



HAL
open science

Méthodes Bayésiennes pour le démixage d'images hyperspectrales

Olivier Eches

► **To cite this version:**

Olivier Eches. Méthodes Bayésiennes pour le démixage d'images hyperspectrales. Traitement du signal et de l'image [eess.SP]. Institut National Polytechnique de Toulouse - INPT, 2010. Français. NNT : 2010INPT0067 . tel-04274734

HAL Id: tel-04274734

<https://theses.hal.science/tel-04274734>

Submitted on 8 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par *l'Institut National Polytechnique de Toulouse (INP de Toulouse)*
Discipline ou spécialité : *Signal, Image, Acoustique et Optimisation*

Présentée et soutenue par *Olivier Eches*
Le 14 Octobre 2010

Titre : *Méthodes Bayésiennes pour le démélange d'images hyperspectrales*

JURY

M. Jocelyn Chanussot.....Professeur E.N.S.E.3 Grenoble.....Rapporteurs
M. Hichem Snoussi.....MdC Université de Troyes.....
M. Jon Atli Benediktsson.....Professeur University of Iceland Reykjavik.....
M. Grégoire Mercier.....Professeur E.N.S.T. Rennes.....Examineurs
Mme Corinne Mailhes.....Professeur E.N.S.E.E.I.H.T. Toulouse.....
M. Jean-Denis Durou.....MdC Université Paul Sabatier Toulouse.....
M. Jean-Yves Tournet.....Professeur E.N.S.E.E.I.H.T. Toulouse.....Encadrants
M. Nicolas Dobigeon.....Mdc E.N.S.E.E.I.H.T. Toulouse.....

Ecole doctorale : *Mathématiques Informatique Télécommunication de Toulouse*

Unité de recherche : *Institut de Recherche en Informatique de Toulouse*

Directeur(s) de Thèse : *Jean-Yves Tournet et Nicolas Dobigeon*

Rapporteurs :

Remerciements

Au cours de ces trois années de thèse, j'ai pu faire la connaissance et échanger avec tellement de personnes, dans et en dehors du cadre de mon travail, qui m'est impossible de me souvenir de tout le monde !

Je commence donc tout naturellement par Jean-Yves Tourneret, mon directeur de thèse qui m'a proposé cette thèse au cours de l'année 2007. Sa grande confiance en mes capacités, ses sages conseils, son expérience, sa patience face à mon caractère ont fortement contribué à la réussite de cette thèse. Je ne peux également oublier Nicolas Dobigeon qui a co-encadré ma thèse à partir de la deuxième année, sans lui, sans son courage et sa détermination (notamment à relire et corriger les fautes de typographie mes articles) rien n'aurait été possible. Je n'ai pas également oublié les moments passés en leur compagnie en conférence à Reykjavik, notamment quand ils m'ont comparé avec un certain personnage porte-poisse d'une assez vieille production cinématographique française.

Je tiens à remercier également Corinne Mailhes pour avoir présidé le jury de soutenance, en plus de m'avoir encadré au cours de la première année. Puis je remercie également messieurs Jocelyn Chanussot et Hichem Snoussi qui ont accepté de rapporter ma thèse. Remercions tout particulièrement Jon Atli Benediktsson qui a accepté de rapporter ma thèse et d'assister à ma soutenance malgré la distance et la barrière de la langue et qui de plus a accepté de m'accueillir en post-doc. Bien entendu, je n'oublie pas messieurs Jean-Denis Durou et Grégoire Mercier pour leur présence au sein du jury.

Au fil de ces trois années, l'IRIT est devenu comme une deuxième maison pour moi. C'est dans ses locaux neufs que je termine ma thèse, après un déménagement au cours de la troisième année depuis un bureau aux teintes orangées et mal éclairé vers un bureau plus vaste à la vue superbe sur la ville de Toulouse mais - disons le franchement - à la décoration du plus mauvais goût. Commençons donc par remercier les secrétaires, Sylvie Armengaud et Sylvie Eichen, très compétentes et toujours prête à aider. Puis l'équipe du service informatique, notamment les deux premières années Stéphane Jene-saisplustonnomdefamilledésolé, remplacé ensuite par Frédéric Peyré. L'équipe Signal et Communications dans laquelle j'ai effectué ma thèse n'est pas en reste ! Merci à Nathalie pour sa gentillesse, à Marie pour les quelques discussions que nous avons eu. Merci à Martial pour ses conseils pédagogiques et sa bonne humeur. Je tiens également à remercier Marco Lops, récemment arrivé d'Italie, pour sa sympathie lors de cette dernière année de thèse. Remercions également Bruno Escrig, ancien responsable de stage de mastère et d'agréable compagnie, et Marie-Laure Boucheret, sympathique motarde. Bien entendu, je n'oublie pas les thésards qui se sont succédés dans mon bureau : Jérôme, Fares, Sokchenda, Cécile et les deux petits derniers Abderrahim et Yohann. Une petite pensée pour Jean-Rémi et Henri qui travaillaient leur thèse dans notre bureau voisin avant le déménagement. Je salue également Emmanuel Christophe qui a suivi avec intérêt mon travail au tout début de ma thèse.

Je tiens à remercier mes amis pour les moments passés ensemble : Vincent, Chahinez, Lionel, Sébastien ainsi que tout ceux que j'ai oublié ! J'ai une très forte pensée pour ma famille, notamment mes parents, mon frère et ma grand-mère.

Enfin, je tiens à remercier plus que tout Aurore sans quoi rien de tout cela ne serait possible...

Résumé

L'imagerie hyperspectrale est très largement employée en télédétection pour diverses applications, dans le domaine civil comme dans le domaine militaire. Une image hyperspectrale est le résultat de l'acquisition d'une seule scène observée dans plusieurs longueurs d'ondes. Par conséquent, chacun des pixels constituant cette image est représenté par un vecteur de mesures (généralement des réflectances) appelé spectre. Une étape majeure dans l'analyse des données hyperspectrales consiste à identifier les composants macroscopiques (signatures) présents dans la région observée et leurs proportions correspondantes (abondances). Les dernières techniques développées pour ces analyses ne modélisent pas correctement ces images. En effet, habituellement ces techniques supposent l'existence de pixels purs dans l'image, c'est-à-dire des pixels constitué d'un seul matériau pur. Or, un pixel est rarement constitué d'éléments purs distincts l'un de l'autre. Ainsi, les estimations basées sur ces modèles peuvent tout à fait s'avérer bien loin de la réalité. Le but de cette étude est de proposer de nouveaux algorithmes d'estimation à l'aide d'un modèle plus adapté aux propriétés intrinsèques des images hyperspectrales. Les paramètres inconnus du modèle sont ainsi déduits dans un cadre Bayésien. L'utilisation de méthodes de Monte Carlo par Chaînes de Markov (MCMC) permet de surmonter les difficultés liées aux calculs complexes de ces méthodes d'estimation.

Abstract

Hyperspectral imagery has been widely used in remote sensing for various civilian and military applications. A hyperspectral image is acquired when a same scene is observed at different wavelengths. Consequently, each pixel of such image is represented as a vector of measurements (reflectances) called spectrum. One major step in the analysis of hyperspectral data consists of identifying the macroscopic components (signatures) that are present in the sensed scene and the corresponding proportions (concentrations). The latest techniques developed for this analysis do not properly model these images. Indeed, these techniques usually assume the existence of pure pixels in the image, i.e. pixels containing a single pure material. However, a pixel is rarely composed of pure spectrally elements, distinct from each other. Thus, such models could lead to weak estimation performance. The aim of this thesis is to propose new estimation algorithms with the help of a model that is better suited to the intrinsic properties of hyperspectral images. The unknown model parameters are then inferred within a Bayesian framework. The use of Markov Chain Monte Carlo (MCMC) methods allows one to overcome the difficulties related to the computational complexity of these inference methods.

Abréviations et notations

Abréviations

ACI	Analyse en Composantes Indépendantes
ACP	Analyse en Composantes Principales
DAG	graphe acyclique orienté (<i>directed acyclic graph</i>)
EM	<i>expectation-maximisation</i>
EQM	erreur quadratique moyenne
FCLS	<i>Fully Constrained Least Squares</i>
MCMC	méthodes de Monte Carlo par chaînes de Markov
MEE	Méthodes d'Extraction d'Endmembers
MML	Modèle de Mélange Linéaire
MNF	<i>Minimum Noise Fraction</i>
MVC-NMF	<i>Minimum Volume Constrained - Nonnegative Matrix Factorization</i>
MAP	maximum <i>a posteriori</i>
MMSE	estimateur qui minimise l'erreur quadratique moyenne (<i>minimum mean square error</i>)
MRF	champ de Markov aléatoire (<i>Markov random field</i>)

NCM	<i>Normal Compositional Model</i>
NN-ICA	<i>Non Negative - Independent Component Analysis</i>
pdf	fonction densité de probabilité (<i>probability density function</i>)
PPI	<i>Pixel Purity Index</i>
RJ-MCMC	méthodes de Monte Carlo par chaînes de Markov à sauts réversibles (<i>reversible jump MCMC</i>)
SNR	rapport signal sur bruit (<i>Signal to Noise Ratio</i>)
VCA	<i>Vertex Component Analysis</i>

Notations standard

\propto	proportionnel à
\ll	très inférieur à
\gg	très supérieur à
$\mathbf{1}_{\mathbb{E}}(\mathbf{x})$	fonction indicatrice définie sur l'espace \mathbb{E} :
	$\mathbf{1}_{\mathbb{E}}(\mathbf{x}) = \begin{cases} 1, & \text{si } \mathbf{x} \in \mathbb{E}; \\ 0, & \text{sinon.} \end{cases}$
$\Gamma(\cdot)$	fonction Gamma,
\mathbb{R}	ensemble des réels
\mathbb{R}^+	ensemble des réels positifs

Notations matricielles

x	scalaire
\mathbf{x}	vecteur
\cdot^T	transposé

\mathbf{X}	matrice
$ \mathbf{X} $	déterminant de la matrice \mathbf{X}
\mathbf{X}_i	i -ième colonne de la matrice \mathbf{X}
\mathbf{X}_{-i}	matrice \mathbf{X} privée de sa i -ième colonne
\mathbf{I}_j	matrice identité de taille $j \times j$
$\mathbf{0}_j$	vecteur nul de \mathbb{R}^j
$\ \mathbf{x}\ ^2$	norme L^2 standard $\ \mathbf{x}\ ^2 = \mathbf{x}^T \mathbf{x}$

Notations

r	indice du composant pur
R	nombre de composants purs
l	indice de la bande spectrale
L	nombre de bandes spectrales
K	nombre de classes
k	indice de classe
P	nombre de pixels dans l'image hyperspectrale
p	indice de pixel

Algorithmes d'estimation statistique

\hat{x}	estimation de x
$E[x y]$	espérance de x conditionnellement à y
$\text{var}[x y]$	variance de x conditionnellement à y
$P[A]$	probabilité de l'événement A
\sim	distribué suivant
N_{MC}	longueur totale de la chaîne de Markov

N_{bi} longueur de la période de chauffage de la chaîne de Markov
 N_r nombre d'itérations de la chaîne de Markov construite pour réaliser les estimations
 $x^{(k)}$ k -ième échantillon du processus $(x^{(t)})_{t=1, \dots, N_{\text{MC}}}$

Distributions de probabilité usuelles

$\mathcal{U}_{\mathbb{E}}$ loi uniforme sur l'ensemble \mathbb{E}
 $\mathcal{Be}(a, b)$ loi Beta de paramètres a et b
 $\mathcal{N}(\mathbf{x}, \Sigma)$ loi normale de vecteur moyenne \mathbf{x} et de matrice de covariance Σ
 $\mathcal{IG}(a, b)$ loi inverse Gamma de paramètres a et b
(voir [RC04, p. 581] pour la définition des paramètres a et b)
 $\mathcal{G}(a, b)$ loi Gamma de paramètres a et b
(voir [RC04, p. 582] pour la définition des paramètres a et b)
 $\mathcal{D}_p(\mathbf{x})$ loi de Dirichlet de vecteur paramètre $\mathbf{x} = [x_1, \dots, x_p]$

Table des matières

Remerciements	iii
Résumé	v
Abstract	vii
Abréviations et notations	ix
Introduction	1
0.1 L'imagerie hyperspectrale	1
0.2 Modèle de mélange linéaire	4
0.3 Etat de l'art des méthodes de démixage linéaire spectral	7
0.4 Contributions	12
1 Démixage spectral à l'aide de modèles NCM	15
1.1 Introduction	15
1.2 Modèle Bayésien hiérarchique	19
1.3 Echantillonneur de Gibbs hybride	22
1.4 Extension aux cas d'endmembers de variances différentes	25
1.5 Algorithme MCMC	29
1.6 Résultats de simulations sur des données synthétiques	31
1.7 Démixage spectral d'une image AVIRIS	39

1.8	Conclusion	46
2	Estimation du nombre de composants purs	47
2.1	Introduction	47
2.2	Modèle Bayésien hiérarchique étendu	49
2.3	Algorithme MCMC à sauts réversibles	51
2.4	Simulations sur des données synthétiques	58
2.5	Démélange spectrale de données réelles	64
2.6	Conclusion	69
3	Démélange spectral avec corrélations spatiales	73
3.1	Introduction	73
3.2	Formulation du problème	75
3.3	Modèle Bayésien hiérarchique	80
3.4	Échantillonneur de Gibbs hybride	85
3.5	Résultats de simulations sur données synthétiques	88
3.6	Résultats de simulations sur une image AVIRIS	94
3.7	Conclusions	96
4	Conclusions et perspectives	99
4.1	Conclusions	99
4.2	Perspectives	101
A	Calcul de la loi conditionnelle $f(w_r^2 w_{-r}, Y, A, M)$	105
B	Calcul des probabilités d'acceptation	107
C	Démélange spectral avec contraintes spatiales	111
C.1	Modèle Bayésien hiérarchique	111
C.2	Algorithme de Gibbs hybride	115
C.3	Résultats de simulations sur des données synthétiques	117

C.4 Résultats de simulations sur une image AVIRIS	122
Liste des publications	125

Table des figures

1	Exemple de cube hyperspectral (Moffet Field, Californie, Etats-Unis).	2
2	Illustration d'un cube hyperspectral dont 3 pixels ont été extraits (avec leurs spectres de réflectances correspondants), issue de [Pla09].	3
3	Modèles de mélange (a) linéaire (b) non-linéaire (figures issues de [KM02]).	5
4	Représentation en deux bandes des observations non bruitées (points gris) contenues dans le simplexe (en rouge).	7
1.1	Représentation en deux bandes des endmembers présents dans l'image (en rouge), des résultats du N-FINDR (en bleu) et des pixels (en noir).	17
1.2	Vraisemblance pour (de haut en bas) $P = 1, 2, 9$ pixels en fonction de (w_1^2, w_2^2) . (a) Vue 3D. (b) Vue de dessus.	27
1.3	Spectres des endmembers : béton de construction (trait continu), herbe verte (trait pointillés).	32
1.4	Loi conditionnelle estimée des abondances $[a_1, a_2]^T$	32
1.5	Lois conditionnelles estimées de la variance w^2	33
1.6	Estimations MAP (croix) et écarts-types (barres verticales) des composantes de \mathbf{a}^+ en fonction du SNR.	34
1.7	Lois conditionnelles estimées des variances pour $P = 3$ pixels.	35
1.8	Lois conditionnelles estimées des abondances pour chaque pixel (haut : pixel 1, centre : pixel 2, bas : pixel 3).	36

1.9	Données hyperspectrales réelles : image de Moffett field acquise par AVIRIS en 1997 (à gauche) et la région d'intérêt représenté en couleurs réelles (à droite).	40
1.10	Les spectres des $R = 3$ endmembers obtenus par l'algorithme N-FINDR. . .	40
1.11	En haut : cartes d'abondances estimées par le MML Bayésien (tiré de [DTC08]). Au milieu : cartes d'abondances estimées par l'algorithme FCLS [HC01]. En bas : cartes d'abondances estimées par l'algorithme proposé (pixel noir (resp. blanc) signifie absence (resp. présence) du matériau). . .	42
1.12	En haut : cartes d'abondances estimées par l'algorithme MVC-NMF. En bas : cartes d'abondances estimées par l'algorithme NN-ICA (pixel noir (resp. blanc) signifie absence (resp. présence) du matériau).	43
1.13	En haut : les spectres des $R = 3$ endmembers obtenus par l'algorithme MVC-NMF [MQ07]. En bas : les spectres des $R = 3$ endmembers obtenus par l'algorithme NN-ICA [PO04].	43
1.14	Lois conditionnelles de w^2 pour les pixels $\#(35, 43)$ (à gauche) et $\#(43, 35)$ (à droite) estimées par l'algorithme proposé.	44
1.15	Régions d'eau, de sol et de végétation analysées pour les probabilités de présence.	44
2.1	Spectres des endmembers de la bibliothèque.	58
2.2	Lois conditionnelles estimée du nombre de composants purs R	59
2.3	Lois conditionnelles des abondances (à gauche) et de la variance w^2 (à droite) estimées.	59
2.4	Estimations MAP (croix) et écarts-types (barres verticales) de R en fonction du SNR.	60
2.5	Histogrammes des chaînes de Markov du paramètre R générées par 50 simulations de l'algorithme proposé avec différentes tailles de blocs.	61

2.6	Les $R = 6$ spectres des endmembers extraits de l'image « Moffett Field » par l'utilisation de l'algorithme PPI sur les classes trouvées par la procédure K-means.	64
2.7	Cartes d'abondances estimées par l'algorithme proposé.	65
2.8	Cartes de présences estimées par l'algorithme proposé.	66
2.9	Nombre d'endmembers estimé par l'algorithme proposé (zones obscures (resp. claires) signifiant $R = 1$ (resp. $R = 6$)).	66
2.10	Données hyperspectrales réelles : image Cuprite acquise par AVIRIS en 1997 (à gauche) et la région d'intérêt représentée en couleurs réelles (à droite).	67
2.11	Les $R = 6$ spectres des endmembers estimés par l'algorithme VCA appliqué sur l'image Cuprite.	68
2.12	Nombre de matériaux purs estimés par l'algorithme proposé (zones obscures (resp. claires) signifiant $R = 1$ (resp. $R = 6$)).	68
2.13	Cartes d'abondances de la région d'intérêt estimées par l'algorithme proposé.	69
2.14	Haut : les 3 spectres des endmembers de la bibliothèque récupérés par l'algorithme VCA depuis l'image Cuprite (en rouge), comparés avec les signatures spectrales extraites de la bibliothèque de l'USGS (en noir). Milieu : cartes d'abondances estimées par l'algorithme Bayésien à sauts réversibles. Bas : cartes d'abondances estimées par l'algorithme Bayésien non supervisé étudié dans [DMC ⁺ 09].	70
3.1	Structures de voisinages 4-connexes (à gauche) et 8-connexes (à droite). Le pixel considéré apparaît comme étant un disque noir alors que ses voisins apparaissent en blanc.	77
3.2	Images synthétiques générés à partir d'un modèle de Potts-Markov avec (de gauche à droite) $\beta = 0.8, 1.4, 2$	80
3.3	DAG pour les lois <i>a priori</i> (les paramètres fixés apparaissent dans les cases en pointillés) pour le NCM.	82

3.4	Les $R = 3$ spectres des endmembers : béton de construction (trait continu), herbe verte (trait semi-continu), terre micacée (trait pointillés).	90
3.5	A gauche : carte des étiquettes réelle. A droite : carte des étiquettes estimée par l'échantillonneur de Gibbs hybride proposé.	90
3.6	En haut : cartes d'abondances des 3 matériaux purs pour le NCM. En bas : cartes d'abondances des 3 matériaux purs estimés par l'échantillonneur de Gibbs hybride basé sur le NCM (de gauche à droite : béton de construction, herbe verte, terre grasse).	91
3.7	Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ estimés par l'algorithme de Gibbs hybride pour les $K = 3$ classes (valeurs réelles en pointillés).	93
3.8	Données hyperspectrales réelles : l'image de Moffett field acquise par AVI- RIS en 1997 (à gauche) et la région d'intérêt représenté en couleurs réelles (à droite).	94
3.9	Les spectres des $R = 3$ endmembers obtenues par l'algorithme N-FINDR.	95
3.10	Cartes d'abondances estimées par l'algorithme basé sur le NCM (de gauche à droite : végétation, eau et terre).	95
3.11	En haut : cartes d'abondances estimées par le MML Bayésien (tiré de [DTC08]). En bas : cartes d'abondances estimées par l'algorithme FCLS [HC01].	96
3.12	Cartes des étiquettes estimées par l'algorithme basé sur le NCM.	97
3.13	Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ (de gauche à droite : végétation, eau et terre) estimés par l'algorithme de Gibbs hybride pour les $K = 4$ classes.	98
C.1	DAG pour les lois <i>a priori</i> (les paramètres fixés apparaissent dans les cases en pointillés) pour le MML.	113
C.2	A gauche : carte des étiquettes réelle. A droite : carte des étiquettes estimée par l'échantillonneur de Gibbs hybride basé sur le MML.	117

C.3	En haut : cartes d'abondances des 3 matériaux purs pour le MML. En bas : cartes d'abondances des 3 matériaux purs estimés par l'échantillonneur de Gibbs hybride basé sur le MML (de gauche à droite : béton de construction, herbe verte, terre grasse).	120
C.4	Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ estimés par l'algorithme de Gibbs hybride basé sur le MML pour (de haut en bas) les $K = 3$ classes (valeurs réelles en pointillés).	121
C.5	Carte des étiquettes estimées par l'algorithme basé sur le MML.	122
C.6	Cartes d'abondances estimées par l'algorithme basé sur le MML (de gauche à droite : végétation, eau et terre).	122
C.7	Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ (de gauche à droite : végétation, eau et terre) estimés par l'algorithme de Gibbs hybride pour les $K = 4$ classes.	123

Liste des tableaux

1.1	EQM globale du vecteur d'abondance pour le NCM avec une unique variance et avec des variances distinctes.	35
1.2	EQM globales de chaque coefficient d'abondance pour différents algorithmes de démélages ($\times 10^{-3}$).	38
1.3	Erreurs de reconstruction pour les algorithmes NCM Bayésien, MML Bayésien et FCLS.	38
1.4	EQM globaux de chaque coefficient d'abondance pour une image synthétique contenant des pixels purs ($\times 10^{-3}$).	39
1.5	Probabilité de présence pour chaque endmember.	45
1.6	Estimation MMSE de w_r^2 ($r = 1, \dots, R$).	45
2.1	Nombre estimé de composants purs (en %) obtenu par la méthode MCMC à sauts réversibles et par d'autres algorithmes de sélection de modèle avec $w^2 = 1.10^{-2}$ and $w^2 = 2.10^{-5}$	63
3.1	Moyennes et variances réelles et estimées des abondances dans chacune des classes.	91
3.2	EQMs globales pour chaque composante des vecteurs d'abondances. . . .	94
C.1	Moyennes et variances réelles et estimées des abondances dans chacune des classes.	119
C.2	EQMs globales pour chaque composante des vecteurs d'abondances. . . .	122

Liste des algorithmes

1.1	Algorithme de Metropolis-within-Gibbs pour la simulation des abondances.	24
1.2	Échantillonneur de Gibbs hybride pour le démélange spectral utilisant le NCM	24
1.3	Démélange spectrale employant le NCM avec différentes variances	31
2.1	Échantillonneur de Metropolis-within-Gibbs hybride pour le démélange spectral semi-supervisé selon le NCM	52
2.2	Mouvement de naissance	54
2.3	Mouvement de mort	54
2.4	Mouvement d'échange	55
3.1	Simulation des MRF	79
3.2	Échantillonneur de Gibbs hybride pour le démélange spectral avec contraintes spatiales pour la modèle NCM	89
C.1	Échantillonneur de Gibbs hybride pour le démélange spectral avec contraintes spatiales pour le modèle MML	118

Introduction

0.1 L'imagerie hyperspectrale

À l'origine les images de la Terre prises par des capteurs spatiaux ou aéroportés furent monochromes, notamment dans le domaine militaire. Il est ensuite survenu l'idée qu'en observant la même scène dans plusieurs bandes spectrales, c'est-à-dire à différentes longueurs d'onde, l'exploitation des données en serait considérablement améliorée. Ainsi, un très grand nombre de capteurs multispectraux ont été développés, le premier étant Landsat au début des années 70. C'est dans les années 80 que Alexander F.H. Goetz du Jet Propulsion Laboratory (JPL) de la NASA a mis en œuvre le concept d'imagerie hyperspectrale [GVSR85]. Dans une telle image, la scène est observée dans plusieurs centaines de bandes spectrales étroites et contiguës (au lieu de trois ou quatre longueurs d'onde pour les images multispectrales). L'image obtenue n'est plus à deux dimensions mais à trois dimensions, la dernière étant la longueur d'onde. On parle alors de « cube » hyperspectral pour décrire l'ensemble des données échantillonnées spatialement et spectralement (voir la figure 1).

Parmi les imageurs hyperspectraux embarqués les plus performants, on peut citer :

- le capteur en 224 bandes AVIRIS du JPL (NASA),
- le capteur en 384 bandes OMEGA embarqué à bord de Mars Express (ESA),
- le capteur en 128 bandes HyMAP (Hyperspectral MAPping) de l'entreprise Integrated Spectronics,

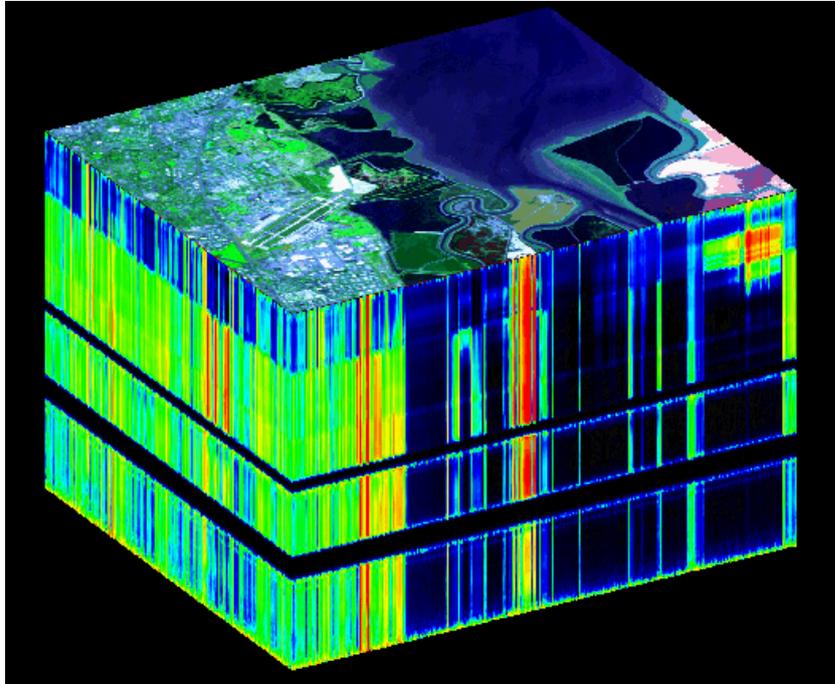


FIGURE 1 – Exemple de cube hyperspectral (Moffet Field, Californie, Etats-Unis).

De nombreuses applications, civiles ou militaires, font appel à l'imagerie hyperspectrale. Depuis plusieurs années, ce type d'imagerie fait l'objet d'études avancées aussi bien au niveau national qu'au niveau international par différents organismes tels que l'ESA ou la NASA ou par de grands groupes industriels. Ainsi, on peut regrouper entre autre les applications liées à :

- l'agriculture [CSB⁺02] : surveillance des cultures, qualité des sols,...
- l'écologie [Jus09, HSY⁺05] : surveillance des forêts, détection de pollutions,...
- la géologie [Kru98] : cartographie de minéraux, détection de minéraux rares,...
- la défense [Pla06] : détection de cibles à caractères militaires (mines, avions),...

La principale source d'éclairage pour les capteurs hyperspectraux embarqués est le Soleil. Le spectre électromagnétique de l'énergie émise par notre étoile (que l'on nommera « spectre solaire ») est modifiée en terme d'intensité et de distribution spectrale par l'atmosphère. Cette énergie est ensuite réémise, transmise et/ou absorbée par les matériaux

éclairés et retransmis par l'atmosphère, subissant de nouvelles modifications. L'énergie reçue par le capteur hyperspectral est caractérisée par un spectre appelé « spectre de luminance », qui est le spectre de l'énergie émise par le Soleil modifiée par deux traversées de l'atmosphère et par la réflectance du pixel observé [MMS03]. Ce dernier s'agissant de notre signal d'intérêt peut dès lors se déduire par correction et inversion du spectre de radiance. Les imageurs hyperspectraux existants réalisent le meilleur compromis entre haute résolution spatiale et haute résolution spectrale. Par exemple, le capteur AVIRIS possède une résolution spatiale de 20 mètres (à une altitude de 20 km) pour une résolution spectrale de l'ordre de 10 nm [Gre98].

Ainsi, chaque pixel d'une image hyperspectrale, représenté par un vecteur de mesure correspondant au spectre de réflectance, est caractéristique des éléments le composant (figure 2 issue de [Pla09]).

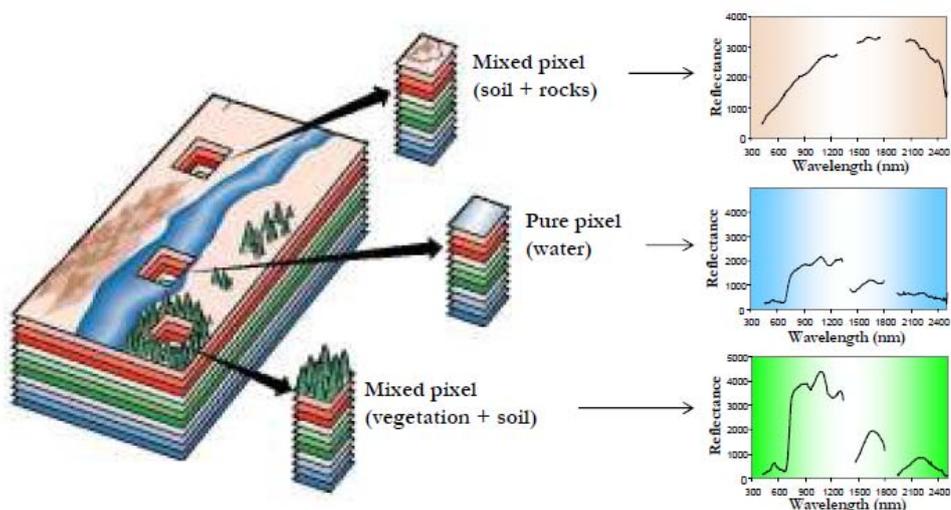


FIGURE 2 – Illustration d'un cube hyperspectral dont 3 pixels ont été extraits (avec leurs spectres de réflectances correspondants), issue de [Pla09].

Dans le cas d'analyses minéralogiques, il est possible par exemple de déterminer très précisément les zones riches en un minéral particulier. Les raies d'absorption de certains

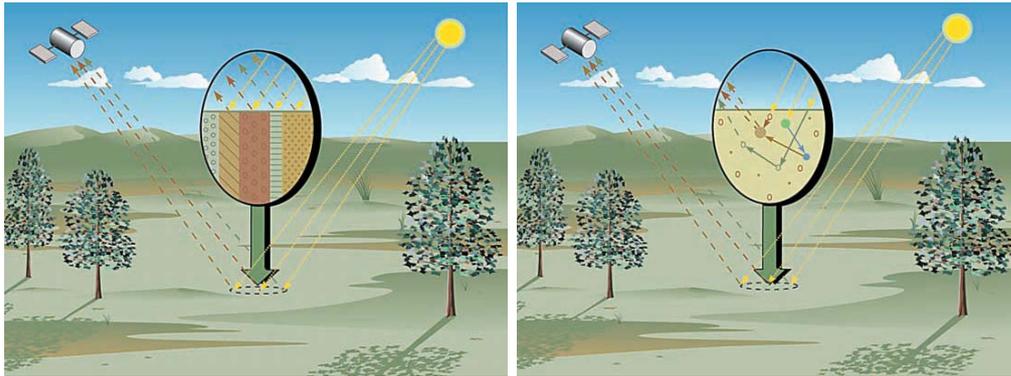
minéraux étant très fines, la résolution de l’outil hyperspectral est donc indispensable pour les séparer. De plus, cet outil peut s’avérer beaucoup moins coûteux qu’une campagne de prélèvements au sol, comme sur Mars. Ainsi, grâce au capteur OMEGA, la présence d’eau et de dioxyde de carbone gelé sur la surface de Mars peut être déterminée avec exactitude [MHS⁺08, SMD09]. Dans un tout autre domaine, les militaires se sont intéressés à la détection de cibles sous-pixelliques, c’est-à-dire de tailles inférieures à celle du pixel [MSS01]. En imagerie hyperspectrale, un pixel n’est pas composé à 100% d’éléments spectralement purs ce qui peut poser de nombreux problèmes. Par exemple, comment classifier un pixel constitué d’un mélange de 2 éléments purs différents? Ou comment détecter des cibles dont la taille est inférieure à la résolution spatiale du capteur (cible sub-pixellique)?

C’est dans ce contexte que le démixage spectral intervient. Largement utilisé pour l’analyse d’images hyperspectrales, le démixage spectral est fondé sur l’hypothèse suivante [KM02] : le spectre d’un pixel d’une image hyperspectrale est le résultat du mélange des spectres de tous les constituants macroscopiques purs, appelés pôles de mélange ou *endmembers*, et de leurs quantités relatives, nommées *abondances*. Même à haute résolution spatiale il est toujours possible d’observer des pixels mélangés. Les phénomènes physiques à l’origine de ce mélange peuvent être représentés à l’aide de plusieurs modèles.

0.2 Modèle de mélange linéaire

Deux modèles de mélange différents ont principalement été étudiés dans la littérature. Le premier décrit un pixel de l’image comme la combinaison linéaire des *endmembers*, pondérés par leurs abondances respectives [SM79]. Ce modèle convient parfaitement lorsque les constituants purs sont spatialement disjoints, à la manière d’un damier par exemple. Le second modèle, introduit par Hapke [Hap81], considère ce même mélange

comme un mélange intime de constituants purs (à la manière de grains de sable de compositions différentes par exemple) supposant ainsi une combinaison non linéaire des pôles de mélange (voir figure 3 issue de [KM02]).



(a) Modèle linéaire

(b) Modèle non-linéaire

FIGURE 3 – Modèles de mélange (a) linéaire (b) non-linéaire (figures issues de [KM02]).

La question de savoir si l'un ou l'autre de ces deux modèles est plus adapté à un quelconque problème est loin d'être résolue. Pour être rigoureux jusqu'à l'extrême, le modèle linéaire ne devrait jamais être utilisé, le cas où les constituants purs sont distincts dans un pixel n'ayant jamais été rencontré. Cependant, comme démontré dans [JSTGA83], le mélange non linéaire peut être linéarisé pour une analyse beaucoup moins complexe. C'est principalement dû à cette simplification que le modèle linéaire a été le plus utilisé pour le démixage spectral.

Mathématiquement, le modèle de mélange linéaire (MML) suppose que le L -spectre $\mathbf{y} = [y_1, \dots, y_L]^T$ d'un pixel mélangé p résulte d'une combinaison linéaire de R spectres purs \mathbf{m}_r , $r = 1, \dots, R$, entaché d'un bruit additif

$$\mathbf{y} = \sum_{r=1}^R \mathbf{m}_r a_r + \mathbf{n}, \quad (1)$$

où

- $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$ représente le spectre du r -ième matériau (le r -ième end-member),

- a_r est la proportion du r -ième matériau dans le pixel (ou r -ième abondance),
- R est le nombre de matériaux purs présents dans toute la scène observée,
- L est le nombre de bandes spectrales disponibles pour l'image,
- $\mathbf{n} = [n_1, \dots, n_L]^T$ est une séquence de bruits additifs. Une hypothèse classique, que nous reprenons également dans la suite, consiste à prendre le bruit \mathbf{n} comme étant blanc et Gaussien [CXAP98, MSS01]. Le bruit est alors noté $\mathbf{n} \sim \mathcal{N}_L(\mathbf{0}_L, s^2 \mathbf{I}_L)$, dont les composantes n_l sont classiquement supposées indépendantes et identiquement distribuées suivant une loi normale centrée de variance s^2 (\mathbf{I}_L est la matrice identité de dimension $L \times L$). Cependant, il est également possible de considérer ce bruit comme coloré [DTH08].

Comme les coefficients a_r ($r = 1, \dots, R$) représentent des proportions, le vecteur $\mathbf{a}^+ = [a_1, \dots, a_R]^T$ est soumis aux contraintes de positivité et d'additivité, soit :

$$\begin{cases} a_r \geq 0, \forall r = 1, \dots, R, \\ \sum_{r=1}^R a_r = 1. \end{cases} \quad (2)$$

Le MML peut se généraliser à toute l'image hyperspectrale. Dans ce cas, en considérant un nombre P de pixels, on obtient :

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N}, \quad (3)$$

où $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_P]$ matrice de taille $L \times P$, $\mathbf{M} = [\mathbf{m}_1, \dots, \mathbf{m}_R]$ matrice de taille $L \times R$, $\mathbf{A} = [\mathbf{a}_1^+, \dots, \mathbf{a}_P^+]$ matrice de taille $R \times P$, $\mathbf{N} = [\mathbf{n}_1, \dots, \mathbf{n}_P]$ matrice de taille $L \times P$ et \mathbf{a}_p^+ est le vecteur d'abondance du p ième pixel.

Une propriété intéressante concerne la géométrie du problème, découlant de la forme du modèle et des contraintes. En effet, dans un sous-espace de \mathbb{R}^L , l'ensemble \mathcal{S} tel que

$$\mathcal{S} = \left\{ \mathbf{x} \mid \mathbf{x} = \sum_{r=1}^R \mathbf{m}_r a_r, \sum_{r=1}^R a_r = 1, a_r \geq 0, \forall r = 1, \dots, R \right\}$$

définit un simplexe de dimension $R - 1$. Ainsi, l'ensemble des observations non bruitées $\mathbf{Y} - \mathbf{N}$ est contenu dans ce simplexe, comme l'illustre la figure 4. Les R sommets du

simplexe correspondent aux différents spectres des endmembers m_1, \dots, m_R présents dans l'image. Ainsi dans cette représentation, plus un pixel est proche d'un des sommets, plus il est pur. Le but du démixage est alors de retrouver les sommets de ce simplexe.

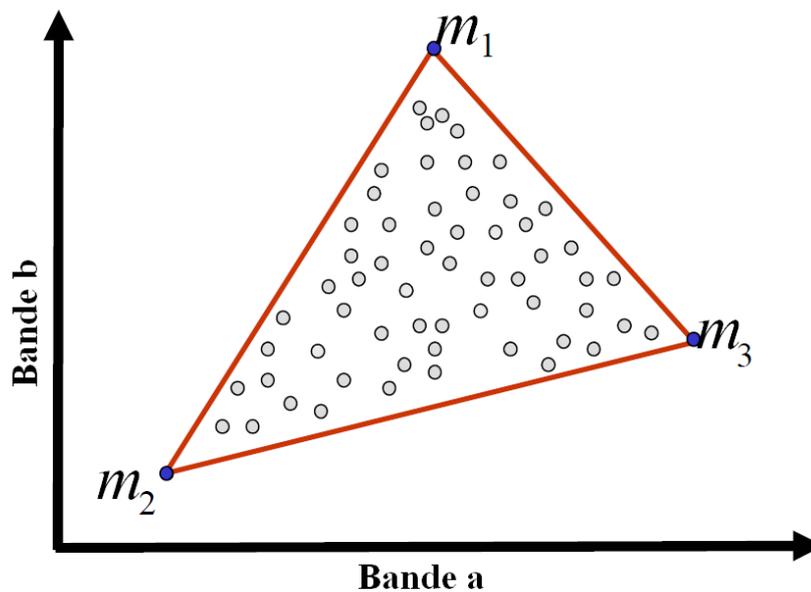


FIGURE 4 – Représentation en deux bandes des observations non bruitées (points gris) contenues dans le simplexe (en rouge).

0.3 Etat de l'art des méthodes de démixage linéaire spectral

Les méthodes de démixage linéaire spectral développées jusqu'à présent ont pour but de retrouver les spectres des endmembers présents dans l'image et leurs abondances correspondantes. Toutefois, une phase de pré-traitement doit précéder les étapes d'estimation. Cette phase a pour but d'augmenter la rapidité de convergence des algorithmes proposés, de réduire l'espace de travail (au vu du volume important de données) et de

filtrer le bruit. Pour ce faire, l'utilisation d'algorithmes de réduction de données est appropriée. Citons notamment l'Analyse en Composante Principale (ACP) [Jol86] projetant les données dans l'espace engendré par les axes de plus grandes variances ou l'algorithme Maximum Noise Fraction (MNF) [GBSC94] cherchant la projection maximisant le rapport signal sur bruit (SNR). Après cette étape de réduction de données, on poursuit par l'étape de détermination des spectres des endmembers puis celle d'estimation des abondances, ou étape d'*inversion*. Les deux étapes se font généralement dans cet ordre, l'étape d'inversion nécessitant la connaissance préalable des spectres des endmembers. Il est aussi possible de combiner les deux étapes en une seule, estimant ainsi conjointement les abondances et les endmembers.

0.3.1 Méthodes d'extraction d'endmembers

Le point clé du démelange spectral linéaire réside dans la recherche de signatures spectrales pures. Ces dernières peuvent être soit extraites de bibliothèques spectrales (obtenues par des mesures en laboratoire), soit obtenues à partir de l'image observée. Il est à noter que, tout en étant conscient du risque de possibles différences avec les données observées (dues aux conditions d'acquisition), l'utilisation de bibliothèques spectrales s'avère très avantageuse pour le démelange spectral lorsqu'elles sont à disposition. C'est ce qui sera abordé dans le deuxième chapitre.

