

Méthodes avancées de séparation de sources applicables aux mélanges linéaires-quadratiques

Lina Jarboui

▶ **To cite this version:**

Lina Jarboui. Méthodes avancées de séparation de sources applicables aux mélanges linéaires-quadratiques. Traitement du signal et de l'image [eess.SP]. Université Paul Sabatier - Toulouse III; Université de Sfax (Tunisie), 2017. Français. NNT : 2017TOU30295 . tel-01988590

HAL Id: tel-01988590

https://theses.hal.science/tel-01988590

Submitted on 21 Jan 2019

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.

République Tunisienne
Ministère de l'Enseignement Supérieur, de la Recherche Scientifique et de la Technologie

Université de Sfax
École Nationale d'Ingénieurs de Sfax

Université Paul Sabatier Toulouse III

Ecole Doctorale
Sciences et Technologies
Thèse de *DOCTORAT en*
Ingénierie des Systèmes Informatiques

Ecole Doctorale MITT
Thèse de *DOCTORAT en*
Signal, Image, Acoustique et Optimisation

N° d'ordre : 2017– 76

THESE

Présentée à

L'École Nationale d'Ingénieurs de Sfax

Pour obtenir le grade de

DOCTEUR DE L'ECOLE NATIONALE D'INGENIEURS DE SFAX
DE L'UNIVERSITE DE SFAX
Discipline : Ingénierie des Systèmes Informatiques
et
DOCTEUR DE L'UNIVERSITE PAUL SABATIER TOULOUSE III
Discipline : Signal, Image, Acoustique et Optimisation

Par

Lina JARBOUI

(Master NTSID et Ingénieur Informatique)

Méthodes avancées de séparation de sources applicables aux mélanges linéaires-quadratiques

Soutenu le 18 Novembre 2017, devant le jury composé de :

M.	**Riadh ABDELFETTAH (Professeur, SUPCOM, Université Tunis, Tunisie)**	Président
Mme	**Amel BEN AZZA (Professeur, SUPCOM, Université Tunis, Tunisie)**	Rapporteur
M.	**Ali MANSOUR (Professeur, ENSTA, Bretagne, France)**	Rapporteur
M.	**Ahmed BEN HAMIDA (Professeur, ENIS, Université de Sfax, Tunisie)**	Directeur de Thèse
M.	**Yannick DEVILLE (Professeur Université Paul Sabatier, Toulouse, France)**	Co-Directeur de Thèse
M.	**Shahram HOSSEINI (Maître de conférences, Université Paul Sabatier, Toulouse, France)**	Encadrant
Mme	**Rima GUIDARA (Maître assistant, ENET'COM, Université de Sfax, Tunisie)**	Invité

République Tunisienne
Ministère de l'Enseignement Supérieur, de la Recherche Scientifique et de la Technologie

Université de Sfax
École Nationale d'Ingénieurs de Sfax

Université Paul Sabatier Toulouse III

Ecole Doctorale
Sciences et Technologies
Thèse de *DOCTORAT en*
Ingénierie des Systèmes Informatiques
Ecole Doctorale MITT
Thèse de *DOCTORAT en*
Signal, Image, Acoustique et Optimisation
N° d'ordre : 2017– 76

Méthodes avancées de séparation de sources applicables aux mélanges linéaires-quadratiques

Lina JARBOUI

Résumé : Dans cette thèse, nous nous sommes intéressés à proposer de nouvelles méthodes de Séparation Aveugle de Sources (SAS) adaptées aux modèles de mélange non-linéaires. La SAS consiste à estimer les signaux sources inconnus à partir de leurs mélanges observés lorsqu'il existe très peu d'informations disponibles sur le modèle de mélange. La contribution méthodologique de cette thèse consiste à prendre en considération les interactions non-linéaires qui peuvent se produire entre les sources en utilisant le modèle linéaire-quadratique (LQ). A cet effet, nous avons développé trois nouvelles méthodes de SAS. La première méthode vise à résoudre le problème du démélange hyperspectral en utilisant un modèle linéaire-quadratique. Celle-ci se repose sur la méthode d'Analyse en Composantes Parcimonieuses (ACPa) et nécessite l'existence des pixels purs dans la scène observée. Dans le même but, nous proposons une deuxième méthode du démélange hyperspectral adaptée au modèle linéaire-quadratique. Elle correspond à une méthode de Factorisation en Matrices Non-négatives (FMN) se basant sur l'estimateur du Maximum A Posteriori (MAP) qui permet de prendre en compte les informations a priori sur les distributions des inconnus du problème afin de mieux les estimer. Enfin, nous proposons une troisième méthode de SAS basée sur l'analyse en composantes indépendantes (ACI) en exploitant les Statistiques de Second Ordre (SSO) pour traiter un cas particulier du mélange linéaire-quadratique qui correspond au mélange bilinéaire.

