regularization [21]. More elaborate approaches also considered non-local redundancy in the images, exploring information that similar small patterns and shapes might occur at different locations of an HI [22]. Despite their success at improving abundance estimation results, such approaches usually lead to SU algorithms having a substantially higher computational complexity.

1.2.2 Computational Efficiency

The use of improved models can naturally increase the performance of SU methods. However, this also results in complex algorithms with typically much higher computational complexity when compared to traditional methods. This happens, for instance, when considering both nonlinearity [23] and endmember variability [16] in the mixing process. Nevertheless, most works are not concerned with massive increases in the computational cost when proposing new solutions.

This is an important issue for HI analysis since a basic characteristic of multiband images is that the volume of data increases linearly with the amount of spectral bands. Therefore, processing or even storing HIs can become challenging very quickly when the amount of data is large. What makes this problem specially prominent nowadays is the vast volumes of data being generated in remote sensing applications, putting this problem into the big data framework [24, 25]. Thus, it is of great interest to devise algorithms which can efficiently process large volumes of data without sacrificing the accuracy of the

results.

1.3 RELATED CHALLENGES IN NONLINEAR AND MULTITEMPORAL SU, AND MULTIMODAL IMAGE FUSION

An important aspect of the challenges mentioned in Section 1.2 is that they also appear in other problems in hyperspectral image analysis. This can open new possibilities to extend approaches aimed at overcoming such challenges to tackle different problems. For instance, while most SU algorithms are based on the LMM, the interaction between light and the different materials in a pixel often happens nonlinearly. This makes the LMM unable to adequately represent the mixing process [23]. Although several nonlinear mixing models have been proposed for SU [26], these usually contain a large amount of parameters, which can also make the SU problem ill-posed. Thus, the use of abundance spatial information can be important to improve the quality of the unmixing results.

Moreover, spectral variability of the EM signatures occurs not only within a single HI, but also among HIs acquired over multiple time instants [11, 10]. Thus, addressing spectral variability is also important when dealing with multitemporal SU, where seasonality can make changes in EM spectra very pronounced. Moreover, properly handling multitemporal data also entails handling additional challenges, such as dealing with abrupt changes in the image sequence [27].

Another popular approach to overcome the low spatial resolution of HIs is to combine them with multispectral images (MIs), which have higher spatial resolution but a small number of spectral bands, to obtain images with higher spatial and spectral resolutions [28]. Most works formulate this so-called image fusion problem considering that both the HI and the MI are acquired under the same conditions. However, in practice, many applications require the continuous monitoring of a given region over multiple time instants (e.g. vegetation monitoring), which can be performed by combining HIs and MIs acquired at different time instants (i.e., on board of different instruments) [29, 30, 31]. This, however, introduces a new challenge to the image fusion problem, since the materials in the scenes may be affected by spectral variability due to the different acquisition or seasonal conditions. Thus, addressing spectral variability can also have a positive impact on the image fusion problem.

1.4 OBJECTIVES, CONTRIBUTIONS AND ORGANIZATION

The objective of this work is to devise novel methods that advance the state of the art by addressing the limitations of current SU methods which consider spectral variability in terms of their modeling accuracy and computational complexity.

The particular objectives can be described as follows:

a) Propose new models which can better represent the abundances and the endmembers in SU when spectral variability is considered;

b) Devise efficient algorithms to solve the SU problem considering endmember variability.

An additional objective is to extend the proposed approaches to address multitemporal data, nonlinear mixtures, and multimodal image fusion.

1.4.1 Organization

We start by presenting an overview of the origins of spectral variability and their effects on SU in Chapter 2. Since a significant portion of this work is concerned with parametric approaches to represent EM spectra in SU, an in-depth overview about how spectral variability behaves in real scenes gives valuable insight for devising or adjusting models later on. Chapters 3–8 concern specific methodological contributions to SU with spectral variability, which are described in the remainder of this section.

Moreover, we note that the contributions generated during the thesis work were not limited to the SU problem, or to considering spectral variability. However, for conciseness and to avoid detracting from the main theme of the document, we present only a brief overview of such contributions in Chapter 9. Finally, Chapter 10 concludes this document. In the following, we describe the main contributions and the content of each chapter.

1.4.2 Multiscale spatial regularizations for fast SU

Using spatial regularization approaches such as the TV [21] or non-local methods [22] often leads to a considerable increase in the computational cost of SU. Moreover, signal-agnostic regularizations such as the TV have limited capability to represent the spatial content of complex HIs. This sets forward two (often conflicting) goals in the design of spatial regularization strategies: better explore spatial information (i.e., improve SU performance) and reduce the computational complexity of SU.

In Chapters 3 and 4, we explore a different idea to design spatially regularized SU methods in both supervised (library aided) and blind SU based on a signal-adaptive, multiscale spatial representation of the HI. Using the superpixels decomposition as a support tool, we divide the image in two spatial scales, one containing only the coarse image content and another one which also contains the fine-scale image details. Based on this representation, we introduce spatial regularity information into the reconstructed abundance maps by using either a two-step unmixing method with an inter-scale regularization (in Chapter 3), or in a single step by controlling the energy of the estimated abundances in each of the two scales (in Chapter 4). Both algorithms are able to improve the abundance reconstruction results at a reduced computational complexity. These chapters are related to the publications [32, 33].

1.4.3 New models for EM spectra: Parametric, tensor, and neural-network-based strategies

Chapters 5, 6, 7 and 8 present new models to represent the variability of EM spectral signatures in both supervised and blind SU. Despite the success achieved by recently proposed parametric EM models in this problem (see, e.g., [20, 19]), there is still need for models which are flexible enough to properly represent the variability seen in many practical applications while also being physically interpretable. Considering the limitations of current models, we propose three new models to represent EM spectra, using parametric, tensor, and neural network-based representations.

In Chapter 5, we propose a new parametric model to represent EM variability at each pixel using a band-dependent scaling of a fixed, reference spectral signature of each endmember. This extends the model previously proposed in [19], which considers spectrally invariant scaling factors. The new model can account for complex spectral distortions where different wavelength intervals can be affected unevenly. Moreover, it also relates the amount of the spectral variability per wavelength with the amplitude of the reference spectral signatures of the endmembers. This chapter is related to the publication [34].

In Chapter 6 we propose to explore the high-order (multidimensional) structure of the pixel dependent EMs in the HI, when organized as a tensor, in order to regularize the resulting SU problem. The endmember tensor is assumed to be of approximately low rank, which greatly reduces the number of degrees of freedom and makes the SU problem well posed. A similar low-rank assumption is considered for the abundance maps, and SU is performed by solving an optimization problem using a block coordinate descent approach. The proposed model was able to effectively reduce the amount of unknowns to be estimated in the problem and explore the spatial-spectral structure of the data, as opposed to adding strict constraints to the problem *a priori*. This chapter is related to the publication [35].

In Chapter 7 and Chapter 8, we represent the spectral signatures of the endmembers using the perspective of deep generative modeling. Although material spectra reside in a high-dimensional space, an important characteristic of spectral variability is that the variation of spectral signatures of a material usually depends on a very small number of parameters [36, 37]. This behavior can be seen in many practical applications, and supplies important information to the design of new models: the spectra of a material is confined to a low-dimensional manifold. This makes the task of learning spectral variability much more tractable. Based on this information, in Chapter 7 we represent the manifold of spectral signatures of each EM as the image of an unknown function acting on a low-dimensional latent domain. This representation of the EM signatures is then used to parametrize the SU problem in a matrix factorization-inspired framework. Functions mapping from the latent domain to the EM signatures are learned from the observed HI using deep generative

models (DGMs) such as variational autoencoders (VAEs) or generative neural networks (GANs), which allows for unsupervised SU (i.e., without the need for spectral libraries). This chapter is related to publication [38].

In Chapter 8, we extend the approach of Chapter 7 to improve supervised (library aided) SU. One of the main downsides of library-based SU is that the quality of the unmixing results depends directly on how well spectra in the library is able to represent the EM signatures that are actually present in the HI. To address this issue, we propose to use DGMs to augment existing spectral libraries with synthetic spectra. Using a statistical interpretation of DGMs, the statistical distribution of EM spectra is first learned from the signatures in the library. Afterwards, new spectra are then sampled and added to the existing library before SU. This allows for the augmentation of existing libraries without requiring a precise physical modeling of EM spectra. This chapter is related to the publication [39].

1.4.4 Extending the contributions to new problems

In Section 9, we detail research performed during the course of this thesis work which extends the methods from Chapters 3–8 to related problems in SU and in image fusion.

1.4.4.1 Multiscale spatial regularization for nonlinear SU

The multiscale regularization strategy proposed in Chapter 3 was also extended to address nonlinear SU with kernels. Although this was based on the same underlying concept of decomposing SU in two different spatial domains (containing the coarse and the original image content, respectively), some important challenges were addressed in this work. After applying the multiscale decomposition to a non-parametric nonlinear mixture model, the SU problem was formulated as non-convex quadratically constrained optimization problems. This allowed for the development of a theoretically founded methodology to select the regularization parameters of the algorithm automatically. Moreover, after showing that strong duality holds for the proposed optimization problems, efficient solutions were devised by exploring their dual formulations and a root finding strategy. This work is related to the publication [40].

1.4.4.2 Multitemporal spectral unmixing and change detection

Recently, significant interest has been dedicated to the problem of SU of multitemporal HI sequences (MTSU – multitemporal SU). Besides the difficulties already experienced in traditional (i.e., single image) SU, this problem poses the additional challenges of handling EM variability in time (which can be more pronounced than within each HI) and abrupt changes between temporally adjacent images. However, MTSU also allows

one to improve SU results by exploring non-redundant information contained in multiple HIs, and can provide more complete information to the practitioner by tracking the time evolution of the EMs and their corresponding abundances. In this context, we proposed a new approach to estimate the EMs in multitemporal HIs by coupling a temporal extension of parametric models to represent spectral variability with a Bayesian filtering methodology. Most of the parameters of the model were estimated from the observed HIs using the Expectation Maximization algorithm. However, this approach needs the abundances to vary slowly over time (during short time intervals) to work satisfactorily. To better deal with abrupt changes between adjacent HIs, we later considered to perform MTSU and change detection jointly using spectral libraries. To this end, we extended a popular combinatorial algorithm that searches for the spectral signatures that best represent each pixel [16] to the multitemporal setting. Since this algorithm has a very high computational cost, a fast approximate solution was proposed to make it scalable. Theoretical results were supplied showing that under certain conditions, the approximate algorithm is able to operate well both with constant abundances and under the presence of sudden changes. This work is related to the publications [41, 42].

1.4.4.3 Addressing spatial and spectral variability in image fusion

Although spectral variability has been addressed mostly from a SU viewpoint, it can also significantly impact the performance of multimodal (i.e., hyperspectral and multispectral) image fusion algorithms. Despite the significant interest raised by this problem, previous algorithms addressed it by considering that the images from both modalities are acquired under the same conditions (e.g., by sensors on-board of the same instrument). However, this is not always verified in practice, what makes the performance of state-of-the-art algorithms degrade considerably when processing images acquired by different sensors or at different time instants. To address this problem, we proposed to incorporate spectral variability in a matrix factorization-based image fusion problem by allowing the image from each modality to have a unique set of spectral signatures or spectral basis vectors, which successfully accounted for spatially invariant changes. This work was later extended in order to also account for spatially localized changes using a tensor-based representation. The image fusion problem with (spatial-spectral) variability was formulated as a coupled tensor factorization problem, and two image fusion algorithms were devised (one purely algebraic, and another based on an optimization algorithm). Moreover, theoretical results were also obtained concerning the identifiability of the model and exact recovery conditions for each of the algorithms. This work is related to the publications [43, 44].

Table 1 – Subjects treated in each paper published or submitted during the duration of this doctorate.

Subject / Paper	SU	Spectral variability	Spectral libraries	Non-linear	Image fusion	Multi-scale	Multi-temporal	Tensors	Neural nets	Change detection
C1	✓	✓								
C2	✓							✓		
C3	✓	✓						✓		
C4							✓			✓
J1	✓		✓			✓				
J2		✓			✓		✓			
J3	✓	✓							✓	
J4	✓	✓						✓		
J5	✓	✓				✓				
J6	✓			✓		✓				
J7	✓	✓	✓						✓	
J8	✓	✓					✓			
J9		✓			✓		✓	✓		✓
J10	✓	✓								
PP1	✓	✓	✓				✓			✓

1.5 LIST OF PUBLICATIONS

The publications related to this thesis are listed below, divided into conference papers (starting with C#), journal papers (J#) and preprints (PP#). They contain the main works [32, 33, 34, 35, 38, 39, 40, 41, 42, 43, 44] discussed above, as well as preliminary work mainly published in conferences [45, 46, 47] and a review article [15]. The main subjects treated in each publication are also briefly described in Table 1.

[C1] T. Imbiriba, **R.A. Borsoi**, J.C.M. Bermudez. Generalized linear mixing model accounting for endmember variability. IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP), 2018.

[C2] T. Imbiriba, **R.A. Borsoi**, J.C.M. Bermudez. A low-rank tensor regularization strategy for hyperspectral unmixing. IEEE Statistical Signal Processing Workshop (SSP), 2018.

[C3] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez. Improved hyperspectral unmixing with endmember variability parametrized using an interpolated scaling tensor. IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP), 2019.

[C4] **R.A. Borsoi**, C. Richard, A. Ferrari, J. Chen, J.C.M. Bermudez. Online Graph-Based Change Point Detection in Multiband Image Sequences. European Signal Processing Conference (EUSIPCO), 2021.

[J1] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez, C. Richard. A fast multiscale spatial regularization for sparse hyperspectral unmixing. IEEE Geoscience and Remote Sensing Letters, 2018.

[J2] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez. Super-resolution for hyperspectral and multispectral image fusion accounting for seasonal spectral variability. IEEE Transactions on Image Processing, 2019.

[J3] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez. Deep generative endmember modeling: An application to unsupervised spectral unmixing. IEEE Transactions on Computational Imaging, 2019.

[J4] T. Imbiriba, **R.A. Borsoi**, J.C.M. Bermudez. Low-rank tensor modeling for hyperspectral unmixing accounting for spectral variability. IEEE Transactions on Geoscience and Remote Sensing, 2019.

[J5] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez. A data dependent multiscale model for hyperspectral unmixing with spectral variability. IEEE Transactions on Image Processing, 2020.

[J6] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez, C. Richard. A blind multiscale spatial regularization framework for kernel-based spectral unmixing. IEEE Transactions on Image Processing, 2020.

[J7] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez, C. Richard. Deep generative models for library augmentation in multiple endmember spectral mixture analysis. IEEE Geoscience and Remote Sensing Letters, 2020.

[J8] **R.A. Borsoi**, T. Imbiriba, P. Closas, J.C.M. Bermudez, C. Richard. Kalman filtering and expectation maximization for multitemporal spectral unmixing. IEEE Geoscience and Remote Sensing Letters, 2020.

[J9] **R.A. Borsoi**, C. Prévost, K. Usevich, D. Brie, J.C.M. Bermudez, C. Richard. Coupled Tensor Decomposition for Hyperspectral and Multispectral Image Fusion with Inter-image Variability. IEEE Journal of Selected Topics in Signal Processing, 2021.

[J10] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez, C. Richard, J. Chanussot, L. Drumetz, J.-Y. Tourneret, A. Zare, C. Jutten. Spectral Variability in Hyperspectral Data Unmixing: A Comprehensive Review. IEEE Geoscience and Remote Sensing Magazine, 2021.

[PP1] **R.A. Borsoi**, T. Imbiriba, J.C.M. Bermudez, C. Richard. Fast Unmixing and Change Detection in Multitemporal Hyperspectral Data. (submitted).

2 ORIGINS OF SPECTRAL VARIABILITY IN HYPERSPECTRAL IMAGES

The variability in the spectral signatures occurs mainly due to (a) atmospheric effects, (b) illumination and topographic changes, and (c) intrinsic variation of the spectral signatures of the materials (i.e., due to physico-chemical variations). Understanding how these conditions affect the spectral signatures of the materials and the unmixing results is important in order to develop informed models and methods to deal with EM variability. In the following, we review each of these effects and comment on their influence on the endmembers in the mixing model.

2.1 ATMOSPHERIC EFFECTS

One of the main sources of spectral variability is the interference by the atmosphere when measuring ground reflectance. Atmospheric gases (such as O_3, O_2, CH_4, CO_2, etc.), aerosols and, most prominently water vapor, absorb significant amounts of radiation, while other molecules and aerosols scatter incoming light [48]. These effects have an impact on the radiance measured at the sensor, which can become significantly different than that corresponding to the desired ground reflectance. Atmospheric absorption from gases is also heavily wavelength dependent, whereas aerosol absorption varies smoothly in spectra. These effects must be compensated to achieve an accurate characterization of surface reflectance.

Atmospheric compensation models can be roughly divided into statistical (empirical) and physics-based models [48]. Statistical models are based on additional information about the atmospheric influence, usually obtained by means of reference objects or calibration panels in the scene. This information is used to find a relationship (e.g., linear) between the radiances observed at the sensor and at the surface of the scene [48]. This results in a gain and an offset factor for each spectral band, which are then uniformly applied to every image pixel to compensate for the atmospheric effects [48]. Sometimes, when a reference object is not present in the scene, naturally occurring objects can be employed as reference spectra, most commonly consisting of smooth bodies of water, which exhibit low reflectance and can be considered as dark objects [6]. The downsides of this approach are that the true reflectance of a reference object must be accurately known, and that it does not account for the spatial variability of the distribution of gases and aerosols. This variability can be very significant, and thus can introduce spatially-dependent residual atmospheric effects. A classical example of statistical methods is the empirical line method (ELM) [6].

Physics-based models, on the other hand, are robust alternatives to empirical methods which do not assume additional information about the scene to be known. These methods are currently mature and widely used, addressing the limitations of empirical

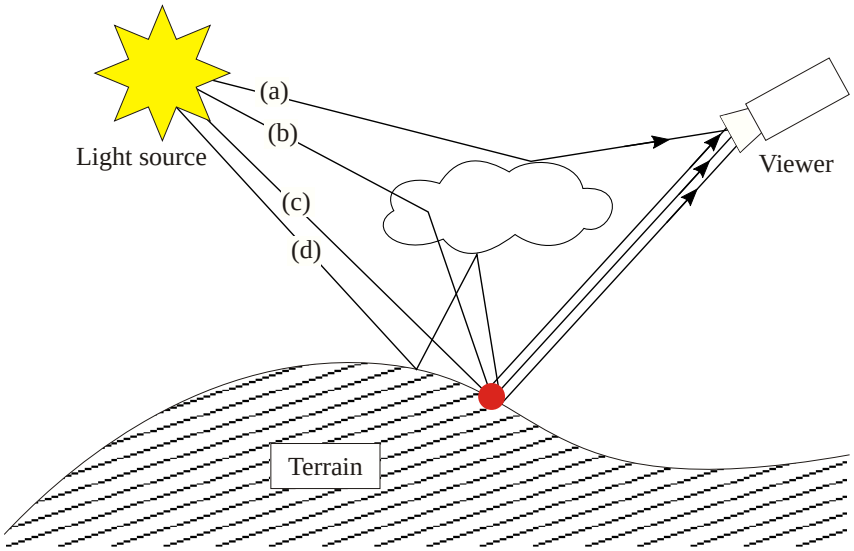


Figure 4 – Illustration of the effects of the atmosphere on the acquired hyperspectral image. The sources of radiation are represented by (a) light directly reflected by the atmosphere to the sensor, (b) light scattered by the atmosphere and reflected by the ground, (c) light directly reflected by the ground and (d) light reflected by surrounding regions on the ground and then scattered to the sensor.

methods by employing a rigorous model that explicitly describes the absorption and scattering effects due to atmospheric gases and aerosols [49]. Popular examples include the Atmospheric Removal (ATREM) and the Fast Line-of-Sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) algorithms [48].

Assuming a ground terrain illuminated by the sun, the light incident on a pixel in the sensor can be roughly characterized by four sources: solar radiation directly reflected off the ground, light directly reflected off the atmosphere into the sensor, light scattered by the atmosphere and reflected off the ground, and light that is reflected off surrounding regions on the ground and then scattered before reaching the sensor (constituting the adjacency effect) [50, 51]. These effects are illustrated in Fig. 4. A model for the reflectance at the sensor y_{sensor} is given by [48]:

$$y_{\text{sensor}} = y_{\text{atm}} T_{\text{g}} + \frac{y_{\text{s}} T_{\text{g}} T_{\downarrow} T_{\uparrow} + (y_{\text{avg}} - y_{\text{s}}) T_{\text{g}} T_{\downarrow} T_{\uparrow} r}{1 - y_{\text{avg}} s} \tag{3}$$

where y_{s} is the surface reflectance, T_{g} is the gaseous transmittance, y_{atm} the reflectance of the atmosphere, $T_{\downarrow}$ and $T_{\uparrow}$ are the upward and downward scattering transmittances, r is the ratio between diffuse and total transmittance for the ground-to-sensor path, s is the spherical albedo of the atmosphere, and y_{avg} is the average surface reflectance in a region around a pixel, which is used to account for scattering (adjacency) effects [48].

Physics-based atmospheric correction algorithms then try to obtain the ground reflectance y_{s} from the at-sensor reflectance y_{sensor} by solving (3). In the overall working of these algorithms the first step for atmospheric compensation consists of retrieving the atmospheric parameters necessary to represent the quantities in (3), mainly consisting of

aerosol description (visibility and type of aerosol) and amount of water vapor for each pixel [52]. They are typically based on variations of the so-called three-band ratio technique, which is an important step used to quantify the amount of water vapor for each pixel. The three-band ratio technique basically compares ratios of radiances measured near the edges of a number of spectral wavelengths which are known to present heavy water-vapor absorption (e.g., at around 0.91 μm, 0.94 μm and 1.14 μm), using this information to derive the column water vapor information for each pixel [6, 53]. After the necessary parameters have been estimated, (3) can be solved for the ground reflectance and an optional post-processing step can be employed (called spectral polishing) to remove artifacts from the correction process [52].

Physics-based models can represent and account for the interaction between solar radiation and the atmosphere very accurately. However, for this accuracy to translate into meaningful surface reflectance estimates, these models require precise information about atmospheric properties, which are very difficult to obtain in practice. This is specially true for scattering and absorption by aerosols, which are hard to characterize accurately due to their spatial and temporal variability [54]. Inaccuracies in the estimation of these parameters (which include the atmospheric visibility, aerosol model type and an atmospheric model) introduce errors in the retrieved surface reflectance spectra that can be significant and spectrally non-uniform [55].

Furthermore, unlike water vapor compensation, which is performed on a pixel-by-pixel basis, most methods assume that individual aerosol and gas concentrations are uniform across the scene (resulting in a single transmittance spectrum being computed for each gas) [55, 52]. While this is true for some gases (such as NH_4, O_2, CH_4, CO_2, etc.) that are fairly constant in the atmosphere [53], it is far from true for aerosols, which may show significant variation in space [56, 57]. Aerosol concentration can vary depending on the environment (e.g., in large cities and rural areas), and thus must be informed by the user to the existing algorithms [53]. Moreover, standard aerosol types often do not adequately represent the scene being processed, leading to inaccuracies in the retrieved spectra [58]. Furthermore, experimental studies have found that aerosol optical thickness has a significant spatial variability within a single scene [56, 59] and is often correlated with cloud concentrations [59].

Some works attempted to estimate aerosol optical thickness for smaller patches of the image individually using shadow detection results [60], which depends on the presence of a large number of shadowed pixels. However, acquiring precise data for an accurate and possibly spatially variable atmospheric correction is generally difficult, which means that the results of common atmospheric compensation methods can be subject to significant errors [56]. For instance, a number of studies have investigated the residual errors in surface reflectance data after the application of atmospheric compensation methods by comparing the processed results with *in situ* data or using simulations. These studies

found that generally there is still an appreciable error in the retrieved reflectances. As an example, errors in the retrieved reflectance by atmospheric corrections due to the spatial variability of aerosol optical thickness over southern England were found to be of up to 1.7%, with 5% errors in the normalized difference vegetation index (NDVI) [56]. This can be significant for practical applications, as it corresponds to errors of up to 30% in biomass production estimates [61, 56]. Furthermore, standard methods for column water vapor retrieval loose accuracy when the aerosol optical thickness is high, leading to errors of up to 10% if aerosol effects are not properly compensated [62]. Note that experimental measurements in a water quality management application found significant differences between the true and retrieved spectral responses. Errors of up to 15% in reflectance spectra were found, more prominently concentrated in short (<450 nm) and long (>750 nm) wavelength intervals [63]. Another study evaluated a number of physics-based atmospheric correction methods in an experiment for a *playa and canola* target and found that although the average relative differences were moderate, ranging between 0.023 and 0.042, larger deviations of up to 0.12 occurred in the near-infrared region [64]. A study with simulated data found that incorrectly supplying input parameters to the model used in the FLAASH algorithm can lead to considerable errors in the retrieved reflectance, with an absolute difference of up to 0.11, and a strong sensitivity to moisture/optical depth (visibility) errors [55]. Also, very large errors can be introduced by a bad specification of the aerosol model type, with higher errors generally present in short wavelengths where scattering processes are most significant [55].

The influence of uncertainties in column water vapor and aerosol optical depth specification on SU was investigated in [57] (given their influence in the retrieved reflectances). The performance degradation was found to be more severe in abundance than in reflectance estimation, with degradation of up to 30% in high scattering conditions. The results were more severely affected due to uncertainties in water vapor amount than in aerosol optical thickness, although the latter showed a strong influence on the quality of the reconstructed abundance maps when the endmembers were spectrally similar.

Finally, it is interesting to highlight that two characteristics were noticed from these studies. First, the errors in the retrieved reflectances are fairly non-uniform in spectral bands, with large spikes often concentrated near bands where there is significant gas/water absorption [55, 57]. Second, errors due to bad aerosol specification are quite significant in short wavelengths (450 nm-750 nm), where they are concentrated [55, 63]. All these effects are illustrated in Fig. 5.

2.2 ILLUMINATION AND TOPOGRAPHIC EFFECTS

"There is a shadow under this red rock"

T.S. Elliot, *The Wasteland*

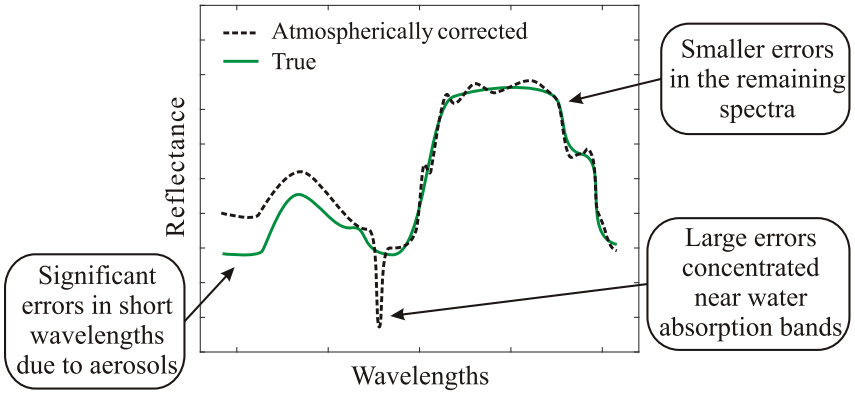


Figure 5 – Illustration of variability caused by atmospheric effects.

Varying illumination conditions are one of the main sources of spectral variability in spectral mixture analysis [65]. Illumination changes are mainly due to two effects: varying terrain topography, which affects the angles of the incident radiation, and occlusion of the light source by other objects (leading to shaded areas).

A number of work handled the presence of heavily shaded areas by considering the presence of an additional endmember representing shadow [66, 67, 68, 69, 16, 70, 71]. Although this approach is very simple, its effectiveness is certainly limited since a single spectral signature can be insufficient to adequately represent all pixels affected by shadow [72]. For instance, there might be many shadow endmembers since shadows in different regions of the image are influenced by both the material that is being shaded and by the absorption properties of the material that is blocking the light, what might lead to significantly different spectral signatures [73]. Furthermore, besides presenting a lower reflectance amplitude, the shadow EM is also usually significantly affected by nonlinear atmospheric scattering and multipath effects, since these areas are illuminated by a large proportion of diffuse irradiation scattered by the atmosphere (i.e., skylight) and by other nearby objects. This implies that the shadow endmember is sensitive to the state of the atmosphere and can vary significantly in space depending on the amount of scattered light being reflected from the sky at each position [74, 75].

When illumination predominantly comes from scattered radiation, the spectrum not only presents a lower amplitude but is also skewed to short (e.g., blue) wavelengths [76, 77]. This means that the signal amplitudes in the shorter (blue) wavelengths are considerably larger than in the rest of the spectra [77].

Furthermore, since the shadow spectral signature is a function of diffuse illumination, it depends on the neighboring image area (where the skylight is scattered) [77] and on the cloud cover. Moreover, variations of ground reflectance may not be easily discernible from atmospheric effects since both effects are observed jointly and are not easily separable [77]. These facts introduce a strong dependence of the shadow signature to the spatial position, and go against the common notion that shadow endmembers can be adequately represented by scaled versions of true endmembers [6] (that is only true for

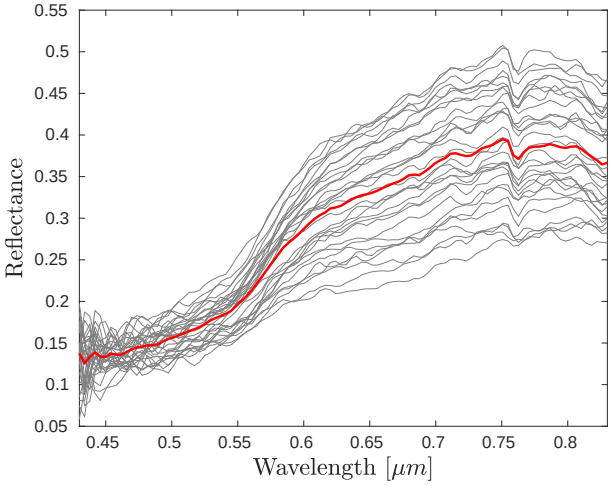


Figure 6 – Examples of 30 pixel instances classified as red roof in the Pavia image (in gray), which are primarily affected by illumination, and their spectral average (in red). The average Pearson correlation coefficient between each signature and the scaled version of the mean spectra that is closest to it is about 0.993, indicating a good agreement between illumination-based spectral variability and the constant scaling model.

small illumination variations).

This makes the detection, correction or quantification of shadow a challenging task, since physical-based inversion of these atmospheric effects turns out to be a hard problem. However, this task is still necessary since linear SU with a single dark endmember usually does not successfully quantifies the presence of shadow in the scene [77].

Although the presence of shadows is common in hyperspectral images, a more prominent source of variability comes from the varying topography of the scene, which introduces complex fluctuations of the relative angles between the incoming light source and the sensor for each pixel of the scene. Topographic variations have been shown to significantly affect spectral reflectance values of soil and green vegetation [78] as well as rocks in lithologic mapping [79], expanding endmember clusters and causing overlap between classes, hindering the endmember identification and unmixing processes.

Considering that only the amplitude of the incident radiation changes along the scene, the reflectance spectra of the observed pixels in the LMM becomes scaled by a constant positive factor. This model agrees with the observation that most of the variability in a hyperspectral image can be represented by a constant scaling of reference endmembers [6]. As a simple empirical verification, we plot a random subset of 30 pixels of red roofs from the Pavia image, which are pure pixels mostly affected by illumination effects. The results, which are depicted in Fig. 6, indicate that these pixels differ mostly by a scaling factor.

Although a constant scaling model is intuitive and simple, a more rigorous conclusion can be achieved by analyzing the dependence of radiative transfer models with the topography of the scene. To this end, one could resort to the model developed by Hapke [37,

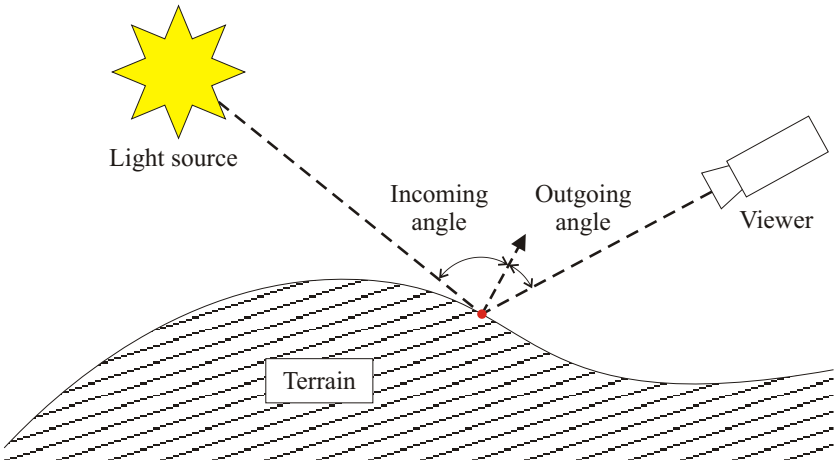


Figure 7 – Hapke's model relates the reflectance to the incidence angles of the light source and observer/viewer shown in this figure, given the material's single scattering albedo and photometric parameters [26].

80], which describes the bidirectional reflectance (i.e., the reflectance as a function of the incidence angles of the light source and observer/viewer depicted in Fig. 7) as a function of the single scattering albedo and of photometric parameters of the material [26].

Hapke's model suggests a more complex relationship between the endmember signatures and the topography. In this context, the mixture of materials is assumed to happen at the macroscopic level, allowing for the consideration of the LMM in the albedo domain, where Hapke's model acts separately on each endmember. Besides the dependency on the spectral signature with photometric parameters, which shall be discussed in the next section, the dependence on the single scattering albedo[1] indicates that changes in incident angles can affect each material in a pixel differently from the others, since the behavior of the reflectance as a function of the angle is different for each material. This indicates that each endmember/material in a pixel can be differently affected by topographic effects. Furthermore, the nontrivial relationship between geometry and the spectral signatures leads to a more complex variation than single scaling for each endmember for high albedo materials [19, 81]. Besides, even small topographic variations can significantly affect the ground reflectance. For instance, in [82] experimental studies found that even small slopes (of less than 10 degrees) originating from irregularities in tree canopy can lead to appreciable (enough to influence the results of subsequent tasks) changes in the measured reflectance of vegetation spectra.

2.3 INTRINSIC SPECTRAL VARIABILITY

Another important source of spectral variability is the intrinsic variation pertaining the definition of a material, which is also called *intrinsic variability*. The characteriza-

[1] i.e. the ratio between reflected and received radiation, as a function of the viewing angle.

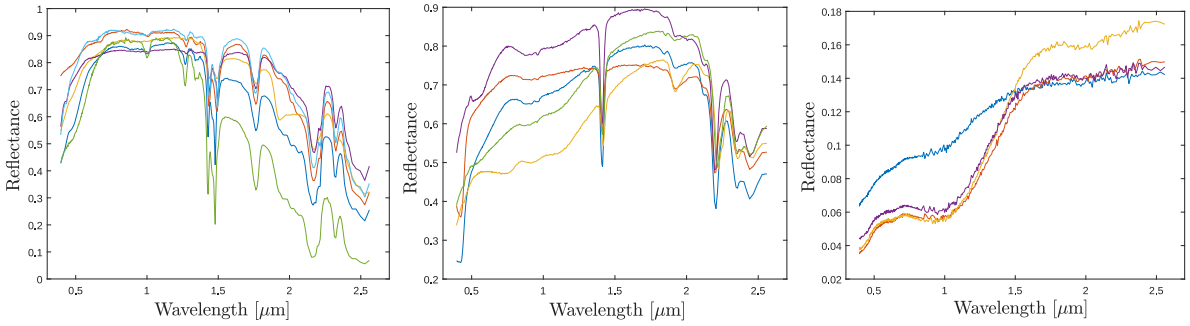


Figure 8 – Samples of variation of spectra from the USGS library. (a) Alunite. (b) Muscovite. (c) Pyrite.

tion of this type of variability has been prominently studied in the area of vegetation monitoring, where it poses a huge challenge to the ability to identify tree species from spectral measurements [83, 84], and also to the characterization of soil and mineral spectra. Vegetation spectral signature can change due to many factors, including micro-climates, soil characteristics, precipitation, presence of heavy metals and drought, foliage age and colonization by leaf pathogens [83]. The spectral signature of soil is also heavily affected by variations in its composition and moisture content [85]. Furthermore, intrinsic spectral variability is also common in mineral spectra due to variations in the grain size distribution and the presence of variable amounts of impurities [86, 87]. Moreover, it also depends on what level of detail is adopted to represent a given material (e.g., a *tree* endmember may possibly be split into *trunk* and *leaf* endmembers), which is generally application dependent [88]. Although imposing a large impact on the endmember spectral signatures, the dependence of intrinsic spectral variability on physico-chemical parameters, which are usually unknown, makes it very hard to tackle.

One characteristic consistently observed in experimental studies is the smoothness of the observed spectra (i.e., the reflectance varies slowly between spectral bands). This behavior can be taken into account when designing SU algorithms. Moreover, unlike spectral changes caused by illumination and topography effects, intrinsic spectral variability frequently presents a considerable dependence of the variability amplitude with the spectral wavelength. For instance, the signatures of different instances of minerals in the United States Geological Survey (USGS) library depicted in Fig. 8 show complex dependence between the reflectance variation and wavelength. The samples from alunite and muscovite show a variability that is far from uniform across the spectrum. Moreover, different instances from pyrite display complex variation, which is not consistent across all samples, occurring independently in different regions of the spectra. This behavior has been verified in similar experimental studies in other works, and poses a significant challenge for differentiating mineral classes based on their spectral signatures [89].

These characteristics are even more prominent and well known in the spectral

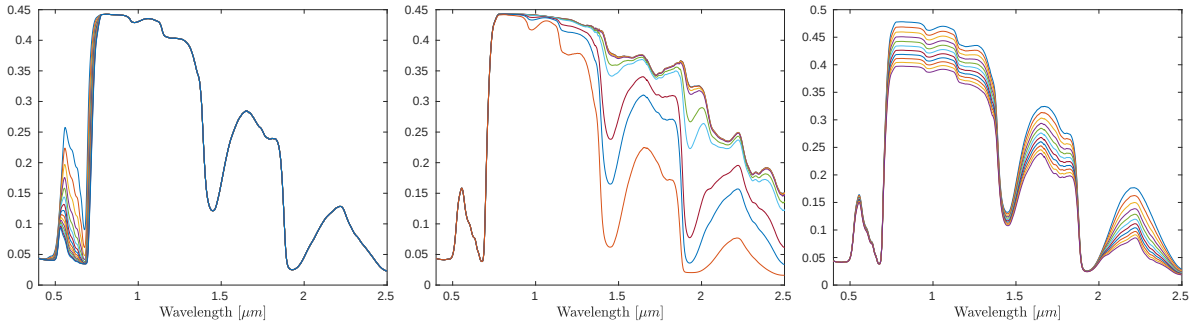
variation of vegetation reflectance, which shows significant dependency on the wavelength and behaves very differently in visible, near-IR, and short-wave-IR ranges [90]. This means that a simple scaling of a reference spectral signature is usually not sufficient to account for variations within tree species [83]. Extensive experimental studies support this claim. In [83] the author found that the variation of spectral reflectance in the visible and near-infrared regions can occur independently when measuring tropical forest canopy in Brazil. Similar inhomogeneity in spectral variation was also observed in other studies with tropical tree species [91] and also in many distinctive environments, including conifer [92] and boreal tree species [93]. Similar non-uniform variation trends are also consistently observed in seasonal changes as indicated by many experiments, including in salt marshes [94], semi-arid environments [95] and boreal tree species [96]. Furthermore, nonuniform spectral variations have also been observed in samples from mineral, soil and rock spectra [89].

Numerous works model the spectral signature of materials as a function of photometric or chemical properties of the medium, being based on either radiative transfer modelling or in empirical approaches. A well known example is Hapke's model, which describes the spectra of a surface composed of particles as a function of parameters such as surface roughness and density and size of the particles [37, 80].

Another prominent line of work models the spectral characteristics of vegetation and soil samples as a function of biophysical parameters [97]. Models of this kind have been applied for the estimation of leaf biochemistry from the observed spectra. An important example consists of the characterization of leaf reflectance spectra as a function of leaf biophysical parameters [97], for which a wide variety of models have been used, ranging from a simple description of leaf scattering and absorption properties to complex models which perform a detailed description of the plant cells' shape, size, position, and biochemical content [97]. Some instances of those models include the characterization of the spectra of broadleaf vegetation as a function of leaf mesophyll structure, pigment and water concentration [98] or as a function of leaf angular profiles [99], and of pine needles as a function of cellulose, lignin and water content [100]. Other works model soil reflectance spectra as functions of moisture conditions [101, 102, 103], and snow albedo as a function of snow grain sizes and liquid equivalent depth [104].

As an illustrative example, we generated spectral signatures of vegetation spectra using the PROSPECT-D model [36] as a function of varying degrees of chlorophyll content, equivalent water thickness and dry matter content. The resulting signatures, depicted in Fig. 9, show that intrinsic spectral variability can present complex patterns and non-uniformity, as it is often concentrated in specific regions of the spectrum.

Through their analytical characterization of EM spectra, these kinds of models confine spectral variability to lie on a low-dimensional manifold. This constitutes important information that can be leveraged to alleviate/reduce the severe ill-posedness of unsupervised SU problems accounting for spectral variability.



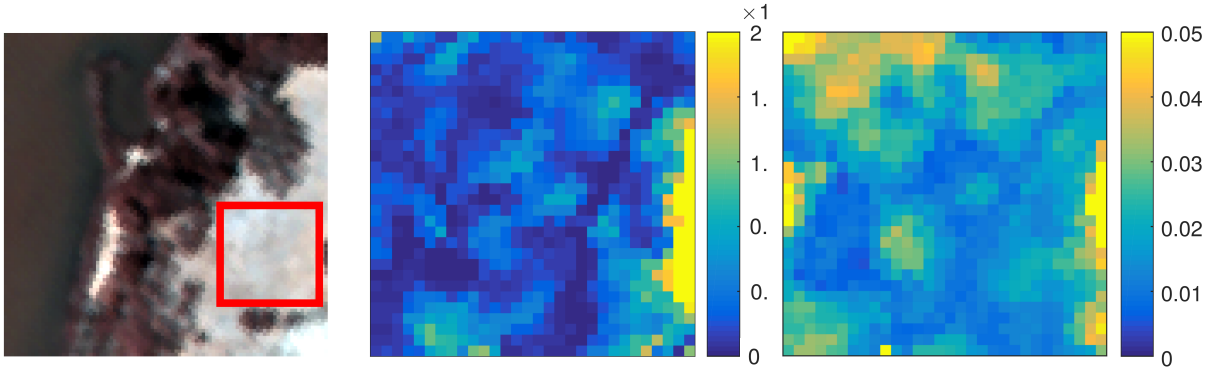
(a) (b) (c)

Figure 9 – Reflectance spectra for vegetation generated with the PROSTECT-D model [36] for varying degrees of (a) chlorophyll content, (b) equivalent water thickness, and (c) dry matter content.

Another important characteristic is that endmembers affected by intrinsic spectral variability usually display significant spatial correlation [105]. For instance, many experimental geostatistical works evaluating the spatial distribution and variability of the physico-chemical properties of the soil (e.g., sand and clay concentration, electrical conductivity, pH, compaction and available elements such as nitrogen, phosphorus and potassium) have reported significant spatial correlation/smoothness in these properties. Reports include measurements performed in Rhodes grass crop terrain [106], calcareous soils [107], rice fields [108] and tobacco plantations [109]. Besides directly impacting the spectral signature of the soil, these characteristics have been widely acknowledged to directly influence vegetation growth (e.g., they show strong correlation with crop productivity [106]), and hence their spectral signature [90, 106]. Therefore, spatial correlation in the variability is expected both in soil/terrain and in vegetation signatures. A similar behavior has also been observed in mineral spectra in the presence of spatially correlated grain size distributions and impurity concentrations [86, 87]. This implies that the variability tends to be small in small spatial neighborhoods, even though it may be large across a large scene. This fact can be leveraged to design SU algorithms since it supplies information that can be used to reduce the severe ill-posedness of the problem.

To illustrate this effect, we performed an experiment by measuring the spectral variability in a homogeneous region (composed by mostly pure pixels) of soil in the Samson image, depicted in Fig. 10-(a). We then computed the Euclidean[2] distance and the spectral angle between each soil pixel and the average spectra of all pixels in the subregion, which was used as a reference material signature. The results are depicted in Figs. 10-(b) and 10-(c), where it can be seen that the variability shows strong spatial correlation, as observed both in the Euclidean distance and spectral angle.

[2] The Euclidean distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is computed as $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{x}_i - \boldsymbol{y}_i)^2}$.



(a) (b) (c)

Figure 10 – Spatial behavior of endmember variability. (a) Soil subregion of the Samson image (highlighted by a red square). (b) Euclidean distance and (c) spectral angle between each pixel and the average spectra of the region.

INTERMEZZO: USING SPATIAL INFORMATION TO ADDRESS SPECTRAL VARIABILITY IN UNMIXING – A MULTISCALE APPROACH

As mentioned in the introduction, the SU problem becomes extremely ill-posed when spectral variability is taken into account. This makes it very important to explore what prior information is available in order to solve this problem effectively. One of the most important types of prior information used in unmixing is related to the (piecewise) spatial smoothness of the abundance maps. Smooth abundance maps are observed in a wide range of practical settings, and the pervasiveness of this hypothesis contributed make regularization approaches such as Tikhonov [110] and Total Variation [21] become a common tool in the design of SU algorithms.

These regularization strategies, however, are not the most effective in exploring information contained in observed images. Subsequent works attempted to improve performance even further by considering, e.g., non-local spatial redundancy [22]. Such strategies improve the SU results at the expense of a considerable increase in computational complexity. This sets forward two (often conflicting) goals in the design of spatial regularization strategies: better explore spatial information (i.e., improve abundance estimation performance) and reduce the computational complexity.

In the following two chapters, we explore a novel idea to design spatially regularized SU methods in both supervised (library aided) and blind SU based on a multiscale representation of the HI. Using the superpixels decomposition as a support tool, we divide the image in two spatial scales, one containing only the coarse image content, and another one which also contains the fine-scale image details. Based on this representation, we introduce spatial regularity information into the reconstructed abundance maps by using either a two-step unmixing method with an inter-scale regularization (in Chapter 3), or in a single step by controlling the energy of the estimated abundances in each of the two scales (in Chapter 4). Both algorithms are able to improve the abundance reconstruction results at a reduced computational complexity.

3 A MULTISCALE SPATIAL REGULARIZATION FOR FAST UNMIXING WITH SPECTRAL LIBRARIES

"All this time the Guard was looking at her, first through a telescope, then through a microscope, and then through an opera-glass."

Lewis Carroll, *Through the Looking Glass*

3.1 INTRODUCTION

Solving the SU problem without supervision requires estimating the EMs in (1) or in (2) directly from the observed HI $\boldsymbol{Y}$. This is not always trivial, even when no spectral variability is present in the scene. Traditional endmember extraction algorithms rely on the existence of pure pixels or on the data not being heavily mixed to obtain satisfactory results [9, 111]. An interesting strategy to circumvent such issues is to model the observed pixel as a linear combination of a large library of endmembers estimated *a priori* [112]. In this case, the number of endmembers in a given scene is usually much smaller than the size of the spectral library. Hence, the unmixing problem becomes a sparse regression problem that consists of finding a small subset of the library endmembers which best represent all the pixels in the image. This problem is often efficiently solved through the use of sparsity promoting regularizations, resulting in the so-called sparse unmixing techniques [113].

Sparse unmixing methods achieved great success in dealing with spectral variability using spectral libraries [113]. However, the use of large spectral libraries leads to the unmixing problem being ill-posed, which makes the solution very sensitive to noise. Regularization techniques have been shown to significantly improve the performance of both non-sparse [114, 45] and sparse unmixing methods [21] by exploiting the correlation between different pixels in the HI. The *Total Variation* (TV) regularization, for instance, promotes solutions that are spatially piecewise homogeneous without compromising sharp discontinuities between neighboring pixels [21].

Most effective spatial regularization techniques, however, require a massive increase in computational cost. For instance, TV [21] leads to a large non-smooth convex optimization problem, which needs to be solved using variable splitting techniques. More recently, regularization strategies exploiting nonlocal redundancy in images were also considered, leading to even larger optimization problems [22]. This is incompatible with recent demands to timely process the vast amounts of remotely sensed data required by many real world applications [24]. Such demands recently sparked significant interest on efficient unmixing strategies with online processing capability [115]. This evidences the need for fast low complexity unmixing strategies that yield state of the art performance.

In this chapter, we introduce a novel multiscale spatial regularization approach for sparse unmixing. We propose a fast *Multiscale sparse Unmixing Algorithm* (MUA) that

promotes piecewise homogeneous abundances without compromising sharp discontinuities among neighboring pixels. The proposed method uses a signal-adaptive spatial multiscale decomposition of the linear mixture model. The unmixing problem is decomposed into two different problems in distinct domains: one in an approximation scale representation constructed using segmentation or over-segmentation algorithms, and another in the original image domain. Spatial contextual information of fractional abundances is initially obtained by solving an unregularized sparse unmixing problem in the approximation scale. This information is then mapped back to the original image domain by means of an appropriately defined conjugate transformation of the multiscale decomposition. The spatial contextual information is then enforced on the solution of the original unmixing problem through a novel and computationally efficient regularization penalty. Simulation results using both synthetic and real data indicate that the proposed method outperforms TV-regularized solutions [21], while requiring a computational time comparable to that of the unregularized algorithm [113].

This chapter is organized as follows. In Section 3.2, we briefly introduce the sparse unmixing problem and present the proposed multiscale formulation. Simulation results using synthetic and real data are presented in Section 3.3. Section 3.4 presents the conclusions.

3.2 SPARSE LINEAR UNMIXING WITH A MULTISCALE SPATIAL REGULARIZATION

Let $\boldsymbol{Y} \in \mathbb{R}^{L\times N}$ denote the observed hyperspectral image with L bands and N pixels, and $\boldsymbol{M}_{\text{Lib}} \in \mathbb{R}^{L\times \widetilde{P}}$ denote a spectral library having $\widetilde{P}$ spectral signatures. Unlike a common EM matrix with P spectral signatures, $\boldsymbol{M}_{\text{Lib}}$ can have many more signatures than EM classes in the scene, leading to $\widetilde{P} \gg P$. Instead of extracting the endmembers directly from the HI $\boldsymbol{Y}$, sparse linear unmixing attempts to find an optimal subset of samples from the spectral library $\boldsymbol{M}_{\text{Lib}}$ that best represents all the mixed pixels in the image, namely,

$$\boldsymbol{Y} = \boldsymbol{M}_{\text{Lib}}\boldsymbol{A}^{\text{Lib}} + \boldsymbol{E}\,, \tag{4}$$

where $\boldsymbol{A}^{\text{Lib}} \in \mathbb{R}^{\widetilde{P}\times N}$ is the (enlarged) fractional abundance matrix, each column of which determines the composition of one image pixel as a linear combination of spectral samples from $\boldsymbol{M}_{\text{Lib}}$, and $\boldsymbol{E} \in \mathbb{R}^{L\times N}$ denotes the joint contribution of modeling errors and noise. The fractional abundance matrix $\boldsymbol{A}^{\text{Lib}}$ is frequently subject to physical constraints imposed to the model, such as the non-negativity (i.e., $\boldsymbol{A}^{\text{Lib}} \geqslant \boldsymbol{0}$) and the sum-to-one constraints (i.e., $\boldsymbol{1}^\top \boldsymbol{A}^{\text{Lib}} = \boldsymbol{1}^\top$). Since only few of the spectral signatures of the library $\boldsymbol{M}_{\text{Lib}}$ are likely to contribute to the observed spectra of each pixel, the abundance matrix $\boldsymbol{A}^{\text{Lib}}$ is usually sparse. A common approach to solve the unmixing problem is to represent it as a spatially regularized sparse regression problem [21]. These techniques, however, are computationally

very expensive. In this section, we propose a multiscale regularization procedure which introduces spatial regularity into the abundance maps at a very low computational cost.

The proposed spatially regularized unmixing scheme consists of two steps. First, we transform the original image from the original domain to an approximation (coarse) scale (denoted by $\mathcal{C}$) to extract the most relevant inter-pixel contextual information. Then, pixels at the coarse scale are unmixed independently from each other. Next, we apply a conjugate transformation to the abundance estimates obtained at the coarse scale to convert the coarse estimate back to the original image domain. This procedure yields an accurate estimate of the low-level image structures, which is then used to regularize the unmixing process applied to the original image to promote the spatial dependency between neighboring pixels.

Consider a linear operator $\boldsymbol{W} \in \mathbb{R}^{N\times S}$, $S < N$ that implements a spatial transformation of both the HI and the abundance map to the approximation domain. Then,

$$\boldsymbol{Y}_{\mathcal{C}} = \boldsymbol{Y}\boldsymbol{W}\,; \qquad \boldsymbol{A}_{\mathcal{C}}^{\text{Lib}} = \boldsymbol{A}^{\text{Lib}}\boldsymbol{W}\,, \tag{5}$$

where $\boldsymbol{Y}_{\mathcal{C}} \in \mathbb{R}^{L\times S}$ and $\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}} \in \mathbb{R}^{\tilde{P}\times S}$ are the coarse approximations of the original image $\boldsymbol{Y}$ and of the abundance matrix $\boldsymbol{A}^{\text{Lib}}$, respectively. A possible choice for $\boldsymbol{W}$ might be a wavelet transform employing the first approximation scales of the wavelet decomposition of $\boldsymbol{Y}$. However, the wavelet transform is feature-agnostic. It does not distinguish between pixels in perceptually different image regions. Its application may result in blurred image edges. Instead, we shall consider a signal-dependent transformation, that is, $\boldsymbol{W} \equiv \boldsymbol{W}(\boldsymbol{Y})$, which groups pixels into perceptually meaningful regions (not necessarily uniform), preserving image contours and leading to sharp transitions.

Multiplying (4) by $\boldsymbol{W}$ from the right, the unmixing problem can be re-cast into the approximation domain. The resulting unmixing problem is as follows:

$$\widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}} \;=\; \underset{\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}\geqslant 0}{\arg\min} \;\; \frac{1}{2}\|\boldsymbol{Y}_{\mathcal{C}} - \boldsymbol{M}_{\text{Lib}}\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}\|_F^2 + \lambda_{\mathcal{C}}\|\boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}\|_{1,1}\,. \tag{6}$$

where $\|\cdot\|_{1,1}$ is the matrix $\mathcal{L}_1$ norm, defined as $\|\boldsymbol{X}\|_{1,1} = \sum_{i,j}\big|[\boldsymbol{X}]_{i,j}\big|$.

We shall now use $\widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}}$ to regularize the original unmixing problem. To this end, we define a conjugate transform $\boldsymbol{W}^* \in \mathbb{R}^{S\times N}$ that converts images from the approximation domain $\mathcal{C}$ back to the original image domain as:

$$\widehat{\boldsymbol{A}}_{\mathcal{C}*}^{\text{Lib}} = \widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}}\boldsymbol{W}^*\,, \tag{7}$$

where $\widehat{\boldsymbol{A}}_{\mathcal{C}*}^{\text{Lib}} \in \mathbb{R}^{\tilde{P}\times N}$ is the low-resolution approximation of the abundances in the original image domain, which captures correlations between neighboring pixels. Note that transformation $\boldsymbol{W}$ is generally not invertible, that is, $\boldsymbol{W}\boldsymbol{W}^* \neq \boldsymbol{I}$.

Finally, we use the coarse abundance matrix $\widehat{\boldsymbol{A}}^{\text{Lib}}_{\mathcal{C}*}$ to regularize a sparse unmixing problem in the original image domain, where $\widehat{\boldsymbol{A}}^{\text{Lib}}$ is obtained as the solution to the following optimization problem:

$$\min_{\boldsymbol{A}^{\text{Lib}} \geqslant 0} \ \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{M}_{\text{Lib}}\boldsymbol{A}^{\text{Lib}}\|_F^2 + \lambda\|\boldsymbol{A}^{\text{Lib}}\|_{1,1} + \frac{\beta}{2}\|\widehat{\boldsymbol{A}}^{\text{Lib}}_{\mathcal{C}*} - \boldsymbol{A}^{\text{Lib}}\|_F^2 \,, \tag{8}$$

where $\beta \geqslant 0$ is a regularization parameter. This formulation requires no explicit consideration of dependencies between pairs of pixels as required by TV. This leads to a simpler optimization problem, reducing both the computational complexity and the convergence time, as will be verified in Section 3.3.

Note that both optimization problems (6) and (8) are particular cases of the following problem

$$\min_{\boldsymbol{x}} \ g_1(\boldsymbol{x}) + g_2(\boldsymbol{x}) \,, \tag{9}$$

where $g_1, g_2 : \mathbb{R}^{N_{\boldsymbol{x}}} \to \mathbb{R}_+ \bigcup \{\infty\}$ are closed, proper and convex functions. For instance, problem (8) can be written in the equivalent form (9) by selecting functions g_1 and g_2 as

$$\begin{aligned} g_1 &\equiv \frac{1}{2}\|\boldsymbol{Y} - \boldsymbol{M}_{\text{Lib}}\boldsymbol{A}^{\text{Lib}}\|_F^2 + \frac{\beta}{2}\|\widehat{\boldsymbol{A}}^{\text{Lib}}_{\mathcal{C}*} - \boldsymbol{A}^{\text{Lib}}\|_F^2 \\ g_2 &\equiv \lambda\|\boldsymbol{A}^{\text{Lib}}\|_{1,1} + \iota_+(\boldsymbol{A}^{\text{Lib}}) \,, \end{aligned} \tag{10}$$

where $\iota_+(\cdot)$ is the indicator function of the set $\mathbb{R}_+^{\tilde{P}\times N}$, that is, $\iota_+(\boldsymbol{A}^{\text{Lib}}) = 0$ if $\boldsymbol{A}^{\text{Lib}} \geqslant 0$ and $\iota_+(\boldsymbol{A}^{\text{Lib}}) = \infty$ otherwise.

The Alternating Direction Method of Multipliers (ADMM) method decomposes a problem in the form (9) into a sequence of simpler problems, which can be solved efficiently [116]. The ADMM method can then be used to solve (8), with the resulting procedure detailed in Algorithm 1 [116, 113], where soft denotes the component-wise soft thresholding operator $\text{soft}(x, \tau) = \text{sign}(x)\max\{|x| - \tau, 0\}$. Note that problem (6) can be solved in the same way by setting $\beta = 0$ and substituting $\boldsymbol{Y} \equiv \boldsymbol{Y}_{\mathcal{C}}$, $\boldsymbol{A}^{\text{Lib}} \equiv \boldsymbol{A}^{\text{Lib}}_{\mathcal{C}}$, and $\lambda \equiv \lambda_{\mathcal{C}}$ in Algorithm 1. The global algorithm of the proposed method, called *Multiscale sparse Unmixing Algorithm* (MUA), is displayed in Algorithm 2.