La connaissance du nombre de composants purs présents dans l'image est généralement obligatoire pour extraire les endmembers. Ce nombre peut être estimé en effectuant une ACP sur les vecteurs d'observations de l'image [KM02]. Considérant les valeurs propres triées λ_i ($i = 1, \dots, L$) de la matrice de covariance des observations (on rappelle que L est la dimension des vecteurs d'observations), la quantité η définie par

$$\eta = \frac{\sum_{i=1}^{L_p} \lambda_i}{\sum_{i=1}^L \lambda_i},$$

où $L_p < L$ peut être utilisée pour rechercher la dimension du sous-espace contenant l'information pertinente liée aux données. Cette dimension est par exemple déterminée

en fixant un seuil de qualité choisi tel que η soit suffisamment proche de 1 ($\eta = 90\%$ ou $\eta = 99\%$). Comme la dimension du simplexe est par définition égale à $R - 1$, on déduit alors facilement le nombre d'endmembers $R = L_p + 1$. Ce nombre des endmembers R peut également être estimé en utilisant par exemple les méthodes HFC ou NWHFC décrites dans [CD04] dont le principe général est de comparer les valeurs propres des matrices de corrélation et de covariances des données en effectuant un test de Neyman-Pearson. À chaque échec du test, le nombre d'endmembers est augmenté de 1. La méthode décrite dans [LCDC09] est basée sur ce principe et permet d'estimer le nombre de pôles de mélanges sans calculer la probabilité de fausse alarme et pour des SNR très faibles.

Le lecteur intéressé pourra consulter [PMPP04] et les références citées pour une liste exhaustive des méthodes d'extraction d'endmembers (MEE) basées sur le modèle linéaire. Parmi celles ci, on peut citer les célèbres algorithmes N-FINDR de Winter [Win99] et Pixel Purity Index (PPI) de Boardman [Boa93]. Ces algorithmes cherchent à retrouver le simplexe d'intérêt à partir de projections sur des espaces appropriés (PPI) ou bien en calculant itérativement le simplexe de plus grand volume (N-FINDR). Néanmoins, le PPI nécessite un nombre élevé d'itérations alors que le N-FINDR dépend fortement des pixels initiaux sélectionnés pour le calcul du volume du premier simplexe. Un autre algorithme, appelé Vertex Component Analysis (VCA) [NB05], exploite la projection dans des sous-espaces orthogonaux afin de retrouver directement les endmembers. Ces trois algorithmes se basent sur l'hypothèse de l'existence d'au moins un pixel pur dans l'image. D'autres approches sont possibles comme la méthode Automated Morphological Endmember Extraction (AMEE) intégrant des informations spatiales par l'utilisation d'opérateurs morphologiques [PMPP02], ou bien la méthode Orthogonal Subspace Projection (OSP) [HC94] qui suit le même principe que la méthode VCA sans toutefois faire l'hypothèse que les endmembers soient les sommets d'un simplexe.

0.3.2 Méthodes d'inversion

Cette étape cherche à estimer au sein de chaque pixel les abondances de chaque endmember retrouvé lors de l'étape précédente. Deux approches peuvent être envisagées pour estimer ces coefficients.

La première approche pour estimer les abondances consiste à minimiser une fonction de coût approprié. Le vecteur d'abondances \mathbf{a}_p^+ étant toujours soumis aux contraintes de positivité et d'additivité (2), il s'agit donc d'un problème d'optimisation sous contraintes d'égalité (additivité) et d'inégalité (positivité). Le recours à des méthodes numériques a été envisagé pour obtenir des solutions optimales tout en incorporant de telles contraintes. C'est le cas notamment pour l'algorithme Fully Constrained Least Squares (FCLS) [HC01]. Cet algorithme itératif cherche à minimiser au sens des moindres carrés le critère suivant (pour un pixel donné p) :

$$J(\mathbf{a}_p^+) = \|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p^+\|^2,$$

sous la contrainte de positivité. La contrainte d'égalité est prise en compte en ajoutant une observation fictive égale à $d = \sum_{i=1}^R a_i + b$ où b est un bruit additif. Cependant, l'algorithme FCLS bien que très rapide ne converge pas nécessairement vers un minimum global. L'algorithme Scaled Gradient Method (SGM) récemment développé dans [TDTL09] propose une alternative à l'algorithme FCLS en se basant sur une méthode de descente de gradient.

La deuxième approche pour estimer ces coefficients d'abondance fait appel aux méthodes Bayésiennes, comme dans [DTC08]. Le principe de l'analyse Bayésienne est basé sur la relation de Bayes, qui en supposant la variance du bruit s_p^2 connue, s'écrit :

$$f(\mathbf{a}_p^+|\mathbf{y}_p) = \frac{f(\mathbf{y}_p|\mathbf{a}_p^+)f(\mathbf{a}_p^+)}{f(\mathbf{y}_p)}.$$

Cette formule permet de calculer la loi *a posteriori* du paramètre d'intérêt (ici \mathbf{a}_p^+) en fonction de la vraisemblance des observations $f(\mathbf{y}_p|\mathbf{a}_p^+)$ et de la loi a priori de \mathbf{a}_p^+ . Après calculs, plusieurs estimateurs peuvent être employés pour approcher \mathbf{a}_p^+ :

- Estimateur MMSE : $\hat{\mathbf{a}}_{\text{MMSE}}^+ = E[\mathbf{a}_p^+ | \mathbf{y}_p]$
- Estimateur MAP : $\hat{\mathbf{a}}_{\text{MAP}}^+ = \arg \max_{\mathbf{a}_p^+} f(\mathbf{a}_p^+ | \mathbf{y}_p)$

0.3.3 Méthodes conjointes

Comme expliqué plus haut, les endmembers et les abondances peuvent être estimés conjointement en une seule étape. Formellement, il s'agit de résoudre le problème de séparation de sources aveugles en estimant \mathbf{M} et \mathbf{A} à partir des observations bruitées \mathbf{Y} , sous les contraintes (2) et également de positivité pour chaque coefficient des spectres de \mathbf{M} . Une méthode communément admise pour résoudre les problèmes de séparation de sources aveugles consiste à faire une Analyse en Composantes Indépendantes (ACI) des données [CJH91]. Une telle méthode a déjà été appliquée à des problèmes de démixage spectral dans [BGC97]. Cependant, cette méthode se base sur l'indépendance des sources, ce qui n'est pas le cas pour des images hyperspectrales puisque les coefficients d'abondance sont liés par les contraintes (2). Il semble en effet que, comme illustré dans [DA05] et [NBD05], cette méthode n'est pas adaptée au problème de démixage des images hyperspectrales. Elle peut cependant donner des résultats assez approximatifs pour l'estimation des endmembers [ND04],[NB05].

Des approches alternatives basées sur des techniques nommées Nonnegative Matrix Factorization (NMF) [PT94] ont été récemment proposées dans la littérature. Ces approches consistent de manière générale à factoriser une matrice en un produit de deux matrices dont les coefficients sont positifs. L'algorithme proposé dans [PT94] met alternativement à jour les matrices des sources suivant l'hypothèse de non-négativité. L'algorithme NMF a été appliqué avec succès pour l'identification de constituants chimiques dans [SDB⁺04] ou pour le démixage d'images hyperspectrales dans [VPPP06]. Cependant, cet algorithme ne tient pas compte de la contrainte d'additivité. Plus récemment, une autre méthode itérative d'optimisation, appelé Minimum Volume Constrained - NMF (MVC-NMF) [MQ07], consiste à rajouter à l'algorithme NMF cette contrainte en

imposant de retrouver le simplexe de volume le plus petit possible. Toutefois, les performances de cet algorithme ont tendance à se dégrader lorsque le niveau de bruit augmente. Une approche Bayésienne sous contraintes de volume minimal a été également proposée dans [ASL10]. Un autre algorithme nommé Spatially Smoothness Constraint - NMF (SSC-NMF) a été décrit dans [AZB07]. Les auteurs proposent de rajouter en plus des contraintes de positivité et d'additivité une contrainte de similarité entre les abondances des pixels voisins dans l'algorithme NMF. Dans [BKL⁺04], un algorithme itératif nommé ICE (iterated constrained endmembers) se propose de minimiser un critère de pénalité, compte tenu des deux contraintes. Comme démontré dans [ND07], les résultats de ICE dépendent fortement des paramètres choisis pour l'algorithme.

Une autre manière de résoudre ce problème d'optimisation consiste à faire appel aux méthodes Bayésiennes. De telles méthodes ont déjà été utilisées pour résoudre les problèmes de séparation de sources aveugles (par exemple dans [Row02]). Les contraintes peuvent être directement incorporées dans le modèle grâce à la formulation Bayésienne. Ainsi l'approche développée dans [PMS⁺00] propose de formaliser par cette méthode le problème d'estimation conjointe des endmembers et des coefficients d'abondance. Elle repose sur un modèle autorégressif *ad hoc* du spectre des endmembers. Ce modèle peut conduire à des spectres estimés négatifs ce qui est clairement un inconvénient majeur de la méthode. La méthode mise en œuvre dans [DMC⁺09] impose une contrainte de positivité pour les spectres en plus des contraintes sur les abondances. Les méthodes Bayésiennes pour la résolution des problèmes de séparation de sources nécessitent une étape de prétraitement. C'est ainsi qu'une ACP est utilisée pour l'algorithme proposé dans [DMC⁺09] ou bien une ACI pour la méthode décrite dans [MHS⁺08].

0.4 Contributions

Nous avons mis en évidence dans ce qui précède que le démixage spectral constitue le pivot de l'analyse d'images hyperspectrales. Nous avons également expliqué l'intérêt

d'utiliser un modèle linéaire sur lequel est basé la quasi-totalité des algorithmes de dé-mélange. Un bref état de l'art a résumé les deux étapes de dé-mélange (extraction des endmembers et inversion) tout en n'oubliant pas de parler des méthodes conjointes de dé-mélange linéaire spectral.

Cette thèse se focalise sur les problèmes d'inversion pour le dé-mélange linéaire spectral par une approche Bayésienne. L'originalité de notre travail repose sur l'utilisation d'un modèle statistique différent du MML en vue de résoudre le problème d'inversion. Ce modèle, proposé par Eismann et Stein dans [ES07] et appelé *Normal Compositional Model* (NCM), considère les spectres du mélange comme des vecteurs aléatoires, contrairement au MML où ces vecteurs sont déterministes. Nous avons montré que ce modèle est très bien adapté à des images hyperspectrales contenant peu de pixels associés à des matériaux purs [EDMT10]. A partir de cette remarque, nous avons développé plusieurs algorithmes Bayésiens associés à ce modèle NCM :

- Une approche supervisée permet d'estimer les coefficients d'abondance en prenant en compte les erreurs d'estimation des méthodes d'extraction d'endmembers. Les endmembers étant supposés aléatoires, on peut considérer que les R vecteurs du modèle partagent la même variance ou possèdent des variances distinctes. Après établissement d'un modèle hiérarchique Bayésien et calcul de la loi *a posteriori*, la forme complexe de cette loi nous impose d'utiliser un algorithme d'échantillonnage sophistiqué afin de pouvoir calculer les estimateurs Bayésiens classiques tels que les estimateurs MAP ou MMSE. C'est pour cette raison que nous faisons appel aux méthodes de Monte Carlo par Chaînes de Markov (MCMC) [RC04]. L'algorithme de Gibbs hybride, développé dans cette étude, va générer des échantillons suivant la loi *a posteriori*. Ces échantillons vont alors être employés pour donner une estimation de nos coefficients d'abondance [EDMT10].
- Une approche semi-supervisée donne une estimation du nombre de composants purs R présents dans un pixel. Cette non-connaissance de ce paramètre implique que la nature des spectres présents dans l'image est inconnue. Notre modèle suppose alors

que ces spectres appartiennent à une bibliothèque spectrale connue. Le paramètre R ajustant la dimension du vecteur d'abondances, il est nécessaire de faire appel à une méthode d'ajustement de dimension. C'est pourquoi l'algorithme de Gibbs hybride développé pour la méthode supervisée sera adapté à cette approche et conjointement utilisé avec une méthode à sauts réversibles [Gre95]. Cet algorithme original va permettre ainsi d'estimer R , le vecteur d'abondances et de donner la nature des spectres impliqués dans le mélange [EDT10b].

- Une généralisation de ce modèle permet de prendre en compte l'existence des corrélations entre les pixels voisins de l'image. Plus précisément, la méthode développée va faire appel aux champs de Markov aléatoires. Cet outil, très utilisé dans la communauté du traitement de l'image, nécessite tout d'abord de définir des étiquettes pour tous les pixels de l'image. Les champs de Markov admettent que la valeur d'une étiquette d'un pixel dépend de la valeur des pixels d'un voisinage donné. La valeur de l'étiquette d'un pixel donné est fonction d'une distribution de probabilité de type Gibbs [Bes74]. Ces étiquettes ont été fixées en fonction des moyennes et variances des vecteurs d'abondances d'un certain groupe de pixels, cherchant ainsi à représenter des zones homogènes. Pour intégrer ces champs de Markov aux algorithmes déjà proposés, quelques modifications ont dû être entreprises. La plus importante concerne la reparamétrisation des coefficients d'abondance (sous les contraintes de positivité et d'additivité) afin que leurs distributions *a priori* soient plus adaptées aux champs de Markov. Comme précédemment, une méthode MCMC va être employée et appliquée à l'ensemble des pixels. Les abondances des pixels, par l'intermédiaire de leur reparamétrisation, seront estimées conjointement avec les étiquettes des pixels [EDT10a].

CHAPITRE 1

Démélange spectral à l'aide de modèles NCM

Sommaire

1.1	Introduction	15
1.2	Modèle Bayésien hiérarchique	19
1.3	Echantillonneur de Gibbs hybride	22
1.4	Extension aux cas d'endmembers de variances différentes	25
1.5	Algorithme MCMC	29
1.6	Résultats de simulations sur des données synthétiques	31
1.7	Démélange spectral d'une image AVIRIS	39
1.8	Conclusion	46

1.1 Introduction

Les méthodes d'inversion dont nous avons rappelé le principe dans la partie précédente sont basées sur le modèle de mélange linéaire (MML). Nous rappelons que ces méthodes opèrent en général en mode supervisé, c'est-à-dire que les signatures des spectres des matériaux purs sont considérés comme connus (extraits par une MEE ou appartenant à une bibliothèque). Deux méthodes peuvent être envisagées pour estimer les coefficients d'abondances. Dans un premier temps il est possible de procéder par optimisation d'un critère, au sens des moindres carrés par exemple. Le plus célèbre algorithme exploitant un tel critère est connu sous le nom de Fully Constrained Least Squares (FCLS) [HC01]. La deuxième approche consiste à exploiter la formule de Bayes pour déduire une estimation des coefficients d'abondances. Une telle approche a été spécifiquement employée pour

ce problème dans [DTC08]. Il faut préciser que des méthodes MCMC ont été utilisées pour contourner le problème de la complexité de la loi jointe *a posteriori* des paramètres d'intérêt.

Le MML bien que très utilisé n'est pas exempt de tout défaut pour l'analyse d'images hyperspectrales [KM02]. En particulier, le rapport entre la variance intra-classe (intrinsèquement aux endmembers) et la variance inter-classe (entre les endmembers) peut conduire à questionner la validité de l'hypothèse d'un spectre déterministe [ES07]. De plus, comme mentionné précédemment, les MEE basées sur le MML supposent que l'image contient au moins un pixel pur. Lorsque cette hypothèse n'est pas satisfaisante, ces MEE peuvent être inefficaces. Ce problème, soulevé dans [NB05], est illustré sur la figure 1.1. La figure montre 1) les projections sur deux bandes spectrales (les deux axes les plus discriminants identifiés par une ACP) des $R = 3$ endmembers (les étoiles rouges correspondent aux sommets du triangle rouge), 2) le domaine en deux bandes contenant toutes les combinaisons linéaires des $R = 3$ endmembers (le triangle rouge) et 3) le simplexe en deux bandes estimés par l'algorithme N-FINDR utilisant les pixels (représentés en noirs). Étant donné qu'aucun pixel n'est proche des sommets du triangle rouge, l'algorithme N-FINDR estime un simplexe beaucoup plus petit (en bleu) que le réel (en rouge).

Un autre modèle statistique connu sous le nom de Normal Compositional Model (NCM) a récemment été proposé par Eismann et Stein dans [ES07]. L'utilisation du NCM permet de contourner certains problèmes inhérents au MML (décrits ci-dessus) en supposant que les pixels de l'image hyperspectrale observée sont une combinaison linéaire d'endmembers aléatoires (au lieu d'être déterministe comme dans le MML) dont les moyennes sont connues (résultant d'une MEE comme les algorithmes N-FINDR ou VCA par exemple). Ce modèle autorise les endmembers à se situer à une certaine distance des pixels observés, ce qui constitue une intéressante propriété pour le problème illustré sur la figure 1.1. En d'autres termes, une certaine « marge d'erreur » concernant l'estimation de ces signatures spectrales pures est modélisée par le caractère aléatoire des endmembers.

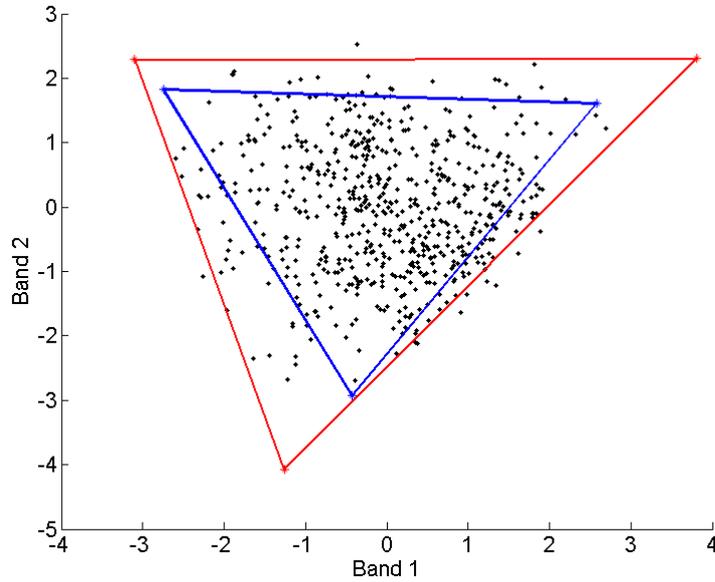


FIGURE 1.1 – Représentation en deux bandes des endmembers présents dans l’image (en rouge), des résultats du N-FINDR (en bleu) et des pixels (en noir).

Ainsi, ce modèle permet une plus grande flexibilité par rapport aux endmembers et aux pixels observés. Le NCM suppose que le spectre d’un pixel mélangé peut être écrit comme suit :

$$\mathbf{y} = \sum_{r=1}^R \mathbf{e}_r a_r \quad (1.1)$$

où les \mathbf{e}_r sont des vecteurs Gaussiens indépendants de moyennes connues. Nous rappelons également les contraintes de positivité et d’additivité du vecteur d’abondances :

$$\begin{cases} a_r \geq 0, \forall r = 1, \dots, R, \\ \sum_{r=1}^R a_r = 1. \end{cases} \quad (1.2)$$

Le bruit additif est absent de l’équation (1.1) puisque la nature aléatoire des endmembers représente déjà une certaine incertitude liée aux endmembers. Sauf mention contraire, nous admettons que la matrice de covariance de chaque vecteur \mathbf{e}_r est proportionnelle à la matrice identité. L’hypothèse simplificatrice d’une même variance dans toutes les

bandes a déjà été utilisée dans plusieurs études comme par exemple [Set96, MSS01]. Il faut cependant noter que des modèles plus sophistiquées avec des variances différentes dans chaque bande spectrale peuvent être étudiés [DTH08].

Ce chapitre décrit un estimateur Bayésien nous permettant d'estimer les abondances dans le cadre du NCM (1.1) sous les contraintes de positivité et d'additivité (1.2). L'approche adoptée définit des lois *a priori* pour les coefficients d'abondance du NCM satisfaisant aux contraintes mentionnées ci-dessus, de la même manière que dans [DTC08]. Plus précisément, ces coefficients sont munis d'une loi *a priori* uniforme sur un simplexe de dimension R . De plus, on définit une loi conjuguée inverse-Gamma pour la variance des endmembers. Les hyperparamètres du modèle sont fixés soit en utilisant les informations appropriées à disposition, soit en les estimant conjointement avec les paramètres inconnus. Une approche classique consiste à prendre des lois vagues *a priori* pour les hyperparamètres, résultant en la construction de plusieurs niveaux d'inférence Bayésienne. On parle alors de modèle Bayésien hiérarchique [RC04, p. 392]. Les paramètres et hyperparamètres d'un modèle Bayésien hiérarchique peuvent être estimés en utilisant leur loi *a posteriori*. Dans le contexte du démélange d'images hyperspectrales, la loi jointe *a posteriori* est trop complexe pour obtenir une expression analytique des estimateurs MMSE et MAP des abondances. Pour pallier cette complexité, il est possible d'employer l'algorithme Expectation-Maximization (EM) [RC04] qui a été justement appliqué au modèle NCM dans [Ste03, ES07]. Mais cet algorithme présente quelques inconvénients, notamment de convergence vers des maxima locaux [DI96, p. 259]. Ces lacunes peuvent être contournées en faisant appel aux méthodes de Monte Carlo par Chaînes de Markov (MCMC) qui permettent de générer des échantillons distribués suivant la loi *a posteriori* d'intérêt (ici la loi jointe *a posteriori* des abondances et des variances des endmembers). Ce chapitre généralise au modèle NCM l'algorithme hybride de Gibbs développé dans [DTC08].

Le chapitre est organisé comme suit. Le paragraphe 1.2 donne la loi *a posteriori* du vecteur des paramètres inconnus résultant du modèle Bayésien proposé. Le paragraphe

1.3 étudie l'échantillonneur de Gibbs hybride qui est employé pour générer des échantillons distribués suivant cette loi *a posteriori*. Les paragraphes 1.4 et 1.5 étendent le modèle proposé à un cas où les spectres des endmembers ont des variances différentes. Des résultats de simulations conduites sur des données synthétiques sont présentés au paragraphe 1.6. En particulier, des comparaisons entre les modèles Bayésiens proposés et des algorithmes de démélages classiques sont présentées. D'autres résultats obtenus sur des images réelles, font l'objet du paragraphe 1.7.

1.2 Modèle Bayésien hiérarchique

Dans ce paragraphe, la fonction de vraisemblance et les lois *a priori* choisies pour le modèle NCM sont définies. Une attention particulière est portée à la définition de la loi *a priori* des abondances respectant les contraintes de positivité et d'additivité.

1.2.1 Vraisemblance

Le modèle NCM suppose que les spectres des endmembers \mathbf{e}_r , $r = 1, \dots, R$ sont des vecteurs Gaussiens indépendant de vecteurs moyenne connus \mathbf{m}_r , $r = 1, \dots, R$. Comme nous l'avons dit en introduction, on suppose dans un premier temps que la matrice de covariance de chaque endmember s'écrit $w^2 \mathbf{I}_L$, où \mathbf{I}_L est la matrice identité de taille $L \times L$ et w^2 est la variance des endmembers dans chaque bande spectrale. Au final, la loi de probabilité pour les endmembers s'écrit :

$$\mathbf{e}_r | \mathbf{m}_r, w^2 \sim \mathcal{N}(\mathbf{m}_r, w^2 \mathbf{I}_L), \quad (1.3)$$

où $\mathcal{N}(\mathbf{m}, \mathbf{\Sigma})$ dénote la loi Gaussienne multivariée de vecteur moyenne \mathbf{m} et de matrice de covariance $\mathbf{\Sigma}$.

Partant de (1.1) et de l'indépendance *a priori* entre les différents spectres des endmembers, la vraisemblance du spectre d'un pixel observé \mathbf{y} peut s'écrire comme :

$$f(\mathbf{y} | \mathbf{a}^+, w^2) = \frac{1}{[2\pi w^2 c(\mathbf{a}^+)]^{\frac{L}{2}}} \exp\left(-\frac{\|\mathbf{y} - \mu(\mathbf{a}^+)\|^2}{2w^2 c(\mathbf{a}^+)}\right) \quad (1.4)$$

où $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ est la norme ℓ^2 standard, $\mathbf{a}^+ = [a_1, \dots, a_R]^T$ et

$$\mu(\mathbf{a}^+) = \sum_{r=1}^R \mathbf{m}_r a_r, \quad c(\mathbf{a}^+) = \sum_{r=1}^R a_r^2. \quad (1.5)$$

Il faut noter que la moyenne et la variance de cette loi de probabilité dépendent du vecteur des abondances \mathbf{a}^+ , ce qui n'était pas le cas pour le MML introduit dans [DTC08].

1.2.2 Lois *a priori* des paramètres

Abondances

Les contraintes d'additivité inhérentes au modèle de mélange permettent de réécrire le vecteur d'abondance : $\mathbf{a}^+ = [\mathbf{a}^T, a_R]^T$ où $\mathbf{a} = [a_1, \dots, a_{R-1}]^T$ et $a_R = 1 - \sum_{r=1}^{R-1} a_r$. De plus, afin de satisfaire aux contraintes de positivité, le vecteur \mathbf{a} doit appartenir au simplexe défini par :

$$\mathbb{S} = \left\{ \mathbf{a} \in \mathbb{R}^{R-1} \left| a_r \geq 0, \forall r = 1, \dots, R-1, \sum_{r=1}^{R-1} a_r \leq 1 \right. \right\}. \quad (1.6)$$

Nous proposons de choisir une loi uniforme sur \mathbb{S} comme loi *a priori* pour le vecteur partiel des abondances \mathbf{a} , étant donné que nous n'avons aucune autre information sur les abondances que la positivité et l'additivité :

$$f(\mathbf{a}) \propto \mathbf{1}_{\mathbb{S}}(\mathbf{a}), \quad (1.7)$$

où \propto signifie « proportionnel à » et $\mathbf{1}_{\mathbb{S}}(\cdot)$ est la fonction indicatrice définie sur l'ensemble \mathbb{S} :

$$\mathbf{1}_{\mathbb{S}}(\mathbf{a}) = \begin{cases} 1, & \text{si } \mathbf{a} \in \mathbb{S}; \\ 0, & \text{sinon.} \end{cases} \quad (1.8)$$

Remarque : N'importe quel coefficient d'abondance peut être retiré du vecteur \mathbf{a}^+ et pas nécessairement le dernier a_R . Pour des raisons de symétries, l'algorithme proposé dans le paragraphe suivant supprimera un coefficient d'abondance du vecteur \mathbf{a}^+ choisi aléatoirement parmi $\{1, \dots, R\}$. Pour simplifier, nous avons supposé que ce coefficient

est a_R . De même, afin d'être le plus concis possible, les notations $\boldsymbol{\mu}(\mathbf{a})$ et $c(\mathbf{a})$ seront utilisées par la suite en lieu et place des quantités $\mu(\mathbf{a}^+)$ et $c(\mathbf{a}^+)$ définies dans (1.5) où a_R a été remplacé par $1 - \sum_{r=1}^{R-1} a_r$.

Variance des spectres

La loi *a priori* de w^2 est une loi conjuguée inverse-Gamma :

$$w^2|\delta \sim \mathcal{IG}(\nu, \delta), \quad (1.9)$$

où ν et δ sont deux hyperparamètres ajustables (respectivement appelés paramètres de forme et d'échelles [RC04, p. 582]). Il est admis dans le reste du chapitre que $\nu = 1$ (hypothèse classique comme en attestent [PADF02] ou [DTD07]).

Hyperparamètre δ

Pour estimer l'hyperparamètre δ , il convient de lui choisir une loi *a priori* non informative, introduisant ainsi un deuxième niveau d'inférence Bayésienne conformément aux modèles Bayésiens hiérarchiques. Une approche naïve consisterait à choisir une loi uniforme pour modéliser cette absence d'information. Le problème est que cette loi n'est pas invariante par reparamétrisation [Dob07], ce qui nous pousse à nous interroger sur le réel caractère non informatif de cette loi. Dans [Jef46], Jeffreys propose comme loi non informative de prendre la racine carrée de l'information de Fisher du modèle [RC04, p.32]. Dans notre cas, on obtient :

$$f(\delta) = \sqrt{-\mathbb{E} \left[\frac{\partial^2 \log f(w^2|\delta)}{\partial \delta^2} \right]}. \quad (1.10)$$

Comme $f(w^2|\delta)$ a été définie au dessus, la loi de probabilité de δ peut être déterminée. Après quelques calculs, on obtient :

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta). \quad (1.11)$$

Cette loi reflète le manque de connaissance sur l'hyperparamètre δ .

1.2.3 Loi *a posteriori* des paramètres

La vraisemblance et les lois *a priori* définies ci-dessus permettent de déterminer la loi *a posteriori* du vecteur des paramètres inconnus $\boldsymbol{\theta} = \{\mathbf{a}, w^2\}$ à l'aide de la formule de Bayes :

$$f(\boldsymbol{\theta}, \delta | \mathbf{y}) \propto f(\mathbf{y} | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \delta) f(\delta) \quad (1.12)$$

où $f(\mathbf{y} | \boldsymbol{\theta})$ et $f(\delta)$ ont été définis respectivement dans (1.4) et (1.11). On suppose que les paramètres inconnus sont *a priori* indépendants, ce qui mène à $f(\boldsymbol{\theta} | \delta) = f(\mathbf{a}) f(w^2 | \delta)$, d'où :

$$f(\boldsymbol{\theta}, \delta | \mathbf{y}) \propto \frac{1}{c(\mathbf{a}^+)^{\frac{L}{2}}} \frac{1}{w^{L+2}} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2w^2 c(\mathbf{a})} - \delta\right) \mathbf{1}_{\mathbb{S}}(\mathbf{a}) \mathbf{1}_{\mathbb{R}^+}(\delta). \quad (1.13)$$

Cette loi jointe *a posteriori* est trop complexe pour pouvoir calculer les estimateurs Bayésiens classiques MMSE et MAP (notons que le numérateur et le dénominateur du terme contenu dans l'exponentielle dépendent tous les deux de \mathbf{a}). Comme expliqué précédemment, la solution que nous envisageons pour cette étude réside dans l'utilisation de techniques d'échantillonnage stochastique de type MCMC. Ces méthodes permettent de générer des échantillons $\boldsymbol{\theta}^{(t)}$ (avec $t = 1, \dots, N_{\text{MCMC}}$, t étant l'indice de l'échantillon) asymptotiquement distribués suivant la loi d'intérêt. Les estimateurs MMSE et MAP peuvent alors être estimés à l'aide de ces échantillons comme suit

$$\hat{\boldsymbol{\theta}}_{\text{MMSE}} \approx \frac{1}{T} \sum_{t=1}^T \boldsymbol{\theta}^{(t)}, \quad \hat{\boldsymbol{\theta}}_{\text{MAP}} \approx \arg \max_{\boldsymbol{\theta}^{(t)}} f(\boldsymbol{\theta}^{(t)} | \mathbf{y}).$$

Le paragraphe suivant montre qu'une stratégie basée sur un échantillonneur de Gibbs hybride permet de générer les abondances et les variances suivant la loi jointe *a posteriori* (1.13).

1.3 Echantillonneur de Gibbs hybride

Ce paragraphe présente un échantillonneur de Gibbs hybride générant des échantillons suivant la loi jointe *a posteriori* $f(\boldsymbol{\theta}, \delta | \mathbf{y})$. Plus exactement, chaque paramètre sera généré

suivant sa loi conditionnelle. Lorsque cette génération n'est pas possible directement, une méthode hybride appelé Metropolis-within-Gibbs est employée : elle combine l'algorithme de Metropolis-Hastings [GRS96] avec l'échantillonneur de Gibbs. L'algorithme complet est décrit dans ALGO. 1.2.

1.3.1 Loi conditionnelle de \mathbf{a}

Le théorème de Bayes donne :

$$f(\mathbf{a}|\mathbf{y}, w^2) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\mathbf{a}) \quad (1.14)$$

ce qui mène facilement à :

$$f(\mathbf{a}|\mathbf{y}, w^2) \propto \frac{1}{[w^2c(\mathbf{a})]^{L/2}} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2w^2c(\mathbf{a})}\right) \mathbf{1}_{\mathbb{S}}(\mathbf{a}). \quad (1.15)$$

Cette loi conditionnelle est définie sur le simplexe \mathbb{S} . Ainsi, le vecteur généré \mathbf{a}^+ vérifie les contraintes de positivité et d'additivité. Pour générer suivant (1.15), un algorithme de Metropolis-within-Gibbs va être employé (voir ALGO. 1.1). Plus précisément, la génération des échantillons est basée sur une marche aléatoire [RC04, p.245] avec une loi Gaussienne $\mathcal{N}(0, u_r^2)$ comme loi de proposition. La variance u_r^2 de cette loi instrumentale est fixée de manière à obtenir un taux d'acceptation entre 0.15 et 0.5 comme recommandé dans [Rob96, p.55]. Une procédure d'acceptation-rejet devant suivre la génération des candidats permet de s'assurer qu'ils appartiennent bien à \mathbb{S} .

1.3.2 Loi conditionnelle de w^2

La loi conditionnelle de w^2 s'écrit $f(w^2|\mathbf{y}, \mathbf{a}, \delta) \propto f(\mathbf{y}|\boldsymbol{\theta})f(w^2|\delta)$. C'est à dire, la loi inverse Gamma telle que :

$$w^2|\mathbf{y}, \mathbf{a}, \delta \sim \mathcal{IG}\left(\frac{L}{2} + 1, \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2c(\mathbf{a})} + \delta\right). \quad (1.16)$$

ALGO. 1.1 Algorithme de Metropolis-within-Gibbs pour la simulation des abondances.

1. Générer \mathbf{a}^* selon loi Gaussienne $\mathcal{N}(0, u_r^2)$
 2. Test d'acceptation/rejet :
 - (a) si $\mathbf{a}^* \in \mathbb{S}$, accepter ce candidat (*accepter*)
 - (b) si $\mathbf{a}^* \notin \mathbb{S}$, retour étape 1 (*rejeter*)
 3. Prendre $\rho = \min \left\{ 1, \frac{f(\mathbf{a}^* | \dots)}{f(\mathbf{a}^{(t)} | \dots)} \right\} = \left\{ 1, \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{L/2} \exp \left[\frac{1}{2w^2} \left(\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{c(\mathbf{a}^*)} \right) \right] \right\}$
 4. Test acceptation/rejet :
 - (a) prendre $\mathbf{a}^{(t+1)} = \mathbf{a}^{(t)}$ avec proba $1 - \rho$ (*rejeter*)
 - (b) prendre $\mathbf{a}^{(t+1)} = \mathbf{a}^*$ avec proba ρ (*accepter*)
 5. Poser $\mathbf{a}^+ = \left[\mathbf{a}^T, 1 - \sum_{r=1}^{R-1} a_r \right]$
-

1.3.3 Loi conditionnelle de l'hyperparamètre

La loi conditionnelle de δ est la loi Gamma telle que :

$$\delta | w^2 \sim \mathcal{G} \left(1, \frac{1}{w^2} \right) \quad (1.17)$$

où $\mathcal{G}(a, b)$ est la loi Gamma de paramètre de forme a et de paramètre d'échelle b [RC04, p. 581].

ALGO. 1.2 Échantillonneur de Gibbs hybride pour le démélange spectral utilisant le NCM

1. Initialisation :
 - Échantillonner $\delta^{(0)}$ à partir de la densité de probabilité définie dans Eq. (1.11),
 - Échantillonner $w^{2(0)}$ à partir de la loi définie dans Eq. (1.9),
 2. Itérations : Pour $t = 1, 2, \dots$, faire
 - Échantillonner $\mathbf{a}^{(t)}$ à partir de la loi définie dans Eq. (1.15) employant une étape de Metropolis-Hastings (voir ALGO. 1.1),
 - Échantillonner $w^{2(t)}$ à partir de la loi définie dans Eq. (1.16),
 - Échantillonner $\delta^{(t)}$ à partir de la loi définie dans Eq. (1.17),
-

1.4 Extension aux cas d'endmembers de variances différentes

Jusqu'à présent, tous les spectres des endmembers partageaient la même variance w^2 . Ce paragraphe explore un cas où les endmembers ont des variances différentes. Ce cas de figure est intéressant lorsque l'on décide d'accorder divers degrés de confiance aux moyennes estimées \mathbf{m}_r ($r = 1, \dots, R$) retrouvées par les algorithmes N-FINDR ou VCA. Ainsi, on introduit un vecteur de variances $\mathbf{w} = [w_1^2, \dots, w_R^2]^T$ où w_r^2 est la variance du r -ième endmember. Tenant compte de cette hypothèse, nous avons :

$$e_r | \mathbf{m}_r, w_r^2 \sim \mathcal{N}(\mathbf{m}_r, w_r^2 \mathbf{I}_L). \quad (1.18)$$

1.4.1 Problèmes d'identifiabilité

Théorie générale

Si les lois *a priori* choisies pour w_r^2 ($r = 1, \dots, R$) ne sont pas suffisamment informatives, des problèmes d'identifiabilité peuvent survenir. En effet, pour R composants purs impliqués dans le mélange, la log-vraisemblance peut s'écrire :

$$\log f(\mathbf{y} | \mathbf{w}, \mathbf{a}) = -\frac{L}{2} \log C(\mathbf{w}) - \frac{K(\mathbf{y}, \mathbf{a})}{C(\mathbf{a}, \mathbf{w})} \quad (1.19)$$

où $K(\mathbf{y}, \mathbf{a}) = \frac{1}{2} \|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2$ et $C(\mathbf{a}, \mathbf{w}) = \sum_{r=1}^R w_r^2 a_r^2$. Recherchant les valeurs du vecteur \mathbf{w} qui maximise la log-vraisemblance, on égalise ses R dérivées partielles à zéro :

$$\begin{cases} \frac{\partial \log f(\mathbf{y} | \mathbf{w})}{\partial w_1^2} = -\frac{La_1^2}{2C(\mathbf{a}, \mathbf{w})} - \frac{K(\mathbf{y}, \mathbf{a})a_1^2}{[C(\mathbf{a}, \mathbf{w})]^2} = 0, \\ \vdots \\ \frac{\partial \log f(\mathbf{y} | \mathbf{w})}{\partial w_R^2} = -\frac{La_R^2}{2C(\mathbf{a}, \mathbf{w})} - \frac{K(\mathbf{y}, \mathbf{a})a_R^2}{[C(\mathbf{a}, \mathbf{w})]^2} = 0, \end{cases} \quad (1.20)$$

ce qui mène facilement à

$$C(\mathbf{a}, \mathbf{w}) = \sum_{r=1}^R w_r^2 a_r^2 = \frac{2K}{L}.$$

Par conséquent, la vraisemblance $f(\mathbf{y} | \mathbf{w}, \mathbf{a})$ possède une infinité de maxima situés sur l'hyperplan \mathcal{H} défini par :

$$\mathcal{H} = \left\{ \mathbf{w} = [w_1^2, \dots, w_R^2]^T \mid C(\mathbf{a}, \mathbf{w}) = \frac{2K}{L} \right\}. \quad (1.21)$$

Ce problème d'identifiabilité peut être contourné lorsque plusieurs pixels avec les mêmes caractéristiques sont considérés. Prenant l'hypothèse que le vecteur des variances \mathbf{w} est le même pour P pixels (avec $P > 1$), on obtient le système de P équations linéaires suivant :

$$\begin{cases} \sum_{r=1}^R w_r^2 a_{r,1}^2 = \frac{2K_1}{L}, \\ \vdots \\ \sum_{r=1}^R w_r^2 a_{r,P}^2 = \frac{2K_P}{L}, \end{cases} \quad (1.22)$$

où $a_{r,p}$ dénote le coefficient d'abondance du r -ième endmember dans le p -ième pixel, $K_p = \frac{1}{2} \|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2$, $\mathbf{a}_p = [a_{1,p}, \dots, a_{R-1,p}]^T$ et \mathbf{y}_p est le spectre mesuré du p -ième pixel (avec $p = 1, \dots, P$). Sous forme matricielle, le système s'écrit :

$$\mathbf{\Lambda} \mathbf{w} = \frac{2}{L} \mathbf{K}$$

avec

$$\mathbf{\Lambda} = \begin{pmatrix} a_{1,1}^2 & \cdots & a_{R,1}^2 \\ \vdots & \ddots & \vdots \\ a_{1,P}^2 & \cdots & a_{R,P}^2 \end{pmatrix}, \quad \mathbf{K} = \begin{pmatrix} K_1 \\ \vdots \\ K_P \end{pmatrix}. \quad (1.23)$$

Ainsi, le vecteur \mathbf{w} maximisant la vraisemblance est unique à la condition que le rang de la matrice $\mathbf{\Lambda}$ soit égal à R .

Exemples

Dans ce paragraphe, la condition d'identifiabilité est illustrée pour différents nombres de pixels générés avec un mélange de 2 endmembers. Comme base de départ, un pixel a été généré avec $\mathbf{w} = [0.006, 0.002]^T$. La figure 1.2 (en haut) représente la log-vraisemblance correspondante en fonction de (w_1^2, w_2^2) pour $P = 1$ pixel. Cette figure montre clairement que les maxima sont atteints pour une infinité de couples (w_1^2, w_2^2) situés sur un hyperplan (ici, une droite).