Abstract: In this thesis, we were interested to propose new Blind Source Separation (BSS) methods adapted to the nonlinear mixing models. BSS consists in estimating the unknown source signals from their observed mixtures when there is little information available on the mixing model. The methodological contribution of this thesis consists in considering the non-linear interactions that can occur between sources by using the linear-quadratic (LQ) model. To this end, we developed three new BSS methods. The first method aims at solving the hyperspectral unmixing problem by using a linear-quadratic model. It is based on the Sparse Component Analysis (SCA) method and requires the existence of pure pixels in the observed scene. For the same purpose, we propose a second hyperspectral unmixing method adapted to the linear-quadratic model. It corresponds to a Non-negative Matrix Factorization (NMF) method based on the Maximum A Posteriori (MAP) estimate allowing to take into account the available prior information about the unknown parameters for a better estimation of them. Finally, we propose a third BSS method based on the Independent Component Analysis (ICA) method by using the Second Order Statistics (SOS) to process a particular case of the linear-quadratic mixture that corresponds to the bilinear one.

Mots clés: Séparation Aveugle de Sources, Analyse en Composantes Parcimonieuses, Factorisation en Matrices Non-négatives, Analyse en Composantes Indépendantes, Démélange Hyperspectral, Modèle linéaire-quadratique.

Key-words: Blind Source Separation, Sparse Component Analysis, Non-Negative Matrix Factorization, Independent Component Analysis, Hyperspectral Unmixing, Linear-Quadratic Model.

“Une idée en appelle toujours une autre. Puis une autre encore.”

Remerciements

Ces travaux de doctorat ont été effectués dans le cadre d'une cotutelle élaborée entre l'Ecole Nationale d'Ingénieurs de Sfax (ENIS), Université de Sfax et l'Universié Paul Sabatier Toulouse III. Cette thése m'a donné l'occasion de vivre une double expérience, extrêmement enrichissante sur le plan professionnel comme personnel.

Je tiens ici à remercier toutes les personnes qui m'ont accompagnée, au cours de ces années, dans cette grande aventure, scientifique et humaine.

Mes premiers remerciements s'adressent à mon directeur de thèse, le professeur Ahmed BEN HAMIDA, pour avoir accepté de diriger mes travaux de thèse, pour ses qualités professionnelles et humaines, sa patience, son écoute et ses conseils judicieux. Je le remercie pour tout son soutien, et aussi pour son aide pendant ces années.

Je tiens à remercier vivement mon co-directeur, le professeur Yannick DEVILLE, pour son accueil au sein de son équipe, pour ses conseils, et pour les riches réunions d'encadrement que nous avons eues pendant mes séjours à l'IRAP. Ses qualités scientifiques et son expérience dans le domaine de la recherche ont été de précieux atouts durant mes années de thèse.

Je remercie chaleureusement mon encadrant, Monsieur Shahram HOSSEINI, pour les discussions, les suggestions, les commentaires et surtout pour m'avoir fait bénéficier de ses idées constructives qui m'ont permis de résoudre la plupart des problèmes rencontrés au cours de la thèse. Je le remercie également de m'avoir toujours aidé à gérer mes documents administratifs, et d'être toujours disponible et à l'écoute quand j'en avais besoin. Ces quelques lignes sont insuffisantes pour lui exprimer ma gratitude.

Je remercie également ma co-encadrante, Madame Rima GUIDARA, pour ses encouragements, ses qualités humaines, ses conseils avisés et ses remarques pertinentes.

Je remercie Monsieur Riadh ABDELFETTAH qui m'a fait l'honneur de présider le jury de cette thèse.