3.2.1 Designing the multiscale transformation

An appropriate choice of transformation $\boldsymbol{W}$ is of paramount importance for the proposed method to achieve a good reconstruction accuracy. The objectives of this transform can be summarized as 1) grouping image pixels that are spatially adjacent and semantically similar, that is, that belong to homogeneous regions, and 2) preserving image contours by not grouping pixels that belong to different image structures or features. Additionally, it must be computationally efficient.

Techniques such as the K-means have been explored for introducing regularity into the solution of inverse problems [114]. However, K-means fails to effectively explore

Algorithme 1 : ADMM method for solving (8)

Input : $\boldsymbol{Y}$, $\boldsymbol{M}_{\text{Lib}}$, parameters λ, β, and $\mu > 0$ and matrices $\boldsymbol{U}^{(0)}, \boldsymbol{V}^{(0)} \in \mathbb{R}^{\widetilde{P} \times N}$.
Output : The estimated abundance matrix $\widehat{\boldsymbol{A}}^{\text{Lib}}$.

Set $i = 0$;
while *stopping criterion is not satisfied* **do**
  $\boldsymbol{B} = \boldsymbol{M}_{\text{Lib}}^{\top}\boldsymbol{Y} + \mu(\boldsymbol{U}^{(i)} + \boldsymbol{V}^{(i)}) + \beta\widehat{\boldsymbol{A}}_{\mathcal{C}*}^{\text{Lib}}$;
  $\boldsymbol{A}^{\text{Lib}} = \left(\boldsymbol{M}_{\text{Lib}}^{\top}\boldsymbol{M}_{\text{Lib}} + (\mu + \beta)\boldsymbol{I}\right)^{-1}\boldsymbol{B}$;
  $\boldsymbol{U}^{(i+1)} = \max\{\boldsymbol{0}, \text{soft}(\boldsymbol{A}^{\text{Lib}} - \boldsymbol{V}^{(i)}, \lambda/\mu)\}$;
  $\boldsymbol{V}^{(i+1)} = \boldsymbol{V}^{(i)} - (\boldsymbol{A}^{\text{Lib}} - \boldsymbol{U}^{(i+1)})$;
  $i = i + 1$;
end
return $\widehat{\boldsymbol{A}}^{\text{Lib}} = \boldsymbol{A}^{\text{Lib}}$;

Algorithme 2 : MUA

Input : $\boldsymbol{Y}$, $\boldsymbol{M}_{\text{Lib}}$, $\boldsymbol{W}$, parameters $\lambda_{\mathcal{C}}$, λ, and β.
Output : The estimated abundance matrix $\widehat{\boldsymbol{A}}^{\text{Lib}}$.

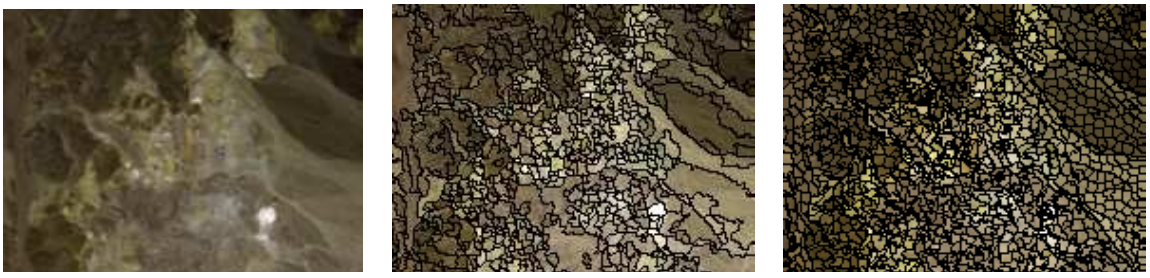
Compute $\boldsymbol{Y}_{\mathcal{C}} = \boldsymbol{Y}\boldsymbol{W}$;
Find $\widehat{\boldsymbol{A}}_{\mathcal{C}}^{\text{Lib}}$ by solving (6) using Algorithm 1 with $\beta = 0$, $\lambda \equiv \lambda_{\mathcal{C}}$, $\boldsymbol{Y} \equiv \boldsymbol{Y}_{\mathcal{C}}$ and $\boldsymbol{A}^{\text{Lib}} \equiv \boldsymbol{A}_{\mathcal{C}}^{\text{Lib}}$;
Compute $\widehat{\boldsymbol{A}}_{\mathcal{C}*}^{\text{Lib}}$ using (7);
Find $\widehat{\boldsymbol{A}}^{\text{Lib}}$ by solving (8) using Algorithm 1;
return $\widehat{\boldsymbol{A}}^{\text{Lib}}$;

local spatial regularity of the image, which is an important contextual information of HIs. Moreover, spectrally similar pixels might share different abundance attributes. Hence, spectral-only methods such as the K-means tend to group pixels that are semantically distinct, especially in noisy scenarios. Therefore, both spatial and spectral information should be explored to obtain good results.

To explore spatial information while grouping semantically similar pixels accounting for image discontinuities, we propose to construct $\boldsymbol{W}$ using image segmentation or over-segmentation algorithms [117, 118]. Image segmentation methods decompose the observed image into a set of contiguous homogeneous regions with contextually similar spatial information, typically consisting of objects which are separated by image borders [117]. Image segmentation often creates groups of pixels of heterogeneous sizes, corresponding to both small and large objects in the same image. Although this allows one to represent large regions with homogeneous abundance characteristics without compromising smaller objects, it can lead to grouping pixels that share different abundance characteristics (even if spectrally similar). As an alternative to circumvent this issue, we also explore over-segmentation techniques, which attempt to divide the observed image into a larger number of regions with relatively homogeneous sizes [118]. Although over-segmentation methods partition large objects into many smaller segments, they provide an increased

ability to adequately represent image borders and reduce the chance of grouping pixels with different contextual information. Superpixel algorithms are a popular and efficient technique for image over-segmentation [118].

We choose the transformation $\boldsymbol{W}$ to be an (over)-segmentation of the image. More precisely, $\boldsymbol{YW}$ computes an (over)-segmentation of the image $\boldsymbol{Y}$, and returns the average of all pixels inside each segmented region or superpixel. Note that the resulting pixels do not lie on a uniform sampling grid. The conjugate transform, $\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^*$, takes each segment in $\boldsymbol{Y}_{\mathcal{C}}$ and attributes its value to all pixels of the uniform image sampling grid that lie inside the corresponding region. The successive application of both transforms, $\boldsymbol{WW}^*$ effectively consists in averaging all pixels inside each segment of the input image. The decomposition of the Cuprite image using a segmentation and an over-segmentation algorithm is illustrated in Fig. 11.



HS Image Segmentation Over-Segmentation

Figure 11 – Coarse-scale decomposition of a section of the Cuprite image for bands 50, 80 and 100 using the segmentation algorithm in [117] and the over-segmentation algorithm in [118], with 950 and 2000 segments, respectively.

3.3 RESULTS

We compare the performances of the proposed MUA, the Total Variation (SUnSAL-TV), the spatially unregularized (SUnSAL) and the S^{2}WSU algorithms [113, 21, 119], both in terms of reconstruction error and computational complexity. The selection of these algorithms comes naturally since MUA, SUnSAL and SUnSAL-TV share the same sparse regression formulation, and S^{2}WSU is considered a state-of-the-art algorithm for library-based sparse unmixing. For the proposed method, we compare two choices for the transformation $\boldsymbol{W}$: 1) a binary partition tree based segmentation algorithm (BPT) [117], and 2) the simple linear iterative clustering (SLIC) over-segmentation method [118]. Finally, we consider also the solution using the K-means algorithm, which is not effective at taking local spatial information into account[1].

[1] SLIC and K-means were implemented using the Euclidean distance between reflectance vectors (HI pixels).

We considered a synthetic library $\boldsymbol{M}'_{\text{Lib}} \in \mathbb{R}^{224\times 240}$ generated by selecting a subset of 240 materials from the USGS library such that the angle between any pair of spectral signatures was at least 4.44 degrees.

Table 2 – SRE results for unmixing data cubes DC1 and DC2.

DC1 data cube						
SNR	SUnSAL	SUnSAL-TV	S^{2}WSU	MUA$_{\text{K-means}}$	MUA$_{\text{BPT}}$	MUA$_{\text{SLIC}}$
20 dB	4.54 dB	9.42 dB	7.70 dB	9.96 dB	**13.39** dB	11.35 dB
30 dB	8.91 dB	14.44 dB	15.49 dB	14.02 dB	**18.26** dB	15.73 dB
DC2 data cube						
SNR	SUnSAL	SUnSAL-TV	S^{2}WSU	MUA$_{\text{K-means}}$	MUA$_{\text{BPT}}$	MUA$_{\text{SLIC}}$
20 dB	4.27 dB	11.61 dB	9.39 dB	12.69 dB	14.08 dB	**14.88** dB
30 dB	10.48 dB	17.97 dB	**21.72** dB	17.42 dB	16.92 dB	18.46 dB

3.3.1 Simulation results using synthetic data sets

For the simulations presented in this section two spatially correlated synthetic data cubes DC1 and DC2 were built using 5 and 9 endmembers, respectively, selected from library $\boldsymbol{M}'_{\text{Lib}}$. DC1 has 75×75 pixels and its abundance map is composed of square regions distributed uniformly over a background in five rows. Data cube DC2 has 100×100 pixels and its abundance maps were sampled according to a Dirichlet distribution centered at a Gaussian random field, leading to piecewise smooth maps that also have steep transitions. For both datacubes, the generated HIs were contaminated by white Gaussian noise, with signal-to-noise ratios (SNR) of 20 and 30 dB. The quality of the reconstruction of the spectral mixtures was evaluated using the signal to reconstruction error, defined as $\text{SRE} = 10\log_{10}(\text{E}\{\|\boldsymbol{A}^{\text{Lib}}\|_F^2\}/\text{E}\{\|\boldsymbol{A}^{\text{Lib}} - \widehat{\boldsymbol{A}}^{\text{Lib}}\|_F^2\})$ [21].

To find the optimal parameters for the selected algorithms we performed a grid search for each dataset, and the parameters leading to the best SRE results for each method were selected. For the MUA method, the parameter search occurred in the intervals $\lambda_{\mathcal{C}} \in [0.0001,\ 0.05]$, $\lambda \in [0.001,\ 0.1]$ and $\beta \in [0.007,\ 30]$, while the cluster sizes were selected among the integer values $\sqrt{N/S} \in \{3, \ldots, 15\}$. For the SUnSAL, SUnSAL-TV and S^{2}WSU algorithms, the parameter ranges were selected according to those reported in the original work in [21]. The SRE achieved by the SUnSAL, SUnSAL-TV, S^{2}WSU, K-means and MUA are shown in Table 2 for both SNR values. Samples of the reconstructed abundance maps for both data cubes and SNRs are shown in Figs. 12 and 13 for a qualitative comparison.

The computational complexity of the algorithms was evaluated through their execution times. SUnSAL, SUnSAL-TV, S^{2}WSU, BPT and SLIC were implemented using the codes made available by the authors. The algorithms were implemented in Matlab

on a desktop computer equipped with an Intel Core I7 processor with 4.2Ghz and 16Gb RAM. The results are shown in Table 3.

Table 3 – Average Execution time (in seconds) of each algorithm

	SUnSAL	SUnSAL-TV	S^2WSU	$MUA_{K\text{-means}}$	MUA_{BPT}	MUA_{SLIC}
DC1	2.57	58.24	24.21	2.88	4.19	2.66
DC2	4.24	92.1	42.41	3.69	4.94	4.04
Real Image	184.8	1145.8	469.5	84.9	77.1	101.5

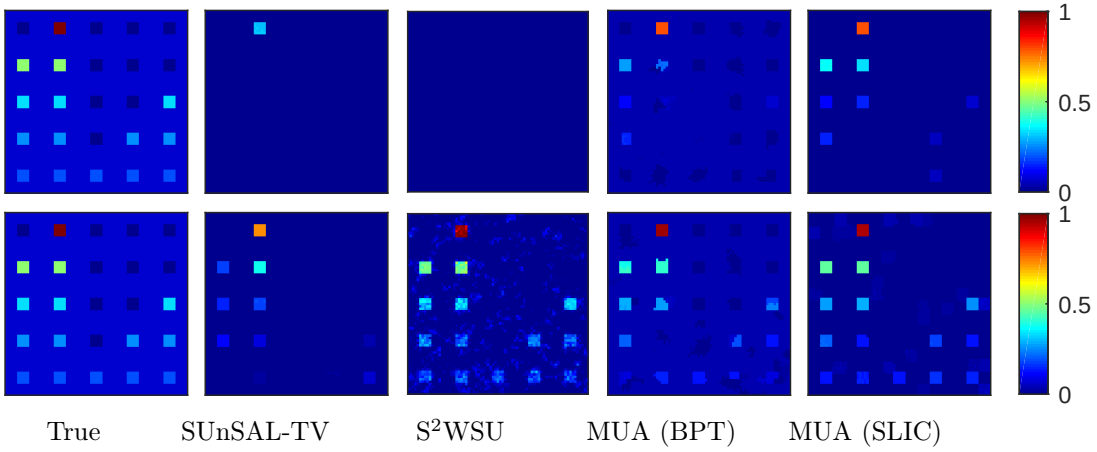


Figure 12 – Abundance maps estimated by the different unmixing methods for the 2nd endmember of data cube DC1. From top to bottom: SNR of 20 and 30 dB.

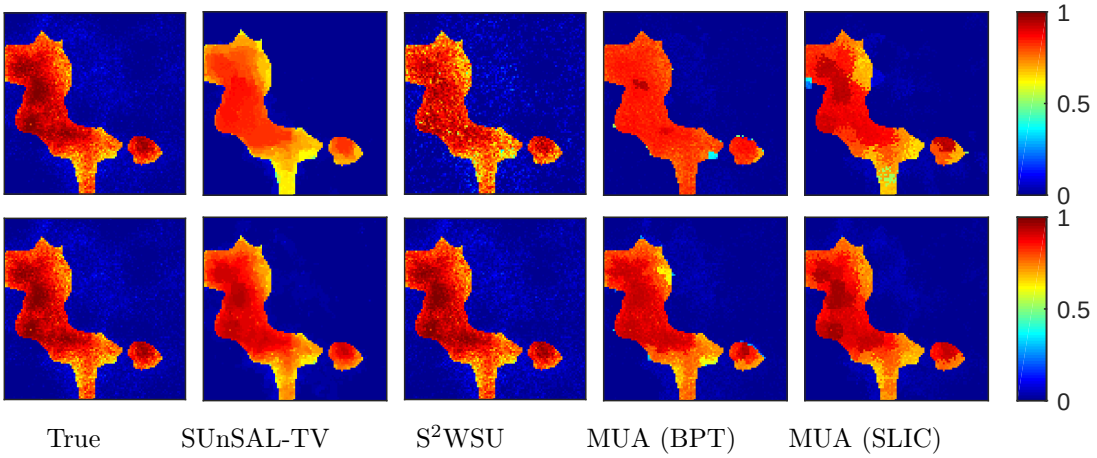


Figure 13 – Abundance maps estimated by the different unmixing methods for the 1st endmember of data cube DC2. From top to bottom: SNR of 20 and 30 dB.

3.3.1.1 Discussion

It can be seen from Table 2 that the proposed algorithm can provide significantly better performance than the SUnSAL-TV algorithm for both data cubes. The

BPT segmentation-based transformation provided a variable performance, yielding very good results for DC1, but a performance closer to SUnSAL-TV for DC2, especially for SNR=30 dB. This indicates a considerable sensitivity to the image content. The results obtained using SLIC, on the other hand, indicate a more regular performance, with significantly better results than SunSAL-TV for both data cubes. Although the S^2WSU presented the best SRE result for DC2 with SNR of 30 dB, the method is very sensitive to variations of the noise level, as can be seen for both datasets. Finally, we note that a regularization based on the K-means algorithm performed only similarly to the SUnSAL-TV method, and significantly worse than the proposed transformations.

Figs. 12 and 13 show samples of the abundance maps of data cubes DC1 and DC2 estimated by the SUnSAL-TV, S^2WSU and MUA algorithms using BPT and SLIC transformation, which provided the best quantitative performance, except for the DC2 at 30 dB where the S^2WSU produced a comparable map. However, the performance degradation of the S^2WSU is clear when the SNR is decreased. The results of the MUA algorithm were significantly better than those of the SUnSAL-TV algorithm. This difference is most noticeable for an SNR of 20 dB, where the resulting abundance maps are much closer to the ground truth than those estimated by the SUnSAL-TV.

In terms of computational cost, MUA performed significantly better than SUnSAL-TV, with execution times comparable to those of SUnSAL algorithm, and, on average, 19 and 10 times smaller than those of SUnSAL-TV and S^2WSU respectively. These results illustrate the effectiveness of the proposed regularization method both in terms of quality and computational cost.

3.3.2 Simulation results using real image

In this experiment, we consider a well-known region of the Cuprite data set with 250×191 pixels. The spectral library $\boldsymbol{M}_{\text{Lib}} \in \mathbb{R}^{188\times498}$ was built using the spectral signatures in the USGS library after removing water absorption and low SNR spectral bands, resulting in 188 bands. The parameters of the algorithms were selected empirically for MUA, and set identically to those reported in [21] for SUnSAL and SUnSAL-TV, and for the S^2WSU we used $\lambda_{\text{swsp}} = 7 \times 10^{-5}$. Since the true abundance maps are unavailable for this HI, we compare the fractional abundance maps of three dominant materials (Alunite, Buddingtonite, and Chalcedony) estimated using the three algorithms. The results are shown in Fig. 14.

Although the unmixing results for SUnSAL-TV, S^2WSU, and MUA were similar, it can be observed that the TV regularization tends to yield an over-smooth visual effect. This is not observed in the results using S^2WSU and MUA (especially for the over-segmentation transformation) which produce spatially consistent abundance maps without compromising the fine variability and the intricate structures in the image. These results again indicate the effectiveness of the proposed spatial regularization. The computational times are shown

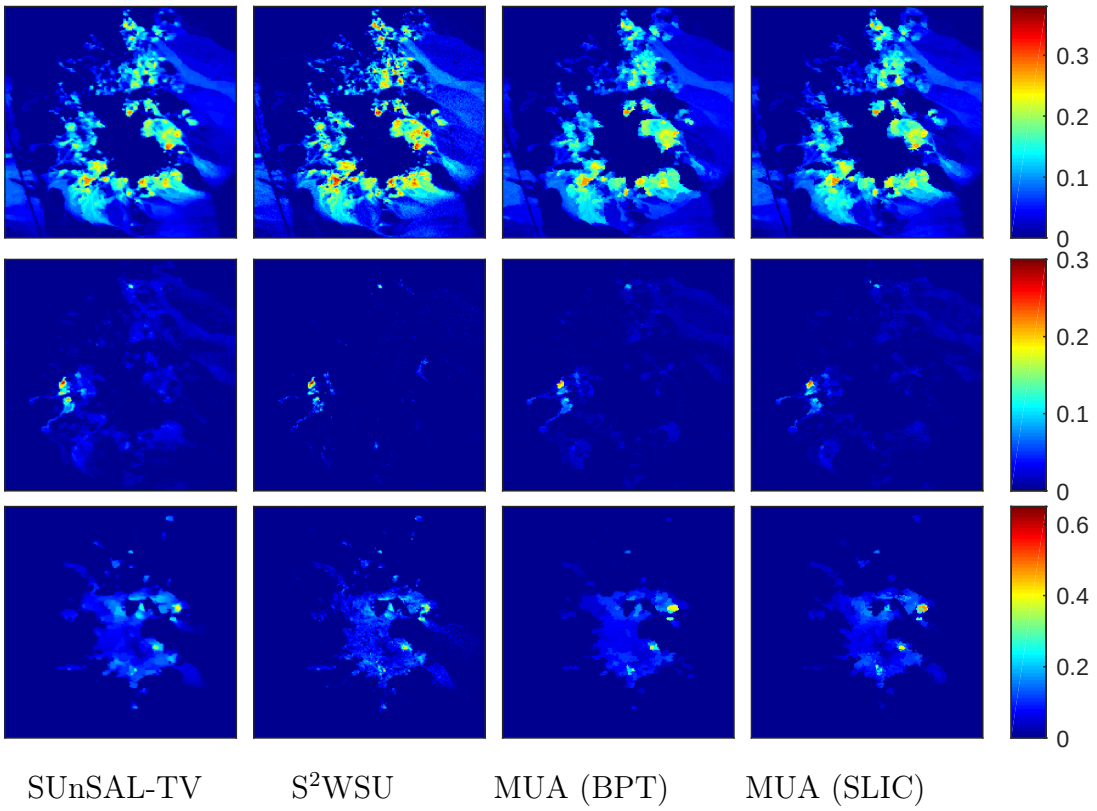


Figure 14 – Fractional abundance maps estimated for the Cuprite image. From top to bottom: Alunite, Buddingtonite, and Chalcedony.

in Table 3, and illustrate again the considerably lower complexity of MUA when compared to SUnSAL-TV and S^{2}WSU. It also runs significantly faster than the SUnSAL algorithm due to the faster convergence rate achieved with the use of proposed regularization.

3.4 CONCLUSIONS

In this chapter, we presented a novel multiscale methodology to introduce spatial information in sparse SU problems. It decomposes the spatially regularized unmixing problem into two simple, low-cost problems in different image domains. Two multiscale domain transformations were proposed based on segmentation and over-segmentation methods, which allow an effective capture of spatial and spectral contextual information at a reasonable computational cost. Simulation results using both synthetic and real data showed that the proposed method outperforms state-of-the-art TV-based sparse SU algorithms. Moreover, it requires execution times that are an order of magnitude lower than the TV-based solution, and comparable to or even smaller than those of unregularized methods.

4 A DATA DEPENDENT MULTISCALE MODEL FOR SPECTRAL UNMIXING WITH SPECTRAL VARIABILITY

4.1 INTRODUCTION

Using parametric models to represent EM spectra is one of the main approaches to deal with spectral variability in SU, due to good trade-off they offer between flexibility, performance and computational complexity. The Perturbed LMM model (PLMM) [20] introduces an additive perturbation to a reference endmember matrix. Such perturbation matrix then needs to be estimated jointly with the abundances. Though the perturbation matrix can model arbitrary endmember variations, it lacks physical motivation. The Extended Linear Mixing Model (ELMM) proposed in [19] increased the flexibility of the LMM by associating a pixel-dependent multiplicative term to each endmember. This generalization can efficiently model changes in the observed reflectances due to illumination, an important effect [19]. This model is physically motivated, with the advantage of estimating a variability parameter vector of much lower dimension when compared with the additive perturbation matrix in PLMM. Although the ELMM performs well in situations where spectral variability is mainly caused by illumination variations, it lacks flexibility when the endmembers are subject to more complex spectral distortions. This motivates the development of more flexible EM models, a topic that will be addressed in detail in Chapters 5, 6 and 7.

Though the above described models were shown to be capable of modeling endmember variability effects with good accuracy, their use in SU leads to severely ill-posed inverse problems, which require sound regularization strategies to yield meaningful solutions. One way to mitigate this ill-posedness is to explore spatial correlations found in typical abundance [120] and EM variability [19] maps. For instance, spatial information has been employed both for endmember extraction [121, 122] and for regularization in linear [123], nonlinear [124], Bayesian [120, 125, 126] and sparse [21] unmixing strategies. Total variation (TV) deserves special mention as a spatial regularization approach that promotes spatially piecewise homogeneous solutions without compromising sharp discontinuities between neighboring pixels. This property is important to handle the type of spatial correlation found in many hyperspectral unmixing applications [110, 127].

Although important to mitigate the ill-posedness of the inverse problem, the use of spatial regularization in spectral-variability-aware SU introduces interdependencies among abundance solutions for different image pixels. This in turn leads to intricate, large scale and computationally demanding optimization problems. Even though some approaches have been investigated to accelerate the minimization of convex TV-regularized functionals [128, 129], this is still a computationally demanding operation which, in the context of SU, have been primarily addressed using variable splitting (e.g. ADMM) techniques [20, 19, 34]. Such complexity is usually incompatible with recent demands for

timely processing of vast amounts of remotely sensed data required by many modern real world applications [25, 24]. Thus, it is desirable to search for faster and lower complexity strategies that yield comparable unmixing performances.

In Chapter 3 (related to the publication [32]) a multiscale spatial regularization approach was proposed for sparse unmixing. The method uses a signal-adaptive spatial multiscale decomposition to break the unmixing problem down into two simpler problems, one in an approximation domain and another in the original domain. The spatial contextual information is obtained by solving an unregularized unmixing problem in the approximation domain. This information is then mapped back to the original image domain and used to regularize the original unmixing problem. The multiscale approach resulted in a fast algorithm that outperformed competing methods, both in accuracy and in execution time, and promoted piecewise homogeneity in the estimated abundances without compromising sharp discontinuities among neighboring pixels.

Motivated by the good results obtained in Chapter 3, we propose in this paper a novel data dependent multiscale mixture model for use in hyperspectral unmixing accounting for spectral variability of the endmembers. The new model uses a multiscale transform to incorporate spatial contextual information into the abundances of a generic mixing model considering spectral variability. The unmixing problem is then formulated as the minimization of a cost function in which a parametric endmember model (e.g. ELMM or PLMM) is used to constraint and reduce the ill-posedness of the endmember estimation problem. However, the dimensionality of this problem is still very high since the spatial regularization ties the abundance solutions of all pixels together. Nevertheless, under a few mild assumptions we are able to devise a computationally efficient solution to the abundance estimation problem that can also be computed separately in the two domains.

The contributions in this chapter include:

1. The proposal of a new regularization strategy based on a multiscale representation of the hyperspectral images and abundance maps. This regularization is significantly different from and improves the one in Chapter 3. While in Chapter 3 the static/fixed endmember matrix for all pixels allowed the easy separation of the abundance estimation process in different domains, the same approach is not applicable to the present case since the variability of the endmember matrix ties the abundances in the approximation and original image domains.

2. A new approximate multiscale decomposition of the generic mixing model considering spectral variability. The new decomposition leads to a separable abundance estimation problem that allows a simple and efficient solution without significantly sacrificing accuracy. Moreover, the solution can be determined in parallel for all image pixels.

When compared with approaches that rely on standard spatial regularization strate-

gies and on variable splitting techniques such as ADMM, the proposed strategy leads to a faster iterative algorithm that at each iteration solves the abundance problem only once in each domain. The new algorithm is named *Multiscale Unmixing Algorithm Accounting for Spectral Variability* (MUA-SV). Simulation results clearly show the advantage of the proposed algorithm, both in accuracy and in execution time, over the competing methods.

The chapter is organized as follows. Section 4.2 briefly reviews the linear mixing models and its variants accounting for spectral variability. In Section 4.3, we present the proposed multiscale formulation for the mixture model. In Section 4.4 we formulate the unmixing problem using the multiscale approach. The optimization of the resulting cost function is presented in Section 4.5. In Section 4.6, we propose an approximate formulation of the abundance estimation problem that leads to a simple and efficient solution. The resulting MUA-SV algorithm is detailed in Section 4.7. Simulation results with synthetic and real data are presented in Section 4.8. Section 4.9 presents the conclusions.

4.2 LINEAR MIXING MODELS CONSIDERING SPECTRAL VARIABILITY

In the Introduction, the most general form of the LMM considering spectral variability was defined by generalizing the LMM to allow a different EM matrix for each pixel. Recalling equation (2), this model represents the n-th pixel as

$$\boldsymbol{y}_n = \boldsymbol{M}_n \boldsymbol{a}_n + \boldsymbol{e}_n \,, \qquad n = 1, \ldots, N \tag{11}$$

where $\boldsymbol{M}_n \in \mathbb{R}^{L \times P}$ is the n-th pixel endmember matrix. This model can also be written for all pixels as

$$\boldsymbol{Y} = \left[\boldsymbol{M}_1 \boldsymbol{a}_1, \ldots, \boldsymbol{M}_N \boldsymbol{a}_N\right] + \boldsymbol{E} \,. \tag{12}$$

Different models have been recently proposed to represent endmember variability as a parametric function of some reference endmember spectral signatures [20, 19, 34, 38]. These models are generically denoted by

$$\boldsymbol{M}_n = f(\boldsymbol{M}_0, \boldsymbol{\theta}_n) \,, \tag{13}$$

where f is a parametric function, $\boldsymbol{M}_0 \in \mathbb{R}^{L \times P}$ is a fixed reference endmember matrix and $\boldsymbol{\theta}_n$ is a vector of parameters used to describe the endmember signatures for the n-th pixel. Although different forms have been proposed for f, two notable examples are given by the Perturbed Linear Mixing Model (PLMM) [20] and the Extended Linear Mixing Model (ELMM) [19].

The PLMM proposed in [20] models $\boldsymbol{M}_n$ as a fixed matrix $\boldsymbol{M}_0$ plus a pixel-dependent variation matrix that can accommodate generic spatial variations. Mathematically,

$$\boldsymbol{y}_n = \big(\boldsymbol{M}_0 + \boldsymbol{d}\boldsymbol{M}_n\big)\boldsymbol{a}_n + \boldsymbol{e}_n \,, \tag{14}$$

where the parameters of this model are related to those of (13) by $\boldsymbol{\theta}_n \equiv \text{vec}(\boldsymbol{dM}_n)$, where vec$(\cdot)$ is the vectorization operator. This model is not physically motivated. Hence, in most cases all elements of $\boldsymbol{dM}_n$ must be included as independent variables in the solution of the ill-posed unmixing problem, making the inverse problem hard to solve. This limitation motivated the development of simpler, physically motivated variability models.

The ELMM is a simpler model proposed in [19]. It incorporates a multiplicative diagonal matrix to LMM, which maintains the directional information of the reference endmembers, but allows them to be independently scaled. The ELMM is expressed as

$$\boldsymbol{y}_n = \boldsymbol{M}_0 \ \text{diag}(\boldsymbol{\phi}_n)\boldsymbol{a}_n + \boldsymbol{e}_n \,, \tag{15}$$

where $\boldsymbol{\phi}_n \in \mathbb{R}^P$ is a vector containing a (positive) scaling factor for each endmember, which is related to the parameters of (13) by $\boldsymbol{\theta}_n \equiv \boldsymbol{\phi}_n$, and diag$(\boldsymbol{x})$ denotes a diagonal matrix whose diagonal elements are given by the elements of vector $\boldsymbol{x}$. This model is a particular case of (11) that can model typical endmember variations, such as those caused by illumination variability due to the topography of the scene [19]. The optimization problem that needs to be solved using model (15) is much less ill-posed than that generated using model (14) due to the reduced number of parameters to be estimated. This simplicity is obtained at the price of reduced generality.

For both the PLMM and ELMM, the problem of estimating the fractional abundances and the spectral signatures of the endmembers was cast as a large scale, non-convex inverse problem, which was solved using variable splitting procedures [20, 19]. The computational cost of these solutions is very high, making them unsuited for processing large amounts of data. Furthermore, the introduction of *a priori* information about the spatial regularity of the abundance maps, which is essential to reduce the ill-posedness of the inverse problem, results in a optimization problem that is not separable per pixel. This significantly increases the computational cost of the solution. Considering this limitation of the models described above, it is of interest to develop new mixture models that combine the generality of the endmember variability patterns that can be considered with the possibility of an efficient solution of the associated inverse problem. In the next section, we introduce a new mixture model that represents separately the image components at different scales using a data-dependent transformation learned from the observed hyperspectral image $\boldsymbol{Y}$. This new multiscale representation can be employed to solve the unmixing problem with any parametric model to represent spectral variability that fits the form (13). The use of this new model results in a method that is able to provide more accurate solutions at a much lower computational cost than the existing methods.

4.3 A MULTISCALE SPATIAL MIXTURE MODEL

To constrain the set of possible solutions, we propose to separately represent the mixture process in two distinct image scales, namely, the coarse scale containing rough

spatial structures, and the fine spatial scale containing small structures and details. By doing so, the conditions for spatial smoothness can be imposed on the relevant parameters in the much simpler coarse scale, and then be translated into the fine scale for further processing.

Note that the decomposition applied in this chapter have an important difference from the one used in Chapter 3. The decomposition in Chapter 3 represented the HI and the abundance maps using a coarse spatial scale and the original image domain, the latter of which contains both coarse and fine image structures. Here, on the other hand, we decompose the image into a coarse and a *detail* spatial scales, the latter containing *only* small structures and details.

As in Chapter 3, we consider a transformation $\boldsymbol{W} \in \mathbb{R}^{N\times S}$ based on relevant contextual inter-pixel information present in the observed image $\boldsymbol{Y}$ to be applied to both $\boldsymbol{Y}$ and $\boldsymbol{A}$ to unveil the coarse image structures. Recalling (5), the transformed matrices are given by

$$\boldsymbol{Y}_{\mathcal{C}} = \boldsymbol{Y}\boldsymbol{W}\,; \quad \boldsymbol{A}_{\mathcal{C}} = \boldsymbol{A}\boldsymbol{W}\,, \tag{16}$$

where $\boldsymbol{Y}_{\mathcal{C}} = [\boldsymbol{y}_{\mathcal{C}_1}, \ldots, \boldsymbol{y}_{\mathcal{C}_S}] \in \mathbb{R}^{L\times S}$ and $\boldsymbol{A}_{\mathcal{C}} = [\boldsymbol{a}_{\mathcal{C}_1}, \ldots, \boldsymbol{a}_{\mathcal{C}_S}] \in \mathbb{R}^{P\times S}$ with $S \ll N$ are, respectively, the hyperspectral image and the abundance matrix in the coarse approximation scale, denoted by $\mathcal{C}$.

The spatial details of the image are represented in the detail scale, denoted by $\mathcal{D}$, which is obtained by computing the complement to the transformation $\boldsymbol{W}$. Mathematically,

$$\boldsymbol{Y}_{\mathcal{D}} = \boldsymbol{Y}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)\,; \quad \boldsymbol{A}_{\mathcal{D}} = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)\,, \tag{17}$$

where $\boldsymbol{Y}_{\mathcal{D}} = [\boldsymbol{y}_{\mathcal{D}_1}, \ldots, \boldsymbol{y}_{\mathcal{D}_N}] \in \mathbb{R}^{L\times N}$ and $\boldsymbol{A}_{\mathcal{D}} = [\boldsymbol{a}_{\mathcal{D}_1}, \ldots, \boldsymbol{a}_{\mathcal{D}_N}] \in \mathbb{R}^{P\times N}$ are the input image and the abundance matrix in the detail scale. Matrix $\boldsymbol{W}^* \in \mathbb{R}^{S\times N}$ is a conjugate transformation to $\boldsymbol{W}$, and takes the images from the coarse domain $\mathcal{C}$ back to the original image domain. $\boldsymbol{Y}_{\mathcal{D}}$ and $\boldsymbol{A}_{\mathcal{D}}$ contain the fine scale details of $\boldsymbol{Y}$ and $\boldsymbol{A}$ in the original image domain. The transformation $\boldsymbol{W}$ captures the spatial correlation of the input image, whereas its complement $(\boldsymbol{I}-\boldsymbol{W}\boldsymbol{W}^*)$ captures existing fine spatial variations. This way it is possible to introduce spatial correlation into the abundance map solutions by separately controlling the regularization strength in each of the scales $\mathcal{C}$ and $\mathcal{D}$. This is computationally much simpler than to use more complex penalties. By imposing a smaller penalty in the coarse scale $\mathcal{C}$ and a larger penalty in the details scale $\mathcal{D}$, we effectively favor smooth solutions to the optimization problem.

We can define a composite transformation as

$$\widetilde{\boldsymbol{W}} = [\boldsymbol{W} \quad \boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*]\,, \tag{18}$$

which decomposes the input image into the coarse approximation $\mathcal{C}$ and its complement $\mathcal{D}$. Note that the transformation is invertible, with a right inverse given by

$$\widetilde{\boldsymbol{W}}^{\dagger} = [\boldsymbol{W}^* \quad \boldsymbol{I}]^{\top}\,. \tag{19}$$

Multiplying $\boldsymbol{Y}$ from the right by $\widetilde{\boldsymbol{W}}$ and considering the generic mixing model for all pixels given in (12) yields $\boldsymbol{Y}\widetilde{\boldsymbol{W}} = \big[\boldsymbol{Y}_{\mathcal{C}}\ \boldsymbol{Y}_{\mathcal{D}}\big]$, with

$$\begin{aligned}\boldsymbol{Y}_{\mathcal{C}} &= \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big]\boldsymbol{W} + \boldsymbol{E}_{\mathcal{C}}\\ \boldsymbol{Y}_{\mathcal{D}} &= \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big](\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*) + \boldsymbol{E}_{\mathcal{D}}\end{aligned} \tag{20}$$

where $\boldsymbol{E}_{\mathcal{C}} = \boldsymbol{E}\boldsymbol{W}$ and $\boldsymbol{E}_{\mathcal{D}} = \boldsymbol{E}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)$ represent the additive noise in the coarse and detail scales, respectively.

The choice of the multiscale transformation $\boldsymbol{W}$ is important for the proposed methodology to achieve a good reconstruction accuracy. Based on the results obtained in Chapter 3, we consider the superpixel decomposition of the image for the transformation $\boldsymbol{W}$, due to its ability to capture fine spatial details. We now recapitulate a few important characteristics of the transformation $\boldsymbol{W}$ and of its conjugate $\boldsymbol{W}^*$ which will be useful in the remainder of this chapter.

The decomposition $\boldsymbol{Y}\boldsymbol{W}$ of the image $\boldsymbol{Y}$ returns a set of superpixels. The value of each superpixel is equal to the average of all original pixel values inside that superpixel region. The conjugate transform, $\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^*$, takes each superpixel in $\boldsymbol{Y}_{\mathcal{C}}$ and attributes its value to all pixels of the uniform image sampling grid that lie inside its corresponding superpixel region. The successive application of both transforms, $\boldsymbol{W}\boldsymbol{W}^*$ effectively consists in averaging all pixels inside each superpixel of the input image.

4.4 THE UNMIXING PROBLEM

The spectral unmixing problem with spectral variability can be formulated as the minimization of the cost function

$$\begin{aligned}\mathcal{J}(\mathscr{M},\boldsymbol{\Theta},\boldsymbol{A}) = \frac{1}{2}\big\|\boldsymbol{Y} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big]\big\|_F^2 + \lambda_A\mathcal{R}(\boldsymbol{A})\\ + \frac{\lambda_M}{2}\sum_{n=1}^{N}\|\boldsymbol{M}_n - f(\boldsymbol{M}_0,\boldsymbol{\theta}_n)\|_F^2 + \lambda_{\boldsymbol{\Theta}}\mathcal{R}(\boldsymbol{\Theta})\\ \text{subject to } \boldsymbol{A} \geqslant \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{A} = \boldsymbol{1}^\top,\\ \boldsymbol{M}_n \geqslant \boldsymbol{0},\ n = 1,\ldots,N.\end{aligned} \tag{21}$$

where $\mathscr{M}$ is an $L \times P \times N$ tensor containing the endmember matrices, with entries given by $[\mathscr{M}]_{:,:,n} = \boldsymbol{M}_n$ and $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N]$ is a matrix containing the parameter vectors of the variability model for all pixels. Note that the generic parametric endmember model $\boldsymbol{M}_n = f(\boldsymbol{M}_0,\boldsymbol{\theta}_n)$ of (13) was included in the cost function (21) in the form of an additive constraint. This decouples the problem of estimating the abundances from that of estimating the parametric endmember model, allowing the application of the multiscale formulation to other endmember models without loss of generality. Furthermore, this also gives more flexibility to the unmixing solution since the parameter λ_M can be adjusted

to either allow matrices $\boldsymbol{M}_n$ to vary more freely or to strictly enforce the endmember variability model (13).

The regularization functionals $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\boldsymbol{\Theta})$ incorporate prior information about the spatial smoothness of the abundance and about the parameters of the spectral variability model. The abundance maps constraint introduces spatial regularity indirectly through the transformation $\boldsymbol{W}$. The constraint is given by

$$\begin{aligned}\mathcal{R}(\boldsymbol{A}) &= \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{W}\|_F^2 + \frac{1}{2}\|\boldsymbol{A}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)\|_F^2 \\ &= \frac{\rho}{2}\|\boldsymbol{A}_{\mathcal{C}}\|_F^2 + \frac{1}{2}\|\boldsymbol{A}_{\mathcal{D}}\|_F^2 \end{aligned} \tag{22}$$

and consists of a quadratic penalization of the multiscale representation of the abundance maps, applied separately to the coarse and detail scales $\mathcal{C}$ and $\mathcal{D}$. Parameter ρ allows the control of the relative weights of each scale in the abundance penalty. For instance, piecewise smooth abundance solutions to the optimization problem can be promoted by imposing a smaller penalty in the coarse scale $\mathcal{C}$ and a larger penalty in the details scale $\mathcal{D}$.

The constraint $\mathcal{R}(\boldsymbol{\Theta})$ is selected according to the endmember variability model that is used, and might encode information such as the amount of spectral variability in a scene or spatial correlation in the variables $\boldsymbol{\theta}_n$. The parameters λ_A and $\lambda_{\boldsymbol{\Theta}}$ control the balance between the different terms in the cost function.

In the following, we employ the ELMM model due to its parsimony and underlying physical motivation [19]. This results in the following concrete forms for f and $\boldsymbol{\Theta}$:

$$\begin{aligned} f(\boldsymbol{M}_0, \boldsymbol{\theta}_n) &\equiv \boldsymbol{M}_0 \ \mathrm{diag}(\boldsymbol{\phi}_n) \\ \boldsymbol{\Theta} &\equiv \boldsymbol{\Phi}\,, \end{aligned} \tag{23}$$

where $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N]$ is a matrix whose n-th column contains scaling factors $\boldsymbol{\phi}_n$ of the ELMM model (15). The scaling maps constraint $\mathcal{R}(\boldsymbol{\Theta})$ is selected to introduce spatial smoothness to the endmember scaling factors, and is given by

$$\begin{aligned}\mathcal{R}(\boldsymbol{\Theta}) &\equiv \mathcal{R}(\boldsymbol{\Phi}) \\ &= \|\mathcal{H}_h(\boldsymbol{\Phi})\|_F^2 + \|\mathcal{H}_v(\boldsymbol{\Phi})\|_F^2\,, \end{aligned} \tag{24}$$

where $\mathcal{H}_h$ and $\mathcal{H}_v$ are linear operators that compute the vertical and horizontal gradients of a bi-dimensional signal, acting separately for each material. In the following, we make the variable substitutions outlined in (23) and (24), which turns the cost function in (21) into $\mathcal{J}(\mathscr{M}, \boldsymbol{\Phi}, \boldsymbol{A})$.

The estimated abundance maps, endmember matrices and scaling factors can be obtained by minimizing (21) with respect to (w.r.t.) these variables, resulting in the following optimization problem

$$\widehat{\mathscr{M}}, \widehat{\boldsymbol{\Phi}}, \widehat{\boldsymbol{A}} = \underset{\mathscr{M}, \boldsymbol{\Phi}, \boldsymbol{A}}{\arg\min} \ \ \mathcal{J}(\mathscr{M}, \boldsymbol{\Phi}, \boldsymbol{A})\,. \tag{25}$$

This problem is non-convex and hard to solve directly due to the interdependence between $\mathscr{M}$, $\boldsymbol{\Phi}$ and $\boldsymbol{A}$. Nevertheless, a local stationary point can be found using an Alternating Least Squares (ALS) strategy, which minimizes (21) successively with respect to one variable at a time [130].

The ALS approach allows us to break (25) into three simpler problems which are solved sequentially, consisting of:

$$\begin{aligned} &a)\,\text{minimize } \mathcal{J}(\mathscr{M}|\boldsymbol{A},\boldsymbol{\Phi}) \text{ w.r.t. } \mathscr{M} \text{ with } \boldsymbol{A} \text{ and } \boldsymbol{\Phi} \text{ fixed} \\ &b)\,\text{minimize } \mathcal{J}(\boldsymbol{\Phi}|\boldsymbol{A},\mathscr{M}) \text{ w.r.t. } \boldsymbol{\Phi} \text{ with } \boldsymbol{A} \text{ and } \mathscr{M} \text{ fixed} \\ &c)\,\text{minimize } \mathcal{J}(\boldsymbol{A}|\,\mathscr{M},\boldsymbol{\Phi}) \text{ w.r.t. } \boldsymbol{A} \text{ with } \mathscr{M} \text{ and } \boldsymbol{\Phi} \text{ fixed} \end{aligned} \tag{26}$$

where $\mathcal{J}(\boldsymbol{B}_1|\boldsymbol{B}_2,\boldsymbol{B}_3)$ denotes a cost function $\mathcal{J}$ in which $\boldsymbol{B}_1$ is considered a variable and $\boldsymbol{B}_2$, $\boldsymbol{B}_3$ are fixed and thus constants.

Although this strategy yields a local minimum of the non-convex problem (25) by solving a sequence of convex optimization problems, it is still computationally intensive, specially due to the abundance estimation problem. This is because the spatial regularization term $\mathcal{R}(\boldsymbol{A})$ in (21) imposes interdependency among the different pixels of $\boldsymbol{A}$, what also happens when the TV regularization is employed [19, 110].

Each of the optimization subproblems of the ALS strategy in (26) will be treated in detail in the next section. Furthermore, in Section 4.6 we will present a multiscale formulation that eliminates the interdependency of the abundance estimation problem between the different image pixels, allowing the solution to be computed faster and in parallel.

4.5 FORMULATION AND SOLUTIONS TO THE OPTIMIZATION PROBLEMS IN (26)

We now detail the solution to each of the optimization problems in the ALS strategy outlined in (26). Although the solutions to the minimization problems w.r.t. $\mathscr{M}$ and $\boldsymbol{\Phi}$ are relatively straightforward and directly amenable to paralel or efficient implementations, optimizing (21) w.r.t. $\boldsymbol{A}$ proves to be significantly more challenging due to the multiscale spatial regularization term $\mathcal{R}(\boldsymbol{A})$. Nevertheless, by making some approximations in Section 4.6, we will reformulate this optimization problem as a function of the multiscale representations $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$ of the abundances. This will allow the extension of the ALS strategy to consider separate minimization steps w.r.t. $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$, leading to a simple and parallelizable solution. The complete algorithm including all optimization steps will be detailed in Section 4.7.

4.5.1 Optimizing with respect to $\mathscr{M}$ at the i-th iteration

The cost function in this case is $\mathcal{J}(\mathscr{M}\,|\boldsymbol{A},\boldsymbol{\Phi})$, where $\mathscr{M}$ is a variable and $\boldsymbol{A}$ and $\boldsymbol{\Phi}$ are fixed at the solutions obtained in the previous iteration. Then,

$$\mathcal{J}(\mathscr{M}\,|\boldsymbol{A},\boldsymbol{\Phi}) = \frac{1}{2}\sum_{n=1}^{N}\Big(\|\boldsymbol{y}_n - \boldsymbol{M}_n\boldsymbol{a}_n\|_2^2 + \lambda_M\|\boldsymbol{M}_n - \boldsymbol{M}_0\ \mathrm{diag}(\boldsymbol{\phi}_n)\|_F^2\Big)$$
$$\text{subject to } \boldsymbol{M}_n \succcurlyeq \boldsymbol{0},\ n = 1,\ldots,N \tag{27}$$

Similarly to [19], we compute an approximate solution to minimize (27) for each image pixel as

$$\widehat{\boldsymbol{M}}_n = \mathcal{P}_+\Big(\big(\boldsymbol{y}_n\boldsymbol{a}_n^\top + \lambda_M\boldsymbol{M}_0\ \mathrm{diag}(\boldsymbol{\phi}_n)\big)\big(\boldsymbol{a}_n\boldsymbol{a}_n^\top + \lambda_M\boldsymbol{I}\big)^{-1}\Big) \tag{28}$$

where $\mathcal{P}_+(\cdot)$ is an operator that projects each element of a matrix onto the nonnegative orthant by thresholding any negative element to zero.

4.5.2 Optimizing with respect to $\boldsymbol{\Phi}$ at the i-th iteration

The cost function in this case is $\mathcal{J}(\boldsymbol{\Phi}\,|\,\mathscr{M},\boldsymbol{A})$, where $\boldsymbol{\Phi}$ is a variable and $\boldsymbol{A}$ and $\mathscr{M}$ are fixed at the solutions obtained in the previous iteration. Then,

$$\mathcal{J}(\boldsymbol{\Phi}|\mathscr{M},\boldsymbol{A}) = \frac{\lambda_M}{2}\sum_{n=1}^{N}\|\boldsymbol{M}_n - \boldsymbol{M}_0\,\mathrm{diag}(\boldsymbol{\phi}_n)\|_F^2 + \lambda_{\boldsymbol{\Theta}}\mathcal{R}(\boldsymbol{\Phi}). \tag{29}$$

We follow the approach detailed in [19, Eqs. (20)-(23)] to minimize (29).

4.5.3 Optimizing with respect to $\boldsymbol{A}$ at the i-th iteration

The cost function in this case is $\mathcal{J}(\boldsymbol{A}\,|\,\mathscr{M},\boldsymbol{\Phi})$, where $\boldsymbol{A}$ is a variable and $\boldsymbol{\Phi}$ and $\mathscr{M}$ are fixed at the solutions obtained in the previous iteration. Then,

$$\mathcal{J}(\boldsymbol{A}\,|\,\mathscr{M},\boldsymbol{\Phi}) = \frac{1}{2}\big\|\boldsymbol{Y} - \big[\boldsymbol{M}_1\boldsymbol{a}_1\ldots\boldsymbol{M}_N\boldsymbol{a}_N\big]\big\|_F^2 + \frac{\rho\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{C}}\|_F^2 + \frac{\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{D}}\|_F^2 \tag{30}$$
$$\text{subject to } \boldsymbol{A} \succcurlyeq \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{A} = \boldsymbol{1}^\top$$
$$\boldsymbol{A}_{\mathcal{C}} = \boldsymbol{A}\boldsymbol{W},\ \boldsymbol{A}_{\mathcal{D}} = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)$$

Using the multiscale transformation $\widetilde{\boldsymbol{W}}$ to write (30) as a function of the observed hyperspectral images $\boldsymbol{Y}_{\mathcal{C}}$ and $\boldsymbol{Y}_{\mathcal{D}}$ represented at the coarse and detail scales yields

$$\begin{aligned}\mathcal{J}(\boldsymbol{A}\,|\,\mathscr{M},\boldsymbol{\Phi}) &= \frac{1}{2}\big\|\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^{*} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big]\boldsymbol{W}\boldsymbol{W}^{*}\big\|_F^2 \\ &+ \frac{1}{2}\big\|\boldsymbol{Y}_{\mathcal{D}} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big](\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^{*})\big\|_F^2 \\ &+ \operatorname{tr}\Big\{\big(\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^{*} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big]\boldsymbol{W}\boldsymbol{W}^{*}\big)^{\top} \\ &\qquad \cdot \big(\boldsymbol{Y}_{\mathcal{D}} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big](\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^{*})\big)\Big\} \\ &+ \frac{\rho\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{C}}\|_F^2 + \frac{\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{D}}\|_F^2 \\ &\text{subject to } \boldsymbol{A} \geqslant \boldsymbol{0},\ \boldsymbol{1}^{\top}\boldsymbol{A} = \boldsymbol{1}^{\top} \\ &\qquad \boldsymbol{A}_{\mathcal{C}} = \boldsymbol{A}\boldsymbol{W},\ \boldsymbol{A}_{\mathcal{D}} = \boldsymbol{A}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^{*})\end{aligned} \tag{31}$$

where $\operatorname{tr}(\cdot)$ is the matrix trace operator.

Cost function (31) is neither separable with respect to the abundance matrices $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$ in the coarse and detail scales, nor with respect to the image pixels. This can severely impact the required computational load and the convergence time to a meaningful result. To mitigate this issue, in the following section we propose to use few reasonable approximations to turn the minimization of (31) into an optimization problem separable in $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$. This will remove the interdependency between the different image pixels, and allow the extension of the ALS strategy to consider the optimization w.r.t. $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$ successively, instead of w.r.t. $\boldsymbol{A}$.

4.6 MODIFICATION AND SOLUTION TO THE OPTIMIZATION PROBLEM W.R.T. $\boldsymbol{A}$

Initially, we note that the cost function (31) does not depend on the endmember variability model $f(\boldsymbol{M}_0, \boldsymbol{\theta}_n)$. Hence, the derivations presented in this section are not limited to the ELMM, and can be equally applied to other models without loss of generality.

4.6.1 Residuals inner product

To proceed, we first denote by $RE_{\mathcal{C}}$ and $RE_{\mathcal{D}}$ the residuals/reconstruction errors in each image scale $\mathcal{C}$ and $\mathcal{D}$, where $RE_{\mathcal{C}}$ and $RE_{\mathcal{D}}$ are given by

$$\begin{aligned}RE_{\mathcal{C}} &= \boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^{*} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big]\boldsymbol{W}\boldsymbol{W}^{*}\,, \\ RE_{\mathcal{D}} &= \boldsymbol{Y}_{\mathcal{D}} - \big[\boldsymbol{M}_1\boldsymbol{a}_1 \ldots \boldsymbol{M}_N\boldsymbol{a}_N\big](\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^{*})\,.\end{aligned} \tag{32}$$

It follows from the above definition that the third term in the cost function (31) consists of the inner product $\langle RE_{\mathcal{C}}, RE_{\mathcal{D}}\rangle$ between the residuals/reconstruction errors at the coarse and detail scales. This inner product, however, usually contributes a small value to the cost function, and can be neglected under the following assumption:

A1 - **Zero-mean, uncorrelated residuals:** We assume that for $\boldsymbol{A}$ a critical point of (31), $RE_{\mathcal{C}}$ and $RE_{\mathcal{D}}$ are spatially zero-mean and uncorrelated across scales. This is reasonable if the observation/mixing model given by the ELMM in (15) represents the data with reasonable accuracy, in which case the main contribution towards the residual error comes from the observation noise $\boldsymbol{e}_n$, which is white and spatially uncorrelated. If A1 is satisfied, then the term $\langle RE_{\mathcal{C}}, RE_{\mathcal{D}} \rangle$ can be neglected when compared to the first two terms without significantly altering the critical point.

Although neglecting the third term of (31) simplifies the optimization problem, the first two terms still encompass intricate relationships between the abundances at different pixels due to the action of the multiscale transformation $\boldsymbol{W}$. Furthermore, the optimization problem still involves terms depending on both $\boldsymbol{A}$ and the pair $(\boldsymbol{A}_{\mathcal{C}}, \boldsymbol{A}_{\mathcal{D}})$, which are related through $\widetilde{\boldsymbol{W}}$, and thus cannot be easily solved in this form.

In order to proceed, we make the following assumption:

A2 - **Spatially smooth endmember signatures:** We assume that the pixel-by-pixel endmember signatures $\boldsymbol{M}_n$ are similar in small, compact spatial neighborhoods. More precisely, if $\mathcal{N}$ is a set of pixels comprising a compact spatial neighborhood, we assume that the endmember signature of any pixel in $\mathcal{N}$ does not deviate significantly from the average signature, so that the quantity

$$\left\| \boldsymbol{M}_j - \frac{1}{|\mathcal{N}|} \sum_{n \in \mathcal{N}} \boldsymbol{M}_n \right\|_F \tag{33}$$

is small for all $j \in \mathcal{N}$, where $|\mathcal{N}|$ is the cardinality of $\mathcal{N}$.

We show in the following that this assumption leads to the separation of the optimization w.r.t. $\boldsymbol{A}$ in (26) into two optimization steps, one w.r.t. $\boldsymbol{A}_{\mathcal{C}}$, and the other w.r.t. $\boldsymbol{A}_{\mathcal{D}}$. For numerical verification of the reasonability of A1 and A2, see Section 4.11 at the end of this chapter.

4.6.2 Approximate Mixture Model

Consider (31) after neglecting its third term. Both $\boldsymbol{W}$ (in the first term) and $\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*$ (in the second term) act upon all the products $\boldsymbol{M}_n \boldsymbol{a}_n$, instead of just upon $\boldsymbol{a}_n$, for $n = 1, \ldots, N$. This precludes the separation of (31) in a sum of non-negative functions exclusively dependent on $\boldsymbol{A}_{\mathcal{C}}$ or $\boldsymbol{A}_{\mathcal{D}}$, which could be independently minimized. However, combining A2 and the fact that the transformation $\boldsymbol{W}$ groups pixels that are in spatially adjacent regions, we now propose an approximate separable mixing model.

We initially express each pixel $\boldsymbol{y}_{\mathcal{C}_i}$ and $\boldsymbol{y}_{\mathcal{D}_i}$ of (20) as

$$\boldsymbol{y}_{\mathcal{C}_i} = \sum_{j=1}^{N} W_{j,i} \, \boldsymbol{M}_j \boldsymbol{a}_j + \boldsymbol{e}_{\mathcal{C}_i} \tag{34}$$

and

$$\boldsymbol{y}_{\mathcal{D}_i} = \boldsymbol{M}_i \boldsymbol{a}_i - \sum_{j=1}^{S} \sum_{\ell=1}^{N} W^*_{j,i} W_{\ell,j} \, \boldsymbol{M}_\ell \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i} \,, \tag{35}$$

where $W_{j,i}$ and $W^*_{j,i}$ are the (j,i)-th elements of $\boldsymbol{W}$ and $\boldsymbol{W}^*$, respectively, and $\boldsymbol{e}_{\mathcal{C}_i}$ and $\boldsymbol{e}_{\mathcal{D}_i}$ denote the i-th columns of $\boldsymbol{E}_{\mathcal{C}}$ and $\boldsymbol{E}_{\mathcal{D}}$. Then, using A2 and the fact that $\boldsymbol{W}$ is a localized decomposition, we approximate every endmember matrix $\boldsymbol{M}_j$ in (34) by

$$\boldsymbol{M}_j \approx \boldsymbol{M}_{\mathcal{C}_i} = \sum_{\ell=1}^{N} \frac{\mathbb{1}_{W_{\ell,i}}}{|\operatorname{supp}_\ell(W_{\ell,i})|} \boldsymbol{M}_\ell \,, \tag{36}$$

where $\mathbb{1}_{W_{j,i}}$ is the indicator function of $W_{j,i}$ (i.e. $\mathbb{1}_{W_{j,i}} = 1$ if $W_{j,i} \neq 0$ and $\mathbb{1}_{W_{j,i}} = 0$ otherwise), and $|\operatorname{supp}_\ell(g)|$ denotes the cardinality of the support of g as a function of ℓ.