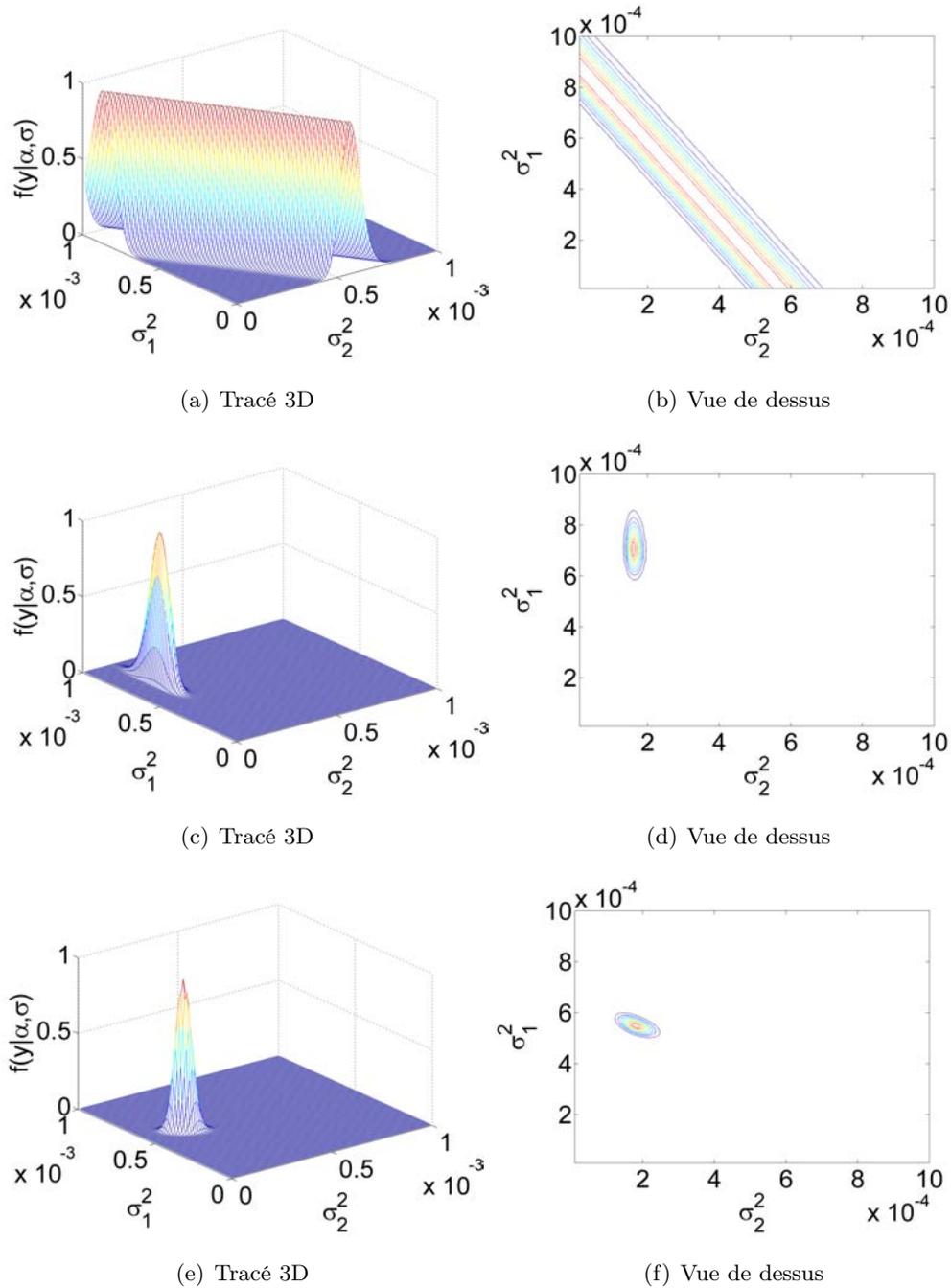


FIGURE 1.2 – Vraisemblance pour (de haut en bas) $P = 1, 2, 9$ pixels en fonction de (w_1^2, w_2^2) . (a) Vue 3D. (b) Vue de dessus.

La figure 1.2 (au centre) représente le cas où le nombre de pixels P est égal à 2. Un unique maximum peut être constaté car le rang de \mathbf{A} est égal à 2 pour cet exemple. Les résultats de la figure 1.2 (en bas) obtenus pour $P = 9$ pixels montrent que la vraisemblance est plus piquée autour des valeurs réelles de \mathbf{w} que dans le cas $P = 2$ pixels.

1.4.2 Modèle Bayésien hiérarchique

Ce paragraphe décrit le modèle Bayésien hiérarchique proposé lorsque les endmembers ont des variances différentes. Compte tenu des conclusions du précédent paragraphe, P pixels sont considérés pour éviter les problèmes d'identifiabilité :

$$\mathbf{y}_p = \sum_{r=1}^R e_{r,p} \alpha_{r,p}, \quad p = 1, \dots, P \quad (1.24)$$

où $e_{r,p} | \mathbf{m}_r, w_r^2 \sim \mathcal{N}(\mathbf{m}_r, w_r^2 \mathbf{I}_L)$, $\mathbf{m}_r = [m_{r,1}, \dots, m_{r,L}]^T$ représente le vecteur moyenne (connu) des endmembers $e_{r,p}$, et $\mathbf{w} = [w_1^2, \dots, w_R^2]^T$ est le vecteur des variances inconnues. La notation $e_{r,p}$ nous permet d'établir l'hypothèse d'indépendance entre les spectres des pixels \mathbf{y}_p ($p = 1, \dots, P$). En effet, si les vecteurs moyennes et les variances sont les mêmes pour chaque pixel, les réalisations des endmembers restent différentes. La vraisemblance jointe pour l'ensemble des P pixels peut donc s'écrire comme le produit des vraisemblances. Cette dernière est décrite ci-dessous en plus des lois *a priori* associées aux différents paramètres du modèle.

Vraisemblance

La fonction de vraisemblance pour le pixel p est :

$$f(\mathbf{y}_p | \mathbf{a}_p, \mathbf{w}) \propto \frac{1}{[c(\mathbf{a}_p)]^{L/2}} \exp \left[-\frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2c(\mathbf{a}_p)} \right]$$

avec

$$\begin{cases} c(\mathbf{a}_p) = \sum_{r=1}^R w_r^2 a_{r,p}^2, \\ \boldsymbol{\mu}(\mathbf{a}_p) = \sum_{r=1}^R \mathbf{m}_r a_{r,p}. \end{cases} \quad (1.25)$$

Comme précisé plus haut, les spectres des pixels \mathbf{y}_p ($p = 1, \dots, P$) sont *a priori* supposés indépendants. La vraisemblance jointe pour l'ensemble des P pixels peut donc s'écrire :

$$f(\mathbf{Y} | \mathbf{A}, \mathbf{w}) \propto \prod_{p=1}^P \frac{1}{[c(\mathbf{a}_p)]^{L/2}} \exp \left[- \sum_{p=1}^P \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2c(\mathbf{a}_p)} \right], \quad (1.26)$$

avec $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_P]$ et $\mathbf{A} = [\mathbf{a}_1^+, \dots, \mathbf{a}_P^+]$.

Lois a priori

Des lois uniformes sur le simplexe défini dans (1.6) ont été choisies comme lois *a priori* pour les vecteurs d'abondances \mathbf{a}_p supposés *a priori* indépendants ($p = 1, \dots, P$) :

$$f(\mathbf{A}) \propto \prod_{p=1}^P \mathbf{1}_{\mathbb{S}}(\mathbf{a}_p). \quad (1.27)$$

Les lois *a priori* pour les variances des endmembers sont des lois conjuguées inverse-Gamma avec comme hyperparamètre δ (similaire à (1.9)). Une loi de Jeffreys est choisie pour l'hyperparamètre δ comme au paragraphe 1.3.3.

1.5 Algorithme MCMC

Comme dans le cas précédent, un échantillonneur hybride de Metropolis-within-Gibbs sera employé pour générer des échantillons asymptotiquement distribués suivant la loi jointe des vecteurs d'abondances et des variances des endmembers. L'échantillonneur génère itérativement \mathbf{a}_p selon $f(\mathbf{a}_p | \mathbf{y}_p, \mathbf{w})$ pour chaque pixel $p = 1, \dots, P$, w_r^2 suivant $f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \delta)$ pour chaque endmember $r = 1, \dots, R$ (\mathbf{w}_{-r} dénote le vecteur des variances \mathbf{w} dont la r -ième composante a été retirée), et δ selon $f(\delta | \mathbf{w})$.

1.5.1 Génération des abondances

La loi conditionnelle *a posteriori* du vecteur d'abondances \mathbf{a}_p est indépendante des autres pixels et peut s'exprimer ainsi :

$$f(\mathbf{a}_p | \mathbf{y}_p, \mathbf{w}) \propto \frac{1}{[c(\mathbf{a}_p)]^{L/2}} \exp \left[- \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2c(\mathbf{a}_p)} \right] \mathbf{1}_{\mathbb{S}}(\mathbf{a}_p). \quad (1.28)$$

La génération de \mathbf{a}_p suivant cette distribution est réalisée par une étape de Metropolis-Hastings, dont l'algorithme est similaire à celui présenté dans ALGO. 1.1.

1.5.2 Génération des variances

La loi conditionnelle *a posteriori* de w_r^2 peut se définir selon la relation suivante :

$$f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \delta) \propto f(\mathbf{Y} | \mathbf{A}, \mathbf{w}) f(w_r^2 | \nu, \delta)$$

avec $\mathbf{w}_{-r} = [w_1^2, \dots, w_{r-1}^2, w_{r+1}^2, \dots, w_r^2]^T$. En posant $c(\mathbf{a}_{-r,p}) = \sum_{\substack{i=1 \\ i \neq r}}^R w_i^2 a_i^2$, quelques calculs élémentaires permettent d'obtenir (voir Annexe A) :

$$f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \delta) \propto \left(\frac{1}{w_r^2} \right)^{\nu+1} \prod_{p=1}^P [w_r^2 a_{r,p}^2 + c(\mathbf{a}_{-r,p})]^{-L/2} \times \exp \left[- \sum_{p=1}^P \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2 [w_r^2 a_{r,p}^2 + c(\mathbf{a}_{-r,p})]} - \frac{\delta}{w_r^2} \right]. \quad (1.29)$$

Échantillonner suivant cette loi peut se faire à l'aide d'une étape de Metropolis-Hastings. La loi *a priori* de w_r^2 étant une loi inverse Gamma, nous avons choisi une loi de proposition inverse Gamma :

$$w_r^2 \sim \mathcal{IG}(\alpha_w, \beta_w), \quad (1.30)$$

où α_w et β_w sont des paramètres réglables. Ces paramètres ont été choisis de manière à obtenir la moyenne et la variance de (1.16), ce qui améliore le taux d'acceptation de l'échantillonneur.

1.5.3 Génération de l'hyperparamètre

La loi conditionnelle *a posteriori* de l'hyperparamètre δ est la loi Gamma :

$$\delta | \mathbf{w} \sim \mathcal{G} \left(R, \sum_{r=1}^R \frac{1}{w_r^2} \right). \quad (1.31)$$

L'algorithme détaillant les étapes décrites ci-dessus est présenté dans ALGO. 1.3.

ALGO. 1.3 Démélange spectrale employant le NCM avec différentes variances

1. Initialisation :

- Échantillonner l'hyperparamètre $\delta^{(0)}$ à partir de la loi dans Eq. (1.11),
- Échantillonner $\mathbf{w}^{(0)} = [w_1^{2(0)}, \dots, w_R^{2(0)}]$ à partir de la loi dans Eq. (1.9),
- Pour chaque pixel p , échantillonner $\mathbf{a}^{(0)}$ suivant une loi uniforme définie sur \mathbb{S} ,

2. Itérations : Pour $t = 1, 2, \dots$, faire

- Pour $p = 1, \dots, P$, échantillonner $\mathbf{a}_p^{(t)}$ à partir de la loi dans Eq. (1.28) utilisant une étape de Metropolis-Hastings,
 - Pour $r = 1, \dots, R$, échantillonner $w_r^{2(t)}$ à partir de la loi dans Eq. (1.29) utilisant une étape de Metropolis-Hastings,
 - Échantillonner $\delta^{(t)}$ à partir de la loi dans Eq. (1.31),
-

1.6 Résultats de simulations sur des données synthétiques

Ce paragraphe illustre les performances des deux algorithmes de démélange proposés via des simulations effectuées sur des données synthétiques. Ces simulations ont été effectuées à l'aide de pixels observés sur $L = 276$ bandes spectrales de longueur d'ondes allant de $0.4\mu m$ à $2.5\mu m$ (du visible au proche infrarouge).

1.6.1 Algorithme NCM avec des endmembers de même variance

Les simulations présentées dans ce paragraphe ont été obtenues à partir de l'algorithme décrit au paragraphe 1.3. Un mélange synthétique avec $R = 2$ composants purs est considéré dans cette expérience. Cet exemple trivial a pour avantage d'avoir peu de paramètres dont les lois conditionnelles peuvent être représentées plus facilement. Les moyennes des endmembers \mathbf{m}_1 et \mathbf{m}_2 ont été extraites des bibliothèques spectrales fournies par le logiciel ENVI [RSI03]. Ces spectres correspondent au béton de construction et à l'herbe verte et sont illustrés sur la figure 1.3.

La variance des endmembers est $w^2 = 0.01$. Le mélange linéaire considéré dans ce paragraphe est construit avec comme valeur des abondances $\mathbf{a}^+ = [0.3, 0.7]^T$. La figure 1.4 montre les lois conditionnelles des abondances générées par l'échantillonneur de Gibbs

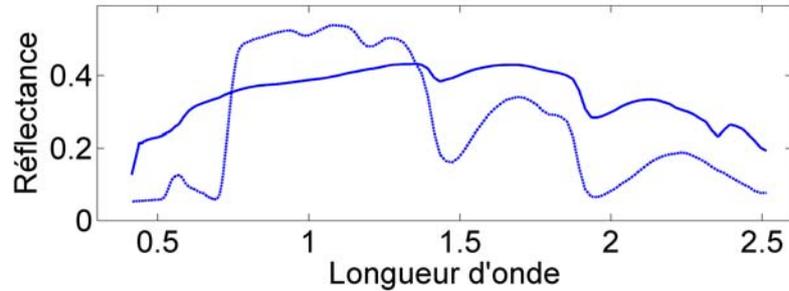


FIGURE 1.3 – Spectres des endmembers : béton de construction (trait continu), herbe verte (trait pointillés).

proposé avec $N_{MC} = 25000$ itérations incluant $N_{bi} = 5000$ itérations de chauffage (qui ne seront pas utilisées pour l'estimation des paramètres). Ces lois sont bien en accord avec les valeurs réelles des abondances. La figure 1.5 montre la loi conditionnelle de w^2 estimée par l'algorithme proposé qui est également en accord avec la valeur de la variance $w^2 = 0.01$.

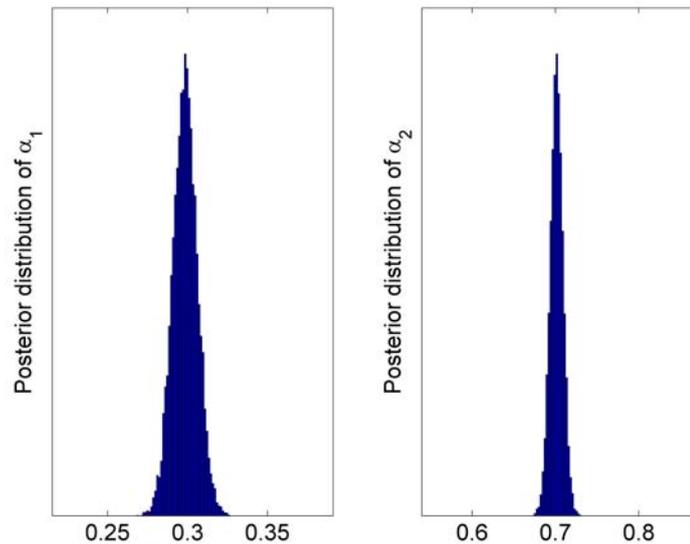
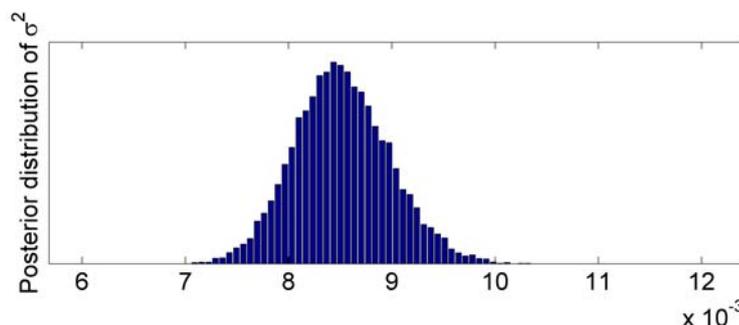


FIGURE 1.4 – Loi conditionnelle estimée des abondances $[a_1, a_2]^T$.

L'algorithme proposé a également été testé pour différentes valeurs du SNR. La figure 1.6 montre les estimations MAP des abondances a_r et les écarts-types correspondants en

FIGURE 1.5 – Lois conditionnelles estimées de la variance w^2 .

fonction du SNR. Il est important de signaler que l’algorithme Bayésien proposé permet d’établir des intervalles de confiance pour les différentes estimations. Ces intervalles sont calculés à partir des échantillons générés par l’algorithme de Gibbs. Il faut noter que les SNR des spectromètres actuels comme AVIRIS sont rarement inférieurs à 20dB après que les bandes d’absorption de vapeur d’eau soient rejetées [Gre98]. Les résultats de la figure 1.6 indiquent que l’algorithme Bayésien proposé permet d’obtenir des résultats satisfaisants pour différents SNR. Cette figure montre également que les estimations de a_r convergent (au sens des moindres carrés) vers les valeurs réelles de a_r quand le SNR augmente.

1.6.2 Algorithme NCM avec différentes variances

Les performances de l’algorithme décrit dans le paragraphe 1.5 sont mises en évidence grâce aux résultats de simulations sur des données synthétiques. Dans ces simulations, $P = 3$ pixels ont été générés en mélangeant $R = 3$ endmembers suivant (1.24). Les valeurs des paramètres sont :

- Pixel 1 : $\mathbf{a}_1^+ = [0.5, 0.3, 0.2]^T$,
- Pixel 2 : $\mathbf{a}_2^+ = [0.4, 0.1, 0.5]^T$,
- Pixel 3 : $\mathbf{a}_3^+ = [0.1, 0.3, 0.6]^T$.

Les variances des $R = 3$ endmembers sont les mêmes pour chacun des pixels et sont égales à :

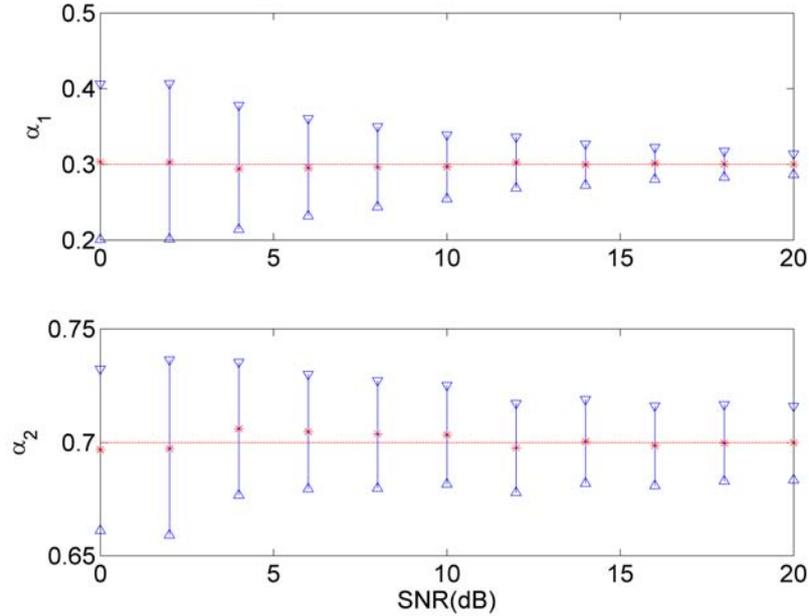
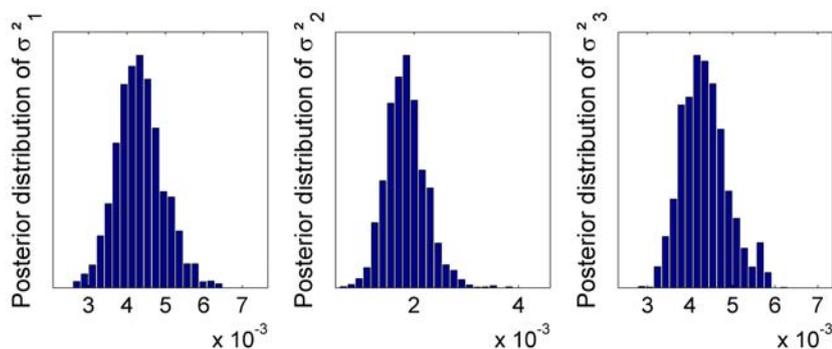


FIGURE 1.6 – Estimations MAP (croix) et écart-types (barres verticales) des composantes de \mathbf{a}^+ en fonction du SNR.

- $w_1^2 = 0.004$,
- $w_2^2 = 0.002$,
- $w_3^2 = 0.0035$.

La figure 1.7 montre les lois conditionnelles estimées des variances w_r^2 ($r = 1, \dots, R$) qui sont clairement centrées autour des valeurs réelles des différentes variances. Les histogrammes des abondances générées pour chaque pixel par l'algorithme de Gibbs proposé sont représentées sur la figure 1.8. Ces résultats sont en accord avec les valeurs réelles des abondances.

Il est possible de comparer les performances de l'algorithme basé sur des variances distinctes (dépeint au paragraphe 1.5) avec celles de l'algorithme basées sur une unique variance pour les endmembers (décrit au paragraphe 1.3). $P = 9$ pixels synthétiques, générés selon le modèle proposé avec des variances distinctes, ont été démêlés par ces deux algorithmes. Les erreurs quadratiques moyennes (EQM) des vecteurs d'abondances

FIGURE 1.7 – Lois conditionnelles estimées des variances pour $P = 3$ pixels.

sont ensuite calculées pour ces algorithmes à l'aide de 100 simulations de Monte Carlo. La table 1.1 affiche les résultats correspondants et montre qu'en tenant compte de variances différentes, les performances d'estimations s'en trouvent améliorées au prix d'un coût calculatoire plus élevé.

TABLE 1.1 – EQM globale du vecteur d'abondance pour le NCM avec une unique variance et avec des variances distinctes.

NCM avec une variance	NCM avec plusieurs variances
1.72×10^{-2}	1.54×10^{-2}

1.6.3 Comparaison avec d'autres algorithmes

Dans cette partie, l'algorithme développé dans le paragraphe 1.5 est comparé avec d'autres méthodes de démixage précédemment proposé dans la littérature. Plus précisément, nous comparons les stratégies de démixage spectral suivantes :

- l'algorithme présenté au paragraphe 1.2 (nommé « NCM Bayésien »),
- un algorithm Bayésien basé sur le MML (nommé « MML Bayésien ») [DTC08],
- l'algorithm fully constrained least-squares (FCLS) [HC01],
- l'algorithme MVC-NMF [MQ07],
- la méthode NN-ICA [PO04],

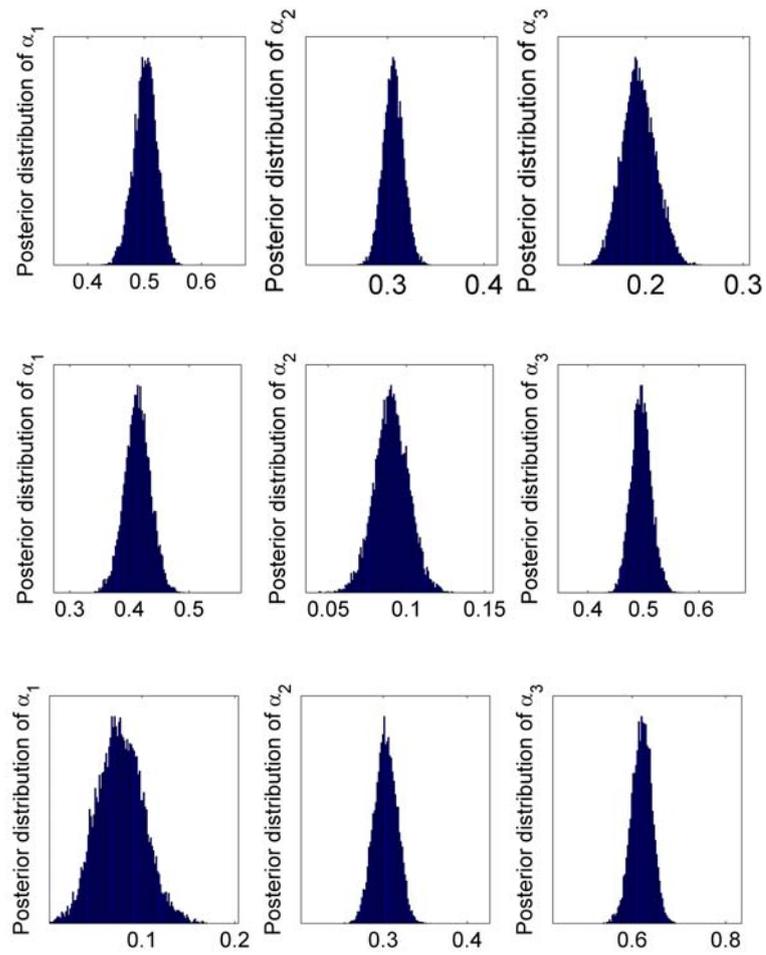


FIGURE 1.8 – Lois conditionnelles estimées des abondances pour chaque pixel (haut : pixel 1, centre : pixel 2, bas : pixel 3).

Les trois premiers algorithmes, à savoir le NCM Bayésien et les deux algorithmes développés dans [DTC08] et [HC01] sont précédés par l'algorithme VCA comme méthode d'extraction d'endmember (MEE). Cependant, d'autres MEE auraient également pu être utilisées à la place du VCA (tels que le N-FINDR ou encore l'algorithme PPI [Boa93])¹. $P = 625$ pixels ont été générés selon le MML avec $R = 6$ endmembers et un bruit blanc Gaussien additif conduisant à un SNR de 20dB. Afin d'évaluer la flexibilité du NCM par rapport à l'absence de pixels purs, nous avons retiré de l'image synthétique les spectres observés proches des moyennes des endmembers, soit $\forall p, r \frac{1}{L} \|\mathbf{y}_p - \mathbf{m}_r\|^2 < \varepsilon$ où le seuil ε est fixé à 6.0×10^{-2} . Dès lors, les pixels purs sont bien enlevés de l'image synthétique. L'EQM globale du r -ième coefficient d'abondance estimée est définie par :

$$\text{EQM}_r^2 = \frac{1}{P} \sum_{p=1}^P (\hat{a}_{r,p} - a_{r,p})^2 \quad (1.32)$$

où $\hat{a}_{r,p}$ dénote l'estimateur MMSE de l'abondance $a_{r,p}$. Le tableau 1.2 montre les EQM globales pour les cinq différentes méthodes de démixage mentionnées précédemment (NCM Bayésien, MML Bayésien, FCLS, MVC-NMF et NN-ICA). L'algorithme NCM Bayésien proposé obtient clairement de meilleurs résultats que les autres algorithmes de démixage. Cette amélioration des performances d'estimations obtenues avec le NCM est due à la robustesse du modèle à l'absence de pixels purs dans l'image.

Nous reportons également dans le tableau 1.3 l'erreur de construction globale pour cet ensemble de pixel, définie ci-dessous :

$$e = \sqrt{\frac{1}{P} \sum_{p=1}^P \|\mathbf{y}_p - \mathbf{M}\hat{\mathbf{a}}_p^+\|^2}. \quad (1.33)$$

Ces résultats ont été obtenus pour les algorithmes NCM Bayésien, MML Bayésien et FCLS². Les algorithmes MML Bayésien et FCLS nécessitent la connaissance *a priori*

1. Remarquons également que les méthodes MVC-NMF et NN-ICA ne sont pas de simples méthodes d'inversion mais des méthodes conjointes. Par conséquent, ces méthodes n'ont pas besoin d'être précédées par une MEE. Nous avons également voulu comparer les performances de l'algorithme proposé dans ce travail avec des méthodes conjointes.

2. Les algorithmes MVC-NMF et NN-ICA ont été ignorés pour cette comparaison puisqu'il s'agit de méthodes conjointes. Du coup, de petites erreurs de reconstructions pour ces algorithmes n'impliquent pas nécessaire un démixage spectral réussi.

TABLE 1.2 – EQM globales de chaque coefficient d'abondance pour différents algorithmes de démélanges ($\times 10^{-3}$).

	NCM Bayésien	MML Bayésien	FCLS	MVC-NMF	NN-ICA
EQM ₁ ²	7.8	13	9.1	7.7	18.2
EQM ₂ ²	9.6	10.4	9.9	24.1	41.4
EQM ₃ ²	8.5	23.2	10.2	45.4	45.2
EQM ₄ ²	8.2	15.9	8.8	26.2	45.3
EQM ₅ ²	10.2	14.8	11.5	12.5	46.8
EQM ₆ ²	10.8	11.7	11.5	35.6	44.9

des spectres déterministes $\mathbf{m}_1, \dots, \mathbf{m}_R$ contenus dans la matrice \mathbf{M} . Par conséquent, la matrice des endmembers \mathbf{M} (en donnée) est également utilisée pour le calcul des erreurs de reconstructions associé au NCM Bayésien pour une comparaison équitable. Comme montré dans la table 1.3, le NCM Bayésien fournit les plus petites erreurs de reconstruction.

TABLE 1.3 – Erreurs de reconstruction pour les algorithmes NCM Bayésien, MML Bayésien et FCLS.

	NCM	LMM	FCLS
e	1.26	1.32	1.28

Enfin, les performances de ces trois algorithmes ont été comparées dans le cas d'une image synthétique comprenant des pixels purs. L'intérêt d'une telle expérience réside dans le cas où l'on doit choisir entre ces trois algorithmes pour analyser une image contenant des pixels purs, alors que précédemment, nous avons cherché à évaluer la robustesse de ces algorithmes par rapport à l'absence de ces pixels purs. Ainsi, nous avons gardé le même nombre de pixels $P = 625$ avec le même nombre d'endmembers $R = 6$ et le même niveau de bruit (20dB). Les EQM ont alors été calculés en utilisant (1.32). Le tableau 1.4

montre que l'algorithme proposé dans notre étude a les mêmes performances d'estimation que l'algorithme MML Bayésien et surpasse l'algorithme FCLS.

TABLE 1.4 – EQM globaux de chaque coefficient d'abondance pour une image synthétique contenant des pixels purs ($\times 10^{-3}$).

	NCM Bayésien	MML Bayésien	FCLS
MSE_1^2	1.2	1.2	1.4
MSE_2^2	1.1	1.1	1.2
MSE_3^2	1.2	1.2	1.4
MSE_4^2	0.34	0.35	0.39
MSE_5^2	1.9	1.9	2.2
MSE_6^2	5.9	5.9	6.4

1.7 Démélange spectral d'une image AVIRIS

Ce paragraphe considère une image hyperspectrale réelle de taille 50×50 représentée sur la figure 1.9 afin d'évaluer les performances des différents algorithmes. Cette image a été extraite d'une plus grande image acquise en 1997 par le spectromètre AVIRIS au dessus de Moffett Field, à la pointe sud de la baie de San Francisco, en Californie. Il faut noter que la présence de vapeur d'eau dans l'atmosphère cause l'absorption de certaines radiations à différentes longueurs d'onde. Le rapport signal-à-bruit à ces bandes d'absorptions étant très bas, toute information sur le spectre de réflectance est perdue. Par conséquent, le jeu de données a été réduit d'un nombre de bandes de 224 à 189 en supprimant les bandes d'absorption de la vapeur d'eau. L'imagette comporte un large point d'eau (une partie d'un lac qui apparaît en foncé en haut de l'image) et une zone côtière composé de terre et de végétation.

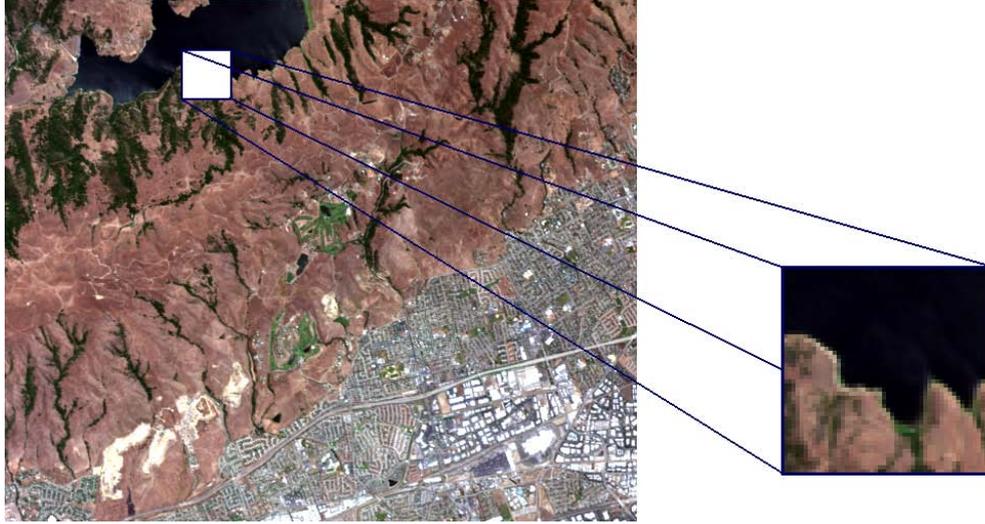


FIGURE 1.9 – Données hyperspectrales réelles : image de Moffett field acquise par AVIRIS en 1997 (à gauche) et la région d'intérêt représenté en couleurs réelles (à droite).

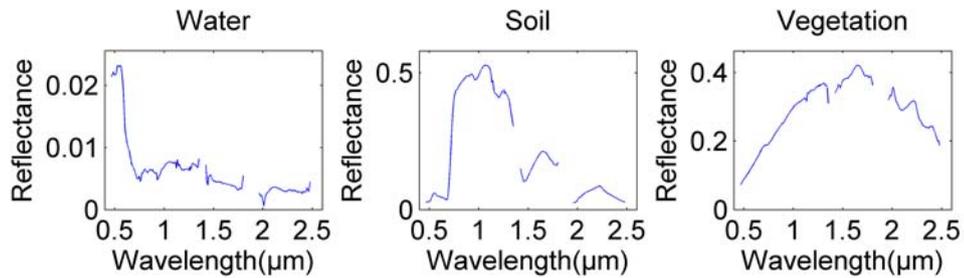


FIGURE 1.10 – Les spectres des $R = 3$ endmembers obtenus par l'algorithme N-FINDR.

Les étapes de démixage expliquées dans le chapitre précédent vont être appliquées pas à pas. Dans un premier temps, une ACP est utilisée en tant que méthode de pré-traitement afin de déterminer le nombre d'endmembers [KM02]. Puis la méthode N-FINDR a été appliquée à cette imagerie pour estimer les spectres des endmembers. Les $R = 3$ endmembers extraits (représentés sur la figure 1.10) correspondent à la végétation, l'eau et la terre et ont été utilisés pour définir les vecteurs moyennes \mathbf{m}_1 , \mathbf{m}_2 et \mathbf{m}_3 .

1.7.1 Algorithme NCM avec une seule variance

Les cartes d'abondances de l'image estimées par l'algorithme proposé au paragraphe 1.2 et 1.3 (pour les $R = 3$ matériaux purs) sont représentés sur la Figure 1.11 (en bas). Un pixel blanc (respectivement noir) sur la carte indique une forte (respectivement faible) valeur des coefficients d'abondance. Ainsi, la zone de lac (représentée par les pixels blancs sur la carte d'abondance de l'eau et en noir sur les autres cartes) est clairement retrouvée. Comme décrit sur la figure 1.11, on retrouve des cartes d'abondances similaires pour chacune des trois méthodes d'inversion. D'autres résultats donnés par les algorithmes MVC-NMF [MQ07] et NN-ICA [PO04] sont donnés sur la figure 1.12. Les endmembers estimés par ces méthodes conjointes, représentés sur la figure 1.13, ne correspondent pas exactement à l'eau, la terre et la végétation, ce qui explique certaines différences avec les cartes obtenues à l'aide des méthodes précédentes (par exemple avec l'algorithme NCM Bayésien).

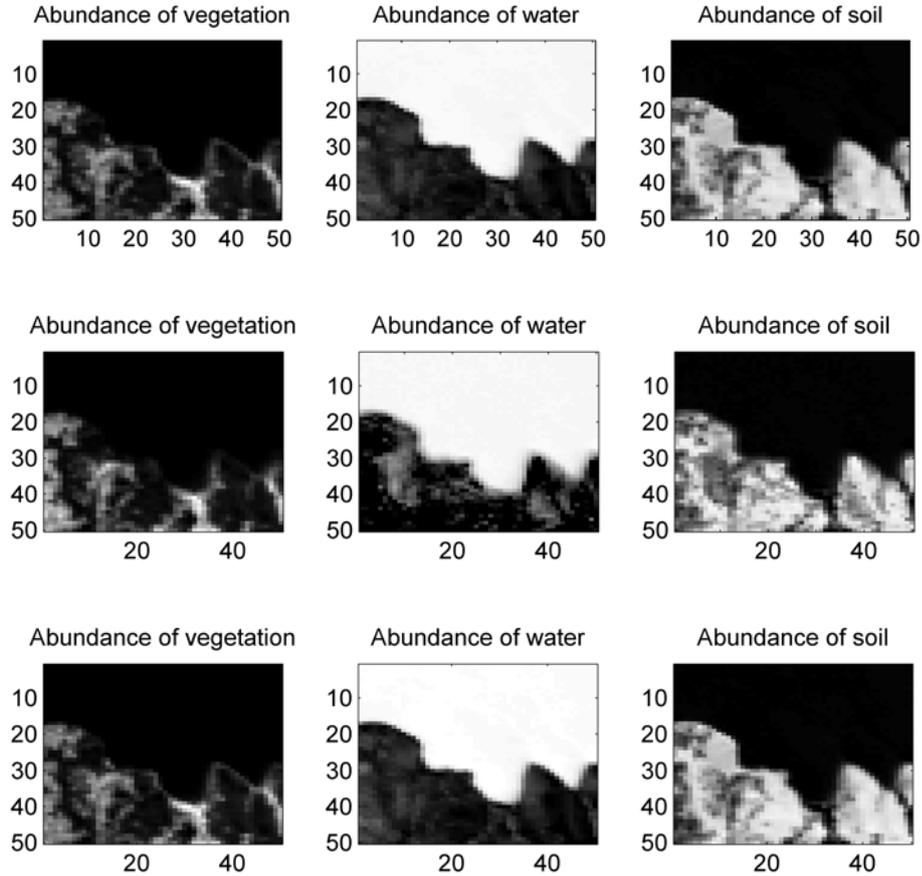


FIGURE 1.11 – En haut : cartes d’abondances estimées par le MML Bayésien (tiré de [DTC08]). Au milieu : cartes d’abondances estimées par l’algorithme FCLS [HC01]. En bas : cartes d’abondances estimées par l’algorithme proposé (pixel noir (resp. blanc) signifie absence (resp. présence) du matériau).

Les résultats concernant l’estimation de la variance des endmembers w^2 sont également présentés. La figure 1.14 montre les lois conditionnelles estimées de w^2 pour les pixels $\#(35, 43)$ (à gauche) et $\#(43, 35)$ (à droite) de l’image ainsi que leurs estimations MAP.

L’algorithme Bayésien proposé dans cette étude peut être employé pour estimer la probabilité de présence d’un matériau pur dans un pixel, défini par $P[a_i > \eta | \mathbf{y}_p]$ où η

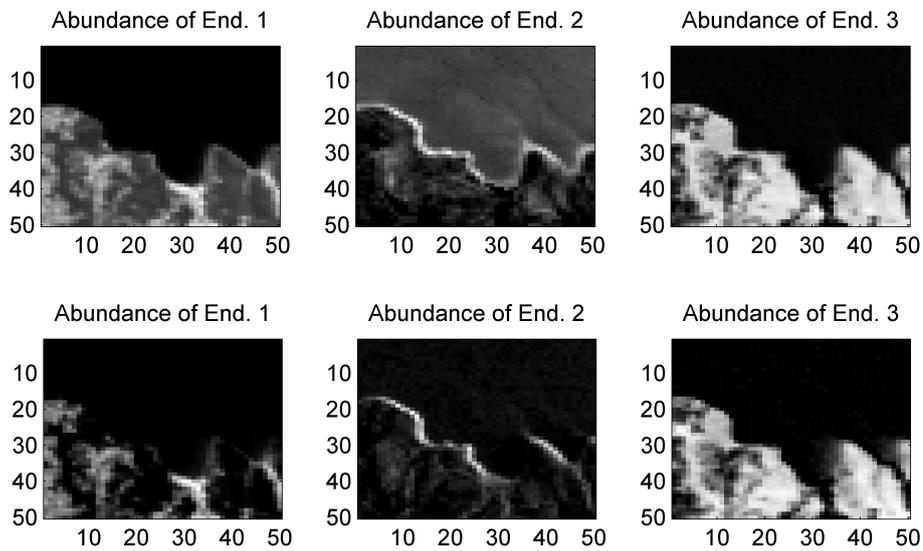


FIGURE 1.12 – En haut : cartes d'abondances estimées par l'algorithme MVC-NMF. En bas : cartes d'abondances estimées par l'algorithme NN-ICA (pixel noir (resp. blanc) signifie absence (resp. présence) du matériau).