Je souhaite remercier particuliérement Madame Amel BEN AZZA et Monsieur Ali MANSOUR pour avoir accepté d'être rapporteurs de ma thèse. Je leur suis reconnaissante pour le temps qu'ils ont consacré et l'intérêt qu'ils ont bien voulu porter à mon travail.

Je remercie mes collègues de l'ATMS et de l'IRAP, avec qui j'ai passé de très bons moments de convivialité et j'ai pu nouer des contacts et des amitiés et avec qui je garderai de très bons souvenirs.

Je profite également de cette occasion pour remercier tous les enseignants et le personnel de l'ESGT-CNAM avec qui j'ai le plaisir de travailler en tant qu'ATER.

Je les remercie pour leurs encouragements, leur écoute, leur gentillesse et leur bonne humeur qui m'ont été d'une grande aide au quotidien.

Enfin, je remercie chaleureusement ma famille pour la confiance, la tendresse, le soutien et l'amour qu'ils m'apportent chaque jour. Un grand merci à ma mère pour les sacrifices qu'elle a faits pour faire de moi ce que je suis aujourd'hui.

Résumé

Dans cette thèse, nous nous sommes intéressés à proposer de nouvelles méthodes de Séparation Aveugle de Sources (SAS) adaptées aux modèles de mélange non-linéaires. La SAS consiste à estimer les signaux sources inconnus à partir de leurs mélanges observés lorsqu'il existe très peu d'informations disponibles sur le modèle de mélange. La plupart des recherches portant sur les méthodes de SAS supposent que le modèle de mélange est linéaire, où les signaux observés résultent de combinaisons linéaires des signaux sources. Cependant, pour certaines applications, le modèle de mélange linéaire n'est pas valide et doit être remplacé par un modèle non-linéaire qui permet de fournir une meilleure description du processus de mélange. La contribution méthodologique de cette thèse consiste à prendre en considération les interactions non-linéaires qui peuvent se produire entre les sources en utilisant le modèle linéaire-quadratique (LQ). À cet effet, nous avons développé trois nouvelles méthodes de SAS. La première méthode vise à résoudre le problème du démélange hyperspectral en utilisant un modèle linéaire-quadratique. Celle-ci se repose sur la méthode d'Analyse en Composantes Parcimonieuses (ACPa) et nécessite l'existence des pixels purs dans la scène observée. Dans le même but, nous proposons une deuxième méthode du démélange hyperspectral adaptée au modèle linéaire-quadratique. Elle correspond à une méthode de Factorisation en Matrices Non-négatives (FMN) se basant sur l'estimateur du Maximum A Posteriori (MAP) qui permet de prendre en compte les informations a priori sur les distributions des inconnues du problème afin de mieux les estimer. Enfin, nous proposons une troisième méthode de SAS basée sur l'analyse en composantes indépendantes (ACI) en exploitant les Statistiques de Second Ordre (SSO) pour traiter un cas particulier du mélange linéaire-quadratique qui correspond au mélange bilinéaire.

Mots clés : Séparation Aveugle de Sources, Analyse en Composantes Parcimonieuses, Factorisation en Matrices Non-négatives, Analyse en Composantes Indépendantes, Démélange Hyperspectral, Modèle linéaire-quadratique.

Abstract

In this thesis, we were interested to propose new Blind Source Separation (BSS) methods adapted to the nonlinear mixing models. BSS consists in estimating the unknown source signals from their observed mixtures when there is little information available on the mixing model. Most research on BSS methods assumes that the mixing model is linear, where the observed signals result from linear combinations of the source signals. However, for some applications, the linear mixing model is not valid and should be replaced by a non-linear model that provides a better description of the mixing process. The methodological contribution of this thesis consists in considering the non-linear interactions that can occur between sources by using the linear-quadratic (LQ) model. To this end, we developed three new BSS methods. The first method aims at solving the hyperspectral unmixing problem by using a linear-quadratic model. It is based on the Sparse Component Analysis (SCA) method and requires the existence of pure pixels in the observed scene. For the same purpose, we propose a second hyperspectral unmixing method adapted to the linear-quadratic model. It corresponds to a Non-negative Matrix Factorization (NMF) method based on the Maximum A Posteriori (MAP) estimate allowing to take into account the available prior information about the unknown parameters for a better estimation of them. Finally, we propose a third BSS method based on the Independent Component Analysis (ICA) method by using the Second Order Statistics (SOS) to process a particular case of the linear-quadratic mixture that corresponds to the bilinear one.