Equivalently, we approximate every matrix $\boldsymbol{M}_\ell$ in (35) by

$$\boldsymbol{M}_\ell \approx \boldsymbol{M}_{\mathcal{C}^*_i} = \sum_{n=1}^{S} \sum_{m=1}^{N} \frac{\mathbb{1}_{W^*_{n,i}} \mathbb{1}_{W_{m,n}}}{|\operatorname{supp}_{n,m}(W^*_{n,i} W_{m,n})|} \boldsymbol{M}_m \,, \tag{37}$$

where $|\operatorname{supp}_{m,n}(g)|$ denotes the cardinality of the support of g as a function of both m and n. Thus, (34) and (35) can be approximated as (details in Appendix 4.10)

$$\boldsymbol{y}_{\mathcal{C}_i} \approx \boldsymbol{M}_{\mathcal{C}_i} \boldsymbol{a}_{\mathcal{C}_i} + \boldsymbol{e}_{\mathcal{C}_i} \tag{38}$$

and

$$\boldsymbol{y}_{\mathcal{D}_i} \approx \boldsymbol{M}_i \boldsymbol{a}_{\mathcal{D}_i} + \boldsymbol{M}_{\mathcal{D}_i} \big[\boldsymbol{A}_{\mathcal{C}} \boldsymbol{W}^* \big]_i + \boldsymbol{e}_{\mathcal{D}_i} \,, \tag{39}$$

where $[\cdot]_i$ denotes the $i-$th column of a matrix, and $\boldsymbol{M}_{\mathcal{D}_i} = \boldsymbol{M}_i - \boldsymbol{M}_{\mathcal{C}^*_i}$ reflects the variability of $\boldsymbol{M}_i$ with respect to $\boldsymbol{M}_{\mathcal{C}^*_i}$, the average endmember matrix of its neighborhood. According to A2, $\boldsymbol{M}_{\mathcal{D}_i} \approx \boldsymbol{0}$. Note that, since the transformation $\boldsymbol{W}$ only groups together pixels that lie inside a single superpixel, we average $\boldsymbol{a}_n$ and $\boldsymbol{M}_n$ only in small spatial neighborhoods where their variability is small.

Selecting $\boldsymbol{W}$ and $\boldsymbol{W}^*$ according to the superpixels decomposition, we have that:

- $\boldsymbol{M}_{\mathcal{C}_i}$ is the average of all $\boldsymbol{M}_j$ inside the i-th superpixel.
- $\boldsymbol{M}_{\mathcal{C}^*_i}$ is the average of all $\boldsymbol{M}_j$ inside the superpixel that contains the i-th pixel. Thus, if pixel i belongs to the k-th superpixel, $\boldsymbol{M}_{\mathcal{C}^*_i}$ is the average of all $\boldsymbol{M}_j$ inside the k-th superpixel.

Note that $\boldsymbol{W}^*$ is also a localized transform, as it attributes the superpixel value to all pixels in the original domain that lie inside that superpixel, which encompasses a compact spatial neighborhood.

Writting (38) and (39) for all pixels, we write (34) and (35) in the matrix form as:

$$\begin{aligned}\boldsymbol{Y}_{\mathcal{C}} &= \left[\boldsymbol{M}_{\mathcal{C}_1}\boldsymbol{a}_{\mathcal{C}_1},\ldots,\boldsymbol{M}_{\mathcal{C}_S}\boldsymbol{a}_{\mathcal{C}_S}\right]+\widetilde{\boldsymbol{E}}_{\mathcal{C}}\\ \boldsymbol{Y}_{\mathcal{D}} &= \left[\boldsymbol{M}_{\mathcal{D}_1}\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_1,\ldots,\boldsymbol{M}_{\mathcal{D}_N}\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_N\right]\\ &\quad+\left[\boldsymbol{M}_1\boldsymbol{a}_{\mathcal{D}_1},\ldots,\boldsymbol{M}_N\boldsymbol{a}_{\mathcal{D}_N}\right]+\widetilde{\boldsymbol{E}}_{\mathcal{D}}\end{aligned} \tag{40}$$

where $\widetilde{\boldsymbol{E}}_{\mathcal{C}}$ and $\widetilde{\boldsymbol{E}}_{\mathcal{D}}$ include additive noise and modeling errors.

4.6.2.1 Abundance constraints

The two constraints in (31) are functions of $\boldsymbol{A}$, and thus must be considered in the optimization with respect to $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$. Assuming $\widetilde{\boldsymbol{W}}$ in (18) to be of full row rank, the sum-to-one constraint can be expressed as

$$\begin{aligned}&\mathbf{1}^\top\boldsymbol{A}\widetilde{\boldsymbol{W}}=\mathbf{1}^\top\widetilde{\boldsymbol{W}}\\ &\Longleftrightarrow \mathbf{1}^\top\boldsymbol{A}_{\mathcal{C}}=\mathbf{1}^\top\boldsymbol{W}\,,\quad \mathbf{1}^\top\boldsymbol{A}_{\mathcal{D}}=\mathbf{1}^\top(\boldsymbol{I}-\boldsymbol{W}\boldsymbol{W}^*).\end{aligned} \tag{41}$$

Considering the positivity constraint we have

$$\begin{aligned}&\boldsymbol{A}\geqslant\mathbf{0}\Rightarrow\boldsymbol{A}\widetilde{\boldsymbol{W}}\widetilde{\boldsymbol{W}}^\dagger\geqslant\mathbf{0}\\ &\Longleftrightarrow[\boldsymbol{A}_{\mathcal{C}}\ \ \boldsymbol{A}_{\mathcal{D}}]\widetilde{\boldsymbol{W}}^\dagger\geqslant\mathbf{0}\Longleftrightarrow\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*+\boldsymbol{A}_{\mathcal{D}}\geqslant\mathbf{0}\,.\end{aligned} \tag{42}$$

If $\boldsymbol{W}^*\geqslant\mathbf{0}$, which is true if $\boldsymbol{W}$ is selected as the superpixel decomposition, we can further state that

$$\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\geqslant\mathbf{0}\Longleftrightarrow\boldsymbol{A}_{\mathcal{C}}\geqslant\mathbf{0}\,, \tag{43}$$

which simplifies the constraint by removing possible interdependencies between different pixels, and makes the problem separable for all pixels in the coarse scale $\mathcal{C}$.

4.6.2.2 The updated optimization problem

Using the results obtained in Sections 4.6.1 to 4.6.2.1, minimizing (31) with respect to $\boldsymbol{A}$ can be restated as determining $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$ that minimize

$$\begin{aligned}\widetilde{\mathcal{J}}(\boldsymbol{A}_{\mathcal{C}},\boldsymbol{A}_{\mathcal{D}}|\,\mathscr{M},\boldsymbol{\Phi}) &= \frac{1}{2}\left\|\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^*-\left[\boldsymbol{M}_{\mathcal{C}_1}\boldsymbol{a}_{\mathcal{C}_1},\ldots,\boldsymbol{M}_{\mathcal{C}_S}\boldsymbol{a}_{\mathcal{C}_S}\right]\boldsymbol{W}^*\right\|_F^2\\ &\quad+\frac{1}{2}\left\|\boldsymbol{Y}_{\mathcal{D}}-\left[\boldsymbol{M}_1\boldsymbol{a}_{\mathcal{D}_1},\ldots,\boldsymbol{M}_N\boldsymbol{a}_{\mathcal{D}_N}\right]-\left[\boldsymbol{M}_{\mathcal{D}_1}\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_1,\ldots,\boldsymbol{M}_{\mathcal{D}_N}\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_N\right]\right\|_F^2\\ &\quad+\frac{\rho\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{C}}\|_F^2+\frac{\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{D}}\|_F^2\\ &\text{subject to } \boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*+\boldsymbol{A}_{\mathcal{D}}\geqslant\mathbf{0},\ \mathbf{1}^\top\boldsymbol{A}_{\mathcal{C}}=\mathbf{1}^\top\boldsymbol{W},\\ &\qquad\mathbf{1}^\top\boldsymbol{A}_{\mathcal{D}}=\mathbf{1}^\top(\boldsymbol{I}-\boldsymbol{W}\boldsymbol{W}^*).\end{aligned} \tag{44}$$

Optimization problem (44) is amenable to an efficient solution, as detailed in the following section.

4.6.3 Solution to the optimization problem (44)

This section details the proposed solution of the optimization problem (44) w.r.t. $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{A}_{\mathcal{D}}$.

4.6.3.1 Optimizing with respect to $\boldsymbol{A}_{\mathcal{C}}$ at the i-th iteration

The cost function in this case is $\widetilde{\mathcal{J}}(\boldsymbol{A}_{\mathcal{C}}|\boldsymbol{A}_{\mathcal{D}}, \mathscr{M}, \boldsymbol{\Phi})$, where $\boldsymbol{A}_{\mathcal{C}}$ is a variable and $\boldsymbol{A}_{\mathcal{D}}$, $\mathscr{M}$ and $\boldsymbol{\Phi}$ are fixed at the solutions obtained in the previous iteration. Note that this problem is still not separable with respect to each pixel in $\boldsymbol{A}_{\mathcal{C}}$ since the second term of (44) includes products between $\boldsymbol{A}_{\mathcal{C}}$ and $\boldsymbol{W}^*$. However, this cost function can be simplified to yield a separable problem by making the following considerations using assumption A2:

1. A2 implies that the entries of $\boldsymbol{M}_{\mathcal{D}_i}$ are small when compared to those of $\boldsymbol{M}_n$;
2. A2 also implies that the entries of $\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^*$ are usually much larger than the entries of $\boldsymbol{Y}_{\mathcal{D}}$.

These considerations imply that the contribution of the terms $\boldsymbol{M}_{\mathcal{D}_i}\big[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\big]$ in the second term of (44) can be neglected when compared to $\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^*$. Using this approximation and (43), the optimization with respect to $\boldsymbol{A}_{\mathcal{C}}$ can be stated as the minimization of

$$\begin{aligned}
&\overline{\mathcal{J}}(\boldsymbol{A}_{\mathcal{C}}|\boldsymbol{A}_{\mathcal{D}}, \mathscr{M}, \boldsymbol{\Phi}) \\
&\quad = \frac{1}{2}\big\|\boldsymbol{Y}_{\mathcal{C}}\boldsymbol{W}^* - \big[\boldsymbol{M}_{\mathcal{C}_1}\boldsymbol{a}_{\mathcal{C}_1}, \ldots, \boldsymbol{M}_{\mathcal{C}_S}\boldsymbol{a}_{\mathcal{C}_S}\big]\boldsymbol{W}^*\big\|_F^2 + \frac{\rho\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{C}}\|_F^2 \\
&\quad \text{subject to } \boldsymbol{A}_{\mathcal{C}} \geqslant \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{A}_{\mathcal{C}} = \boldsymbol{1}^\top\boldsymbol{W}\,.
\end{aligned} \tag{45}$$

For $\boldsymbol{W}$ based on the superpixel decomposition, $\boldsymbol{W}^*$ assigns to each pixel in the original image domain the value of the superpixel to which it belongs. Using this property, the cost function (45) simplifies to

$$\begin{aligned}
\overline{\mathcal{J}}(\boldsymbol{A}_{\mathcal{C}}|\boldsymbol{A}_{\mathcal{D}}, \mathscr{M}, \boldsymbol{\Phi}) = \frac{1}{2}\sum_{n=1}^{S}\Omega_s^2(n)\Big(\|\boldsymbol{y}_{\mathcal{C}_n} - \boldsymbol{M}_{\mathcal{C}_n}\boldsymbol{a}_{\mathcal{C}_n}\|_2^2 + \frac{\widetilde{\rho}(n)\lambda_A}{2}\|\boldsymbol{a}_{\mathcal{C}_n}\|_2^2\Big) \\
\text{subject to } \boldsymbol{a}_{\mathcal{C}_n} \geqslant \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{a}_{\mathcal{C}_n} = \boldsymbol{1}^\top[\boldsymbol{W}]_n,\ n = 1, \ldots, S
\end{aligned} \tag{46}$$

where $[\boldsymbol{W}]_n$ is the n-th column of $\boldsymbol{W}$, $\Omega_s(n)$ is the number of pixels contained in the n-th superpixel and $\widetilde{\rho}(n) = \rho\Omega_s^{-2}(n)$, $n = 1, \ldots, S$ is a superpixel-dependent regularization parameter that controls the balance between both terms in the cost function for each superpixel.

For simplicity, in the following we replace $\widetilde{\rho}(n)$ by a weighting term $\widetilde{\rho}_0 = \rho S^2/N^2$ that is constant for all superpixels. This further simplifies the optimization problem since S is specified a priori by the user. Furthermore, since the optimization is independent for each pixel, we can also move the $\Omega_s^2(n)$ factor outside the summation in (46) without changing the critical point of the cost function.

Doing this results in the following cost function that can be minimized individually for each pixel:

$$\widehat{\mathcal{J}}(\boldsymbol{A}_{\mathcal{C}}|\boldsymbol{A}_{\mathcal{D}},\mathscr{M},\boldsymbol{\Phi}) = \frac{N^2}{2S^2}\sum_{n=1}^{S}\left(\|\boldsymbol{y}_{\mathcal{C}_n} - \boldsymbol{M}_{\mathcal{C}_n}\boldsymbol{a}_{\mathcal{C}_n}\|_2^2 + \frac{\widetilde{\rho}_0\lambda_A}{2}\|\boldsymbol{a}_{\mathcal{C}_n}\|_2^2\right) \tag{47}$$
$$\text{subject to } \boldsymbol{a}_{\mathcal{C}_n} \geqslant \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{a}_{\mathcal{C}_n} = \boldsymbol{1}^\top[\boldsymbol{W}]_n,\ n = 1,\ldots,S.$$

Note that (47) is equivalent to a a standard fully constrained least squares (FCLS) problem, which can be solved efficiently.

4.6.3.2 Optimizing with respect to $\boldsymbol{A}_{\mathcal{D}}$ at the i-th iteration

The cost function in this case is $\widetilde{\mathcal{J}}(\boldsymbol{A}_{\mathcal{D}}|\boldsymbol{A}_{\mathcal{C}},\mathscr{M},\boldsymbol{\Phi})$, where $\boldsymbol{A}_{\mathcal{D}}$ is a variable and $\boldsymbol{A}_{\mathcal{C}}$, $\mathscr{M}$ and $\boldsymbol{\Phi}$ are fixed at the solutions obtained in the previous iteration. Then, considering only the terms and constraints in (44) that depend on $\boldsymbol{A}_{\mathcal{D}}$ yields

$$\begin{aligned}
&\widetilde{\mathcal{J}}(\boldsymbol{A}_{\mathcal{D}}|\boldsymbol{A}_{\mathcal{C}},\mathscr{M},\boldsymbol{\Phi}) \\
&\quad= \frac{1}{2}\Big\|\boldsymbol{Y}_{\mathcal{D}} - \big[\boldsymbol{M}_1\boldsymbol{a}_{\mathcal{D}_1},\ldots,\boldsymbol{M}_N\boldsymbol{a}_{\mathcal{D}_N}\big] - \big[\boldsymbol{M}_{\mathcal{D}_1}\big[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\big]_1,\ldots,\boldsymbol{M}_{\mathcal{D}_N}\big[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\big]_N\big]\Big\|_F^2 \\
&\quad+ \frac{\lambda_A}{2}\|\boldsymbol{A}_{\mathcal{D}}\|_F^2 \qquad (48)\\
&\quad\text{subject to } \boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^* + \boldsymbol{A}_{\mathcal{D}} \geqslant 0, \\
&\qquad\qquad \boldsymbol{1}^\top\boldsymbol{A}_{\mathcal{D}} = \boldsymbol{1}^\top(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*).
\end{aligned}$$

Since matrix $\boldsymbol{A}_{\mathcal{C}}$ is fixed, this problem can be decomposed for each pixel. This results in the minimization of the following cost function:

$$\begin{aligned}
&\widetilde{\mathcal{J}}(\boldsymbol{A}_{\mathcal{D}}|\boldsymbol{A}_{\mathcal{C}},\mathscr{M},\boldsymbol{\Phi}) \\
&\quad= \frac{1}{2}\sum_{n=1}^{N}\left(\big\|\boldsymbol{y}_{\mathcal{D}_n} - \boldsymbol{M}_n\boldsymbol{a}_{\mathcal{D}_n} - \boldsymbol{M}_{\mathcal{D}_n}\big[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\big]_n\big\|_2^2 + \lambda_A\|\boldsymbol{a}_{\mathcal{D}_n}\|_2^2\right) \qquad (49)\\
&\quad\text{subject to } \big[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\big]_n + \boldsymbol{a}_{\mathcal{D}_n} \geqslant 0 \\
&\qquad\qquad \boldsymbol{1}^\top\boldsymbol{a}_{\mathcal{D}_n} = \boldsymbol{1}^\top\big[\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*\big]_n \\
&\qquad\qquad n = 1,\ldots,N
\end{aligned}$$

where matrices $\boldsymbol{M}_{\mathcal{D}_n}$ are given in (38) and (39). Note that this cost function is again equivalent to a standard FCLS problem, which can be solved efficiently.

4.7 THE MUA-SV UNMIXING ALGORITHM

Considering the solutions to the optimization subproblems derived in the previous sections, the global unmixing procedure can be directly derived by setting the fixed variables of each subproblem with the estimates obtained from the previous iteration. The MUA-SV algorithm is presented in Algorithm 3.

Algorithme 3 : Global MUA-SV algorithm

Input : Image $\boldsymbol{Y}$, parameters λ_M, λ_A, λ_Θ, ρ and matrices $\boldsymbol{A}^{(0)}$, $\boldsymbol{\Phi}^{(0)}$ and $\boldsymbol{M}_0$.
Output : Estimated matrices $\widehat{\boldsymbol{A}}$, $\widehat{\boldsymbol{\Phi}}$ and tensor $\widehat{\mathscr{M}}$.

Compute the superpixel decomposition of the hyperspectral image $\boldsymbol{Y}$ and the corresponding transformation matrices $\boldsymbol{W}$, $\boldsymbol{W}^*$, $\widetilde{\boldsymbol{W}}$ and $\widetilde{\boldsymbol{W}}^\dagger$ using the SLIC algorithm [118];
Compute the decomposition of $\boldsymbol{Y}$ into approximation and detail domains $\boldsymbol{Y}_\mathcal{C}$, and $\boldsymbol{Y}_\mathcal{D}$ using (16) and (17) ;
Set $\boldsymbol{A}_\mathcal{D}^{(0)} = \boldsymbol{A}^{(0)}(\boldsymbol{I} - \boldsymbol{W}\boldsymbol{W}^*)$;
Set $i = 1$;
while *stopping criterion is not satisfied* **do**
  $\mathscr{M}^{(i)} = \underset{\mathscr{M}}{\arg\min}\ \mathcal{J}(\mathscr{M}|\boldsymbol{A}^{(i-1)}, \boldsymbol{\Phi}^{(i-1)})$;
  $\boldsymbol{A}_\mathcal{C}^{(i)} = \underset{\boldsymbol{A}_\mathcal{C}}{\arg\min}\ \widehat{\mathcal{J}}(\boldsymbol{A}_\mathcal{C}|\boldsymbol{A}_\mathcal{D}^{(i-1)}, \mathscr{M}^{(i)}, \boldsymbol{\Phi}^{(i-1)})$;
  $\boldsymbol{A}_\mathcal{D}^{(i)} = \underset{\boldsymbol{A}_\mathcal{D}}{\arg\min}\ \widetilde{\mathcal{J}}(\boldsymbol{A}_\mathcal{D}|\boldsymbol{A}_\mathcal{C}^{(i)}, \mathscr{M}^{(i)}, \boldsymbol{\Phi}^{(i-1)})$;
  $\boldsymbol{A}^{(i)} = \left[\boldsymbol{A}_\mathcal{C}^{(i)}\ \boldsymbol{A}_\mathcal{D}^{(i)}\right]\widetilde{\boldsymbol{W}}^\dagger$;
  $\boldsymbol{\Phi}^{(i)} = \underset{\boldsymbol{\Phi}}{\arg\min}\ \mathcal{J}(\boldsymbol{\Phi}\,|\,\mathscr{M}^{(i)}, \boldsymbol{A}^{(i)})$;
  $i = i + 1$;
end
return $\widehat{\boldsymbol{A}} = \boldsymbol{A}^{(i-1)}$, $\widehat{\mathscr{M}} = \mathscr{M}^{(i-1)}$, $\widehat{\boldsymbol{\Phi}} = \boldsymbol{\Phi}^{(i-1)}$;

4.8 RESULTS

In this section, we compare the unmixing performances achieved using the proposed MUA-SV algorithm, the Fully Constrained Least Squares (FCLS), the Scaled Constrained Least Squares (SCLS), the PLMM-based solution [20] and the ELMM-based solution [19], the latter two designed to tackle spectral variability. The SCLS algorithm is a particular case of the ELMM model that employs the same scaling factors $\boldsymbol{\phi}_n$ for all endmembers in each pixel (i.e. $\boldsymbol{M}_n = \phi_n \boldsymbol{M}_0$, where $\phi_n \in \mathbb{R}_+$) [51]. It is a low complexity algorithm that can be used as a baseline method to account for spectral variability.

For all simulations, the reference endmember signatures $\boldsymbol{M}_0$ were extracted from the observed image using the Vertex Component Analysis (VCA) algorithm [131]. The abundance maps were initialized with the SCLS result for all algorithms. The scaling factors $\boldsymbol{\Phi}$ for ELMM and MUA-SV were initialized with ones. The matrix $\boldsymbol{M}$ for the PLMM was initialized with the results from the VCA. The alternating least squares loop in Algorithm 3 is terminated when the norm of the relative variation of the three variables between two successive iterations is smaller than $\epsilon_A = \epsilon_\Phi = \epsilon_M = 2 \times 10^{-3}$.

Experiments were performed for three synthetic and two real data sets. For the synthetic data, the regularization parameters were selected for each algorithm to provide the best abundance estimation performance. The complete set of parameters, comprising the SLIC (S and γ) and the regularization param-

eters (ρ, λ_M, λ_A, and λ_Θ), were searched in appropriate intervals. For instance, $\gamma \in \{0.001, 0.0025, 0.005, 0.01, 0.025, 0.05\}$, S assumed an integer value in the interval $[2, 9]$, ρ was selected so that $\rho S^2/N^2 \in \{0.001, 0.01, 0.025, 0.05, 0.1, 0.15, 0.2, 0.25, 0.35, 0.5\}$, while λ_M, λ_A, and λ_Θ were searched in the range $[5 \times 10^{-4}, 100]$, with 12 points sampled uniformly.

The algorithms were implemented on a desktop computer equipped with an Intel I7 4.2 Ghz processor with 4 cores and 16 Gb RAM. ELMM, PLMM and SLIC were implemented using the codes made available by the respective authors. We did not employ parallelism when implementing the MUA-SV algorithm, so as to reduce the influence of the hardware platform when evaluating the performance gains achieved through the proposed simplifications. If parallelism is employed, the execution times can be even smaller.

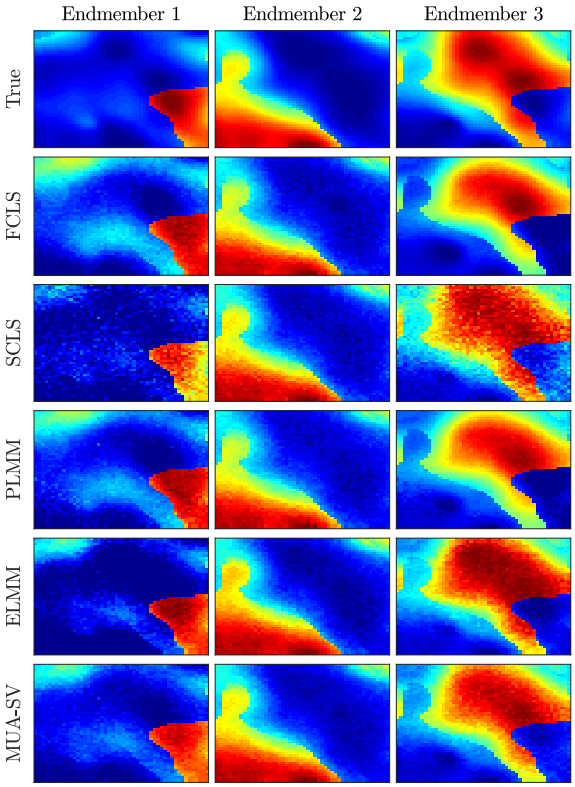
4.8.1 Synthetic data sets

Three synthetic data sets were built. The first data cube (DC1) was built from the ELMM model to verify how MUA-SV performs when the actual endmembers closely follows the adopted model. The second data cube (DC2) was built using the more challenging additive perturbation model of [20]. The third data cube (DC3) was based on a realistic simulation of endmember variability caused by illumination conditions following the Hapke's model [37].

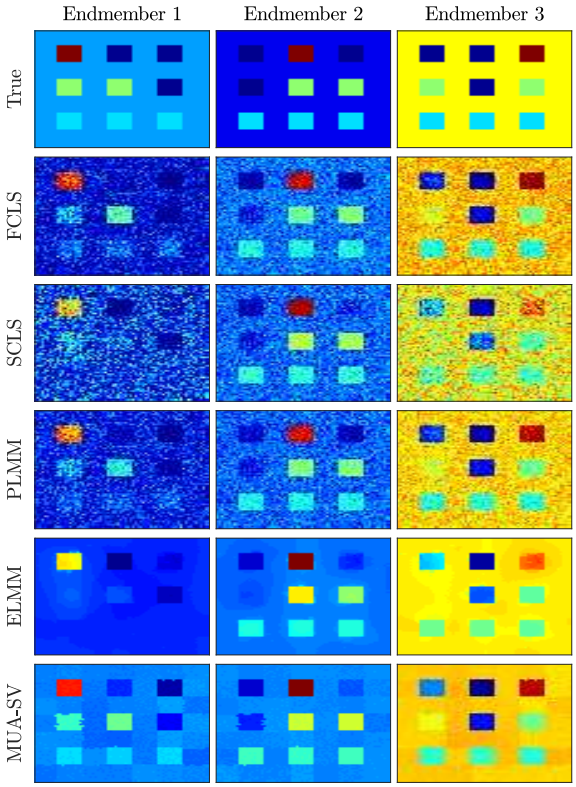
The data cube DC1 contains 50×50 pixels and three materials selected randomly from the USGS library and used as the reference endmember matrix $\boldsymbol{M}_0$, with 224 spectral bands. The abundance maps are piecewise smooth images generated by sampling from a Gaussian Random field [1] [132], and are depicted in Fig. 15a. Spectral variability was added to the reference endmembers using the same model as in [19], where the endmember instances for each pixel were generated by applying a constant scaling factor to the reference endmembers with amplitude limited to the interval $[0.75, 1.25]$. Finally, a white Gaussian noise with a 25 dB SNR was added to the already scaled endmembers. The true scaling factors applied to each endmember were generated using a Gaussian Random field, and thus exhibit spatial correlation.

The data cube DC2 contains 70×70 pixels and three materials, also randomly selected from the USGS spectral library to compose matrix $\boldsymbol{M}_0$ with 224 spectral bands. The abundance maps (shown in Fig. 15b) are composed by square regions distributed uniformly over a background, containing pure pixels (first row) and mixtures of two and three endmembers (second and third rows). The background pixels are mixtures of the same three endmembers, with abundances 0.2744, 0.1055 and 0.62. Spectral variability was added following the model proposed in [20], which considered a per-pixel variability given by random piecewise linear functions to scale individually the spectrum of each

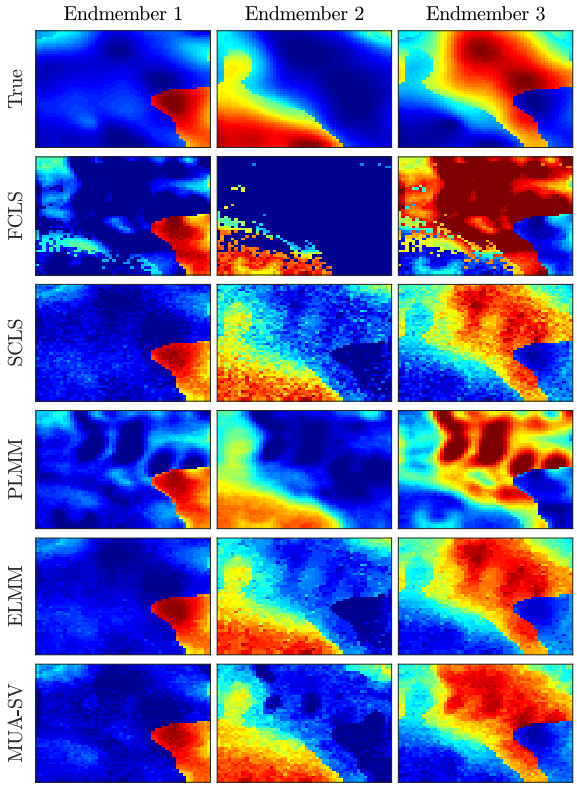
[1] Generated using the code in `http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB`



(a) DC1.



(b) DC2.



(c) DC3.

Figure 15 – True and reconstructed abundance maps for the synthetic data cubes for SNR=30 dB.

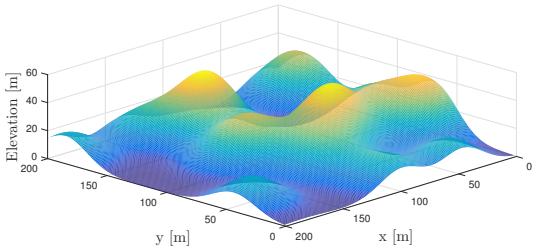


Figure 16 – Discrete terrain model used with the Hapke model in the data cube DC3, provided by [19].

endmember by a factor in the interval $[0.8, 1.2]$. Such a variability model does not match the ELMM, as it yields different variabilities across the spectral bands, and is not designed to produce spatial correlation. Nevertheless, it provides a good ground for comparison with more flexible models such as the PLMM.

The data cube DC3 contains 50×50 pixels and three materials, and is based on a simulation originally presented in [19][2]. This data cube is devised to realistically represent the spectral variability introduced due to changes in the illumination conditions caused by the topography of the scene, and is generated according to a physical model proposed by Hapke [37]. Hapke's model is able to represent the reflectance of a material as a function of its single scattering albedo, photometric parameters and geometric characteristics of the scene, namely, the incidence, emergence and azimuth angles during acquisition [19, 37]. Thus, pixel dependent reflectance signatures for each endmember can be obtained given its single scattering albedo and the scene topography.

In this example, the scene was composed of three materials, namely, basalt, palagonite and tephra, which are frequently present on small bodies of the Solar System, and contained 16 spectral bands. Afterwards, a digital terrain model simulating a hilly region was generated, which is shown in Fig. 16, and from this model the acquisition angles associated with each pixel were derived (as a function of the scene topography) by considering the angle between the sun and the horizontal plane as 18°, and the sensor to be placed vertically downward. Finally, the pixel dependent endmember signatures for the scene were generated from the single scattering albedo of the materials, and from the geometric characteristics of the scene using Hapke's model. The abundance maps used for DC2 were the same used for DC1, as shown in Fig. 15c.

The resulting hyperspectral images for all data cubes were generated from the pixel-dependent endmember signatures and abundance maps following the LMM, and were later contaminated by white Gaussian noise, with SNRs of 20, 30, and 40 dB. The regularization parameters for all algorithms and all examples were selected using a grid search procedure in order to provide best abundance estimation performance.

The unmixing accuracy metrics used are the abundances mean squared error (MSE)

[2] Most of the data for this simulation was generously provided by Lucas Drumetz and his collaborators.

$$\mathrm{MSE}_{\boldsymbol{A}} = \frac{1}{NP}\|\boldsymbol{A} - \widehat{\boldsymbol{A}}\|_F^2\,, \tag{50}$$

the mean squared error of the estimated spectra

$$\mathrm{MSE}_{\mathscr{M}} = \frac{1}{NLP}\sum_{n=1}^{N}\|\boldsymbol{M}_n - \widehat{\boldsymbol{M}}_n\|_F^2\,, \tag{51}$$

and the mean squared reconstruction error

$$\mathrm{MSE}_{\boldsymbol{Y}} = \frac{1}{NL}\sum_{n=1}^{N}\|\boldsymbol{y}_n - \widehat{\boldsymbol{M}}_n\widehat{\boldsymbol{a}}_n\|^2\,. \tag{52}$$

We also evaluate the estimates of the endmember signatures using the average Spectral Angle Mapper (SAM), defined by

$$\mathrm{SAM}_{\mathscr{M}} = \frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{P}\arccos\left(\frac{\boldsymbol{m}_{k,n}^{\top}\widehat{\boldsymbol{m}}_{k,n}}{\|\boldsymbol{m}_{k,n}\|\|\widehat{\boldsymbol{m}}_{k,n}\|}\right). \tag{53}$$

where $\boldsymbol{m}_{k,n}$ and $\widehat{\boldsymbol{m}}_{k,n}$ are the k-th columns of $\boldsymbol{M}_n$ and $\widehat{\boldsymbol{M}}_n$, respectively.

The quantitative results achieved by all algorithms are displayed in Table 4 for all tested SNR values. The reconstructed abundance maps for the three data cubes and an SNR of 30 dB are shown in Figs. 15a, 15b and 15c for a qualitative comparison. The computational complexity of the algorithms was evaluated through their execution times, which are shown in Table 5.

Table 4 – Quantitative results of all algorithms for data cubes DC1, DC2 and DC3 (with parameters selected to yield best abundance estimates). All values are multiplied by 10^3.

		DC1 data cube				DC2 data cube				DC3 data cube			
SNR	Method	$\mathrm{MSE}_{\boldsymbol{A}}$	$\mathrm{MSE}_{\mathscr{M}}$	$\mathrm{SAM}_{\mathscr{M}}$	$\mathrm{MSE}_{\boldsymbol{Y}}$	$\mathrm{MSE}_{\boldsymbol{A}}$	$\mathrm{MSE}_{\mathscr{M}}$	$\mathrm{SAM}_{\mathscr{M}}$	$\mathrm{MSE}_{\boldsymbol{Y}}$	$\mathrm{MSE}_{\boldsymbol{A}}$	$\mathrm{MSE}_{\mathscr{M}}$	$\mathrm{SAM}_{\mathscr{M}}$	$\mathrm{MSE}_{\boldsymbol{Y}}$
20 dB	FCLS	21.97	–	–	6.91	66.47	–	–	6.45	74.14	–	–	2.63
	SCLS	28.79	6.87	190.5	6.86	73.35	4.07	171.0	6.20	73.18	3.02	214.6	0.50
	PLMM	24.64	5.42	188.8	3.50	85.65	3.19	174.4	3.33	39.07	**1.44**	**122.7**	0.39
	ELMM	17.81	5.34	**186.7**	5.59	65.11	**3.09**	**170.9**	6.69	59.54	2.80	317.4	**0.0001**
	MUA-SV	**12.90**	**5.24**	212.2	**1.56**	**29.80**	3.36	185.7	**3.28**	**28.11**	1.84	308.6	0.0002
30 dB	FCLS	28.10	–	–	1.76	60.28	–	–	0.93	172.3	–	×	1.41
	SCLS	12.37	4.53	187.6	1.63	62.23	3.84	**161.2**	0.71	21.41	2.42	68.73	0.05
	PLMM	19.61	4.88	173.0	0.86	49.38	3.95	162.5	0.41	38.00	**1.53**	**68.53**	0.10
	ELMM	10.71	3.70	170.2	0.59	40.16	3.05	177.9	**0.001**	18.47	1.73	101.5	**0.00002**
	MUA-SV	**7.07**	**3.46**	**166.9**	**0.35**	**24.30**	**2.83**	161.5	0.33	**14.70**	1.75	68.62	0.07
40 dB	FCLS	20.04	–	–	1.23	71.37	–	–	0.44	256.2	–	–	1.39
	SCLS	7.38	3.88	186.3	1.10	69.48	3.52	160.1	0.17	8.98	2.40	30.90	**0.01**
	PLMM	13.44	3.64	170.3	0.56	44.73	3.02	**140.7**	0.11	34.38	1.47	74.15	0.08
	ELMM	5.36	**2.51**	**149.7**	**0.02**	46.83	**2.63**	159.2	**0.0002**	8.12	**1.28**	43.14	**0.01**
	MUA-SV	**3.98**	2.52	149.9	**0.02**	**26.01**	2.97	156.0	0.31	**7.94**	1.81	**30.66**	0.02

Table 5 – Execution time (in seconds) of the unmixing algorithms, averaged for all SNR values considered

	FCLS	SCLS	ELMM	PLMM	MUA-SV
DC1	0.14	0.42	14.76	16.17	2.57
DC2	0.27	0.83	37.52	149.91	18.29
DC3	0.17	0.35	15.82	63.07	9.59
Houston	0.82	2.31	174.53	484.02	36.29
Cuprite	6.63	15.61	527.89	7998.02	95.54

Table 6 – Reconstruction errors ($\text{MSE}_{\boldsymbol{Y}}$) for the Houston and Cuprite data sets (all values are multiplied by 10^3).

	FCLS	SCLS	ELMM	PLMM	MUA-SV
Houston	2.283	0.037	0.010	0.190	0.014
Cuprite	0.050	0.044	0.040	0.079	0.050

4.8.1.1 Discussion

Table 4 shows a significantly better $\text{MSE}_{\boldsymbol{A}}$ performance of MUA-SV for all three data cubes and SNR values when compared with the other algorithms. This indicates that MUA-SV effectively exploits the spatial properties of the abundance maps, even when the actual spectral variability does not follow exactly the model in (15).

Figs. 15a, 15b and 15c show the true and reconstructed abundance maps for all algorithms and 30 dB SNR. As expected, models accounting for spectral variability tend to yield better reconstruction quality than FCLS, with ELMM yielding piecewise smooth solutions. In general, the solution provided by MUA-SV approaches better the ground-truth, in that it estimates the intensity of the abundance maps with better accuracy than the other algorithms. This can be most clearly seen for the results for DC2 (Fig. 15b), where the regions with pure pixels are better represented by MUA-SV.

Regarding the spectral performances, as measured by the $\text{MSE}_{\mathscr{M}}$ and $\text{SAM}_{\mathscr{M}}$, the results varied among the algorithms, with no method performing uniformly better than the others. There is also a significant discrepancy between the Euclidean metric and the spectral angle in many examples, highlighting the different characteristics of the two metrics.

The ELMM model yielded the smallest reconstruction error $\text{MSE}_{\boldsymbol{Y}}$ in most cases (6), followed by MUA-SV (4 cases). However, the connection between the reconstruction error $\text{MSE}_{\boldsymbol{Y}}$ and the abundance estimation performance $\text{MSE}_{\boldsymbol{A}}$ of the unmixing methods that address spectral variability is not clear, as can be attested from Table 4.

The execution times shown in Table 5 indicate that MUA-SV is 2.2 times faster than ELMM and 7.5 times faster than PLMM, a significant gain in computational efficiency. This difference is more accentuated when processing larger datasets, as will be verified in the following.

4.8.2 Sensitivity analysis

To evaluate the sensitivity of MUA-SV $\text{MSE}_{\boldsymbol{A}}$ to variations in the algorithm parameters[3], we initially set all regularization parameters (λ_M, λ_A, λ_Θ and ρ) equal to their optimal values[4]. Then, we varied one parameter at a time within a range from -95% to $+95\%$ of its optimal value. Fig. 17 presents the $\text{MSE}_{\boldsymbol{A}}$ values obtained by varying each parameter. It can be seen that small variations about the optimal values do not affect the $\text{MSE}_{\boldsymbol{A}}$ significantly, and that the maximum values obtained for the whole parameter ranges tested are still lower than those achieved by the other algorithms.

To evaluate the sensitivity of the MUA-SV results to variations in the SLIC parameters, we plotted the resulting $\text{MSE}_{\boldsymbol{A}}$ as a function of $\sqrt{N/S}$ and γ, with the algorithm parameters λ_M, λ_A, λ_Θ and ρ fixed at their optimal values. The results are shown in Fig. 17. It is seen that the $\text{MSE}_{\boldsymbol{A}}$ performance does not deviate significantly from its optimal value unless the superpixel size $\sqrt{N/S}$ becomes too large. This is expected since very large superpixels may contain semantically different pixels, hindering the capability of the transform $\boldsymbol{W}$ to adequately capture coarse scale information. Furthermore, large values of $\sqrt{N/S}$ may violate hypothesis A2, which has been used thoroughly in the derivation of the MUA-SV algorithm, and thus represent a bad design choice.

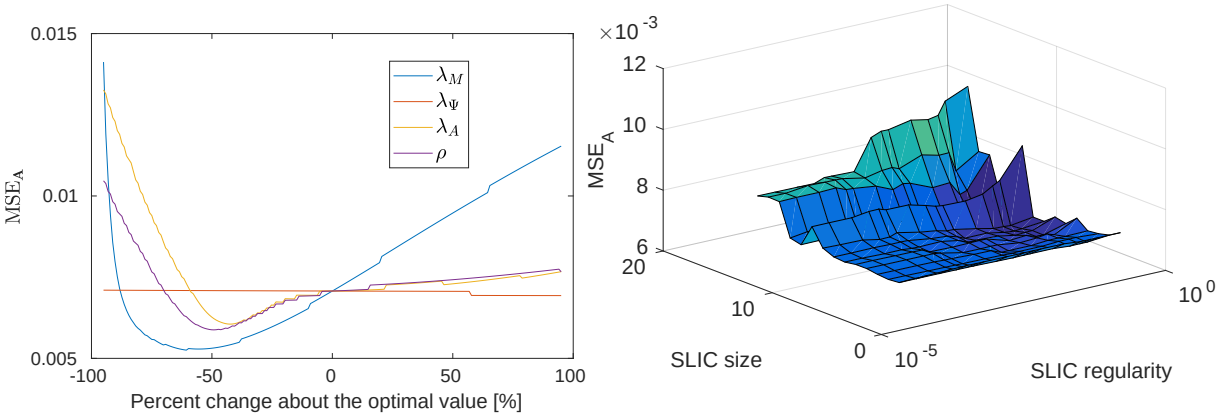


Figure 17 – $\text{MSE}_{\boldsymbol{A}}$ variation due to relative changes in each parameter value about its optimal value (left) and $\text{MSE}_{\boldsymbol{A}}$ as a function of SLIC parameters $\sqrt{N/S}$ and γ (right).

4.8.3 Simulations with real images

In this experiment, we consider two data sets obtained from real hyperspectral images. The first data set is comprised of a 152×108 pixels subset of the Houston hyperspectral image, with 144 spectral bands. The second data set is a 250×191 pixels subregion of the Cuprite image, with 188 spectral bands. Spectral bands presenting water absorption

[3] For conciseness, we present only the results for the DC1 data cube with a 30 dB SNR.

[4] Note that the operating point of MUA-SV is not optimal for this case due to the relatively coarse grid employed in the parameter search procedure.

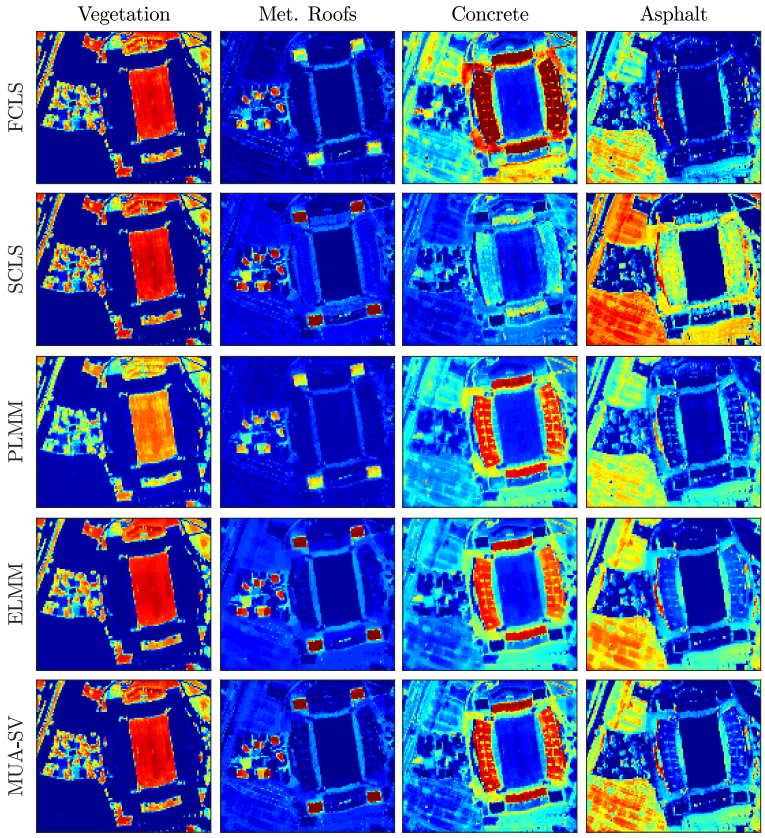


Figure 18 – Reconstructed fractional abundance maps for the Houston data set.

and low SNR were removed from both images. The parameters of the algorithms were selected empirically for the proposed method, and set identically to those reported in [19] for the ELMM and PLMM. The number of endmemebrs was selected as $P = 4$ for the Houston data set, and as $P = 14$ for the Cuprite data set, following the observations in [19]. The endmembers were extracted using the VCA algorithm [131].

Since the true abundance maps are unavailable for those hyperspectral images, we make a qualitative assessment of the recovered abundance maps based on knowledge of materials present in prominent fashion in those scenes. The reconstructed abundance maps for the Houston data set are depicted in Fig. 18. The four materials which are prominently present in this dataset are vegetation, red metallic roofs, concrete stands, and asphalt. It can be seen that ELMM and MUA-SV yield the best results for the overall abundances of all materials, with smaller proportion indeterminacy in regions known to have mostly pure materials such as the football field, the square metallic roofs and the concrete stands in the stadium. However, MUA-SV provides better results, more clearly observed in the purer areas such as the concrete stands of the stadium, which appear to be more mixed with the asphalt abundances in the ELMM results. This evidences the better performance of the MUA-SV algorithm.

The reconstructed abundance maps for the Alunite, Sphene, Buddingtonite and Muscovite materials of the Cuprite data set are depicted in Fig. 19. Although all methods

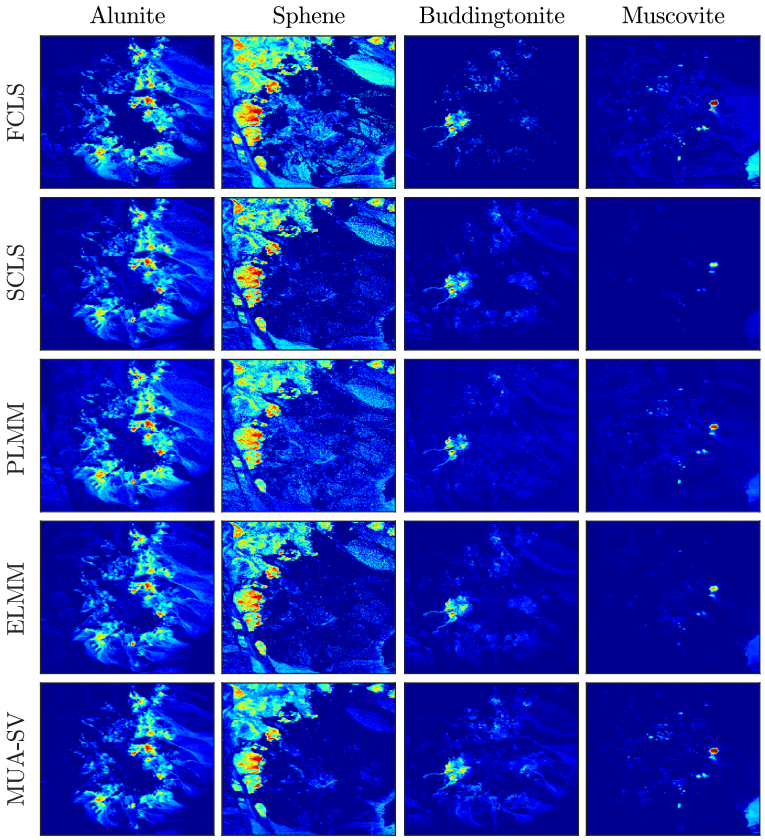


Figure 19 – Reconstructed fractional abundance maps for the Cuprite data set.

provide abundance maps which generally agree with previous knowledge about their distribution in this image [131], the MUA-SV results show abundances for all endmembers in Fig. 19 that are more homogeneous and clearly delineated in the regions where the materials are present. Moreover, these results show significantly smaller contributions due to outliers in the background regions of the abundance maps.

The reconstruction errors for all algorithms and both data sets are shown in Table 6. For the Houston data, the ELMM and MUA-SV results are very close and significantly smaller than those of the other methods, what agrees with their better representation of the abundance maps. For the Cuprite data, the errors are small and comparable for all algorithms, except for a slightly larger PLMM error. This goes in line with the fact that the abundance maps generally agree with the known distribution of these materials in the scene. However, reconstruction error results should be taken with proper care, as observed in the examples using synthetic data. Their correlation with the quality of abundance estimation is far from straightforward.

The execution times for all methods, shown in Table 5, illustrate again the significantly smaller computational load of MUA-SV when compared to other methods addressing spectral variability, as it performed, on average, 5.3 times faster than ELMM and 64.3 times faster than PLMM.

4.9 CONCLUSIONS

In this chapter we proposed a new data-dependent multiscale model for spectral unmixing accounting for spectral variability of the endmembers. Using a multiscale transformation based on the superpixel decomposition, spatial contextual information was incorporated into the unmixing problem through the decomposition of the observation model into two models in different domains, one capturing coarse image structures and another representing fine scale details. This facilitated the characterization of spatial regularity. Under reasonable assumptions, the proposed method yields a fast iterative algorithm, in which the abundance estimation problem is solved only once in each scale. Simulation results with both synthetic and real data show that the proposed MUA-SV algorithm outperforms other methods addressing spectral variability, both in accuracy of the reconstructed abundance maps and in computational complexity.

4.10 APPENDIX FROM CHAPTER 4: DERIVATION OF THE APPROXIMATED MIXING MODEL

Given the coarse pixel model in (34) can be approximated using hypothesis A2 as

$$\begin{aligned}
\boldsymbol{y}_{\mathcal{C}_i} &\approx \sum_{\ell=1}^{N} \frac{\mathbb{1}_{W_{\ell,i}}}{|\operatorname{supp}_\ell(W_{\ell,i})|} \boldsymbol{M}_\ell \sum_{j=1}^{N} W_{j,i}\, \boldsymbol{a}_i + \boldsymbol{e}_{\mathcal{C}_i} \\
&= \sum_{\ell=1}^{N} \frac{\mathbb{1}_{W_{\ell,i}}}{|\operatorname{supp}_\ell(W_{\ell,i})|} \boldsymbol{M}_\ell\ \boldsymbol{a}_{\mathcal{C}_i} + \boldsymbol{e}_{\mathcal{C}_i} \\
&= \boldsymbol{M}_{\mathcal{C}_i} \boldsymbol{a}_{\mathcal{C}_i} + \boldsymbol{e}_{\mathcal{C}_i}
\end{aligned} \tag{54}$$

where $\boldsymbol{a}_{\mathcal{C}_i} = \sum_{j=1}^{N} W_{j,i} \boldsymbol{a}_i$. The detail model in (35) can be approximated as

$$\begin{aligned}
\boldsymbol{y}_{\mathcal{D}_i} &= \boldsymbol{M}_i \boldsymbol{a}_i - \sum_{j=1}^{S} \sum_{\ell=1}^{N} W^*_{j,i} W_{\ell,j}\, \boldsymbol{M}_\ell\, \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i} \\
&\approx \boldsymbol{M}_i \boldsymbol{a}_i - \left(\sum_{n=1}^{S} \sum_{m=1}^{N} \frac{\mathbb{1}_{W^*_{n,i}} \mathbb{1}_{W_{m,n}}}{|\operatorname{supp}_{n,m}(W^*_{n,i} W_{m,n})|} \boldsymbol{M}_m \right) \sum_{j=1}^{S} \sum_{\ell=1}^{N} W^*_{j,i} W_{\ell,j}\, \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i \boldsymbol{a}_i - \boldsymbol{M}_{\mathcal{C}^*_i} \sum_{j=1}^{S} \sum_{\ell=1}^{N} W^*_{j,i} W_{\ell,j}\, \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i}
\end{aligned} \tag{55}$$

and straightforward computations leads to

$$
\begin{aligned}
\boldsymbol{y}_{\mathcal{D}_i} &\approx \boldsymbol{M}_i\boldsymbol{a}_i - \boldsymbol{M}_{\mathcal{C}_i^*} \sum_{j=1}^{S}\sum_{\ell=1}^{N} W_{j,i}^* W_{\ell,j}\, \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i(\boldsymbol{a}_{\mathcal{D}_i} + \sum_{j=1}^{S} W_{j,i}^*\, \boldsymbol{a}_{\mathcal{C}_j}) - \boldsymbol{M}_{\mathcal{C}_i^*} \sum_{j=1}^{S}\sum_{\ell=1}^{N} W_{j,i}^* W_{\ell,j}\, \boldsymbol{a}_\ell + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i\left(\boldsymbol{a}_{\mathcal{D}_i} + \sum_{j=1}^{S} W_{j,i}^*\, \boldsymbol{a}_{\mathcal{C}_j}\right) - \boldsymbol{M}_{\mathcal{C}_i^*} \sum_{j=1}^{S} W_{j,i}^*\, \boldsymbol{a}_{\mathcal{C}_j} + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i\boldsymbol{a}_{\mathcal{D}_i} + \left(\boldsymbol{M}_i - \boldsymbol{M}_{\mathcal{C}_i^*}\right) \sum_{j=1}^{S} W_{j,i}^*\, \boldsymbol{a}_{\mathcal{C}_j} + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i\boldsymbol{a}_{\mathcal{D}_i} + \left(\boldsymbol{M}_i - \boldsymbol{M}_{\mathcal{C}_i^*}\right)\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_i + \boldsymbol{e}_{\mathcal{D}_i} \\
&= \boldsymbol{M}_i\boldsymbol{a}_{\mathcal{D}_i} + \boldsymbol{M}_{\mathcal{D}_i}\left[\boldsymbol{A}_{\mathcal{C}}\boldsymbol{W}^*\right]_i + \boldsymbol{e}_{\mathcal{D}_i}. \qquad (56)
\end{aligned}
$$

where $\boldsymbol{a}_{\mathcal{D}_i} = \boldsymbol{a}_i - \sum_{j=1}^{S} W_{j,i}^* \boldsymbol{a}_{\mathcal{C}_j}$.

4.11 APPENDIX FROM CHAPTER 4: NUMERICAL VERIFICATION OF THE SIMPLIFYING HYPOTHESIS

Although hypotheses A1 and A2 impose some limitation to the MUA-SV algorithm, they are reasonable and are satisfied in many practical circumstances. Below, we present a more thorough analysis of each of these hypotheses.

Hypothesis A1 consists of assuming that the inner product $\langle RE_{\mathcal{C}}, RE_{\mathcal{D}}\rangle$ between the residuals/reconstruction errors $RE_{\mathcal{C}}$ and $RE_{\mathcal{D}}$ in the coarse and detail image scales is comparatively small, when compared to the first two terms of the cost function (31). To illustrate the validity of this claim, we compare here the values of $\langle RE_{\mathcal{C}}, RE_{\mathcal{D}}\rangle$ with those of the first two terms of the cost function, given by $\|RE_{\mathcal{C}}\|_F^2$ and $\|RE_{\mathcal{D}}\|_F^2$, for some practical examples. We considered the result of unmixing DC1, DC2 and DC3 with an SNR of 30 dB presented in Section 4.8 using the ELMM model. The results are presented below in Table 7. It can be seen that the quadratic norms exceed this inner product in value by several orders of magnitude. Thus, the latter can be reasonably neglected, i.e., $\langle RE_{\mathcal{C}}, RE_{\mathcal{D}}\rangle \approx 0$.

Table 7 – Comparison between the residuals inner product and the first two terms of the cost function

	$\|RE_{\mathcal{C}}\|_F^2 + \|RE_{\mathcal{D}}\|_F^2$	$\langle RE_{\mathcal{C}}, RE_{\mathcal{D}}\rangle$
DC1	328.35	-1.316×10^{-15}
DC2	0.5605	2.845×10^{-16}
DC3	7.105×10^{-4}	1.948×10^{-19}

Hypothesis A2 basically states that the endmember signatures for each pixel $\boldsymbol{M}_n$ do not deviate much from the average endmember signature in its neighborhood, i.e. $\boldsymbol{M}_n$ is similar to $\frac{1}{|\mathcal{N}_n|}\sum_{j\in\mathcal{N}_n}\boldsymbol{M}_j$ where $\mathcal{N}_n$ contains indexes of pixels that are spatially close to pixel n. This is an assumption about the underlying physical model that is reasonable in practical scenarios. To illustrate this, we consider two experiments, one based on the Hapke model and another based on real data, based on the discussion of Chapter 2.

For instance, using synthetic data generated using the Hapke model [37] we can represent spectral variability due to topographic variations of the scene. Consider the discrete terrain model and reference endmember signatures presented in Figure 20 below, extracted from [19]. From this data and using the Hapke model, one can generate a set of pixel dependent endmember signatures which can be used to evaluate the spatial characteristics of spectral variability. For simplicity, we measure the similarity between the reference and the pixel dependent endmember signatures using both the Euclidean distance and the spectral angle, for all materials. The results are shown in Figure 21 below, where it can be seen that there these deviations show significant spatial correlation.

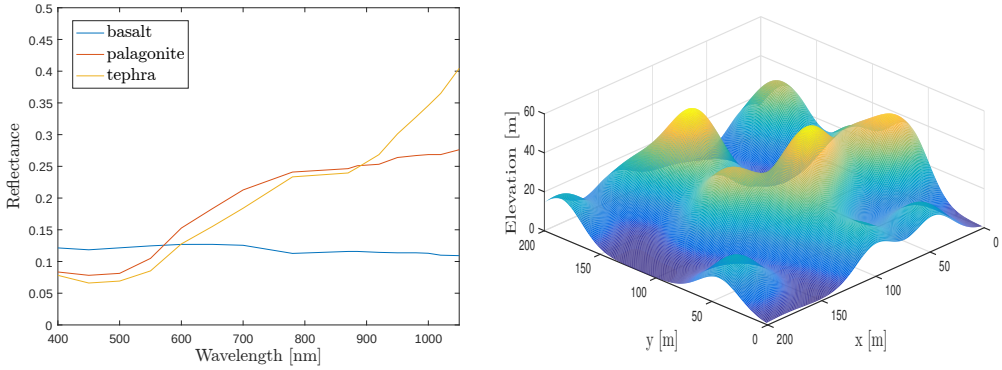


Figure 20 – Reference endmember signatures (left) and discrete terrain model (right) used with the Hapke model in the data cube DC3 to generate the pixel-dependent endmember signatures (data provided by [19]).

Spectral variability occurring due to intrinsic variations of the material spectra (e.g. soil or vegetation) can also show significant spatial correlation (see [105]), since endmember spectra usually depends on physical quantities that are correlated in space. Many experimental studies support this claim, including geostatistical works evaluating the spatial distribution and variability of soil's physico-chemical properties (e.g. for grass crop terrain (see [106]), calcareous soils (see [107]), rice fields (see [108]) and tobacco plantations (see [109]), and also measurements of mineral spectra due to the presence of spatially correlated grain sizes and impurity concentrations (see [86, 87]).