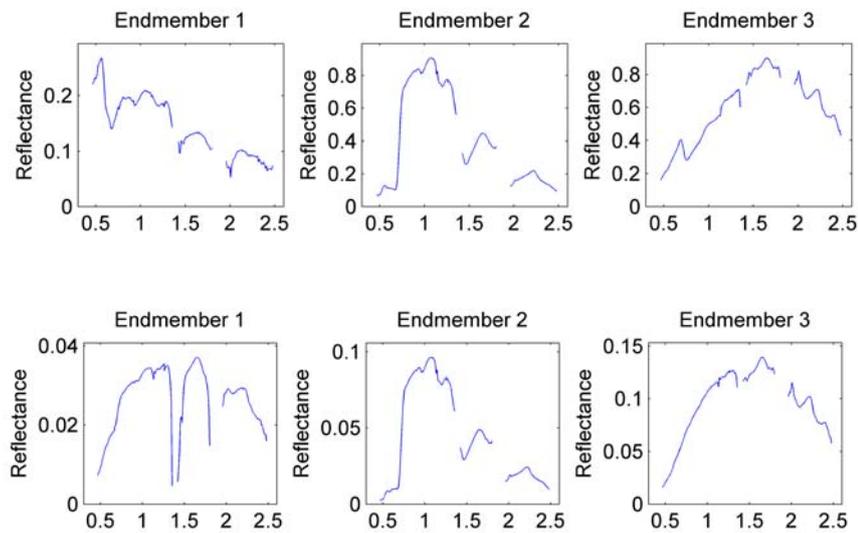


FIGURE 1.13 – En haut : les spectres des $R = 3$ endmembers obtenus par l'algorithme MVC-NMF [MQ07]. En bas : les spectres des $R = 3$ endmembers obtenus par l'algorithme NN-ICA [PO04].

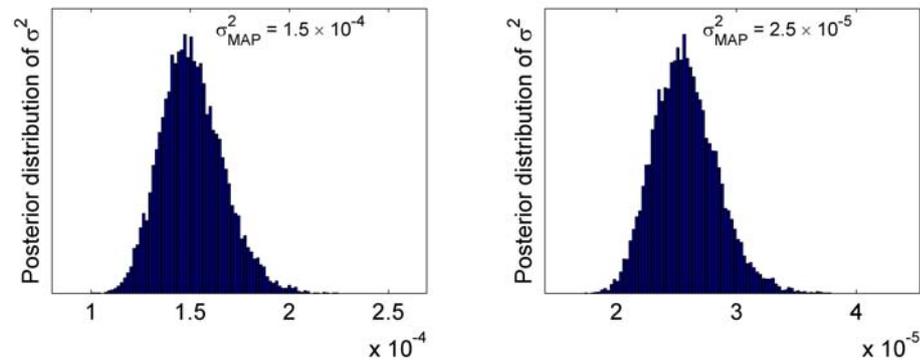


FIGURE 1.14 – Lois conditionnelles de w^2 pour les pixels $\#(35, 43)$ (à gauche) et $\#(43, 35)$ (à droite) estimées par l'algorithme proposé.

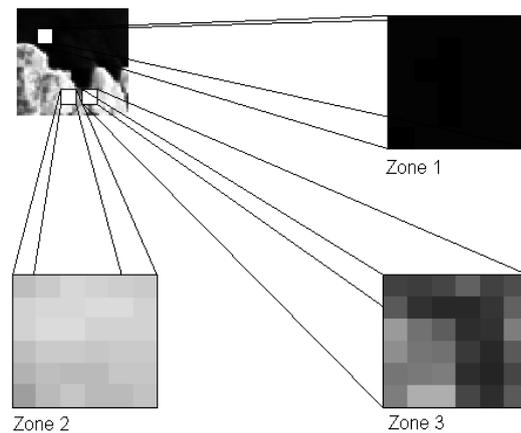


FIGURE 1.15 – Régions d'eau, de sol et de végétation analysées pour les probabilités de présence.

est un seuil donné. Trois zones distinctes de 6×6 pixels, représentées sur la figure 1.15, ont été analysées pour estimer ces probabilités. La première région (zone 1) est extraite de la zone du lac et de ce fait contient une majorité de pixels d'eau. Inversement, les deux autres régions (zone 2 et 3) sont des zones côtières contenant de la terre et de la végétation. Le tableau 1.5 montre les résultats obtenus pour différents seuils dans chaque zone analysée. Ces probabilités reflètent bien la composition de chaque zone, ainsi la

zone 1 est pratiquement composée à 100% d'eau tandis que les autres zones ont des compositions plus nuancées (par exemple, il n'y a que 50% de chances de trouver une abondance de terre de plus de 90% dans la zone 2).

TABLE 1.5 – Probabilité de présence pour chaque endmember.

	Zone 1 $\eta = 0.98$	Zone 2 $\eta = 0.9$	Zone 3 $\eta = 0.8$
$P[a_{\text{eau}} > \eta \mathbf{y}_p]$	0.9922	0	0
$P[a_{\text{terre}} > \eta \mathbf{y}_p]$	0	0.5147	0.0556
$P[a_{\text{végétation}} > \eta \mathbf{y}_p]$	0	0	0.2774

1.7.2 Algorithme NCM avec des variances distinctes

L'image hyperspectrale AVIRIS a également été analysée à l'aide de l'algorithme détaillé au paragraphe 1.4 afin d'évaluer ses performances d'estimation. Comme l'algorithme requiert plus d'un pixel pour fonctionner correctement, l'image a été découpée en 256 blocs de 3×3 pixels. Dès lors, la région étudiée³ a été réduite à 48×48 pixels. Les variances estimées pour les endmembers associés au bloc centré autour du pixel $\#(35, 43)$ sont données dans le tableau 1.6.

TABLE 1.6 – Estimation MMSE de w_r^2 ($r = 1, \dots, R$).

	Terre	Végétation	Eau
Estimations MMSE	1×10^{-4}	6.9×10^{-3}	1×10^{-4}

3. Seulement les bords à droite et en bas de l'image n'ont pas été inclus, ce qui est une très petite portion comparé à la taille de l'image entière.

1.8 Conclusion

Un nouvel algorithme hiérarchique Bayésien de démixage pour les images hyperspectrales a été développé dans ce chapitre. Cet algorithme se base sur le Normal Compositional Model introduit par Eismann et Stein [ES07]. L'algorithme proposé génère des échantillons distribués selon la loi *a posteriori* jointe des abondances, des variances des endmembers et d'un hyperparamètre. Ces échantillons ont ensuite été utilisés pour estimer les paramètres d'intérêt. Cet algorithme présente de nombreux avantages vis à vis des algorithmes basés sur le MML. Il permet notamment d'étendre le modèle standard à un cas où les spectres des endmembers ont des variances différentes. Les résultats de simulations sur des données synthétiques et réelles ont démontré l'efficacité de ce modèle et de la méthode d'estimation associée.

CHAPITRE 2

Estimation du nombre de composants purs dans une image hyperspectrale

Sommaire

2.1	Introduction	47
2.2	Modèle Bayésien hiérarchique étendu	49
2.3	Algorithme MCMC à sauts réversibles	51
2.4	Simulations sur des données synthétiques	58
2.5	Démélange spectrale de données réelles	64
2.6	Conclusion	69

2.1 Introduction

Dans le chapitre précédent, nous avons abordé une nouvelle manière d'estimer les abondances des endmembers dans chaque pixel d'une image hyperspectrale. Il nous a paru pertinent d'utiliser un autre modèle de mélange appelé Normal Compositional Model (NCM) [ES07] pour effectuer le démélange spectral. En effet, alors que le modèle linéaire classique présente quelques inconvénients liés à l'absence de pixels purs dans l'image, le NCM compense ces défauts en supposant que les spectres des endmembers sont aléatoires. Les résultats de simulations sur des données synthétiques et réelles ont montré que l'algorithme basé sur ce modèle donnait de meilleurs résultats d'estimation que les méthodes classiques lors de l'absence de pixels purs dans l'image.

La méthode précédente a été mise en œuvre dans un cadre supervisé, c'est-à-dire avec des endmembers connus. Ces endmembers peuvent être obtenus à partir d'une bibliothèque spectrale ou bien extraits à l'aide d'algorithmes d'extraction d'endmembers tels que le N-FINDR développé par Winter [Win99] ou bien l'algorithme Vertex Component Analysis (VCA) proposé par Nascimento *et al.* [NB05] (voir Introduction pour plus de détails). Dans ce chapitre, nous abordons le problème de l'estimation du nombre de ces endmembers (noté R) présents dans un ou plusieurs pixels de l'image hyperspectrale. Il s'agit d'un problème de sélection de modèle¹ puisque chaque valeur de R conduit à un modèle dont le nombre de paramètres dépend de R . Notons qu'une approche similaire a déjà été appliquée au MML dans [DTC08]. L'objectif de ce chapitre est de généraliser cette approche au modèle NCM.

Les moyennes des endmembers sont supposés appartenir à une bibliothèque spectrale. En revanche, contrairement à l'approche précédente, la nature et le nombre des moyennes des endmembers impliqués dans le NCM ne sont pas connus. Pour insister sur la connaissance partielle des endmembers présents dans le mélange, le terme « semi-supervisé » est employé pour désigner l'algorithme développé dans ce chapitre. La méthode proposée est basée sur un modèle Bayésien hiérarchique comme dans le chapitre précédent. Les lois *a priori* identifiés au chapitre 1 sont choisies pour les paramètres et hyperparamètre inconnus du modèle. Cependant, nous définissons une nouvelle loi *a priori* concernant le nombre de matériaux purs R du modèle NCM. La loi jointe *a posteriori* de ces paramètres et hyperparamètres est comme précédemment trop complexe pour en déduire facilement les expressions d'estimateurs Bayésiens classiques. De plus, la dimension du vecteur des abondances est dépendante de R . Pour résoudre à la fois ce problème de complexité et de sélection de modèles, nous proposons d'employer un algorithme à sauts réversibles combiné avec des méthodes de Monte Carlo par chaînes de Markov (RJ-MCMC) [Gre95]. Cet algorithme propose des mouvements entre des espaces de paramètres de différentes

1. Une introduction à ces problèmes de sélection de modèle peut être trouvé dans [Che98]

dimensions, ce qui permet d'estimer le paramètre R . Notons que l'utilisation de sauts-réversibles a déjà fait ses preuves dans de nombreuses applications du traitement du signal et des images comme par exemple en segmentation [PADF02, DTD07], analyse de signaux audio musicaux [DGI06] et en analyse spectrale [AD99].

2.2 Modèle Bayésien hiérarchique étendu

Le modèle Bayésien hiérarchique développé dans le chapitre précédent va être adapté à l'estimation du nombre d'endmembers présents dans l'image. Les expressions de la fonction de vraisemblance et des lois *a priori* des paramètres sont données, avec une attention toute particulière sur le nombre des endmembers R .

2.2.1 Vraisemblance

On pose $\mathbf{M}_R = [\mathbf{m}_1, \dots, \mathbf{m}_R]$ comme étant la matrice contenant les moyennes des vecteurs des endmembers présents dans le mélange dont la taille dépend de la dimension R . Le modèle étant toujours le même (excepté que R et \mathbf{M}_R sont inconnus), la vraisemblance du spectre d'un pixel observé \mathbf{y} peut s'écrire comme :

$$f(\mathbf{y} | \mathbf{a}^+, w^2, R, \mathbf{M}_R) = \frac{1}{[2\pi w^2 c(\mathbf{a}^+)]^{\frac{L}{2}}} \exp\left(-\frac{\|\mathbf{y} - \mu(\mathbf{a}^+)\|^2}{2w^2 c(\mathbf{a}^+)}\right) \quad (2.1)$$

où l'on rappelle que $\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}}$ est la norme ℓ^2 de \mathbf{x} , $\mathbf{a}^+ = [a_1, \dots, a_R]^T$ et

$$\mu(\mathbf{a}^+) = \sum_{r=1}^R \mathbf{m}_r a_r, \quad c(\mathbf{a}^+) = \sum_{r=1}^R a_r^2. \quad (2.2)$$

On remarque ainsi qu'à travers les quantités $\mu(\mathbf{a}^+)$ et $c(\mathbf{a}^+)$, les dimensions de \mathbf{a}^+ et \mathbf{M}_R dépendent du paramètre inconnu R .

2.2.2 Lois a priori des paramètres

Grâce aux contraintes de positivité et d'additivité, le vecteur d'abondances \mathbf{a}^+ peut être réécrit de la manière suivante : $\mathbf{a}^+ = [\mathbf{a}^T, a_R]^T$ où $a_R = 1 - \sum_{r=1}^{R-1} a_r$. Les lois a

priori de \mathbf{a} , de w^2 et δ respectivement détaillées dans (1.7), (1.9) et (1.11) sont conservées dans ce chapitre. Les lois de R et de \mathbf{M}_R sont détaillées dans la suite de ce chapitre.

Nombre de matériaux purs

Une loi discrète uniforme sur $\{1, \dots, R_{\max}\}$ est choisie comme loi a priori pour le nombre de matériaux purs du mélange R :

$$P(R = k) = \frac{1}{R_{\max}}, \quad k = 1, \dots, R_{\max} \quad (2.3)$$

où R_{\max} est le nombre maximum de matériaux purs pouvant être présents dans le pixel (par exemple, le nombre de composantes présents dans la bibliothèque spectrale). On peut remarquer que cette loi uniforme ne favorise aucun modèle parmi $\{1, \dots, R_{\max}\}$.

Moyennes des endmembers

Soit une bibliothèque spectrale \mathcal{S} . Toutes les combinaisons des R spectres appartenant à cette bibliothèque sont supposées être équiprobables conditionnellement aux nombres d'endmembers R , soit :

$$P(\mathbf{M}_R = [\mathbf{s}_{i_1}, \dots, \mathbf{s}_{i_R}] | R) = \binom{R_{\max}}{R}^{-1} = \frac{\Gamma(R+1)\Gamma(R_{\max}-R+1)}{\Gamma(R_{\max}+1)} \quad (2.4)$$

où i_1, \dots, i_R est une séquence non-ordonnée de R éléments distincts de $\{1, \dots, R_{\max}\}$.

2.2.3 Loi *a posteriori* des paramètres

La loi jointe *a posteriori* du vecteur des paramètres inconnus $\boldsymbol{\theta} = \{\mathbf{a}, w^2, \mathbf{M}_R, R\}$ peut être calculée comme suit :

$$f(\boldsymbol{\theta}, \delta | \mathbf{y}) \propto f(\mathbf{y} | \mathbf{a}, w^2, \mathbf{M}_R, R) f(\mathbf{a} | R) P(\mathbf{M}_R | R) P(R) f(w^2 | \delta) f(\delta), \quad (2.5)$$

ce qui mène à :

$$f(\boldsymbol{\theta}, \delta | \mathbf{y}) \propto \frac{\Gamma(R+1)\Gamma(R_{\max}-R+1)}{R_{\max}\Gamma(R_{\max}+1)} \frac{1}{w^{2R+L} [c(\mathbf{a})]^{L/2}} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2w^2 c(\mathbf{a})} - \delta\right) \mathbf{1}_{\mathcal{S}}(\mathbf{a}) \mathbf{1}_{\mathbb{R}^+}(\delta). \quad (2.6)$$

Deux points sont à noter à la suite de cette relation. Tout d'abord, le vecteur des paramètres inconnus a une dimension inconnue dépendant de R : $\boldsymbol{\theta} \in \mathbb{S}_R \times \mathbb{R}^+ \times \{1, \dots, R_{\max}\} \times \mathcal{S}^R$ où \mathbb{S}_R dénote le simplexe (1.6) à R sommets :

$$\mathbb{S}_R = \left\{ \mathbf{a} \in \mathbb{R}^{R-1} \left| a_r \geq 0, \forall r = 1, \dots, R-1, \sum_{r=1}^{R-1} a_r \leq 1 \right. \right\}. \quad (2.7)$$

Ensuite, la loi *a posteriori* (2.6) est trop complexe pour pouvoir déterminer analytiquement les estimateurs MMSE ou MAP de $\boldsymbol{\theta}$. L'utilisation de méthodes MCMC permet de contourner le problème de la complexité de la loi *a posteriori*, comme ce fut le cas pour 1.3. Ces méthodes peuvent s'appliquer à des problèmes de sélection de modèle (résultant de la non-connaissance du paramètre R) en utilisant des sauts réversibles (« reversible jumps ») dont les bases ont été posées par Green dans [Gre95]. La méthode permet alors la génération d'échantillons asymptotiquement distribués selon $f(\boldsymbol{\theta}, \delta|\mathbf{y})$.

2.3 Algorithme MCMC à sauts réversibles

L'algorithme décrit dans ce paragraphe génère des échantillons distribués suivant la loi *a posteriori* $f(\mathbf{a}, w^2, \mathbf{M}_R, R, \delta|\mathbf{y})$. Les vecteurs à échantillonner appartiennent à un espace dont la dimension dépend de R , ce qui requiert l'utilisation d'une stratégie d'ajustement de dimensions comme décrit dans [RG97, RG98]. Plus précisément, l'algorithme proposé se décompose en 4 étapes (voir ALGO. 2.1) :

1. mise à jour des moyennes des endmembers présents dans la matrice \mathbf{M}_R ,
2. mise à jour du vecteur partiel d'abondances \mathbf{a} ,
3. mise à jour de la variance w^2 ,
4. mise à jour de l'hyperparamètre δ .

Ces mouvements possibles sont réalisés systématiquement comme dans [RG97] et sont décrits ci-dessous.

ALGO. 2.1 Échantillonneur de Metropolis-within-Gibbs hybride pour le démélange spectral semi-supervisé selon le NCM

– Initialisation :

- Échantillonner le paramètre $R^{(0)}$,
- Choisir $R^{(0)}$ spectres dans la bibliothèque \mathcal{S} pour construire $\mathbf{M}_{R^{(0)}}^{(0)}$,
- Échantillonner les paramètres $w^{2(0)}$ et $\mathbf{a}^{(0)}$,

– Itérations : Pour $t = 1, 2, \dots$, faire

– Mise à jour de $\mathbf{M}_{R^{(t)}}^{(t)}$:

tirer $u_1 \sim \mathcal{U}_{[0,1]}$,

si $u_1 \leq b_{R^{(t-1)}}$ **alors**

proposer un mouvement de *naissance* (cf ALGO. 2.2),

sinon si $b_{R^{(t-1)}} < u_1 \leq b_{R^{(t-1)}} + d_{R^{(t-1)}}$ **alors**

proposer un mouvement de *mort* (cf ALGO. 2.3),

sinon si $u_1 > b_{R^{(t-1)}} + d_{R^{(t-1)}}$ **alors**

proposer un mouvement d'*échange* (cf ALGO. 2.4),

fin si

tirer $u_2 \sim \mathcal{U}_{[0,1]}$,

si $u_2 < \rho$ (cf Eq. (2.13)) **alors**

poser $(\mathbf{a}^{(t)}, w^{2(t)}, \mathbf{M}_{R^{(t)}}^{(t)}, R^{(t)}) = (\mathbf{a}^*, w^{2*}, \mathbf{M}_{R^*}^*, R^*)$,

sinon

poser $(\mathbf{a}^{(t)}, w^{2(t)}, \mathbf{M}_{R^{(t)}}^{(t)}, R^{(t)}) = (\mathbf{a}^{(t-1)}, w^{2(t-1)}, \mathbf{M}_{R^{(t-1)}}^{(t-1)}, R^{(t-1)})$,

fin si,

- Échantillonner $\mathbf{a}^{(t)}$ d'après la loi Eq. (2.15),
 - Échantillonner $w^{2(t)}$ d'après la loi Eq. (2.17) ,
 - Échantillonner $\delta^{(t)}$ d'après la loi Eq. (2.18) ,
-

2.3.1 Mise à jour de la matrice des endmembers \mathbf{M}_R

Trois types de mouvements, appelés mouvement de « Naissance », « Mort » et « Échange » (comme dans [DTC08]), permettent de mettre à jour les moyennes des spectres des endmembers impliqués dans le mélange. Les deux premiers mouvements consistent en l'augmentation ou la diminution du nombre de composants purs R de 1. Ainsi, ces mouvements nécessitent l'utilisation de méthodes MCMC à sauts réversibles introduites par Green [Gre95]. Dans le troisième mouvement, la dimension de R demeure inchangée, ce qui requiert l'utilisation d'une procédure d'acceptation-rejet standard de Metropolis-Hastings. On suppose qu'à l'itération t , le modèle est défini par $(\mathbf{a}^{(t)}, \mathbf{M}_{R^{(t)}}^{(t)}, R^{(t)})$. Les mouvements de « Naissance », « Mort » et « Échange » sont définis comme suit (les algorithmes détaillés pas à pas sont respectivement décrits dans ALGO. 2.2, ALGO. 2.3 et ALGO. 2.4) :

- naissance : un mouvement de *naissance* $R^* = R^{(t)} + 1$ est proposé avec une probabilité $b_{R^{(t)}}$. Un nouveau spectre \mathbf{s}^* est aléatoirement choisi parmi les moyennes des endmembers disponibles dans la librairie \mathcal{S} afin de construire $\mathbf{M}_{R^*}^* = [\mathbf{M}_{R^{(t)}}^{(t)}, \mathbf{s}^*]$. Comme le nombre de matériaux purs a augmenté de 1, un nouveau vecteur d'abondances \mathbf{a}^{+*} est proposé suivant une règle inspirée de [RG97]. Cependant, en cas d'absence de loi de proposition pour w^2 , de mauvais rapports d'acceptation-rejet peuvent survenir lors du choix du modèle. C'est pourquoi une nouvelle variance w^{2*} est proposée suivant une loi inverse Gamma (même loi que la loi *a priori* de w^2) de paramètres α_w et β_w , tels que $w^{2*} \sim \mathcal{IG}(\alpha_w, \beta_w)$. Ces paramètres ont été choisis tels que la moyenne et la variance de w^{2*} soient égales à :

$$\mathbb{E}[w^{2*}] = w^{2(t)} \frac{\sum_{r=1}^R a_r^{2(t)}}{\sum_{r=1}^R a_r^{2*}}, \quad \text{Var}(w^{2*}) = D \times w^{2(t)}, \quad (2.8)$$

où D est un hyperparamètre.

- mort : un mouvement de *mort* $R^* = R^{(t)} - 1$ est proposé avec la probabilité $d_{R^{(t)}}$. Un des spectres de $\mathbf{M}_{R^{(t)}}^{(t)}$ est retiré, ainsi que son abondance correspondante. Les coefficients d'abondance restants sont re-normalisés pour sommer à 1. On propose

ALGO. 2.2 Mouvement de naissance

- 1: poser $R^* = R^{(t)} + 1$,
- 2: choisir \mathbf{s}^* parmi \mathcal{S} tel que $\mathbf{s}^* \neq \mathbf{m}_r^{(t)}$, $r = 1, \dots, R^{(t)}$,
- 3: poser $\mathbf{M}_{R^*}^* = [\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)}, \mathbf{s}^*]$,
- 4: tirer $\omega^* \sim \mathcal{Be}(1, R^{(t)})$,
- 5: ajouter ω^* à $\mathbf{a}^{+(t)}$ et re-normaliser les autres coefficients tels que

$$\mathbf{a}^{+*} = \left[\frac{a_1^{(t)}}{C}, \dots, \frac{a_{R^{(t)}}^{(t)}}{C}, \omega^* \right]^T$$

avec $C = \frac{1}{(1-\omega^*)}$,

- 6: tirer $w^{2*} \sim \mathcal{IG}(\alpha_w, \beta_w)$.
-

également une nouvelle variance w^{2*} suivant une loi inverse Gamma de paramètres α_w et β_w (identiques à ceux choisis précédemment),

ALGO. 2.3 Mouvement de mort

- 1: poser $R^* = R^{(t)} - 1$,
- 2: tirer $j \sim \mathcal{U}_{\{1, \dots, R^{(t)}\}}$,
- 3: supprimer $\mathbf{m}_j^{(t)}$ de $\mathbf{M}_{R^{(t)}}^{(t)}$, soit poser

$$\mathbf{M}_{R^*}^* = [\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{(j-1)}^{(t)}, \mathbf{m}_{(j+1)}^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)}],$$

- 4: supprimer $a_j^{(t)}$ de $\mathbf{a}^{+(t)}$ et remise à l'échelle des coefficients restants, soit poser

$$\mathbf{a}^{+*} = \left[\frac{a_1^{(t)}}{C}, \dots, \frac{a_{(j-1)}^{(t)}}{C}, \frac{a_{(j+1)}^{(t)}}{C}, \dots, \frac{a_{R^{(t)}}^{(t)}}{C} \right]^T,$$

avec $C = \sum_{r \neq j} a_r^{(t)}$.

- 5: tirer $w^{2*} \sim \mathcal{IG}(\alpha_w, \beta_w)$.
-

– échange : un mouvement d'échange est proposé avec la probabilité $u_{R^{(t)}}$. Un spectre aléatoirement choisi dans $\mathbf{M}_{R^{(t)}}^{(t)}$ est remplacé par un autre spectre choisi aléatoirement dans la bibliothèque \mathcal{S} .

ALGO. 2.4 Mouvement d'échange

-
- 1: **si** $R_{\max} - R^{(t)} \neq 0$ **alors**
 - 2: tirer $j \sim \mathcal{U}_{\{1, \dots, R^{(t)}\}}$,
 - 3: choisir \mathbf{s}^* dans \mathcal{S} tels que $\mathbf{s}^* \neq \mathbf{m}_r^{(t)}$, $r = 1, \dots, R^{(t)}$,
 - 4: remplacer $\mathbf{m}_j^{(t)}$ dans $\mathbf{M}_{R^{(t)}}^{(t)}$ par \mathbf{s}^* , soit

$$\mathbf{M}_{R^{(t)}}^{(t)} = \left[\mathbf{m}_1^{(t)}, \dots, \mathbf{m}_{(j-1)}^{(t)}, \mathbf{s}^*, \mathbf{m}_{(j+1)}^{(t)}, \dots, \mathbf{m}_{R^{(t)}}^{(t)} \right],$$

- 4: poser $\mathbf{a}^{+*} = \mathbf{a}^{+(t)}$ et $R^* = R^{(t)}$.
 - 5: **fin si**
-

À chaque itération, un de ces mouvements est aléatoirement choisi avec les probabilités $b_{R^{(t)}}, d_{R^{(t)}}$ et $u_{R^{(t)}}$. Ces probabilités doivent respecter trois conditions² :

- $b_{R^{(t)}} + d_{R^{(t)}} + u_{R^{(t)}} = 1$,
- $d_1 = 0$ (un mouvement de *mort* est interdit pour $R = 1$),
- $b_{R_{\max}} = 0$ (un mouvement de *naissance* est impossible pour $R = R_{\max}$),

Ainsi, $b_{R^{(t)}} = d_{R^{(t)}} = u_{R^{(t)}} = \frac{1}{3}$ pour $R \in \{2, R_{\max} - 1\}$ et $b_1 = d_{R_{\max}} = u_1 = u_{R_{\max}} = \frac{1}{2}$.

La loi a priori du vecteur d'abondance \mathbf{a} est une loi uniforme sur le simplexe \mathbb{S}_R défini dans (2.7), ce qui est équivalent à choisir une loi de Dirichlet $\mathcal{D}_R(1, \dots, 1)$ comme loi a priori pour le vecteur \mathbf{a}^+ . Ainsi, la probabilité d'acceptation pour le mouvement de *naissance* est $\rho = \min\{1, A_b\}$ où :

$$A_b = \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{\frac{L}{2}} \frac{d_{R^{(t)}+1}}{b_{R^{(t)}}} \frac{R^{(t)}}{g_{1, R^{(t)}}(\omega^*)} (1 - \omega^*)^{R^{(t)}-1} \frac{h_{\alpha_{w^*}, \beta_{w^*}}(w^{2(t)})}{h_{\alpha_{w^{(t)}}, \beta_{w^{(t)}}}(w^{2*})} \times \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{2c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{2c(\mathbf{a}^*)} \right], \quad (2.9)$$

où $g_{a,b}(\cdot)$ et $h_{a,b}(\cdot)$ représentent respectivement les densités de probabilités des lois Beta $\mathcal{B}e(a, b)$ et inverse Gamma $\mathcal{IG}(a, b)$ (voir Annexe B pour le détail des calculs).

Deux cas doivent être abordés pour la probabilité d'acceptation du mouvement de *mort* $\rho = \min\{1, A_d\}$. Si l'ordre du modèle à l'instant t est tel que $R^{(t)} \neq 2$, la probabilité

2. Le cas où $R = 1$ a été pris en compte pour l'algorithme semi-supervisé puisqu'un pixel peut être spectralement pur.

d'acceptation est :

$$A_d = \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{2c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{2c(\mathbf{a}^*)} \right] \\ \times \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{\frac{L}{2}} \frac{b_{R^{(t)}-1}}{d_R^{(t)}} \frac{g_{1,R^{(t)}-1}(\mathbf{a}_{R^{(t)}}^*)}{R^{(t)}-1} \frac{h_{\alpha_{w^*}, \beta_{w^*}}(w^{2(t)})}{h_{\alpha_{w^{(t)}}, \beta_{w^{(t)}}}(w^{2*})} (1 - \alpha_{R^{(t)}}^*)^{R^{(t)}-2}. \quad (2.10)$$

Quand $R^{(t)} = 2$, le mouvement de mort donne $R^* = 1$, ce qui implique qu'un unique élément pur est présent dans le modèle proposé, ce qui mène évidemment à $\mathbf{a}^* = 1$. Ainsi, \mathbf{a}^* est déterministe et le taux d'acceptation est :

$$A_d = \frac{f(\mathbf{M}_{R^*}^*, R^* | \mathbf{y})}{f(\mathbf{M}_{R^{(t)}}^{(t)}, R^{(t)} | \mathbf{y})} \frac{p_{R^* \rightarrow R^{(t)}}}{p_{R^{(t)} \rightarrow R^*}} \frac{q(\mathbf{M}_{R^{(t)}}^{(t)}, w^{2(t)} | \mathbf{M}_{R^*}^*, w^{2*})}{q(\mathbf{M}_{R^*}^*, w^{2*} | \mathbf{M}_{R^{(t)}}^{(t)}, w^{2(t)})}. \quad (2.11)$$

Dans ce cas précis, le Jacobien de la transformation est égal à 1 étant donné que \mathbf{a}^* est indépendant de $\mathbf{a}^{(t)}$. Par conséquent, l'équation précédente donne

$$A_d = \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{L/2} \frac{b_{R^{(t)}-1}}{d_R^{(t)}} \frac{h_{\alpha_{w^*}, \beta_{w^*}}(w^{2(t)})}{h_{\alpha_{w^{(t)}}, \beta_{w^{(t)}}}(w^{2*})} \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{2c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{2c(\mathbf{a}^*)} \right]. \quad (2.12)$$

En ce qui concerne le mouvement d'échange, la probabilité d'acceptation est le rapport classique de Metropolis-Hastings $\rho = \min\{1, A_s\}$ avec :

$$A_s = \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{L/2} \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{2c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{2c(\mathbf{a}^*)} \right]. \quad (2.13)$$

On peut noter que le rapport des lois de proposition lié au mouvement d'échange est égal à 1, puisque dans chaque direction la probabilité de sélectionner un spectre de la bibliothèque est de $1/(R_{\max} - R^{(t)})$. Une fois que la matrice des endmembers \mathbf{M}_R est obtenue, les abondances, la variance des endmembers et l'hyperparamètre δ sont générés suivant les lois détaillées dans les paragraphes suivant.

2.3.2 Loi conditionnelle de \mathbf{a}

De la même manière qu'au chapitre précédent, la loi conditionnelle de \mathbf{a} et des autres paramètres s'obtient à partir de l'équation de Bayes, soit :

$$f(\mathbf{a}|\mathbf{y}, R, w^2, \mathbf{M}_R) \propto f(\mathbf{y}|\boldsymbol{\theta})f(\mathbf{a}) \quad (2.14)$$

ce qui donne facilement :

$$f(\mathbf{a}|\mathbf{y}, R, w^2, \mathbf{M}_R) \propto \frac{1}{[w^2c(\mathbf{a})]^{L/2}} \exp\left(-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2w^2c(\mathbf{a})}\right) \mathbf{1}_{\mathbb{S}(\mathbf{a})}. \quad (2.15)$$

La génération d'un vecteur \mathbf{a} suivant cette loi sera faite à l'aide d'une étape de Metropolis-within-Gibbs avec une marche aléatoire, comme précédemment pour (1.15).

2.3.3 Loi conditionnelle de w^2

On peut déterminer la loi conditionnelle de la variance w^2 de la manière suivante :

$$f(w^2|\mathbf{y}, R, \mathbf{a}, \mathbf{M}_R, \delta) \propto f(\mathbf{y}|\boldsymbol{\theta})f(w^2|\delta). \quad (2.16)$$

Par conséquent, la loi conditionnelle de la variance est la loi inverse Gamma suivante :

$$w^2|\mathbf{y}, R, \mathbf{a}, \mathbf{M}_R, \delta \sim \mathcal{IG}\left(\frac{L}{2} + 1, \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a})\|^2}{2c(\mathbf{a})} + \delta\right). \quad (2.17)$$

2.3.4 Loi conditionnelle de δ

La loi conditionnelle de δ est la loi Gamma suivante :

$$\delta|w^2, R \sim \mathcal{G}\left(R, \frac{R}{w^2}\right) \quad (2.18)$$

où $\mathcal{G}(a, b)$ est la loi Gamma avec le paramètre de forme a et d'échelle b [RC04, p. 581].

2.4 Simulations sur des données synthétiques

La performance de l'algorithme semi-supervisé proposé est mesurée en étudiant le spectre d'un pixel, ou d'un groupe de pixels, généré de la même manière qu'au paragraphe 1.6.1. Dans un premier temps, un pixel est généré avec $R = 3$ endmembers (suivant le NCM) de variances $w^2 = 0.002$ et de vecteur d'abondances $\mathbf{a}^+ = [0.5, 0.3, 0.2]^T$. Les moyennes des endmembers correspondent aux spectres de béton de construction, d'herbe verte et de terre grasse micacée. Les résultats de simulations ont été obtenus pour $N_{MC} = 20000$ itérations, dont $N_{bi} = 1500$ itérations de chauffage, utilisant l'algorithme résumé dans ALGO. 2.1. La bibliothèque spectrale \mathcal{S} utilisée dans cette simulation,

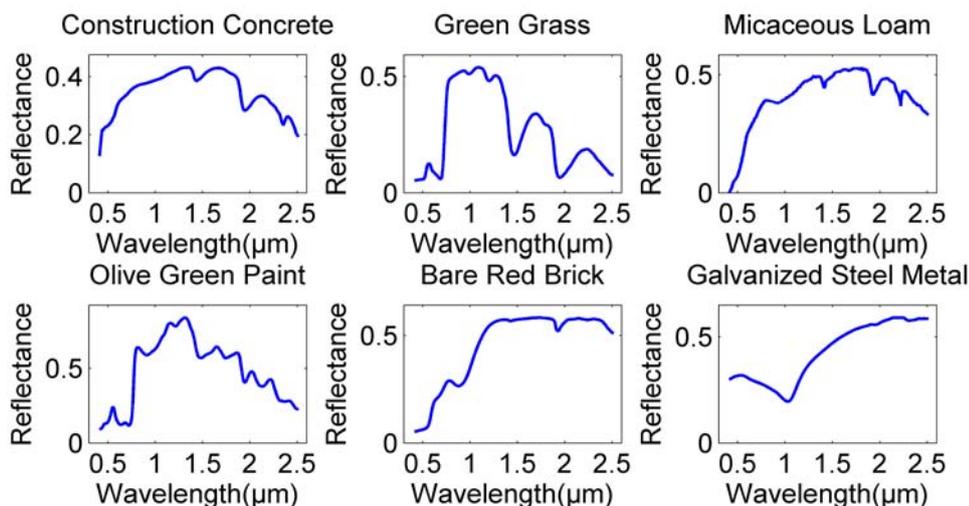


FIGURE 2.1 – Spectres des endmembers de la bibliothèque.

donnée figure 2.1, et extraite du logiciel ENVI [RSI03] contient six spectres : béton de construction, herbe verte, terre grasse micacée, peinture verte olive, brique rouge et acier galvanisé. La première étape de l'analyse consiste en l'estimation de l'ordre du modèle, c'est-à-dire du nombre de composants purs utilisés pour le mélange R . L'histogramme des échantillons $R^{(t)}$ générés est représenté sur la figure 2.2 est clairement en accord avec la valeur réelle de $R = 3$ puisqu'il atteint son maximum pour cette valeur. La seconde étape

de l'analyse estime les probabilités *a posteriori* des combinaisons des spectres conditionnellement à $R = 3$. Pour cette simulation, 100% des combinaisons de spectres générés sont composées des trois premiers spectres de la bibliothèque \mathcal{S} qui sont les spectres réellement employés pour réaliser le mélange. Les lois *a posteriori* des abondances et va-

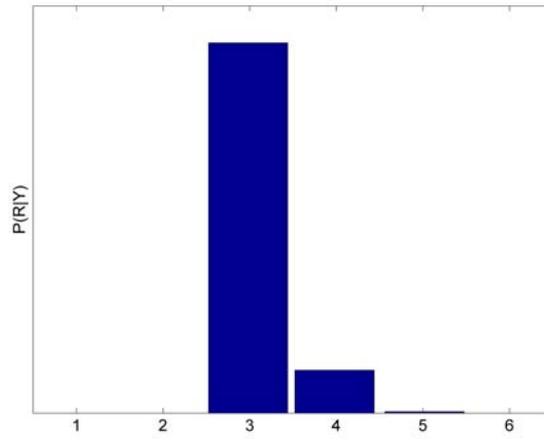


FIGURE 2.2 – Lois conditionnelles estimée du nombre de composants purs R .

riance w^2 correspondantes sont finalement estimées et représentées sur la figure 2.3. Ces distributions sont en accord avec les valeurs réelles de ces paramètres $\mathbf{a}^+ = [0.5 \ 0.3 \ 0.2]^T$ et $w^2 = 0.002$.

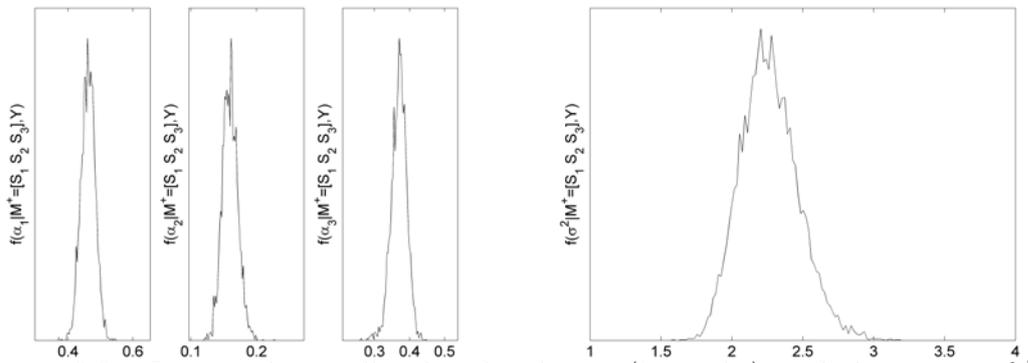


FIGURE 2.3 – Lois conditionnelles des abondances (à gauche) et de la variance w^2 (à droite) estimées.

La robustesse de l'algorithme vis à vis de l'estimation du nombre de composants purs R a été testée pour différents rapports signal sur bruit (SNR) et pour un nombre différent de pixels. Dans un premier temps, un pixel a été généré suivant le MML classique avec $R = 3$ composants purs et des valeurs du SNR variant de 0dB à 20dB. La figure 2.4 montre les estimations MAP du nombre de composants purs estimé, avec les écarts types correspondants, en fonction des valeurs du SNR. Ces estimations ont été obtenues par moyennage de 50 simulations de Monte Carlo. La figure montre qu'au delà d'une certaine valeur du SNR, les estimations sont très précises.

Une autre série de simulations a été faite à SNR fixé et en faisant varier le nombre de pixels à analyser. En effet, l'algorithme semi-supervisé développé permet également de donner une estimation du nombre de composants purs pour un bloc de pixels. Dans cette simulation, $R = 3$ composants purs ont été mélangés pour des blocs de 1×1 , 3×3 et 9×9 pixels, avec un SNR fixé à 15dB. Après 50 simulations de Monte-Carlo de notre algorithme, un histogramme des cinquantes chaînes de Markov résultantes du paramètre R a été effectué pour chaque bloc, représenté sur la figure 2.5. Une telle représentation permet notamment de rendre compte des propriétés de l'algorithme à sauts réversibles. Ces résultats restent en accord avec la valeur théorique de R compte tenu de la proportion d'échantillons $R = 3$ par rapport aux autres valeurs.

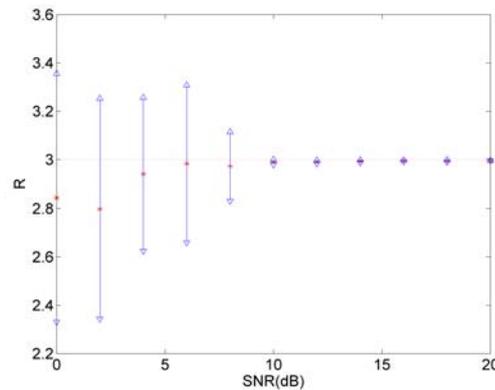


FIGURE 2.4 – Estimations MAP (croix) et écarts-types (barres verticales) de R en fonction du SNR.