Keywords : Blind Source Separation, Sparse Component Analysis, Non-Negative Matrix Factorization, Independent Component Analysis, Hyperspectral unmixing, Linear-Quadratic Model.

Table des Matières

Remerciements ii

Résumé iv

Abstract v

Liste des Figures viii

Liste des Tableaux ix

Liste des abréviations x

Introduction Générale 1

1 Séparation Aveugle de Sources 4
1.1 Introduction . . . 5
1.2 Principe général de la séparation aveugle de sources . . . 5
1.3 Mélanges linéaires . . . 9
1.4 Méthodes de séparation linéaires instantanés . . . 11
1.5 Mélange linéaire-quadratique . . . 26
1.6 Méthodes linéaires-quadratiques . . . 29
1.7 Conclusions . . . 33

2 Démélange spectral 34
2.1 Introduction . . . 35
2.2 Principe de la télédétection hyperspectrale . . . 36
2.3 Modèle linéaire pour le démélange hyperspectral . . . 39
2.4 Méthodes de démélange linéaire . . . 41
2.5 Modèles non-linéaires pour le démélange hyperspectral . . . 46
2.6 Méthodes de démélange non-linéaires . . . 53
2.7 Conclusions . . . 59

3 Une méthode d'ACPa pour le démélange hyperspectral des mélanges linéaires-quadratiques 60
3.1 Introduction . . . 61
3.2 Modèle de mélange linéaire-quadratique . . . 62
3.3 Méthode non supervisée de démélange hyperspectral . . . 63
3.4 Étude expérimentale . . . 72
3.5 Conclusions . . . 84

4 Une méthode de NMF basée sur l'estimateur du maximum a posteriori (MAP) pour le démélange hyperspectral des mélanges linéaires-quadratiques **85**
4.1 Introduction 86
4.2 Formulation du problème 87
4.3 Méthode proposée 89
4.4 Résultats de simulations 101
4.5 Conclusions 109

5 Une méthode d'ACI basée sur les Statistiques de Second Ordre pour les mélanges bilinéaires **110**
5.1 Introduction 111
5.2 Modèle de mélange bilinéaire 112
5.3 Étude de la corrélation mutuelle des sources étendues 112
5.4 Reformulation du modèle bilinéaire 117
5.5 Méthode proposée 120
5.6 Résultats des simulations 125
5.7 Conclusions 135

Conclusion Générale et Perspectives **136**

Liste des publications **139**

Bibliographie **140**

Liste des Figures

1.1 Exemple de séparation de sources. 6
1.2 Configuration générale de SAS 7

2.1 La télédétection passive par des capteurs satellitaires. 36
2.2 Exemple d'une image hyperspectrale. 37
2.3 Illustration du principe des mélanges dans l'imagerie hyperspectrale. . . 38
2.4 Principe du modèle linéaire. 39
2.5 Projection en dimension 2 d'un simplexe de dimension 4 contenant les données hyperspectrales. 41
2.6 Illustration des réflexions multiples 47

3.1 L'organigramme de la méthode proposée 64
3.2 Spectres des matériaux purs 75
3.3 Image satellitaire d'une région de Toulouse 78
3.4 Cartes d'abondances déduites à partir de l'image de Toulouse 79
3.5 Cartes d'abondances estimées 80
3.6 Les 8 cartes d'abondance originales 81
3.7 Les 8 cartes d'abondance estimées 82
3.8 Les 8 spectres estimés 83

4.1 Spectres utilisés dans le Test 2 105
4.2 Image satellitaire de Tournefeuille. 107
4.3 Spectres utilisés dans le Test 3 108

5.1 Comparaison entre les vraies sources synthétiques et leurs estimées, dans le cas d'un mélange artificiel non-bruité, en utilisant 3 méthodes différentes : B-SO-BSS ($\tau_i = 1$) (haut), B-SO-BSS ($\tau_i = 1, 2, 3, 4$) (milieu) et NMF-Grd-LQ (en bas). Seule une partie des signaux correspondant à $n \in [200, 249]$ est présentée ici. 129
5.2 Activités des ions Na+ et K+. 130
5.3 Comparaison entre les sources chimiques et leurs estimées, dans le cas d'un mélange artificiel non-bruité, en utilisant 3 méthodes différentes : B-SO-BSS ($\tau_i = 1$) (haut), B-SO-BSS ($\tau_i = 1, 2, 3, 4$) (milieu) et NMF-Grd-LQ (en bas). 131
5.4 RSI_s en fonction du RSB dans le cas de l'Expérience 3. 133
5.5 RSI_s en fonction du RSB dans le cas de l'Expérience 4. 133