To illustrate this effect, we performed an experiment considering real data using the Samsom image. We considered a subregion containing pure pixels of the soil material, shown in Figure 22-(a) below. We considered these pixels as pixel dependent endmember signatures and evaluated the similarity between them and the average endmember spectra

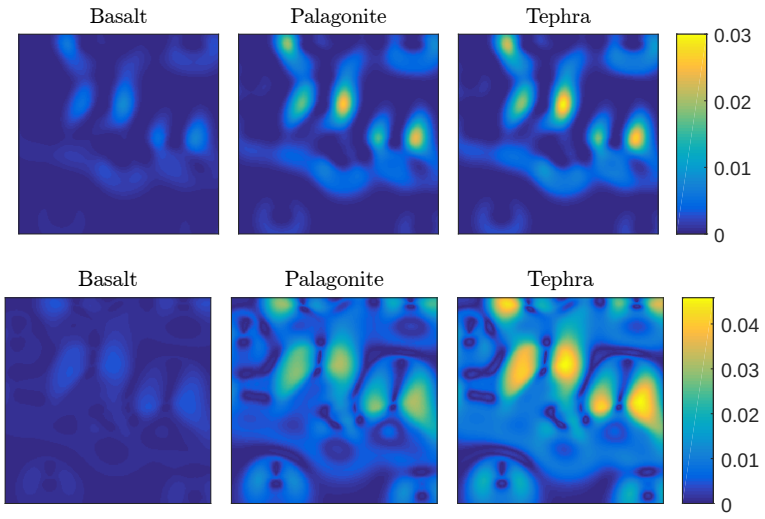


Figure 21 – Measures of endmember spatial variability in the Hapke model. Top row: Euclidean distance between the soil spectral signature of each pixel and the reference signature. Bottom row: Spectral angle between the soil spectral signature of each pixel and the reference signature.

for all these pixels. The results, shown in Figures 22-(b) and 22-(c), are similar to the Hapke data, and illustrate that the variability shows considerable spatial correlation.

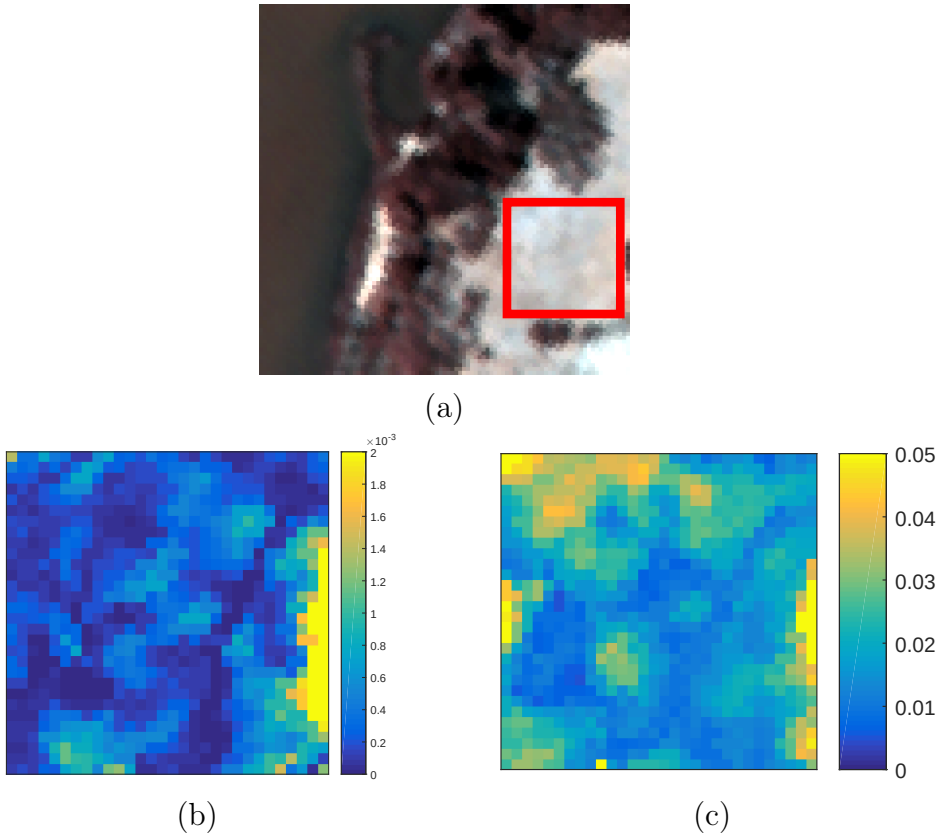


Figure 22 – (a) Samson hyperspectral image with a subimage containing soil highlighted. (b) Euclidean distance between the soil spectral signature of each pixel and their average value. (c) Spectral angle between the soil spectral signature of each pixel and their average value.

INTERMEZZO: NEW MODELS FOR ENDMEMBER SPECTRA: FROM TENSOR TO NEURAL-NETWORK-BASED REPRESENTATIONS

In Chapters 3 and 4, spatial information about the abundances was considered in order to improve the SU results with spectral variability, both in blind and in library-aided scenarios. A complementary approach consists of devising appropriate, low-dimensional models to represent the variability of the endmember spectral signatures. Devising appropriate models for the EMs (or, alternatively, regularizations for the optimization problems) is critical in order to effectively address spectral variability in SU, since the spatially varying EM matrices in model (2) make the problem severely ill-posed in general.

Although statistical and library-based models were the predominant approaches to represent the endmembers in SU [10], parametric EM models attracted significant interest more recently (see, e.g., [19, 20]). Such models can be physically motivated and yield good abundance estimation performance. Moreover, SU can often be performed using computationally efficient algorithms. However, parametric models also pose a significant challenge, as they must be carefully designed for their potential to be fully explored.

The following chapters are dedicated to devising new parametric EM models and applying them to the SU problem. We begin by extending the ELMM model [19] to consider band dependent scaling factors in Chapter 5, giving it more flexibility to represent complex spectral variability. In Chapter 6 we propose to explore the high-order (multidimensional) structure of the endmembers in the HI, when organized as tensor. The endmember tensor is assumed to be of approximately low rank, which greatly reduces the number of degrees of freedom and regularizes the SU problem.

In Chapter 7, we represent the spectral signatures of the endmembers using the perspective of deep generative modeling. We represent the manifold of spectral signatures of each EM as the image of an unknown function (learned from the observed HI) acting on a low-dimensional latent domain. This representation of the EM signatures is then used to parametrize the SU problem in a matrix factorization-inspired framework. This approach is later extended in Chapter 8 to augment spectral libraries known a priori, blending deep learning with library-based SU approaches.

5 GENERALIZED LINEAR MIXING MODEL ACCOUNTING FOR ENDMEMBER VARIABILITY

5.1 INTRODUCTION

Using parametric representations for the endmember spectra is becoming a popular approach to address spectral variability since this approach does not require spectral libraries and can provide computationally efficient solutions. The Perturbed LMM (PLMM) [20] introduced an additive perturbation to the endmember matrix for each pixel that needs to be estimated jointly with the abundances. Though the perturbation matrix can model arbitrary endmember variations, it lacks physical motivation. The Extended LMM (ELMM) [19] uses one pixel-dependent multiplicative term for each endmember, a generalization that can efficiently model changes in the observed reflectances due to illumination, an important effect [19]. This model addresses a physically motivated problem, with the advantage of estimating a variability parameter vector of much lower dimension when compared with the additive perturbation matrix in the PLMM.

Although the ELMM performs well in situations where spectral variability is mainly caused by illumination variations, it lacks a necessary flexibility when the endmembers are subject to more complex spectral distortions. For instance, experimental measurements on vegetation spectra under different conditions have shown a significant dependence of the spectral variation on wavelength intervals [94, 95, 93]. This type of variability is not supported by the ELMM model [19], which assumes a fixed scaling across all wavelengths.

In this chapter we introduce a generalization of the ELMM model proposed in [19] to account for endmember variability in arbitrary regions of the measured spectrum. We call the resulting model the Generalized Linear Mixing Model (GLMM). The estimation of the required parameters is realized by generalizing the methodology used in [19] through the use of three-dimensional tensors to accommodate the new model without significantly effecting the simplicity of the proposed solution or its computational complexity. Simulation results using synthetic and real data indicate that the extra flexibility introduced by the GLMM model can improve the results of existing methods for different types of endmember variability.

This chapter is organized as follows. In Section 5.2 we briefly revisit the ELMM model. Section 5.3 introduces the proposed GLMM. In Section 5.4 we define new tensor variables and extend the solution in [19] to the GLMM. The performance of the proposed method is compared with competing algorithms in Section 5.5. Finally, the conclusions are presented in Section 5.6.

5.2 EXTENDED LINEAR MIXING MODEL REVISITED

The LMM assumes that the endmember spectra are fixed for all pixels $\boldsymbol{y}_n$, $n = 1, \ldots, N$, in the HI. This assumption can jeopardize the accuracy of estimated abundances

in many circumstances due to the spectral variability existing in a typical scene. The ELMM [19] partially mitigates such limitation by including a vector of multiplicative weights $\boldsymbol{\phi}_n = \left[\phi_{1,n}, \ldots, \phi_{P,n}\right]^\top$ in the LMM such that

$$\boldsymbol{y}_n = \boldsymbol{M}_0 \operatorname{diag}(\boldsymbol{\phi}_n)\boldsymbol{a}_n + \boldsymbol{e}_n \,, \tag{57}$$

with $\phi_{k,n} \in \mathbb{R}_+$, $k = 1, \ldots, P$. Each coefficient $\phi_{k,n}$ scales the whole spectrum of endmember $\boldsymbol{m}_{0,k}$ in pixel n, leading to a simple strategy to model variability resulting from illumination effects.

5.3 GENERALIZED LINEAR MIXING MODEL (GLMM)

As explained in Section 5.1, we propose a generalization of the ELMM model to allow for spectral variabilities per wavelength intervals. To this end, we propose to employ a band-dependent scaling factor, enabling the new model to adapt to arbitrary variations of the endmember spectra. In the new GLMM model each pixel $\boldsymbol{y}_n$ is written as

$$\boldsymbol{y}_n = \big(\boldsymbol{M}_0 \odot \boldsymbol{\Psi}_n\big)\boldsymbol{a}_n + \boldsymbol{e}_n \,, \tag{58}$$

where $\boldsymbol{\Psi}_n \in \mathbb{R}^{L\times P}$ is a scaling matrix with nonnegative entries, and $\odot$ is the Hadamard (elementwise) product. This model is a generalization of the ELMM where the scaling matrix $\boldsymbol{\Psi}_n$ acts on each wavelength of each endmember individually. Such feature leads to a more flexible model that allows to consider variabilities that are not uniform along each endmember spectrum. ELMM is clearly a particular case of GLMM, and the new model can be employed for any level of granularity of variability per wavelength ranges, to the limit of an independent scaling of each wavelength component of each endmember in each pixel. Moreover, model (58) directly relates the amount of spectral variability at each wavelength of each EM to the corresponding amplitude of its spectrum, which agrees with experimental observations in many applications [94, 93].

5.4 THE UNIMIXING PROBLEM

Assuming the availability of a reference endmember matrix $\boldsymbol{M}_0$ (which can be obtained using any endmember extraction method), the SU problem reduces to estimating the free parameters minimizing a given risk functional defined for the whole HI $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]$. For this purpose the methodology presented in [19] can be extended for the GLMM by defining three-dimensional tensors. For the interested reader, a more in-depth overview of tensors and multilinear albegra will be presented in Section 6.2. Thus, we

propose to minimize the following regularized cost functional:

$$\begin{aligned}\mathcal{J}(\boldsymbol{A},\mathscr{M},\Psi) = \frac{1}{2}\sum_{n=1}^{N}\Big(&\|\boldsymbol{y}_n - \boldsymbol{M}_n\boldsymbol{a}_n\|^2 \\ &+ \lambda_M\|\boldsymbol{M}_n - \boldsymbol{M}_0 \odot \boldsymbol{\Psi}_n\|_F^2\Big) \\ &+ \lambda_A\mathcal{R}(\boldsymbol{A}) + \lambda_\Psi\mathcal{R}(\Psi).\end{aligned} \tag{59}$$

where $\mathscr{M}$ and Ψ are $L \times P \times N$ tensors with entries $[\mathscr{M}]_{:,:,n} = \boldsymbol{M}_n$, and $[\Psi]_{:,:,n} = \boldsymbol{\Psi}_n$ respectively, $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N]$ is the abundance matrix, $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\Psi)$ are spatial regularizations over $\boldsymbol{A}$ and Ψ, and the parameters λ_M, λ_A and λ_Ψ control the contribution of each term in the cost function. Thus, the optimization problem becomes

$$\begin{aligned}\big(\widehat{\boldsymbol{A}}, \widehat{\mathscr{M}}, \widehat{\Psi}\big) = \underset{\boldsymbol{A},\,\mathscr{M},\,\Psi}{\arg\min}\ \ &\mathcal{J}(\boldsymbol{A},\mathscr{M},\Psi) \\ &\text{subject to } \ \boldsymbol{A} \geqslant \boldsymbol{0},\ \boldsymbol{1}^\top\boldsymbol{A} = \boldsymbol{1}^\top,\ \mathscr{M} \geqslant \mathbb{0}\,.\end{aligned} \tag{60}$$

The problem defined in (60) is non-smooth and non-convex with respect to all variables $\boldsymbol{A}$, $\mathscr{M}$, and Ψ, but is convex with respect to each one of them. Thus, we follow the same approach used in [19] and find a local stationary point minimizing (60) iteratively with respect to each variable, leading to the strategy presented in Algorithm 4.

Algorithme 4 : Global algorithm for solving (59)

Input : $\boldsymbol{Y}$, λ_M, λ_A, λ_Ψ, $\boldsymbol{A}^{(0)}$, $\Psi^{(0)}$ and $\boldsymbol{M}_0$.
Output : $\widehat{\boldsymbol{A}}$, $\widehat{\mathscr{M}}$ and $\widehat{\Psi}$.

Set $i = 0$;
while *stopping criterion is not satisfied* **do**
  $i = i + 1$;
  $\mathscr{M}^{(i)} = \underset{\mathscr{M}}{\arg\min}\ \ \mathcal{J}(\boldsymbol{A}^{(i-1)}, \mathscr{M}, \Psi^{(i-1)})$;
  $\boldsymbol{A}^{(i)} = \underset{\boldsymbol{A}}{\arg\min}\ \ \mathcal{J}(\boldsymbol{A}, \mathscr{M}^{(i)}, \Psi^{(i-1)})$;
  $\Psi^{(i)} = \underset{\Psi}{\arg\min}\ \ \mathcal{J}(\boldsymbol{A}^{(i)}, \mathscr{M}^{(i)}, \Psi)$;
end
return $\widehat{\boldsymbol{A}} = \boldsymbol{A}^{(i)}$, $\widehat{\mathscr{M}} = \mathscr{M}^{(i)}$, $\widehat{\Psi} = \Psi^{(i)}$;

The regularization functionals $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\Psi)$ in (59) are selected in order to provide spatial smoothness to the abundances and scaling factors. They are selected as

$$\mathcal{R}(\boldsymbol{A}) = \|\mathcal{H}_h(\boldsymbol{A})\|_{2,1} + \|\mathcal{H}_v(\boldsymbol{A})\|_{2,1} \tag{61}$$

and

$$\mathcal{R}(\Psi) = \frac{1}{2}\sum_{\ell=1}^{L}\sum_{k=1}^{P}\left(\|\mathcal{H}_h([\Psi]_{\ell,k,:})\|_F^2 + \|\mathcal{H}_v([\Psi]_{\ell,k,:})\|_F^2\right), \tag{62}$$

where $[\cdot]_{\ell,k,:}$ is a slice of a tensor for band ℓ, endmember k and all N pixels. The linear operators $\mathcal{H}_h$ and $\mathcal{H}_v$ compute the first-order horizontal and vertical gradients of a bidimensional signal, acting separately for each material of $\boldsymbol{A}$. The spatial regularization in the abundances is promoted by a mixed $\mathcal{L}_{2,1}$ norm of their gradient, where $\|\boldsymbol{X}\|_{2,1} = \sum_{n=1}^{N} \|\boldsymbol{x}_n\|_2$. This norm is used to promote sparsity of the gradient across different materials (i.e. to force neighboring pixels to be homogeneous in all constituent endmembers). The $\mathcal{L}_1$ norm can also be used, leading to the Total Variation regularization [133].

5.4.1 Optimization with respect to $\mathscr{M}$

Rewriting the problem (60) using only the terms in (59) that depend on $\mathscr{M}$, the problem becomes

$$\widehat{\mathscr{M}} = \underset{\mathscr{M} \geqslant \mathbf{0}}{\arg\min} \frac{1}{2} \sum_{n=1}^{N} \left(\|\boldsymbol{y}_n - \boldsymbol{M}_n \boldsymbol{a}_n\|^2 - \lambda_M \|\boldsymbol{M}_n - \boldsymbol{M}_0 \odot \boldsymbol{\Psi}_n\|_F^2 \right). \tag{63}$$

The problem in (63) can be solved individually for each pixel $\boldsymbol{y}_n$. Thus, relaxing the positivity constraint on the elements of $\mathscr{M}$, the solution can be found as

$$\widehat{\boldsymbol{M}}_n = \left(\boldsymbol{y}_n \boldsymbol{a}^\top + \lambda_M \boldsymbol{M}_0 \odot \boldsymbol{\Psi}_n \right) \left(\boldsymbol{a}_n \boldsymbol{a}_n^\top + \lambda_M \boldsymbol{I}_P \right)^{-1}, \tag{64}$$

where $\boldsymbol{I}_P$ is the $P \times P$ identity matrix. Then, an approximate solution to the constrained problem can be obtained by projecting $\widehat{\boldsymbol{M}}_n$ onto the nonnegative orthant $\mathbb{R}_+^{L \times P}$ by thresholding the negative entries to zero [19].

5.4.2 Optimization with respect to $\boldsymbol{A}$

By restating the problem (60) only considering the terms in (59) that depend on $\boldsymbol{A}$ and introducing the non-negativity and sum-to-one constraints as additive penalties in the cost function, the abundance estimation problem can be written as follows

$$\begin{aligned} \widehat{\boldsymbol{A}} = \; & \underset{\boldsymbol{A}}{\arg\min} \max_{\boldsymbol{\mu}} \frac{1}{2} \sum_{n=1}^{N} \|\boldsymbol{y}_n - \boldsymbol{M}_n \boldsymbol{a}_n\|^2 \\ & + \lambda_A \left(\|\mathcal{H}_h(\boldsymbol{A})\|_{2,1} + \|\mathcal{H}_v(\boldsymbol{A})\|_{2,1} \right) \\ & + \iota_+(\boldsymbol{A}) + \boldsymbol{\mu}^\top (\boldsymbol{A}^\top \mathbf{1}_{P \times 1} - \mathbf{1}_{N \times 1}), \end{aligned} \tag{65}$$

where $\iota_+(\cdot)$ is the indicator function of $\mathbb{R}_+$ (i.e. $\iota_+(a) = 0$ if $a \geqslant 0$ and $\iota_+(a) = \infty$ if $a < 0$) acting component-wise on its input, and enforces the abundances nonnegativity constraint, and $\boldsymbol{\mu} \in \mathbb{R}^N$ is a vector of Lagrange multipliers associated with the sum-to-one constraint.

This problem is clearly not separable with respect to the pixels in the image. However, problem (65) can be efficiently solved using the ADMM [134]. The procedure is well described in [19] and will be suppressed here for conciseness.

5.4.3 Optimization with respect to Ψ

Rewriting the optimization problem (60) considering only the terms in (59) that depend on Ψ leads to

$$\begin{aligned}\widehat{\Psi} = \arg\min_{\Psi} \frac{\lambda_M}{2} \sum_{n=1}^{N} \|\boldsymbol{M}_n - \boldsymbol{M}_0 \odot \boldsymbol{\Psi}_n\|_F^2 \\ + \frac{\lambda_\Psi}{2}(\|\mathcal{H}_h(\Psi)\|_F^2 + \|\mathcal{H}_v(\Psi)\|_F^2)\,,\end{aligned} \tag{66}$$

which can be rewritten as

$$\begin{aligned}\widehat{\Psi} = \arg\min_{\Psi} \frac{\lambda_M}{2} \sum_{\ell=1}^{L}\sum_{k=1}^{P} \left(\|[\mathscr{M}]_{\ell,k,:} - [\boldsymbol{M}_0]_{\ell,k}[\Psi]_{\ell,k,:}\|^2\right) \\ + \frac{\lambda_\Psi}{2}\left(\|\mathcal{H}_h([\Psi]_{\ell,k,:})\|_F^2 + \|\mathcal{H}_v([\Psi]_{\ell,k,:})\|_F^2\right)\,,\end{aligned} \tag{67}$$

where $[\cdot]_{\ell,k,:}$ is a slice of a tensor for band ℓ, endmember k and all N pixels. The problem can be solved for each endmember k and band ℓ individually, and its solution is given by

$$[\widehat{\Psi}]_{\ell,k,:} = \boldsymbol{B}^{-1}\left(\lambda_M[\boldsymbol{M}_0]_{\ell,k}[\mathscr{M}]_{\ell,k,:}\right)\,, \tag{68}$$

where $\boldsymbol{B} = \lambda_M\left([\boldsymbol{M}_0]_{\ell,k}\right)^2\boldsymbol{I}_N + \lambda_\Psi(\boldsymbol{H}_h^\top\boldsymbol{H}_h + \boldsymbol{H}_v^\top\boldsymbol{H}_v)$ and $\boldsymbol{H}_h$ and $\boldsymbol{H}_v$ are matrix-based representations of the operators $\mathcal{H}_h$ and $\mathcal{H}_v$, respectively. The solution in (68) involves the inverse of the $N \times N$ matrix $\boldsymbol{B}$ which can be computationally intensive or intractable. However, if we assume periodic boundary conditions for the differential operators $\mathcal{H}_h$ and $\mathcal{H}_v$, the corresponding matrices $\boldsymbol{H}_v$, $\boldsymbol{H}_h$ and consequently $\boldsymbol{B}$ will have the structure of a block circulant matrix with circulant blocks (BCCB). Since BCCB matrices can be diagonalized using the bi-dimensional Discrete Fourier Transform, problem (68) can be solved efficiently as follows [135]

$$[\widehat{\Psi}]_{\ell,k,:} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\lambda_M[\boldsymbol{M}_0]_{\ell,k}[\mathscr{M}]_{\ell,k,:})}{\lambda_M\left([\boldsymbol{M}_0]_{\ell,k}\right)^2\mathbf{1}_{N_1\times N_2} + \lambda_\Psi(|\mathcal{F}(\boldsymbol{h}_h)|^2 + |\mathcal{F}(\boldsymbol{h}_v)|^2)}\right)\,, \tag{69}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ represents the bi-dimensional discrete Fourier transform and its inverse respectively, $\mathbf{1}_{N_1\times N_2}$ is an $N_1 \times N_2$ matrix of ones, where N_1 and N_2 are the number of rows and columns of the HI cube, and $\boldsymbol{h}_h$ and $\boldsymbol{h}_v$ are convolution masks corresponding to operators $\mathcal{H}_v$ and $\mathcal{H}_h$, represented as $N_1 \times N_2$ matrices.

5.5 SIMULATIONS

In this section, the performance of the proposed methodology is illustrated through simulations with both synthetic and real data. We compare the proposed method based on the GLMM with the the FCLS, the SCLS, the ELMM [19], and the PLMM [20], which

Table 8 – Simulations with synthetic data.

	Data Cube 0 – DC0				Data Cube 1 – DC1				Houston Data
	$\text{RMSE}_{\boldsymbol{A}}$	$\text{RMSE}_{\mathscr{M}}$	$\text{SAM}_{\mathscr{M}}$	$\text{RMSE}_{\boldsymbol{Y}}$	$\text{RMSE}_{\boldsymbol{A}}$	$\text{RMSE}_{\mathscr{M}}$	$\text{SAM}_{\mathscr{M}}$	$\text{RMSE}_{\boldsymbol{Y}}$	$\text{RMSE}_{\boldsymbol{Y}}$
FCLS	0.0968	–	–	0.0420	0.0243	–	–	0.0385	0.006082
SCLS	0.0642	0.0673	0.0625	0.0403	0.0509	0.0457	0.0617	0.0383	0.006082
PLMM	0.0641	0.0689	**0.0566**	0.0250	0.0476	0.0401	0.0578	0.0257	0.002918
ELMM	0.0540	0.0608	0.0568	0.0254	0.0209	0.0425	0.0609	0.0367	0.003217
GLMM	**0.0512**	**0.0587**	0.0601	**0.0011**	**0.0193**	**0.0391**	**0.0564**	**0.0202**	**0.000313**

is based on an NMF approach. To measure the accuracy of the unmixing methods we consider the Root Mean Squared Error (RMSE)

$$\text{RMSE}_{\mathcal{X}} = \sqrt{\frac{1}{N_{\mathcal{X}}} \| \text{vec}(\mathcal{X}) - \text{vec}(\widehat{\mathcal{X}}) \|^2} \tag{70}$$

where $\text{vec}(\cdot)$ is the vectorization operator $\mathcal{X} \to \boldsymbol{x}$, $\mathbb{R}^{a \times b \times c} \mapsto \mathbb{R}^{abc}$, $N_{\mathcal{X}} = abc$. In this work we apply the RMSE to evaluate the estimates of the abundances ($\text{RMSE}_{\boldsymbol{A}}$), of the endmembers tensor ($\text{RMSE}_{\mathscr{M}}$) and of the reconstructed images ($\text{RMSE}_{\boldsymbol{Y}}$). We also consider the Spectral Angle Mapper for the endmembers tensor

$$\text{SAM}_{\mathscr{M}} = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{P} \arccos \left(\frac{\boldsymbol{m}_{k,n}^{\top} \widehat{\boldsymbol{m}}_{k,n}}{\|\boldsymbol{m}_{k,n}\| \|\widehat{\boldsymbol{m}}_{k,n}\|} \right). \tag{71}$$

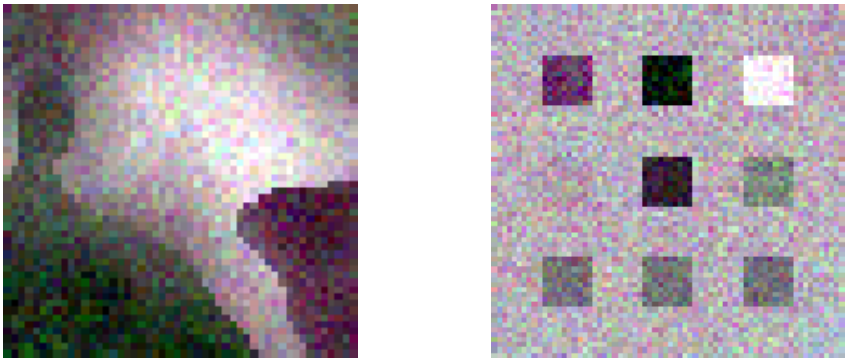


Figure 23 – Synthetic data cubes DC0, left, and DC1, right.

5.5.1 Synthetic data

For a comprehensive comparison among the different methods we created two synthetic datasets, namely Data Cube 0 (DC0) and Data Cube 1 (DC1), represented in Fig. 23. These datasets were built using endmembers extracted from the USGS Spectral Library [136], and different strategies were used to generate the abundance maps, which exhibit spatial correlation between neighboring pixels. For DC0, we adopted the variability model used in [19] (a multiplicative factor acting in each endmember), while for DC1 we considered the variability following the GLMM where correlation was imposed over $\boldsymbol{\Psi}_n$

using a 3-D Gaussian filter. White Gaussian noise was added to both datasets resulting in a SNR of 30 dB.

To find the optimal parameters for the selected algorithms we performed a grid search for each dataset. The parameter ranges were chosen based on the ranges tested and discussed by the authors in the original publication of each algorithm. For the PLMM we used $\gamma = 1$, since the authors fixed this parameter in all simulations, and searched for α and β in the range [0.35, 0.7, 1.4, 25] and [10^{-9}, 10^{-5}, 10^{-4}, 10^{-3}], respectively. For both the ELMM and the GLMM, the parameters were selected among the following values: λ_S, $\lambda_M \in$ [0.01, 0.1, 1, 5, 10, 15], $\lambda_A \in$ [0.001, 0.01, 0.05], and λ_ψ, $\lambda_\Psi \in$ [10^{-6}, 10^{-3}, 10^{-1}].

The results are presented in Table 8. In terms of RMSE for the abundance vectors, $\text{RMSE}_{\boldsymbol{A}}$, the proposed strategy clearly outperformed the competing algorithms for both datasets. This behavior can be verified for almost all metrics considered. The only exception is the $\text{SAM}_{\mathscr{M}}$ for DC0 where PLMM and ELMM presented smaller spectral angles. Regarding the increase of computational complexity introduced by the GLMM when compared with the ELMM, the simulations point out that the GLMM approach demanded $3.62 \times \text{Time}_{\text{ELMM}}$ for DC0 and $1.92 \times \text{Time}_{\text{ELMM}}$ for DC1, where $\text{Time}_{\text{ELMM}}$ is the CPU time elapsed during the ELMM unmixing process. The results show that the extra flexibility of the GLMM can be beneficial for the SU problem at the expense of a reasonable increase in the computational complexity.

5.5.2 Real data

For simulations with real data we considered the Houston dataset discussed in [19]. This dataset is known to have four endmembers which were extracted using the VCA algorithm [131]. Fig. 24 shows the reconstructed abundance maps for all tested methods while Table 8 presents the results in terms of $\text{RMSE}_{\boldsymbol{Y}}$ and $\text{SAM}_{\boldsymbol{Y}}$. Fig. 24 shows that the proposed GLMM method provided smooth and accurate abundance estimation, comparable with the results obtained using the ELMM. In fact, for the Concrete endmember, the GLMM abundance map shows stronger components in the stadium stands when compared with the other methods considering spectral variability. Although the results presented in Table 8 indicate better fitting for the GLMM method, these results should be taken with the proper care, since the connection of reconstruction error and abundance estimation is not straightforward. The GLMM demanded a computational time of $0.92 \times \text{Time}_{\text{ELMM}}$.

5.6 CONCLUSIONS

In this chapter, we proposed a new Generalized Linear Mixing Model (GLMM) that accounts for endmember spectral variability. The new model generalizes the ELMM to allow for the consideration of band dependent scaling factors for the endmember signatures. This way the GLMM model can represent a larger variety of realistic spectral variations

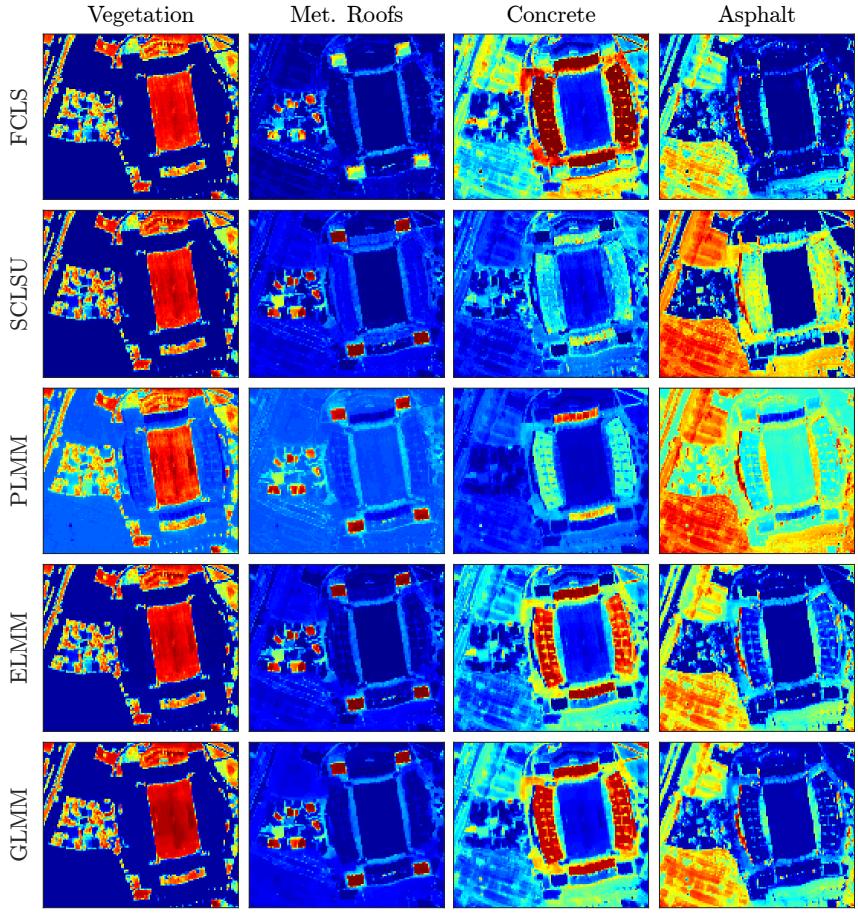


Figure 24 – Abundance maps of the Houston dataset for all tested algorithms where the abundance values are represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

of the endmembers, generalizing the representation capability of the ELMM. To solve the resulting optimization problem, we extended the variable splitting methodology used in [19] by including new tensor variables. Simulation results with both synthetic and real data suggest that the extra flexibility introduced by the GLMM can be beneficial for the unmixing process, resulting in improvements in both the abundance estimation and the reconstruction error.

6 LOW-RANK TENSOR MODELING FOR SPECTRAL UNMIXING ACCOUNTING FOR SPECTRAL VARIABILITY

"Wintermute was a simple cube of white light, that very simplicity suggesting extreme complexity."

William Gibson, *Neuromancer*

6.1 INTRODUCTION

SU algorithms that use parametric EM models (such as, e.g., [20, 19, 34]) or additive residual terms (such as, e.g., [137, 138]) to address endmember variability resort to different strategies to regularize the ill-posed optimization problem which leads to the estimation of abundances and endmembers. The regularization is achieved by introducing into the unmixing problem additional information based on common knowledge about the low-dimensionality of structures embedded in hyperspectral images.

Possible ways to recover lower-dimensional structures from noisy and corrupted data include the imposition of low-rank matrix constraints on the estimation process [139], or the low-rank decomposition of the observed data [140, 141]. The facts that HIs are naturally represented and treated as tensors, and that low-rank decompositions of higher-order (>2) tensors tend to capture homogeneities within the tensor structure make such strategies even more attractive for SU. Low-rank tensor models have been successfully employed in various tasks involving HIs, such as recovery of missing pixels [142], anomaly detection [143], classification [144], compression [145], dimensionality reduction [146] and analysis of multi-angle images [147]. More recently, [148] and [147] considered low-rank tensor decompositions applied to standard and multitemporal SU, respectively.

In [148] the HI is treated as three-dimensional tensor, and spatial regularity is enforced through a nonnegative tensor factorization (NTF) strategy that imposes a low-rank tensor structure. In [147], nonnegative canonical polyadic decomposition were used to unmix multitemporal HIs represented as three-dimensional tensors built by stacking multiple temporal matricized HIs. Though a low-rank tensor representation may naturally describe the regularity of HIs and abundance maps, the forceful introduction of stringent rank constraints may prevent an adequate representation of some image structures that are important for accurate unmixing. Another limitation of the approach proposed in [148] is the lack of guarantee that endmembers and abundances will be correctly factorized into their respective tensors. In [45], we proposed a new low-rank SU method called *Unmixing with Low-rank Tensor Regularization Algorithm* (ULTRA), which accounts for highly correlated endmembers. The SU problem was formulated using tensors and a low-rank abundance tensor regularization term was introduced. Differently, from the strict tensor decomposition considered in [148, 147], ULTRA allowed important flexibility to the rank

of the estimated abundance tensor to adequately represent fine scale structure and details that lie beyond a low-rank structure, but without compromising the regularity of the solution.

In this work we extend the strategy proposed in [45] to account for the important effect of endmember variability as well as a novel method to estimate the sufficient rank of a tensor for accurately solving the unmixing problem. Instead of introducing a rigid parametric model for the endmembers, we employ a more general tensor model which exploits the high dimensional structure of the problem, using a well-devised low-rank constraint to introduce regularity to the estimated endmember tensor. The main novel contributions of this chapter are:

a) We extend the strategy proposed in [45] by imposing a new low-rank regularization on the four-dimensional endmember tensor, which contains one endmember matrix for each pixel, to account for endmember variability. The new cost function results in an iterative algorithm, named *Unmixing with Low-rank Tensor Regularization Algorithm accounting for endmember Variability* (ULTRA-V). At each iteration, ULTRA-V updates the estimations of the abundance and endmember tensors as well as their low-rank approximations.

b) We propose a novel non-trivial strategy to determine the smallest rank representation that contains most of the variation of multilinear singular values [149].

Simulation results using synthetic and real data illustrate the performance improvement obtained using ULTRA-V when compared to competing methods, as well as its competitive computational complexity for relatively small images.

The chapter is organized as follows. Section 6.2 briefly reviews important background, definitions and notation used for tensors. Section 6.3 presents the proposed solution and the strategy to estimate tensor ranks. Section 6.4 presents the simulation results and comparisons. Finally, Section 6.5 presents the conclusions.

6.2 TENSORS BACKGROUND AND NOTATION

6.2.1 Notation

An order-R tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times \cdots \times N_R}$ $(R > 2)$ is an $N_1 \times \cdots \times N_R$ array with elements indexed by $\mathcal{T}_{n_1,n_2,\ldots,n_R}$. The R dimensions of a tensor are called *modes*. A mode-ℓ fiber of tensor $\mathcal{T}$ is the one-dimensional subset of $\mathcal{T}$ obtained by fixing all but the ℓ-th dimension, and is indexed by $\mathcal{T}_{n_1,\ldots,n_{\ell-1},:,n_{\ell+1},\ldots,n_R}$. A slab or slice of tensor $\mathcal{T}$ is a two-dimensional subset of $\mathcal{T}$ obtained by fixing all but two of its modes. An HI is often conceived as a three dimensional data cube, and can be naturally represented by an order-3 tensor $\mathcal{Y} \in \mathbb{R}^{N_1 \times N_2 \times L}$, containing $N_1 \times N_2$ pixels represented by the tensor fibers $\mathcal{Y}_{n_1,n_2,:} \in \mathbb{R}^L$. Analogously, the abundances can also be collected in an order-3 tensor $\mathcal{A} \in \mathbb{R}^{N_1 \times N_2 \times P}$.

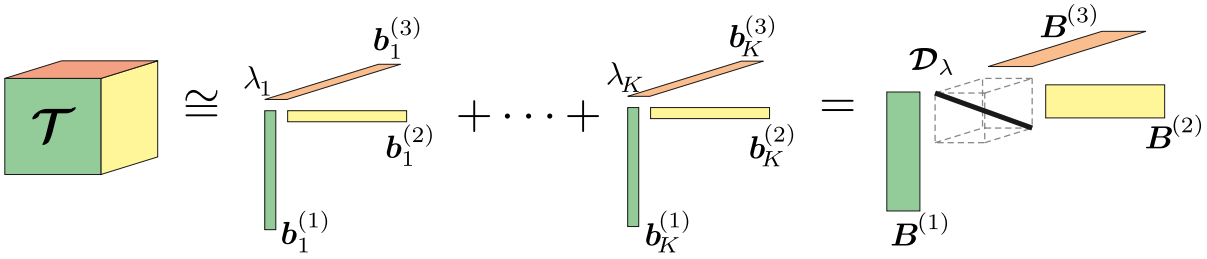


Figure 25 – Polyadic decomposition of a three-dimensional tensor, written as both outer products and mode$-n$ products.

Thus, given a pixel $\mathcal{Y}_{n_1,n_2,:}$, the respective abundance vector $\boldsymbol{a}_{n_1,n_2}$ is represented by the mode-3 fiber $\mathcal{A}_{n_1,n_2,:}$. Similarly, the endmember matrices for each pixel can be represented as an order-4 tensor $\mathcal{M} \in \mathbb{R}^{N_1 \times N_2 \times L \times P}$, where $\mathcal{M}_{n_1,n_2,:,:} = \boldsymbol{M}_{n_1,n_2}$. We now review some operations of multilinear algebra (the algebra of tensors) that will be used in the following sections (more details can be found in [150]).

6.2.2 Tensor product definitions

Definition 1. ***Outer product:*** *The outer product between vectors $\boldsymbol{b}^{(1)} \in \mathbb{R}^{N_1}, \boldsymbol{b}^{(2)} \in \mathbb{R}^{N_2}, \ldots, \boldsymbol{b}^{(R)} \in \mathbb{R}^{N_R}$ is defined as the order-R tensor $\mathcal{T} = \boldsymbol{b}^{(1)} \circ \boldsymbol{b}^{(2)} \circ \cdots \circ \boldsymbol{b}^{(R)} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_R}$, where $\mathcal{T}_{n_1,n_2,\ldots,n_R} = b^{(1)}_{n_1} b^{(2)}_{n_2} \cdots b^{(R)}_{n_R}$ and $b^{(i)}_{n_i}$ is the n_i-th position of $\boldsymbol{b}^{(i)}$. It generalizes the outer product between two vectors.*

Definition 2. ***Mode-k product:*** *The mode-k product, denoted $\mathcal{U} = \mathcal{T} \times_k \boldsymbol{B}$, of a tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times \cdots \times N_k \times \cdots \times N_R}$ and a matrix $\boldsymbol{B} \in \mathbb{R}^{M_k \times N_k}$ is evaluated such that each mode-k fiber of $\mathcal{T}$ is multiplied by matrix $\boldsymbol{B}$, yielding $\mathcal{U}_{n_1,n_2,\ldots,m_k,\ldots,n_R} = \sum_{i=1}^{N_k} \mathcal{T}_{\ldots,n_{n-1},i,n_{n+1},\ldots} \boldsymbol{B}_{m_k,i}$.*

Definition 3. ***Multilinear product:*** *The full multilinear product, denoted by $[\![\mathcal{T}; \boldsymbol{B}^{(1)}, \boldsymbol{B}^{(2)}, \ldots, \boldsymbol{B}^{(R)}]\!]$, consists of the successive application of mode-k products between $\mathcal{T}$ and matrices $\boldsymbol{B}^{(i)}$, represented as $\mathcal{T} \times_1 \boldsymbol{B}^{(1)} \times_2 \boldsymbol{B}^{(2)} \times_3 \ldots \times_R \boldsymbol{B}^{(R)}$.*

Definition 4. ***Mode-(M,1) contracted product:*** *The contracted mode-M product, denoted by $\mathcal{U} = \mathcal{T} \times^M \boldsymbol{b}$, is a product between a tensor $\mathcal{T}$ and a vector $\boldsymbol{b}$ in mode-M, where the resulting singleton dimension is removed, given by $\mathcal{U}_{\ldots,n_{n-1},n_{n+1},\ldots} = \sum_{i=1}^{N_n} \mathcal{T}_{\ldots,n_{n-1},i,n_{n+1},\ldots} \boldsymbol{b}_i$.*

6.2.3 The Canonical Polyadic Decomposition

An order-R rank-1 tensor is obtained as the outer product of R vectors. The rank of an order-R tensor $\mathcal{T}$ is defined as the minimum number of order-R rank-1 tensors that must be added to obtain $\mathcal{T}$ [141]. Thus, any tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_R}$ with $\text{rank}(\mathcal{T}) = K$ can be decomposed as a linear combination of at least K outer products of R rank-1 tensors. This so-called *polyadic decomposition* is illustrated in Fig. 25. When this decomposition

involves exactly K terms, it is called the canonical polyadic decomposition (CPD) [150] of a rank-K tensor $\mathcal{T}$, and is given by

$$\mathcal{T} = \sum_{i=1}^{K} \delta_i \boldsymbol{b}_i^{(1)} \circ \boldsymbol{b}_i^{(2)} \circ \cdots \circ \boldsymbol{b}_i^{(R)} . \tag{72}$$

It has been shown that this decomposition is essentially unique under mild conditions [141]. The CPD can be written alternatively using mode-k products as

$$\mathcal{T} = \mathcal{D}_\delta \times_1 \boldsymbol{B}^{(1)} \times_2 \boldsymbol{B}^{(2)} \cdots \times_R \boldsymbol{B}^{(R)} , \tag{73}$$

or using the full multilinear product as

$$\mathcal{T} = \left[\!\left[\mathcal{D}_\delta; \boldsymbol{B}^{(1)}, \boldsymbol{B}^{(2)}, \ldots, \boldsymbol{B}^{(R)} \right]\!\right] , \tag{74}$$

where $\mathcal{D}_\delta = \mathrm{TDiag}_R\left(\delta_1, \ldots, \delta_K\right)$ is the R-dimensional diagonal tensor and $\boldsymbol{B}^{(r)} = [\boldsymbol{b}_1^{(r)}, \ldots, \boldsymbol{b}_K^{(r)}]$, for $r = 1, \ldots, R$. Given a tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \times \cdots \times N_R}$, the CPD can be obtained as the solution to the following optimization problem [141]

$$\min_{\mathcal{D}_\delta, \boldsymbol{B}^{(1)}, \ldots, \boldsymbol{B}^{(K)}} \quad \frac{1}{2} \left\| \mathcal{T} - \sum_{i=1}^{K} \delta_i \boldsymbol{b}_i^{(1)} \circ \cdots \circ \boldsymbol{b}_i^{(R)} \right\|_F^2 . \tag{75}$$

A widely used strategy to compute an approximate solution to (75) is to use an alternating least-squares technique [141], which optimizes the cost function with respect to one term at a time, while keeping the others fixed, until convergence. Although optimization problem (75) is generally non-convex, its solution is unique under relatively mild conditions, which is an important advantage of tensor-based methods [141].

6.2.4 Tensor rank bounds

Finding the rank of an arbitrary tensor $\mathcal{T}$ is NP-hard [151]. In [141], upper and lower bounds on tensor ranks are presented for arbitrary tensors. Let $\mathcal{T}$ be an order-3 tensor and

$$\begin{aligned} F_1 &\equiv \dim \operatorname{span}\{\mathcal{T}_{:,j,k}\}_{\forall j,k} \\ F_2 &\equiv \dim \operatorname{span}\{\mathcal{T}_{i,:,k}\}_{\forall i,k} \\ F_3 &\equiv \dim \operatorname{span}\{\mathcal{T}_{i,j,:}\}_{\forall i,j} \end{aligned} \tag{76}$$

be the mode-1 (column), mode-2 (row) and mode-3 (fiber) ranks, respectively, of $\mathcal{T}$. Thus, the $K \equiv \mathrm{rank}(\mathcal{T})$ which is able to represent an arbitrary tensor is limited in the interval

$$\max(F_1, F_2, F_3) \leqslant K \leqslant \min(F_1 F_2, F_1 F_3, F_2 F_3). \tag{77}$$

The reader can note that the bounds presented above often lead to very large tensor ranks. In many practical applications, however, the "useful signal" rank is often much less than

the actual tensor rank [141]. Hence, when low-rank decompositions are employed to extract low-dimensional structures from the signal, the ranks that lead to meaningful results are usually much smaller than $\max(F_1, F_2, F_3)$. In Section 6.3.6, we propose a strategy to estimate the rank of tensor CPDs based on the variation of the multilinear singular values of $\mathcal{T}$.

6.3 LOW-RANK UNMIXING PROBLEM

An effective strategy to capture the low-dimensional structures of HIs for solving the SU problem is to impose a low-rank structure to the abundance tensor [148]. The same strategy can also be applied to the endmember tensor if one considers the endmember variabilities to be small or highly correlated in low-dimensional structures within the HI. The low-rank property of HI tensors has been an important tool in the design of hyperspectral image completion [152] and restoration algorithms [153], consisting in one of the main low-dimensional structures that are currently being considered in hyperspectral imaging applications. Thus, assuming that $\mathcal{A}$ has a low-rank $K_{\mathcal{A}}$, and that $\mathcal{M}$ has a low-rank $K_{\mathcal{M}}$ the global cost functional for the unmixing problem can be written as

$$\begin{aligned} \mathcal{J}(\mathcal{A}, \mathcal{M}) &= \frac{1}{2} \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \|\mathcal{Y}_{n_1,n_2,:} - \mathcal{M}_{n_1,n_2,:,:}\mathcal{A}_{n_1,n_2,:}\|_F^2 \\ \text{subject to} \quad & \operatorname{rank}(\mathcal{M}) = K_{\mathcal{M}},\ \mathcal{M} \geqslant \mathbf{0} \\ & \operatorname{rank}(\mathcal{A}) = K_{\mathcal{A}},\ \mathcal{A} \geqslant \mathbf{0},\ \mathcal{A} \times^3 \mathbf{1}_P = \mathbf{1}_{N_1 \times N_2}. \end{aligned} \tag{78}$$

Defining the SU problem as in (78) with fixed data independent ranks $K_{\mathcal{M}}$ and $K_{\mathcal{A}}$ limits its flexibility to adequately represent the desired abundance maps and endmember variability. Though fixing low ranks for $\mathcal{A}$ and $\mathcal{M}$ tends to capture the most significant part of the tensors energy [154], one may incur in a loss of fine and small scale details that may be relevant for specific data. On the other hand, using large values for $K_{\mathcal{A}}$ and $K_{\mathcal{M}}$ makes the solution sensitive to noise, undermining the purpose of regularization. Thus, an important issue is how to effectively impose the low-rank constraint to achieve regularity in the solution without undermining its flexibility to adequately model small variations and details.

We propose to modify (78) by introducing new regularization terms, controlled by two low-rank tensors $\mathcal{Q} \in \mathbb{R}^{N_1 \times N_2 \times P}$ and $\mathcal{P} \in \mathbb{R}^{N_1 \times N_2 \times L \times P}$, to impose non-strict constraints on $K_{\mathcal{A}}$ and $K_{\mathcal{M}}$. Doing that, tensors $\mathcal{Q}$ and $\mathcal{P}$ work as *a priori* information, and the strictness of the low-rank constraint is controlled by two additional parameters

$\lambda_{\mathscr{A}}$, $\lambda_{\mathscr{M}} \in \mathbb{R}_+$. The proposed cost function is given by

$$\begin{aligned} \mathcal{J}(\mathscr{A}, \mathscr{M}, \mathscr{P}, \mathbb{Q}) = \frac{1}{2} \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \|\mathscr{Y}_{n_1,n_2,:} - \mathscr{M}_{n_1,n_2,:,:} \mathscr{A}_{n_1,n_2,:}\|_F^2 \\ + \frac{\lambda_{\mathscr{M}}}{2} \|\mathscr{M} - \mathscr{P}\|_F^2 + \frac{\lambda_{\mathscr{A}}}{2} \|\mathscr{A} - \mathbb{Q}\|_F^2 \\ \text{subject to} \quad \mathscr{M} \geqslant \mathbf{0},\ \mathscr{A} \geqslant \mathbf{0},\ \mathscr{A} \times^3 \mathbf{1}_P = \mathbf{1}_{N_1 \times N_2} \end{aligned} \tag{79}$$

with $\text{rank}(\mathscr{P}) = K_{\mathscr{M}}$ and $\text{rank}(\mathbb{Q}) = K_{\mathscr{A}}$. The optimization problem becomes

$$(\widehat{\mathscr{A}}, \widehat{\mathscr{M}}, \widehat{\mathscr{P}}, \widehat{\mathbb{Q}}) = \underset{\mathscr{A}, \mathscr{M}, \mathscr{P}, \mathbb{Q}}{\arg\min} \ \mathcal{J}(\mathscr{A}, \mathscr{M}, \mathscr{P}, \mathbb{Q}) \,. \tag{80}$$

To solve (80), we propose to find a local stationary point by minimizing (79) iteratively with respect to each variable. The resulting algorithm is termed the *Unmixing with Low-rank Tensor Regularization Algorithm accounting for spectral Variability* (ULTRA-V), and is presented in Algorithm 1. The intermediate steps are detailed in the following.

Algorithme 5 : Global algorithm for solving (79)

Input : $\mathscr{Y}$, $\lambda_{\mathscr{M}}$, $\lambda_{\mathscr{A}}$, $\mathscr{A}^{(0)}$, and $\mathscr{M}^{(0)}$.
Output : $\widehat{\mathscr{A}}$ and $\widehat{\mathscr{M}}$.

```
K_Q = estimateTensorRank(A^(0));
K_P = estimateTensorRank(M^(0));
Set i = 0 ;
while stopping criterion is not satisfied do
    i = i + 1 ;
    P^(i) = argmin_P  J(A^(i-1), M^(i-1), P) ;
    Q^(i) = argmin_P  J(A^(i-1), M^(i-1), Q) ;
    M^(i) = argmin_M  J(A^(i-1), M, P^(i), Q^(i)) ;
    A^(i) = argmin_A  J(A, M^(i), P^(i), Q^(i)) ;
end
return Â = A^(i),  M̂ = M^(i);
```

6.3.1 Solving with respect to $\mathscr{A}$

To solve problem (80) with respect to $\mathscr{A}$ we use only the terms in (79) that depend on $\mathscr{A}$, leading to the cost function

$$\begin{aligned} \mathcal{J}(\mathscr{A}) = \frac{1}{2} \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} \|\mathscr{Y}_{n_1,n_2,:} - \mathscr{M}_{n_1,n_2,:,:} \mathscr{A}_{n_1,n_2,:}\|_F^2 + \frac{\lambda_{\mathscr{A}}}{2} \|\mathscr{A} - \mathbb{Q}\|_F^2 \\ \text{subject to} \quad \mathscr{A} \geqslant \mathbf{0},\ \mathscr{A} \times^3 \mathbf{1}_P = \mathbf{1}_{N_1 \times N_2} \,, \end{aligned} \tag{81}$$

which results in a standard regularized fully constrained least-squares problem that can be solved efficiently.

6.3.2 Solving with respect to $\mathscr{M}$

Analogously to the previous section, to solve problem (80) with respect to $\mathscr{M}$, we use only the terms in (79) that depend on $\mathscr{M}$, leading to

$$\begin{aligned} \mathcal{J}(\mathscr{M}) = \frac{1}{2}\sum_{n_1=1}^{N_1}\sum_{n_2=1}^{N_2} \|\mathscr{Y}_{n_1,n_2,:} - \mathscr{M}_{n_1,n_2,:,:}\mathscr{A}_{n_1,n_2,:}\|_F^2 + \frac{\lambda_{\mathscr{M}}}{2}\|\mathscr{M} - \mathscr{P}\|_F^2 \\ \text{subject to} \quad \mathscr{M} \geqslant \mathbf{0}. \end{aligned} \tag{82}$$

which results in a regularized nonnegative least-squares problem. An approximate solution can be obtained ignoring the positivity constraint over the endmember tensor and projecting the least-squares result onto the positive orthant as [34]

$$\begin{aligned} \widehat{\mathscr{M}}_{n_1,n_2,:,:} = \mathcal{P}_+\Bigg(\Big(\mathscr{Y}_{n1,n2,:}\mathscr{A}_{n1,n2,:}^\top + \lambda_{\mathscr{M}}\mathscr{P}_{n1,n2,:,:}\Big) \\ \Big(\mathscr{A}_{n1,n2,:}\mathscr{A}_{n1,n2,:}^\top + \lambda_{\mathscr{M}}\boldsymbol{I}\Big)^{-1}\Bigg) \end{aligned} \tag{83}$$

where $\mathcal{P}_+ : \mathbb{R}^{N_1\times N_1\times L} \to \mathbb{R}_+^{N_1\times N_1\times L}$ is the projection operator that maps every negative element to zero. Although this solution is approximate, it is significantly faster than directly solving (82) and the algorithm still demonstrated good empirical convergence in our experiments.

6.3.3 Solving with respect to $\mathscr{P}$

Rewriting the terms in (79) that depend on $\mathscr{P}$ leads to

$$\mathcal{J}(\mathscr{P}) = \frac{\lambda_{\mathscr{M}}}{2}\|\mathscr{M} - \mathscr{P}\|_F^2 . \tag{84}$$

Assuming that most of the energy of $\mathscr{M}$ lies in a low-rank structure, we write the tensor $\mathscr{P}$ as a sum of a small number $K_{\mathscr{P}}$ of rank-1 components, such that

$$\mathscr{P} = \sum_{i=1}^{K_{\mathscr{P}}} \delta_i \boldsymbol{x}_i^{(1)} \circ \boldsymbol{x}_i^{(2)} \circ \boldsymbol{x}_i^{(3)} \circ \boldsymbol{x}_i^{(4)} . \tag{85}$$

This introduces a low-rank *a priori* condition on $\mathscr{P}$ whose strictness can be controlled by the regularization constant $\lambda_{\mathscr{M}}$. Using (85) in (84) leads to the optimization problem

$$\begin{aligned} \Big(\widehat{\boldsymbol{\Delta}}, \widehat{\boldsymbol{X}}^{(1)}, \widehat{\boldsymbol{X}}^{(2)}, \widehat{\boldsymbol{X}}^{(3)}, \widehat{\boldsymbol{X}}^{(4)}\Big) = \\ \underset{\boldsymbol{\Delta},\boldsymbol{X}^{(1)},\boldsymbol{X}^{(2)},\boldsymbol{X}^{(3)},\boldsymbol{X}^{(4)}}{\arg\min} \frac{\lambda_{\mathscr{M}}}{2}\left\|\mathscr{M} - \sum_{i=1}^{K_{\mathscr{P}}} \delta_i \boldsymbol{x}_i^{(1)} \circ \boldsymbol{x}_i^{(2)} \circ \boldsymbol{x}_i^{(3)} \circ \boldsymbol{x}_i^{(4)}\right\|_F^2 , \end{aligned} \tag{86}$$

where $\boldsymbol{\Delta} = \text{TDiag}_4(\delta_1, \ldots, \delta_{K_{\mathscr{P}}})$ is a 4-dimensional diagonal tensor with $\boldsymbol{\Delta}_{i,i,i,i} = \delta_i$. Problem (86) can be solved using an alternating least-squares strategy [141].