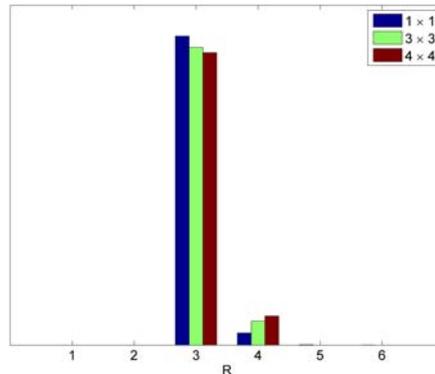


FIGURE 2.5 – Histogrammes des chaînes de Markov du paramètre R générées par 50 simulations de l’algorithme proposé avec différentes tailles de blocs.

2.4.1 Comparaison avec d’autres méthodes de sélection de modèle

L’algorithme Bayésien développé dans ce chapitre est comparé avec six autres méthodes de sélection de modèles, ayant déjà été appliquées à des modèles de mélanges. Dans un premier temps, il est important de noter que toutes les méthodes de sélection de modèles supposent qu’un nombre conséquent de pixels contiennent les mêmes endmembers, contrairement à l’algorithme proposé. De plus, ces méthodes ne fournissent aucune information quant à la nature des endmembers présents dans le mélange observé (ils permettent seulement d’estimer leur nombre R définissant (1.1)). On suppose ici que plusieurs pixels synthétiques partagent les R mêmes spectres des endmembers. Le premier algorithme, présenté dans [LWC⁺07], étudie un estimateur du maximum de vraisemblance (MLE) de la dimension intrinsèque des données observées. Les cinq autres méthodes dépendent principalement de la matrice de covariance des données. Plus précisément, on considère les versions améliorées des critères d’information d’Akaike AIC (*Akaike information criterion*) et de Rissanen MDL (*minimum distance length*) développés au départ dans [Aka73, Aka74], [Ris78] et repris dans [WK85] pour la détection de signaux noyés dans un bruit blanc. Le nouveau critère basé sur AIC construit pour

la théorie des matrices aléatoires (RMT) dans [NE08] (connu sous le nom de *RMT estimator*) et le F-test de Malinowski défendu dans [Mal88] sont également utilisés pour la comparaison. Une technique de réduction de dimension basée sur une ACP décrite dans [LWC⁺07] et très largement utilisée dans la communauté hyperspectrale [KM02] a également été évaluée. Enfin, la méthode HySime développée par Bioucas dans [BDN08] est aussi considérée.

Afin de proposer un estimateur de dimension pertinent pour tous les ensembles de données, un grand nombre de pixels contenant les mêmes R spectres des endmembers doit être considéré. Pour les algorithmes décrits ci-dessus, le nombre d'observations (le nombre de pixels) doit être plus grand que la dimension des observations (le nombre de bandes). Les paramètres de simulations ont été fixés à $L = 207$ bandes et $P = 225$ pixels. Ces pixels ont été simulés suivant le NCM (1.1) avec deux variances différentes ($w^2 = 2.10^{-5}$ et $w^2 = 0.01$) et des nombres différents d'endmembers ($R = 3$, $R = 4$ et $R = 5$). Pour chaque scénario, les vecteurs d'abondances \mathbf{a}^+ de chaque pixel ont été générés suivant une loi Gaussienne tronquée sur le simplexe (1.6) avec les vecteurs moyennes suivant :

- 3 endmembers : $\bar{\mathbf{a}}_{R=3} = [0.4, 0.25, \bar{a}_3]$, $\bar{a}_3 = 1 - \sum_{r=1}^{R-1} \bar{a}_r$,
- 4 endmembers : $\bar{\mathbf{a}}_{R=4} = [0.3, 0.15, 0.2, \bar{a}_4]$, $\bar{a}_4 = 1 - \sum_{r=1}^{R-1} \bar{a}_r$,
- 5 endmembers : $\bar{\mathbf{a}}_{R=5} = [0.3, 0.15, 0.1, 0.1, \bar{a}_5]$, $\bar{a}_5 = 1 - \sum_{r=1}^{R-1} \bar{a}_r$.

Les valeurs de seuil par défaut recommandées dans [LWC⁺07] ont été prises pour la méthode basée sur l'ACP et le test de Malinowski. Le nombre estimé de composants purs R obtenu par la méthode MCMC à sauts réversibles proposée est comparé avec les autres méthodes de sélection de modèles dans le tableau 2.1 pour les différents paramètres de simulations. Pour un bloc de 225 pixels, la méthode proposée donne clairement de meilleurs résultats, tout particulièrement pour de grandes valeurs de w^2 . Ces résultats peuvent être expliqués par le fait que l'algorithme Bayésien proposé tire pleinement avantage de la structure du modèle NCM, contrairement aux autres techniques de sélection de modèles.

TABLE 2.1 – Nombre estimé de composants purs (en %) obtenu par la méthode MCMC à sauts réversibles et par d'autres algorithmes de sélection de modèle avec $w^2 = 1.10^{-2}$ and $w^2 = 2.10^{-5}$.

R estimé	$w^2 = 1.10^{-2}$															$w^2 = 2.10^{-5}$																	
	$R = 3$					$R = 4$					$R = 5$					$R = 3$					$R = 4$					$R = 5$							
	2	3	4	3	4	4	5	4	3	4	5	6	5	4	3	2	3	4	3	4	3	4	5	4	3	2	3	4	3	4	4	5	6
RJ-MCMC	0	100	0	0	100	0	100	0	0	100	0	100	0	0	100	0	100	0	0	100	0	100	0	0	100	0	100	0	0	100	0	100	0
MLE [LB05]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	100	0	100	0	100	0	0	0	100	0	100	0	0	0
AIC [WK85]	0	100	0	0	95	5	45	5	55	0	0	55	0	0	0	0	100	0	0	99	1	0	95	5	0	100	0	0	100	0	95	5	
MDL [WK85]	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	100	0	0	100	0	100	0	100	0	0	100	0	0	0	
RMT [NE08]	32	38	20	1	8	21	0	18	20	20	20	18	20	20	30	34	27	0	6	17	0	12	23	0	100	0	0	100	0	100	0		
F-test [Mal88]	100	0	0	7	0	0	0	0	0	0	0	0	0	0	100	0	0	100	0	0	100	0	100	0	100	0	0	100	0	100	0		
PCA [KM02]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	100	0	0	100	0	83	0	0	0	0	100	0	0	0		
HySime [BDN08]	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		

2.5 Démélange spectrale de données réelles

2.5.1 Première image : « Moffett field »

Ce paragraphe considère l'image hyperspectrale étudiée au chapitre précédent au paragraphe 1.7. Nous rappelons qu'il s'agit d'une imagerie de taille 50×50 pixels extraite d'une plus grande prise par l'imageur AVIRIS en 1997 et que les bandes d'absorption de vapeur d'eau ont été supprimées, donnant ainsi un jeu de données de 189 bandes.

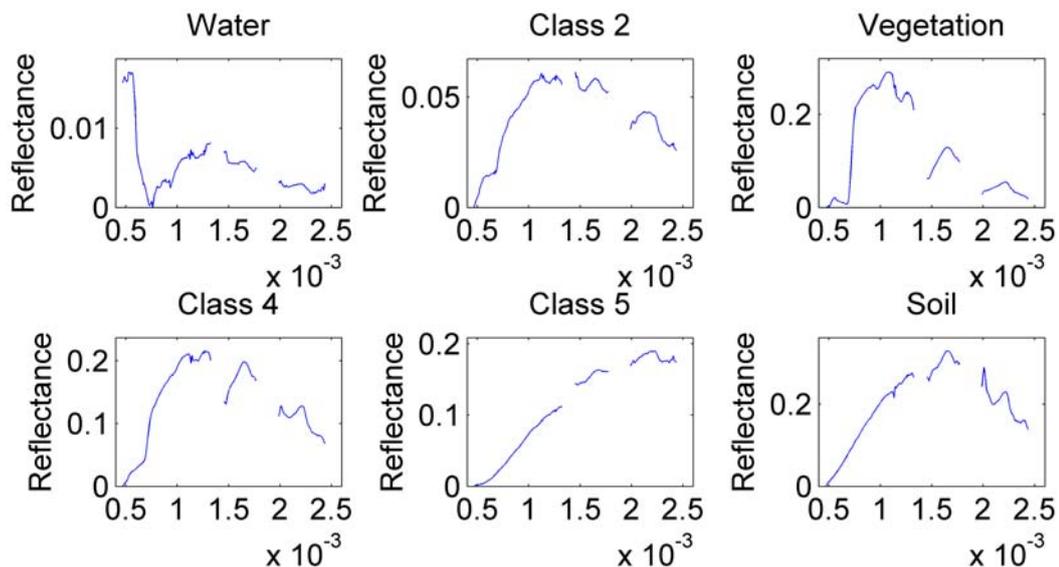


FIGURE 2.6 – Les $R = 6$ spectres des endmembers extraits de l'image « Moffett Field » par l'utilisation de l'algorithme PPI sur les classes trouvées par la procédure K-means.

Comme précisé au paragraphe 2.3, l'algorithme MCMC à saut réversibles requiert l'utilisation d'une bibliothèque spectrale contenant les spectres des endmembers pouvant apparaître dans l'image à analyser. Comme aucune vérité terrain n'est disponible pour cette image AVIRIS, nous avons construit la bibliothèque requise directement à partir de l'image. Dans un premier temps, une procédure K-means a été employé pour classifier l'image en 6 régions. Puis les pixels les plus purs de chaque classe, extraits grâce à l'algorithme PPI (*pixel purity index*) [Boa93], ont été choisis comme spectres de la

bibliothèque. Cette nouvelle bibliothèque spectrale dont les spectres sont représentés sur la figure 2.6 contient six endmembers. Trois d'entre eux ont été clairement identifiés :

- endmember 1 : eau,
- endmember 3 : végétation,
- endmember 6 : terre,

alors que les endmembers 2, 4 et 5 sont plus difficiles à interpréter. L'image « Moffett field » est analysée par l'algorithme proposé, en utilisant cette bibliothèque spectrale. Pour chaque pixel, le vecteur d'abondance est estimé conditionnellement à la matrice des endmembers (et par extension à l'estimateur MAP de R) comme décrit au paragraphe 2.4. Les cartes d'abondances pour chaque endmember et les cartes de présence (avec un seuil de 15% pour les abondances) obtenus par l'algorithme proposé sont donnés sur les figures 2.7 et 2.8, respectivement.

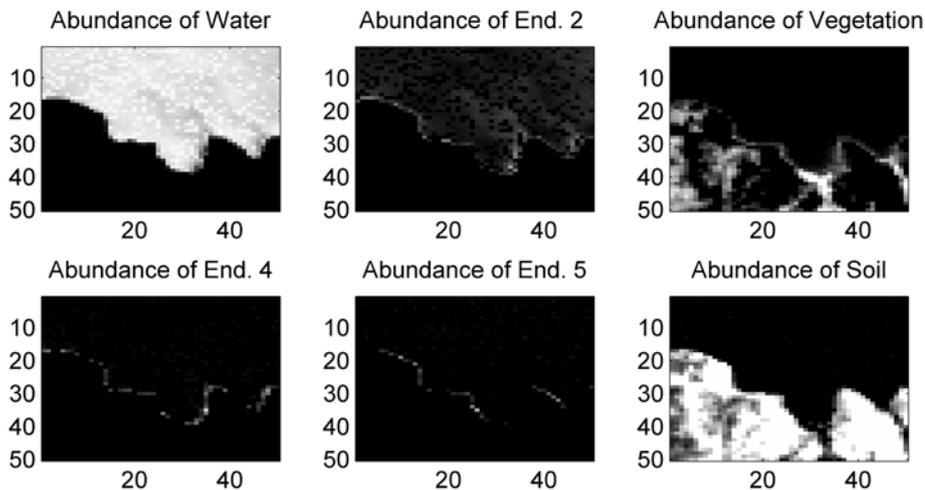


FIGURE 2.7 – Cartes d'abondances estimées par l'algorithme proposé.

La figure 2.9 montre la carte du nombre estimé de matériaux purs dans la zone considérée. La zone du lac comporte au moins deux endmembers comme étant « l'eau » et le spectre non identifié « endmember 2 ». Comme l'énergie du spectre de l'eau est très faible, l'algorithme doit compenser le spectre mesuré de l'eau en ajoutant au moins une

autre contribution, ce qui explique la présence de plus d'un endmember dans cette zone. La zone côtière contient le plus grand nombre de matériaux purs (jusqu'à 6), incluant les endmembers 4 et 5. La zone terrestre est exclusivement composée de terre et de végétation, ce qui est clairement en accord avec les résultats obtenus au paragraphe 1.7 et dans [DTC08].

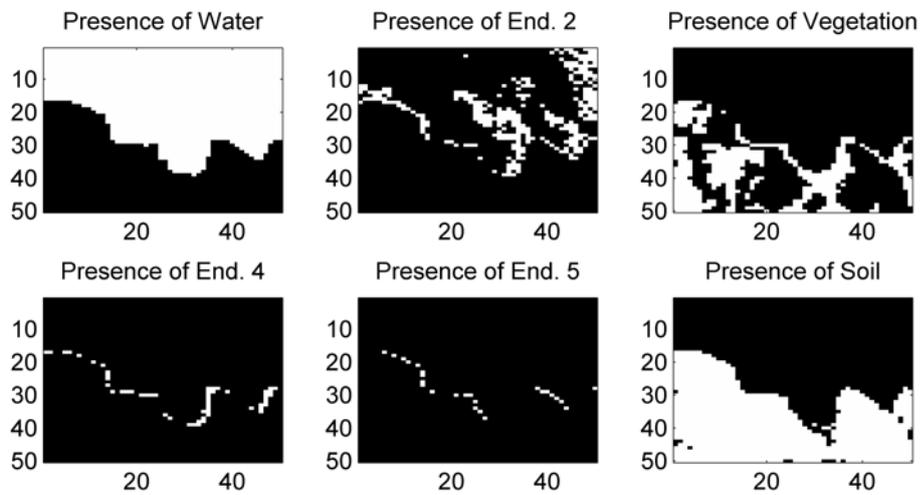


FIGURE 2.8 – Cartes de présences estimées par l'algorithme proposé.

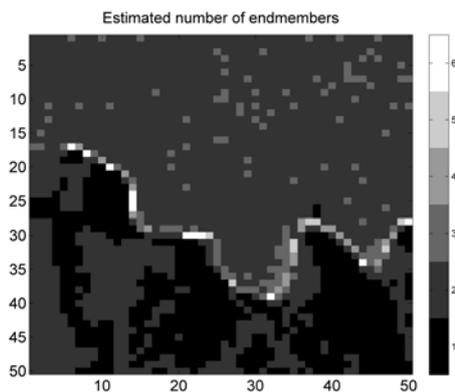


FIGURE 2.9 – Nombre d'endmembers estimé par l'algorithme proposé (zones obscures (resp. claires) signifiant $R = 1$ (resp. $R = 6$)).

2.5.2 Deuxième image : « Cuprite »

Une deuxième image a également été étudiée pour l'évaluation des performances de notre algorithme semi-supervisé. Connue sous le nom de « Cuprite », cette image représente une zone riche en minerais de natures différentes. Elle a également été prise par l'imager AVIRIS en 1997. Les études décrites dans [CSG93, Cla03] ont établi une carte géologique de la région (représentée figure 2.10).

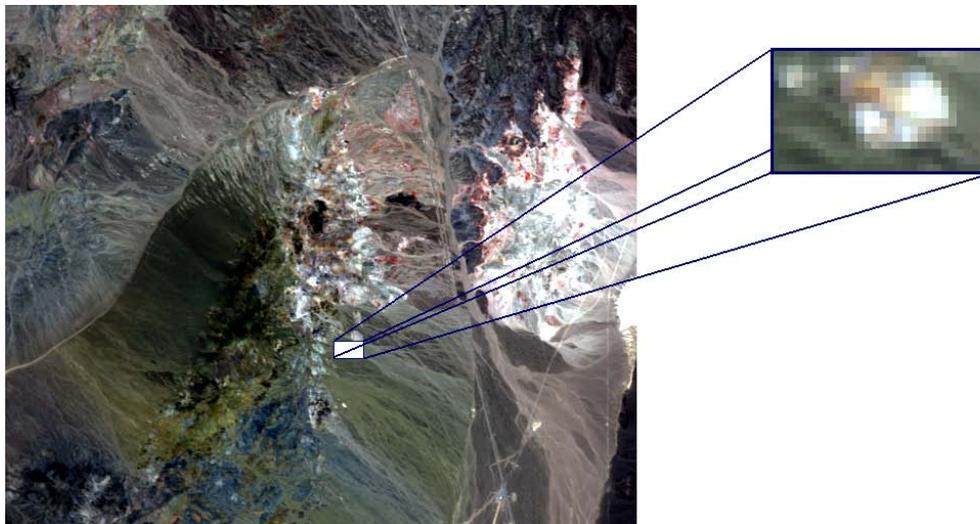


FIGURE 2.10 – Données hyperspectrales réelles : image Cuprite acquise par AVIRIS en 1997 (à gauche) et la région d'intérêt représentée en couleurs réelles (à droite).

Comme ci-dessus, une bibliothèque spectrale de 6 endmembers est construite directement à partir de l'image. Nous avons utilisé la méthode VCA pour extraire les spectres de cette bibliothèque (représentés sur la figure 2.11) car cette méthode a déjà été utilisée pour cette région d'intérêt dans [DMC⁺09].

L'algorithme MCMC à sauts réversibles a été appliqué sur une imagerie de taille 28×16 pixels, extraite de l'image Cuprite. L'imagerie à analyser peut être schématiquement décrite comme un « amas » d'alunite. Le nombre estimé de matériaux purs et les cartes d'abondances sont données figure 2.12 et 2.13.

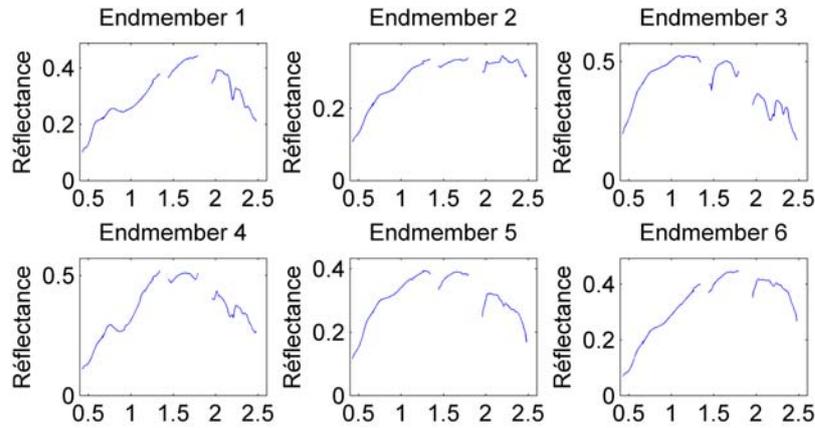


FIGURE 2.11 – Les $R = 6$ spectres des endmembers estimés par l’algorithme VCA appliqué sur l’image Cuprite.

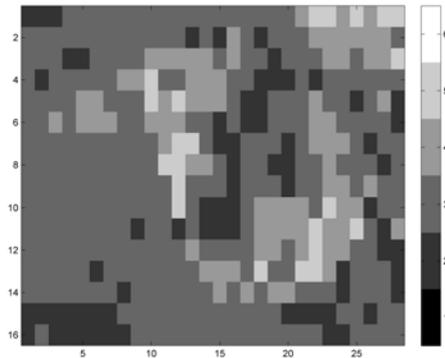


FIGURE 2.12 – Nombre de matériaux purs estimés par l’algorithme proposé (zones obscures (resp. claires) signifiant $R = 1$ (resp. $R = 6$)).

L’« amas » d’alunite a clairement été identifié : jusqu’à cinq endmembers, et surtout les endmembers 3 et 4, sont présents dans cette partie centrale de l’image. La région entourant l’« amas » est principalement composé de l’endmember 1. En comparant ces spectres avec les signatures spectrales extraites de la bibliothèque spectrale de l’*United States Geological Survey* (USGS) [Cla03], les endmembers 1, 3 et 4 peuvent être identifiés comme étant respectivement des minerais de muscovite, alunite et kaolinite. Afin d’illustrer cette dernière analyse, les spectres estimés et leurs signatures correspondantes tirées de la bibliothèque USGS sont donnés sur la figure 2.14. Ces résultats, en accord avec les

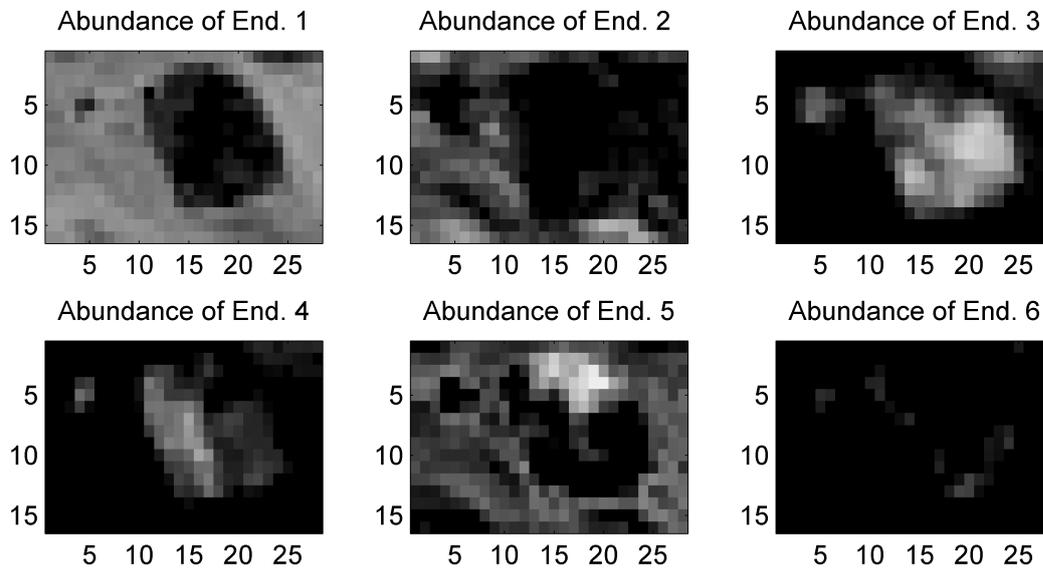


FIGURE 2.13 – Cartes d’abondances de la région d’intérêt estimées par l’algorithme proposé.

résultats rapportés dans [DMC⁺09], confirment la performance correcte de l’algorithme à sauts réversibles.

2.6 Conclusion

Alors que le chapitre précédent a présenté un algorithme Bayésien supervisé de dé-mélange spectral basé sur le NCM (introduit par Eismann et Stein [ES07]), ce chapitre a abordé le cas où le nombre de composants purs R de ce modèle est inconnu. La nature des endmembers impliqués dans le mélange spectral est aussi inconnue. Nous avons supposé que les spectres de ces endmembers appartiennent à une bibliothèque spectrale donnée reflétant la connaissance partielle des paramètres pour le dé-mélange.

La complexité de la forme de la loi *a posteriori* et la dimension variable du vecteur des paramètres inconnus nous a conduit à utiliser une méthode MCMC à sauts réversibles. En plus des générations d’échantillons suivant les différentes lois conditionnelles grâce à

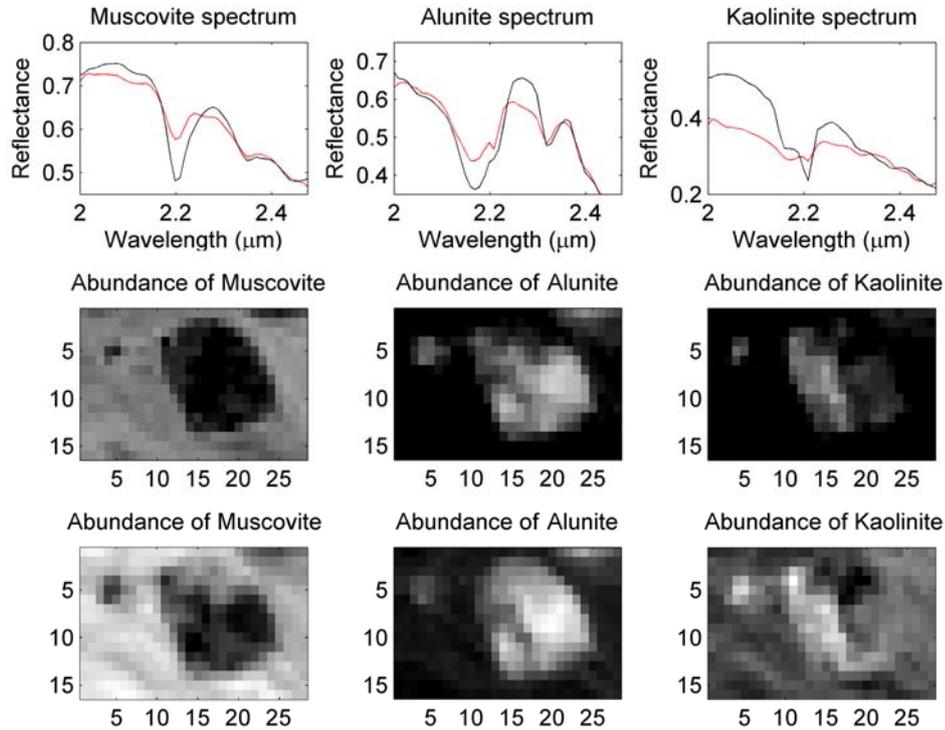


FIGURE 2.14 – Haut : les 3 spectres des endmembers de la bibliothèque récupérés par l’algorithme VCA depuis l’image Cuprite (en rouge), comparés avec les signatures spectrales extraites de la bibliothèque de l’USGS (en noir). Milieu : cartes d’abondances estimées par l’algorithme Bayésien à sauts réversibles. Bas : cartes d’abondances estimées par l’algorithme Bayésien non supervisé étudié dans [DMC⁺09].

la méthode de Metropolis-within-Gibbs, l’algorithme permet d’estimer le nombre de composants purs du mélange et leur nature. Cette nouvelle approche présente de nombreux avantages par rapport aux méthodes standards. En particulier, elle permet de tester quels composants d’une bibliothèque spectrale interviennent dans le mélange associé à un pixel de l’image. Les résultats obtenus sur des données synthétiques et réelles ont démontré les bonnes performances de notre algorithme.

Les algorithmes développés jusqu'à présent n'ont pas pris en compte l'existence de corrélations spatiales entre des pixels voisins de l'image. C'est ce que propose d'aborder le chapitre suivant.

CHAPITRE 3

Démélange spectral avec corrélations spatiales

Sommaire

3.1	Introduction	73
3.2	Formulation du problème	75
3.3	Modèle Bayésien hiérarchique	80
3.4	Échantillonneur de Gibbs hybride	85
3.5	Résultats de simulations sur données synthétiques	88
3.6	Résultats de simulations sur une image AVIRIS	94
3.7	Conclusions	96

3.1 Introduction

Les algorithmes d'inversion développés dans les chapitres précédents ont montré des estimations des coefficients d'abondances très satisfaisantes. Cependant, ces algorithmes comme la plupart des méthodes d'inversion, analysent les pixels de l'image indépendamment. Or une image réelle présente dans bien des cas des zones homogènes (zones de lac, de terres agricoles, etc). L'hypothèse d'indépendance des pixels appartenant à de telles zones paraît donc très éloignée de la réalité. C'est pourquoi nous pensons que l'introduction de corrélations entre les pixels voisins d'une image pourrait sérieusement améliorer les performances d'estimation. Plus précisément l'algorithme développé dans le chapitre 1 va être modifié dans ce sens.

Tout en restant dans un cadre d'estimation Bayésienne, une stratégie de modélisation de contraintes spatiales très utilisée dans la littérature est basée sur les champs de Markov

ou *Markov Random Field* en anglais (MRF). Les champs de Markov ont été introduits et modélisés formellement par Besag dans [Bes74]. En utilisant le théorème de Hammersley-Clifford, il a pu établir une forme générale de la pseudo-vraisemblance des champs de Markov, ouvrant également la possibilité d'estimer leurs paramètres. Les travaux de Geman [GG84] ont popularisé les champs de Markov par le développement d'algorithmes de simulation en exploitant la similarité de ces champs avec la distribution de Gibbs. La communauté du traitement d'image a très largement employé ces modèles pour la segmentation de textures [KH95] ou bien pour des algorithmes de restauration aveugles d'images [TBS06]. Les champs de Markov ont également montré leur utilité pour l'analyse d'images hyperspectrales aussi bien pour des problèmes de segmentation [RK03, MFM04, NS05, RDFZ04] que pour des problèmes de séparation de sources [BMD08]. Un autre approche exploitant les corrélations spatiales a été proposé très récemment pour résoudre pour un problème de démélange linéaire spectral [MH10].

L'image hyperspectrale à analyser est découpée en régions homogènes (ou classes) dans lesquelles les vecteurs d'abondances partagent les mêmes propriétés statistiques du premier et deuxième ordre (moyennes et covariances). Cette hypothèse sous-entend qu'une classification implicite de l'image soit mise en œuvre. On modélise ainsi l'image par des variables cachées (ou étiquettes) dont les dépendances spatiales (par rapport aux étiquettes voisines) sont représentées par un champ de Potts-Markov [Wu82] détaillé plus bas. Conformément à l'approche Bayésienne, une loi *a priori* appropriée sera choisie conditionnellement aux étiquettes pour les vecteurs d'abondances avec des moyennes et des variances inconnues mais dépendants de la classe du pixel. Grâce à une reparamétrisation des vecteurs d'abondances, ces lois *a priori* respectent les contraintes d'additivité et de positivité (2) déjà mentionnées auparavant. En effet, l'introduction de contraintes spatiales nous a conduit à choisir une nouvelle reparamétrisation (similaire à [GBJ96]) afin de pouvoir utiliser une loi *a priori* plus adaptée à un découpage de l'image que la loi utilisée précédemment (uniforme sur le simplex \mathbb{S} , voir (1.7)). Comme toujours dans un cadre Bayésien, la précision de l'estimation des coefficients d'abondance dépend des

hyperparamètres associés. Dans ce chapitre, un second niveau de hiérarchie au sens de l'inférence Bayésienne va être introduit afin de pouvoir estimer tous les hyperparamètres par lesquels des lois a priori non informatives seront choisies. Une fois encore, les méthodes MCMC sont employées pour surmonter les problèmes numériques liés à la forme de la loi *a posteriori*, pour le modèle NCM sous les contraintes spatiales. Ces dernières vont nous permettre d'approcher les estimateurs classiques Bayésiens, eux mêmes donnant une estimation des paramètres (étiquettes et vecteurs d'abondances) et hyperparamètres (moyenne et variance *a priori* des vecteurs d'abondances de chaque classe). Il est à noter que les problèmes de classification et d'estimation sont résolus conjointement.

Le chapitre est organisé comme suit. Le problème de démélange lié au NCM avec les corrélations spatiales est formulé en 3.2. Le paragraphe 3.3 introduit un modèle Bayésien hiérarchique approprié à ces deux problèmes de démélange. L'algorithme MCMC requis pour le calcul des estimateurs liés au modèle Bayésien basé sur le NCM est décrit au paragraphe 3.4. Les résultats de simulations entreprises sur des données synthétiques et réelles sont respectivement présentés aux paragraphes 3.5 et 3.6. En fin de compte, les conclusions de ce chapitre sont présentés au paragraphe 3.7.

3.2 Formulation du problème

3.2.1 Introduction de dépendances spatiales entre les abondances

Étant donné que nous avons choisi d'exploiter les corrélations entre les pixels de l'image, il est intéressant de considérer que les abondances d'un pixel donné sont similaires à celles des pixels voisins. Formellement, on suppose que l'image hyperspectrale est découpée en K régions ou *classes*. Soit $\mathcal{I}_k \subset \{1, \dots, P\}$ le sous-ensemble des indices des pixels appartenant à la k ème classe. Un vecteur d'étiquettes de taille $P \times 1$ noté $\mathbf{z} = [z_1, \dots, z_P]^T$ avec $z_p \in \{1, \dots, K\}$ est considéré pour permettre d'identifier la classe à laquelle chaque pixel p appartient ($p = 1, \dots, P$). En d'autres termes :

$$z_p = k \Leftrightarrow p \in \mathcal{I}_k. \quad (3.1)$$

Dans chaque classe, les vecteurs des abondances à estimer sont supposés aléatoires, comme pour le modèle NCM étudié au chapitre 1. De plus, nous supposons que ces vecteurs partagent les mêmes statistiques du premier et second ordre, c'est-à-dire, $\forall k \in \{1, \dots, K\}$, $\forall (p, p') \in \mathcal{I}_k \times \mathcal{I}_k$

$$\begin{aligned} \mathbb{E}[\mathbf{a}_p] &= \mathbb{E}[\mathbf{a}_{p'}] = \boldsymbol{\mu}_k \\ \mathbb{E}\left[(\mathbf{a}_p - \boldsymbol{\mu}_k)(\mathbf{a}_p - \boldsymbol{\mu}_k)^T\right] &= \mathbb{E}\left[(\mathbf{a}_{p'} - \boldsymbol{\mu}_k)(\mathbf{a}_{p'} - \boldsymbol{\mu}_k)^T\right]. \end{aligned} \quad (3.2)$$

Dès lors, la k ième classe de l'image hyperspectrale à analyser est pleinement caractérisée par la moyenne du vecteur d'abondance $\boldsymbol{\mu}_k$ et la matrice de covariance des abondances des pixels appartenant à cette classe.

3.2.2 Les champs de Markov

Des explications techniques nécessaires à la compréhension et l'utilisation des champs de Markov (MRF) vont être décrites dans ce paragraphe.

Dans un premier temps, pour représenter les contraintes spatiales entre les pixels, il est important de définir convenablement une structure de voisinage. La relation de voisinage entre deux pixels i et j , doit être symétrique : si i est un voisin de j alors j est un voisin de i . Cette relation peut s'appliquer aux plus proches voisins d'un pixel considéré, par exemple les quatre, huit ou bien douze plus proches pixels. La figure 3.1 montre deux exemples de structures de voisinages. La structure à quatre pixels ou *voisinage du premier ordre* est la structure considérée pour le reste du chapitre. L'ensemble des pixels voisins associés, ou *cliques*, n'admet alors que des configurations verticales ou horizontales (voir [Bes74, GG84] pour des détails complémentaires).

Une fois que la structure de voisinage a été clairement établie, les MRF peuvent être facilement définis. Notons z_p une variable aléatoire associée au p ième pixel d'une image de P pixels. Dans notre contexte du démélange hyperspectral, cette variable fait référence à la classe à laquelle le pixel appartient, et prend des valeurs dans un ensemble fini, par exemple dans $\{1, \dots, K\}$. L'ensemble des variables aléatoires $\{z_1, \dots, z_P\}$ forme alors un

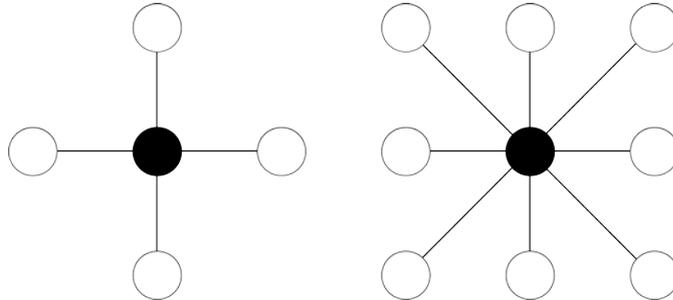


FIGURE 3.1 – Structures de voisinages 4-connexes (à gauche) et 8-connexes (à droite). Le pixel considéré apparaît comme étant un disque noir alors que ses voisins apparaissent en blanc.

champ aléatoire. Un MRF est ensuite défini lorsque la loi conditionnelle de z_i par rapport aux autres pixels \mathbf{z}_{-i} dépend seulement de ses voisins $\mathbf{z}_{\mathcal{V}(i)}$, soit :

$$f(z_i | \mathbf{z}_{-i}) = f(z_i | \mathbf{z}_{\mathcal{V}(i)}). \quad (3.3)$$

Comme dit auparavant, les MRF constituent un outil efficace pour modéliser les corrélations entre des pixels. En effet, en plus de représenter les dépendances spatiales, ils permettent d'ajuster l'influence qu'un pixel étend sur ses voisins. Ainsi, lorsque les données étudiées sont organisées de telles sortes que des « régions » apparaissent (comme les images hyperspectrales), l'état d'un pixel d'une zone sera lié d'une certaine façon à celui de chacun de ses voisins. De plus, les appareils d'acquisition de ce type d'image peuvent introduire des phénomènes de superposition entre des pixels voisins, ce qui justifie l'utilisation des champs de Markov. Il existe deux types de champs de Markov appropriés pour l'analyse d'image [MR07] : le modèle d'Ising pour des variables aléatoires binaires et le modèle de Potts qui est une simple généralisation du modèle d'Ising à plus de deux variables aléatoires [Wu82]. La méthode de démixage que nous développons dans ce chapitre a fait exclusivement appel au modèle de Potts étant donné qu'il est tout à fait adapté à la segmentation d'images hyperspectrales [BMD08, RDFZ04]. Pour un champ aléatoire

discret \mathbf{z} associé à une image de P pixels, le théorème de Hammersley-Clifford donne :

$$f(\mathbf{z}) = \frac{1}{G(\beta)} \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \quad (3.4)$$

où β est le coefficient de *granularité*, $G(\beta)$ est la constante de normalisation ou *fonction de partition* [KS80] et $\delta(\cdot)$ est la fonction de Kronecker

$$\delta(x) = \begin{cases} 1, & \text{si } x = 0, \\ 0, & \text{sinon.} \end{cases}$$

Il est possible de générer un vecteur d'étiquettes $\mathbf{z} = [z_1, \dots, z_P]$ à partir de la loi (3.4) sans la connaissance de la fonction de partition $G(\beta)$ en utilisant un échantillonneur de Gibbs (l'algorithme détaillé est présenté dans ALGO. 3.1). Cependant, un problème majeur de la loi (3.4) provient de $G(\beta)$ qui n'a aucune expression analytique et dépend du paramètre inconnu β . Cet hyperparamètre règle le degré d'homogénéité de chaque région de l'image. Des simulations ont été entreprises afin de montrer l'influence de ce paramètre sur l'homogénéité des régions de l'image. Des images synthétiques ont été générées à partir d'un modèle de Potts-Markov avec $K = 3$ (correspondant à trois niveaux de gris dans l'image) et un voisinage du premier ordre. La figure 3.2 indique qu'une petite valeur de β donne une image « bruitée » avec un très grand nombre de régions, contrairement à une grande valeur de β menant à peu de régions mais de grandes tailles et homogènes. Il n'est pas nécessaire d'étudier des cas où $\beta > 2$ puisque selon [MR07, p. 237] l'image sera vraisemblablement constituée d'une seule région. Nous considérons pour cette étude que le coefficient de granularité β sera fixé *a priori*. Il faut noter cependant que l'estimation de ce coefficient a fait l'objet de nombreuses études et peut par conséquent être effectuée suivant les méthodes décrites dans [ZLQ97], [DMZB99] ou [CFP03].

3.2.3 Reparamétrisation des coefficients d'abondances

Jusqu'à présent dans un cadre Bayésien, les contraintes (1.2) du vecteur d'abondances \mathbf{a} étaient représentées par une loi *a priori* uniforme distribuée sur un simplexe approprié (satisfaisant les contraintes) comme dans [DTC08, DMC⁺09] et les chapitres 1 et

ALGO. 3.1 Simulation des MRF

- 1: **pour** $n = 1$ à N_{MC} **faire**
- 2: **pour** $p = 1$ à P **faire**
- 3: **pour** $k = 1$ à K **faire**
- 4: Calculer $\omega_k \propto P[z_p = k | \mathbf{z}_{-p}]$ selon (3.4)
- 5: **fin pour**
- 6: Calculer la constante de normalisation :

$$G(\beta) = \sum_{k=1}^K \omega_k$$

- 7: Définir le vecteur de probabilité :

$$\tilde{\omega} = \left[\frac{\omega_1}{G(\beta)}, \dots, \frac{\omega_K}{G(\beta)} \right]$$

- 8: Générer z_p dans $\{1, \dots, K\}$ avec les probas. $\{\tilde{\omega}_1, \dots, \tilde{\omega}_K\}$.
 - 9: **fin pour**
 - 10: **fin pour**
-

2. Toutefois, cette loi uniforme définie sur le simplexe n'est pas adaptée pour définir des caractéristiques qui permettent de discriminer les différentes zones de l'image. C'est pourquoi nous avons considéré une reparamétrisation pour des paramètres positifs et sous contraintes d'additivité, introduite dans [GBJ96] pour un problème de pharmacocinétique et appliquée à l'imagerie hyperspectrale dans [TR08]. Elle consiste en une réécriture des abondances en fonction de variables aléatoires que l'on dénommera *coefficients logistiques* dans le reste du chapitre. Un vecteur de coefficients logistiques $\mathbf{t}_p = [t_{1,p} \dots, t_{R,p}]^T$ est affecté à chaque vecteur d'abondance a_p selon la relation suivante :

$$a_{r,p} = \frac{\exp(t_{r,p})}{\sum_{r=1}^R \exp(t_{r,p})}. \quad (3.5)$$

Au départ, les dépendances spatiales résultant du découpage de l'image décrites au paragraphe 3.2 sont basées sur les statistiques du premier et du second ordre du vecteur d'abondances a_p . Toutefois, les contraintes spatiales définies dans (3.2) peuvent être facilement adaptées en utilisant les coefficients logistiques. En effet, dans chacune des classes,

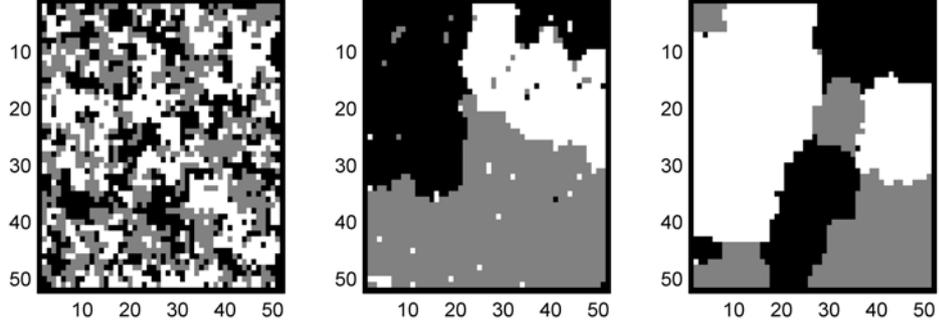


FIGURE 3.2 – Images synthétiques générés à partir d’un modèle de Potts-Markov avec (de gauche à droite) $\beta = 0.8, 1.4, 2$.

les vecteurs des coefficients logistiques inconnus sont supposés partager les mêmes statistiques du premier et deuxième ordres, soit, au sein d’une région k et pour un pixel $p \in \mathcal{I}_k$:

$$\begin{aligned} \mathbb{E}[\mathbf{t}_p] &= \mathbb{E}[\mathbf{t}_{p'}] = \boldsymbol{\psi}_k \\ \mathbb{E}[(\mathbf{t}_p - \boldsymbol{\psi}_k)(\mathbf{t}_p - \boldsymbol{\psi}_k)^T] &= \mathbb{E}[(\mathbf{t}_{p'} - \boldsymbol{\psi}_k)(\mathbf{t}_{p'} - \boldsymbol{\psi}_k)^T] \\ &= \boldsymbol{\Sigma}_k. \end{aligned} \quad (3.6)$$

Avec cette reparamétrisation, la k ième classe est pleinement caractérisée par les hyperparamètres inconnus $\boldsymbol{\psi}_k$ et $\boldsymbol{\Sigma}_k$.