Liste des Tableaux

3.1 EQMN pour un mélange de 2 à 5 matériaux purs 75
3.2 EQMN pour un mélange de 2 matériaux purs avec un *RSB* variant de 30 à 80 dB . 77
3.3 EQMN pour des mélanges de 2 à 5 matériaux purs avec *RSB*=50 dB 77
3.4 EQMN pour l'**Expérience 3** avec un *RSB* variant de 30 à 80 dB 80

4.1 Résultats du Test 1 104
4.2 Résultats du Test 2 106
4.3 Résultats du Test 3 108

5.1 Résultats des simulations en utilisant (B-SO-BSS) avec un et quatre décalages, et en utilisant NMF-Grd-LQ dans le cas de deux signaux synthétiques. 128
5.2 Résultats des simulations en utilisant (B-SO-BSS) avec un et quatre décalages, et en utilisant NMF-Grd-LQ dans le cas de deux signaux chimiques. 130
5.3 Résultats des simulations obtenus pour l'Expérience 3 en utilisant B-SO-BSS avec un et quatre décalages lorsque les mélanges sont bruités. . 132
5.4 Résultats des simulations obtenus pour l'Expérience 4 en utilisant B-SO-BSS avec un et quatre décalages lorsque les mélanges sont bruités. . 134

Liste des abréviations

ACI	Analyse en Composantes Indépendantes
ACPa	Analyse en Composantes Parcimonieuses
ADMM	Alternating Direction Method of Multipliers
AMUSE	Algorithm for Multiple Unknown Signal Extraction
BL	BiLinéaire
B-SO-BSS	Bilinear Second Order Blind Source Separation
BP	Basis Pursuit
BPDN	Basis Pursuit Denoising
COM2	COntrast Maximization 2
DEMIX	Direction Estimation of Mixing matrIX
DUET	Degenerate Unmixing and Estimation Technique
EGM	Eigenvalue Grads Methods
ENVI	ENvironment for Visualizing Images
EQMN	Erreur Quadratique Moyenne Normalisée
FCLS	Fully Constrained Least Squares
FMN	Factorisation en Matrices Non-négatives
	ou NMF pour Non-negative Matrix Factorization
FMP	Factorisation en Matrices Positives
FOBIUM	Fourth Order Blind Identification of Underdetermined Mixtures of sources
GBM	Generalized Bilinear Model
GPLVM	Gaussian Process Latent Variable Model
i.i.d.	indépendants et identiquement distribués
Isomap	Isometric feature mapping
JADE	Joint Approximate Diagonalization of Eigenmatrices
KFCLS	Kernel Fully Constrained Least Squares
LI-TEMPCORR	Linear Instantaneous-TEMPoral CORRelation
LI-TEMPROM	Linear Instantaneous-TEMPoral Ratio Of Mixtures

LI-TIFCOHERE	Linear Instantaneous-TIme-Frequency COHERENce
LI-TIFCORR	Linear Instantaneous-TIme-Frequency CORRelation
LI-TIFROM	Linear-Instantaneous-TIme-Frequency Ratio Of Mixtures
LLE	Locally Linear Embedding
LQ	Linéaire-Quadratique
MAP	Maximum A Posteriori
MCMC	Markov Chain Monte Carlo
MCNN	Moindres Carrés avec contrainte de Non-Négativité
MM	Macroscopiques et Microscopiques
MNF	Maximum Noise Fraction
MP	Moyenne a Posteriori
MV	Minimisation du Volume
MVC-NMF	Minimum Volume Constrained Non-negative Matrix Factorization
MVSA	Minimum Volume Simplex Analysis
NCM	Normal Compositional Model
OMP	Orthogonal Matching Pursuit
PP	Pixels Purs
PPI	Pixel Purity Index
PPN	Pôlynomial Post-Non-linéaire
SAS	Séparation Aveugle de Sources
	ou BSS pour Blind Source Separation
SEONS	SEcond Order Nonstationary Source Separation
SISAL	Simplex Identification via variable Splitting and Augmented Lagrangian
SIR	Signal-to-Interference Ratio
SMM	Stochastic Mixing Model
SOBI	Second Order Blind Identification
SPICE	Sparsity Promoting Iterated Constrained Endmember detection
SSO	Statistiques de Second Ordre
	ou SOS pour Second Order Statistics
STOTD	Simultaneous Third Order Tensor Diagonalization
SUnSAL	Sparse Unmixing by variable Splitting and Augmented Lagrangian
SUnSAL-TV	SUnSAL Total Variation