Finally, the solution $\widehat{\mathscr{P}}$ is obtained from $\widehat{\boldsymbol{\Delta}}, \widehat{\boldsymbol{X}}^{(1)}, \widehat{\boldsymbol{X}}^{(2)}, \widehat{\boldsymbol{X}}^{(3)}$, and $\widehat{\boldsymbol{X}}^{(4)}$ using the full multilinear product as

$$\widehat{\mathscr{P}} = [\![\widehat{\boldsymbol{\Delta}}; \widehat{\boldsymbol{X}}^{(1)}, \widehat{\boldsymbol{X}}^{(2)}, \widehat{\boldsymbol{X}}^{(3)}, \widehat{\boldsymbol{X}}^{(4)}]\!]. \tag{87}$$

6.3.4 Solving with respect to $\mathcal{Q}$

Analogous to the previous section, the cost function to be optimized for $\mathcal{Q}$ can be written as

$$\mathcal{J}(\mathcal{Q}) = \frac{\lambda_{\mathcal{A}}}{2}\|\mathcal{A} - \mathcal{Q}\|_F^2. \tag{88}$$

Assuming that most of the energy of $\mathcal{A}$ lies in a low-rank structure, we write tensor $\mathcal{Q}$ as a sum of a small number $K_{\mathcal{Q}}$ of rank-1 components, such that

$$\mathcal{Q} = \sum_{i=1}^{K_{\mathcal{Q}}} \xi_i \boldsymbol{u}_i^{(1)} \circ \boldsymbol{u}_i^{(2)} \circ \boldsymbol{u}_i^{(3)}. \tag{89}$$

This introduces a low-rank *a priori* condition on $\mathcal{A}$, which will be more or less enforced depending on the regularization constant $\lambda_{\mathcal{A}}$. Using (89) in (88) leads to the optimization problem

$$\left(\widehat{\boldsymbol{\Xi}}\ , \widehat{\boldsymbol{U}}^{(1)}, \widehat{\boldsymbol{U}}^{(2)}, \widehat{\boldsymbol{U}}^{(3)}\right) = \underset{\boldsymbol{\Xi}, \boldsymbol{U}^{(1)}, \boldsymbol{U}^{(2)}, \boldsymbol{U}^{(3)}}{\arg\min} \frac{\lambda_{\mathcal{A}}}{2} \left\| \mathcal{A} - \sum_{i=1}^{K_{\mathcal{Q}}} \xi_i \boldsymbol{u}_i^{(1)} \circ \boldsymbol{u}_i^{(2)} \circ \boldsymbol{u}_i^{(3)} \right\|_F^2 \tag{90}$$

where $\boldsymbol{\Xi} = \text{TDiag}_3\left(\xi_1, \ldots, \xi_{K_{\mathcal{Q}}}\right)$ is an order-3 diagonal tensor with $\boldsymbol{\Xi}_{i,i,i} = \xi_i$. Problem (90) can be solved using an alternating least-squares strategy [141]. Finally, the solution $\widehat{\mathcal{Q}}$ is obtained from $\widehat{\boldsymbol{\Xi}}, \widehat{\boldsymbol{U}}^{(1)}, \widehat{\boldsymbol{U}}^{(2)}$ and $\widehat{\boldsymbol{U}}^{(3)}$ using the full multilinear product as

$$\widehat{\mathcal{Q}} = \left[\!\left[\widehat{\boldsymbol{\Xi}}\ ; \widehat{\boldsymbol{U}}^{(1)}; \widehat{\boldsymbol{U}}^{(2)}; \widehat{\boldsymbol{U}}^{(3)}\right]\!\right]. \tag{91}$$

6.3.5 Computational complexity of Algorithm 5

The computational complexity of each iteration of Algorithm 5 can be measured as follows. The optimizations w.r.t. $\mathcal{A}$ and $\mathcal{M}$ both consist of regularized constrained least squares problems with N_1N_2P and N_1N_2LP variables respectively. Thus, these problems can be solved with a complexity of $\mathcal{O}\big((N_1N_2P)^3\big)$ and $\mathcal{O}\big((N_1N_2LP)^3\big)$, respectively. The optimizations w.r.t. variables $\mathcal{P}$ and $\mathcal{Q}$ consist of CPDs of these tensors with ranks $K_{\mathcal{P}}$ and $K_{\mathcal{Q}}$, respectively. Considering an alternating least squares approach for the CPD, these optimization problems will have computational complexities of $\mathcal{O}\big(K_{iter}K_{\mathcal{P}}N_1N_2LP\big)$ and $\mathcal{O}\big(K_{iter}K_{\mathcal{Q}}N_1N_2P\big)$, respectively, where K_{iter} is the number of ALS iterations [155]. Thus, the overall complexity of the algorithm scales linearly with the number of ALS iterations and with the tensor ranks, and cubically in the problem dimensions. When processing large datasets, the extra complexity could be partially mitigated by applying image segmentation or band selection [156] strategies. This analysis is beyond the scope of the present work and will be addressed in the future.

6.3.6 Estimating tensor ranks

In Section 6.2.4 we have recalled important results relating bounds for order-3 tensor ranks to the span of the matricized versions of tensors. We have also noted, from

our own experience, that those bounds tend to indicate tensor ranks that are larger than the rank associated with the information relevant for SU. Our interest in SU is to model low-dimensional structures of the HI using low-rank tensors. At the same time, this low-rank representation should be rich enough to include relevant information of the original HI tensor. Therefore, although the literature presents many rank estimation strategies (see [157] and references therein), in this work we exploit the rank bounds discussed in Section 6.2.4 to approximate the "useful rank" of a tensor by the number of the largest singular values of their matricized versions required to represent most of the tensor energy.

Let $\boldsymbol{T}_i = \text{mat}_i(\mathcal{T}) \in \mathbb{R}^{N_i \times (N_1 \ldots N_{i-1} N_{i+1}, \ldots N_R)}$ be the matricization of an arbitrary tensor $\mathcal{T} \in \mathbb{R}^{N_1 \times N_2 \ldots \times N_R}$ obtained by stacking all tensor fibers along the i-th tensor dimension. Let $\boldsymbol{s}_i = \text{SVD}(\boldsymbol{T}_i)$ be the set of singular values of $\boldsymbol{T}_i$, sorted descending in value. Also, let $\boldsymbol{d}_i = \text{diff}(\boldsymbol{s}_i)$ be the vector of first order differences of the elements of $\boldsymbol{s}_i$, such that, $d_j^{(i)} = s_j^{(i)} - s_{j+1}^{(i)}$. Then, we define the i-th candidate for rank of $\mathcal{T}$ as the smallest index j such that $|d_j^{(i)}|$ sufficiently small, namely,

$$\widehat{F}_i = \min j, \text{ subject to } , |d_j^{(i)}| < \varepsilon \,, \tag{92}$$

where ε is a parameter limiting the singular value variation. In all experiments reported here we used $\varepsilon = 0.15$. We have experimentally verified that the resulting abundance MSE has very low sensitivity to the choice of ε. Finally, we approximate the rank of tensor $\mathcal{T}$ as

$$K = \max\{\widehat{F}_1, \ldots, \widehat{F}_R\}. \tag{93}$$

For the experiments reported in this paper, we have used definition (93) to estimate $K_{\mathcal{P}}$ and $K_{\mathbb{Q}}$ in (85) and (89) from the abundance and endmember tensors estimated using simple unmixing strategies such as the SCLS [19].

6.4 SIMULATIONS

In this section, the performance of the proposed methodology is illustrated through simulations with both synthetic and real data. We compare the proposed ULTRA-V method with the the FCLS, the SCLS [19], the PLMM [20], the ELMM [19], and the GLMM [34]. To highlight the differences between ULTRA-V and ULTRA [45], we also consider ULTRA for simulations with synthetic data.

To measure the accuracy of the unmixing methods we consider the Mean Squared Error

$$\text{MSE}_{\mathcal{X}} = \frac{1}{N_{\mathcal{X}}} \| \text{vec}(\mathcal{X}) - \text{vec}(\widehat{\mathcal{X}}) \|^2 \tag{94}$$

where $\text{vec}(\cdot)$ is the vectorization operator $\mathcal{X} \to \boldsymbol{x}$, $\mathbb{R}^{a \times b \times c} \mapsto \mathbb{R}^{abc}$, $N_{\mathcal{X}} = abc$, and the Spectral Angle Mapper for the endmembers tensor

$$\text{SAM}_{\mathcal{M}} = \frac{1}{N} \sum_{n=1}^{N} \sum_{k=1}^{P} \arccos \left(\frac{\boldsymbol{m}_{k,n}^{\top} \widehat{\boldsymbol{m}}_{k,n}}{\|\boldsymbol{m}_{k,n}\| \|\widehat{\boldsymbol{m}}_{k,n}\|} \right). \tag{95}$$

Table 9 – Simulation results using synthetic data.

	$\text{MSE}_{\mathcal{A}}$	$\text{MSE}_{\mathcal{M}}$	$\text{SAM}_{\mathcal{M}}$	$\text{MSE}_{\mathcal{Y}}$	Time
Data Cube 0 – DC0					
FCLS	1.81	-	-	16.11	0.42
SCLS	0.68	175.79	6.19	59.38	0.38
PLMM	0.76	94.57	5.58	4.89	81.89
ELMM	0.35	106.17	5.63	4.78	17.15
GLMM	0.34	101.51	5.87	**5.7e-3**	20.23
ULTRA-V	**0.23**	**92.39**	**5.56**	0.73	14.46
ULTRA	1.81	-	-	16.11	1.42
Data Cube 1 – DC1					
FCLS	2.01	-	-	6.93	0.74
SCLS	2.07	92.16	5.37	24.78	0.76
PLMM	1.58	157.00	8.49	2.75	120.48
ELMM	1.29	69.15	5.95	**0.01**	23.18
GLMM	1.20	68.11	6.09	**0.01**	29.81
ULTRA-V	**1.12**	**60.12**	**5.21**	4.26	29.47
ULTRA	1.17	-	-	10.93	4.44
Data Cube 2 – DC2					
FCLS	1.90	-	-	2.03	0.30
SCLS	0.71	**1.66**	2.29	1.07	0.30
PLMM	1.27	2.45	**2.28**	2.12	61.13
ELMM	0.63	2.84	3.32	2.31	11.15
GLMM	0.59	1.84	2.79	2.04	11.60
ULTRA-V	**0.46**	2.91	2.94	**9e-5**	5.19
ULTRA	0.83	-	-	1.02	0.36

All the algorithms were implemented in Matlab on a desktop computer equipped with an Intel Core I7 processor with 4.2Ghz and 16Gb of RAM. In all cases, we used endmembers extracted using the VCA [131] either to build the reference endmember matrix or to initialize the different methods, with the number of endmembers P assumed to be known a priori. The abundance maps were initialized using the maps estimated by the SCLS.

6.4.1 Synthetic data

For a comprehensive comparison among the different methods we created three synthetic datasets, namely Data Cube 0 (DC0), Data Cube 1 (DC1) and Data Cube 2

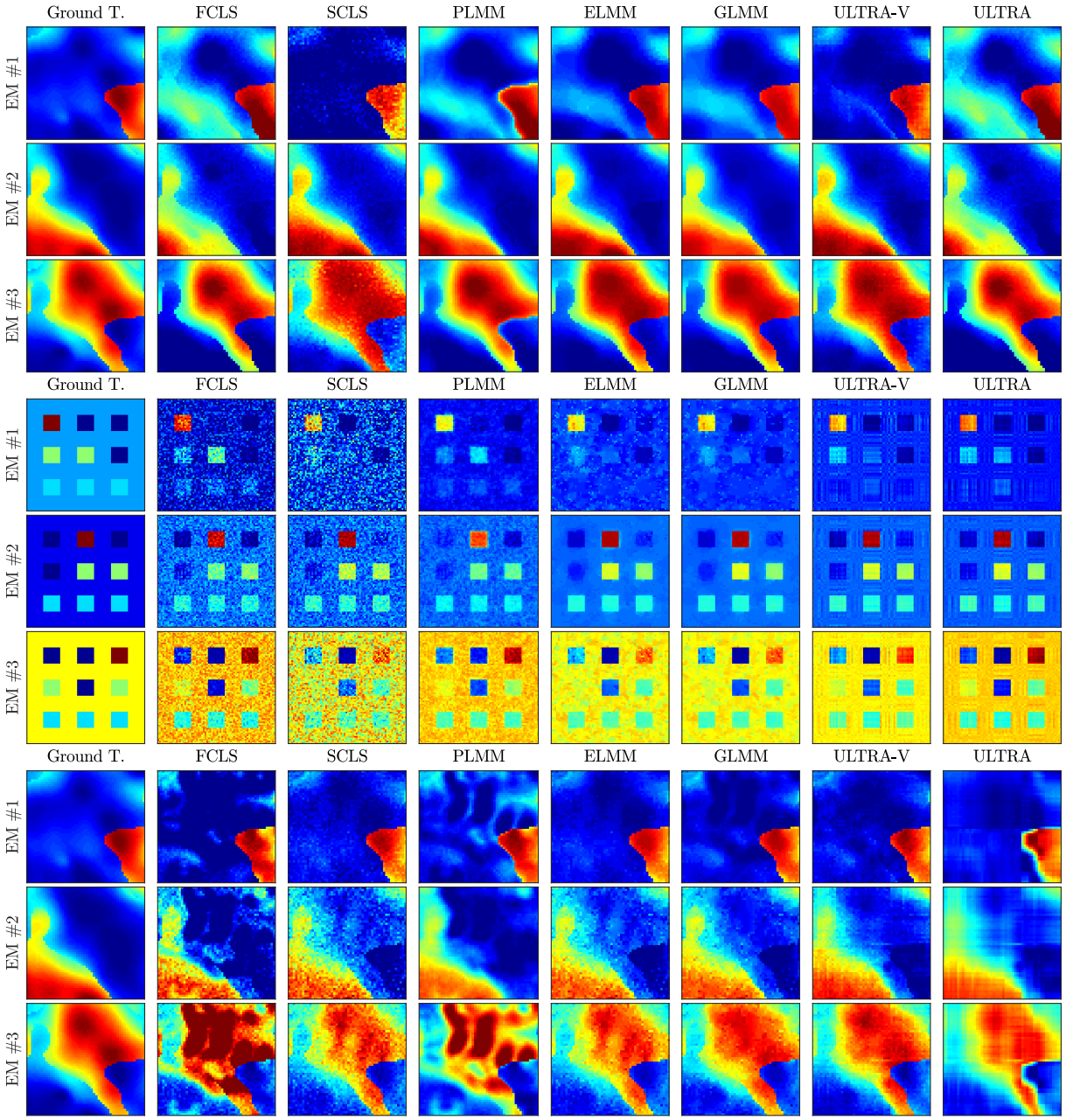


Figure 26 – Abundance maps of (top–down) DC0, DC1, and DC2 for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

(DC2), with 50×50 pixels (DC0 and DC2) and 70x70 pixels (DC1). DC0 and DC1 were built using three 224-band endmembers extracted from the USGS Spectral Library [136], while DC2 was built using three 16-band minerals often found in bodies of the Solar System [19]. For the three datasets, spatially correlated abundance maps were used, as depicted in the first column of Fig. 26. For DC0, we adopted the variability model used in [19] (a multiplicative factor acting on each endmember). For DC1, we used the variability model according to the PLMM [20]. For DC2, we used the Hapke model [37] devised to realistically represent the spectral variability introduced due to changes in the illumination

conditions caused by the topography of the scene [19]. White Gaussian noise was added to all datasets to yield a 30 dB SNR.

To select the optimal parameters for each algorithm, we performed grid searches for each dataset. We used parameter search ranges based on the ranges tested and discussed by the authors in the original publications. For the PLMM we used $\gamma = 1$, since the authors fixed this parameter in all simulations, and searched for α and β in the ranges $[10^{-6}, 10^{-3}, 0.1, 0.35, 0.7, 1.4, 5, 25]$ and $[10^{-9}, 10^{-5}, 10^{-4}, 10^{-3}]$, respectively. For both ELMM and GLMM, we selected the parameters among the following values: $\lambda_S, \lambda_M \in [0.01, 0.1, 1, 5, 10, 15]$, $\lambda_A \in [10^{-6}, 10^{-3}, 0.01, 0.05, 0.1, 1, 10,$ and $\lambda_\psi, \lambda_\Psi \in [10^{-6}, 10^{-3}, 10^{-1}]$, while for the proposed ULTRA-V we selected the parameters in the intervals $\lambda_{\mathscr{A}} \in [0.001, 0.01, 0.1, 1, 10, 100]$ and $\lambda_{\mathscr{M}} \in [0.1, 0.2, 0.4, 0.6, 0.8, 1]$. For the ULTRA we searched $\lambda_{\mathscr{A}}$ in the same interval used for the ULTRA-V.

The results are shown in Table 9, were the best values for each metric are marked in bold. ULTRA-V clearly outperformed the competing algorithms for all datasets in terms of $\text{MSE}_{\mathscr{A}}$. For the other metrics, the best results depended on the datasets. In terms of $\text{MSE}_{\mathscr{M}}$ and $\text{SAM}_{\mathscr{M}}$, ULTRA-V yielded the best results for DC0 and DC1. Finally, ULTRA-V results for $\text{MSE}_{\mathscr{Y}}$ were the second best for DC0 and the best for DC2. The execution times, shown in the rightmost columns of Table 9, show that ULTRA-V required the smallest execution time among the more sophisticated algorithms (PLMM, ELMM and GLMM) for DC0 and DC2, and comparable execution time for DC1. As expected, the ULTRA method provided results that were often better than the FCLS but significantly worse than those obtained from methods accounting for EM variability. This happens because ULTRA imposes a low-rank structure over the abundances but keeps the EMs fixed for all pixels, what greatly limits the algorithm capacity to adapt to EM variations along the image.

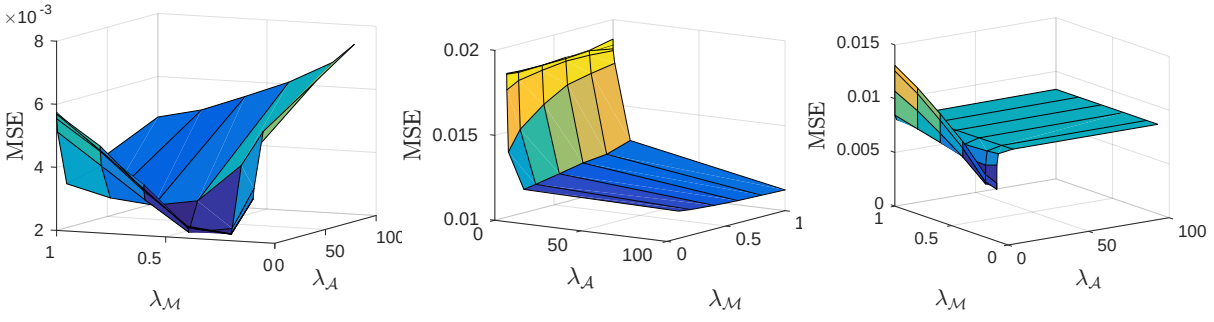


Figure 27 – Parameters sensitivity to changes around the optimal values. Left: DC0, Middle: DC1, and Right: DC2.

6.4.1.1 Parameters sensitivity

While we have proposed a strategy to determine rank values for tensors $\mathscr{P}$ and $\mathbb{Q}$, the parameters $\lambda_{\mathscr{M}}$ and $\lambda_{\mathscr{A}}$ need to be selected by the user. We now study the sensitivity

of the ULTRA-V performance to variations of the parameters within the parameter search intervals presented in the previous section. Fig. 27 shows the values of $\text{MSE}_{\mathcal{A}}$ resulting from unmixing the data using each combination of the parameter values. The sensitivity clearly tends to increase when values less than 1 are used for both parameters. Our practical experience indicates that good $\text{MSE}_{\mathcal{A}}$ results can be obtained using $\lambda_{\mathcal{M}}$ in $[0, 1]$, and large values about 100 for $\lambda_{\mathcal{A}}$. Moreover, some insensitivity is verified for small changes in $\lambda_{\mathcal{A}}$ about large values. Thus, searching $\lambda_{\mathcal{A}}$ in $[0.001, 100]$ with values spaced by decades as done for the examples in the previous section seems reasonable.

6.4.1.2 Discussion

A close look at Fig. 26 reveals the abundance maps estimated using ULTRA-V look noisier than those obtained using ELMM and GLMM. This is because the proposed approach does not impose the local smoothness imposed by total variation, but emphasizes abundance regularity by enforcing a low-rank property. We note, however, that the spacial smoothness imposed by total variation is not necessarily mandatory for a good abundance estimation, as can be verified from the results in Table 9.

6.4.2 Real data

For the simulations with real data, we considered three datasets, consisting of the Houston, Samson and Jasper Ridge images. All datasets were captured by the AVIRIS, which originally has 224 spectral bands. For all images, the water absorption bands were removed resulting in 188 bands for the Houston image, 156 bands for the Samson image and 198 bands for the Jasper Ridge image. The Houston data set is known to have four predominant endmembers [19, 46]. The Samson and Jasper Ridge images are known to have three and four endmembers, respectively [32]. For all images the endmembers were extracted using the VCA [131]. Fig. 28 shows the reconstructed abundance maps for all images and for all tested methods. The quantitative results are shown in Table 10. Note that since the ground truth (correct) abundance values are not available for these images, only the reconstruction error $\text{MSE}_{\mathcal{Y}}$ has been used as a sort of quality verification.

The last column in Fig. 28 shows that the proposed ULTRA-V method provided an accurate abundance estimation, clearly outperforming the competing algorithms[1]. In fact, for the Concrete and Metallic Roofs endmembers, the ULTRA-V abundance map presents stronger Concrete and Metallic Roofs components in the stadium stands and stadium towers, respectively, when compared with the other methods equipped for dealing with spectral variability. The performance improvement provided by ULTRA-V is clearer for the Samson and Jasper Ridge images. For instance, there is significantly less confusion between the Water, Tree and Soil endmembers in the ULTRA-V results for the Samson

[1] The differences between the ELMM and ULTRA-V results are less significant for the Houston image.

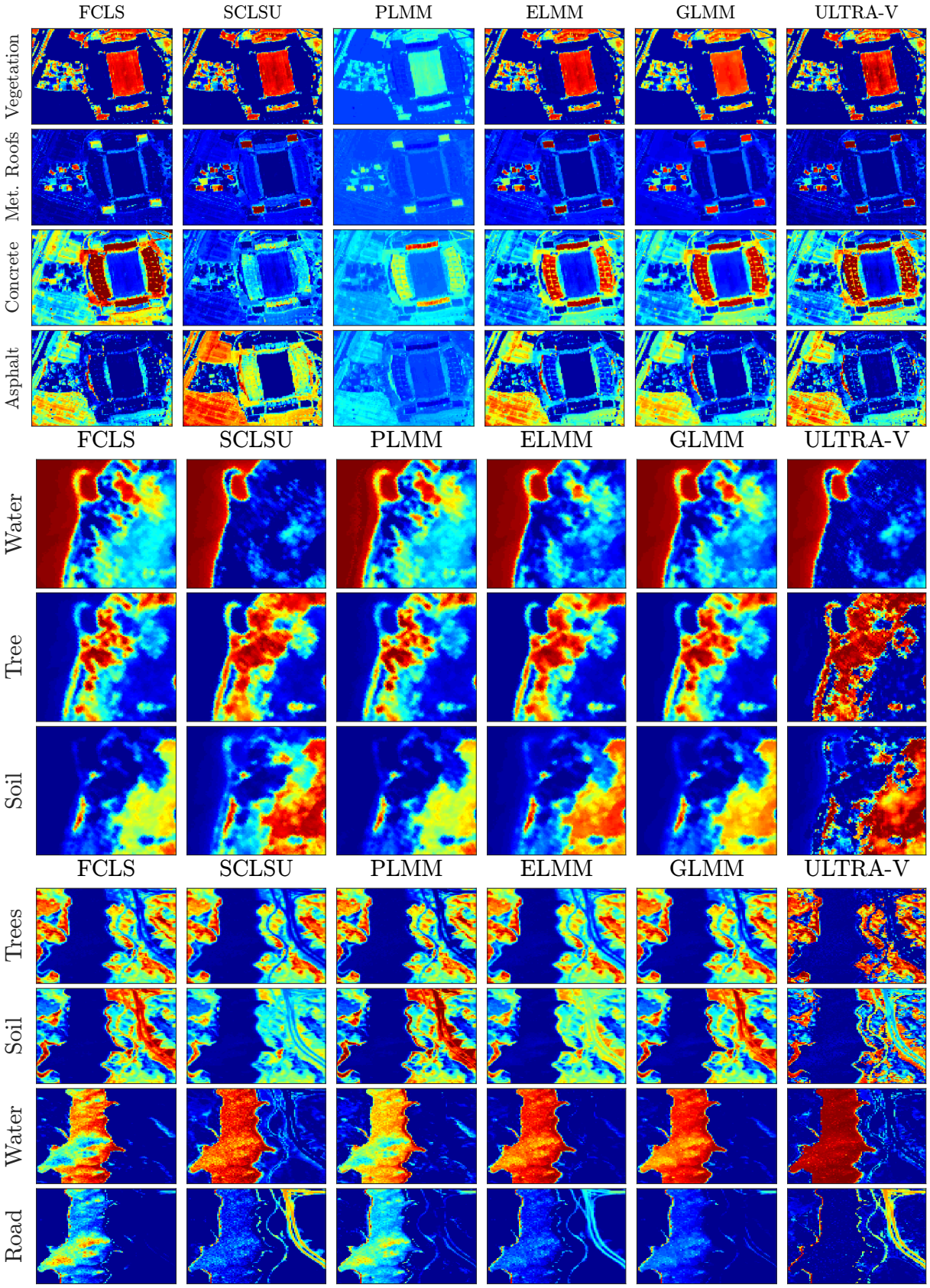


Figure 28 – Abundance maps of the Houston (upper panel), Samson (middle painel), and Jasper Ridge (bottom painel) data sets for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

image when compared to those of the PLMM, ELMM, and GLMM methods. Similarly, the ULTRA-V reconstructed abundance maps of the Jasper Ridge image show a much stronger Water component in the river and less confusion between the Tree, Soil and Road endmembers.

The objective metrics presented in Table 10 indicate that ULTRA-V yields competitive reconstruction errors in terms of MSE. These results, however, should be interpreted with proper care, as the connection of reconstruction error and abundance estimation is not straightforward.

The execution times in Table 10 indicate that, as discussed in Section 6.3.5, ULTRA-V did not scale well with the larger image sizes and higher number of endmembers, which directly impacted the CPD stage of ULTRA-V. Moreover, the more complex images resulted in higher rank estimates using the strategy discussed in Section 6.3.6. This indicates that there is still room for improving the proposed method by either providing a segmentation strategy or using faster CPD methods. This, however, is an open problem that will be addressed in future works.

To assess the estimated endmember variabilities, we analyzed the results for the Samson data set. We considered two approaches. The first approach consisted in averaging the projection of the estimated endmembers on the three eigenvectors associated to the three largest eigenvalues for each endmember. The results are shown in Fig. 31. These plots illustrate the endmember variances for each pixel, with red implying a large variance and blue a small variance. The second approach consisted in directly comparing the endmembers estimated with ULTRA-V and VCA. The results are shown in Fig. 32. These figures illustrate the ability of the proposed method to characterize the spectral variability while enforcing a spatial structure for the estimated endmembers.

To illustrate the role of the low-rank tensors $\mathscr{P}$ and $\mathbb{Q}$, we compare them to the abundances $\mathscr{A}$ and endmembers $\mathscr{M}$ in Figs. 29 and 30, for the Jasper Ridge dataset. Fig. 29 shows the estimated abundances $\mathscr{A}$ and their low-rank counterpart $\mathbb{Q}$. One can verify that $\mathbb{Q}$ (bottom row) has a very coarse spatial distribution when compared with $\mathscr{A}$ (top row). This shows that imposing the low-rank structure through a regularization constraint gives the resulting abundances enough flexibility to model fine-scale spatial details while maintaining most of its spatial distribution. Fig. 30 leads to similar conclusions for the endmembers. One can note that the low-rank tensor $\mathscr{P}$ has a coarser structure when compared with the estimated endmembers $\mathscr{M}$. This distinction can be seen very clearly for the Water and Road endmembers.

6.5 CONCLUSIONS

In this chapter, we proposed a new low-rank regularization strategy for introducing low-dimensional spatial-spectral structure into the abundance and endmember tensors for hyperspectral unmixing considering spectral variability. The resulting iterative algorithm,

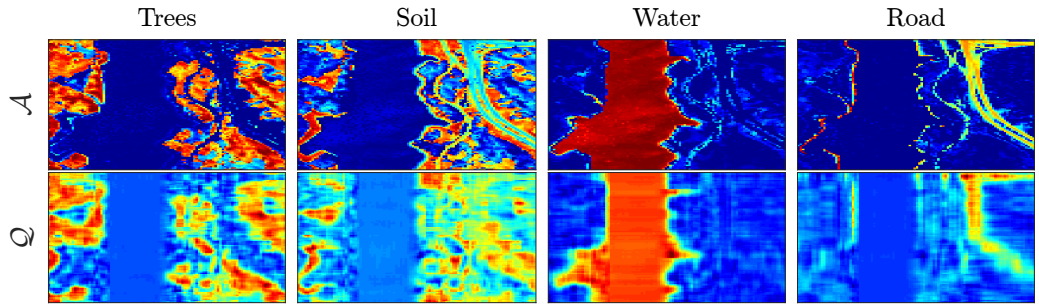


Figure 29 – Comparison of tensors $\mathscr{A}$ and $\mathbb{Q}$ after ULTRA-V convergence for the Jasper Ridge data set.

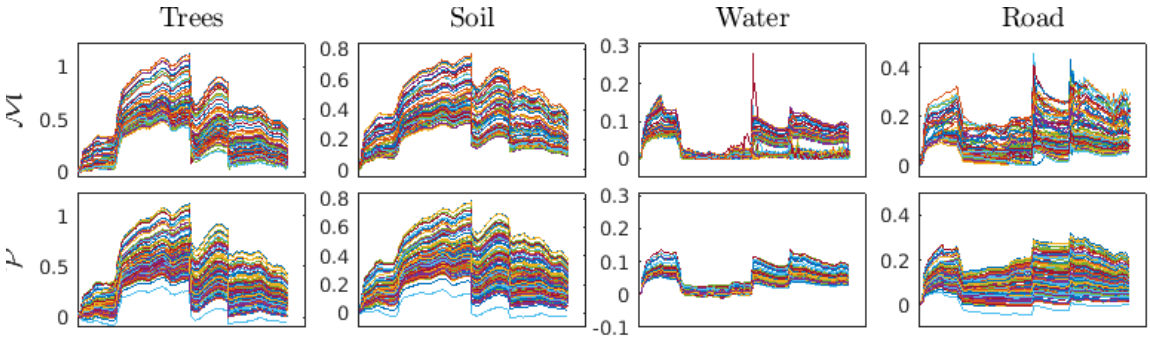


Figure 30 – Comparison of tensors $\mathscr{M}$ and $\mathscr{P}$ after ULTRA-V convergence for the Jasper Ridge data set.

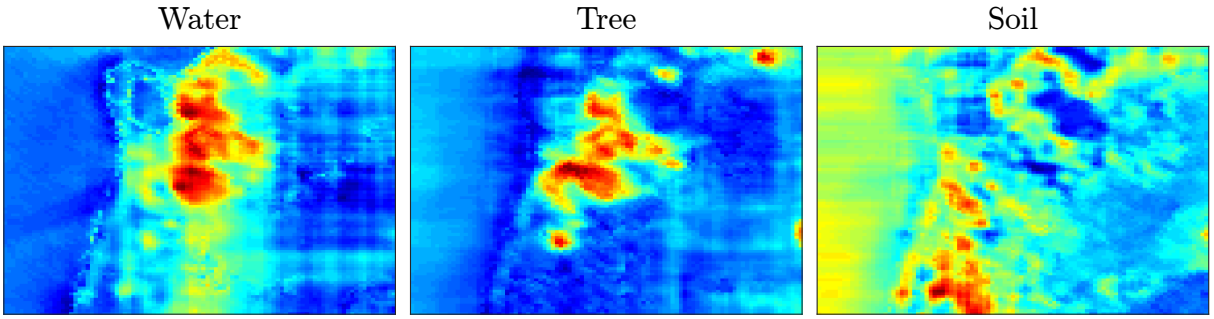


Figure 31 – Average of the ULTRA-V endmembers tensor projection over the 3 principal components for the Samson data set.

called ULTRA-V, imposes low-rank structure by means of regularizations that force most of the energy of the estimated abundances and endmembers to lay within a low-dimensional structure. The proposed approach does not confine the estimated abundances and endmembers to a strict low-rank structure, which would not adequately account for the complexity experienced in real-world scenarios. It includes also a strategy to estimate the rank of the regularization tensors $\mathscr{P}$ and $\mathbb{Q}$, leaving only two parameters to be adjusted within a relatively reduced search space. Simulation results using both synthetic and real data showed that the ULTRA-V can outperform state-of-the-art unmixing algorithms accounting for spectral variability.

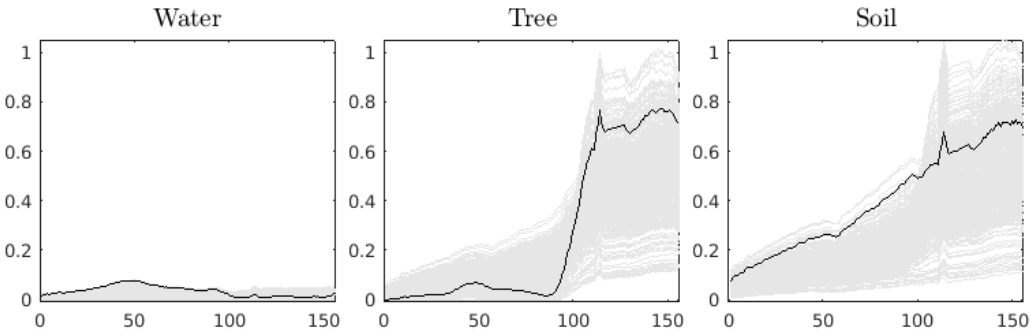


Figure 32 – VCA result (black) and ULTRA-V (gray) endmembers for each pixel of the Samson data set.

Table 10 – Simulation results with real data.

Algorithm	Houston Data		Samson		Jasper Ridge	
	$\text{MSE}_{\mathcal{Y}}$	Time	$\text{MSE}_{\mathcal{Y}}$	Time	$\text{MSE}_{\mathcal{Y}}$	Time
FCLS	0.2283	1.90	0.0177	1.38	0.3567	1.59
SCLS	0.0037	2.04	0.0041	1.29	0.0271	1.79
PLMM	0.0190	454.90	0.0034	105.36	0.0257	72.86
ELMM	0.0010	474.45	7.82e-4	40.50	0.0058	100.49
GLMM	**1.0e-5**	1326.85	**0.2e-5**	50.62	**2.5e-5**	214.07
ULTRA-V	0.0018	264.71	6.4e-5	148.46	15.0e-5	120.91

7 DEEP GENERATIVE ENDMEMBER MODELING: AN APPLICATION TO UNSUPERVISED SPECTRAL UNMIXING

7.1 INTRODUCTION

Although SU methods based on extended parametric models offer different trade-offs between representation capacity and model complexity, they still fail to achieve a desirable balance between a low-dimensional representation and enough flexibility to represent complex EM variability. Specifically, they fail to properly explore the fact that, although being complex and spectrally non-homogeneous, spectral variability in real scenes is often confined to low-dimensional manifolds [37, 97, 101]. This property is due to the fact that the spectral signature of many materials is a function of only a few photometric or chemical properties of the medium. Prominent examples include packed particle spectra as a function of its roughness, size and density [37], leaf reflectance spectra as a function of various biophysical parameters [97], and soil reflectance as a function of moisture conditions [101]. Thus, existing models tend to be either too restrictive in their modeling capability or to lead to severely ill-posed estimation problems.

SU considering EM variability has also been formulated as a supervised learning problem, which is then solved without the need for an accurate physical model using neural networks (NNs) or support vector machines (SVMs) [158, 159, 160, 161]. However, these strategies depend on the availability of vast amounts of training data to adequately capture the spectral diversity of real scenes. This makes the training process computationally intensive and often intractable for large EM libraries, which must also be known *a priori*. Some works attempt to reduce the computational cost of these solutions by modifying learning algorithms to use hybrid soft-hard classification [162, 163, 164]. However, the resulting reconstructed abundance fractions do not have a clear physical interpretation due to the lack of a direct relationship to a physically motivated mixing model.

More recently, unsupervised SU approaches have also emerged by using autoencoders (AEC), which consist of encoder-decoder structured NNs originally devised for nonlinear dimensionality reduction [165]. These methods attempt to associate the decoder structure of the network with the LMM and the low-dimension representation of the input spectral vectors to the fractional abundances [166]. Different variations have been proposed, using pre-processing steps to reduce noise and outliers [167, 168], untiying the decoder from the encoder weights [169], using spectral angle distances to address nonlinear SU [170], or using denoising autoencoders to generate a robust initialization to matrix factorization-based SU strategies [171].

Despite their popularity, supervised learning-based SU algorithms are still not able to properly address the spectral variability problem, as they depend on extremely large amounts of labeled training data, leading to a computationally unfeasible learning process. Furthermore, the lack of a clear connection between AEC-based strategies and the physical

mixing process makes one skeptic when concerning the robustness of AEC-based SU in face of more complex phenomena such as spectral variability.

In this work, we propose a novel SU formulation that leverages the advantages of deep learning methods to address EM variability while still maintaining a strong connection to the physical mixing process, and using limited amounts of training data. Specifically, we adopt a deep generative NN to represent the manifold of EM spectra, which is then incorporated within the LMM. Generative models such as variational autoencoders (VAE) [172] and generative adversarial networks (GAN) [173] have recently obtained excellent performance at learning the probability distribution of complex data sets in very high dimensional spaces (e.g. natural images) from relatively small amounts of training data. The structure of generative models allow one to find a low-dimensional latent representation that parsimonously describes the variability of complex high-dimensional data sets. This leads to a low-dimensional parametrization of the training data distribution.

We formulate a novel unmixing strategy that can be cast as the problem of estimating the latent representations of the generative endmember models and the corresponding fractional abundances for each pixel in the HI. Specifically, we break down the SU problem in two steps. In the first step, we learn the latent EM variability manifold for each material in the scene using a deep generative EM model. The learning process uses pure pixel information directly extracted from the observed HI, which makes the proposed strategy suitable for unsupervised SU. In the second step, an alternating least-squares strategy is employed to estimate the parameters of an extended version of the LMM parametrized using the generative EM models obtained in the first step. The corresponding optimization problem is solved iteratively with respect to the abundances and to the low-dimensional representations of the EMs in the latent space of the deep generative models.

As a result, the proposed approach benefits from the reduced dimension of the latent space. Moreover, unlike current approaches, the new method does not depend on the careful selection of regularization parameters to yield a good performance. The resulting algorithm is named *Deep Generative Unmixing algorithm* (DeepGUn). The proposed method is strongly related to parametric models and leverages the learning and generalization capability of deep neural networks to properly represent the manifold of EM variability. Hence, DeepGUn leads to a model that is both low-dimensional and physically accurate, better describing the variability actually present in the scene.

Experimental results performed with both synthetic and real data indicate that the proposed strategy leads to more accurate abundance estimations than standard state-of-art SU methods accounting for EM variability. Qualitative analysis of the estimated abundance maps confirms these results. The improved accuracy comes at the expense of a small increase in the computational cost when compared to the best competing strategies.

This chapter is organized as follows. Section 7.2 briefly reviews the LMM and its parametric extended versions. Section 7.3 discusses the basic properties of generative

models in the context of VAE and GAN. Section 7.4 introduces the proposed generative EM model and its learning strategy. In Section 7.5 we formulate the resulting SU problem, present the DeepGUn algorithm, and discuss aspects of the proposed optimization strategy. The neural network architecture is discussed in Section 7.6. The performance of the proposed method is compared with that of competing algorithms in Section 7.7. Finally, the conclusions are presented in Section 7.8.

7.2 LINEAR MIXING MODELS WITH SPECTRAL VARIABILITY, REVISITED

The LMM (1) assumes that the EM spectra are fixed for all HI pixels $\boldsymbol{y}_n$, $n = 1, \ldots, N$. This assumption jeopardizes the accuracy of estimated abundances in many circumstances due to the presence of spectral variability. Thus, different parametric models have been recently proposed to account for variable EM spectra within a given scene [19, 20, 34]. These models can be generically described as

$$\boldsymbol{y}_n = f(\boldsymbol{M}_0, \boldsymbol{\theta}_n)\boldsymbol{a}_n + \boldsymbol{e}_n \,, \tag{96}$$

where f is a parametric function, $\boldsymbol{M}_0 \in \mathbb{R}^{L \times P}$ is a reference EM matrix, and $\boldsymbol{\theta}_n$ is a vector of parameters describing the manifold of EM variability.

Different functional forms have been proposed for $f(\boldsymbol{M}_0, \boldsymbol{\theta}_n)$ to account for EM variability in this framework, such as additive [20] or multiplicative [19, 34] variability factors acting upon the reference EM matrix $\boldsymbol{M}_0$. However, these models fail to achieve a desirable balance between a low-dimensional representation and enough flexibility to represent complex variability patterns. They tend to be either too restrictive in their modeling capability, or to lead to ill-posed optimization problems [46]. Instead of using a pre-defined parametric model, we propose to address this issue by learning a parametric function $f(\boldsymbol{M}_0, \boldsymbol{\theta}_n)$ using a generative model.

7.3 GENERATIVE MODELS

Generative models attempt to estimate the probability distribution $p(X)$ of a random variable $X \in \mathbb{R}^L$ based on a set of observations $\boldsymbol{x}_i$, $i = 1, \ldots, M_x$ in such a way that allows one to generate new samples that look similar to new realizations of X. The main characteristic of this problem, which sets it apart of other unsupervised learning methods such as density estimation, is the fact that we must to be able to sample from the estimated model $\hat{p}(X)$.

In many practical applications of interest, the dimensionality L of the variable of interest X is very high. This makes the general problem very difficult, as it amounts to estimating and sampling from an arbitrary high-dimensional probability density function [174, 175]. Nonetheless, the distributions of interest are often supported at a low-dimensional manifold of a set of so-called latent variables, and this fact can be explored to make the

problem more tractable. A convenient way to address this problem is to define a new random variable $\mathbb{R}^H \ni Z \sim p(Z)$ with a known distribution in a low-dimensional space (e.g. an isotropic Gaussian distribution with $H \ll L$), and a parametric function (e.g. a neural network) $\mathcal{G}_\eta$ mapping $Z \mapsto \widehat{X} \in \mathbb{R}^L$ such that the image of Z by $\mathcal{G}_\eta$ is a random variable whose distribution is very close to $p(X)$. In other words, the goal becomes to learn the parameters η of $\mathcal{G}_\eta$ such that the distribution of $\widehat{X} = \mathcal{G}_\eta(Z)$ is as close to $p(X)$ as possible. Then, samples of $\widehat{X}$ can be generated by sampling from $Z \sim p(Z)$ and using the mapping $\mathcal{G}_\eta(Z)$.

Although estimating η may still seem difficult at first, recent advances in machine learning such as VAEs [172] and GANs [173] have shown formidable performance at learning complex distributions such as those of natural images.

VAEs address this problem by assuming that the distribution of the observed data X follows a directed graphical model $p(X|Z)$, which is represented by the function $\mathcal{G}_\eta$. The parameters of $\mathcal{G}_\eta$ are learned by maximizing a lower bound on the log-likelihood of $p(X)$ [172]:

$$\log\, p(X) \geqslant\, \mathrm{E}_{q_\tau(Z|X)} \left\{ \log p(X|Z) \right\} - \mathrm{KL}\left(q_\tau(Z|X) \| p(Z) \right), \tag{97}$$

where $\mathrm{KL}(\cdot\|\cdot)$ is the Kullback-Leibler divergence between two distributions, $\mathrm{E}_\varsigma\{\cdot\}$ is the expected value operator with respect to the distribution ς and $q_\tau(Z|X)$ is a variational approximation to the intractable posterior $p(Z|X)$, which is modeled by a function $\mathcal{D}_\tau$ (e.g. another neural network) parameterized in τ. Note that $q_\tau(Z|X)$ must be a high-capacity distribution[1], so that it can provide a good approximation of the posterior $p(Z|X)$, which then allows the lower bound in (97) to be close to the true value of $\log\, p(X)$ [176].

GANs, on the other hand, attempt to learn the distribution $p(X)$ by searching for the Nash equilibrium of a two-player adversarial game [173]. A generator network $\mathcal{G}_\eta$ tries map the distribution of the latent variables Z into the data distribution of X, and a discriminator network $\mathcal{C}_\tau$ tries to predict the probability of a random sample $\boldsymbol{x}_i$ coming from the true distribution $p(X)$ instead of being generated through $\mathcal{G}_\eta$. The generator $\mathcal{G}_\eta$ is trained by maximizing the probability of the discriminator making a mistake. This is formulated as the minimax optimization problem

$$\min_{\mathcal{G}_\eta} \max_{\mathcal{C}_\tau} \;\; \mathrm{E}_{p(X)} \left\{ \log \mathcal{C}_\tau(X) \right\} + \mathrm{E}_{p(Z)} \left\{ (1 - \mathcal{C}_\tau(\mathcal{G}_\eta(Z))) \right\}. \tag{98}$$

GANs are more flexible and have shown better performance at approximating complex distributions such as natural images (leading to sharper results) when compared to VAEs [173]. However, GANs are also much harder to train [175]. Moreover, VAEs naturally offer a way to obtain the latent representations corresponding to samples $\boldsymbol{x}_i \sim p(X)$ by mapping $X \mapsto Z$ using the function $\mathcal{D}_\tau$, which is also called an encoder model. This property and their stable training have motivated us to use VAEs in this work.

[1] Capacity of a distribution is a generic term to describe how complex a relationship it can model.

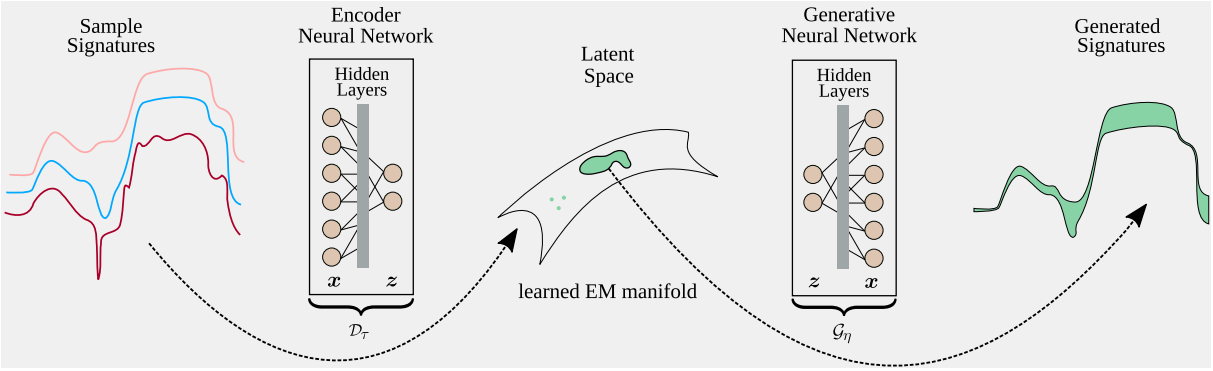


Figure 33 – Illustration of the proposed Deep Generative Endmember Model.

7.4 A DEEP GENERATIVE ENDMEMBER MODEL

In this section, we propose to model the distribution of EM spectral variability using a deep generative model. By doing so, we can explicitly explore a common property of spectral variability: the EM spectra are usually confined to a low-dimensional manifold. This property is due to the fact that the spectral signature of many materials is a function of a few photometric or chemical properties of the medium. Prominent examples include packed particle spectra as a function of its roughness, size and density [37], leaf reflectance spectra as a function of various biophysical parameters [97], and soil reflectance as a function of moisture conditions [101].

7.4.1 The steps of the proposed SU method

We assume the existence of nonlinear functions $\mathcal{G}_{\eta_p}$, $p = 1, \ldots, P$ (the generative model) that map latent representations $\boldsymbol{z}_p$ into their corresponding spectral signatures $\boldsymbol{m}_p$. We assume also the existence of encoder models $\mathcal{D}_{\tau_p}$ that map spectral signatures into their latent representations. In other words, we assume that any arbitrary observation $\boldsymbol{m}_p$ of a spectral signature of a material belongs to the set

$$\boldsymbol{m}_p \in \left\{ \mathcal{G}_{\eta_p}(\boldsymbol{z}_p) \ : \ \boldsymbol{z}_p \in \mathbb{R}^H \right\} \tag{99}$$

and thus can be equivalently represented by a corresponding low-dimensional vector $\boldsymbol{z}_p \in \mathbb{R}^H$ in the latent space of the generative model $\mathcal{G}_{\eta_p}$. This reasoning is illustrated in Fig. 33, where the encoder function $\mathcal{D}_\tau$ maps the input EM signature to the low-dimensional manifold. Reciprocally, low-dimensional vectors in the latent space can be mapped (decoded) to their corresponding spectral signatures using $\mathcal{G}_\eta$.

As such, we can formulate EM estimation in the SU problem in the latent domain (as opposed to the input spectral space), which is of a much lower order. Moreover, this approach will keep the physical interpretation of the model, provided that we have relevant training data to learn the generative models [177, 178, 179, 180]. This strategy relies on the existence of a priori training data for each material in the image, which might come

in the form of, e.g., spectral libraries of laboratory measurements [181]. Nevertheless, we propose a more practical and effective approach to train the generators $\mathcal{G}_{\eta_p}$ by exploring information contained in multiple pure pixels extracted from the observed HI. The presence of multiple pure pixels in an observed HI is a characteristic of many real scenes, and can be leveraged to help in estimating the EM models, thus, reducing the ill-posedness of the SU problem[2].

Therefore, we propose to break the unmixing problem into a sequence of two problems:

i) Using pure pixel information extracted from the HI by standard EM extraction methods, learn the generative and encoder models, $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$, for all EMs in the scene ($p = 1, \ldots, P$).

ii) Using the learned generative models, solve the SU problem by estimating the latent EM representations $\boldsymbol{Z}_n = [\boldsymbol{z}_{1,n}, \ldots, \boldsymbol{z}_{P,n}]$ and the fractional abundance vectors $\boldsymbol{a}_n$ that can best represent the observed hyperspectral data, for all pixels in the scene ($n = 1, \ldots, N$).

7.4.2 Learning the generative and encoder models $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$

The objective of this first problem is to estimate the generative and encoding models $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$, for $p = 1, \ldots, P$. We assume the knowledge of a set $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ of pure pixels for the p-th EM, for all $p = 1, \ldots, P$. Multiple pure pixels exist in many scenes, and can be directly extracted from the observed HI using automated EM extraction techniques [182, 183]. The sets of pure pixels $\boldsymbol{\Upsilon}_{\mathcal{P},p}$, which can be seen as observations from the statistical distribution of each EM, are then used in the form of training data to learn the models $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$ using a VAE [172]. If the set $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ is representative of the variability of the p-th material, the learned generative model $\mathcal{G}_{\eta_p}$ will be able to accurately describe the manifold of the p-th EM variability. Doing the same for all $p = 1, \ldots, P$ yields a set of variability models for all the EM spectra.

Although the extraction of multiple pure pixels from observed HIs is a well-established technique used to produce EM libraries [182], mixed pixels can sometimes be mistakenly identified as a pure pixel of some of the EMs. This constitutes a problem for library-based SU applications (e.g. sparse SU) since some of the library spectra may end up not being representative of their EM class (material).

The smooth nature of the latent representation of VAEs allows the mitigation of this problem in the proposed approach. Assuming the availability of a reference EM matrix $\boldsymbol{M}_0$ of correctly identified signatures (which can be obtained using any EM extraction method) and of a set of encoder models $\mathcal{D}_{\tau_p}$, we can compute the latent representation of

[2] Pure pixels are defined here as a set of pixels whose spectral distance relative to the reference EMs in $\boldsymbol{M}_0$ is less than a specified threshold.

these reference signatures of each EM as

$$\begin{aligned}\boldsymbol{Z}_0 &= \left[\boldsymbol{z}_{1,0}, \ldots, \boldsymbol{z}_{P,0}\right] \\ &= \left[\mathcal{D}_{\tau_1}(\boldsymbol{m}_{1,0}), \ldots, \mathcal{D}_{\tau_P}(\boldsymbol{m}_{P,0})\right], \end{aligned} \tag{100}$$

where $\boldsymbol{m}_{p,0}$ is the p-th column of $\boldsymbol{M}_0$. The latent representation $\boldsymbol{z}_{p,0}$ can be used as a reference latent code for the p-th material. Thus, we can measure how close an estimated EM latent representation $\boldsymbol{z}_p$ is to the latent representation of a pure pixel by evaluating its Euclidean distance to $\boldsymbol{z}_{p,0}$. This can be performed since the output of VAEs have been shown to vary smoothly with changes of the latent variable [172]. Thus, we can use $\boldsymbol{Z}_0$ to regularize the SU problem to prevent $\mathcal{G}_{\eta_p}(\boldsymbol{z}_p)$ from representing mixed pixels. This increases the robustness of the proposed approach.

7.4.3 Extracting sets of pure pixels from the observed HI

An important part of the proposed methodology consists in the extraction of the sets of pure pixels $\boldsymbol{\Upsilon}_{\mathcal{P},p}$, $p = 1, \ldots, P$ from the observed hyperspectral image $\boldsymbol{Y}$. Although different strategies have been proposed for image-based library construction (see e.g. [184, 185]), these techniques depend on multiple parameters that must be carefully adjusted in order to obtain good results. Instead of these approaches, we adopt a very simple strategy to select pure pixels from an HI that makes use of the reference matrix $\boldsymbol{M}_0$ extracted from the image using a pure-pixel-based endmember extraction algorithm (e.g. VCA [131]), which will also later be used to construct $\boldsymbol{Z}_0$ in (100). We simply select as the elements of $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ the U_p image pixels that have the smallest spectral angle to the reference signature in the p-th column of $\boldsymbol{M}_0$, where U_p is the cardinality of $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ for $p = 1, \ldots, P$. Although the success of this strategy depends on having a reasonably accurate estimation of $\boldsymbol{M}_0$, we experimentally found it to be more robust and easier to adjust than, for instance, the one in [182].

7.5 THE UNMIXING ALGORITHM

Given a set of generative models $\mathcal{G}_{\eta_p} : \mathbb{R}^H \to \mathbb{R}^L$, $p = 1, \ldots, P$ for each EM in the scene, a latent space representation $\boldsymbol{Z}_0$ of a reference EM matrix $\boldsymbol{M}_0$, and an HI $\boldsymbol{Y} = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N]$, the SU problem can be cast as the minimization of a risk functional of the form

$$\mathcal{J}(\boldsymbol{A}, \mathfrak{Z}) = \frac{1}{2} \sum_{n=1}^{N} \|\boldsymbol{y}_n - \widetilde{\mathcal{G}}(\boldsymbol{Z}_n)\boldsymbol{a}_n\|_F^2 + \mathcal{R}(\boldsymbol{A}) + \mathcal{R}(\mathfrak{Z}), \tag{101}$$

where $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N] \in \mathbb{R}^{P \times N}$ is the abundance matrix, $\mathfrak{Z} \in \mathbb{R}^{N \times P \times H}$ is a 3-D tensor obtained by stacking all pixel-dependent latent EM representations $\boldsymbol{Z}_n$, such that $[\mathfrak{Z}]_{n,:,:} = \boldsymbol{Z}_n$, $\mathcal{R}(\boldsymbol{A})$ and $\mathcal{R}(\mathfrak{Z})$ are regularization terms to improve the problem conditioning, and

the matrix-valued function $\widetilde{\mathcal{G}}(\boldsymbol{Z}_n)$ defined as

$$\widetilde{\mathcal{G}}(\boldsymbol{Z}_n) = \big[\mathcal{G}_{\eta_1}(\boldsymbol{z}_{1,n}), \ldots, \mathcal{G}_{\eta_P}(\boldsymbol{z}_{P,n})\big], \;\; n = 1, \ldots, N\,, \tag{102}$$

is the concatenation of the generative functions for each EM.

The term $\mathcal{R}(\boldsymbol{A})$ is a regularization functional that aims to provide spatial smoothness and to enforce positivity and sum-to-one constraints to the abundances. It is given by [34]

$$\mathcal{R}(\boldsymbol{A}) = \lambda_A\big(\|\mathcal{H}_h(\boldsymbol{A})\|_{2,1} + \|\mathcal{H}_v(\boldsymbol{A})\|_{2,1}\big) + \iota_{\mathcal{S}^1}(\boldsymbol{A})\,, \tag{103}$$

where parameter λ_A controls the contribution of this term to the cost function. The first two terms are a spatial regularizers over $\boldsymbol{A}$, where $\mathcal{H}_h$ and $\mathcal{H}_v$ are linear operators that compute the first-order horizontal and vertical gradients of a bidimensional signal, acting separately for each material of $\boldsymbol{A}$, and $\|\cdot\|_{2,1}$ is the $\mathcal{L}_{2,1}$ norm, defined as $\|\boldsymbol{X}\|_{2,1} = \sum_{n=1}^{N}\|\boldsymbol{x}_n\|_2$. The term $\iota_{\mathcal{S}^1}(\boldsymbol{A})$ is the indicator function of the unity simplex, i.e. $\iota_{\mathcal{S}^1}(\boldsymbol{A}) = 0$ if $\boldsymbol{A} \in \mathcal{S}^1$ and $\iota_{\mathcal{S}^1}(\boldsymbol{A}) = \infty$ otherwise, where

$$\mathcal{S}^1 = \big\{\boldsymbol{A} \in \mathbb{R}^{P\times N} \,:\, \boldsymbol{A} \geqslant 0,\, \mathbf{1}^\top\boldsymbol{A} = \mathbf{1}^\top\big\}\,. \tag{104}$$

The term $\mathcal{R}(\mathfrak{Z})$ constrains the EM latent representations $\mathfrak{Z}$ to be close to the latent representation $\boldsymbol{Z}_0$ of the reference EM matrix $\boldsymbol{M}_0$. It is given by

$$\mathcal{R}(\mathfrak{Z}) = \frac{\lambda_Z}{2}\sum_{n=1}^{N}\|\boldsymbol{Z}_n - \boldsymbol{Z}_0\|_F^2\,, \tag{105}$$

where parameter λ_Z controls the contribution of this term to the cost function. This regularization makes the estimation problem more robust to the selection of the training data $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ by assuring the closeness of the estimated latent codes $\mathfrak{Z}$ and the representations of pure pixels of each class. However, it relies indirectly on the reference EM signatures $\boldsymbol{M}_0$ (which are extracted from the observed HI with endmember extraction algorithms) being adequate representatives of their material classes in order to provide a good performance.

The optimization problem then becomes

$$(\widehat{\boldsymbol{A}}, \widehat{\mathfrak{Z}}) = \underset{\boldsymbol{A},\,\mathfrak{Z}}{\arg\min}\;\; \mathcal{J}(\boldsymbol{A}, \mathfrak{Z})\,. \tag{106}$$

The problem defined in (106) is non-smooth and non-convex if solved simultaneously with respect to both variables $\boldsymbol{A}$, and $\mathfrak{Z}$. However, an approximate solution can be found by minimizing (106) iteratively with respect to each variable, leading to the Deep Generative Unmixing (DeepGUn) method described in Algorithm 6. The DeepGUn algorithm consists of two distinctive steps. First, the generative endmember models generative and encoder models $\mathcal{G}_{\eta_p}$, $\mathcal{D}_{\tau_p}$, $p = 1, \ldots, P$ are trained based on the pure pixels $\boldsymbol{\Upsilon}_{\mathcal{P},p}$, $p = 1, \ldots, P$ extracted from the observed HI and $\boldsymbol{Z}_0$ is computed. Afterwards, the alternating minimization approach is applied to compute the abundance maps and the latent representations of the endmembers for each pixel. We next describe the details of each optimization step. Implementation details are described in Sections 7.6 and 7.7.

Algorithme 6 : DeepGUn algorithm for solving (106)

Input : $\boldsymbol{Y}$, λ_Z, and λ_A.
Output : $\widehat{\boldsymbol{A}}$ and $\widehat{\mathscr{M}}$.