3.3 Modèle Bayésien hiérarchique

Ce paragraphe décrit les vraisemblances et les lois *a priori* inhérentes au NCM en généralisant l’algorithme développé au chapitre 1. On rappelle que la forme du modèle NCM est donné dans l’équation (1.1). Toutefois, il peut être également intéressant de généraliser l’algorithme basé sur le modèle standard [DTC08] pour des cas tenant compte des corrélations entre les pixels de l’image. Le modèle ainsi proposé est développé dans l’annexe C.

3.3.1 Vecteurs des paramètres inconnus

Dans un premier temps, nous définissons convenablement le vecteur des paramètres inconnus associé au NCM : $\Theta = \{\mathbf{T}, \mathbf{z}, \mathbf{w}\}$, où $\mathbf{w} = [w_1^2, \dots, w_P^2]^T$ est le vecteur des variances du bruit et des endmembers, \mathbf{z} est le vecteur des étiquettes et $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_P]$ avec $\mathbf{t}_p = [\mathbf{t}_{1,p}, \dots, \mathbf{t}_{R,p}]^T$ ($p = 1, \dots, P$) est la matrice de coefficients logistiques employés pour la reparamétrisation des abondances.

3.3.2 Fonction de vraisemblance

Conformément au modèle NCM, chaque endmember $\mathbf{e}_{p,r}$ ($r = 1, \dots, R, p = 1, \dots, P$) suit une loi Gaussienne :

$$\mathbf{e}_{p,r} | w_p^2 \sim \mathcal{N}(\mathbf{m}_r, w_p^2 \mathbf{I}_L). \quad (3.7)$$

de vecteur moyenne \mathbf{m}_r et de matrice de covariance $w_p^2 \mathbf{I}_L$ pouvant changer d'un pixel p à un autre. Dès lors, la vraisemblance d'un pixel observé \mathbf{y}_p pour le modèle NCM peut s'écrire de la façon suivante :

$$f(\mathbf{y}_p | \mathbf{t}_p, w_p^2) = \frac{1}{[2\pi w_p^2 c(\mathbf{t}_p)]^{\frac{L}{2}}} \exp\left[-\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{t}_p)\|^2}{2w_p^2 c(\mathbf{t}_p)}\right], \quad (3.8)$$

avec

$$\boldsymbol{\mu}(\mathbf{t}_p) = \sum_{r=1}^R \mathbf{m}_r a_{r,p}(\mathbf{t}_p), \quad c(\mathbf{t}_p) = \sum_{r=1}^R a_{r,p}^2(\mathbf{t}_p).$$

En supposant l'indépendance *a priori* entre les spectres des pixels observés, la vraisemblance des P pixels de l'image est :

$$f(\mathbf{Y} | \mathbf{T}, \mathbf{w}) = \prod_{p=1}^P f(\mathbf{y}_p | \mathbf{t}_p, w_p^2). \quad (3.9)$$

3.3.3 Lois *a priori* des paramètres

Ce paragraphe présente les lois *a priori* des paramètres inconnus et leurs hyperparamètres associés qui seront utilisées pour le modèle NCM. Le graphe orienté (*directed acyclic graph* en anglais ou DAG) pour les lois *a priori* des paramètres et hyperparamètres pour ce modèle statistique est représenté sur la figure 3.3.

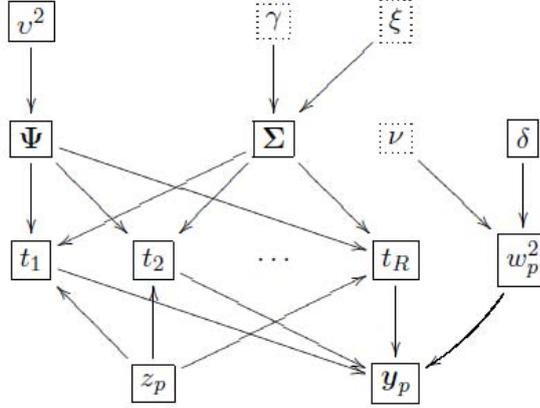


FIGURE 3.3 – DAG pour les lois *a priori* (les paramètres fixés apparaissent dans les cases en pointillés) pour le NCM.

Loi *a priori* des étiquettes

La loi *a priori* du vecteur des étiquettes $\mathbf{z} = [z_1, \dots, z_P]^T$ présentée au paragraphe 3.2.2 est un champ de Potts-Markov avec un voisinage d'ordre 1 et un coefficient de granularité β connu (fixé *a priori*). La loi *a priori* résultante peut s'écrire :

$$f(\mathbf{z}) \propto \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \quad (3.10)$$

où $\mathcal{V}(p)$ représente le voisinage d'ordre 1 du pixel p donné à la figure 3.1.

Loi *a priori* des coefficients logistiques

Suivant l'approche décrite au paragraphe 3.2.1, chaque composante de \mathbf{t}_p est supposée être distribuée selon une loi Gaussienne. De plus, comme démontré au paragraphe 3.2.3 (voir (3.6)), la moyenne et la variance des coefficients logistique dépendent de la classe à laquelle le pixel considéré appartient. Donc, la loi *a priori* de \mathbf{t}_p est explicitement définie conditionnellement à l'étiquette du pixel :

$$t_{r,p} | z_p = k, \psi_{r,k}, \sigma_{r,k}^2 \sim \mathcal{N}(\psi_{r,k}, \sigma_{r,k}^2) \quad (3.11)$$

où les hyperparamètres $\psi_{r,k}$ et $\sigma_{r,k}^2$ dépendent de la classe k du pixel associé. L'estimation de ces hyperparamètres se fera en construisant un second niveau de hiérarchie d'inférence Bayésienne. Pour un pixel p donné, en supposant l'indépendance *a priori* entre les coefficients $t_{1,p}, \dots, t_{R,p}$, la loi *a priori* pour le vecteur $\mathbf{t} = [t_{1,p}, \dots, t_{R,p}]^T$ est :

$$f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \sim \mathcal{N}(\boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \quad (3.12)$$

où $\boldsymbol{\psi}_k = [\psi_{1,k}, \dots, \psi_{R,k}]^T$ et $\boldsymbol{\Sigma}_k = \text{diag}(\sigma_{r,k}^2)$ est la matrice diagonale de taille $R \times R$ dont les éléments diagonaux sont $\sigma_{r,k}^2$.

En supposant l'indépendance entre les P vecteurs $\mathbf{t}_1, \dots, \mathbf{t}_P$, la loi *a priori* complète de la matrice des coefficients logistiques \mathbf{T} est alors :

$$f(\mathbf{T} | \boldsymbol{\Psi}, \boldsymbol{\Sigma}) = \prod_{k=1}^K \prod_{p \in \mathcal{I}_k} f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) \quad (3.13)$$

avec $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \dots, \boldsymbol{\psi}_K]$ et $\boldsymbol{\Sigma} = \{\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K\}$.

Lois *a priori* de la variance des endmembers

Une loi conjuguée inverse Gamma a été choisie pour la variance des endmembers, soit :

$$w_p^2 | \nu, \delta \sim \mathcal{IG}(\nu, \delta) \quad (3.14)$$

où ν et δ sont des hyperparamètres ajustables. Comme aux chapitres précédents, on suppose que $\nu = 1$ et estime δ conjointement avec les autres paramètres et hyperparamètres (par l'utilisation d'un modèle Bayésien hiérarchique).

Prenant comme hypothèse l'indépendance entre les différentes lois *a priori* des variances et en notant $\mathbf{w} = [w_1^2, \dots, w_P^2]^T$, la loi *a priori* de \mathbf{w} est alors :

$$f(\mathbf{w} | \delta) = \prod_{p=1}^P f(w_p^2 | \delta) \quad (3.15)$$

3.3.4 Lois *a priori* des hyperparamètres

Pour la définition des lois *a priori* des hyperparamètres, les hyperparamètres $\psi_{r,k}$ et $\sigma_{r,k}^2$ ont reçu une attention toute particulière étant donné qu'ils décrivent en intégralité les différentes classes découpant l'image. Les lois *a priori* pour $\psi_{r,k}$ et $\sigma_{r,k}^2$ sont des lois conjuguées. Plus précisément, une loi vague inverse Gamma a été choisie pour la variance des coefficients logistiques $\sigma_{r,k}^2$, soit

$$\sigma_{r,k}^2 | \xi, \gamma \sim \mathcal{IG}(\xi, \gamma) \quad (3.16)$$

où l'on a imposé $\xi = 1$ et $\gamma = 5$ (afin d'obtenir une grande variance pour cette loi). De plus, une loi Gaussienne centrée de variance inconnue a été prise comme loi *a priori* pour la moyenne des coefficients logistiques :

$$\psi_{r,k} | v^2 \sim \mathcal{N}(0, v^2) \quad (3.17)$$

où v^2 est un autre hyperparamètre ajustable. Comme précédemment, on suppose que les différents vecteurs moyennes ψ_k , ($k = 1, \dots, K$) ainsi que les matrices de covariance Σ_k , ($k = 1, \dots, K$) sont indépendants. Les lois *a priori* complètes pour ces deux hyperparamètres sont alors :

$$f(\Psi | v^2) \propto \prod_{k=1}^K \prod_{r=1}^R \left(\frac{1}{v^2} \right)^{\frac{1}{2}} \exp \left(-\frac{\psi_{r,k}^2}{2v^2} \right) \quad (3.18)$$

$$f(\Sigma | \xi, \gamma) \propto \prod_{k=1}^K \prod_{r=1}^R \frac{\gamma^\xi}{\Gamma(\xi)} (\sigma_{r,k}^2)^{-(\xi+1)} \exp \left(-\frac{\gamma}{\sigma_{r,k}^2} \right). \quad (3.19)$$

Des lois non informatives de Jeffreys sont choisies pour les hyperparamètres δ et v^2 :

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta), \quad f(v^2) \propto \frac{1}{v^2} \mathbf{1}_{\mathbb{R}^*}(v^2). \quad (3.20)$$

Ces lois reflètent bien le manque de connaissance sur ces deux hyperparamètres. S'agissant des deux derniers niveaux de hiérarchie dans notre inférence Bayésienne, le vecteur des hyperparamètres peut être définie comme suit : $\Omega = \{\Psi, \Sigma, v^2, \delta\}$.

3.3.5 Loi jointe

La loi jointe *a posteriori* des paramètres inconnus et des hyperparamètres est définie suivant : $f(\Theta, \Omega | \mathbf{Y}) = f(\mathbf{Y} | \Theta) f(\Theta | \Omega) f(\Omega)$.

Après calculs, la loi *a posteriori* pour le modèle NCM peut s'écrire comme suit :

$$\begin{aligned}
 f(\Theta, \Omega | \mathbf{Y}) \propto & \prod_{p=1}^P \frac{1}{[w_p^2 c(\mathbf{t}_p)]^{L/2}} \exp \left[- \sum_{p=1}^P \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{t}_p)\|^2}{2w_p^2 c(\mathbf{t}_p)} \right] \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \\
 & \times \delta^{P-1} \prod_{p=1}^P \left(\frac{1}{w_p^2} \right)^{\nu+1} \exp \left(- \frac{\delta}{w_p^2} \right) \left(\frac{1}{v^2} \right)^{\frac{RK}{2}+1} \\
 & \times \prod_{r,k} \frac{1}{\sigma_{r,k}^{n_k+1}} \exp \left[- \left(\frac{\psi_{r,k}^2}{2v^2} + \frac{2\gamma + \sum_{p \in \mathcal{I}_k} (t_{r,p} - \psi_{r,k})^2}{2\sigma_{r,k}^2} \right) \right] \quad (3.21)
 \end{aligned}$$

avec $n_k = \text{card}(\mathcal{I}_k)$. Comme ce fut le cas pour les chapitres précédents, la loi *a posteriori* est trop complexe pour obtenir des expressions analytiques des estimateurs MMSE et MAP du vecteur des paramètres inconnus Θ . Encore une fois nous allons faire appel aux méthodes MCMC pour pallier le problème de la complexité des lois *a posteriori*. Le paragraphe suivant étudie les deux échantillonneurs de Gibbs générant des échantillons asymptotiquement distribués suivant la loi *a posteriori* (3.21).

3.4 Échantillonneur de Gibbs hybride

Ce paragraphe décrit l'échantillonneur de Metropolis-within-Gibbs générant des échantillons selon $f(\Theta, \Omega | \mathbf{Y})$. L'algorithme est résumé dans ALGO. 3.2. Les lois conditionnelles des différents paramètres et hyperparamètres sont détaillées ci-dessous.

3.4.1 Loi conditionnelle du vecteur des étiquettes \mathbf{z}

Pour chaque pixel p ($p = 1, \dots, P$), l'étiquette z_p est une variable aléatoire discrète dont la loi conditionnelle est pleinement caractérisée par les probabilités :

$$P[z_p = k, \mathbf{z}_{-p}, \mathbf{t}_p, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k] \propto f(\mathbf{t}_p | z_p = k, \mathbf{z}_{-p}, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) P[z_p = k, \mathbf{z}_{-p}, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k],$$

ce qui donne par la relation de Bayes :

$$P[z_p = k | \mathbf{z}_{-p}, \mathbf{t}_p, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k] \propto f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k) P[z_p = k | \mathbf{z}_{-p}] \quad (3.22)$$

où $k = 1, \dots, K$ (K étant le nombre de classes), $f(z_p | \mathbf{z}_{-p})$ a été définie dans l'équation 3.10 et \mathbf{z}_{-p} dénote le vecteur \mathbf{z} dont le p ème élément a été supprimé. Ces probabilités *a posteriori* peuvent s'exprimer ainsi :

$$P[z_p = k | \mathbf{z}_{-p}, \mathbf{t}_p, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k] \propto |\boldsymbol{\Sigma}_k|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{t}_p - \boldsymbol{\psi}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{t}_p - \boldsymbol{\psi}_k) \right] \\ \times \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \quad (3.23)$$

où $|\boldsymbol{\Sigma}_k| = \prod_{r=1}^R \sigma_{r,k}^2$. La loi *a posteriori* du vecteur des étiquettes \mathbf{z} dans (3.23) définit un champ de Markov. Par conséquent, on peut échantillonner suivant cette loi conditionnelle en utilisant l'algorithme détaillé dans ALGO. 3.1, c'est-à-dire en générant des valeurs discrètes suivant les probabilités en (3.23). Un échantillonneur de Gibbs est proposé pour générer z_p .

3.4.2 Loi conditionnelle de la matrice des coefficients logistiques \mathbf{T}

Pour chaque pixel p , la loi jointe peut s'écrire¹ :

$$f(\mathbf{t}_p, z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k | \mathbf{y}_p) \propto f(\mathbf{y}_p | \mathbf{t}_p) f(\mathbf{t}_p, z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k),$$

et le théorème de Bayes donne :

$$f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k, \mathbf{y}_p) \propto f(\mathbf{y}_p | \mathbf{t}_p) f(\mathbf{t}_p | \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k).$$

Après calculs, la loi conditionnelle s'écrit :

$$f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k, \mathbf{y}_p, w_p^2) \propto \frac{1}{[w_p^2 c(\mathbf{t}_p)]^{L/2}} \exp \left\{ -\frac{1}{2w_p^2 c(\mathbf{t}_p)} \|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{t}_p)\|^2 \right\} \\ \times |\boldsymbol{\Sigma}_k|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{t}_p - \boldsymbol{\psi}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{t}_p - \boldsymbol{\psi}_k) \right]. \quad (3.24)$$

1. La variance des endmembers a été omise pour des raisons de concisions.

Malheureusement, il est trop difficile de générer des échantillons suivant (3.24). Une étape de Metropolis-Hastings va donc être utilisée, basée sur une marche aléatoire, similaire à celle utilisée au chapitre 1, au paragraphe 1.3.1. La loi de proposition est une Gaussienne $\mathcal{N}(0, u_r^2)$, centrée de variance u_r^2 qui a été fixée de manière à obtenir un taux d'acceptation entre 0.15 et 0.5, comme recommandé dans [Rob96].

3.4.3 Lois conditionnelles des variances des endmembers

Pour chaque pixel p , le théorème de Bayes donne :

$$f(s_p^2 | \mathbf{y}_p, \mathbf{t}_p, \delta) \propto f(\mathbf{y}_p | \mathbf{t}_p, s_p^2) f(s_p^2 | \delta).$$

Dès lors, on obtient la loi inverse Gamma :

$$w_p^2 | \mathbf{y}_p, \mathbf{t}_p, \delta \sim \text{IG} \left(\frac{L}{2} + 1, \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{t}_p)\|^2}{2c(\mathbf{t}_p)} + \delta \right). \quad (3.25)$$

3.4.4 Loi conditionnelle de Ψ et Σ

Pour chaque endmember r ($r = 1, \dots, R$) et chaque classe k ($k = 1, \dots, K$) grâce à l'équation de Bayes la loi conditionnelle de $\psi_{r,k}$ peut s'écrire selon :

$$f(\psi_{r,k} | \mathbf{z} = k, \mathbf{t}_r, \sigma_{r,k}^2, v^2) \propto f(\psi_{r,k} | v^2) \prod_{p \in \mathcal{I}_k} f(t_{r,p} | z_p = k, \psi_{r,k}, \sigma_{r,k}^2). \quad (3.26)$$

De façon similaire, la loi conditionnelle de $\sigma_{r,k}^2$ s'écrit :

$$f(\sigma_{r,k}^2 | \mathbf{t}_r, \mathbf{z} = k, \psi_{r,k}) \propto f(\sigma_{r,k}^2) \prod_{p \in \mathcal{I}_k} f(t_{r,p} | z_p = k, \psi_{r,k}, \sigma_{r,k}^2). \quad (3.27)$$

Après calcul, les résultats suivants sont obtenus :

$$\psi_{r,k} | \mathbf{z} = k, \mathbf{t}_r, \sigma_{r,k}^2, v^2 \sim \mathcal{N} \left(\frac{v^2 n_k \bar{t}_{r,k}}{\sigma_{r,k}^2 + v^2 n_k}, \frac{v^2 \sigma_{r,k}^2}{\sigma_{r,k}^2 + v^2 n_k} \right) \quad (3.28)$$

avec $\bar{t}_{r,k} = \frac{1}{n_k} \sum_{p \in \mathcal{I}_k} t_{r,p}$ et

$$\sigma_{r,k}^2 | \mathbf{z} = k, \mathbf{t}_r, \psi_{r,k} \sim \text{IG} \left(\frac{n_k}{2} + 1, \gamma + \sum_{p \in \mathcal{I}_k} \frac{(t_{r,p} - \psi_{r,k})^2}{2} \right). \quad (3.29)$$

3.4.5 Lois conditionnelles de v^2 et δ

La loi conditionnelle de v^2 est la loi inverse-Gamma suivante :

$$v^2 | \Psi \sim \mathcal{IG} \left(\frac{RK}{2}, \frac{1}{2} \Psi^T \Psi \right), \quad (3.30)$$

alors que la loi conditionnelle de δ est donnée par :

$$\delta | \mathbf{w} \sim \mathcal{G} \left(P, \sum_{p=1}^P \frac{1}{w_p^2} \right). \quad (3.31)$$

3.5 Résultats de simulations sur données synthétiques

De nombreuses simulations ont été conduites afin d'illustrer l'efficacité de l'algorithme proposé. La première expérience considère une image synthétique de 25×25 pixels avec $K = 3$ classes différentes. L'image contient $R = 3$ composants purs mélangés dont les spectres ($L = 413$ bandes spectrales) ont été extraits des bibliothèques spectrales du logiciel ENVI [RSI03]. Les spectres extraits sont représentés sur la figure 3.4, les deux premiers étant identiques à ceux pris pour les simulations du chapitre 1 (béton de construction et herbe verte). Une carte des étiquettes, représentée sur la figure 3.5 (à gauche) a été générée en utilisant (3.10) avec $\beta = 1.1$.

Les moyennes et variances des abondances ont été imposées pour chaque classe comme résumés dans la Table 3.1. Ces valeurs reflètent bien le fait que le premier endmember est prépondérant dans la Classe 1 (avec une concentration moyenne de 60%), le deuxième endmember est plus présent dans la Classe 2 (avec une concentration moyenne de 50%) et le troisième prépondérant dans la Classe 3 (avec une concentration moyenne de 50%). Dans cette simulation, la variance des abondances a été fixée à une valeur de 0.005 quelque soient les endmembers, pixels et classes. Les cartes d'abondances générées suivant le NCM sont données sur la figure 3.6 (haut). Les variances des endmembers ont été générées suivant leurs lois *a priori* (3.15) avec $\delta = 1 \times 10^{-3}$ conduisant à un SNR égal à SNR = 12dB.

ALGO. 3.2 Échantillonneur de Gibbs hybride pour le démélange spectral avec contraintes spatiales pour la modèle NCM

– % *Initialisation* :

- 1: Échantillonner $\mathbf{z}^{(0)}$ à partir de la densité de probabilité (pdf) dans Eq. (3.10),
- 2: Échantillonner $\delta^{(0)}$ et $v^{2(0)}$ à partir des pdfs dans Eq. (3.20),
- 3: Échantillonner $\mathbf{w}^{(0)}$ à partir de la pdf dans Eq. (3.15),
- 4: Échantillonner $\Psi^{(0)}$ à partir de la pdf dans Eq. (3.18),
- 5: Échantillonner $\Sigma^{(0)}$ à partir de la pdf dans Eq. (3.19),

– % *Itérations* :

- 1: **pour** $t = 1, 2, \dots$ **faire**
 - 2: **pour** chaque pixels $p = 1, \dots, P$ **faire**
 - 3: Échantillonner $\mathbf{z}^{(t)}$ à partir de la pdf dans Eq. (3.23),
 - 4: Échantillonner $\mathbf{t}_p^{(t)}$ à partir de la pdf dans Eq. (3.24) utilisant une étape de Metropolis-within-Gibbs,
 - 5: Calculer $\mathbf{a}_p^{(t)}$ à partir de Eq. (3.5),
 - 6: Échantillonner $w_p^{2(t)}$ à partir de la pdf dans Eq. (3.25),
 - 7: **fin pour**
 - 8: **pour** chaque classe $k = 1, \dots, K$ **faire**
 - 9: **pour** chaque endmember $r = 1, \dots, R$ **faire**
 - 10: Échantillonner $\psi_{r,k}^{(t)}$ à partir de la pdf dans Eq. (3.28),
 - 11: Échantillonner $\sigma_{r,k}^{2(t)}$ à partir de la pdf dans Eq. (3.29),
 - 12: **fin pour**
 - 13: **fin pour**
 - 14: Échantillonner $v^{2(t)}$ à partir de la pdf dans in Eq. (3.30)
 - 15: Échantillonner $\delta^{(t)}$ à partir de la pdf dans Eq. (3.31),
 - 16: **fin pour**
-

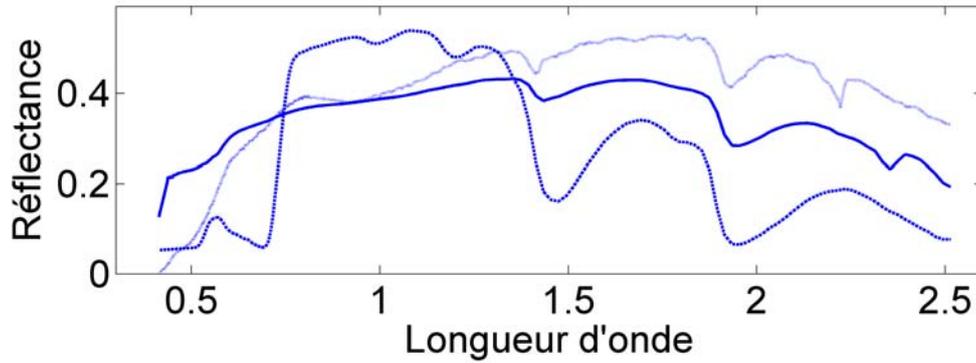


FIGURE 3.4 – Les $R = 3$ spectres des endmembers : béton de construction (trait continu), herbe verte (trait semi-continu), terre micacée (trait pointillés).

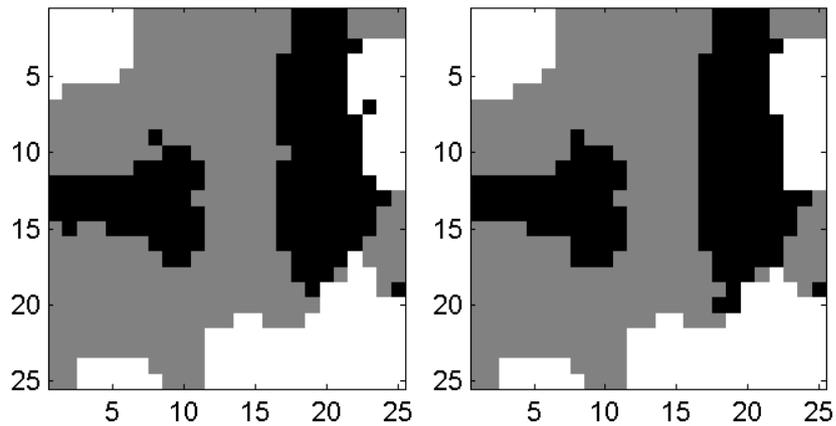


FIGURE 3.5 – A gauche : carte des étiquettes réelle. A droite : carte des étiquettes estimée par l'échantillonneur de Gibbs hybride proposé.

Les estimateurs MMSE et MAP des paramètres inconnus peuvent être calculés à partir des échantillons générés par l'échantillonneur de Gibbs présenté au paragraphe 3.4. Dans un premier temps, l'estimateur MAP marginal du vecteur d'étiquettes \hat{z}_{MAP} est représenté sur la figure 3.5. Les estimateurs MMSE des abondances conditionnellement à \hat{z}_{MAP} ont également été représentées sur la figure 3.6 (en bas). Il a fallu un nombre de $N_{\text{MC}} = 5000$ itérations (avec 500 itérations de chauffage) pour arriver à ces résultats. De

TABLE 3.1 – Moyennes et variances réelles et estimées des abondances dans chacune des classes.

		Valeurs réelles	NCM
Classe 1	$\boldsymbol{\mu}_1 = \mathbf{E}[a_p, p \in \mathcal{I}_1]$	$[0.6, 0.3, 0.1]^T$	$[0.59, 0.29, 0.12]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_1] (\times 10^{-3})$	$[5, 5, 5]^T$	$[4.7, 4.5, 7.3]^T$
Classe 2	$\boldsymbol{\mu}_2 = \mathbf{E}[a_p, p \in \mathcal{I}_2]$	$[0.3, 0.5, 0.2]^T$	$[0.31, 0.49, 0.2]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_2] (\times 10^{-3})$	$[5, 5, 5]^T$	$[7, 4.2, 11.7]^T$
Classe 3	$\boldsymbol{\mu}_3 = \mathbf{E}[a_p, p \in \mathcal{I}_3]$	$[0.3, 0.2, 0.5]^T$	$[0.31, 0.2, 0.49]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_3] (\times 10^{-3})$	$[5, 5, 5]^T$	$[4.5, 4.3, 5.5]^T$

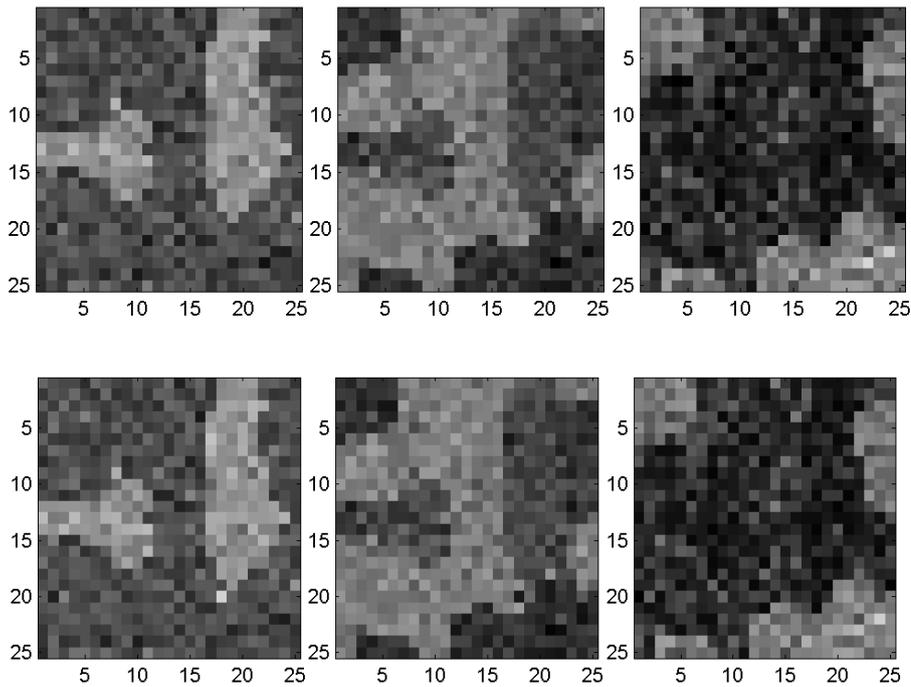


FIGURE 3.6 – En haut : cartes d'abondances des 3 matériaux purs pour le NCM. En bas : cartes d'abondances des 3 matériaux purs estimés par l'échantillonneur de Gibbs hybride basé sur le NCM (de gauche à droite : béton de construction, herbe verte, terre grasse).

plus, les vecteurs moyennes $\boldsymbol{\mu}_k = \mathbb{E}[\mathbf{a}_p]$ ($k = 1, \dots, K$, $p \in \mathcal{I}_k$), introduits dans (3.2), présentent un très grand intérêt dans la mesure où ils caractérisent chaque classe selon leurs définitions au paragraphe 3.2.1. Dès lors, il peut être judicieux de retrouver les lois *a posteriori* de ces vecteurs moyennes à partir des échantillons générés par l'algorithme de Gibbs hybride proposé. Les lois *a posteriori* des moyennes des abondances $\boldsymbol{\mu}_k$ ($k = 1, \dots, 3$) estimés par l'algorithme proposé sont représentés sur la figure 3.7 pour chacune des classes. Enfin, les résultats des estimations MMSE des moyennes et variances des abondances pour chaque endmember pour chacune des classes ont été présentés dans la table 3.1. Les différents paramètres (classes, coefficients d'abondances et vecteurs des moyennes des abondances) estimés par l'algorithme sont clairement en accord avec les valeurs réelles.

Cet algorithme est comparé aux algorithmes Bayésiens spatialement indépendants, respectivement développés dans [DTC08] et au chapitre 1. L'image synthétique montrée figure 3.5 a été analysée par ces deux derniers algorithmes avec le même nombre d'itérations N_{MC} en plus de l'algorithme FCLS [HC01]. Comme critère, les EQMs (déjà utilisées au premier chapitre au paragraphe 1.6.3) des abondances estimées ont été calculées pour chacun des algorithmes. Nous rappelons l'expression de l'EQM de la r ième abondance définie par :

$$\text{EQM}_r^2 = \frac{1}{P} \sum_{p=1}^P (\hat{a}_{r,p} - a_{r,p})^2 \quad (3.32)$$

où $\hat{a}_{r,p}$ est l'estimateur MMSE de l'abondance $a_{r,p}$. Les résultats obtenus en calculant ces EQMs, reportés dans la table 3.2, montrent que l'algorithme développé dans ce chapitre (appelé « Bayésien corrélé » dans la table) présente de meilleures performances que les algorithmes spatialement indépendants (nommés « Bayésien indépendant » et « FCLS »).

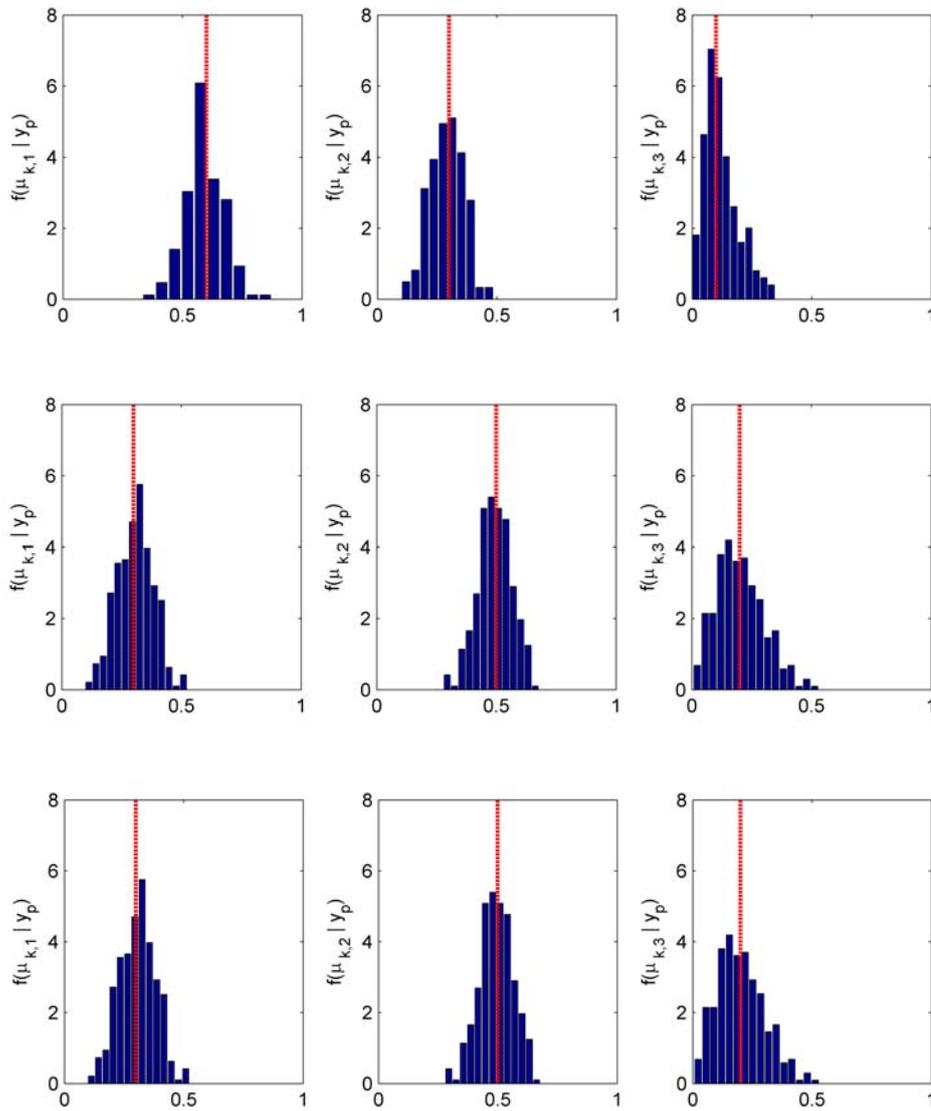


FIGURE 3.7 – Histogrammes des moyennes d’abondances $\mu_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ estimés par l’algorithme de Gibbs hybride pour les $K = 3$ classes (valeurs réelles en pointillés).

TABLE 3.2 – EQMs globales pour chaque composante des vecteurs d’abondances.

	NCM		MML	
	Bayésien indépen- dant	Bayésiën corrélé	FCLS	Bayésiën indépen- dant
EQM_1^2	0.0027	0.0013	0.0019	0.0016
EQM_2^2	8.9×10^{-4}	5.9×10^{-4}	4.3×10^{-4}	4.1×10^{-4}
EQM_3^2	9.5×10^{-4}	5.5×10^{-4}	0.0014	0.0013

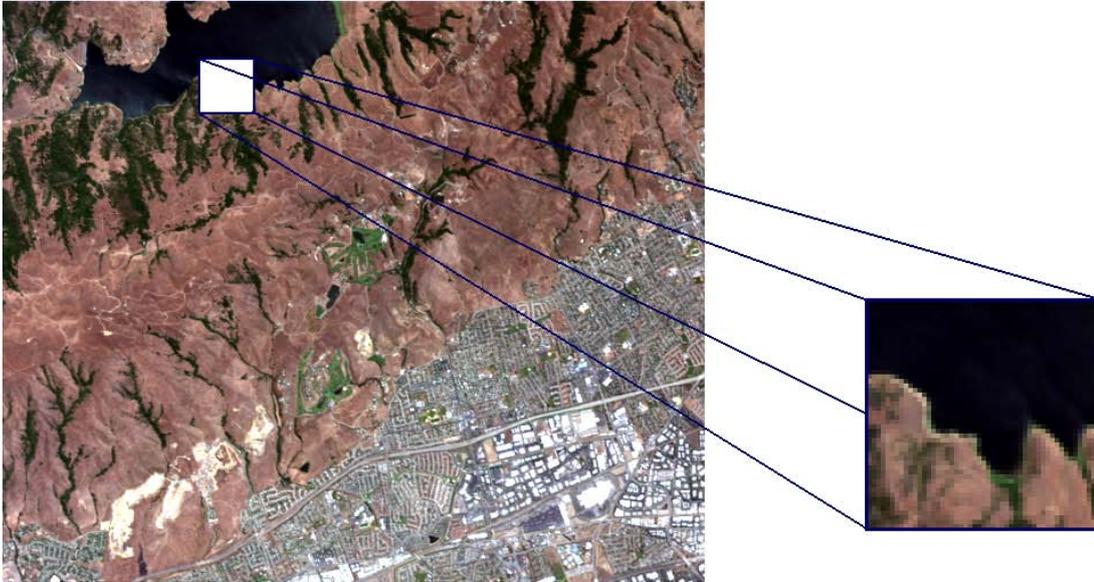


FIGURE 3.8 – Données hyperspectrales réelles : l’image de Moffett field acquise par AVIRIS en 1997 (à gauche) et la région d’intérêt représenté en couleurs réelles (à droite).

3.6 Résultats de simulations sur une image AVIRIS

Ce paragraphe illustre les performances de l’algorithme exploitant les corrélations entre pixels voisins à partir d’une image hyperspectrale utilisée pour les simulations des chapitres 1 (paragraphe 1.7) et 2 (paragraphe 2.5), à savoir, l’image Moffett Field, prise

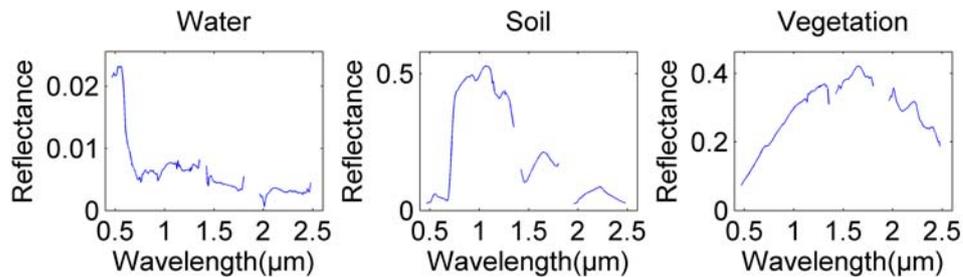


FIGURE 3.9 – Les spectres des $R = 3$ endmembers obtenues par l’algorithme N-FINDR.

par l’imageur AVIRIS en 1997. La région d’intérêt concerne toujours l’image de 50×50 pixels. Les endmembers utilisés pour les simulations ont été extraits suivant la méthode décrite au paragraphe 1.7. À titre de rappel, l’image complète et la région d’intérêt ainsi que les spectres des endmembers extraits sont respectivement représentés sur la figure 3.8 et 3.9. Ainsi, les algorithmes proposés au paragraphe 3.4 ont été appliqués sur cette image avec $K = 4$ classes et $N_{MC} = 5000$ itérations (avec 500 itérations de chauffage).