Estimate the reference EM signatures $\boldsymbol{M}_0$ using an EM extraction method (e.g. VCA);
Estimate $\boldsymbol{A}^{(0)}$ using a standard LMM-based SU method;
Extract sets of pure pixels $\boldsymbol{\Upsilon}_{\mathcal{P},p}$, $p = 1, \ldots, P$ from the HI using a bundle extraction strategy;
Train the generative and encoder models $\mathcal{G}_{\eta_p}$, $\mathcal{D}_{\tau_p}$, $p = 1, \ldots, P$;
Compute the latent representation of $\boldsymbol{M}_0$ as $\boldsymbol{Z}_0 = [\mathcal{D}_{\tau_1}(\boldsymbol{m}_{1,0}), \ldots, \mathcal{D}_{\tau_P}(\boldsymbol{m}_{P,0})]$;
Set $i = 0$;
while *stopping criterion is not satisfied* **do**
  $i = i + 1$;
  $\mathfrak{Z}^{(i)} = \arg\min_{\mathfrak{Z}} \ \mathcal{J}(\boldsymbol{A}^{(i-1)}, \mathfrak{Z})$;
  $\boldsymbol{A}^{(i)} = \arg\min_{\boldsymbol{A}} \ \mathcal{J}(\boldsymbol{A}, \mathfrak{Z}^{(i)})$;
end
for $n = 1, \ldots, N$, **do**, $[\widehat{\mathscr{M}}]_{:,:,n} = \widetilde{\mathcal{G}}([\mathfrak{Z}]_{:,:,n})$, **end**;
return $\widehat{\boldsymbol{A}} = \boldsymbol{A}^{(i)}$, $\widehat{\mathscr{M}}$;

7.5.1 Optimization with respect to $\mathfrak{Z}$

Rewriting (106) considering only the terms in (101) that depend on $\mathfrak{Z}$, the problem becomes

$$\min_{\mathfrak{Z}} \ \frac{1}{2} \sum_{n=1}^{N} \left(\|\boldsymbol{y}_n - \widetilde{\mathcal{G}}(\boldsymbol{Z}_n)\boldsymbol{a}_n\|_F^2 + \lambda_Z \|\boldsymbol{Z}_n - \boldsymbol{Z}_0\|_F^2 \right). \tag{107}$$

This is a regularized nonlinear least squares problem, which can be solved individually for each pixel $\boldsymbol{y}_n$. Thus, (107) can be decomposed into N non-convex, nonlinear optimization problems with dimensionality $H \times P$ by denoting each summand in (107) by $\widetilde{\mathcal{J}}^{(n)}$, $n = 1, \ldots, N$. We solve each of those problems $\widetilde{\mathcal{J}}^{(n)}$ using a quasi-Newton algorithm, described in Algorithm 7, which provides an efficient solution for high-dimensional functions $\widetilde{\mathcal{G}}$ [186].

Although problem (107) is generally non-convex, recent research [187] has proven that, under suitable assumptions on the generator network $\widetilde{\mathcal{G}}$, the problem of recovering the latent variable $\boldsymbol{Z}_n$ does not have any stationary point (e.g. local minima or saddle points) outside a small neighborhood of the desired solution and its negative scalar multiple. This indicates the existence of a favorable global geometry of (107).

Note that $\widetilde{\mathcal{G}}$ is not necessarily differentiable with respect to the latent representations $\boldsymbol{Z}_n$, which can make the optimization problem more challenging. Nonetheless, quasi-Newton algorithms show excellent performance at non-smooth problems [188], where convergence is generally observed as long as the line search procedure does not return a point at which the objective function is non-differentiable. This allows quasi-Newton

algorithms to be directly applied to obtain approximate solutions to non-smooth problems with good computational efficiency [188, 189].

Algorithme 7 : Quasi-Newton algorithm for solving (107)

Input : $\boldsymbol{a}_n$, $\boldsymbol{y}_n$, λ_Z, $\boldsymbol{Z}_0$ and $\widetilde{\mathcal{J}}^{(n)}$.
Output : $\boldsymbol{Z}_n$.

Set $i = 0$ and $\boldsymbol{B}_1 = \boldsymbol{I}$;
while *stopping criterion is not satisfied* **do**
  $i = i + 1$;
  Compute search direction $\boldsymbol{p}_i = -\boldsymbol{B}_i \nabla \widetilde{\mathcal{J}}_i^{(n)}$;
  Set $\boldsymbol{z}_{i+1} = \mu_i \boldsymbol{p}_i$, where μ_i is computed using a line search procedure to satisfy the Wolfe conditions;
  Define $\boldsymbol{s}_i = \boldsymbol{z}_{i+1} - \boldsymbol{z}_i$ and $\boldsymbol{u}_i = \nabla \widetilde{\mathcal{J}}_{i+1}^{(n)} - \nabla \widetilde{\mathcal{J}}_i^{(n)}$;
  $\boldsymbol{B}_{i+1} = \boldsymbol{B}_i - \frac{\boldsymbol{B}_i \boldsymbol{s}_i \boldsymbol{s}_i^\top \boldsymbol{B}_i}{\boldsymbol{s}_i^\top \boldsymbol{B}_i \boldsymbol{s}_i} + \frac{\boldsymbol{u}_i \boldsymbol{u}_i^\top}{\boldsymbol{u}_i^\top \boldsymbol{s}_i}$;
end
Reorder $\boldsymbol{z}_{i+1}$ as a matrix $\boldsymbol{Z}_n$;
return $\widehat{\boldsymbol{Z}}_n = \boldsymbol{Z}_n$;

7.5.2 Optimization with respect to the abundances

Restating (106) considering only the terms in (101) that depend on $\boldsymbol{A}$ leads to

$$\begin{aligned} \min_{\boldsymbol{A}} \; & \frac{1}{2} \sum_{n=1}^{N} \|\boldsymbol{y}_n - \widetilde{\mathcal{G}}(\boldsymbol{Z}_n)\boldsymbol{a}_n\|_F^2 + \iota_{\mathcal{S}^1}(\boldsymbol{A}) \\ & + \lambda_A (\|\mathcal{H}_h(\boldsymbol{A})\|_{2,1} + \|\mathcal{H}_v(\boldsymbol{A})\|_{2,1}) \,. \end{aligned} \tag{108}$$

Since the latent variables $\boldsymbol{Z}_n$ are fixed, (108) consists of a SU problem with a pixel-dependent EM matrix and an edge-preserving spatial regularization. Although this problem is not separable with respect to each pixel in the image, the ADMM can be used to obtain an efficient solution [134]. The solution of (108) using the ADMM is well described elsewhere (e.g. [19]) and will thus be omitted here for conciseness.

7.6 NEURAL NETWORK ARCHITECTURE

As discussed before, we used a VAE [172] to learn the generative and encoder models $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$ from the sets of pure pixels $\boldsymbol{\Upsilon}_{\mathcal{P},p}$. Compared to GANs, the training of VAEs is much simpler and more stable [175]. Moreover, VAEs naturally return the encoder model $\mathcal{D}_{\tau_p}$ as an approximation to the posterior distribution when learning $\mathcal{G}_{\eta_p}$. We have selected a dimension $H = 2$ for the latent space, as it was experimentally verified to be sufficient to adequately capture the variability of each single material in a scene.

Table 11 – Encoder network architecture.

Layer	Activation Function	Number of units
Input	—	L
Hidden # 1	ReLU	$\lceil 1.2 \times L \rceil + 5$
Hidden # 2	ReLU	$\max\left\{\lceil L/4 \rceil, H + 2\right\} + 3$
Hidden # 3	ReLU	$\max\left\{\lceil L/10 \rceil, H + 1\right\}$

For the network architectures, we selected the number of layers and neurons according to the autoencoder implementation in [190, 165], with three hidden layers using ReLU activations (defined as $ReLU(x) = \max(x, 0)$) in the hidden layers, which are described in more detail in Tables 11 and 12.

We found that this configuration led to spectrally smooth generated signatures, and was effective at generalizing well with small training sample sizes. We trained the network for 50 epochs with the Adam optimizer [191] in TensorFlow, using a batch optimization with mini-batch size equal to one third of the total amount of training data for each EM.

Table 12 – Decoder network architecture.

Layer	Activation Function	Number of units
Hidden # 1	ReLU	$\max\left\{\lceil L/10 \rceil, H + 1\right\}$
Hidden # 2	ReLU	$\max\left\{\lceil L/4 \rceil, H + 2\right\} + 3$
Hidden # 3	ReLU	$\lceil 1.2 \times L \rceil + 5$
Output	Sigmoid	L

7.7 EXPERIMENTAL RESULTS

In this section, simulation results using both synthetic and real data illustrate the performance of the proposed method. We compare the proposed DeepGUn method with the FCLS, the PLMM [20], the ELMM [19], and the GLMM [34]. In all experiments, the VCA algorithm [131] was used to extract the reference EM matrix $\boldsymbol{M}_0$ from the observed HI and to initialize the different methods. The abundance maps of all methods were initialized using the results obtained by the FCLS algorithm. The sets $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ of pure pixels were constructed by selecting the 100 image pixels $\boldsymbol{y}_n$ with the smallest spectral angles relative to the reference EMs in $\boldsymbol{M}_0$. We ran the alternating optimization process in Algorithm 6 for at most 10 iterations or until the relative change of $\boldsymbol{A}$ and $\mathfrak{L}$ was less than 10^{-3}. The iterative procedure in Algorithm 7 was run until the relative change of $\boldsymbol{z}_i$ was less than 10^{-3}. The performances were evaluated using the Normalized Root Means Squared Error (NRMSE) between the estimated abundance maps ($\text{NRMSE}_{\boldsymbol{A}}$), between the EM matrices ($\text{NRMSE}_{\mathcal{M}}$) and between the reconstructed images ($\text{NRMSE}_{\boldsymbol{Y}}$). The

Table 13 – Simulation results using synthetic data.

	$\text{NRMSE}_{\boldsymbol{A}}$	$\text{NRMSE}_{\mathscr{M}}$	$\text{SAM}_{\mathscr{M}}$	$\text{NRMSE}_{\boldsymbol{Y}}$	Time [s]
Data Cube 1 – DC1					
FCLS	0.2854	—	—	0.0350	0.71
PLMM	0.2604	0.1075	0.0440	**0.0007**	122.09
ELMM	0.2554	0.1032	0.0398	0.0321	8.82
GLMM	0.2480	0.1036	0.0355	0.0235	23.74
DeepGUn	**0.0566**	**0.0944**	**0.0233**	0.0448	75.20
Data Cube 2 – DC2					
FCLS	0.1294	—	—	0.0393	0.38
PLMM	0.1197	0.0481	0.0378	0.0336	41.31
ELMM	0.1110	0.0566	0.0382	0.0231	20.25
GLMM	0.1146	0.0534	0.0367	**0.0226**	17.03
DeepGUn	**0.0969**	**0.0463**	**0.0323**	0.0384	36.40
Data Cube 3 – DC3					
FCLS	0.2606	—	—	0.0542	0.34
PLMM	0.2028	0.0928	0.0385	0.0302	59.88
ELMM	0.1997	0.0640	0.0188	0.0238	17.99
GLMM	0.1841	0.0638	0.0185	**0.0226**	25.74
DeepGUn	**0.1613**	**0.0600**	**0.0172**	0.0457	48.96
Data Cube 4 – DC4					
FCLS	0.5109	—	—	0.1712	0.50
PLMM	0.5066	0.6245	0.4874	0.0320	269.48
ELMM	0.4385	0.4712	0.1451	**0.0106**	18.36
GLMM	0.4371	0.4855	0.1972	0.0108	21.15
DeepGUn	**0.2550**	**0.2918**	**0.0873**	0.1403	99.94

NRMSE between a true, generic tensor $\mathcal{X}$ and its estimate $\widehat{\mathcal{X}}$ is defined as

$$\text{NRMSE}_{\mathcal{X}} = \sqrt{\frac{\|\mathcal{X} - \widehat{\mathcal{X}}\|_F^2}{\|\mathcal{X}\|_F^2}}\,. \tag{109}$$

Note that for the case of $\text{NRMSE}_{\boldsymbol{Y}}$, the reconstructed image $\widehat{\boldsymbol{Y}}$ is given by $[\widehat{\boldsymbol{Y}}]_{:,n} = [\widehat{\mathscr{M}}]_{:,:,n}[\widehat{\boldsymbol{A}}]_{:,n}$, $n = 1, \ldots, N$.

We consider also the Spectral Angle Mapper to evaluate the estimated EMs

$$\text{SAM}_{\mathscr{M}} = \frac{1}{N}\sum_{n=1}^{N}\sum_{p=1}^{P} \arccos\left(\frac{\boldsymbol{m}_{p,n}^{\top}\widehat{\boldsymbol{m}}_{p,n}}{\|\boldsymbol{m}_{p,n}\|\|\widehat{\boldsymbol{m}}_{p,n}\|}\right). \tag{110}$$

where $\boldsymbol{m}_{p,n}$ and $\widehat{\boldsymbol{m}}_{p,n}$ are the true and the estimated signatures of the p-th endmember in the n-th pixel, respectively.

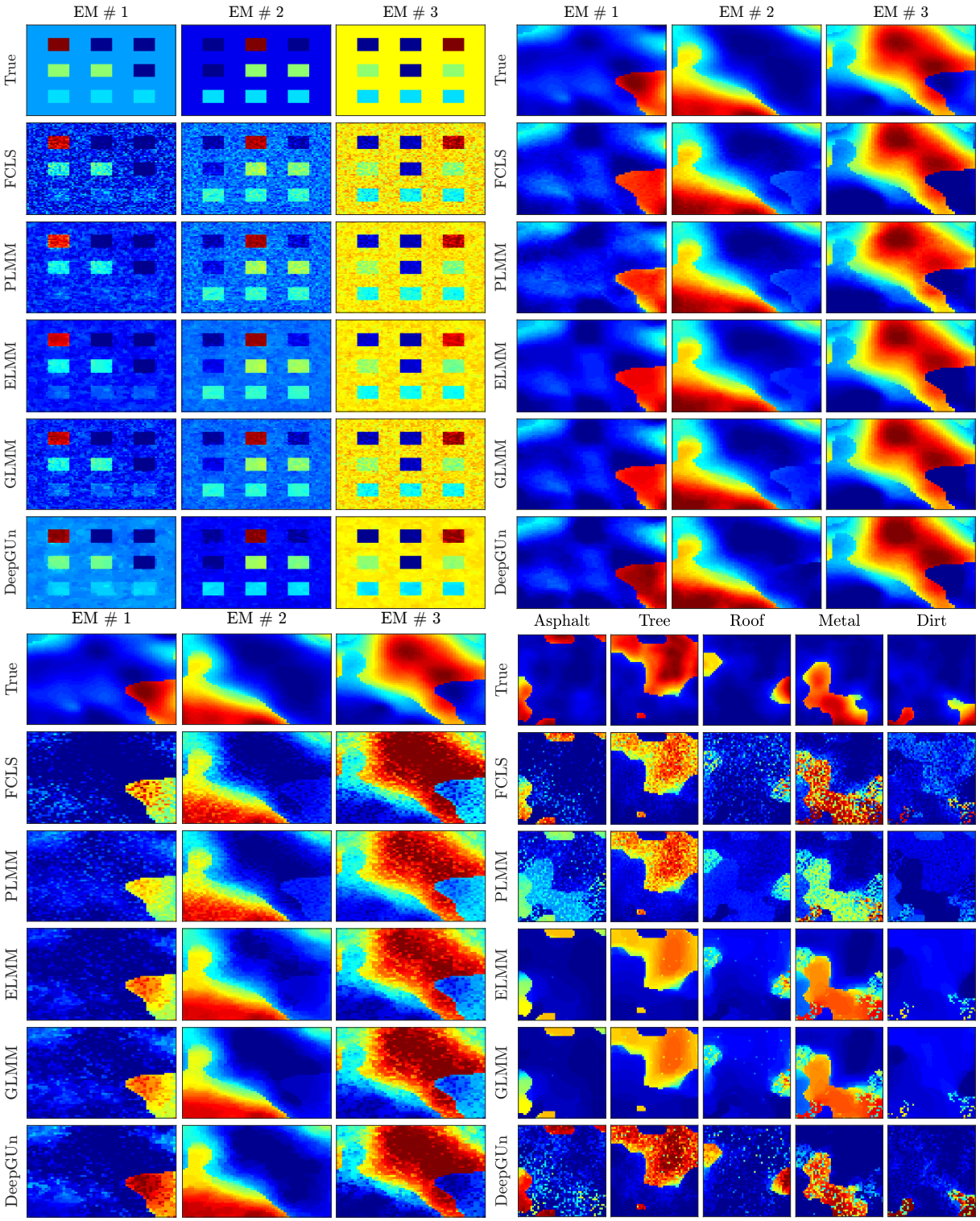


Figure 34 – Abundance maps of DC1 (top left), DC2 (top right), DC3 (bottom left) and DC4 (bottom right) for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

7.7.1 Synthetic data

To quantitatively compare the different algorithms, four synthetic datasets were created, namely Data Cubes 1–4 (DC1–DC4), with 70×70 pixels (DC1) and 50×50 pixels

(DC2, DC3 and DC4). These datasets were built using three (DC1, DC2 and DC3) and five (DC4) 224-band EMs extracted from the USGS Spectral Library [136] and spatially correlated abundance maps, as depicted in Fig. 34.

Spectral variability of the EMs was imposed using four different models. For the DC1 datacube, we adopted the variability model used in [20], consisting of pixelwise multiplicative spectral factors given by random piecewise-linear functions. For DC2, the variability model of [34] was used, consisting of band dependent scaling factors that varied smoothly in both the spatial and spectral dimensions. For DC3, we considered a simple model introduced in [184, Section IV-A-1] to emulate errors in atmospheric compensation as a function of the viewing geometry given the direct and diffuse light on the scene, and the solar path transmittance. For datacube DC4, we used as endmembers pure pixels of five materials (asphalt, tree, roof, metal and dirt) which were manually extracted from a real hyperspectral image, thus depicting realistic spectral variability. White Gaussian noise was finally added to all datasets to yield a 30 dB SNR.

The optimal parameters for each algorithm were selected by performing grid searches for each dataset. The ranges in which the parameters were searched were selected according to those discussed by the authors in the original publications. For the PLMM we searched for α, β and γ in the ranges [0.01, 0.1, 0.35, 0.7, 1.4, 25], [10^{-9}, 10^{-5}, 10^{-4}, 10^{-3}] and [10^{-2}, 0.1, 1, 10, 10^{2}], respectively. For both ELMM and GLMM, the parameters were selected among the following values: λ_S, $\lambda_M \in$ [0.01, 0.1, 1, 5, 10, 50], $\lambda_A \in$ [0.001, 0.01, 0.05], and λ_ψ, $\lambda_\Psi \in$ [10^{-6}, 10^{-3}, 1, 10^{3}]. For the proposed DeepGUn algorithm, we fixed $\lambda_Z = 0.1$ and selected λ_A among the values [0.005, 0.01, 0.05]. For the proposed method, the sets of pure pixels for each EM $\boldsymbol{\Upsilon}_{\mathcal{P},p}$ were constructed by selecting the 100 pixels closest to the reference materials $\boldsymbol{M}_0$.

The quantitative results are shown in Table 13, with the best results for each metric marked in bold. The proposed method clearly outperformed the competing algorithms in terms of NRMSE$_{\boldsymbol{A}}$ for all four datasets. Qualitatively, the abundance maps provided by DeepGUn, displayed in Fig. 34, are clearly much closer to the true abundance maps than those provided by the other methods. These are important results, as accuracy in abundance estimation is the main objective of SU.

For the EM reconstruction metrics NRMSE$_{\mathcal{M}}$ and SAM$_{\mathcal{M}}$, DeepGUn gave the best results for all data cubes. This indicates that the proposed endmember model used by DeepGUn allows for precise material identification from the observed hyperspectral scenes.

The reconstruction error NRMSE$_{\boldsymbol{Y}}$ of the DeepGUn algorithm was comparable to the FCLS and significantly larger than that of the GLMM. This is natural since the GLMM has more degrees of freedom. However, the connection between NRMSE$_{\boldsymbol{Y}}$ and the abundance reconstruction error is far from being direct, as can be seen in Table 13.

The execution times, at the rightmost column of Table 13, indicate that the computational complexity of DeepGUn is somewhere between the complexities of GLMM and

PLMM, the two major competing algorithms. Hence, the DeepGUn method yielded superior SU performance, with easier parameter tuning, and at a reasonable computational cost.

7.7.2 Influence of the latent dimension H

An important parameter in the design of the DeepGUn method is the dimensionality H of the latent space of the generative endmember models $\mathcal{G}_i$, $i = 1, \ldots, P$. The individual dimensions of the latent space are used to represent changes in the endmember signatures due to spectral variability. Since in a given scene the endmembers are likely to be affected only by a small number of effects (hence the hypothesis that they are supported at a low-dimensional manifold), H should be small in order to avoid introducing spurious effects and increasing the computational complexity of the solution.

To illustrate this, we performed a simulation with DeepGUn where we varied the dimensionality of the latent space H and measured the normalized abundance reconstruction error $\text{NRMSE}_{\boldsymbol{A}}$. For this, we considered the data cubes DC1 and DC2 described above in Section 7.7.1 . The results are depicted in Fig. 35, and show that the abundance estimation error tends to increase with H. It can also be seen that there is a sharper increase in $\text{NRMSE}_{\boldsymbol{A}}$ for DC1 when compared to DC2. This is likely due to the fact that the dataset DC2 uses a more complex model to generate endmember variability. This indicates that the selection of a small value for H is important to obtain good unmixing results.

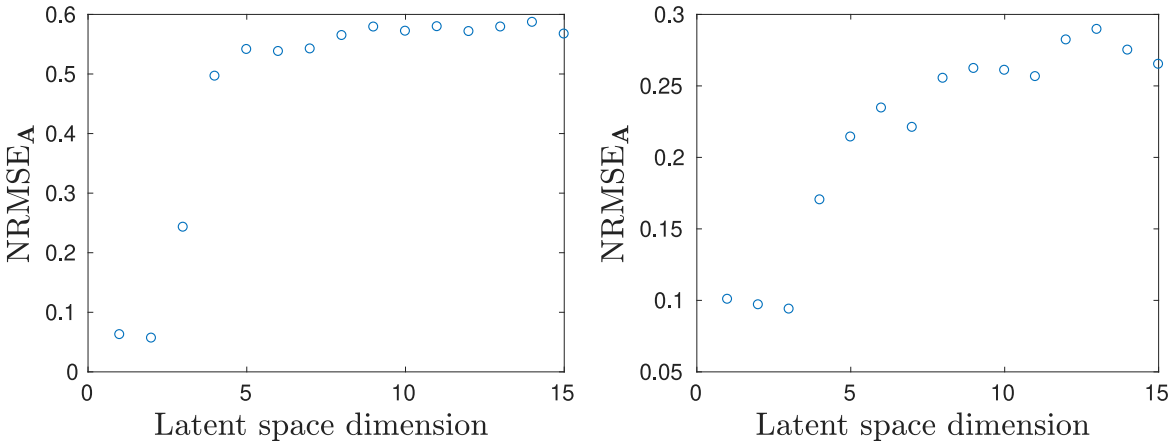


Figure 35 – Abundance NRMSE as a function of the latent space dimension H for datacubes DC1 (left) and DC2 (right).

7.7.3 Real data

We considered the Houston, Samson and the Jasper Ridge datasets for the simulations with real data. These datasets were captured by the AVIRIS instrument, and originally had 224 bands. The spectral bands corresponding to water absorption and low

Table 14 – Simulation results using real data.

	Houston HI		Samson HI		Jasper Ridge HI	
	$\text{NRMSE}_{\boldsymbol{Y}}$	Time [s]	$\text{NRMSE}_{\boldsymbol{Y}}$	Time [s]	$\text{NRMSE}_{\boldsymbol{Y}}$	Time [s]
FCLS	0.2470	2.56	0.0545	1.38	0.2057	1.52
PLMM	0.0713	663.25	0.0239	103.84	0.0553	220.84
ELMM	0.0171	38.30	0.0119	14.76	0.0278	27.08
GLMM	**0.0016**	48.53	**0.0006**	46.69	**0.0019**	86.33
DeepGUn	0.2355	259.61	0.0862	121.88	0.1094	209.64

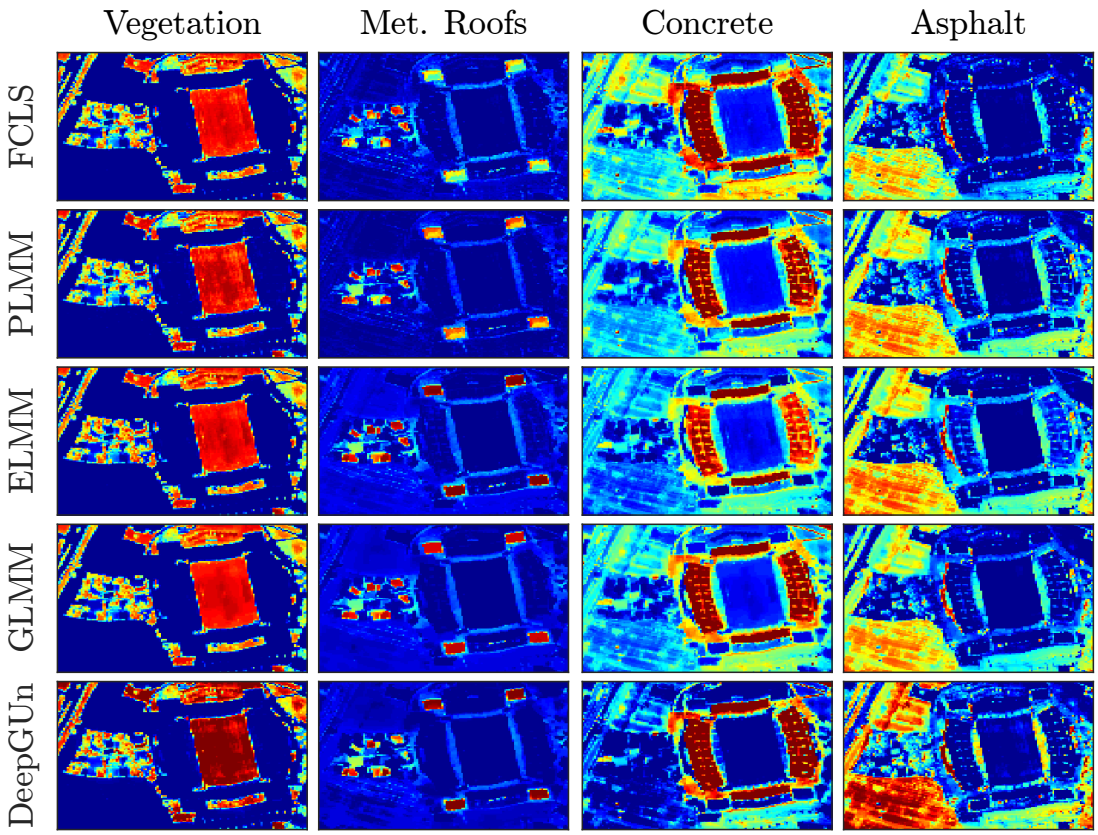


Figure 36 – Abundance maps of the Houston dataset for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

SNR regions were removed, resulting in 188 bands for the Houston image, 156 bands for the Samson image and 198 bands for the Jasper Ridge image. Previous studies indicate that the Houston HI has four predominant EMs [19], while the Samson and Jasper Ridge HIs are known to have three and four EMs, respectively [32].

The reconstructed abundance maps for both datasets and all algorithms are shown in Figs. 36, 37 and 38. For the Houston dataset, the last row of Fig. 36 shows that the abundance maps provided by the DeepGUn method better evidence the strong vegetation and concrete abundances at the stadium field and stands, respectively, as well as the stronger asphalt abundances in the parking lot. For the Samson and Jasper Ridge images, a clear performance improvement can be seen for the DeepGUn algorithm. Note, for instance,

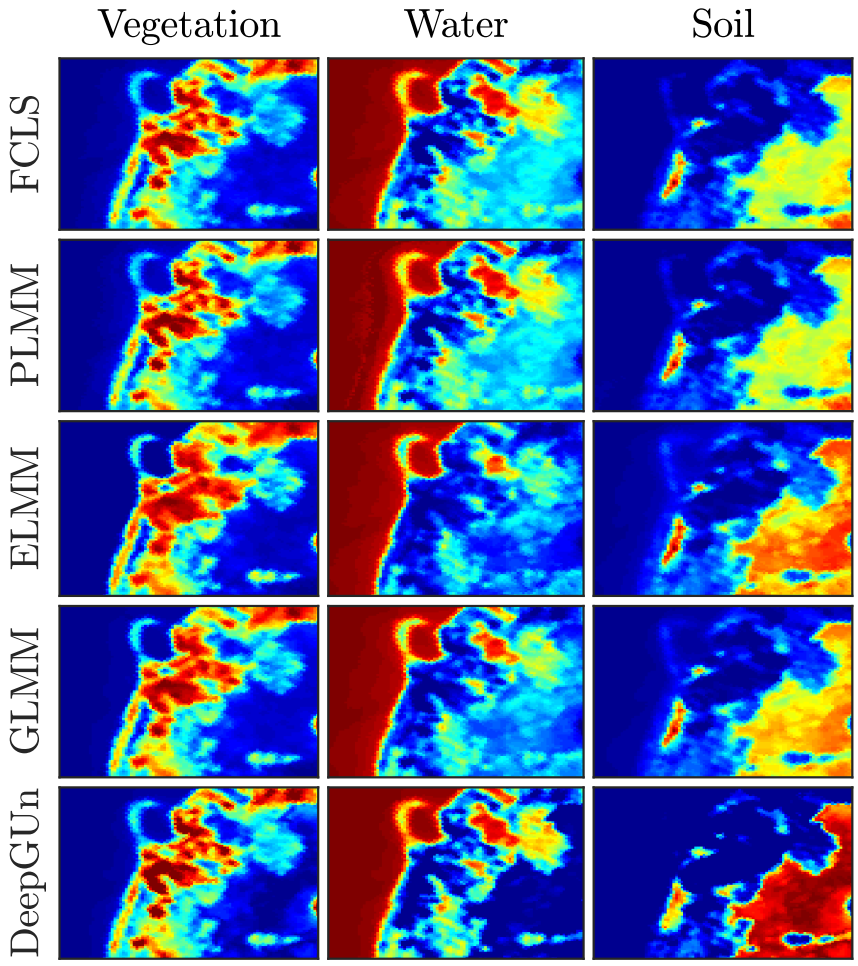


Figure 37 – Abundance maps of the Samson dataset for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

a smaller confusion between the Water and Soil EMs in the Samson HI when compared to the other methods. Similarly, for the Jasper Ridge HI, the DeepGUn method leads to considerably stronger Water abundances in the region containing the river. Moreover, although the ELMM provided a better estimation of the road in the scene when compared to the remaining methods, it also resulted in a greater confusion between the Vegetation and Soil EMs, especially in the right part of the scene.

The quantitative results for all algorithms and datasets are shown in Table 14. Since the correct abundance values (the ground truth) are not available for most real images, the reconstruction error $\text{NRMSE}_{\boldsymbol{Y}}$ has been used as a sort of quality verification. As was the case for the synthetic data, the DeepGUn reconstruction errors are higher than those yielded by other methods that address spectral variability. However, the reconstruction error is definitely not a good performance measure for abundance estimation in real images, which is the main objective of unmixing algorithms. The higher reconstruction errors of DeepGUn in this case are just due the fact that DeepGUn has much fewer degrees of freedom than the ELMM, PLMM and GLMM algorithms. In fact, the DeepGUn has only $H \times P$ degrees of freedom for each pixel, which is comparable to the FCLS (P)

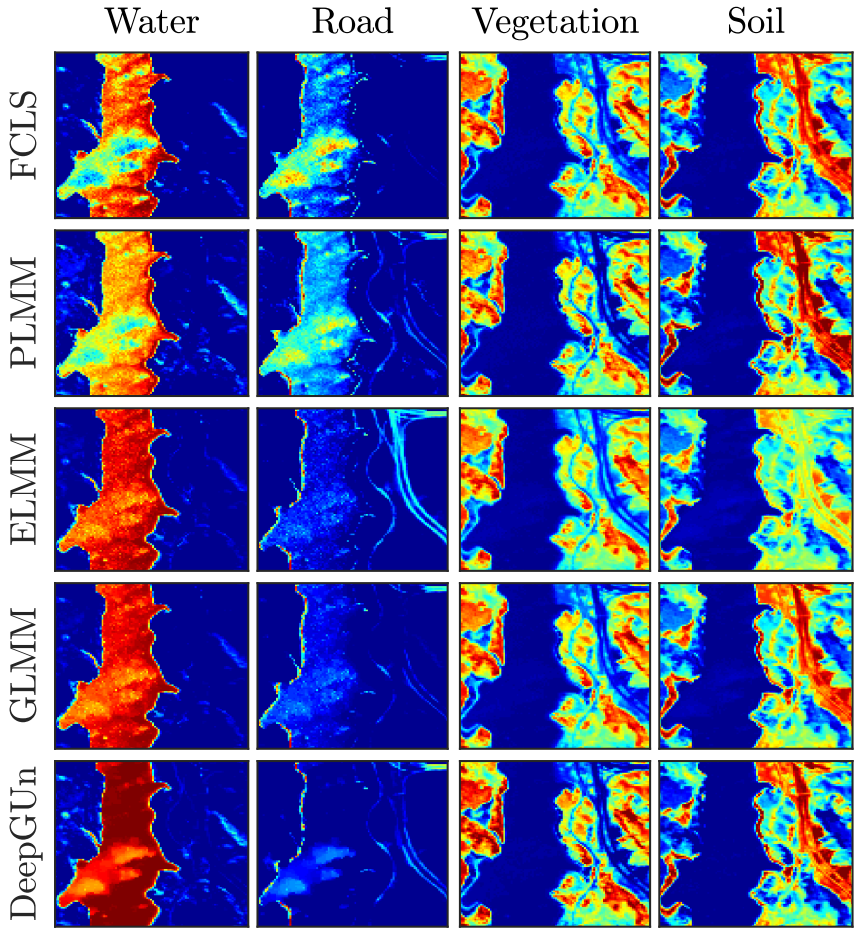


Figure 38 – Abundance maps of the Jasper Ridge dataset for all tested algorithms. Abundance values represented by colors ranging from blue ($a_k = 0$) to red ($a_k = 1$).

much smaller than the ELMM, GLMM and PLMM methods ($> L \times P$). Although this means that the ELMM, GLMM and PLMM can achieve arbitrarily small reconstruction errors $\text{NRMSE}_{\boldsymbol{Y}}$, this is not necessarily reflected as good abundance estimation results. The execution times of the proposed DeepGUn method were again comparable to those of the other algorithms addressing spectral variability, which indicates that it scales well with larger image sizes.

7.8 CONCLUSIONS

In this chapter, a deep generative EM model was proposed to address spectral variability in SU of HIs. Instead of relying on user-defined parametric EM models which have shown to be very hard to estimate in practical scenes, the proposed methodology leveraged the generalization capability of deep neural networks to accurately model EM spectra while still maintaining a strong connection to the physical mixing process. A deep generative model for each EM was trained prior to unmixing by using pure pixel information extracted directly from the observed HI, which allowed for an unsupervised formulation. The proposed EM model was then applied to solve the SU problem, which

was cast as the estimation of the low-dimensional representations of the EMs in the latent space of the deep generative models and their corresponding fractional abundances, for each pixel. The resulting DeepGUn algorithm presented excellent performance despite the simple strategy used for selecting the training data for learning the generative model. Simulations using synthetic and real data indicate that the proposed method can lead to significant improvements in abundance estimation accuracy.

8 DEEP GENERATIVE MODELS FOR LIBRARY AUGMENTATION IN MULTIPLE ENDMEMBER SPECTRAL MIXTURE ANALYSIS

8.1 INTRODUCTION

Despite the good performance of parametric EM models discussed in previous chapters, the most prominent approach to deal with EM variability in SU still considers spectral libraries acquired a priori through laboratory or *in situ* measurements [11]. Within this framework, The SU problem then becomes equivalent to selecting a subset of signatures in the libraries that can best represent the observed HI under the LMM. The methods that attempt to solve this problem can be roughly divided between sparse SU [21, 32] (which was discussed in Chapter 3) and *Multiple Endmember Spectral Mixture Analysis* (MESMA) [16] algorithms. Although sparse SU is computationally more efficient, MESMA has been widely employed in practice due to its greater simplicity and interpretability [11].

However, the quality of the MESMA results is strongly dependent on how well the spectral libraries represent the EM signatures actually present in the scene. This is a problem since spectral libraries are often small, are usually not acquired under the same conditions as the observed HI, and *in situ* measurements can be costly or impractical.

One approach to alleviate this problem consists of generating multiple synthetic samples of an endmember using a physical model (radiative transfer function – RTF) describing the variability of the spectra as a function of atmospheric or biophysical parameters [11], such as e.g. the PROSPECT of Hapke models [98, 37] for vegetation or mineral spectra. These additional signatures are then included in the library to augment it before performing SU. The use of RTFs to generate spectral libraries has great potential since it can represent spectral variability caused by different effects which are unlike to be captured by laboratory or field measurements [192, 193, 102]. However, physics-based models require accurate knowledge of the physical process governing the observation of the materials spectra by the sensor, which is hard to obtain in practice. This limits the practical interest of these methods.

Recently, deep generative models (DGMs) have seen remarkable advances in the form of variational autoencoders (VAEs) and generative adversarial networks (GANs) [172, 173]. This have made it possible to learn the distribution of complex data (e.g., natural images) efficiently, and from a limited amount of samples [194]. DGMs have been considered for data augmentation in few-sample settings for image classification problems [194]. In Chapter 7, we proposed to use DGMs learned from observed HIs in order to parametrize the variable EM spectra in the optimization step of a matrix factorization-based blind SU problem, where the EMs are estimated from the HI. This showed that using generative neural networks is a promising approach to represent the EMs in SU.

In this chapter, we propose a spectral library augmentation method for MESMA-based algorithms by leveraging the power of DGMs to represent the EMs. The main

contribution of the proposed method is that it works blindly, what allows for augmentation of the spectral libraries used with MESMA even when RTFs or physical models are unknown. The overall strategy can be divided in three steps. First we learn the statistical distribution of each EM in the scene using the spectral signatures contained in the existing spectral library and a DGM. Then, we sample new spectral signatures using the DGMs and augment their respective spectral libraries. Finally, we unmix the observed HI using MESMA and the augmented library. Simulations with synthetic and real data show a substantial accuracy gain in abundance estimation when comparing the proposed method with competing strategies.

This chapter is organized as follows. The MESMA algorithm is briefly revisited in Section 8.2. The proposed library augmentation strategy is presented in Section 8.3. Simulations with synthetic and real data are presented in Section 8.4. Finally, concluding remarks are presented in Section 8.5.

8.2 REVISITING SPECTRAL UNMIXING WITH MESMA

SU with MESMA considers multiple spectra libraries, or bundles, one for each endmember, and performs a search for the best fitting model within all possible combinations of endmembers. Thus, assuming prior knowledge of spectral bundles for each EM in the scene, the set $\mathcal{M}$ of endmember matrices that can be drawn from the library can be defined as

$$\mathcal{M} = \Big\{ \big[\boldsymbol{m}_1, \ldots, \boldsymbol{m}_P\big] \,:\, \boldsymbol{m}_i \in \mathcal{M}_i \,,\, i = 1, \ldots, P \Big\} \,, \tag{111}$$

where $\mathcal{M}_i = \{\boldsymbol{m}_{i,1}, \ldots, \boldsymbol{m}_{i,C_i}\}$, $\boldsymbol{m}_{i,j} \in \mathbb{R}^L$ is a set of C_i spectral signatures of the i-th material. The MESMA SU problem can be formulated as

$$\min_{\boldsymbol{M} \in \mathcal{M},\, \boldsymbol{a}_n} \big\| \boldsymbol{y}_n - \boldsymbol{M}\boldsymbol{a}_n \big\|_2^2 \;\text{ subject to }\; \boldsymbol{a}_n \geqslant \boldsymbol{0},\, \boldsymbol{1}^\top \boldsymbol{a}_n = 1 \,. \tag{112}$$

Although the MESMA algorithm has shown excellent performance when dealing with spectral variability in many practical scenarios, its performance is strongly effected by the quality of the spectral library $\mathcal{M}$ [11]. In order for MESMA to perform well, the library must be representative of the spectral library observed in a given scene. Previous works tried to address this issue by augmenting the spectral libraries using physics-based models that describe well the variability of the endmembers. See, e.g., the PROSPECT or Hapke models [98, 37].

However, a major drawback of physics-based models is the requirement of accurate knowledge of the physical process governing the observation of the materials spectra by the sensor. This detailed information is rarely available in practice, which limits the applicability of these methods. In the following, we will present a new approach for spectral library augmentation that is based on deep generative models such as VAEs and GANs.

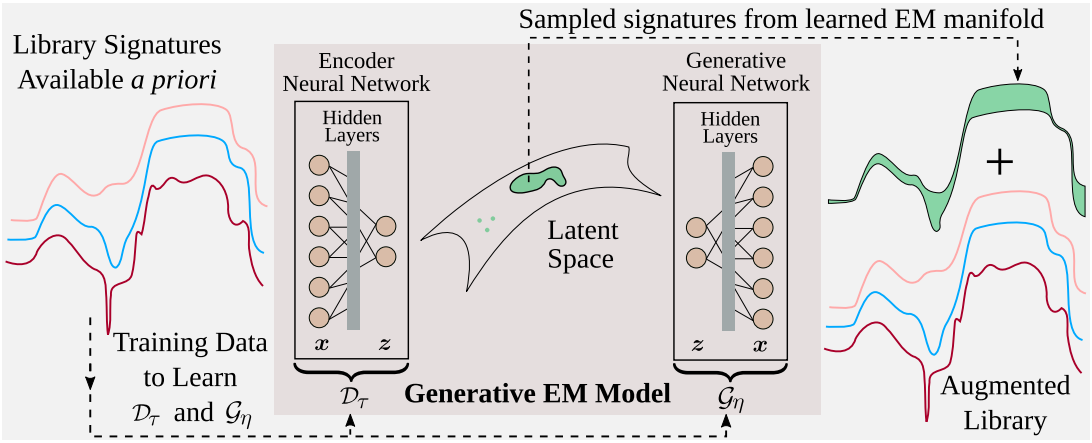


Figure 39 – Outline of the proposed approach: deep generative models are used to approximate the distribution of spectra belonging to a library. Then, new spectral samples (right) can be obtained by propagating samples drawn from the EM submanifold through $\mathcal{G}_\eta$ and used to augment the spectral library.

These approaches allows one to learn the statistical distribution of the endmembers from very few training samples, making it effective in practical scenarios.

8.3 LIBRARY AUGMENTATION WITH DEEP GENERATIVE MODELS

Physics-based models describing the variations of the spectral signatures in a scene reveal an important characteristic of spectral variability: that EM spectra usually lies on a low-dimensional submanifold of the high-dimensional spectral space $\mathbb{R}^L$. This assumption is in agreement with most physical models, such as the PROSPECT or Hapke's [98, 37], which represent the spectral signature of the materials as a function of only a small number of photometric or chemical properties.

Instead of employing physics-based models, we propose in this paper to augment the spectral libraries by using deep generative models. Generative models aim to estimate the probability distribution $p(X)$ of a random variable $X \in \mathbb{R}^L$ based on a set of M_x observations $\boldsymbol{x}_i$. Then, they allow one to generate new samples that look similar to new realizations of X. A more in-depth overview of (deep) generative models (including VAEs and GANs) can be found in Section 7.3. Such models have shown good performance at representing endmember spectra in blind unmixing applications [38]. Here we propose to use the signatures in existing spectral libraries to learn the generative models describing the distributions of EM spectra. Then, to enhance the ability of MESMA to adapt to a wider range of spectral variability, we augment the libraries by sampling from the estimated distributions. An illustrative outline of this strategy is shown in Fig. 39.

Even though the spectral dimensionality L is high compared to the small number of signatures often found in typical spectral libraries (making this problem very hard in general [174, 175]), the low-dimensinality of the manifolds to which the EM spectra is confined, allied with recent advances in generative models, have made this problem

Algorithme 8 : MESMA with spectral library augmentation

Input : $\boldsymbol{Y}$, $\mathcal{M}_i$, $i = 1, \ldots, P$ and Q_s.

for $i = 1, \ldots, P$ **do**
  Set $\widetilde{\mathcal{M}}_i = \mathcal{M}_i$ and train a DGM $\mathcal{G}_{\eta_i}$ using the samples in $\mathcal{M}_i$;
  for $j = 1, \ldots, Q_s$ **do**
    Sample $\boldsymbol{z} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$ and compute $\widehat{\boldsymbol{m}} = \mathcal{G}_{\eta_i}(\boldsymbol{z})$;
    $\widetilde{\mathcal{M}}_i \leftarrow \widetilde{\mathcal{M}}_i \bigcup \{\widehat{\boldsymbol{m}}\}$;
  end
end
Set $\widetilde{\mathcal{M}} = \left\{[\boldsymbol{m}_1, \ldots, \boldsymbol{m}_P] \, : \, \boldsymbol{m}_i \in \widetilde{\mathcal{M}}_i \, , \, i = 1, \ldots, P\right\}$;
Run MESMA with the augmented library $\widetilde{\mathcal{M}}$ to compute $\widehat{\boldsymbol{A}}$;
return $\widehat{\boldsymbol{A}}$, $\widetilde{\mathcal{M}}$;

tractable. This framework has shown success in capturing the distribution of complex data such as natural images from very few training samples [194], which illustrates its appropriateness for our application.

8.3.1 Library augmentation

Consider a small spectral library $\mathcal{M}$ known a priori containing a set of spectral signatures $\mathcal{M}_i$ for each material $i = 1, \ldots, P$. Each signature $\boldsymbol{m}_{i,j} \in \mathcal{M}_i$, $j = 1, \ldots, C_i$, can be viewed as a sample drawn from the statistical distribution of the i-th EM spectra. Thus, these libraries can be employed as training data to learn a set of generative models $\mathcal{G}_{\eta_i}$ that represents the probability distribution function $p_i(M)$ of each EM $i = 1, \ldots, P$ using a VAE [172].

Given the learned generative models $\mathcal{G}_{\eta_i}$, we can then generate new spectral signatures from each EM class by sampling from the distribution of $\mathcal{G}_{\eta_i}(Z)$, where $Z \sim \mathcal{N}(0, \boldsymbol{I}_H)$. These new signatures can then be used to augment into the original library $\mathcal{M}$, yielding a new spectral library $\widetilde{\mathcal{M}}$ which is more comprehensive and better accounts for different spectral variations of each material. Finally, the MESMA algorithm can be applied to unmix each image pixel $\boldsymbol{y}_n$ using the augmented library $\widetilde{\mathcal{M}}$. This procedure is described in detail in Algorithm 8, where the spectral library is augmented by adding Q_s samples to each EM set. Note that although this increases the complexity of SU with MESMA, approximate strategies can be used to obtain an efficient solution when the augmented library has many signatures [195].

8.3.2 Network architecture

To learn the generative models $\mathcal{G}_{\eta_p}$, we used a VAE [172] due to its stable training [175] and because it behaved well with small spectral libraries. The network architectures for $\mathcal{G}_{\eta_p}$ and $\mathcal{D}_{\tau_p}$ and the dimension of the latent spaces were selected as the same

Table 15 – Encoder and Decoder network architectures.

	Layer	Activation	Number of units
$\mathcal{D}_\tau$	Input	—	L
	Hidden # 1	ReLU	$\lceil 1.2 \times L \rceil + 5$
	Hidden # 2	ReLU	$\max\left\{\lceil L/4 \rceil,\, H+2\right\} + 3$
	Hidden # 3	ReLU	$\max\left\{\lceil L/10 \rceil,\, H+1\right\}$
$\mathcal{G}_\eta$	Hidden # 1	ReLU	$\max\left\{\lceil L/10 \rceil,\, H+1\right\}$
	Hidden # 2	ReLU	$\max\left\{\lceil L/4 \rceil,\, H+2\right\} + 3$
	Hidden # 3	ReLU	$\lceil 1.2 \times L \rceil + 5$
	Output	Sigmoid	L

Table 16 – Simulations with synthetic and real data (values $\times 10^3$).

	Synthetic HI		Alunite Hill	Gulfport
	$\text{RMSE}_{\boldsymbol{A}}$	$\text{RMSE}_{\boldsymbol{Y}}$	$\text{RMSE}_{\boldsymbol{Y}}$	$\text{RMSE}_{\boldsymbol{Y}}$
FCLS	50.0 ± 32.2	0.73 ± 0.87	0.47 ± 0.60	1.00 ± 2.06
GLMM	45.3 ± 31.2	0.30 ± 0.22	0.001 ± 0.002	0.002 ± 0.003
MESMA	18.2 ± 13.7	0.41 ± 0.45	19.2 ± 14.0	1.31 ± 2.02
Proposed	15.3 ± 11.0	0.26 ± 0.25	18.4 ± 12.8	1.16 ± 1.86

$\text{RMSE}_{\boldsymbol{A}}$ of Algorithm 8 as a function of Q_s							
Q_s	0	1	2	3	4	5	6
$\text{RMSE}_{\boldsymbol{A}}$	18.18	16.23	15.65	15.34	15.21	15.09	15.01

ones used in Chapter 7 and in [38], since they resulted in a good experimental performance and showed sufficient capacity to capture the spectral variability of a given library. For convenience, the network architectures are also shown in Table 15 and the latent spaces dimension was set to $H = 2$. Finally, the network training was performed with the Adam optimizer [191] in TensorFlow for 50 epochs.

8.4 EXPERIMENTAL RESULTS

In this section, simulation results using both synthetic and real data illustrate the performance of the proposed method. We compare the performance of MESMA using the augmented library with that of the traditional MESMA algorithm. We also present results obtained with the FCLS and the the GLMM [34], which estimate the endmembers from the observed HI (without using a spectral library). The VCA algorithm [131] was used to extract EMs used by the FCLS and GLMM methods. The performances were evaluated using the RMSE between the estimated abundance maps ($\text{RMSE}_{\boldsymbol{A}}$) and between the reconstructed images ($\text{RMSE}_{\boldsymbol{Y}}$). The RMSE between two matrices is defined as $\text{RMSE}_{\boldsymbol{X}} = \sqrt{\|\boldsymbol{X} - \widehat{\boldsymbol{X}}\|_F^2 \,/\, N_{\boldsymbol{X}}}$, where $N_{\boldsymbol{X}}$ denotes the number of elements in $\boldsymbol{X}$.

8.4.1 Synthetic data with library mismatch

In this example, we evaluate the performance of the proposed approach quantitatively using a synthetic data set with three endmembers and $L = 198$ spectral bands. The goal is to simulate a typical library mismatch scenario often found when considering library-based unmixing [21]. To generate and process this dataset, we first obtained two disjoint sets of endmember spectra $\mathcal{M}_i^1$ and $\mathcal{M}_i^2$, with $\mathcal{M}_i^1 \cap \mathcal{M}_i^2 = \varnothing$, $i \in \{1, 2, 3\}$ by manually extracting pure pixels of soil, vegetation and water from a real hyperspectral scene (the Jasper Ridge HI [35]). The sets $\mathcal{M}_i^1$ contained 20 signatures each and were used to compose the synthetic pixel spectra $\boldsymbol{y}_n$, while each of the sets $\mathcal{M}_i^2$ contained 14 signatures that were employed to construct the spectral libraries used by MESMA to perform SU. We simulated a library mismatch by applying a random affine transformation (a gain and an additive scaling in the intervals $[0.75, 1.25]$ and $[-0.15, 0.15]$, respectively) to each element of $\mathcal{M}_i^1$, $i \in \{1, 2, 3\}$. To generate each pixel, we used the LMM considering abundance fractions $\boldsymbol{a}_n$ sampled from a Dirichlet distribution with concentration parameters selected such as to have a heavily mixed data in order to evaluate the methods in a challenging scenario, and pixel-dependent endmember matrices obtained by randomly (uniformly) selecting one spectral signature from each of the sets $\mathcal{M}_i^1$, $i \in \{1, 2, 3\}$. White Gaussian noise with an SNR of 30 dB was added to the data.

The final library $\mathcal{M}$ available for the MESMA-based methods was created by sampling five signatures at random of each material from $\mathcal{M}_i^2$, and no other preprocessing or adequacy strategy was used to mitigate mismatch between the available library and the true endmembers used to construct the scene. Only the spectra in $\mathcal{M}$ was used to learn the DGMs, and $Q_s = 3$ additional signatures were sampled for each material. Finally, in order to provide a proper statistical evaluation, this whole procedure was repeated for 10^4 Monte Carlo realizations. The mean values and standard deviations are shown in Table 16. It can be seen that despite only a small number of signatures being available to train the DGMs, the proposed strategy provided a substantial (16%) improvement in the abundance estimation RMSE when compared to the MESMA algorithm. This shows

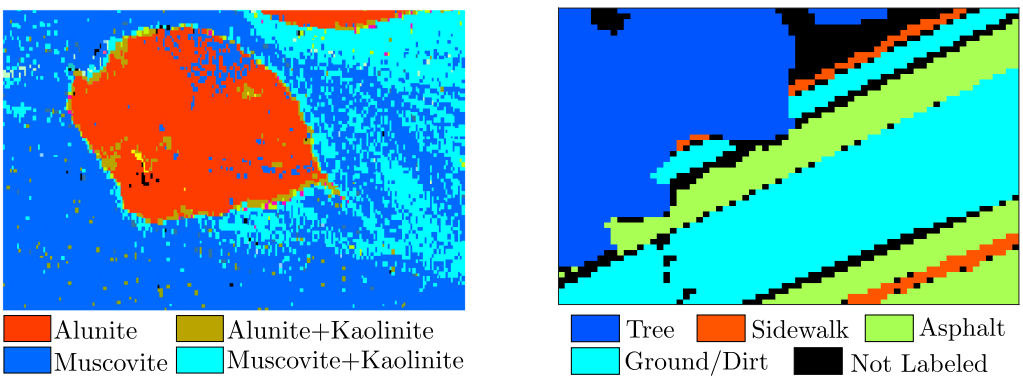


Figure 40 – "Ground truth" for the Alunite Hill (left) and Gulfport (right) HIs.

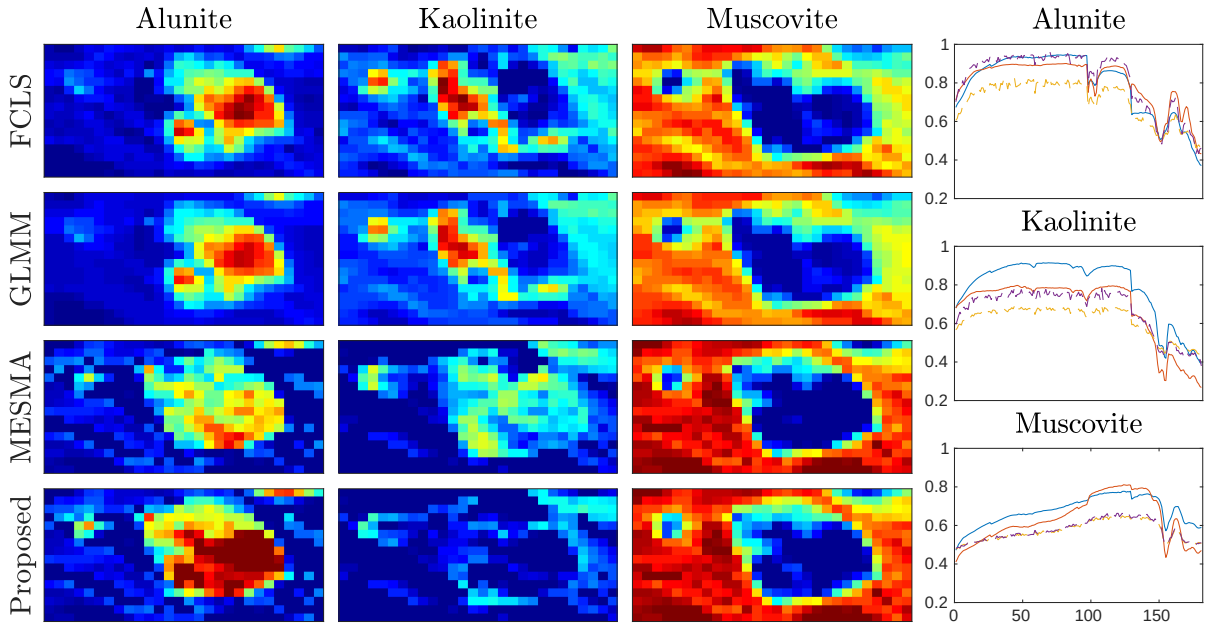


Figure 41 – Left: abundance maps for the Alunite Hill subscene. Right: original endmembers (solid line) and synthetically generated signatures (dashed line).

Figure 42 – Abundance maps of MESMA for the Alunite Hill HI with a spectral library augmented using the Hapke model with known acquisition conditions.

that with a careful selection of the neural network architecture, the proposed method can work even under such challenging conditions. When compared with the other methods the proposed solution improvement is even more significant obtaining gains of 70% (FCLS) and 67% (GLMM). These experiments show that the proposed data augmentation strategy can lead to significant performance gains when compared to the plain MESMA algorithm.

To investigate the influence of the parameter Q_s on the performance of the proposed method, we repeated this experiment for different values of $Q_s \in \{0, \ldots, 8\}$ and evaluated the behavior of $\text{RMSE}_{\boldsymbol{A}}$. The results, also seen in Table 16, show that $\text{RMSE}_{\boldsymbol{A}}$ decreases with Q_s. However, the performance improvements get small after about $Q_s > 3$, which indicates that a value of $Q_s \leqslant 3$ can yield a good compromise between abundance estimation performance and computational complexity.

8.4.2 Real data

For the simulations with real data, we considered the Alunite Hill subscene of the Cuprite HI with 16×28 pixels, and a subscene of the Gulfport HI with 54×70 pixels [195]. Water absorption or low SNR bands were removed and both the images and the spectral libraries were rescaled to have the same number of bands, resulting in $L = 181$ for the

Alunite Hill and $L = 192$ for Gulfport. These images were selected since the unmixing results can be evaluated using high-resolution classification maps available a priori, shown in Fig. 40. The libraries $\mathcal{M}$ were built by selecting two signatures of each endmember from the USGS library and from field surveys (for the Alunite Hill and Gulfport HIs, resp.) such that the MESMA results closely approached (visually) the ground truth. $Q_s = 2$ additional signatures per EM were generated.