Les cartes d’abondances estimées représentés sur la figure 3.10. Ces résultats sont très similaires et en accord avec ceux obtenus sur cette image par l’utilisation de l’algorithme Bayésien basé sur le MML [DTC08] ou avec l’algorithme FCLS [HC01] et représentés, pour rappel, sur la figure 3.11.

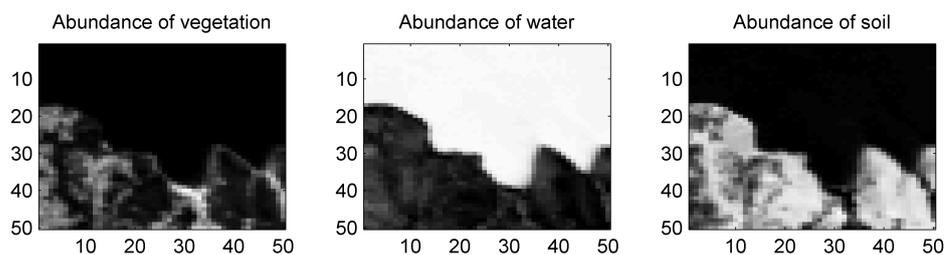


FIGURE 3.10 – Cartes d’abondances estimées par l’algorithme basé sur le NCM (de gauche à droite : végétation, eau et terre).

Les résultats des classifications obtenus par l'échantillonneur de Gibbs hybride basé sur le NCM sont donnés sur la figure 3.12. Les lois *a posteriori* des moyennes des abondances $\mu_k = E[\mathbf{a}_p]$ ($k = 1, \dots, K$, $p \in \mathcal{I}_k$) ont également été représentées sous formes d'histogrammes sur la figure 3.13. L'approche spatiale du démelange spectral initié par notre algorithme permet l'établissement d'une carte de classification à partir de laquelle il est possible d'examiner la répartition des matériaux purs de l'image. Par exemple, la zone du lac est aisément identifiable (la zone en blanc) où l'eau est bien évidemment en très forte quantité (moyenne très proche de 1 sur la figure 3.13 pour la quatrième classe).

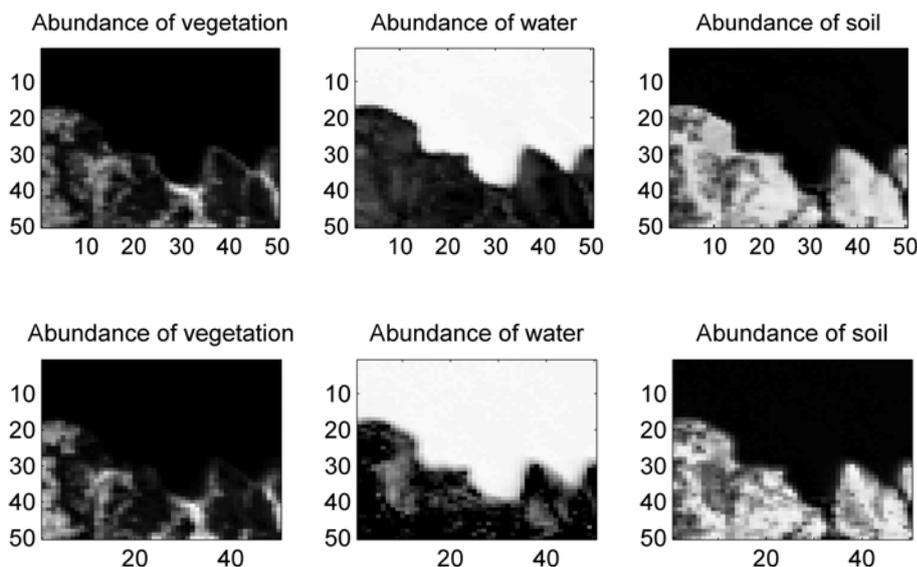


FIGURE 3.11 – En haut : cartes d'abondances estimées par le MML Bayésien (tiré de [DTC08]). En bas : cartes d'abondances estimées par l'algorithme FCLS [HC01].

3.7 Conclusions

Ce chapitre présente un nouvel algorithme de démelange spectral Bayésien et tenant compte des corrélations entre les pixels de l'image. Ces corrélations ont été modélisées à l'aide des champs de Markov. Une variable discrète supplémentaire (étiquette) a été

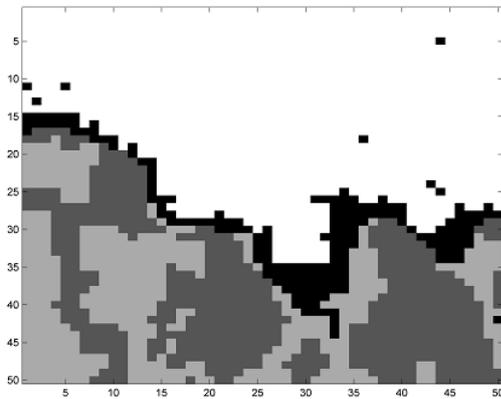


FIGURE 3.12 – Cartes des étiquettes estimées par l'algorithme basé sur le NCM.

introduite afin d'identifier des régions (ou classes) définies par des abondances homogènes (avec des statistiques du premier et deuxième ordre communes). Les contraintes de positivité et d'additivité des abondances ont été assurées à l'aide d'une reparamétrisation appropriée définie par des vecteurs de coefficients logistiques. La loi jointe *a posteriori* des paramètres inconnus et hyperparamètres associés a été calculée pour le modèle NCM. Ainsi une méthode MCMC a été étudiée afin de générer des échantillons asymptotiquement distribués suivant les lois conditionnelles associés à ce modèle. Ces échantillons ont ensuite été utilisés pour estimer les cartes d'abondances ainsi que les étiquettes de l'image sous-jacentes. Les résultats obtenus sur des données simulées et sur une image réelle AVIRIS se sont révélés intéressants, ce qui encourage la poursuite des études sur ce type d'algorithme spatial.

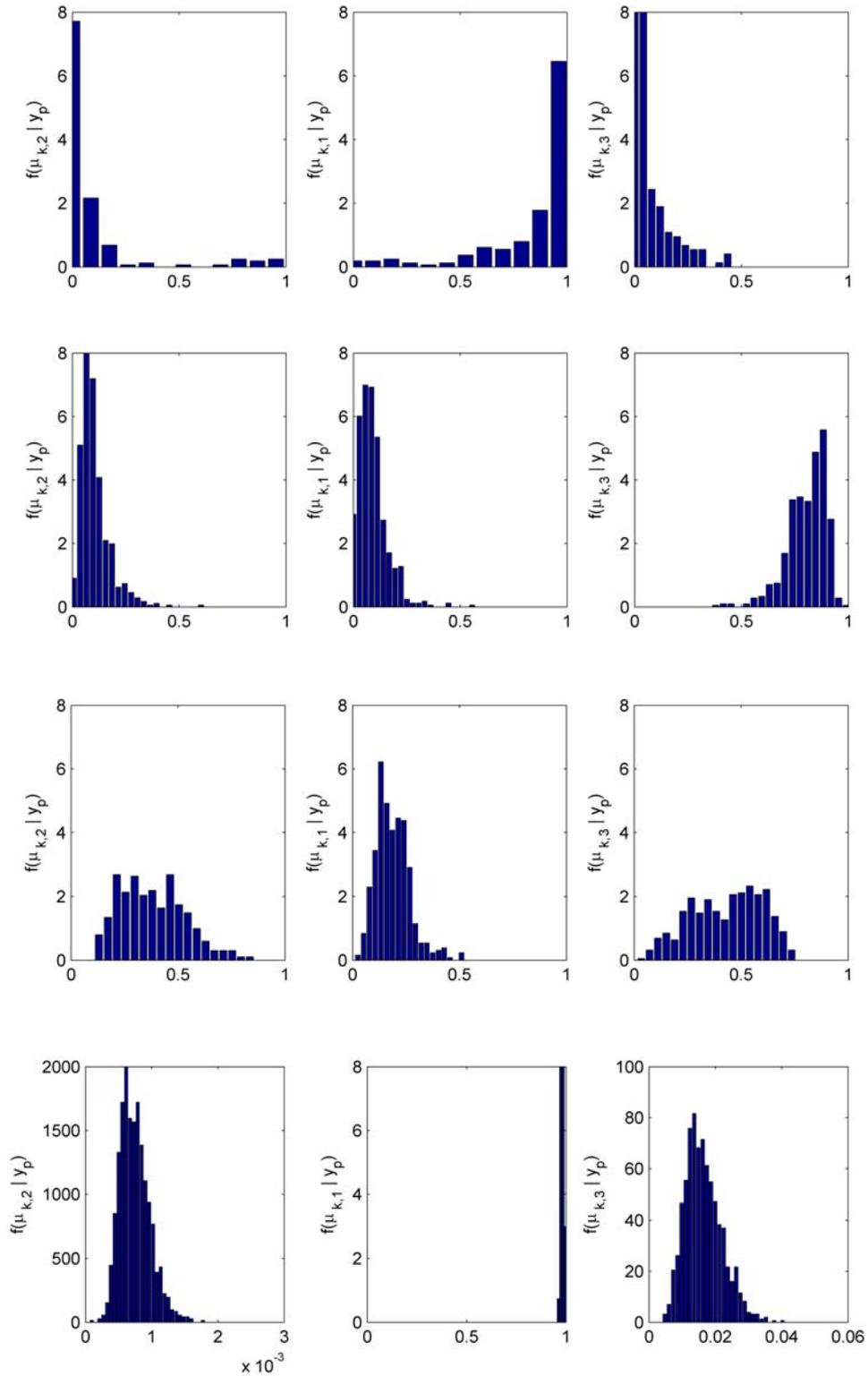


FIGURE 3.13 – Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ (de gauche à droite : végétation, eau et terre) estimés par l'algorithme de Gibbs hybride pour les $K = 4$ classes.

CHAPITRE 4

Conclusions et perspectives

Sommaire

4.1	Conclusions	99
4.2	Perspectives	101

4.1 Conclusions

Cette thèse avait pour objet d'étudier des méthodes de démixtion spectrale linéaire pour les images hyperspectrales. Nous nous sommes focalisés sur le problème d'inversion, étape consistant à estimer les abondances des matériaux purs d'un pixel mélangé. Nous avons considéré dans un premier temps que les spectres des matériaux purs, ou endmembers, étaient considérés comme connus, c'est à dire extraits d'une bibliothèque spectrale ou bien estimés par une méthode d'extraction d'endmembers. Ensuite, nous avons pu constater que le modèle de mélange linéaire, modèle couramment utilisé pour le démixtion, présente des inconvénients en l'absence de pixel purs. Dès lors nous avons considéré le modèle NCM présenté dans [ES07] dont la caractéristique principale est de considérer les spectres des matériaux purs comme étant aléatoires, au lieu de l'hypothèse déterministe présente dans le modèle MML standard. Nous avons choisi d'évoluer dans un cadre Bayésien en définissant des lois *a priori* appropriées pour les paramètres inconnus et leurs hyperparamètres associés. La complexité des lois *a posteriori* obtenues a été dépassée grâce à des méthodes de Monte Carlo par chaînes de Markov, consistant à

générer des échantillons asymptotiquement distribués suivant la loi d'intérêt. Ces échantillons sont ensuite employés pour approcher les estimateurs Bayésiens classiques (MAP ou MMSE).

L'utilisation du modèle NCM pour le démélange spectral linéaire a permis d'améliorer grandement les résultats d'estimation des abondances (par rapport au modèle MML) lorsque l'image analysée ne contient pas un nombre suffisant de pixels purs. Pour une image contenant un grand nombre de pixels purs en revanche, les résultats restent sensiblement identiques à ceux obtenus avec le modèle MML. Le modèle NCM est donc particulièrement recommandé si l'image hyperspectrale considérée est susceptible de contenir des pixels très hétérogènes. Inversement, il peut être préférable de garder le modèle linéaire standard si la présence de pixels purs représentatifs de tous les endmembers est assurée.

À partir de l'algorithme développé au chapitre précédent, nous avons cherché à estimer le nombre des composants purs présents dans un pixel, ou dans un groupe de pixels. Pour cela, nous avons supposé que les spectres des endmembers impliqués dans le mélange appartiennent à une bibliothèque spectrale. L'utilisation d'un algorithme à sauts réversibles s'est révélée tout à fait appropriée pour estimer le nombre d'endmembers présents dans le pixel (ou groupe de pixels) considéré mais également leur nature.

Les algorithmes d'inversion développés jusqu'à présent analysent les pixels de l'image indépendamment. Il est cependant possible de prendre en compte les corrélations entre les pixels d'une image grâce à l'utilisation des champs de Potts-Markov aléatoires. L'image est alors implicitement découpée en différentes zones ou « classes » dont les caractéristiques sont fonctions des valeurs des coefficients d'abondances des pixels correspondants. Grâce à une reparamétrisation des abondances, il a été possible d'établir un modèle Bayésien hiérarchique incluant ces corrélations spatiales. Une fois encore, les résultats obtenus en termes d'estimation des abondances se sont avérés très prometteurs. De plus, il a été possible d'extraire une carte de classification de l'image. Dès lors, un algorithme avec de telles hypothèses pourrait être employé à la fois comme algorithme de démélange

mais également comme classifieur, notamment si l'on souhaite analyser l'image par une approche « région ».

4.2 Perspectives

Ces travaux ouvrent la voie à des perspectives de recherches intéressantes pour l'amélioration du démélange spectral.

Le chapitre 1 a montré que le NCM est un modèle intéressant pour le démélange spectral d'une image ne contenant pas de pixels purs. Ce fut le cas notamment au paragraphe 1.6.3 où une image synthétique a été générée en supprimant les pixels purs pour l'analyse. Cependant, nous pensons que d'autres modèles que NCM prenant en compte l'absence de pixels purs dans l'image pourraient être envisagés. Plus précisément, il s'agirait d'intégrer directement leur absence dans le modèle statistique. Une piste envisageable serait de modifier les contraintes inhérentes aux coefficients d'abondances. Les coefficients, au lieu d'être strictement positifs, devraient être strictement supérieur à un réel positif (inférieur à 1) alors que les contraintes d'additivité seraient conservées. Un autre type de modèle consisterait à supposer que les données d'une image hyperspectrale ne sont pas contenues dans un seul ensemble convexe mais dans plusieurs ensembles convexes. C'est ce qui a été présenté dans [ZG10] où les données ont été regroupés dans plusieurs ensembles convexes et où les spectres des matériaux purs et abondances correspondantes ont été estimés.

Un autre algorithme semi-supervisé basé sur le modèle NCM peut également être envisagé. Ainsi au chapitre 2, le nombre de composants purs a été estimé à l'aide de méthodes de sélection de modèles. Il est possible d'éviter d'estimer directement le nombre de composants purs en faisant appel aux méthodes parcimonieuses. Ces méthodes supposent que le vecteur d'abondances a un grand nombre de composantes toutes nulles,

excepté un petit nombre. Cette approche a été présentée dans [TRK10] mais n'a cependant été appliquée que pour le modèle MML. Il serait alors intéressant de voir si de telles méthodes pourraient être applicable au modèle NCM.

Enfin, concernant la prise en compte des dépendances spatiales, il serait judicieux d'approfondir l'algorithme décrit au chapitre 3. En effet, l'algorithme a été développé pour des endmembers ayant la même variance. Comme ce fut le cas au paragraphe 1.4, cet algorithme pourrait être adapté au cas où les variances des endmembers ont des valeurs différentes. De même, nous avons supposé que le coefficient de granularité était connu pour des raisons de simplification. De futures études pourraient être réalisées sur l'estimation de ce coefficient comme dans [ZLQ97], [DMZB99] ou encore [CFP03]. Enfin, les corrélations spatiales ont été modélisée à l'aide de champs de Potts-Markov. Nous pensons qu'il est possible de modéliser autrement ces dépendances spatiales. Ainsi, nous pourrions faire appel, par exemple, aux champs aléatoires « discriminatifs » [KH03]. L'utilisation d'opérateurs morphologiques, autres outils d'intégration d'informations spatiales et utilisés avec succès pour la classification d'images hyperspectrales urbaines [BPS05] reste également envisageable.

De manière générale, l'utilisation de méthodes MCMC s'est avérée très coûteuse en temps de calcul. Les méthodes variationnelles [JJ00] s'avèrent être une alternative intéressante, réduisant de façon drastique le temps de calcul. Des résultats concluants ont été obtenus avec ces méthodes appliquées au modèle MML en négligeant la contrainte d'additivité des abondances. Cependant, le problème de la prise en compte de ces contraintes pour le MML et l'emploi du modèle NCM (avec ou sans les contraintes d'additivité) avec les méthodes variationnelles reste ouvert.

Annexes

ANNEXE A

Calcul de la loi conditionnelle

$$f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \mathbf{M})$$

En utilisant le théorème de Bayes, la loi conditionnelle $f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \mathbf{M})$ s'écrit comme suit :

$$f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \mathbf{M}) \propto f(\mathbf{Y} | \mathbf{A}, \mathbf{w}, \mathbf{M}) f(w_r^2 | \nu, \delta) \quad (\text{A.1})$$

ce qui mène à :

$$\begin{aligned} f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \mathbf{M}) &\propto \prod_{p=1}^P \left(\frac{1}{c(\mathbf{a}_p)} \right)^{L/2} \\ &\times \exp \left(- \sum_{p=1}^P \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2c(\mathbf{a}_p)} \right) \\ &\times \left(\frac{1}{w_r^2} \right)^{\nu+1} \exp \left(- \frac{\delta}{w_r^2} \right) \end{aligned}$$

Cette loi conditionnelle peut se réécrire de la façon suivante :

$$\begin{aligned} f(w_r^2 | \mathbf{w}_{-r}, \mathbf{Y}, \mathbf{A}, \mathbf{M}) &\propto \left(\frac{1}{w_r^2} \right)^{\nu+1} \prod_{p=1}^P (w_r^2 a_{r,p}^2 + c(\mathbf{a}_{-r,p}))^{-L/2} \\ &\times \exp \left(- \sum_{p=1}^P \frac{\|\mathbf{y}_p - \boldsymbol{\mu}(\mathbf{a}_p)\|^2}{2(w_r^2 a_{r,p}^2 + c(\mathbf{a}_{-r,p}))} - \frac{\delta}{w_r^2} \right). \quad (\text{A.2}) \end{aligned}$$

ANNEXE B

Probabilités d'acceptation pour les mouvements de « naissance » et « mort » pour l'algorithme à sauts réversibles

Les calculs des probabilités d'acceptation pour les mouvements de « naissance » et « mort » décrites au paragraphe 2.3 sont détaillés dans cette partie annexe.

A l'itération t , on considère le mouvement de naissance de l'état $\{\mathbf{a}^{(t)}, w^{2(t)}, \mathbf{M}^{(t)}, R^{(t)}\}$ au nouvel état $\{\mathbf{a}^*, w^{2*}, \mathbf{M}^*, R^*\}$ avec $\mathbf{a}^* = [(1 - \omega^*)a_1, \dots, (1 - \omega^*)a_{R^{(t)}}]^T$, $\mathbf{M}^* = [\mathbf{M}^{(t)}, \mathbf{s}^*]$ et $R^* = R^{(t)} + 1$. Le rapport d'acceptation pour ce mouvement de naissance est :

$$A_b = \frac{f(\mathbf{a}^*, \mathbf{M}^*, R^* | \mathbf{y})}{f(\mathbf{a}^{(t)}, \mathbf{M}^{(t)}, R^{(t)} | \mathbf{y})} \frac{p_{R^* \rightarrow R^{(t)}}}{p_{R^{(t)} \rightarrow R^*}} \frac{q(\mathbf{M}^{(t)}, w^{2(t)}, \mathbf{a}^{(t)} | \mathbf{M}^*, w^{2*}, \mathbf{a}^*)}{q(\mathbf{M}^*, w^{2*}, \mathbf{a}^* | \mathbf{M}^{(t)}, w^{2(t)}, \mathbf{a}^{(t)})} |\mathbf{J}(\omega^*)| \quad (\text{B.1})$$

où $q(\cdot | \cdot)$ fait référence aux lois de proposition, $|\mathbf{J}(\omega^*)|$ est le Jacobien de la transformation et p_{\rightarrow} est la probabilité de transition, c'est-à-dire, $p_{R^* \rightarrow R^{(t)}} = d_{R^*}$ et $p_{R^{(t)} \rightarrow R^*} = b_{R^{(t)}}$. D'après les mouvements précisés au paragraphe 2.3, le rapport des lois de propositions est :

$$\frac{q(\mathbf{M}^{(t)}, w^{2(t)}, \mathbf{a}^{(t)} | \mathbf{M}^*, w^{2*}, \mathbf{a}^*)}{q(\mathbf{M}^*, w^{2*}, \mathbf{a}^* | \mathbf{M}^{(t)}, w^{2(t)}, \mathbf{a}^{(t)})} = \frac{1}{g_{1, R^{(t)}}(w^*)} \frac{h_{\alpha_{w^*}, \beta_{w^*}}(w^{2(t)})}{h_{\alpha_{w^{(t)}}, \beta_{w^{(t)}}}(w^{2*})} \frac{R_{\max} - R^{(t)}}{R^{(t)} + 1} \quad (\text{B.2})$$

où $g_{a,b}(\cdot)$ et $h_{a,b}(\cdot)$ sont respectivement les fonctions densités de probabilité d'une loi Beta $\mathcal{B}e(a, b)$ et d'une loi inverse Gamma $\mathcal{IG}(a, b)$. En effet, la probabilité de choisir un

nouvel élément (mouvement de « naissance ») est $1/(R_{\max} - R^{(t)})$ et la probabilité de supprimer un élément (mouvement de « mort ») est $1/(R^{(t)} + 1)$. La matrice Jacobienne de la transformation est donnée par la relation suivante

$$\begin{aligned} \mathbf{J}(\omega^*) &= \begin{pmatrix} \frac{\partial a_1^*}{\partial a_1^{(t)}} & \cdots & \frac{\partial a_1^*}{\partial a_{R-1}^{(t)}} & \frac{\partial a_1^*}{\partial \omega^*} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial a_R^*}{\partial a_1^{(t)}} & \cdots & \frac{\partial a_R^*}{\partial a_{R-1}^{(t)}} & \frac{\partial a_R^*}{\partial \omega^*} \end{pmatrix} \\ &= \begin{pmatrix} 1 - \omega^* & 0 & \cdots & 0 & -a_1^{(t)} \\ 0 & 1 - \omega^* & \cdots & 0 & -a_2^{(t)} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 1 - \omega^* & -a_{R-1}^{(t)} \\ -(1 - \omega^*) & -(1 - \omega^*) & \cdots & -(1 - \omega^*) & -a_R^{(t)} \end{pmatrix} \end{aligned}$$

ce qui implique

$$|\mathbf{J}(\omega^*)| = (1 - \omega^*)^{R^{(t)} - 1}. \quad (\text{B.3})$$

Le rapport des lois *a posteriori*

$$\frac{f(\mathbf{a}^*, \mathbf{M}^*, R^* | \mathbf{y})}{f(\mathbf{a}^{(t)}, \mathbf{M}^{(t)}, R^{(t)} | \mathbf{y})} = \frac{f(\mathbf{y} | \mathbf{a}^*, \mathbf{M}^*, R^*)}{f(\mathbf{y} | \mathbf{a}^{(t)}, \mathbf{M}^{(t)}, R^{(t)})} \frac{f(\mathbf{a}^* | R^*)}{f(\mathbf{a}^{(t)} | R^{(t)})} \frac{f(\mathbf{M}^* | R^*)}{f(\mathbf{M}^{(t)} | R^{(t)})} \frac{f(R^*)}{f(R^{(t)})}. \quad (\text{B.4})$$

La loi *a priori* du vecteur d'abondances \mathbf{a}^+ est une loi de Dirichlet $\mathcal{D}_R(1, \dots, 1)$, ce qui mène à

$$\frac{f(\mathbf{a}^* | R^*)}{f(\mathbf{a}^{(t)} | R^{(t)})} = \frac{\Gamma(R^{(t)} + 1)}{\Gamma(1)\Gamma(R^{(t)})} = R^{(t)} + 1. \quad (\text{B.5})$$

En choisissant des configurations équiprobables *a priori* pour \mathbf{M} conditionnellement à R , le rapport des lois *a priori* de la matrice des spectres :

$$\frac{f(\mathbf{M}^* | R^*)}{f(\mathbf{M}^{(t)} | R^{(t)})} = \frac{R^{(t)} + 1}{R_{\max} - R^{(t)}}. \quad (\text{B.6})$$

Le rapport des lois *a priori* uniformes liées au nombre de composants purs dans le mélange R se réduit à 1. Au final, le rapport d'acceptation pour le mouvement de «

naissance » peut s'écrire

$$A_b = \left[\frac{c(\mathbf{a}^{(t)})}{c(\mathbf{a}^*)} \right]^{\frac{L}{2}} \frac{d_{R^{(t)}+1}}{b_{R^{(t)}}} \frac{R^{(t)}}{g_{1,R^{(t)}}(\omega^*)} (1 - \omega^*)^{R^{(t)}-1} \frac{h_{\alpha_{w^*}, \beta_{w^*}}(w^{2(t)})}{h_{\alpha_{w^{(t)}}, \beta_{w^{(t)}}}(w^{2*})} \\ \times \exp \left[\frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^{(t)})\|^2}{2c(\mathbf{a}^{(t)})} - \frac{\|\mathbf{y} - \boldsymbol{\mu}(\mathbf{a}^*)\|^2}{2c(\mathbf{a}^*)} \right]. \quad (\text{B.7})$$

ANNEXE C

Démélange spectral avec contraintes spatiales : application au modèle de mélange MML standard

Sommaire

C.1	Modèle Bayésien hiérarchique	111
C.2	Algorithme de Gibbs hybride	115
C.3	Résultats de simulations sur des données synthétiques	117
C.4	Résultats de simulations sur une image AVIRIS	122

L'algorithme développé dans [DTC08] va être modifié de manière à intégrer les corrélations entre les pixels de l'image.

C.1 Modèle Bayésien hiérarchique

Ce paragraphe décrit la vraisemblance et les lois a priori inhérentes au MML. La forme du modèle MML est donnée pour rappel :

$$\mathbf{y} = \sum_{r=1}^R \mathbf{m}_r a_r + \mathbf{n}. \quad (\text{C.1})$$

C.1.1 Vecteur des paramètres inconnus

Dans un premier temps, le vecteur des paramètres inconnus est défini ci-dessous

$$\Theta = \{\mathbf{T}, \mathbf{z}, \mathbf{s}\},$$

où $\mathbf{s} = [s_1^2, \dots, s_P^2]^T$ représente le vecteur des variances du bruit, \mathbf{z} est le vecteur des étiquettes et $\mathbf{T} = [\mathbf{t}_1, \dots, \mathbf{t}_P]$ avec $\mathbf{t}_p = [t_{1,p}, \dots, t_{R,p}]^T$ ($p = 1, \dots, P$) est la matrice de coefficients logistiques employés pour la reparamétrisation des abondances.

C.1.2 Fonction de vraisemblance

Le bruit blanc Gaussien additif du MML permet d'écrire¹ $\mathbf{y}_p | \mathbf{t}_p, s_p^2 \sim \mathcal{N}(\mathbf{M}\mathbf{a}_p(\mathbf{t}_p), s_p^2 \mathbf{I}_L)$ ($p = 1, \dots, P$). La fonction de vraisemblance de \mathbf{y}_p peut alors s'exprimer comme suit :

$$f(\mathbf{y}_p | \mathbf{t}_p, s_p^2) = \left(\frac{1}{2\pi s_p^2} \right)^{\frac{L}{2}} \exp \left[-\frac{\|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p(\mathbf{t}_p)\|^2}{2s_p^2} \right]. \quad (\text{C.2})$$

En supposant l'indépendance entre les diverses séquences de bruit \mathbf{n}_p ($p = 1, \dots, P$), la vraisemblance des P pixels de l'image est :

$$f(\mathbf{Y} | \mathbf{T}, \mathbf{s}) = \prod_{p=1}^P f(\mathbf{y}_p | \mathbf{t}_p, s_p^2). \quad (\text{C.3})$$

C.1.3 Lois a priori des paramètres

Ce paragraphe présente les lois *a priori* des paramètres inconnus et leur hyperparamètres associés qui seront utilisés pour le modèle MML. Le graphe orienté correspondant est représenté sur la figure C.1.

1. La dépendance du vecteur d'abondance en fonction du vecteur des coefficients logistiques suivant (3.5) a été explicitement mentionnés en notant $\mathbf{a}_p = \mathbf{a}_p(\mathbf{t}_p)$.

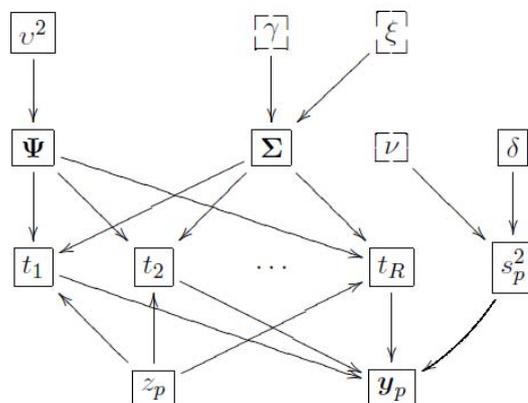


FIGURE C.1 – DAG pour les lois *a priori* (les paramètres fixés apparaissent dans les cases en pointillés) pour le MML.

Loi *a priori* des étiquettes

La loi *a priori* du vecteur des étiquettes $\mathbf{z} = [z_1, \dots, z_P]^T$ est la même que celle donnée à l'équation (3.10) :

$$f(\mathbf{z}) \propto \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \quad (\text{C.4})$$

où $\mathcal{V}(p)$ représente le voisinage d'ordre 1 du pixel p donné à la figure 3.1.

Loi *a priori* des coefficients logistiques

La loi *a priori* des coefficients logistiques restent la même que celle développée au paragraphe 3.3.3 :

$$t_{r,p} | z_p = k, \psi_{r,k}, \sigma_{r,k}^2 \sim \mathcal{N}(\psi_{r,k}, \sigma_{r,k}^2). \quad (\text{C.5})$$

Ce qui donne pour la matrice des coefficients logistiques \mathbf{T} (en supposant l'indépendance entre les pixels de l'image) :

$$f(\mathbf{T} | \Psi, \Sigma) = \prod_{k=1}^K \prod_{p \in \mathcal{I}_k} f(t_p | z_p = k, \psi_k, \Sigma_k). \quad (\text{C.6})$$

Lois *a priori* des variances du bruit

Une loi conjuguée inverse Gamma a été choisi pour la variance du bruit, soit :

$$s_p^2 | \nu, \delta \sim \mathcal{IG}(\nu, \delta), \quad (\text{C.7})$$

où $\nu = 1$ et δ hyperparamètre qui sera estimé conjointement avec les autres paramètres et hyperparamètres. Prenant comme hypothèse l'indépendance entre les différentes lois *a priori* des variances et en notant $\mathbf{s} = [s_1^2, \dots, s_P^2]^T$, la loi *a priori* de \mathbf{s} est alors :

$$f(\mathbf{s} | \delta) = \prod_{p=1}^P f(s_p^2 | \delta). \quad (\text{C.8})$$

C.1.4 Lois *a priori* des hyperparamètres

Les lois *a priori* complètes pour les deux hyperparamètres $\boldsymbol{\psi}_k$, ($k = 1, \dots, K$) et $\boldsymbol{\Sigma}_k$, ($k = 1, \dots, K$) restent inchangées par rapport aux lois dans (3.18) et (3.19) :

$$f(\boldsymbol{\Psi} | v^2) \propto \prod_{k=1}^K \prod_{r=1}^R \left(\frac{1}{v^2} \right)^{\frac{1}{2}} \exp \left(-\frac{\psi_{r,k}^2}{2v^2} \right) \quad (\text{C.9})$$

$$f(\boldsymbol{\Sigma} | \xi, \gamma) \propto \prod_{k=1}^K \prod_{r=1}^R \frac{\gamma^\xi}{\Gamma(\xi)} (\sigma_{r,k}^2)^{-(\xi+1)} \exp \left(-\frac{\gamma}{\sigma_{r,k}^2} \right). \quad (\text{C.10})$$

Des lois non informatives de Jeffreys sont choisies pour les hyperparamètres δ et v^2 :

$$f(\delta) \propto \frac{1}{\delta} \mathbf{1}_{\mathbb{R}^+}(\delta), \quad f(v^2) \propto \frac{1}{v^2} \mathbf{1}_{\mathbb{R}^+}(v^2). \quad (\text{C.11})$$

On définit le vecteur des hyperparamètres comme suit : $\Omega = \{\boldsymbol{\Psi}, \boldsymbol{\Sigma}, v^2, \delta\}$.

C.1.5 Loi jointe

La loi jointe *a posteriori* des paramètres inconnus et des hyperparamètres est définie suivant $f(\boldsymbol{\Theta}, \Omega | \mathbf{Y}) = f(\mathbf{Y} | \boldsymbol{\Theta}) f(\boldsymbol{\Theta} | \Omega) f(\Omega)$. Après calculs, la loi *a posteriori* pour le

modèle MML peut s'écrire comme suit :

$$\begin{aligned}
f(\Theta, \Omega | \mathbf{Y}) &\propto \prod_{p=1}^P \left(\frac{1}{s_p^2} \right)^{\frac{L}{2}} \exp \left[-\frac{\|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p(\mathbf{t}_p)\|^2}{2s_p^2} \right] \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \\
&\quad \times \delta^{P-1} \prod_{p=1}^P \left(\frac{1}{w_p^2} \right)^{\nu+1} \exp \left(-\frac{\delta}{w_p^2} \right) \left(\frac{1}{v^2} \right)^{\frac{RK}{2}+1} \\
&\quad \times \prod_{r,k} \frac{1}{\sigma_{r,k}^{n_k+1}} \exp \left[-\left(\frac{\psi_{r,k}^2}{2v^2} + \frac{2\gamma + \sum_{p \in \mathcal{I}_k} (t_{r,p} - \psi_{r,k})^2}{2\sigma_{r,k}^2} \right) \right] \quad (\text{C.12})
\end{aligned}$$

avec $n_k = \text{card}(\mathcal{I}_k)$. Cette loi étant trop complexe pour retrouver les expressions analytiques des estimateurs MMSE et MAP du vecteur des paramètres inconnus Θ , un algorithme de Gibbs hybride va permettre de générer des échantillons asymptotiquement distribués suivant (C.12).

C.2 Algorithme de Gibbs hybride

Ce paragraphe décrit l'échantillonneur de Metropolis-within-Gibbs générant des échantillons selon $f(\Theta, \Omega | \mathbf{Y})$. L'algorithme est résumé dans ALGO. C.1. Les lois conditionnelles des différents paramètres et hyperparamètres sont détaillés ci-dessous.

C.2.1 Loi conditionnelle du vecteur des étiquettes \mathbf{z}

La loi conditionnelle du vecteur des étiquettes \mathbf{z} est identique à celle développée pour le modèle NCM dans (3.23), soit :

$$\begin{aligned}
\mathbb{P}[z_p = k | \mathbf{z}_{-p}, \mathbf{t}_p, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k] &\propto |\boldsymbol{\Sigma}_k|^{-1/2} \exp \left[-\frac{1}{2} (\mathbf{t}_p - \boldsymbol{\psi}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{t}_p - \boldsymbol{\psi}_k) \right] \\
&\quad \times \exp \left[\sum_{p=1}^P \sum_{p' \in \mathcal{V}(p)} \beta \delta(z_p - z_{p'}) \right] \quad (\text{C.13})
\end{aligned}$$

où $|\boldsymbol{\Sigma}_k| = \prod_{r=1}^R \sigma_{r,k}^2$. L'échantillonnage suivant cette loi peut se faire de la même manière que la loi (3.23) (génération de valeurs discrètes suivant les probabilités en (C.13)).

C.2.2 Loi conditionnelle de la matrice des coefficients logistiques T

Pour chaque pixel p , le théorème de Bayes donne :

$$f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k, \mathbf{y}_p) \propto f(\mathbf{y}_p | \mathbf{t}_p) f(\mathbf{t}_p | \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k).$$

Ainsi, pour un pixel donné p , la loi conditionnelle de \mathbf{t}_p est :

$$f(\mathbf{t}_p | z_p = k, \boldsymbol{\psi}_k, \boldsymbol{\Sigma}_k, \mathbf{y}_p, s_p^2) \propto \left(\frac{1}{s_p^2} \right)^{\frac{L}{2}} \exp \left\{ -\frac{1}{2s_p^2} \|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p(\mathbf{t}_p)\|^2 \right\} \\ \times |\boldsymbol{\Sigma}_k|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{t}_p - \boldsymbol{\psi}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{t}_p - \boldsymbol{\psi}_k) \right]. \quad (\text{C.14})$$

Une étape de Metropolis-Hastings va être employée de la même manière que pour (3.24) afin de générer des échantillons suivant la loi (C.14).

C.2.3 Lois conditionnelles des variances du bruit

Pour chaque pixel p , le théorème de Bayes donne :

$$f(s_p^2 | \mathbf{y}_p, \mathbf{t}_p, \delta) \propto f(\mathbf{y}_p | \mathbf{t}_p, s_p^2) f(s_p^2 | \delta).$$

Dès lors, la loi conditionnelle de s_p^2 est une inverse Gamma :

$$s_p^2 | \mathbf{y}_p, \mathbf{t}_p, \delta \sim \mathcal{IG} \left(\frac{L}{2} + 1, \frac{\|\mathbf{y}_p - \mathbf{M}\mathbf{a}_p(\mathbf{t}_p)\|^2}{2} + \delta \right). \quad (\text{C.15})$$

C.2.4 Lois conditionnelles de Ψ et Σ

Les lois conditionnelles des deux hyperparamètres Ψ et Σ sont les mêmes que celles obtenues par le modèle NCM :

$$\psi_{r,k} | \mathbf{z} = k, \mathbf{t}_r, \sigma_{r,k}^2, v^2 \sim \mathcal{N} \left(\frac{v^2 n_k \bar{t}_{r,k}}{\sigma_{r,k}^2 + v^2 n_k}, \frac{v^2 \sigma_{r,k}^2}{\sigma_{r,k}^2 + v^2 n_k} \right) \quad (\text{C.16})$$

avec $\bar{t}_{r,k} = \frac{1}{n_k} \sum_{p \in \mathcal{I}_k} t_{r,p}$ et

$$\sigma_{r,k}^2 | \mathbf{z} = k, \mathbf{t}_r, \psi_{r,k} \sim \mathcal{IG} \left(\frac{n_k}{2} + 1, \gamma + \sum_{p \in \mathcal{I}_k} \frac{(t_{r,p} - \psi_{r,k})^2}{2} \right). \quad (\text{C.17})$$

C.2.5 Lois conditionnelles de v^2 et δ

La loi conditionnelle de v^2 est la loi inverse Gamma suivante :

$$v^2 | \Psi \sim \mathcal{IG} \left(\frac{RK}{2}, \frac{1}{2} \Psi^T \Psi \right), \quad (\text{C.18})$$

alors que la loi conditionnelle de δ est une loi Gamma :

$$\delta | \mathbf{s} \sim \mathcal{G} \left(P, \sum_{p=1}^P \frac{1}{s_p^2} \right) \quad (\text{C.19})$$

C.3 Résultats de simulations sur des données synthétiques

Un schéma de simulation identique à celui présenté au paragraphe 3.5 va être utilisée afin d'évaluer les performances de cet algorithme basé sur le modèle de mélange standard.

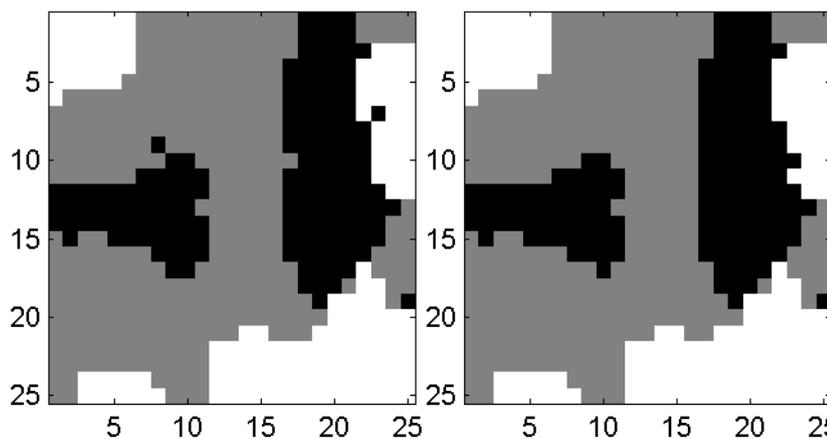


FIGURE C.2 – À gauche : carte des étiquettes réelle. À droite : carte des étiquettes estimée par l'échantillonneur de Gibbs hybride basé sur le MML.

Là encore, les étiquettes sous-jacentes de l'image ont été générées en utilisant (C.4) avec $\beta = 1.1$ (voir figure C.2). Puis les moyennes et variances des abondances ont été choisies suivant les valeurs données dans la Table C.1.