The abundance maps reconstructed by all algorithms are provided in Figs. 41 and 43. It can be seen that the abundance maps of the MESMA-based methods are significantly closer to the ground truth when compared to the GLMM and FCLS results. Furthermore, the proposed library augmentation strategy led to a much better representation of the alunite and kaolinite endmembers when compared to the competing approaches in the Alunite Hill HI. Similar results were obtained for the Gulfport HI, where the abundances obtained by the proposed method for the sidewalk and asphalt EMs approach the ground truth more closely when compared to those estimated by FCLS, GLMM and by MESMA with the original library. The spectral signatures generated using the DGMs, also seen in Figs. 41 and 43, show that the proposed strategy is able to generate signatures that accommodate variability seen in typical scenes from its representation in the original library. Specifically, a generally agreeable shape but different scaling variations that act nonuniformly over the spectral space can be seen in all cases except for the ground/dirt EM in the Gulfport HI, whose original field surveyed spectra (contained in $\mathcal{M}$) did not contain a meaningful amount of spectral variability.

The quantitative $\text{RMSE}_{\boldsymbol{Y}}$ results in Table 16 show that the FCLS achieves smaller reconstruction errors in the real datasets when compared to the synthetic one, which contains more heavily mixed pixels and thus results in a worse data fitting for the FCLS (which is based on the VCA). However, we note that $\text{RMSE}_{\boldsymbol{Y}}$ is not a good measure of unmixing performance, as an infinite number of combinations (endmembers, abundances) often leads to the same reconstructed HI.

To compare the proposed method with physics-based library augmentation, we considered a Lambertian scattering approximation of the Hapke model to augment the library used with the Alunite Hill HI. Given prior knowledge about the laboratory acquisition conditions of the spectra in the USGS library, we can generate different variations of these mineral spectra by considering different viewing geometries as detailed in [26]. The abundances estimated by MESMA using the augmented library are shown in Fig. 42. Although a clear improvement can be seen in the alunite and muscovite EMs when compared to the original library, the kaolinite abundances were completely absorbed into the muscovite abundance map. Moreover, the alunite region is smaller than what is indicated in the ground truth, which is more closely matched by the results obtained using the proposed method. This shows that the proposed strategy can be competitive with physics-based models in practice.

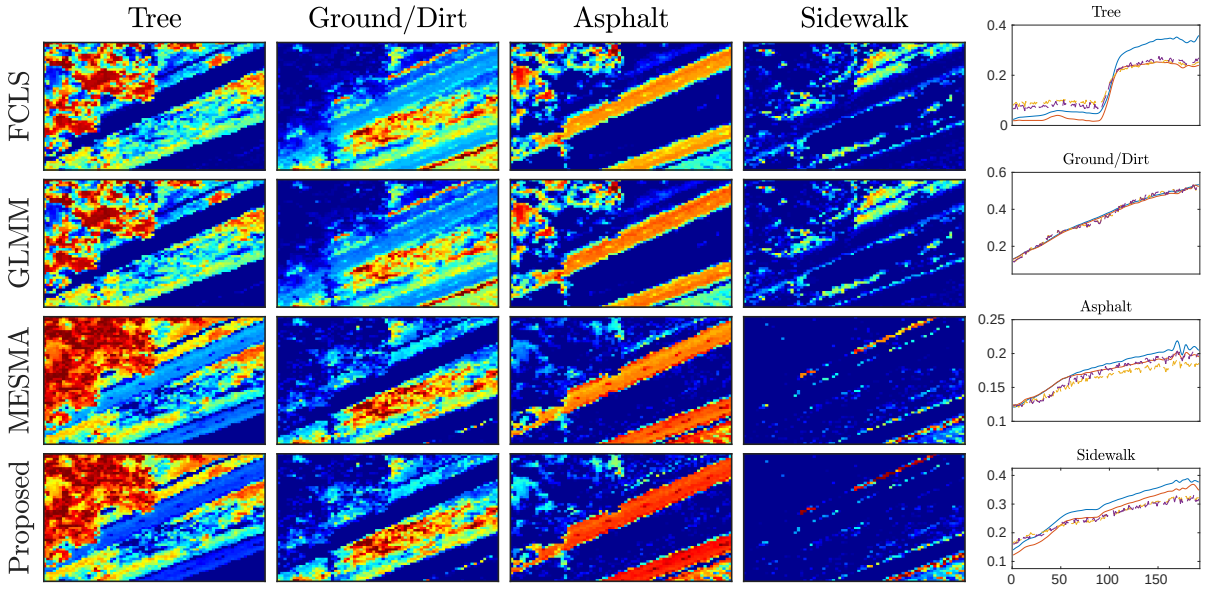


Figure 43 – Left: abundance maps for the Gulfport subscene. Right: original endmembers (solid line) and synthetically generated signatures (dashed line).

8.5 CONCLUSIONS

In this chapter, a novel spectral library augmentation strategy was proposed for MESMA-like algorithms. Using the spectral signatures present in existing libraries as training samples, we applied deep generative models to learn the statistical distribution of endmember spectra. This allowed us to sample new spectral signatures from the estimated endmember distribution, which were then included in the augmented library, improving its ability to properly represent the materials present in practical scenes. Simulation results with both synthetic and real data showed that the proposed methodology can significantly improve the performance of the MESMA algorithm.

INTERMEZZO: SPECTRAL VARIABILITY, REDUX: IMAGE FUSION, MULTITEMPORAL SU AND CHANGE DETECTION

The challenges associated with spectral variability are also present in other applications such as multimodal image fusion and target detection. Moreover, the difficulties which underlie SU in practice also go beyond spectral variability, and include nonlinearity in the mixing process as well as the presence of outliers. In the next chapter, we propose to address a different challenge of the SU problem, and a different problem to which spectral variability applies, by extending the methods proposed in Chapters 3–8. For the sake of brevity and to avoid deviating from the main scope of this thesis, we only provide brief overviews of these works; complete versions can be found in the related publications.

We begin by extending the multiscale regularization of Chapter 3 to address nonlinear SU with kernels in Section 9.1. The HI is decomposed in two spatial domains (containing the coarse and the original image content). SU is then formulated as non-convex quadratically constrained optimization problems to allow for an automated parameter design strategy. An efficient solution to these problems is then devised by exploring their strong duality property. This section is related to the publication [40].

Section 9.2 moves on to introduce the time variable into SU. Unmixing in the multitmeporal setting introduces both challenges as well as opportunities: while spectral variability between different HIs can be very pronounced and abrupt changes can be present, the availability of multiple HIs also allows unmixing results to be improved. In this context, two new multitemporal SU approaches are proposed. First, we extend parametric EM models to the multitemporal setting and use a Bayesian filter framework to perform SU under slowly varying abundances. Afterwards, we consider a spectral library-based multitemporal SU approach which is well adapted to abrupt abundance changes. This section is related to references [41, 42].

Finally, in Section 9.3 we address spectral variability in the hyperspectral and multispectral image fusion problem, where the images in both modalities can be subjected to spectral variations and changes. To solve this problem, we first propose to address spatially invariant variability between the images in a matrix factorization framework. Spectrally and spatially localized changes are later accounted for by using a tensor-based representation, in which image fusion is formulated as a coupled tensor factorization problem. This section is related to publications [43, 44].

9 AND NOW FOR SOMETHING DIFFERENT...

"And now for something completely different."
Monty Python

Besides the content presented in Chapters 3–8, several other works have been conducted in the duration of the thesis, which we do not include in full form for the sake of brevity. In the following, we briefly detail the main contributions of these works.

9.1 MULTISCALE NONLINEAR UNMIXING WITH KERNELS

The LMM is effective in accurately modeling mixtures occurring in scenes where the materials of interest cover a large portion of the pixel area [23]. However, it disregards non-linearity in the mixing process, which often results in estimation errors being propagated to the abundance maps [23, 111]. Nonlinear interactions between materials occur in many scenes where there is complex radiation scattering among several endmembers, such as in some vegetation areas [196]. In such situations, nonlinear mixing models must be considered [23, 26]. Several nonlinear SU strategies have been proposed in the literature, which can be roughly divided between model-based and model-free methods. Most model-based nonlinear SU algorithms assume that the mixing process that occurs in the scene is known a priori [197, 198, 199, 26, 23]. However, real mixing mechanisms can be very complex and prior knowledge about them is seldom available in practice. This led to the consideration of more flexible model-free nonlinear SU, which employ more flexible nonlinear mixing models that are able to represent generic functions. Prominent model-free strategies include the estimation of abundances as posterior class probabilities of a nonlinear classifier [200], the use of graph-based approximate geodesic distances [201], and kernel-based algorithms [202, 203, 204, 205]. Kernel-based methods provide non-parametric representations of functional spaces that are able to model arbitrary nonlinear mixtures [202, 26, 23, 203, 204, 156]. This flexibility, allied to a good experimental performance has led to the wide application of kernel methods.

In [203] the authors considered a semi-parametric kernel-based model consisting of a linear trend parameterized by the abundance vector plus an additive nonlinear fluctuation. The model, which allows the quantification of the abundance vectors during the unmixing process, is given by

$$\boldsymbol{y}_n = \boldsymbol{M}_0\boldsymbol{a}_n + \boldsymbol{\zeta}_n(\boldsymbol{M}_0) + \boldsymbol{e}_n\,, \tag{113}$$

where $\boldsymbol{\zeta}_n(\boldsymbol{M}_0) = [\zeta_n([\boldsymbol{M}_0]_{1,:}), \ldots, \zeta_n([\boldsymbol{M}_0]_{L,:})]^\top$, $\zeta_n : \mathbb{R}^P \to \mathbb{R}$ is an arbitrary smooth function belonging to a reproducing kernel Hilbert space, and $\boldsymbol{y}_n$, $\boldsymbol{a}_n$ and $\boldsymbol{M}_0$ are as defined in (1).

In [203] the authors proposed to solve the unmixing problem accounting for the model in (113) by considering a multi-kernel generalization of standard least-squares

support vector regression (LS-SVR) methods [206]. Such methods, however, significantly increase the amount of parameters to be estimated from the data, rendering the SU problem more ill-posed. Thus, the total variation regularization has been considered in [204] to devise a regularized LS-SVR-based SU algorithm. However, using the TV regularization introduces the same challenges in terms of computational complexity and limitations concerning the quality of the SU results as discussed in Chapter 3. Moreover, it also introduces an additional regularization parameter which must be adjusted by the user to provide good results.

Motivated by the results in Chapter 3, we propose to introduce spatial information into SU by representing this problem separately in two spatial scales, which significantly reduces the computational complexity of SU. Moreover, this also makes the inter-scale interaction between the abundances explicit in the resulting optimization problem, which is essential for devising a theoretically principled automatic parameter determination strategy.

9.1.1 Imaging model

We consider the multiscale transformation $\boldsymbol{W} \in \mathbb{R}^{N\times S}$, $S < N$ as defined in Chapter 3, that maps both the HI and the abundance maps to the approximation domain based on the superpixel decomposition. We recall that the transformed matrices in the coarse spatial domain are given by

$$\boldsymbol{Y}_{\mathcal{C}} = \boldsymbol{Y}\boldsymbol{W}\,, \quad \boldsymbol{A}_{\mathcal{C}} = \boldsymbol{A}\boldsymbol{W}\,. \tag{114}$$

Considering the nonlinear observation model (113), the transformed image in (114) leads to an equivalent mixing model in the coarse spatial domain, which is given by

$$\boldsymbol{y}_{\mathcal{C}_i} = \boldsymbol{M}_0\boldsymbol{a}_{\mathcal{C}_i} + \boldsymbol{\zeta}_{\mathcal{C}_i}(\boldsymbol{M}_0) + \boldsymbol{e}_{\mathcal{C}_i}\,, \tag{115}$$

with

$$\boldsymbol{\zeta}_{\mathcal{C}_i}(\boldsymbol{M}_0) = \sum_{n=1}^{N} W_{n,i}\,\boldsymbol{\zeta}_n(\boldsymbol{M}_0) \tag{116}$$

denoting the nonlinear contributions at the coarse spatial scale, for $i = 1, \ldots, S$.

9.1.2 Optimization problem and parameter selection strategy

Given models (113) and (115), we can formulate optimization problems whose solutions will be the estimated abundance maps in the coarse and original spatial scales. However, an important challenge associated with the framework proposed in Chapter 3 is the need for adjusting multiple regularization parameters to obtain a good performance.

One principled way to select regularization parameters in inverse problems consists on the Chi-squared method [207, 208, 209] and the closely related discrepancy principle [210]. The underlying idea behind these methods is to evaluate the statistical properties

of the data reconstruction term in a cost function when the estimated solution is equal to the desired one. The regularization constant is then selected so that the estimated solution yields a residual with the same statistical properties of the observation noise, which are assumed to be known [207, 208, 209]. Thus, we formulate the optimization problems using equality constraints, which allows us to characterize their corresponding statistical properties to provide a principled choice of the regularization constants.

Following the observation model (115) and using the Chi-squared principle, the SU problem at the coarse spatial scale can be formulated using the LS-SVR framework [203] as:

$$\begin{aligned}\{\widehat{\boldsymbol{a}}_{\mathcal{C}_i}, \widehat{\zeta}_{\mathcal{C}_i}\} = \underset{\{\boldsymbol{a}_{\mathcal{C}_i}, \zeta_{\mathcal{C}_i}, \boldsymbol{\xi}_{\mathcal{C}_i}\}}{\arg\min} \; & \frac{1}{2}\sum_{i=1}^{S}\left(\|\zeta_{\mathcal{C}_i}\|_{\mathcal{H}}^2 + \|\boldsymbol{a}_{\mathcal{C}_i}\|_2^2\right) \\ \text{subject to } \; & \boldsymbol{a}_{\mathcal{C}_i} \geqslant \mathbf{0}\,,\; \mathbf{1}^\top \boldsymbol{a}_{\mathcal{C}_i} = 1\,,\;\; i = 1, \ldots, S\,, \\ & \boldsymbol{\xi}_{\mathcal{C}_i} = \boldsymbol{y}_{\mathcal{C}_i} - \boldsymbol{M}_0 \boldsymbol{a}_{\mathcal{C}_i} - \boldsymbol{\zeta}_{\mathcal{C}_i}(\boldsymbol{M}_0), \quad i = 1, \ldots, S\,, \\ & \frac{1}{S}\sum_{i=1}^{S}\|\boldsymbol{\xi}_{\mathcal{C}_i}\|_2^2 = V_0\,,\end{aligned} \tag{117}$$

where V_0 is a constant related to the coarse scale reconstruction error, which can be approximated by (see [40] for more details):

$$V_0 \approx \frac{1}{S}\operatorname{tr}\{\boldsymbol{\Sigma}_{\boldsymbol{e}}\} + \nu_{\text{err}}^2\,, \tag{118}$$

with $\boldsymbol{\Sigma}_{\boldsymbol{e}} = \mathrm{E}\{\boldsymbol{e}_n \boldsymbol{e}_n^\top\} \in \mathbb{R}^{L\times L}$ being the noise covariance matrix (which is assumed to be the same for all pixels) and ν_{err}^2 denoting modeling errors.

The abundance maps estimated at the coarse spatial scale, denoted by $\widehat{\boldsymbol{A}}_{\mathcal{C}} = [\widehat{\boldsymbol{a}}_{\mathcal{C}_1}, \ldots, \widehat{\boldsymbol{a}}_{\mathcal{C}_S}]$, can be used to regularize the original unmixing problem. To this end, we convert the abundance map from the coarse approximation domain $\mathcal{C}$ back to the original image domain as

$$\widehat{\boldsymbol{A}}_{\mathcal{C}*} = \widehat{\boldsymbol{A}}_{\mathcal{C}} \boldsymbol{W}^*\,. \tag{119}$$

Applying the Chi-squared principle to determine the relationship between the abundances at two spatial scales is, however, not direct. We need to describe the desired properties of the inter-scale regularization term $\left\|\boldsymbol{A} - \widehat{\boldsymbol{A}}_{\mathcal{C}*}\right\|_F$. To do so, we explore the statistical properties of the following term:

$$\mathrm{E}\left\{\frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{M}_0^\dagger(\boldsymbol{y}_n - \boldsymbol{y}_{\mathcal{C}_n^*})\|^2\right\}, \tag{120}$$

where $\dagger$ denotes the pseudoinverse operator. Using some simplifying assumptions and approximations (see [40] for more details), we obtain the following constraint:

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{a}_n - \widehat{\boldsymbol{a}}_{\mathcal{C}_n^*}\|^2 \approx \frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{M}_0^\dagger(\boldsymbol{y}_n - \boldsymbol{y}_{\mathcal{C}_n^*})\|^2 - \frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{M}_0^\dagger\big(\boldsymbol{\zeta}_n(\boldsymbol{M}_0) - \widehat{\boldsymbol{\zeta}}_{\mathcal{C}_n^*}(\boldsymbol{M})\big)\|^2 \\ - \|\boldsymbol{M}_0^\dagger \boldsymbol{\Sigma}_{\boldsymbol{e}}^{1/2}\|_F^2 \frac{S-1}{S}\,.\end{aligned} \tag{121}$$

where $\widehat{\boldsymbol{\zeta}}_{\mathcal{C}_n^*}$ is the function obtained as the solution to (117) for the superpixel which contains pixel n.

Using relation (121) to regularize the abundances, the SU problem at the original spatial scale can be written as:

$$\begin{aligned}
&\underset{\{\boldsymbol{a}_n,\zeta_n,\boldsymbol{\xi}_n,\boldsymbol{\xi}_{\zeta,n}\}}{\arg\min} \ \frac{1}{2}\sum_{n=1}^{N}\|\zeta_n\|_{\mathcal{H}}^2 \qquad (122)\\
&\text{subject to } \ \boldsymbol{a}_n \succcurlyeq \mathbf{0}\,,\ \mathbf{1}^\top\boldsymbol{a}_n = 1\,,\ \ n=1,\ldots,N\,,\\
&\qquad \boldsymbol{\xi}_n = \boldsymbol{y}_n - \boldsymbol{M}_0\boldsymbol{a}_n - \boldsymbol{\zeta}_n(\boldsymbol{M}_0)\,,\ \ n=1,\ldots,N\,,\\
&\qquad \boldsymbol{\xi}_{\zeta,n} = \boldsymbol{M}_0^\dagger\big(\boldsymbol{\zeta}_n(\boldsymbol{M}_0) - \widehat{\boldsymbol{\zeta}}_{\mathcal{C}_n^*}(\boldsymbol{M}_0)\big)\,,\ \ n=1,\ldots,N\,,\\
&\qquad \frac{1}{2}\sum_{n=1}^{N}\|\boldsymbol{\xi}_n\|_2^2 = \frac{N}{2}V_1\,,\\
&\qquad \frac{1}{2}\sum_{n=1}^{N}\big(\|\boldsymbol{a}_n - \widehat{\boldsymbol{a}}_{\mathcal{D}_n}\|_2^2 + \|\boldsymbol{\xi}_{\zeta,n}\|^2\big)\\
&\qquad\qquad = \frac{N}{2}\Big(\frac{1}{N}\sum_{n=1}^{N}\|\boldsymbol{M}_0^\dagger(\boldsymbol{y}_n - \boldsymbol{y}_{\mathcal{D}_n})\|^2 - \|\boldsymbol{M}_0^\dagger\boldsymbol{\Sigma}_e^{1/2}\|_F^2\frac{S-1}{S}\Big)\,,
\end{aligned}$$

where constant $V_1 \approx \mathrm{tr}\{\boldsymbol{\Sigma}_e\} + \nu_{\mathrm{err}}^2$ is related to the fine scale reconstruction error. Problem (122) can now be used to perform unmixing in the original spatial scale.

The quadratic equality constraints in (117) and (122) make these optimization problems non-convex. To provide an efficient solution, we resorted to a Lagrangian relaxation and solved the dual optimization problem, which is concave and finite-dimensional [211, 212]. Although the non-convexity of the constraints implies the possibility of a non-zero duality gap, by building upon results from non-convex optimization [213] we showed that strong duality holds under mild conditions for problems (117) and (122). Thus, this approach incurs no loss of performance.

Although the Lagrangian duals of (117) and (122) are concave, they are nonlinear and costly to solve due to the high-dimensional nature of the problem. To obtain an efficient solution, we rewrote the dual problems as low-dimensional root finding problems, which can be solved in very few iterations using a multidimensional bisection algorithm. For more details, please see the original publication [40].

9.1.3 Conclusions

In this section, we presented an overview of a new multiscale spatial regularization approach for kernel-based nonlinear unmixing. Building upon the MUA algorithm of Chapter 3, we employed a multiscale representation to divide the unmixing problem into two simpler problems in different scales. Besides dealing with a non-parametric nonlinear mixing model, we also addressed the parameter adjustment problem to reduce the required amount of user supervision. The SU problem was reformulated at multiple scales by

statistically characterizing not only the algorithm reconstruction error in both scales, but also the inter-scale interaction between the abundances and the nonlinear mixing contributions across the coarse and fine image domains. This formulation led to physically motivated constraints which were leveraged to devise a method in which all the parameters were determined automatically from the observed data. Thus, the proposed strategy benefits from an improved quality without the need for *ad hoc* parameter adjustment such as in TV-based works. The unmixing problem was then cast using quadratically constrained optimization problems, for which efficient solutions were obtained by exploring their strong duality and a reformulation of their dual representations as root-finding problems. Simulation results with both synthetic and real data indicate that the proposed strategy leads to a consistent performance improvement when compared to the classical Total Variation regularization, even though no parameter adjustment is necessary.

9.2 MULTITEMPORAL UNMIXING AND CHANGE DETECTION

Multitemporal SU (MTSU) has recently become a subject of great interest due to the possibility of leveraging time information in HI sequences, allowing for monitoring the dynamical evolution of the materials and their distributions [214, 215, 216]. However, the influence of spectral variability in multitemporal scenarios can be significantly stronger than in the case of a single HI. This introduces a challenge to multitemporal SU since EM variability must be carefully modeled to achieve a good performance. Previous works have considered different strategies to incorporate dynamical information about the EMs. One popular strategy is based on parametric models originally devised to account for variations within a single HI, such as the ones discussed in Chapters 4, 5 and 7. These include constraining the EMs at adjacent time instants to be scaled versions of each other [217], or to be represented as a mean EM matrix with small, additive perturbations [218, 219, 220]. However, these works disregard important information as they do not account for the low-dimensional structure that often underlies the changes observed in EM spectra when representing their evolution.

Another line of work makes use of spectral libraries to constrain EM spectra by applying the MESMA algorithm (discussed in Section 8.2) to MTSU. This approach was successfully applied to the monitoring of rainforests [214, 216] and shrublands [215]. Although MESMA can give good results if the spectral libraries are representative of the observed scenes, it has a very high computational complexity due to its combinatorial nature. Moreover, existing algorithms do not explicitly explore temporal correlation between the abundance maps at adjacent time instants.

To address the issues with existing multitemporal SU based on parametric EM models and spectral libraries, two new MTSU methods are proposed. First, we propose an algorithm for multitemporal SU which is based on a dynamical parametric model for the EM time variability. Specifically, we couple the representation power of parametric

EM models originally devised to operate within a single HI, such as the GLMM, with a Bayesian filtering methodology to estimate the EMs in HI sequences. Instead of operating directly on the EM spectral space, we make use of a parametric EM model to represent EM dynamics indirectly through vectors of parameters that capture the time variations of each material. Bayesian filtering and smoothing are combined with the Expectation Maximization algorithm to estimate the required parameters given a window of observations in time. The initialization of the resulting Kalman filter is also estimated in the process, which improves convergence for short image sequences. Under some approximations and for small temporal variation of the abundances, the proposed algorithm is able to blindly estimate the EMs, the average abundances, and the remaining model parameters from the observed HI data.

Afterwards, we consider an approach based on MESMA in order to benefit from the information available in the form of spectral libraries and to address abrupt abundance changes. Contrary to previous works, we formulate the multitemporal unmixing problem by explicitly characterizing the temporal correlation and abrupt changes in the abundances. Then, we propose to disjoin the EM selection problem from the abundance estimation problem, what significantly reduces the computational complexity of the algorithm without sacrificing the accuracy of the results. We also introduce a simple strategy to detect pixels characterized by abrupt abundance changes. Theoretical guarantees are provided on how our strategy compares to MESMA in terms of solution quality, that is, when it recovers the correct endmembers, and also in terms of robustness, that is, when it correctly detects abrupt abundance changes. This provides important insight into the practical performance of the method. The proposed algorithm provides robust, high-quality abundances reconstruction over time at a low computational cost.

9.2.1 Multitemporal Spectral Unmixing Using a Dynamic EM Model

The LMM (1) can be easily extended to the multitemporal setting to represent an HI with L bands and N pixels at time t as:

$$\boldsymbol{Y}^{(t)} = \boldsymbol{M}^{(t)}\boldsymbol{A}^{(t)} + \boldsymbol{E}^{(t)}, \text{ subject to } \mathbf{1}^\top \boldsymbol{A}^{(t)} = \mathbf{1}^\top, \ \boldsymbol{A}^{(t)} \geqslant \mathbf{0}\,, \tag{123}$$

where $\boldsymbol{Y}^{(t)} \in \mathbb{R}^{L\times N}$ is the observed HI, the columns of $\boldsymbol{M}^{(t)} \in \mathbb{R}^{L\times P}$ are the P endmember spectral signatures, $\boldsymbol{A}^{(t)} \in \mathbb{R}^{P\times N}$ contains the abundances for each pixel, and $\boldsymbol{E}^{(t)}$ represents additive noise, all indexed at time $t \in \{1, \ldots, T\}$.

An important challenge related to the use of model (123) regards the consideration of spectral variability, which causes the signatures of the endmembers to change both in space (within the same HI) and in time. Since we have addressed spatial domain spectral variability in Chapters 3–8, for simplicity in this section we assume only variations of the EMs in time. EM variation within the same HI can be later incorporated to the proposed model.

Different SU algorithms accounting for EM time variability have been recently proposed, most of them inspired by models designed to account for spectral variability within a single image. For instance, in [217] the authors constrain the EM matrices at each time instant to be scaled versions of a reference EM matrix. In [218], the authors model the EMs at each time instant by a mean EM matrix plus small perturbations, which are assumed to be temporally smooth. All variables are then estimated using a stochastic approach. This latter model was later extended for distributed unmixing with additional sparsity constraints in [219], and to include sparse additive residual terms to represent abrupt spectral variations in the HI using a hierarchical Bayesian framework in [220]. However, these works operate directly in the input spectral space, ignoring the fact that spectral variability can often be represented more accurately using physically meaningful parametrizations of EM spectra.

As discussed in Chapters 4, 5, and 7, parametric models have become widely used to represent spatial domain spectral variability as:

$$\boldsymbol{M} = f(\boldsymbol{M}_0, \boldsymbol{\theta}) \,, \tag{124}$$

where f is a parametric function, $\boldsymbol{M}_0 \in \mathbb{R}^{L \times P}$ contains reference spectral signatures of the EMs and $\boldsymbol{\theta}$ is a vector of parameters of the variability model. Such parametric models are specially interesting for building a dynamical model to consider EM time variability.

We propose a multitemporal extension of the parametric EM model (124). We assume a fixed reference EM matrix $\boldsymbol{M}_0$, and model the time variations in $\boldsymbol{M}^{(t)}$ through a time varying $\boldsymbol{\theta}^{(t)}$, $t = 1, \ldots, T$. Assuming that temporally adjacent images are acquired at reasonably short time intervals, we model the difference $\boldsymbol{\theta}^{(t)} - \boldsymbol{\theta}^{(t-1)}$ as a small zero-mean vector. Thus, we assume the following model for $\boldsymbol{\theta}^{(t)}$:

$$\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)} + \boldsymbol{c}^{(t)} \,, \tag{125}$$

where $\boldsymbol{c}^{(t)} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{C})$ contains the innovations which describe its dynamical evolution. Note that the Gaussian assumption is only made in the model parameters $\boldsymbol{\theta}^{(t)}$ and not on the EM signatures themselves, which allows for the use of complex EM distributions through the pushforward measure obtained using the function f, as done in Chapter 7. This generalizes parametric EM model (124) to the multitemporal setting as $\boldsymbol{M}^{(t)} = f(\boldsymbol{M}_0, \boldsymbol{\theta}^{(t)})$, where the parametric function f now relates the EM matrices and the vectors of parameters at each time instant. Considering this model, the multitemporal LMM can be represented as

$$\boldsymbol{Y}^{(t)} = f(\boldsymbol{M}_0, \boldsymbol{\theta}^{(t)}) \boldsymbol{A}^{(t)} + \boldsymbol{E}^{(t)} \,. \tag{126}$$

Next, one must choose a function f for (126) that establishes a good compromise between mathematical tractability and performance. The GLMM from Chapter 5 is able to represent arbitrary spectral variability by considering spectrally varying multiplicative

scaling factors, introducing a connection between the amount of spectral variability and the amplitude of endmember reflectance spectra at each band. By considering the GLMM model and using the vectorization property, (123) can be expressed as

$$\operatorname{vec}\left(\boldsymbol{Y}^{(t)}\right) = \left[\left(\boldsymbol{A}^{(t)}\right)^{\top} \otimes \boldsymbol{I}\right] \operatorname{diag}\left(\operatorname{vec}(\boldsymbol{M}_0)\right) \operatorname{vec}\left(\boldsymbol{\Psi}^{(t)}\right) + \operatorname{vec}\left(\boldsymbol{E}^{(t)}\right). \tag{127}$$

with $\boldsymbol{\Psi}^{(t)} \in \mathbb{R}^{L \times P}$ containing the scaling factors at instant t.

We write the abundance matrix $\boldsymbol{A}^{(t)}$ as $\boldsymbol{A}^{(t)} = \boldsymbol{A}_{\text{avg}} + \boldsymbol{\Delta A}^{(t)}$, where $\boldsymbol{\Delta A}^{(t)}$ represents small random fluctuations over the average abundance matrix $\boldsymbol{A}_{\text{avg}}$. Considering $\boldsymbol{\Delta A}^{(t)}$ to be small for a time window $t \in \{t_0, \ldots, t_1\}$, these variations can be incorporated into the observation noise, leading to the following model:

$$\operatorname{vec}\left(\boldsymbol{Y}^{(t)}\right) = \left[\boldsymbol{A}_{\text{avg}}^{\top} \otimes \boldsymbol{I}\right] \operatorname{diag}\left(\operatorname{vec}(\boldsymbol{M}_0)\right) \operatorname{vec}\left(\boldsymbol{\Psi}^{(t)}\right) + \boldsymbol{r}^{(t)}, \tag{128}$$

where $\boldsymbol{r}^{(t)} = \operatorname{vec}(\boldsymbol{E}^{(t)}) + \left[\left(\boldsymbol{\Delta A}^{(t)}\right)^{\top} \otimes \boldsymbol{I}_L\right] \operatorname{diag}(\operatorname{vec}(\boldsymbol{M}_0)) \operatorname{vec}(\boldsymbol{\Psi}^{(t)})$. Note that the observation noise $\boldsymbol{r}^{(t)}$ in (128) is correlated with the state $\boldsymbol{\Psi}^{(t)}$. In the following, we will use a signal-independent noise approximation, which provides competitive performance at a modest computational cost. Further discussion on the impact of such an approximation can be found in the original publication [41].

By considering $\boldsymbol{\theta}^{(t)} \equiv \operatorname{vec}(\boldsymbol{\Psi}^{(t)})$ to be the state vector, $\operatorname{vec}(\boldsymbol{Y}^{(t)})$ the output vector, and assuming $\boldsymbol{r}^{(t)}$ to be independent of $\boldsymbol{c}^{(t)}$, zero-mean and Gaussian, equations (125) and (128) form a linear state space model conditioned on $\boldsymbol{M}_0$ and $\boldsymbol{A}_{\text{avg}}$. Thus, by additionally considering the abundances fixed over a time window $t \in \{t_0, \ldots, t_0 + T\}$, we proposed to use the Kalman smoother to estimate the state vector $\operatorname{vec}(\boldsymbol{\Psi}^{(t)})$, coupled with the Expectation Maximization algorithm to estimate the model parameters $\boldsymbol{A}_{\text{avg}}$, the covariance matrices of $\boldsymbol{c}^{(t)}$ and $\boldsymbol{r}^{(t)}$, and the model initialization $\operatorname{vec}(\boldsymbol{\Psi}^{(0)})$ [221]. A regularized Least Squares problem was finally used to estimate the abundances $\boldsymbol{A}^{(t)}$ for each time instant. For more details, please see the complete publication [41].

9.2.2 Library-based multitemporal SU and change detection

Although requiring little supervision, the approach described in the previous section does not account for spatial EM variability and relies on the abundances varying slowly in time. A simple way which have been used to address spectral variability in multitemporal SU is to apply the MESMA algorithm described in Section 8.2, separately to each HI as:

$$\begin{aligned} \min_{\boldsymbol{M}^{(t)} \in \mathcal{M}} \; \min_{\boldsymbol{a}_n^{(t)}} \; & \left\| \boldsymbol{y}_n^{(t)} - \boldsymbol{M}^{(t)} \boldsymbol{a}_n^{(t)} \right\|^2 \\ & \text{subject to} \;\; \boldsymbol{a}_n^{(t)} \geqslant 0, \, \mathbf{1}^{\top} \boldsymbol{a}_n^{(t)} = 1. \end{aligned} \tag{129}$$

for each $n = 1, \ldots, N$ and $t \in \{1, \ldots, T\}$, where $\mathcal{M}$ is a spectral library as defined in Section 8.2. This strategy is widely used in practical applications [214, 215, 216], and can naturally address both the spatial and temporal variations of EM spectra.

Despite its widespread use and good performance in practical scenarios, the computational cost of the MESMA algorithm is extremely high due to the combinatorial nature of (129). Recent strategies attempt to provide low-complexity alternatives to the MESMA problem with without negatively impacting the unmixing results. For instance, an alternative approach to MESMA in [195] employs an angle minimization strategy. This lead to a significant reduction of the computational complexity for cases where P is small and C_k (the size of the libraries) is possibly very large. Another approach formulates problem (129) as a mixed-integer optimization problem in order to benefit from advanced software packages [222]. However, these works do not consider multitemporal information which can help reduce the computational complexity for larger values of P.

To leverage time information, we propose to model the evolution of the abundance maps by considering its changes to be composed of a small additive signal, and of large sparse changes. That is, the observation model represents the pixels at time t and $t+1$ (for all $t \in \{1, \ldots, T-1\}$) as follows:

$$\boldsymbol{y}_n^{(t)} = \boldsymbol{M}_n^{(t)} \boldsymbol{a}_n^{(t)} + \boldsymbol{e}_n^{(t)}\,, \tag{130a}$$

$$\begin{aligned} \boldsymbol{y}_n^{(t+1)} &= \boldsymbol{M}_n^{(t+1)} \big(\boldsymbol{a}_{t,n} + \boldsymbol{\delta}_{\text{small},n}^{(t)} + \boldsymbol{\delta}_{\text{large},n}^{(t)} \big) + \boldsymbol{e}_n^{(t+1)} \\ &= \boldsymbol{M}_n^{(t+1)} \boldsymbol{a}_n^{(t+1)} + \boldsymbol{e}_n^{(t+1)}\,, \end{aligned} \tag{130b}$$

where $\boldsymbol{M}_n^{(t)}$ is the (true) endmember matrix for pixel n at time instant t. Changes taking place in the abundances between time t and time $t+1$ are modeled as a combination of a small magnitude term $\boldsymbol{\delta}_{\text{small},n}^{(t)}$ and a spatially sparse, high magnitude term $\boldsymbol{\delta}_{\text{large},n}^{(t)}$, which represents abrupt variations taking place in a small number of image pixels.

The structure outlined in the model (130) can be explored in order to devise an efficient MESMA-based MTSU algorithm. We propose to use an online strategy to estimate the abundances and the EM matrices at time instant $t+1$ based on an estimate $\widehat{\boldsymbol{a}}_{t,n}$ of the abundances at time instant t. The procedure is a two-step one:

- Considering $\boldsymbol{a}_n^{(t)} \equiv \widehat{\boldsymbol{a}}_n^{(t)}$, estimate $\boldsymbol{M}_n^{(t+1)}$, $\boldsymbol{\delta}_{\text{small},n}^{(t)}$ and $\boldsymbol{\delta}_{\text{large},n}^{(t)}$ that best represent pixel $\boldsymbol{y}_n^{(t+1)}$ in the model (130);
- Set $\widehat{\boldsymbol{a}}_n^{(t+1)} = \widehat{\boldsymbol{a}}_n^{(t)} + \boldsymbol{\delta}_{\text{small},n}^{(t)} + \boldsymbol{\delta}_{\text{large},n}^{(t)}$ and repeat for the next image.

The prior information stated about the properties of $\boldsymbol{\delta}_{\text{small},n}^{(t)}$ and $\boldsymbol{\delta}_{\text{large},n}^{(t)}$, and the knowledge of the spectral library $\mathcal{M}$ could then be used to formulate an optimization problem based on (129) to solve the multitemporal SU problem. Unfortunately, this problem would still be computationally very expensive when compared to MESMA.

Rather than estimating the abundances, the endmembers and the changes jointly at the same time, we adopt an alternative strategy to obtain an efficient solution. First, let us assume that $\boldsymbol{\delta}_{\text{large},n}^{(t)} = \boldsymbol{0}$. If $\boldsymbol{\delta}_{\text{small},n}^{(t)}$ is sufficiently small and $\widehat{\boldsymbol{a}}_n^{(t)}$ is a good estimate

of the true abundance $\boldsymbol{a}_n^{(t)}$, we have:

$$\begin{aligned} \boldsymbol{a}_n^{(t)} + \boldsymbol{\delta}_{\text{small},n}^{(t)} &\approx \boldsymbol{a}_n^{(t)} \\ &\approx \widehat{\boldsymbol{a}}_n^{(t)} . \end{aligned} \tag{131}$$

In that case, we can isolate the problem of estimating $\boldsymbol{M} \in \mathcal{M}$ from the full SU problem at time $t+1$ in order to solve it separately in a much simpler manner. We formulate the optimization problem as follows:

$$RE_{M,n}^{(t+1)} = \min_{\boldsymbol{M} \in \mathcal{M}} \left\| \boldsymbol{y}_n^{(t+1)} - \boldsymbol{M}\, \widehat{\boldsymbol{a}}_n^{(t)} \right\| . \tag{132}$$

The endmember matrix $\widehat{\boldsymbol{M}}_n^{(t+1)}$ obtained by solving problem (132) can then be used to compute abundance vector $\widehat{\boldsymbol{a}}_n^{(t+1)}$ with a single run of the FCLS algorithm.

However, this strategy relies on a strong hypothesis, namely, $\boldsymbol{\delta}_{\text{large},n}^{(t)} = \mathbf{0}$ and $\widehat{\boldsymbol{a}}_n^{(t)} \approx \boldsymbol{a}_n^{(t)}$, in order for the approximation in (131) to hold. This hypothesis may not be satisfied for all pixels. Fortunately, it turns out that we can devise a simple strategy to address those cases without significantly compromising the performance of the algorithm. Specifically, by evaluating the magnitude of the reconstruction error $RE_{M,n}^{(t+1)}$ in (132), we can indirectly identify if there were any significant changes in the abundance vector by testing whether $RE_{M,n}^{(t+1)}$ is larger than a given threshold RE_0 and, if so, estimate the corresponding abundance vector from scratch using MESMA. The threshold RE_0 offers a trade-off between accuracy (when small) and computational performance (when large), and can be set based on the typical reconstruction error achieved by MESMA.

Intuitively, the reason this works is that, if the spectral library $\mathcal{M}$ is not too large, and if $\boldsymbol{a}_n^{(t+1)} \approx \widehat{\boldsymbol{a}}_n^{(t)}$ is not satisfied, then we cannot accurately reconstruct pixel $\boldsymbol{y}_n^{(t+1)}$ from problem (132). This notion was formalized mathematically, which provided a deeper theoretical analysis of this method. This led to abundance reconstructions matching that of MESMA, which a much smaller computational cost. For more details, please see the original publication [42].

9.2.3 Conclusions

In this section, we presented an overview of two new multitemporal spectral unmixing methods, one based on a parametric model for the EMs, and another using a spectral library (for more details, see [41, 42]). In the first method, we used a parametric EM model to represent EM dynamics indirectly through parameters that capture the time variations of each material, instead of operating directly in the spectral space. Bayesian filtering and smoothing was then used to estimate the parameters of the endmembers. Assuming small abundance variations in short time intervals, the Expectation Maximization algorithm was employed to blindly estimate the remaining parameters, including the average fractional abundances, directly from the observed HI data. Simulation results indicate that

the proposed method can outperform state-of-the-art multitemporal spectral unmixing algorithms for slowly varying abundances.

To address the case of abrupt abundance variations and to benefit from a priori knowledge in spectral libraries, we proposed a second multitemporal SU method based on an efficient dynamical extension of the MESMA algorithm. The proposed strategy exploits the high temporal correlation of the abundance maps in order to improve both the accuracy and the computational complexity of the algorithm, which is one of the main drawbacks of MESMA. Specifically, we approximated the solution to the multitemporal unmixing problem by separating it into two sub-problems, namely, endmember selection and abundance estimation, which are much easier to solve individually. A simple strategy was proposed to detect abrupt abundance changes by analyzing residuals of the endmember selection problem. Theoretical results demonstrated how the proposed method compares to MESMA in terms of quality, and effectiveness in detecting abrupt abundance changes. Besides, these results also provide valuable insight into the conditions under which the approximate algorithm succeeds. Simulation results showed that the proposed method gives results with quality similar to, or better than, both MESMA and parametric models at a reduced computational complexity when an accurate spectral library is available.

9.3 IMAGE FUSION WITH SPECTRAL VARIABILITY

The low spatial resolution of HIs arises due to the delicate trade-off on which hyperspectral sensors operate in terms of pixel size and SNR. Since the radiated light observed at the sensor must be divided into a large number of spectral bands, the size of each HI pixel must be large enough to attain a minimum signal to noise ratio [6]. Multispectral sensors, on the other hand, provide images with much higher spatial resolution, albeit with a small number of spectral bands. One approach to obtain images with high spatial and spectral resolutions consists in combining hyperspectral and multispectral images (MI) of the same scene, resulting in the so-called HI-MI image fusion problem [28].

A large number of algorithms have been proposed to solve this problem (see, e.g., [28] and references therein). Most existing algorithms, however, share a common limitation: they assume that the HI and the MI are acquired under the same conditions. However, despite the short revisit cycles provided by the increasing number of optical satellites orbiting the Earth (e.g. Sentinel, Orbview, Landsat and Quickbird missions), the number of platforms carrying both hyperspectral and multispectral sensors is still considerably limited [223, 224]. This makes combining hyperspectral and multispectral observations acquired on board of different satellites of great interest to obtain high-resolution (HR) images [29, 225]. Images acquired at different time instants can be impacted by, e.g., illumination, atmospheric or seasonal changes. This may result in significant variations between the HI and the MI [15], negatively impacting traditional image fusion algorithms.

To address this issue, we propose two new image fusion methods accounting for

spectral and spatial variability between the images, which can be caused by, e.g., acquisition (atmospheric, illumination) or seasonal changes. Differently from previous approaches, we allow the high-resolution images underlying the observed HI and MI to be different from one another. In the first algorithm, we allow the spectral signatures of the materials in the HI and in the MI to be different, which accounts for spatially uniform variability. Afterwards, to account for spatially localized variations, we propose a tensor-based formulation using an additive variability term, which accounts for arbitrary changes between the images. By assuming that the images and variability tensors have low rank, two algorithms are then proposed. Theoretical results were also derived to demonstrate the exact recovery capability of the method.

9.3.1 The imaging model

Let $\boldsymbol{Y} \in \mathbb{R}^{L \times N}$ be an observed low spatial resolution HI with L bands and N pixels, and $\boldsymbol{Y}_m \in \mathbb{R}^{L_m \times \widetilde{N}}$ an observed low spectral resolution MI with L_m bands and $\widetilde{N}$ pixels, with $L_m < L$ and $N < \widetilde{N}$. The image fusion problem consists of estimating an underlying image $\boldsymbol{S} \in \mathbb{R}^{L \times \widetilde{N}}$ with high spatial and spectral resolutions, given $\boldsymbol{Y}$ and $\boldsymbol{Y}_m$. Traditional methods assume that the observed HI and MI are generated according to [28]

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{S}\boldsymbol{D}_{\text{spat}} + \boldsymbol{E} \\ \boldsymbol{Y}_m &= \boldsymbol{D}_{\text{spec}}\boldsymbol{S} + \boldsymbol{E}_m \end{aligned} \tag{133}$$

where $\boldsymbol{D}_{\text{spat}} \in \mathbb{R}^{\widetilde{N} \times N}$ accounts for optical blurring due to the sensor point spread function and for spatial downsampling, $\boldsymbol{D}_{\text{spec}} \in \mathbb{R}^{L_m \times L}$ is a matrix containing the spectral response functions of each band of the multispectral instrument. $\boldsymbol{E} \in \mathbb{R}^{L \times N}$ and $\boldsymbol{E}_m \in \mathbb{R}^{L_m \times \widetilde{N}}$ represent additive noise.

However, this model does not account for any type of variability between the images, what can propagate estimation errors throughout the fusion process when the images are acquired under different conditions. We propose to consider distinct HR images $\boldsymbol{S}_h$ and $\boldsymbol{S}_m$, both with L bands and $\widetilde{N}$ pixels, corresponding to the observed HI and MI, respectively. These HR images can be different due to variability effects. This leads to the following observation model:

$$\begin{aligned} \boldsymbol{Y} &= \boldsymbol{S}_h\boldsymbol{D}_{\text{spat}} + \boldsymbol{E}\,, \\ \boldsymbol{Y}_m &= \boldsymbol{D}_{\text{spec}}\boldsymbol{S}_m + \boldsymbol{E}_m\,. \end{aligned} \tag{134}$$

In the following, we propose two formulations for image fusion accounting for spectral, and for spectral and spatial variability.

9.3.2 A matrix factorization formulation for spatially invariant variability

A common approach to solve this problem is based on a low-rank assumption, which is directly connected to the LMM [28, 226]. Traditionally, this amounts to considering

that the HR image $\boldsymbol{S}$ can be represented $\boldsymbol{S} = \boldsymbol{M}_0\boldsymbol{A}_{\text{HR}}$, where $\boldsymbol{M}_0 \in \mathbb{R}^{L\times P}$ contain the spectral signatures of the image, and $\boldsymbol{A}_{\text{HR}} \in \mathbb{R}^{P\times \tilde{N}}$ are the HR abundance maps. In our case, we allow the HR images to be different due to seasonal variability effects. We then decompose $\boldsymbol{S}_h$ and $\boldsymbol{S}_m$ as

$$\boldsymbol{S}_h = \boldsymbol{M}_0\boldsymbol{A}_{\text{HR}} \tag{135}$$

$$\boldsymbol{S}_m = \boldsymbol{M}_m\boldsymbol{A}_{\text{HR}}\,, \tag{136}$$

where the abundances $\boldsymbol{A}_{\text{HR}}$ are assumed to be the same for both images, and $\boldsymbol{M}_0, \boldsymbol{M}_m \in \mathbb{R}^{L\times P}$ are the endmember matrices of the HI and for the MI, respectively. $\boldsymbol{M}_0$ and $\boldsymbol{M}_m$ can be different to account for spectral variability.

Matrix $\boldsymbol{M}_0$ can be estimated from the observed HI $\boldsymbol{Y}$ using endmember extraction or subspace decomposition. However, the same is not true for the estimation of $\boldsymbol{M}_m$ since the MI $\boldsymbol{Y}_m$ has low spectral resolution. To address this issue, we propose to write $\boldsymbol{M}_m$ as a function of $\boldsymbol{M}_0$ using a specific model for spectral variability.

Considering the GLMM proposed in Chapter 5, we propose to model the multispectral endmember matrix $\boldsymbol{M}_m$ as a function of the endmembers extracted from the HI as

$$\boldsymbol{M}_m = \boldsymbol{\Psi} \odot \boldsymbol{M}_0\,, \tag{137}$$

where $\boldsymbol{\Psi} \in \mathbb{R}^{L\times P}$ is a matrix of positive scaling factors and $\odot$ denotes the Hadamard product. Then, using models in (137), (135) and (134), the image fusion problem can finally be formulated as the problem of recovering the matrices $\boldsymbol{M}_0$, $\boldsymbol{\Psi}$ and $\boldsymbol{A}_{\text{HR}}$ from the observed HI $\boldsymbol{Y}$ and MI $\boldsymbol{Y}_m$. This was formulated as a coupled matrix factorization-inspired problem, which also considered different regularizations to control the amount of variability between the images, the spectral smoothness of $\boldsymbol{\Psi}$, and the spatial smoothness of $\boldsymbol{A}_{\text{HR}}$. Its solution was obtained by using an ALS strategy and the ADMM algorithm. For more details, please see the complete publication [43].

9.3.3 A tensor formulation

The natural representation of HIs and MIs as 3-dimensional tensors has been successfully exploited for HI-MI image fusion by formulating this problem as a coupled tensor approximation problem [227, 228]. Algorithms based on, e.g., the CPD [227] or Tucker decomposition [228] were proposed. The use of tensor-based strategies provided advantages including a superior experimental performance and theoretical recoverability guarantees. In order to extend the strategy outlined in Section 9.3.2 to account for spatially localized changes and to provide theoretical recovery guarantees, in this section we consider a tensor-based image fusion strategy, considering a more general, additive variability term.

Let us represent the HI as the order-3 tensor $\mathcal{Y} \in \mathbb{R}^{N_1\times N_2\times L}$, where N_1 and N_2 are the spatial and L the spectral dimensions, and the MI as the order-3 tensor $\mathcal{Y}_m \in$

$\mathbb{R}^{\tilde{N}_1 \times \tilde{N}_2 \times L_m}$, where $\tilde{N}_1 > N_1$ and $\tilde{N}_2 > N_2$ are the spatial and $L_m < L$ the spectral dimensions. Previous tensor-based works also assume a single HR image to underlie the observed HI and MI. We extend this approach by considering two different HR images $\mathcal{S}_h \in \mathbb{R}_+^{\tilde{N}_1 \times \tilde{N}_2 \times L}$ and $\mathcal{S}_m \in \mathbb{R}_+^{\tilde{N}_1 \times \tilde{N}_2 \times L}$, both with high spectral and spatial resolutions, corresponding to a tensor representation of $\boldsymbol{S}_h$ and $\boldsymbol{S}_m$ in (134).

By assuming that the spatial degradation (blurring and downsampling) operator $\boldsymbol{D}_{\text{spat}}$ is separable for each spatial dimension (see e.g., [227, 228] for more details), this leads to the following representation of model (134):

$$\mathcal{Y} = \mathcal{S}_h \times_1 \boldsymbol{D}_{1,\text{spat}} \times_2 \boldsymbol{D}_{2,\text{spat}} + \mathcal{E}\,, \tag{138}$$

$$\mathcal{Y}_m = \mathcal{S}_m \times_3 \boldsymbol{D}_{\text{spec}} + \mathcal{E}_m\,. \tag{139}$$

where $\times_k$ is the mode-k product between a tensor and a matrix (defined in detail in Section 6.2), tensors $\mathcal{E}_m \in \mathbb{R}^{\tilde{N}_1 \times \tilde{N}_2 \times L_m}$ and $\mathcal{E} \in \mathbb{R}^{N_1 \times N_2 \times L}$ represent additive noise and matrices $\boldsymbol{D}_{1,\text{spat}} \in \mathbb{R}^{N_1 \times \tilde{N}_1}$ and $\boldsymbol{D}_{2,\text{spat}} \in \mathbb{R}^{N_2 \times \tilde{N}_2}$ represent the spatial blurring and downsampling in the hyperspectral sensor for each spatial dimension.

Without loss of generality, we assume that both HR images $\mathcal{S}_h$ and $\mathcal{S}_m$ are related to each other as follows:

$$\mathcal{S}_m = \mathcal{S}_h + \mathbb{\Delta}\,, \tag{140}$$

where $\mathbb{\Delta} \in \mathbb{R}^{\tilde{N}_1 \times \tilde{N}_2 \times L}$ is an additive variability tensor representing changes between the scenes. Considering the variability model (140) along with (138)–(139), we obtain the following observation model for the acquired HI and MI:

$$\mathcal{Y} = \mathcal{S}_h \times_1 \boldsymbol{D}_{1,\text{spat}} \times_2 \boldsymbol{D}_{2,\text{spat}} + \mathcal{E}\,, \tag{141}$$

$$\mathcal{Y}_m = \left(\mathcal{S}_h + \mathbb{\Delta}\right) \times_3 \boldsymbol{D}_{\text{spec}} + \mathcal{E}_m\,. \tag{142}$$

The image fusion problem in this case consists in recovering $\mathbb{\Delta}$ and $\mathcal{S}_h$ from the observed images $\mathcal{Y}$ and $\mathcal{Y}_m$. More precisely,

$$\begin{cases} \text{find } \mathcal{S}_h \in \Omega_S \text{ and } \mathbb{\Delta} \in \Omega_\Delta \\ \text{such that equations (141)–(142) are satisfied.} \end{cases} \tag{143}$$

where the sets $\Omega_S \subseteq \mathbb{R}^{M_1 \times M_2 \times L}$ and $\Omega_\Delta \subseteq \mathbb{R}^{M_1 \times M_2 \times L}$ denote prior information about the HR image and the variability factor, respectively.

Since the number of unknowns is significantly greater than the number of available data, problem (143) is severely ill-posed and additional a priori information about the structure of $\mathcal{S}_h$ and $\mathbb{\Delta}$ must be introduced through the sets Ω_S and Ω_Δ in order to obtain a stable recovery. We considered a Tucker-based low-rank tensor model for both $\mathcal{S}_h$ and $\mathbb{\Delta}$. This structure made it possible to obtain exact recovery guarantees for problem (143), where spatial and spectral variabilities are present. Moreover, it also makes the problem

well-posed and easier to solve since the number of unknowns becomes smaller than the amount of available data. Two algorithms were proposed, one being algebraic (i.e., non iterative) and another based on a coupled tensor approximation optimization problem. For more details, see the related reference [44].

9.3.4 Conclusions

In this section, we presented an overview of two image fusion methods accounting for spectral and spatial-spectral variability between the images (for more details, see [43, 44]). In the first method, which employs a subspace/unmixing-based formulation, we used a unique set of endmembers for each image, with a parametric model to represent the variability of the spectral signatures occurring between them. The proposed algorithm estimated the subspace components (endmembers and abundance maps) of the HR images using an alternating optimization approach, making use of the ADMM to solve each subproblem. Simulation results with synthetic and real data showed that the proposed approach performed similarly or better than state-of-the-art methods for images acquired under the same conditions, and much better when there was spatially uniform seasonal or acquisition variations between them.

To address the case of spatially varying variability (i.e., both spatial and spectral variability) and to offer theoretical guarantees for the recovery of the HR image, we used a tensor based formulation for the second method. A more general observation model was considered, in which the HR image underlying the MI admits an additive variability term to account for changes between the scenes. Studying the general identifiability of this model, we showed that this variability term can only be identified in general up to its smooth structure (which is defined according to the degradation operators). To introduce additional prior information and mitigate the ambiguity associated with the proposed model, both the HR image and the additive perturbations were assumed to have low multilinear rank (i.e., to admit a Tucker decomposition). Two algorithms were then proposed, one algebraic and another based on an optimization procedure. Theoretical guarantees for the exact recovery of the HR image were provided for both. Simulation results showed that the proposed optimization-based algorithm yielded superior performance at a considerably lower computational cost when compared to the previous algorithm, especially when spatially localized variability was considered.

"where Saint Mary Woolnoth kept the hours,
With a dead sound on the final stroke of nine."
T.S. Elliot, *The Wasteland*

10 CONCLUSIONS

In this thesis, new models and algorithms were proposed to address spectral variability in spectral unmixing. To this end, we approached the unmixing problem from two perspectives, one in which we focused at modeling the abundances, and another in which we focused at modeling the endmembers.

In Chapter 3, we presented a multiscale model based on (over)-segmentation algorithms to represent the hyperspectral image and the abundance maps for sparse spectral unmixing using spectral libraries. The unmixing problem was divided into two separate problems: one in a coarse spatial scale, and another in the original image domain. The solution to SU at the coarse scale was able to characterize the relationship between the abundances at neighboring pixels, and was then used to regularize the SU problem at the original spatial scale. Differently from traditional spatial regularization techniques such as the Total Variation, the proposed multiscale regularization led to optimization problems without explicit dependency between the solutions at different pixels. This resulted in better abundance estimation performance at a much reduced computational complexity.

This multiscale regularization strategy was then extended in Chapter 4 to consider the SU problem with a parametric model to represent spatial endmember variability. Instead of separating SU into two distinct problems, we proposed a solution based on a single optimization problem. In particular, a different multiscale transformation was employed to decompose the image into two domains; one containing the coarse spatial information, and another containing only the fine spatial details. Spatial regularity of the abundance maps was then controlled by penalizing the energy of the abundances separately in each scale during SU, which was formulated as a non-convex optimization problem. Using an alternating least squares approach and a few simplifying hypotheses, an efficient solution was obtained, leading to improved abundance estimation at a lower computational complexity.

After that, in Chapters 5–8, we developed new SU methods by modeling the spectral signatures of the endmembers using parametric, tensor, and neural network-based representations. First, a parametric model was proposed in Chapter 5 to represent the spectral signatures of the endmembers in the scene as wavelength and pixel-dependent scalings of a set of reference spectral signatures. The abundance maps and scaling tensor were obtained as the solution to a non-convex optimization problem, which was performed using an alternating least squares approach. Simulations indicated that the proposed model can improve the quality of SU.

In Chapter 6, the multidimensional nature of the abundance maps and endmember matrices for each pixel was explored to address spectral variability in SU. Specifically, no direct model for the endmember signatures was employed. Instead, the natural representation of the data as high-order tensors and their underlying low-rank structure were used

to regularize the SU problem, which was formulated as a low-rank tensor approximation method. A simple strategy was also proposed to select the approximate ranks of the abundance and endmember tensors. Simulations with synthetic and real images showed that the proposed strategy can lead to good SU results, while offering a competitive computational cost for small hyperspectral images.

Despite obtaining a good practical performance, strategies based on parametric or tensor-based models do not exploit the fact that the spectral signatures of the endmembers are confined to a low-dimensional manifold. To make use of this property, in Chapter 7 we proposed to model the endmembers using generative neural networks, such as variational autoencoders and generative adversarial networks. This allowed us to represent the endmember manifold as the image of a function on a low-dimensional Euclidean space. A two-step SU procedure was then proposed. First, pure pixels were extracted from the observed image and used to train the generative networks, one for each material in the scene. Then, the abundances and the latent representations of the endmembers were estimated by means of a matrix factorization-inspired optimization problem. Experimental results showed excellent performance on scenes containing pure pixels. Chapter 8 extended this approach to aid library-based SU for cases where the library has a small quantity of spectral signatures. In this case, generative models for each material were first learned using the spectral signatures contained in the original library as training data. Afterwards, the learned models were used to generate additional synthetic signatures of each material which were then included in the library before SU, following a data augmentation framework. Experimental results showed the method's performance to be comparable to that of data augmentation approaches based on physics models specified a priori.