ALGO. C.1 Échantillonneur de Gibbs hybride pour le démélange spectral avec contraintes spatiales pour le modèle MML

– % *Initialisation* :

- 1: Échantillonner $\mathbf{z}^{(0)}$ à partir de la densité de probabilité (pdf) dans Eq. (C.4),
- 2: Échantillonner $\delta^{(0)}$ et $v^{2(0)}$ à partir des pdfs dans Eq. (C.11),
- 3: Échantillonner $\mathbf{s}^{(0)}$ à partir de la pdf dans Eq. (C.8),
- 4: Échantillonner $\Psi^{(0)}$ à partir de la pdf dans Eq. (C.9),
- 5: Échantillonner $\Sigma^{(0)}$ à partir de la pdf dans Eq. (C.10),

– % *Itérations* :

- 1: **pour** $t = 1, 2, \dots$ **faire**
 - 2: **pour** chaque pixel $p = 1, \dots, P$ **faire**
 - 3: Échantillonner $\mathbf{z}^{(t)}$ à partir de la pdf dans Eq. (C.13),
 - 4: Échantillonner $\mathbf{t}_p^{(t)}$ à partir de la pdf dans Eq. (C.14) using Metropolis-within-Gibbs step,
 - 5: Calculer $\mathbf{a}_p^{(t)}$ d'après Eq. (3.5),
 - 6: Échantillonner $s_p^{2(t)}$ à partir de la pdf dans Eq. (C.15),
 - 7: **fin pour**
 - 8: **pour** chaque classe $k = 1, \dots, K$ **faire**
 - 9: **pour** chaque endmember $r = 1, \dots, R$ **faire**
 - 10: Échantillonner $\psi_{r,k}^{(t)}$ à partir de la pdf dans Eq. (C.16),
 - 11: Échantillonner $\sigma_{r,k}^{2(t)}$ à partir de la pdf dans Eq. (C.17),
 - 12: **fin pour**
 - 13: **fin pour**
 - 14: Échantillonner $v^{2(t)}$ à partir de la pdf dans Eq. (C.18)
 - 15: Échantillonner $\delta^{(t)}$ à partir de la pdf dans Eq. (C.19),
 - 16: **fin pour**
-

TABLE C.1 – Moyennes et variances réelles et estimées des abondances dans chacune des classes.

		Valeurs réelles	MML
Classe 1	$\boldsymbol{\mu}_1 = \mathbf{E}[\mathbf{a}_p, p \in \mathcal{I}_1]$	$[0.6, 0.3, 0.1]^T$	$[0.58, 0.29, 0.13]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_1] (\times 10^{-3})$	$[5, 5, 5]^T$	$[4.5, 4.3, 5.5]^T$
Classe 2	$\boldsymbol{\mu}_2 = \mathbf{E}[\mathbf{a}_p, p \in \mathcal{I}_2]$	$[0.3, 0.5, 0.2]^T$	$[0.29, 0.49, 0.2]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_2] (\times 10^{-3})$	$[5, 5, 5]^T$	$[4.5, 4.7, 5.3]^T$
Classe 3	$\boldsymbol{\mu}_3 = \mathbf{E}[\mathbf{a}_p, p \in \mathcal{I}_3]$	$[0.3, 0.2, 0.5]^T$	$[0.31, 0.19, 0.49]^T$
	$\text{Var}[a_{p,r}, p \in \mathcal{I}_3] (\times 10^{-3})$	$[5, 5, 5]^T$	$[7, 4.2, 11.7]^T$

Les cartes d'abondances générées suivant le MML sont données sur la figure C.3 (haut). Les variances du bruit ont été générées suivant la loi *a priori* (C.8) avec $\delta = 1 \times 10^{-3}$. Ce qui implique que le SNR est de 19dB.

L'estimateur MAP marginal du vecteur d'étiquettes $\hat{\mathbf{z}}_{\text{MAP}}$ est représentée sur la figure 3.5 pour l'algorithme de Gibbs hybride basé sur le MML (à droite). Les estimateurs MMSE des abondances conditionnellement à $\hat{\mathbf{z}}_{\text{MAP}}$ ont également été représentés sur la figure C.3 (en bas). Un nombre de $N_{\text{MC}} = 5000$ itérations (avec 500 itérations de chauffage) a été nécessaire pour obtenir ces résultats. Comme précédemment, les histogrammes des vecteurs moyennes $\boldsymbol{\mu}_k = \mathbf{E}[\mathbf{a}_p]$ ($k = 1, \dots, K$, $p \in \mathcal{I}_k$) estimés par l'algorithme basé sur le MML ont été déterminés seulement et représentés sur la figure C.4. La table 3.1 donne les résultats des estimateurs MMSE des moyennes et variances des abondances pour chaque endmember pour chacune des classes. Les différents paramètres (classes, coefficients d'abondance et vecteurs des moyennes des abondances) estimés par les deux algorithmes sont clairement en accord avec les valeurs réelles.

Nous comparons ensuite l'algorithme de Gibbs hybride basé sur le MML au NCM « Bayésien corrélé », à leurs équivalents spatialement indépendants ainsi qu'à l'algorithme FCLS. Nous reprenons la table 3.2 (au paragraphe 3.5) à laquelle nous ajoutons une colonne. Celle-ci contient les résultats de calculs des EQM globales résultant des simulations effectués sur l'image synthétique (montrée figure C.2) à l'aide du MML « Bayésien

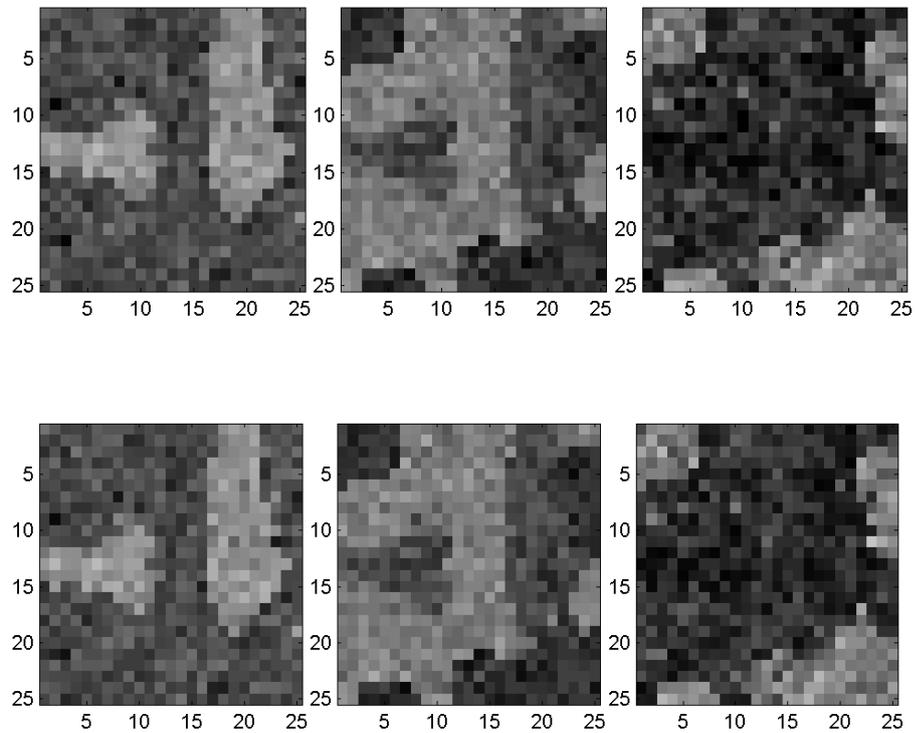


FIGURE C.3 – En haut : cartes d’abondances des 3 matériaux purs pour le MML. En bas : cartes d’abondances des 3 matériaux purs estimés par l’échantillonneur de Gibbs hybride basé sur le MML (de gauche à droite : béton de construction, herbe verte, terre grasse).

corrélé ». Ainsi, ces résultats, reportés dans la table C.2, montrent dans un premier temps que l’algorithme MML « Bayésien corrélé » présente de meilleures performances que les algorithmes FCLS et « Bayésiens indépendants ». Ensuite et de manière générale, ils démontrent l’intérêt d’exploiter les contraintes spatiales aussi bien pour le modèle standard de mélange que pour le modèle NCM.

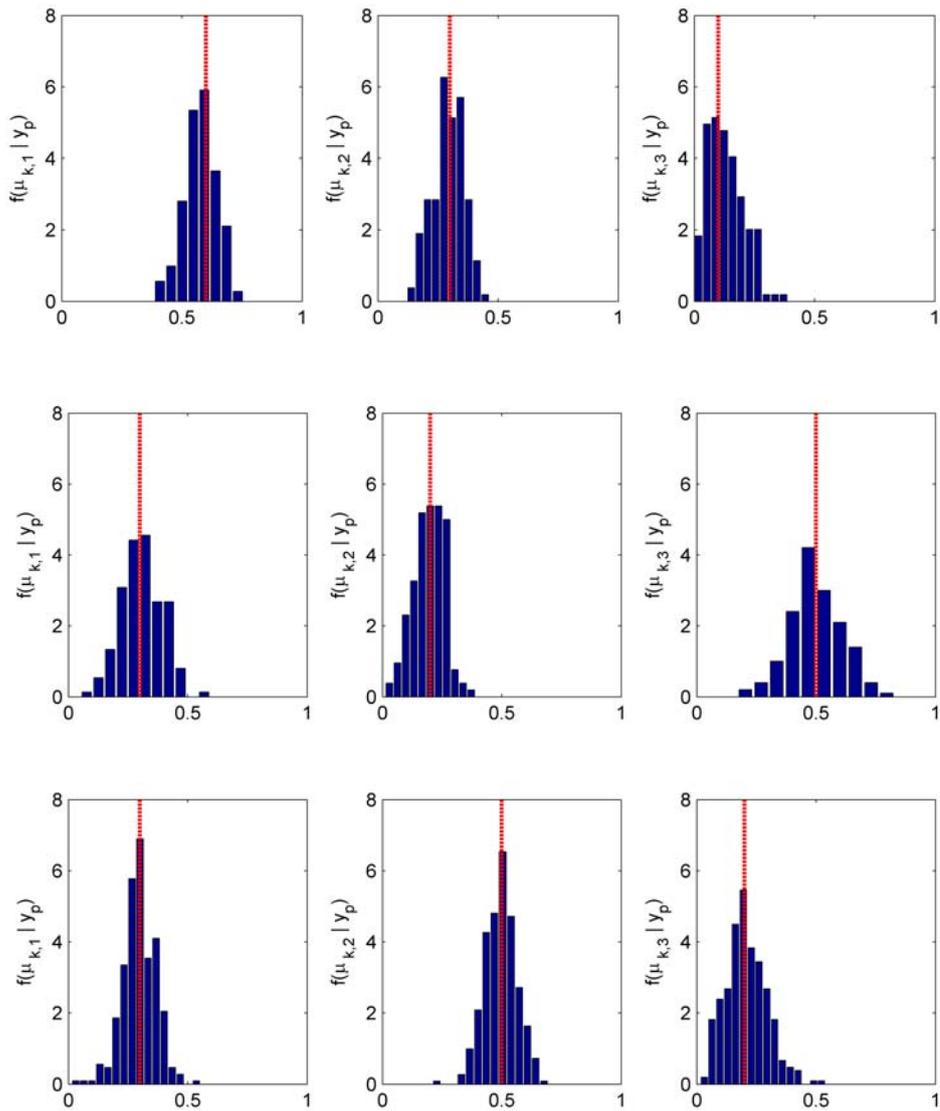


FIGURE C.4 – Histogrammes des moyennes d’abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ estimés par l’algorithme de Gibbs hybride basé sur le MML pour (de haut en bas) les $K = 3$ classes (valeurs réelles en pointillés).

TABLE C.2 – EQMs globales pour chaque composante des vecteurs d’abondances.

	NCM		MML		
	Bayésien in-dépendant	Bayésien corrélé	FCLS	Bayésien in-dépendant	Bayésien corrélé
EQM_1^2	0.0027	0.0013	0.0019	0.0016	0.001
EQM_2^2	8.9×10^{-4}	5.9×10^{-4}	4.3×10^{-4}	4.1×10^{-4}	3.1×10^{-4}
EQM_3^2	9.5×10^{-4}	5.5×10^{-4}	0.0014	0.0013	8.6×10^{-4}

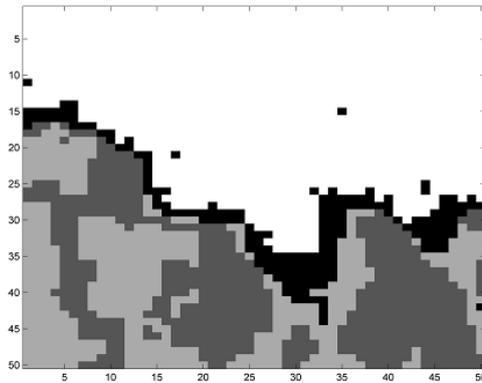


FIGURE C.5 – Carte des étiquettes estimées par l’algorithme basé sur le MML.

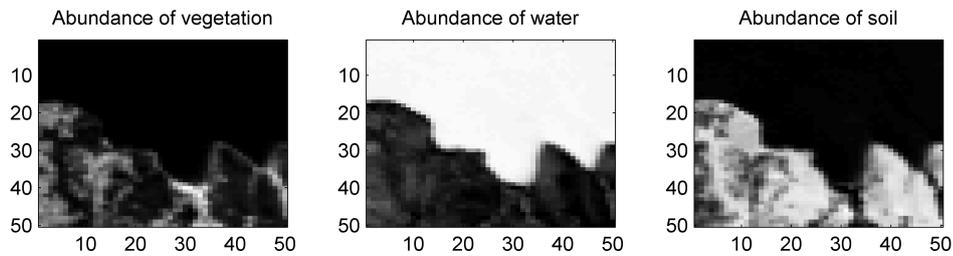


FIGURE C.6 – Cartes d’abondances estimées par l’algorithme basé sur le MML (de gauche à droite : végétation, eau et terre).

C.4 Résultats de simulations sur une image AVIRIS

L’image Moffett field va encore être mise à contribution pour évaluer les performances sur des données réelles de l’échantillonneur de Gibbs hybride basé sur le MML. Le schéma de simulation reste identique au paragraphe 3.6.

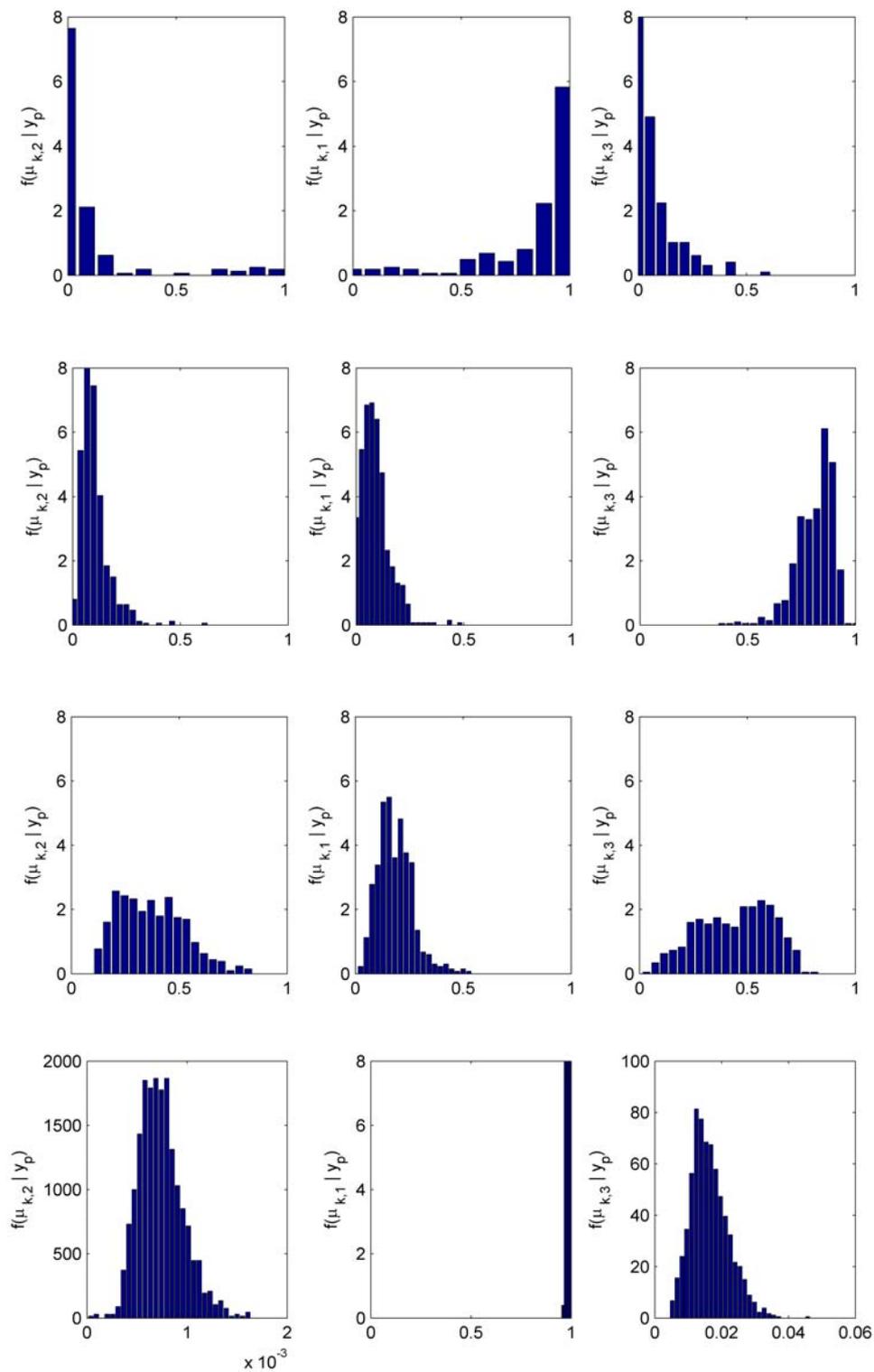


FIGURE C.7 – Histogrammes des moyennes d'abondances $\boldsymbol{\mu}_k = [\mu_{k,1}, \mu_{k,2}, \mu_{k,3}]^T$ (de gauche à droite : végétation, eau et terre) estimés par l'algorithme de Gibbs hybride pour les $K = 4$ classes.

La carte de classification obtenue par l'estimation des étiquettes est représentée sur la figure C.5, alors que les cartes d'abondances estimés sont données sur la figure C.6. Les lois *a posteriori* des moyennes des abondances $\boldsymbol{\mu}_k = \mathbf{E}[\mathbf{a}_p]$ ($k = 1, \dots, K$, $p \in \mathcal{I}_k$) ont également été représentées sous formes d'histogrammes sur la figure C.7.

Ces résultats restent très comparables à ceux obtenus pour l'algorithme basé sur le NCM. Dès lors, si l'image contient un nombre largement suffisant de pixels purs, il sera possible d'utiliser l'algorithme développé ci-dessus.

Liste des publications

Revues

1. O. Eches, N. Dobigeon and J.-Y. Tournet, “Enhancing hyperspectral image unmixing with spatial correlations,” submitted to *IEEE Transactions on Geoscience and Remote Sensing*, 2009. Technical report : http://olivier.eches.free.fr/papers/TechReport_TGARS2010.pdf
2. O. Eches, N. Dobigeon and J.-Y. Tournet, “Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm.”, *IEEE Journal of Selected Topics on Signal Processing*, vol. 3, no. 3, pp. 582-591, June 2010. Technical report : http://olivier.eches.free.fr/papers/TechReport_JSTSP2009.pdf
3. O. Eches, N. Dobigeon, Corinne Mailhes and J.-Y. Tournet “Bayesian Estimation of Linear Mixtures using the Normal Compositional Model. Application to Hyperspectral Imagery.”, *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1403-1413, June 2010. Technical report : http://olivier.eches.free.fr/papers/TechReport_2009.pdf

Conférences

1. O. Eches, N. Dobigeon and J.-Y. Tournet, “Markov random fields for joint unmixing and segmentation of hyperspectral images.” *Int. Workshop on Bayesian Inference and Maximum Entropy (MaxEnt)*, 2010, submitted.

2. O. Eches, N. Dobigeon and J.-Y. Tournet, "Markov random fields for joint unmixing and segmentation of hyperspectral images." to appear in *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Reykjavik,Iceland, 2010.
3. O. Eches, N. Dobigeon and J.-Y. Tournet, "A reversible-jump algorithm for estimating the number of endmembers in the normal compositional model. Application to the unmixing of hyperspectral images." in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, Dallas, USA, 2010,pp.1222-1225.
4. O. Eches, N. Dobigeon and J.-Y. Tournet, "NCM-based Bayesian algorithm for hyperspectral unmixing." in *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Grenoble, France, 2009, pp.1-4.
5. O. Eches, N. Dobigeon and J.-Y. Tournet, "Analyse d'images hyperspectrales à l'aide d'un modèle de mélange de spectres aléatoires." in *Actes du XXIIIème Colloque GRETSI*, Dijon, France, 2009.
6. O. Eches, N. Dobigeon, C. Mailhes and J.-Y. Tournet, "Unmixing hyperspectral images using a Normal Compositional Model and MCMC methods", in *Proc. IEEE Workshop on Statistical Signal Processing (SSP)*, Cardiff, Wales, UK, 2009, pp.646-649.

Bibliographie

- [AD99] C. Andrieu and A. Doucet. Joint Bayesian model selection and estimation of noisy sinusoids via reversible jump MCMC. *IEEE Trans. Signal Processing*, 47(10) :2667–2676, Oct. 1999.
- [Aka73] H. Akaike. Information theory and an extension of the maximum likelihood principle. In *Proc. 2nd Int. Symp. Inf. Theory*, pages 267–281, 1973.
- [Aka74] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Autom. Contr.*, 19(6) :716–723, Dec. 1974.
- [ASL10] M. Arngren, M. N. Schmidt, and J. Larsen. Unmixing of hyperspectral images using bayesian nonnegative matrix factorization with volume prior. In *Journal of Signal Processing Systems*, Jan. 2010.
- [AZB07] J. Skaf M. Parente A. Zymnis, S.-J. Kim and S. Boyd. Hyperspectral image unmixing via alternating projected subgradients. In *Proc. of the 41st IEEE Asilomar Conference on Signals, Systems and Computers*, pages 1164–1168, 2007.
- [BDN08] J. M. Bioucas-Dias and J. M. P. Nascimento. Hyperspectral subspace identification. *IEEE Trans. Geosci. and Remote Sensing*, 46(8) :2435–2445, Aug. 2008.
- [Bes74] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Stat. Soc. Ser. B*, 36(2) :192–236, 1974.

- [BGC97] J. Bayliss, J. A. Gualtieri, and R. F. Cromp. Analyzing hyperspectral data with independent component analysis. In *Proc. SPIE AIPR Workshop*, volume 9, pages 133–143. SPIE, 1997.
- [BKL⁺04] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington. ICE : a statistical approach to identifying endmembers in hyperspectral images. *IEEE Trans. Geosci. and Remote Sensing*, 42(10) :2085 – 2095, Oct. 2004.
- [BMD08] N. Bali and A. Mohammad-Djafari. Bayesian approach with hidden Markov modeling and mean field approximation for hyperspectral data analysis. *IEEE Trans. Image Processing*, 17(2) :217–225, Feb. 2008.
- [Boa93] J. Boardman. Automating spectral unmixing of AVIRIS data using convex geometry concepts. In *Summaries 4th Annu. JPL Airborne Geoscience Workshop*, volume 1, pages 11–14, Washington, D.C., 1993. JPL Pub.
- [BPS05] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson. Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Trans. Geosci. and Remote Sensing*, 43(3) :480–491, March 2005.
- [CD04] C.-I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Trans. Geosci. and Remote Sensing*, 42(3) :608–619, March 2004.
- [CFP03] G. Celeux, F. Forbes, and N. Peyrard. EM procedures using mean field-like approximations for Markov model-based image segmentation. *Pattern Recognition*, 36(1) :131–144, Jan. 2003.
- [Che98] R. C. H. Cheng. Bayesian model selection when the number of components is unknown. *Winter Simulation Conference*, 1 :653–659, 1998.
- [CJH91] P. Comon, C. Jutten, and J. Herault. Blind separation of sources, part ii : problems statement. *Signal Processing*, 24(1) :11–20, 1991.

- [Cla03] R. N. Clark *et al.* Imaging spectroscopy : Earth and planetary remote sensing with the USGS Tetracorder and expert systems. *J. Geophys. Res.*, 108(E12) :5–1–5–44, Dec. 2003.
- [CSB⁺02] C. Champagne, K. Staenz, A. Bannari, H. P. White, J.-C. Deguise, and H. McNairn. Estimation of plant water content of agricultural canopies using hyperspectral remote sensing. In *1st International Symposium on Recent Advances in Quantitative Remote Sensing*, Valencia, Spain, Sept. 2002.
- [CSG93] R. N. Clark, G. A. Swayze, and A. Gallagher. Mapping minerals with imaging spectroscopy, U.S. Geological Survey. *Office of Mineral Resources Bulletin*, 2039 :141–150, 1993.
- [CXAP98] C.-I Chang, X.-L Zhao, M. L. G. Althouse, and J. J. Pan. Least squares subspace projection approach to mixed pixel classification for hyperspectral images. *IEEE Trans. Geosci. and Remote Sensing*, 36(3) :898–912, May 1998.
- [DA05] N. Dobigeon and V. Achard. Performance comparison of geometric and statistical methods for endmembers extraction in hyperspectral imagery. volume 5982, pages 335–344. SPIE, Oct. 2005.
- [DGI06] M. Davy, S. Godsill, and J. Idier. Bayesian analysis of polyphonic western tonal music. *J. Acoust. Soc. Am.*, 119(4) :2498–2517, April 2006.
- [DI96] J. Diebolt and E. H. S. Ip. Stochastic EM : method and application. In W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors, *Markov chain Monte Carlo in Practice*. Chapman & Hall, London, 1996.
- [DMC⁺09] N. Dobigeon, S. Moussaoui, M. Coulon, J.-Y. Tourneret, and A. O. Hero. Joint Bayesian endmember extraction and linear unmixing for hyperspectral imagery. *IEEE Trans. Signal Processing*, 57(11) :4355–4368, Nov. 2009.
- [DMZB99] X. Descombes, R. Morris, J. Zerubia, and M. Berthod. Estimation of Markov random field prior parameters using Markov chain Monte Carlo maximum likelihood. *IEEE Trans. Image Processing*, 8(7) :945–963, July 1999.

- [Dob07] N. Dobigeon. *Modèles Bayésiens hiérarchiques pour le traitement multi-capteur*. PhD thesis, INP-ENSEEIH, Toulouse, France, October 2007.
- [DTC08] N. Dobigeon, J.-Y. Tournet, and C.-I Chang. Semi-supervised linear spectral using a hierarchical Bayesian model for hyperspectral imagery. *IEEE Trans. Signal Processing*, 56(7) :2684–2696, July 2008.
- [DTD07] N. Dobigeon, J.-Y. Tournet, and M. Davy. Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach. *IEEE Trans. Signal Processing*, 55(4) :1251–1263, April 2007.
- [DTH08] N. Dobigeon, J.-Y. Tournet, and A. O. Hero. Bayesian linear unmixing of hyperspectral images corrupted by colored Gaussian noise with unknown covariance matrix. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, pages 3433–3436, Las Vegas, USA, March 2008.
- [EDMT10] O. Eches, N. Dobigeon, C. Mailhes, and J.-Y. Tournet. Bayesian estimation of linear mixtures using the normal compositional model. Application to hyperspectral imagery,. *IEEE Trans. Image Processing*, 19(6) :1403–1413, Jun 2010. See also technical report available at <http://eches.perso.enseeiht.fr>.
- [EDT10a] O. Eches, N. Dobigeon, and J.-Y. Tournet. Enhancing hyperspectral image unmixing with spatial correlations. *IEEE Trans. Geosci. and Remote Sensing*, 2010. submitted.
- [EDT10b] O. Eches, N. Dobigeon, and J.-Y. Tournet. Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical Bayesian algorithm. *IEEE J. Sel. Topics Signal Processing*, 3(3) :582–591, Jun. 2010.
- [ES07] M. T. Eismann and D. Stein. Stochastic mixture modeling. In C.-I Chang, editor, *Hyperspectral Data Exploitation : Theory and Applications*, chapter 5. Wiley, 2007.

- [GBJ96] A. Gelman, F. Bois, and J. Jiang. Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *J. Amer. Math. Soc.*, 91(436) :1400–1412, Dec. 1996.
- [GBSC94] A. Green, M. Berman, P. Switzer, and M. D. Craig. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. Geosci. and Remote Sensing*, 26(1) :65–74, Jan. 1994.
- [GG84] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 6(6) :721–741, Nov. 1984.
- [Gre95] P. J. Green. Reversible jump Markov chain Monte Carlo methods computation and Bayesian model determination. *Biometrika*, 82(4) :711–732, Dec. 1995.
- [Gre98] R. O. Green *et al.* Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sens. Environ.*, 65(3) :227–248, Sept. 1998.
- [GRS96] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. Introducing Markov chain Monte Carlo. In W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors, *Markov chain Monte Carlo in Practice*, pages 5–8. Chapman & Hall, London, 1996.
- [GVSR85] A. F.H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock. Imaging spectrometry for Earth remote sensing. *Science*, 228(4704) :1147–1153, 1985.
- [Hap81] B. Hapke. Bidirectional reflectance spectroscopy 1. theory. *J. Geophys. Res.*, 86 :3039–3054, 1981.
- [HC94] J. Harsanyi and C.-I Chang. Hyperspectral image classification and dimensionality reduction : An orthogonal subspace projection approach. *IEEE Trans. Geosci. and Remote Sensing*, 32 :779–785, July 1994.
- [HC01] D. C. Heinz and C.-I Chang. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. and Remote Sensing*, 39(3) :529–545, March 2001.

- [HSY⁺05] K. Homma, M. Shibayama, H. Yamamoto, K. Sugahara, and H. Shingu. Water pollution monitoring using a hyperspectral imaging spectropolarimeter. volume 5655, pages 419–426. SPIE, Jan. 2005.
- [Jef46] H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proc. of the Royal Society of London. Series A*, 186(1007) :453–461, 1946.
- [JJ00] T. S. Jaakola and M. Jordan. Bayesian parameter estimation via variational methods. *Statistics and Computing*, 10(1) :25–37, Jan. 2000.
- [Jol86] I. T. Jolliffe. *Principal Component Analysis*. Springer Verlag, New York, 1986.
- [JSTGA83] P. E. Johnson, M. O. Smith, S. Taylor-George, and J. B. Adams. A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures. *J. Geophys. Res.*, 88 :3557–3562, April 1983.
- [Jus09] K. Jusoff. Precision forestry using airborne hyperspectral imaging sensor. *Journal of Agricultural Science*, 1(1) :142–147, June 2009.
- [KH95] C. Kevrann and F. Heitz. A Markov random field model-based approach to unsupervised texture segmentation using local and global statistics. *IEEE Trans. Image Processing*, 4(6) :856–862, June 1995.
- [KH03] S. Kumar and M. Hebert. Discriminative random fields : A discriminative framework for contextual interaction in classification. In *International Conference on Computer Vision*, pages 1150–1157, 2003.
- [KM02] N. Keshava and J.F. Mustard. Spectral unmixing. *IEEE Signal Processing Magazine*, pages 44–56, Jan. 2002.
- [Kru98] F. A. Kruse. Advances in hyperspectral remote sensing for geologic mapping and exploration. In *Proceedings 9th Australasian Remote Sensing Conference*, Sydney, Australia, July 1998.
- [KS80] R. Kindermann and J. L. Snell. *Markov random fields and their applications*. Amer. Math. Soc., Providence, RI, 1980.

- [LB05] E. Levina and P. J. Bickel. Maximum likelihood estimation of intrinsic dimension. In Lawrence K. Saul, Yair Weiss, and Léon Bottou, editors, *Advances in Neural Information Processing Systems 17*, pages 777–784. MIT Press, Cambridge, MA, 2005.
- [LCDC09] B. Luo, J. Chanussot, S. Douté, and X. Ceamanos. Unsupervised endmember extraction of Martian hyperspectral images. In *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, Grenoble, France, Aug. 2009.
- [LWC⁺07] E. Levina, A. S. Wagaman, A. F. Callender, G. S. Mandair, and M. D. Morris. Estimating the number of pure chemical components in a mixture by maximum likelihood. *Journal of Chemometrics*, 21(1–2) :24–34, Jan. 2007.
- [Mal88] E. R. Malinowski. Statistical F-tests for abstract factor analysis and target testing. *Journal of Chemometrics*, 3(1) :49–60, 1988.
- [MFM04] A. Mohammadpour, O. Féron, and A. Mohammad-Djafari. Bayesian segmentation of hyperspectral images. In R. Fischer, R. Preuss, and U. V. Toussaint, editors, *Proc. Workshop on Bayesian Inference and Maximum Entropy Methods (MaxEnt)*, volume 735 of *AIP Conference Series*, pages 541–548, Max Plank Inst., Germany, Nov. 2004.
- [MH10] R. Mittelman and A. O. Hero. Hyperspectral image segmentation and unmixing using hidden markov trees. In *Proc. IEEE Int. Conf. Image Processing*, 2010. submitted.
- [MHS⁺08] S. Moussaoui, H. Hauksdóttir, F. Schmidt, C. Jutten, J. Chanussot, D. Brie, S. Douté, and J. A. Benediksson. On the decomposition of Mars hyperspectral data by ICA and Bayesian positive source separation. *Neurocomputing*, 71(10–12) :2194–2208, 2008.
- [MMS03] D. Manolakis, D. Marden, and G. Shaw. Hyperspectral image processing for

- automatic target detection applications. *Lincoln Laboratory Journal*, 14(1) :79–116, 2003.
- [MQ07] L. Miao and H. Qi. Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. *IEEE Trans. Geosci. and Remote Sensing*, 45(3) :765–776, March 2007.
- [MR07] J.-M. Marin and C. P. Robert. *Bayesian core : a practical approach to computational Bayesian statistics*. Springer, New-York, 2007.
- [MSS01] D. Manolakis, C. Siracusa, and G. Shaw. Hyperspectral subpixel target detection using the linear mixing model. *IEEE Trans. Geosci. and Remote Sensing*, 39(7) :1392–1409, July 2001.
- [NB05] J. M. P. Nascimento and J. M. Bioucas-Dias. Vertex component analysis : a fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. and Remote Sensing*, 43(4) :898–910, April 2005.
- [NBD05] J. M. P. Nascimento and J. M. Bioucas-Dias. Does Independent Component Analysis play a role in unmixing hyperspectral data? *IEEE Trans. Geosci. and Remote Sensing*, 43(1) :175–187, Jan. 2005.
- [ND04] J. M. P. Nascimento and J. M. B. Dias. Does independent component analysis play a role in unmixing hyperspectral data? *Pattern Recognition and Image Analysis, ser. Lecture Notes in Computer Science*, 43 :175–187, 2004.
- [ND07] J. M. P. Nascimento and J. M. B. Dias. Hyperspectral unmixing algorithm via dependent component analysis. In *Proc. IEEE Int. Conf. Geoscience and Remote Sensing*, pages 4033–4036, July 2007.
- [NE08] R. R. Nadakuditi and A. Edelman. Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Trans. Signal Processing*, 56(7) :2625–2638, July 2008.

- [NS05] R. Neher and A. Srivastava. A Bayesian MRF framework for labeling terrain using hyperspectral imaging. *IEEE Trans. Geosci. and Remote Sensing*, 43(6) :1363–1374, June 2005.
- [PADF02] E. Punsakaya, C. Andrieu, A. Doucet, and W. Fitzgerald. Bayesian curve fitting using MCMC with applications to signal segmentation. *IEEE Trans. Signal Processing*, 50(3) :747–758, March 2002.
- [Pla06] N. Playle. Detection of landmines using hyperspectral imaging. volume 6217. SPIE, May 2006.
- [Pla09] A. Plaza. High performance computing in remote sensing applications. In *First ComplexHPC Meeting*, Lisbon, Portugal, Oct. 2009. Technical University of Lisbon (IST). <http://complexhpc.org/events/lisbon/PDF/Plaza.pdf>.
- [PMPP02] A. Plaza, P. Martinez, R. Perez, and J. Plaza. Spatial/Spectral endmember extraction by multidimensional morphological operations. *IEEE Trans. Geosci. and Remote Sensing*, 40(9) :2025–2041, Sept. 2002.
- [PMPP04] A. Plaza, P. Martinez, R. Perez, and J. Plaza. A quantitative and comparative analysis of endmember extraction algorithms for hyperspectral data. *IEEE Trans. Geosci. and Remote Sensing*, 42(3) :650–663, March 2004.
- [PMS⁺00] L. Parra, K.-R. Mueller, C. Spence, A. Ziehe, and P. Sajda. Unmixing hyperspectral data. In *Advances in Neural Information Processing Systems (NIPS)*, volume 12, pages 942–948. MIT Press, 2000.
- [PO04] M. D. Plumbley and E. Oja. A “nonnegative-PCA” algorithm for independent component analysis. *IEEE Trans. Neural Networks*, 15(1) :66–76, Jan. 2004.
- [PT94] P. Paatero and U. Tapper. Positive matrix factorization : a non-negative factor model with optimal utilization of error estimates of data values. In *Fourth International Conference on Statistical methods for the Environmental Sciences (Environmetrics)*, volume 5, pages 111–126, Aug. 1994.

- [RC04] C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer Verlag, New York, second edition, 2004.
- [RDFZ04] G. Rellier, X. Descombes, F. Falzon, and J. Zerubia. Texture feature analysis using a Gauss-Markov model in hyperspectral image classification. *IEEE Trans. Geosci. and Remote Sensing*, 42(7) :1543–1551, July 2004.
- [RG97] S. Richardson and P. J. Green. On Bayesian analysis of mixtures with an unknown number of components. *J. Roy. Stat. Soc. Ser. B*, 59(4) :731–792, 1997.
- [RG98] S. Richardson and P. J. Green. Corrigendum : On Bayesian analysis of mixtures with an unknown number of components. *J. Roy. Stat. Soc. Ser. B*, 60(3) :661, 1998.
- [Ris78] J. Rissanen. Modeling by shortest data description. *Automatica*, 14(5) :465–471, 1978.
- [RK03] R. S. Rand and D. M. Keenan. Spatially smooth partitioning of hyperspectral imagery using spectral/spatial measures of disparity. *IEEE Trans. Geosci. and Remote Sensing*, 41(6) :1479–1490, June 2003.
- [Rob96] G. O. Roberts. Markov chain concepts related to sampling algorithms. In W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors, *Markov Chain Monte Carlo in Practice*, pages 45–57. Chapman & Hall, London, 1996.
- [Row02] J. B. Rowe. A Bayesian approach to blind source separation. *J. of Interdisciplinary Mathematics*, 5(1) :49–76, 2002.
- [RSI03] RSI (Research Systems Inc.). *ENVI User's guide Version 4.0*. Boulder, CO 80301 USA, Sept. 2003.
- [SDB⁺04] P. Sajda, S. Du, T. R. Brown, R. Stoyanova, D. C. Shungu, M. Xiangling, and L.C. Parra. Nonnegative matrix factorization for rapid recovery of constituent spectra in magnetic resonance chemical shift imaging of the brain. *IEEE Transactions on Medical Imaging*, 23(12) :1453–1465, Dec. 2004.

- [Set96] J. J. Settle. On the relationship between spectral unmixing and subspace projection. *IEEE Trans. Geosci. and Remote Sensing*, 34(4) :1045–1046, July 1996.
- [SM79] R. B. Singer and T. B. McCord. Mars : Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. In *Proc. Lunar Planet. Sci. Conf.*, volume 2, pages 1835–1848, 1979.
- [SMD09] F. Schmidt, S. Moussaoui, and N. Dobigeon. Material identification on martian hyperspectral images using Bayesian source separation. In *Proc. IEEE GRSS Workshop on Hyperspectral Image and Signal Processing : Evolution in Remote Sensing (WHISPERS)*, pages 1–4, Grenoble, France, Aug. 2009.
- [Ste03] D. Stein. Application of the normal compositional model to the analysis of hyperspectral imagery. In *Proc. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 44–51, Oct. 2003.
- [TBS06] A. Tonazzini, L. Bedini, and E. Salerno. A Markov model for blind image separation by a mean-field EM algorithm. *IEEE Trans. Image Processing*, 15(2) :473–481, Feb. 2006.
- [TDTL09] C. Theys, N. Dobigeon, J.-Y. Tourneret, and H. Lantéri. Linear unmixing of hyperspectral images using a scaled gradient method. In *Proc. IEEE-SP Workshop Stat. and Signal Processing (SSP)*, pages 729–732, Cardiff, UK, Aug. 2009.
- [TR08] K. Themelis and A. A. Rontogiannis. A soft constrained MAP estimator for supervised hyperspectral signal unmixing. In *Proc. of European Signal Processing Conf. (EUSIPCO)*, Lausanne, Switzerland, Aug. 2008.
- [TRK10] K. Themelis, A. Rontogiannis, and K. Koutroumbas. Semi-supervised hyperspectral unmixing via the weighted lasso. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc. (ICASSP)*, Dallas, USA, April 2010.
- [VPPP06] J. Piper V. P. Pauca and R. J. Plemmons. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra and its Applications*, 416 :29–47, 2006.

- [Win99] M. E. Winter. Fast autonomous spectral endmember determination in hyperspectral data. In *Proc. 13th Int. Conf. on Applied Geologic Remote Sensing*, volume 2, pages 337–344, Vancouver, April 1999.
- [WK85] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Trans. Acoust., Speech, Signal Processing*, 33(2) :387–392, Feb. 1985.
- [Wu82] F. Y. Wu. The Potts model. *Rev. Modern Phys.*, 54(1) :235–268, Jan. 1982.
- [ZG10] A. Zare and P. Gader. Pce : Piece-wise convex endmember detection. *IEEE Trans. Geosci. and Remote Sensing*, 2010. to appear.
- [ZLQ97] Z. Zhou, R. Leahy, and J. Qi. Approximate maximum likelihood hyperparameter estimation for Gibbs prior. *IEEE Trans. Image Processing*, 6(6) :844–861, June 1997.