Finally, Chapter 9 presented a brief overview of work that extended the contributions of Chapters 3–8 to address other challenges related to spectral unmixing and to hyperspectral image analysis. First, the multiscale spatial regularization method proposed in Chapter 3 was extended to address nonlinear unmixing with kernels. Besides extending the multiscale decomposition to a non-parametric mixing model, a framework was also developed to automatically adjust the regularization parameters of the algorithm. SU was performed by using a computationally efficient solution to a non-convex optimization problem.

The parametric model developed in Chapter 5 was later extended to address spectral variability found in images acquired at different time instants. Using a Bayesian filtering strategy and the Expectation Maximization algorithm, a method was proposed to perform SU in multitemporal image sequences with slowly varying abundances. To mitigate this latter limitation and address image sequences containing sudden changes, a library-based method was also proposed to perform SU and change detection jointly based on an approximate solution to a combinatorial optimization problem.

Finally, spectral and spatial variability were addressed in the multimodal (i.e.,

hyperspectral and multispectral) image fusion problem. By considering that the spectral signatures of the materials can be different for each of the images, the parametric model of Chapter 5 was first extended to represent the spectral variability present between the images in the different modalities. This led to a considerable performance improvement for image pairs in which the variability is spatially invariant. To address spatially and spectrally varying variability, image fusion was latter cast as a coupled tensor factorization problem, subject to an arbitrary additive variability term between the images. Assuming that the high-resolution images have low rank, two algorithms are proposed, each of which with theoretical performance guarantees. Simulations with synthetic and real data showed a considerable performance improvement when spatially localized variability is present.

Several possibilities remain to further extend the work in this thesis, among which:

- Develop new multiscale decompositions to represent the spatial information contained within each HI more effectively, which could improve the performance of the methods presented in Chapters 3 and 4;
- Improve the computational complexity of the ULTRA-V algorithm of Chapter 6 by using a patch-based spatial decomposition, making this method scalable and more amenable to processing large datasets;
- Develop strategies to automatically adjust the regularization parameters of SU algorithms that account for spectral variability, allowing a more widespread use of these techniques in practice;
- Provide theoretical guarantees for the recoverability of the endmembers and of the abundances in SU when spectral variability is considered;
- Further investigate the use of deep learning tools to aid in solving the SU problem with spectral variability.

"Well now that's done"
T.S. Elliot, *The Wasteland*

REFERENCES

[1] BIOUCAS-DIAS, J. M.; PLAZA, A.; CAMPS-VALLS, G.; SCHEUNDERS, P.; NASRABADI, N.; CHANUSSOT, J. Hyperspectral Remote Sensing Data Analysis and Future Challenges. **IEEE Geoscience and Remote Sensing Magazine**, v. 1, n. 2, p. 6–36, 2013.

[2] IMBIRIBA, T. C. d. O. **Nonlinear hyperspectral unmixing: strategies for nonlinear mixture detection, endmember estimation and band-selection**. 2016. PhD thesis – Federal University of Santa Catarina, Florianópolis.

[3] KOUYAMA, T.; YOKOTA, Y.; ISHIHARA, Y.; NAKAMURA, R.; YAMAMOTO, S.; MATSUNAGA, T. Development of an application scheme for the SELENE/SP lunar reflectance model for radiometric calibration of hyperspectral and multispectral sensors. **Planetary and Space Science**, Elsevier, v. 124, p. 76–83, 2016.

[4] MANOLAKIS, D. Detection algorithms for hyperspectral imaging applications. **IEEE Signal Processing Magazine**, v. 19, n. 1, p. 29–43, 2002.

[5] LU, G.; FEI, B. Medical hyperspectral imaging: a review. **Journal of biomedical optics**, International Society for Optics and Photonics, v. 19, n. 1, p. 010901, 2014.

[6] SHAW, G. A.; BURKE, H.-h. K. Spectral imaging for remote sensing. **Lincoln laboratory journal**, v. 14, n. 1, p. 3–28, 2003.

[7] KESHAVA, N.; MUSTARD, J. F. Spectral unmixing. **IEEE Signal Processing Magazine**, v. 19, n. 1, p. 44–57, 2002.

[8] BIOUCAS-DIAS, J. M.; PLAZA, A.; DOBIGEON, N.; PARENTE, M.; DU, Q.; G., P.; CHANUSSOT, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 5, n. 2, p. 354–379, 2012.

[9] MA, W.-K.; BIOUCAS-DIAS, J. M.; CHAN, T.-H.; GILLIS, N.; GADER, P.; PLAZA, A. J.; AMBIKAPATHI, A.; CHI, C.-Y. A Signal Processing Perspective on Hyperspectral Unmixing: Insights from Remote Sensing. **IEEE Signal Processing Magazine**, v. 31, n. 1, p. 67–81, Jan. 2014.

[10] ZARE, A.; HO, K. C. Endmember Variability in Hyperspectral Analysis: Addressing Spectral Variability During Spectral Unmixing. **IEEE Signal Processing Magazine**, v. 31, p. 95–104, Jan. 2014.

[11] SOMERS, B.; ASNER, G. P.; TITS, L.; COPPIN, P. Endmember variability in spectral mixture analysis: A review. **Remote Sensing of Environment**, Elsevier, v. 115, n. 7, p. 1603–1616, 2011.

[12] GARCÍA-HARO, F.; SOMMER, S.; KEMPER, T. A new tool for variable multiple endmember spectral mixture analysis (VMESMA). **International Journal of Remote Sensing**, Taylor & Francis, v. 26, n. 10, p. 2135–2162, 2005.

[13] DRUMETZ, L.; CHANUSSOT, J.; JUTTEN, C. Variability of the endmembers in spectral unmixing: recent advances. In: 8TH IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. Los Angeles, USA: [s.n.], Aug. 2016. P. 1–5.

[14] DRUMETZ, L.; CHANUSSOT, J.; JUTTEN, C. Chapter 2.7 - Variability of the endmembers in spectral unmixing. In: AMIGO, J. M. (Ed.). **Hyperspectral Imaging**. [S.l.]: Elsevier, 2020. v. 32. (Data Handling in Science and Technology). P. 167–203. DOI: https://doi.org/10.1016/B978-0-444-63977-6.00009-2.

[15] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C.; CHANUSSOT, J.; DRUMETZ, L.; TOURNERET, J.-Y.; ZARE, A.; JUTTEN, C. Spectral Variability in Hyperspectral Data Unmixing: A Comprehensive Review. **IEEE Geoscience and Remote Sensing Magazine**, 2021.

[16] ROBERTS, D. A.; GARDNER, M.; CHURCH, R.; USTIN, S.; SCHEER, G.; GREEN, R. Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. **Remote Sensing of Environment**, Elsevier, v. 65, n. 3, p. 267–279, 1998.

[17] STEIN, D. Application of the normal compositional model to the analysis of hyperspectral imagery. In: PROC. IEEE Workshop on Adv. Tech. for Anal. of Remotely Sens. Data. Greenbelt, MD, USA: [s.n.], 2003. P. 44–51.

[18] EISMANN, M. T.; HARDIE, R. C. Stochastic spectral unmixing with enhanced endmember class separation. **Applied Optics**, Optical Society of America, v. 43, n. 36, p. 6596–6608, 2004.

[19] DRUMETZ, L.; VEGANZONES, M.-A.; HENROT, S.; PHLYPO, R.; CHANUSSOT, J.; JUTTEN, C. Blind hyperspectral unmixing using an Extended Linear Mixing Model to address spectral variability. **IEEE Transactions on Image Processing**, IEEE, v. 25, n. 8, p. 3890–3905, 2016.

[20] THOUVENIN, P.-A.; DOBIGEON, N.; TOURNERET, J.-Y. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. **IEEE Transactions on Signal Processing**, v. 64, n. 2, p. 525–538, Feb. 2016.

[21] IORDACHE, M.-D.; BIOUCAS-DIAS, J. M.; PLAZA, A. Total variation spatial regularization for sparse hyperspectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 50, n. 11, p. 4484–4502, 2012.

[22] WANG, R.; LI, H.-C.; LIAO, W.; HUANG, X.; PHILIPS, W. Centralized Collaborative Sparse Unmixing for Hyperspectral Images. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 10, n. 5, p. 1949–1962, 2017.

[23] DOBIGEON, N.; TOURNERET, J.-Y.; RICHARD, C.; BERMUDEZ, J. C. M.; MCLAUGHLIN, S.; HERO, A. O. Nonlinear Unmixing of Hyperspectral Images: Models and Algorithms. **IEEE Signal Processing Magazine**, v. 31, n. 1, p. 82–94, Jan. 2014.

[24] CHI, M.; PLAZA, A.; BENEDIKTSSON, J. A.; SUN, Z.; SHEN, J.; ZHU, Y. Big data for remote sensing: challenges and opportunities. **Proceedings of the IEEE**, IEEE, v. 104, n. 11, p. 2207–2219, 2016.

[25] MA, Y.; WU, H.; WANG, L.; HUANG, B.; RANJAN, R.; ZOMAYA, A.; JIE, W. Remote sensing big data computing: Challenges and opportunities. **Future Generation Computer Systems**, Elsevier, v. 51, p. 47–60, 2015.

[26] HEYLEN, R.; PARENTE, M.; GADER, P. A review of nonlinear hyperspectral unmixing methods. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 7, n. 6, p. 1844–1868, June 2014.

[27] LIU, S.; MARINELLI, D.; BRUZZONE, L.; BOVOLO, F. A Review of Change Detection in Multitemporal Hyperspectral Images: Current Techniques, Applications, and Challenges. **IEEE Geoscience and Remote Sensing Magazine**, IEEE, v. 7, n. 2, p. 140–158, 2019.

[28] YOKOYA, N.; GROHNFELDT, C.; CHANUSSOT, J. Hyperspectral and Multispectral Data Fusion: A comparative review of the recent literature. **IEEE Geoscience and Remote Sensing Magazine**, IEEE, v. 5, n. 2, p. 29–56, 2017.

[29] HILKER, T.; WULDER, M. A.; COOPS, N. C.; LINKE, J.; MCDERMID, G.; MASEK, J. G.; GAO, F.; WHITE, J. C. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. **Remote Sensing of Environment**, Elsevier, v. 113, n. 8, p. 1613–1627, 2009.

[30] GAO, F.; MASEK, J.; SCHWALLER, M.; HALL, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. **IEEE Transactions on Geoscience and Remote sensing**, IEEE, v. 44, n. 8, p. 2207–2218, 2006.

[31] ROY, D. P.; JU, J.; LEWIS, P.; SCHAAF, C.; GAO, F.; HANSEN, M.; LINDQUIST, E. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. **Remote Sensing of Environment**, Elsevier, v. 112, n. 6, p. 3112–3130, 2008.

[32] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C. A Fast Multiscale Spatial Regularization for Sparse Hyperspectral Unmixing. **IEEE Geoscience and Remote Sensing Letters**, v. 16, n. 4, p. 598–602, Apr. 2019. ISSN 1545-598X. DOI: 10.1109/LGRS.2018.2878394.

[33] BORSOI, R. A.; IMBIRIBA, T.; MOREIRA BERMUDEZ, J. C. A Data Dependent Multiscale Model for Hyperspectral Unmixing With Spectral Variability. **IEEE Transactions on Image Processing**, v. 29, p. 3638–3651, 2020.

[34] IMBIRIBA, T.; BORSOI, R. A.; BERMUDEZ, J. C. M. Generalized linear mixing model accounting for endmember variability. In: PROC. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Calgary, Canada: [s.n.], 2018. P. 1862–1866.

[35] IMBIRIBA, T.; BORSOI, R. A.; BERMUDEZ, J. C. M. Low-Rank Tensor Modeling for Hyperspectral Unmixing Accounting for Spectral Variability. **IEEE Transactions on Geoscience and Remote Sensing**, v. 58, n. 3, p. 1833–1842, 2020.

[36] FÉRET, J.-B.; GITELSON, A.; NOBLE, S.; JACQUEMOUD, S. PROSPECT-D: towards modeling leaf optical properties through a complete lifecycle. **Remote Sensing of Environment**, Elsevier, v. 193, p. 204–215, 2017.

[37] HAPKE, B. Bidirectional reflectance spectroscopy, 1, Theory. **Journal of Geophysical Research**, v. 86, B4, p. 3039–3054, 1981.

[38] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M. Deep Generative Endmember Modeling: An Application to Unsupervised Spectral Unmixing. **IEEE Transactions on Computational Imaging**, v. 6, p. 374–384, 2019.

[39] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C. Deep Generative Models for Library Augmentation in Multiple Endmember Spectral Mixture Analysis. **IEEE Geoscience and Remote Sensing Letters**, 2020. DOI: 10.1109/LGRS.2020.3007161.

[40] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C. A blind multiscale spatial regularization framework for kernel-based spectral unmixing. **IEEE Transactions on Image Processing**, IEEE, v. 29, p. 4965–4979, 2020.

[41] BORSOI, R. A.; IMBIRIBA, T.; CLOSAS, P.; BERMUDEZ, J. C. M.; RICHARD, C. Kalman filtering and expectation maximization for multitemporal spectral unmixing. **IEEE Geoscience and Remote Sensing Letters**, 2020. DOI: 10.1109/LGRS.2020.3025781.

[42] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C. Fast Unmixing and Change Detection in Multitemporal Hyperspectral Data. **arXiv preprint (submitted)**, 2021.

[43] BORSOI, R. A.; IMBIRIBA, T.; BERMUDEZ, J. C. M. Super-Resolution for Hyperspectral and Multispectral Image Fusion Accounting for Seasonal Spectral Variability. **IEEE Transactions on Image Processing**, v. 29, n. 1, p. 116–127, 2020. ISSN 1057-7149. DOI: 10.1109/TIP.2019.2928895.

[44] BORSOI, R. A.; PRÉVOST, C.; USEVICH, K.; BRIE, D.; BERMUDEZ, J. C. M.; RICHARD, C. Coupled Tensor Decomposition for Hyperspectral and Multispectral Image Fusion with Inter-image Variability. **IEEE Journal of Selected Topics in Signal Processing**, p. 1–16, 2021. DOI: 10.1109/JSTSP.2021.3054338.

[45] IMBIRIBA, T.; BORSOI, R. A.; BERMUDEZ, J. C. M. A Low-rank Tensor Regularization Strategy for Hyperspectral Unmixing. In: PROC. IEEE Statistical Signal Processing Workshop (SSP). [S.l.: s.n.], 2018. P. 373–377.

[46] BORSOI, R. A.; IMBIRIBA, T.; MOREIRA BERMUDEZ, J. C. Improved Hyperspectral Unmixing with Endmember Variability Parametrized Using an Interpolated Scaling Tensor. In: PROC. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton, UK: [s.n.], 2019. P. 2177–2181. DOI: 10.1109/ICASSP.2019.8683155.

[47] BORSOI, R. A.; RICHARD, C.; FERRARI, A.; CHEN, J.; BERMUDEZ, J. C. M. Online Graph-Based Change Point Detection in Multiband Image Sequences. In: IEEE. PROC. European Signal Processing Conference (EUSIPCO). Amsterdam, Netherlands: [s.n.], 2020. P. 850–854.

[48] GRIFFIN, M. K.; BURKE, H.-h. K. Compensation of hyperspectral data for atmospheric effects. **Lincoln Laboratory Journal**, v. 14, n. 1, p. 29–54, 2003.

[49] GAO, B.-C.; MONTES, M. J.; DAVIS, C. O.; GOETZ, A. F. Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. **Remote Sensing of Environment**, Elsevier, v. 113, s17–s24, 2009.

[50] HEALEY, G.; SLATER, D. Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 37, n. 6, p. 2706–2717, 1999.

[51] NASCIMENTO, J. M. P.; DIAS, J. M. B. Does Independent Component Analysis Play a Role in Unmixing Hyperspectral Data? **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, p. 175–187, Jan. 2005.

[52] MATTHEW, M. W.; ADLER-GOLDEN, S. M.; BERK, A.; FELDE, G.; ANDERSON, G. P.; GORODETZKY, D.; PASWATERS, S.; SHIPPERT, M. Atmospheric correction of spectral imagery: evaluation of the FLAASH algorithm with AVIRIS data. In: PROC. of the 31st Applied Imagery Pattern Recognition Workshop. Washington, DC: [s.n.], 2002. P. 157–163.

[53] LAU, I. C. Application of atmospheric correction to hyperspectral data: Comparisons of different techniques on Hymap data. In: PROC. of the 12th Australasian Remote Sensing and Photogrammetry Conference (ARSPC). Freemantle, Australia: [s.n.], 2004. P. 1–15.

[54] SONG, C.; WOODCOCK, C. E.; SETO, K. C.; LENNEY, M. P.; MACOMBER, S. A. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? **Remote sensing of Environment**, Elsevier, v. 75, n. 2, p. 230–244, 2001.

[55] GRIFFIN, M. K.; BURKE, H.; VAIL, J.; ADLER-GOLDEN, S.; MATTHEW, M. Sensitivity of atmospheric compensation model retrievals to input parameter specification. In: PROC. AVIRIS Earth Science and Applications Workshop. Pasadena, California: [s.n.], 1999. P. 99–17.

[56] WILSON, R.; MILTON, E.; NIELD, J. M. Spatial variability of the atmosphere over southern England, and its effect on scene-based atmospheric corrections. **International Journal of Remote Sensing**, Taylor & Francis, v. 35, n. 13, p. 5198–5218, 2014.

[57] BHATIA, N.; IORDACHE, M.-D.; STEIN, A.; REUSEN, I.; TOLPEKIN, V. A. Propagation of uncertainty in atmospheric parameters to hyperspectral unmixing. **Remote Sensing of Environment**, Elsevier, v. 204, p. 472–484, 2018.

[58] BASSANI, C.; MANZO, C.; BRAGA, F.; BRESCIANI, M.; GIARDINO, C.; ALBEROTANZA, L. The impact of the microphysical properties of aerosol on the atmospheric correction of hyperspectral data in coastal waters. **Atmospheric Measurement Techniques**, Copernicus GmbH, v. 8, n. 3, p. 1593–1604, 2015.

[59] KAUFMAN, Y. J.; GOBBI, G. P.; KOREN, I. Aerosol climatology using a tunable spectral variability cloud screening of AERONET data. **Geophysical research letters**, Wiley Online Library, v. 33, n. 7, 2006.

[60] SCHLÄPFER, D.; HUENI, A.; RICHTER, R. Cast Shadow Detection to Quantify the Aerosol Optical Thickness for Atmospheric Correction of High

Spatial Resolution Optical Imagery. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 10, n. 2, p. 200, 2018.

[61] KAUFMAN, Y.; HOLBEN, B. Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert reflection. **International Journal of Remote Sensing**, Taylor & Francis, v. 14, n. 1, p. 21–52, 1993.

[62] LARSEN, N. F.; STAMNES, K. Use of shadows to retrieve water vapor in hazy atmospheres. **Applied optics**, Optical Society of America, v. 44, n. 32, p. 6986–6994, 2005.

[63] MARKELIN, L.; SIMIS, S. G.; HUNTER, P. D.; SPYRAKOS, E.; TYLER, A. N.; CLEWLEY, D.; GROOM, S. Atmospheric correction performance of hyperspectral airborne imagery over a small eutrophic lake under changing cloud cover. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 9, n. 1, p. 2, 2016.

[64] STAENZ, K.; SECKER, J.; GAO, B.-C.; DAVIS, C.; NADEAU, C. Radiative transfer codes applied to hyperspectral data for the retrieval of surface reflectance. **ISPRS Journal of Photogrammetry and Remote Sensing**, Elsevier, v. 57, n. 3, p. 194–203, 2002.

[65] MURPHY, R. J.; MONTEIRO, S. T.; SCHNEIDER, S. Evaluating classification techniques for mapping vertical geology using field-based hyperspectral sensors. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 50, n. 8, p. 3066–3080, 2012.

[66] PLAZA, A.; MARTÍNEZ, P.; PÉREZ, R.; PLAZA, J. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 42, n. 3, p. 650–663, 2004.

[67] KESHAVA, N.; KEREKES, J.; MANOLAKIS, D.; SHAW, G. Algorithm taxonomy for hyperspectral unmixing. In: ALGORITHMS for Multispectral, Hyperspectral, and Ultraspectral Imagery VI. Orlando, FL: [s.n.], 2000. P. 42–63.

[68] MCGWIRE, K.; MINOR, T.; FENSTERMAKER, L. Hyperspectral mixture modeling for quantifying sparse vegetation cover in arid environments. **Remote Sensing of Environment**, Elsevier, v. 72, n. 3, p. 360–374, 2000.

[69] BOARDMAN, J. W. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In: PROC. 4th Annu. JPL Airbone Geosci. Workshop, Jet Propulsion Lab. Pasadena, USA: [s.n.], 1993. P. 11–14.

[70] DENNISON, P. E.; ROBERTS, D. A. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. **Remote sensing of environment**, Elsevier, v. 87, n. 2-3, p. 123–135, 2003.

[71] ROPER, T.; ANDREWS, M. Shadow modelling and correction techniques in hyperspectral imaging. **Electronics Letters**, IET, v. 49, n. 7, p. 458–460, 2013.

[72] ZHANG, Q.; PAUCA, V. P.; PLEMMONS, R. J.; NIKIC, D. D. Detecting objects under shadows by fusion of hyperspectral and LiDAR DATA: A physical model approach. In: PROC. 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). Gainesville, FL: [s.n.], 2013. P. 1–4.

[73] FITZGERALD, G. J.; PINTER, P. J.; HUNSAKER, D. J.; CLARKE, T. R. Multiple shadow fractions in spectral mixture analysis of a cotton canopy. **Remote Sensing of Environment**, Elsevier, v. 97, n. 4, p. 526–539, 2005.

[74] CHOI, K.; MILTON, E. An investigation into the properties of the dark endmember in spectral feature space. In: PROC. of the 25th IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Seoul, South Korea: [s.n.], 2005. P. 25–29.

[75] LYNCH, D. K. Shadows. **Applied optics**, Optical Society of America, v. 54, n. 4, b154–b164, 2015.

[76] ADLER-GOLDEN, S.; MATTHEW, M. W.; ANDERSON, G. P.; FELDE, G. W.; GARDNER, J. A. An algorithm for de-shadowing spectral imagery. In: PROC. 11th JPL Airborne Earth Science Workshop. Pasadena, USA: [s.n.], 2000. P. 1–8.

[77] SCHLÄPFER, D.; RICHTER, R.; DAMM, A. Correction of shadowing in imaging spectroscopy data by quantification of the proportion of diffuse illumination. In: PROC. 8th Imaging Spectroscopy Workshop (SIG-EARSeL). Nantes, France: [s.n.], 2013. P. 8–10.

[78] RICHTER, R.; KELLENBERGER, T.; KAUFMANN, H. Comparison of topographic correction methods. **Remote Sensing**, Molecular Diversity Preservation International, v. 1, n. 3, p. 184–196, 2009.

[79] FENG, J.; RIVARD, B.; SANCHEZ-AZOFEIFA, A. The topographic normalization of hyperspectral data: implications for the selection of spectral end members and lithologic mapping. **Remote Sensing of Environment**, Elsevier, v. 85, n. 2, p. 221–231, 2003.

[80] HAPKE, B. **Theory of Reflectance and Emittance Spectroscopy**. [S.l.]: Cambridge University Press, 1993.

[81] DRUMETZ, L.; CHANUSSOT, J.; JUTTEN, C. Spectral Unmixing: A Derivation of the Extended Linear Mixing Model from the Hapke Model. **IEEE Geoscience and Remote Sensing Letters**, v. 17, n. 11, p. 1866–1870, 2020.

[82] COMBAL, B.; ISAKA, H. The effect of small topographic variations on reflectance. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 40, n. 3, p. 663–670, 2002.

[83] COCHRANE, M. Using vegetation reflectance variability for species level classification of hyperspectral data. **International Journal of Remote Sensing**, Taylor & Francis, v. 21, n. 10, p. 2075–2087, 2000.

[84] ZHANG, J.; RIVARD, B.; SÁNCHEZ-AZOFEIFA, A.; CASTRO-ESAU, K. Intra-and inter-class spectral variability of tropical tree species at La Selva, Costa Rica: Implications for species identification using HYDICE imagery. **Remote Sensing of Environment**, Elsevier, v. 105, n. 2, p. 129–141, 2006.

[85] BAUMGARDNER, M. F.; SILVA, L. F.; BIEHL, L. L.; STONER, E. R. Reflectance properties of soils. In: ADVANCES in agronomy. [S.l.]: Elsevier, 1986. v. 38. P. 1–44.

[86] CROWLEY, J. K. Visible and near-infrared spectra of carbonate rocks: Reflectance variations related to petrographic texture and impurities. **Journal of Geophysical Research: Solid Earth**, Wiley Online Library, v. 91, B5, p. 5001–5012, 1986.

[87] CLARK, R. N. Spectroscopy of rocks and minerals, and principles of spectroscopy. In: RENCZ, A. N. (Ed.). **Remote Sensing for the Earth Sciences: Manual of Remote Sensing**. New York, NY, USA: Wiley, 1999. v. 3. P. 3–58.

[88] FRANKE, J.; ROBERTS, D. A.; HALLIGAN, K.; MENZ, G. Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. **Remote Sensing of Environment**, Elsevier, v. 113, n. 8, p. 1712–1723, 2009.

[89] PRICE, J. C. How unique are spectral signatures? **Remote Sensing of Environment**, Elsevier, v. 49, n. 3, p. 181–186, 1994.

[90] ASNER, G. P. Biophysical and biochemical sources of variability in canopy reflectance. **Remote Sensing of Environment**, Elsevier, v. 64, n. 3, p. 234–253, 1998.

[91] FERREIRA, M. P.; GRONDONA, A. E. B.; ROLIM, S. B. A.; SHIMABUKURO, Y. E. Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. **Journal of**

Applied Remote Sensing, International Society for Optics and Photonics, v. 7, n. 1, p. 073502–073502, 2013.

[92] GONG, P.; PU, R.; YU, B. Conifer species recognition: An exploratory analysis of in situ hyperspectral data. **Remote Sensing of Environment**, Elsevier, v. 62, n. 2, p. 189–200, 1997.

[93] LUKEŠ, P.; STENBERG, P.; RAUTIAINEN, M.; MÕTTUS, M.; VANHATALO, K. M. Optical properties of leaves and needles for boreal tree species in Europe. **Remote Sensing Letters**, Taylor & Francis, v. 4, n. 7, p. 667–676, 2013.

[94] GAO, Z.; ZHANG, L. Multi-seasonal spectral characteristics analysis of coastal salt marsh vegetation in Shanghai, China. **Estuarine, Coastal and Shelf Science**, Elsevier, v. 69, n. 1, p. 217–224, 2006.

[95] SCHMIDT, H.; KARNIELI, A. Remote sensing of the seasonal variability of vegetation in a semi-arid environment. **Journal of arid environments**, Elsevier, v. 45, n. 1, p. 43–59, 2000.

[96] MÕTTUS, M.; SULEV, M.; HALLIK, L. Seasonal course of the spectral properties of alder and birch leaves. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 7, n. 6, p. 2496–2505, 2014.

[97] JACQUEMOUD, S.; USTIN, S. L. Leaf optical properties: A state of the art. In: PROC. 8th International Symposium of Physical Measurements & Signatures in Remote Sensing. Aussois, France: [s.n.], 2001. P. 223–332.

[98] JACQUEMOUD, S.; BARET, F. PROSPECT: A model of leaf optical properties spectra. **Remote Sensing of Environment**, Elsevier, v. 34, n. 2, p. 75–91, 1990.

[99] VERHOEF, W. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. **Remote Sensing of Environment**, Elsevier, v. 16, n. 2, p. 125–141, 1984.

[100] DAWSON, T. P.; CURRAN, P. J.; PLUMMER, S. E. LIBERTY—Modeling the effects of leaf biochemical concentration on reflectance spectra. **Remote Sensing of Environment**, Elsevier, v. 65, n. 1, p. 50–60, 1998.

[101] LOBELL, D. B.; ASNER, G. P. Moisture effects on soil reflectance. **Soil Science Society of America Journal**, Soil Science Society, v. 66, n. 3, p. 722–727, 2002.

[102] SOMERS, B.; DELALIEUX, S.; VERSTRAETEN, W. W.; COPPIN, P. A conceptual framework for the simultaneous extraction of sub-pixel spatial extent and spectral characteristics of crops. **Photogrammetric Engineering &**

Remote Sensing, American Society for Photogrammetry and Remote Sensing, v. 75, n. 1, p. 57–68, 2009.

[103] SADEGHI, M.; JONES, S. B.; PHILPOT, W. D. A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. **Remote Sensing of Environment**, Elsevier, v. 164, p. 66–76, 2015.

[104] WISCOMBE, W. J.; WARREN, S. G. A model for the spectral albedo of snow. I: Pure snow. **Journal of the Atmospheric Sciences**, v. 37, n. 12, p. 2712–2733, 1980.

[105] WEBSTER, R.; CURRAN, P.; MUNDEN, J. Spatial correlation in reflected radiation from the ground and its implications for sampling and mapping by ground-based radiometry. **Remote Sensing of Environment**, Elsevier, v. 29, n. 1, p. 67–78, 1989.

[106] TOLA, E.; AL-GAADI, K.; MADUGUNDU, R.; ZEYADA, A.; KAYAD, A.; BIRADAR, C. Characterization of spatial variability of soil physicochemical properties and its impact on Rhodes grass productivity. **Saudi Journal of Biological Sciences**, Elsevier, v. 24, n. 2, p. 421–429, 2017.

[107] NAJAFIAN, A.; DAYANI, M.; MOTAGHIAN, H. R.; NADIAN, H. Geostatistical assessment of the spatial distribution of some chemical properties in calcareous soils. **Journal of Integrative Agriculture**, Elsevier, v. 11, n. 10, p. 1729–1737, 2012.

[108] WEI, Y.-C.; BAI, Y.-L.; JIN, J.-Y.; ZHANG, F.; ZHANG, L.-P.; LIU, X.-Q. Spatial variability of soil chemical properties in the reclaiming marine foreland to Yellow Sea of China. **Agricultural Sciences in China**, Elsevier, v. 8, n. 9, p. 1103–1111, 2009.

[109] HOU-LONG, J.; GUO-SHUN, L.; XIN-ZHONG, W.; WEN-FENG, S.; RUI-NA, Z.; CHUN-HUA, Z.; HONG-CHAO, H.; YAN-TAO, L. Spatial variability of soil properties in a long-term tobacco plantation in central China. **Soil Science**, LWW, v. 175, n. 3, p. 137–144, 2010.

[110] SHI, C.; WANG, L. Incorporating spatial information in spectral unmixing: A review. **Remote Sensing of Environment**, Elsevier, v. 149, p. 70–87, 2014.

[111] IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C.; TOURNERET, J.-Y. Nonparametric detection of nonlinearly mixed pixels and endmember estimation in hyperspectral images. **IEEE Transactions on Image Processing**, v. 25, n. 3, p. 1136–1151, Mar. 2016. ISSN 1057-7149. DOI: 10.1109/TIP.2015.2509258.

[112] BIOUCAS-DIAS, J. M.; PLAZA, A.; DOBIGEON, N.; PARENTE, M.; DU, Q.; GADER, P.; CHANUSSOT, J. Hyperspectral Unmixing Overview: Geometrical,

Statistical, and Sparse Regression-Based Approaches. **IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.**, v. 5, n. 2, p. 354–379, 2012.

[113] BIOUCAS-DIAS, J. M.; FIGUEIREDO, M. A. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In: PROC. WHISPERS. Reykjavik, Iceland: [s.n.], 2010. P. 1–4.

[114] LU, X.; WU, H.; YUAN, Y. Double constrained NMF for hyperspectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 52, n. 5, p. 2746–2758, 2014.

[115] BERNABÉ, S.; JIMÉNEZ, L. I.; GARCÍA, C.; PLAZA, J.; PLAZA, A. Multicore Real-Time Implementation of a Full Hyperspectral Unmixing Chain. **IEEE Geosci. Remote Sens. Lett.**, IEEE, v. 15, n. 5, p. 744–748, 2018.

[116] ECKSTEIN, J.; BERTSEKAS, D. P. On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators. **Mathematical Programming**, Springer, v. 55, n. 1, p. 293–318, 1992.

[117] VEGANZONES, M. A.; TOCHON, G.; DALLA-MURA, M.; PLAZA, A. J.; CHANUSSOT, J. Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation. **IEEE Transactions on Image Processing**, IEEE, v. 23, n. 8, p. 3574–3589, 2014.

[118] ACHANTA, R.; SHAJI, A.; SMITH, K.; LUCCHI, A.; FUA, P.; SÜSSTRUNK, S. SLIC superpixels compared to state-of-the-art superpixel methods. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 34, n. 11, p. 2274–2282, 2012.

[119] ZHANG, S.; LI, J.; LI, H.-C.; DENG, C.; PLAZA, A. Spectral-Spatial Weighted Sparse Regression for Hyperspectral Image Unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, 2018.

[120] ECHES, O.; DOBIGEON, N.; TOURNERET, J.-Y. Enhancing hyperspectral image unmixing with spatial correlations. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 49, n. 11, p. 4239–4247, 2011.

[121] ZORTEA, M.; PLAZA, A. Spatial preprocessing for endmember extraction. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 47, n. 8, p. 2679–2693, 2009.

[122] TORRES-MADRONERO, M. C.; VELEZ-REYES, M. Integrating spatial information in unsupervised unmixing of hyperspectral imagery using multiscale representation. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 7, n. 6, p. 1985–1993, 2014.

[123] ZYMNIS, A.; KIM, S.-J.; SKAF, J.; PARENTE, M.; BOYD, S. Hyperspectral image unmixing via alternating projected subgradients. In: IEEE. CONFERENCE Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, 2007. ACSSC 2007. [S.l.: s.n.], 2007. P. 1164–1168.

[124] CHEN, J.; RICHARD, C.; HONEINE, P. Nonlinear Estimation of Material Abundances in Hyperspectral Images With ℓ_1-Norm Spatial Regularization. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 52, n. 5, p. 2654–2665, 2014.

[125] CHEN, P.; NELSON, J. D.; TOURNERET, J.-Y. Toward a Sparse Bayesian Markov Random Field Approach to Hyperspectral Unmixing and Classification. **IEEE Transactions on Image Processing**, IEEE, v. 26, n. 1, p. 426–438, 2017.

[126] ALTMANN, Y.; MCLAUGHLIN, S.; HERO, A. Robust linear spectral unmixing using anomaly detection. **IEEE Transactions on Computational Imaging**, IEEE, v. 1, n. 2, p. 74–85, 2015.

[127] AFONSO, M. V.; BIOUCAS-DIAS, J. M.; FIGUEIREDO, M. A. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. **IEEE Transactions on Image Processing**, IEEE, v. 20, n. 3, p. 681–695, 2011.

[128] CHAMBOLLE, A.; TAN, P.; VAITER, S. Accelerated alternating descent methods for Dykstra-like problems. **Journal of Mathematical Imaging and Vision**, Springer, v. 59, n. 3, p. 481–497, 2017.

[129] CHAMBOLLE, A.; POCK, T. A first-order primal-dual algorithm for convex problems with applications to imaging. **Journal of mathematical imaging and vision**, Springer, v. 40, n. 1, p. 120–145, 2011.

[130] XU, Y.; YIN, W. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. **SIAM Journal on imaging sciences**, SIAM, v. 6, n. 3, p. 1758–1789, 2013.

[131] NASCIMENTO, J. M. P.; BIOUCAS-DIAS, J. M. Vertex Component Analysis: A fast algorithm to unmix hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, v. 43, n. 4, p. 898–910, Apr. 2005. ISSN 0196-2892. DOI: 10.1109/TGRS.2005.844293.

[132] KOZINTSEV, B. **Computations with Gaussian random fields**. 1999. PhD thesis – University of Maryland.

[133] IORDACHE, M.-D.; BIOUCAS-DIAS, J. M.; PLAZA, A. Total variation spatial regularization for sparse hyperspectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 50, n. 11, p. 4484–4502, 2012.

[134] BOYD, S.; PARIKH, N.; CHU, E.; PELEATO, B.; ECKSTEIN, J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. **Found. Trends Mach. Learn.**, Now Publishers Inc., Hanover, MA, USA, v. 3, n. 1, p. 1–122, Jan. 2011. ISSN 1935-8237. DOI: 10.1561/2200000016.

[135] HANSEN, P. C.; NAGY, J. G.; O'LEARY, D. P. **Deblurring images: matrices, spectra, and filtering**. [S.l.]: SIAM, 2006.

[136] CLARK, R. N.; SWAYZE, G. A.; LIVO, K. E.; KOKALY, R. F.; SUTLEY, S. J.; DALTON, J. B.; MCDOUGAL, R. R.; GENT, C. A. Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems. **Journal of Geophysical Research: Planets**, Wiley Online Library, v. 108, E12, 2003. doi: 10.1029/2002JE001847.

[137] HALIMI, A.; BIOUCAS-DIAS, J.; DOBIGEON, N.; BULLER, G. S.; MCLAUGHLIN, S. Fast hyperspectral unmixing in presence of nonlinearity or mismodelling effects. **IEEE Transactions Computational Imaging**, v. 3, n. 2, p. 146–159, Apr. 2017.

[138] HONG, D.; YOKOYA, N.; CHANUSSOT, J.; ZHU, X. X. An augmented linear mixing model to address spectral variability for hyperspectral unmixing. **IEEE Transactions on Image Processing**, IEEE, v. 28, n. 4, p. 1923–1938, 2019.

[139] TAO, M.; YUAN, X. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. **SIAM Journal on Optimization**, SIAM, v. 21, n. 1, p. 57–81, 2011.

[140] BOUSSE, M.; DEBALS, O.; DE LATHAUWER, L. A tensor-based method for large-scale blind source separation using segmentation. **IEEE Transactions on Signal Processing**, IEEE, v. 65, n. 2, p. 346–358, 2017.

[141] SIDIROPOULOS, N. D.; DE LATHAUWER, L.; FU, X.; HUANG, K.; PAPALEXAKIS, E. E.; FALOUTSOS, C. Tensor decomposition for signal processing and machine learning. **IEEE Transactions on Signal Processing**, IEEE, v. 65, n. 13, p. 3551–3582, 2017.

[142] NG, M. K.-P.; YUAN, Q.; YAN, L.; SUN, J. An Adaptive Weighted Tensor Completion Method for the Recovery of Remote Sensing Images with Missing Data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 55, n. 6, p. 3367–3381, 2017.

[143] ZHANG, X.; WEN, G.; DAI, W. A tensor decomposition-based anomaly detection algorithm for hyperspectral image. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 10, p. 5801–5820, 2016.

[144] GUO, X.; HUANG, X.; ZHANG, L.; ZHANG, L.; PLAZA, A.; BENEDIKTSSON, J. A. Support tensor machines for classification of hyperspectral remote sensing imagery. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 6, p. 3248–3264, 2016.

[145] YANG, S.; WANG, M.; LI, P.; JIN, L.; WU, B.; JIAO, L. Compressive hyperspectral imaging via sparse tensor and nonlinear compressed sensing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 53, n. 11, p. 5943–5957, 2015.

[146] ZHANG, L.; ZHANG, L.; TAO, D.; HUANG, X. Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 51, n. 1, p. 242–256, 2013.

[147] VEGANZONES, M. A.; COHEN, J. E.; FARIAS, R. C.; CHANUSSOT, J.; COMON, P. Nonnegative tensor CP decomposition of hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 5, p. 2577–2588, 2016.

[148] QIAN, Y.; XIONG, F.; ZENG, S.; ZHOU, J.; TANG, Y. Y. Matrix-Vector Nonnegative Tensor Factorization for Blind Unmixing of Hyperspectral Imagery. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 55, n. 3, p. 1776–1792, 2017.

[149] DE LATHAUWER, L.; DE MOOR, B.; VANDEWALLE, J. A multilinear singular value decomposition. **SIAM journal on Matrix Analysis and Applications**, SIAM, v. 21, n. 4, p. 1253–1278, 2000.

[150] CICHOCKI, A.; MANDIC, D.; DE LATHAUWER, L.; ZHOU, G.; ZHAO, Q.; CAIAFA, C.; PHAN, H. A. Tensor decompositions for signal processing applications: From two-way to multiway component analysis. **IEEE Signal Processing Magazine**, IEEE, v. 32, n. 2, p. 145–163, 2015.

[151] HÅSTAD, J. Tensor rank is NP-complete. **Journal of Algorithms**, Elsevier, v. 11, n. 4, p. 644–654, 1990.

[152] GANDY, S.; RECHT, B.; YAMADA, I. Tensor completion and low-n-rank tensor recovery via convex optimization. **Inverse Problems**, IOP Publishing, v. 27, n. 2, p. 025010, 2011.

[153] FAN, H.; CHEN, Y.; GUO, Y.; ZHANG, H.; KUANG, G. Hyperspectral image restoration using low-rank tensor recovery. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 10, n. 10, p. 4589–4604, 2017.

[154] MEI, S.; HOU, J.; CHEN, J.; CHAU, L. P.; DU, Q. Simultaneous Spatial and Spectral Low-Rank Representation of Hyperspectral Images for Classification. **IEEE Transactions on Geoscience and Remote Sensing**, v. 56, n. 5, p. 2872–2886, May 2018. ISSN 0196-2892. DOI: 10.1109/TGRS.2017.2785359.

[155] SORBER, L.; VAN BAREL, M.; DE LATHAUWER, L. Optimization-based algorithms for tensor decompositions: Canonical polyadic decomposition, decomposition in rank-(Lr,Lr,1) terms, and a new generalization. **SIAM Journal on Optimization**, SIAM, v. 23, n. 2, p. 695–720, 2013.

[156] IMBIRIBA, T.; BERMUDEZ, J. C. M.; RICHARD, C. Band Selection for Nonlinear Unmixing of Hyperspectral Images as a Maximal Clique Problem. **IEEE Transactions on Image Processing**, v. 26, n. 5, p. 2179–2191, May 2017. ISSN 1057-7149. DOI: 10.1109/TIP.2017.2676344.

[157] YOKOTA, T.; LEE, N.; CICHOCKI, A. Robust multilinear tensor rank estimation using higher order singular value decomposition and information criteria. **IEEE Transactions on Signal Processing**, IEEE, v. 65, n. 5, p. 1196–1206.

[158] OKUJENI, A.; LINDEN, S. van der; TITS, L.; SOMERS, B.; HOSTERT, P. Support vector regression and synthetically mixed training data for quantifying urban land cover. **Remote Sensing of Environment**, Elsevier, v. 137, n. 1, p. 184–197, 2013.

[159] WANG, L.; LIU, D.; WANG, Q.; WANG, Y. Spectral unmixing model based on least squares support vector machine with unmixing residue constraints. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 10, n. 6, p. 1592–1596, 2013.

[160] MIANJI, F. A.; ZHANG, Y. SVM-based unmixing-to-classification conversion for hyperspectral abundance quantification. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 49, n. 11, p. 4318–4327, 2011.

[161] PARENTE, M.; GEMP, I.; DURUGKAR, I. Unmixing in the presence of nuisances with deep generative models. In: IEEE. PROC. IEEE International Geoscience and Remote Sensing Symposium (IGARSS). [S.l.: s.n.], 2017. P. 5189–5192.

[162] WANG, L.; JIA, X. Integration of soft and hard classifications using extended support vector machines. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 6, n. 3, p. 543–547, 2009.

[163] GU, Y.; WANG, S.; JIA, X. Spectral unmixing in multiple-kernel Hilbert space for hyperspectral imagery. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 51, n. 7, p. 3968–3981, 2013.

[164] LI, X.; JIA, X.; WANG, L.; ZHAO, K. On spectral unmixing resolution using extended support vector machines. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 53, n. 9, p. 4985–4996, 2015.

[165] VAN DER MAATEN, L.; POSTMA, E.; VAN DEN HERIK, J. Dimensionality reduction: A comparative review. **Journal of Machine Learning Research**, v. 10, p. 1–41, 2009.

[166] PALSSON, B.; SIGURDSSON, J.; SVEINSSON, J. R.; ULFARSSON, M. O. Hyperspectral Unmixing Using a Neural Network Autoencoder. **IEEE Access**, IEEE, v. 6, p. 25646–25656, 2018.

[167] GUO, R.; WANG, W.; QI, H. Hyperspectral image unmixing using autoencoder cascade. In: IEEE. 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). Tokyo, Japan: [s.n.], June 2015. P. 1–4.

[168] SU, Y.; MARINONI, A.; LI, J.; PLAZA, J.; GAMBA, P. Stacked Nonnegative Sparse Autoencoders for Robust Hyperspectral Unmixing. **IEEE Geoscience and Remote Sensing Letters**, IEEE, v. 15, n. 9, p. 1427–1431, 2018.

[169] QU, Y.; QI, H. uDAS: An Untied Denoising Autoencoder With Sparsity for Spectral Unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 57, n. 3, p. 1698–1712, Mar. 2019.

[170] OZKAN, S.; KAYA, B.; AKAR, G. B. Endnet: Sparse autoencoder network for endmember extraction and hyperspectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 57, n. 1, p. 482–496, 2018.

[171] SU, Y.; LI, J.; PLAZA, A.; MARINONI, A.; GAMBA, P.; CHAKRAVORTTY, S. DAEN: Deep Autoencoder Networks for Hyperspectral Unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 57, n. 7, p. 4309–4321, 2019.

[172] KINGMA, D. P.; WELLING, M. Auto-encoding variational bayes. In: PROCEEDINGS of the International Conference on Learning Representations (ICLR). [S.l.: s.n.], 2014. Available at: `https://arxiv.org/pdf/1312.6114.pdf`.

[173] GOODFELLOW, I.; POUGET-ABADIE, J.; MIRZA, M.; XU, B.; WARDE-FARLEY, D.; OZAIR, S.; COURVILLE, A.; BENGIO, Y. Generative adversarial nets. In: ADVANCES in neural information processing systems. [S.l.: s.n.], 2014. P. 2672–2680.

[174] NEAL, R. M. Annealed importance sampling. **Statistics and computing**, Springer, v. 11, n. 2, p. 125–139, 2001.

[175] ARJOVSKY, M.; CHINTALA, S.; BOTTOU, L. Wasserstein generative adversarial networks. In: INTERNATIONAL Conference on Machine Learning. [S.l.: s.n.], 2017. P. 214–223.

[176] DOERSCH, C. Tutorial on variational autoencoders. **arXiv preprint arXiv:1606.05908**, 2016.

[177] SHAH, V.; HEGDE, C. Solving linear inverse problems using GAN priors: An algorithm with provable guarantees. In: IEEE. PROC. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP). [S.l.: s.n.], 2018. P. 4609–4613.

[178] BORA, A.; JALAL, A.; PRICE, E.; DIMAKIS, A. Compressed sensing using generative models. In: PROC. Int. Conf. Machine Learning. [S.l.: s.n.], 2017. P. 537–546.

[179] ANIRUDH, R.; THIAGARAJAN, J. J.; KAILKHURA, B.; BREMER, T. An Unsupervised Approach to Solving Inverse Problems using Generative Adversarial Networks. **arXiv preprint arXiv:1805.07281**, 2018.

[180] ASIM, M.; SHAMSHAD, F.; AHMED, A. Solving bilinear inverse problems using deep generative priors. **arXiv preprint arXiv:1802.04073**, 2018.

[181] IORDACHE, M.-D.; BIOUCAS-DIAS, J. M.; PLAZA, A. Sparse unmixing of hyperspectral data. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 49, n. 6, p. 2014–2039, 2011.

[182] SOMERS, B.; ZORTEA, M.; PLAZA, A.; ASNER, G. P. Automated extraction of image-based endmember bundles for improved spectral unmixing. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 5, n. 2, p. 396–408, 2012.

[183] SOMERS, B.; TITS, L.; ROBERTS, D.; WETHERLEY, E. Endmember library approaches to resolve spectral mixing problems in remotely sensed data: Potential, challenges, and applications. In: DATA Handling in Science and Technology. [S.l.]: Elsevier, 2016. v. 30. P. 551–577.

[184] UEZATO, T.; MURPHY, R. J.; MELKUMYAN, A.; CHLINGARYAN, A. A novel endmember bundle extraction and clustering approach for capturing

spectral variability within endmember classes. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 11, p. 6712–6731, 2016.

[185] ANDREOU, C.; ROGGE, D.; MÜLLER, R. A New Approach for Endmember Extraction and Clustering Addressing Inter-and Intra-Class Variability via Multiscaled-Band Partitioning. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 9, n. 9, p. 4215–4231, 2016.

[186] NOCEDAL, J.; WRIGHT, S. **Numerical Optimization**. [S.l.]: Springer Science & Business Media, 2006.

[187] HAND, P.; VORONINSKI, V. Global Guarantees for Enforcing Deep Generative Priors by Empirical Risk. In: BUBECK, S.; PERCHET, V.; RIGOLLET, P. (Eds.). **Proceedings of the 31st Conference On Learning Theory**. [S.l.]: PMLR, June 2018. (Proceedings of Machine Learning Research), p. 970–978.

[188] LEWIS, A. S.; OVERTON, M. L. Nonsmooth optimization via quasi-Newton methods. **Mathematical Programming**, Springer, v. 141, n. 1-2, p. 135–163, 2013.

[189] CURTIS, F. E.; QUE, X. A quasi-Newton algorithm for nonconvex, nonsmooth optimization with global convergence guarantees. **Mathematical Programming Computation**, Springer, v. 7, n. 4, p. 399–428, 2015.

[190] VAN DER MAATEN, L. **Matlab toolbox for dimensionality reduction**. [S.l.: s.n.], 2008. Available at: `https://lvdmaaten.github.io/drtoolbox/code/drtoolbox.tar.gz`.

[191] KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. In: INTERNATIONAL Conference on Learning Representations (ICLR). [S.l.: s.n.], 2015. Available at: `https://arxiv.org/pdf/1412.6980.pdf`.

[192] PEDDLE, D. R.; HALL, F. G.; LEDREW, E. F. Spectral mixture analysis and geometric-optical reflectance modeling of boreal forest biophysical structure. **Remote Sensing of Environment**, Elsevier, v. 67, n. 3, p. 288–297, 1999.

[193] DENNISON, P. E.; CHAROENSIRI, K.; ROBERTS, D. A.; PETERSON, S. H.; GREEN, R. O. Wildfire temperature and land cover modeling using hyperspectral data. **Remote Sensing of Environment**, Elsevier, v. 100, n. 2, p. 212–222, 2006.

[194] ANTONIOU, A.; STORKEY, A.; EDWARDS, H. Data augmentation generative adversarial networks. **arXiv preprint arXiv:1711.04340**, 2017.

[195] HEYLEN, R.; ZARE, A.; GADER, P.; SCHEUNDERS, P. Hyperspectral unmixing with endmember variability via alternating angle minimization. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 54, n. 8, p. 4983–4993, 2016.

[196] RAY, T. W.; MURRAY, B. C. Nonlinear spectral mixing in desert vegetation. **Remote Sensing of Environment**, Elsevier, v. 55, n. 1, p. 59–64, 1996. ISSN 0034-4257.

[197] MUSTARD, J. F.; PIETERS, C. M. Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra. **Journal of Geophysical Research: Solid Earth**, Wiley Online Library, v. 94, B10, p. 13619–13634, 1989.

[198] GUILFOYLE, K. J.; ALTHOUSE, M. L.; CHANG, C.-I. A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 10, p. 2314–2318, 2001.

[199] ALTMANN, Y.; HALIMI, A.; DOBIGEON, N.; TOURNERET, J. Y. Supervised Nonlinear Spectral Unmixing Using a Postnonlinear Mixing Model for Hyperspectral Imagery. **IEEE Transactions on Image Processing**, v. 21, n. 6, p. 3017–3025, June 2012. ISSN 1057-7149. DOI: 10.1109/TIP.2012.2187668.

[200] LI, P.-X.; WU, B.; ZHANG, L. Abundance estimation from hyperspectral image based on probabilistic outputs of multi-class support vector machines. In: 2005 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). [S.l.: s.n.], July 2005. P. 4315–4318. DOI: 10.1109/IGARSS.2005.1525873.

[201] HEYLEN, R.; BURAZEROVIC, D.; SCHEUNDERS, P. Non-linear spectral unmixing by geodesic simplex volume maximization. **IEEE Journal of Selected Topics in Signal Processing**, v. 5, n. 3, p. 534–542, 2011.

[202] LI, X.; CUI, J.; ZHAO, L. Blind nonlinear hyperspectral unmixing based on constrained kernel nonnegative matrix factorization. **Signal, Image and Video Processing**, Springer, v. 8, n. 8, p. 1555–1567, 2012. DOI: 10.1007/s11760-012-0392-3.

[203] CHEN, J.; RICHARD, C.; HONEINE, P. Nonlinear Unmixing of Hyperspectral Data Based on a Linear-Mixture/Nonlinear-Fluctuation Model. **IEEE Transactions on Signal Processing**, v. 61, p. 480–492, Jan. 2013.

[204] CHEN, J.; RICHARD, C.; HONEINE, P. Nonlinear Estimation of Material Abundances in Hyperspectral Images With ℓ_1-Norm Spatial Regularization. **IEEE Transactions on Geoscience and Remote Sensing**, v. 52, n. 5, p. 2654–2665, May 2014. ISSN 0196-2892. DOI: 10.1109/TGRS.2013.2264392.

[205] ALTMANN, Y.; DOBIGEON, N.; MCLAUGHLIN, S.; TOURNERET, J.-Y. Nonlinear Spectral Unmixing of Hyperspectral Images Using Gaussian Processes. **IEEE Transactions on Signal Processing**, v. 61, p. 2442–2453, May 2013.

[206] SUYKENS, J. A. K.; GESTEL, T. V.; BRABANTER, J. D.; MOOR, B. D.; VANDEWALLE, J. **Least Squares Support Vector Machines**. Singapore: World Scientific, 2002.

[207] HALL, P.; TITTERINGTON, D. M. Common structure of techniques for choosing smoothing parameters in regression problems. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 49, n. 2, p. 184–198, 1987.

[208] THOMPSON, A. M.; BROWN, J. C.; KAY, J. W.; TITTERINGTON, D. M. A study of methods of choosing the smoothing parameter in image restoration by regularization. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 4, p. 326–339, 1991.

[209] GALATSANOS, N. P.; KATSAGGELOS, A. K. Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation. **IEEE Transactions on image processing**, v. 1, n. 3, p. 322–336, 1992.

[210] HEINZ WERNER ENGL MARTIN HANKE, A. N. **Regularization of Inverse Problems (Mathematics and Its Applications)**. 1. ed. [S.l.]: Springer, 1996. (Mathematics and Its Applications). ISBN 9780792341574,0792341570.

[211] BOYD, S.; VANDENBERGHE, L. **Convex optimization**. [S.l.]: Cambridge university press, 2004.

[212] SCHÖLKOPF, B.; SMOLA, A. J. **Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond**. [S.l.]: The MIT Press, 2001.

[213] TUY, H.; TUAN, H. D. Generalized S-lemma and strong duality in nonconvex quadratic programming. **Journal of Global Optimization**, Springer, v. 56, n. 3, p. 1045–1072, 2013.

[214] SOMERS, B.; ASNER, G. P. Invasive species mapping in Hawaiian rainforests using multi-temporal Hyperion spaceborne imaging spectroscopy. **IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing**, IEEE, v. 6, n. 2, p. 351–359, 2013.

[215] LIPPITT, C. L.; STOW, D. A.; ROBERTS, D. A.; COULTER, L. L. Multidate MESMA for monitoring vegetation growth forms in southern California

shrublands. **International Journal of Remote Sensing**, Taylor & Francis, v. 39, n. 3, p. 655–683, 2018.

[216] SOMERS, B.; ASNER, G. P. Multi-temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests. **Remote Sensing of Environment**, Elsevier, v. 136, n. 1, p. 14–27, 2013.

[217] HENROT, S.; CHANUSSOT, J.; JUTTEN, C. Dynamical spectral unmixing of multitemporal hyperspectral images. **IEEE Transactions on Image Processing**, IEEE, v. 25, n. 7, p. 3219–3232, 2016.

[218] THOUVENIN, P.-A.; DOBIGEON, N.; TOURNERET, J.-Y. Online unmixing of multitemporal hyperspectral images accounting for spectral variability. **IEEE Transactions on Image Processing**, IEEE, v. 25, n. 9, p. 3979–3990, 2016.

[219] SIGURDSSON, J.; ULFARSSON, M. O.; SVEINSSON, J. R.; BIOUCAS-DIAS, J. M. Sparse distributed multitemporal hyperspectral unmixing. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 55, n. 11, p. 6069–6084, 2017.

[220] THOUVENIN, P.-A.; DOBIGEON, N.; TOURNERET, J.-Y. A hierarchical Bayesian model accounting for endmember variability and abrupt spectral changes to unmix multitemporal hyperspectral images. **IEEE Transactions on Computational Imaging**, IEEE, v. 4, n. 1, p. 32–45, 2018.

[221] SÄRKKÄ, S. **Bayesian filtering and smoothing**. [S.l.]: Cambridge University Press, 2013. v. 3.

[222] MHENNI, R.; BOURGUIGNON, S.; NININ, J.; SCHMIDT, F. Spectral unmixing with sparsity and structuring constraints. In: PROC. 9th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing. Amsterdam, The Netherlands: [s.n.], 2018. P. 1–5.

[223] ECKARDT, A.; HORACK, J.; LEHMANN, F.; KRUTZ, D.; DRESCHER, J.; WHORTON, M.; SOUTULLO, M. DESIS (DLR Earth sensing imaging spectrometer for the ISS-MUSES platform). In: IEEE. PROC. IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Milan, Italy: [s.n.], 2015. P. 1457–1459.

[224] KAUFMANN, H.; SEGL, K.; CHABRILLAT, S.; HOFER, S.; STUFFLER, T.; MUELLER, A.; RICHTER, R.; SCHREIER, G.; HAYDN, R.; BACH, H. EnMAP a hyperspectral sensor for environmental mapping and analysis. In: IEEE. PROC. IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Denver, Colorado, USA: [s.n.], 2006. P. 1617–1619.

[225] EMELYANOVA, I. V.; MCVICAR, T. R.; VAN NIEL, T. G.; LI, L. T.; DIJK, A. I. van. Assessing the accuracy of blending Landsat–MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection. **Remote Sensing of Environment**, Elsevier, v. 133, p. 193–209, 2013.

[226] SIMÕES, M.; BIOUCAS-DIAS, J.; ALMEIDA, L. B.; CHANUSSOT, J. A convex formulation for hyperspectral image superresolution via subspace-based regularization. **IEEE Transactions on Geoscience and Remote Sensing**, IEEE, v. 53, n. 6, p. 3373–3388, 2015.

[227] KANATSOULIS, C. I.; FU, X.; SIDIROPOULOS, N. D.; MA, W.-K. Hyperspectral super-resolution: A coupled tensor factorization approach. **IEEE Transactions on Signal Processing**, IEEE, v. 66, n. 24, p. 6503–6517, 2018.

[228] PRÉVOST, C.; USEVICH, K.; COMON, P.; BRIE, D. Hyperspectral super-resolution with coupled Tucker approximation: Recoverability and SVD-based algorithms. **IEEE Transactions on Signal Processing**, IEEE, v. 68, p. 931–946, 2020.