RSB	Rapport Signal sur Bruit
RSI	Rapport Signal sur Interférence
	ou SIR pour Signal-to-Interference Ratio
TDSEP	Temporal Decorrelation SEParation
VCA	Vertex Component Analysis
WDO	W-Disjoint Orthogonality

Introduction Générale

Cette thèse rentre dans le cadre d'une convention de cotutelle conclue entre l'Université de Sfax, l'Ecole Nationale d'Ingénieurs de Sfax (Tunisie) et l'Université Toulouse III Paul Sabatier (France). Les travaux ont été effectués alternativement au sein du laboratoire de recherche Advanced Technologies for Medicine and Signals (ATMS) et l'Institut de Recherche en Astrophysique et Planétologie (IRAP).

Problématique

La présente thèse de doctorat s'inscrit dans le cadre de la séparation aveugle de sources (SAS) appliquée aux mélanges non-linéaires. La SAS consiste à estimer les signaux sources inconnus à partir de leurs mélanges observés quand il y a très peu d'informations disponibles sur le modèle de mélange. De nos jours, les méthodes de séparation de sources sont utilisées dans plusieurs applications de traitement de signal et de traitement d'images. La plupart des recherches portant sur les méthodes de SAS supposent que le modèle de mélange est linéaire, où les signaux observés résultent de combinaisons linéaires des signaux sources. Cela est dû à la simplicité et à la facilité de l'interprétation du modèle linéaire et à la variété de ses applications. Néanmoins, pour certaines applications, le modèle de mélange linéaire n'est pas valide et doit être remplacé par un modèle non-linéaire. Ce modèle non linéaire fournit une meilleure description du processus de mélange et des interactions entre les sources. À titre d'exemple, le mélange défini en télédétection est linéaire dans le cas d'un champ plat, mais un effet non-linéaire, dû aux réflexions multiples de la lumière, est introduit dès qu'il y a des éléments en relief. Ces phénomènes physiques intervenant dans le milieu peuvent être décrits par un modèle de mélange non-linéaire. Cependant, les méthodes traitant les mélanges non-linéaires restent relativement peu étudiées à ce jour à cause de la difficulté et de la complexité du développement théorique du problème. Cette complexité peut être réduite en contraignant la structure des modèles de mélange

non-linéaires par l'ajout d'hypothèses simplificatrices. Parmi ces modèles, le modèle linéaire-quadratique (LQ) a attiré beaucoup d'attention grâce à sa capacité à décrire le processus complexe de mélange dans certaines applications.

Le travail de cette thèse vise à explorer de nouvelles voies pour la SAS appliquée aux modèles non-linéaires et en particulier aux modèles linéaires-quadratiques. La contribution méthodologique de ce travail consiste à prendre en considération les interactions multiples qui peuvent se produire entre les sources afin d'améliorer la qualité de l'analyse et du traitement des méthodes de séparation des sources.

Organisation du manuscrit

Le manuscrit est organisé en deux parties. La première partie présente un tour d'horizon du problème de SAS et du démélange spectral. La deuxième partie est consacrée au développement théorique et à l'étude expérimentale de trois nouvelles méthodes de SAS adaptées aux mélanges linéaires-quadratiques/bilinéaires.

Le plan du manuscrit est structuré comme suit :

Première partie : état de l'art

- **Chapitre 1 :** méthodes de séparation aveugle de sources.
  Ce chapitre est consacré à la présentation d'un état de l'art sur les méthodes de séparation aveugle de sources les plus répandues. Nous décrivons l'ensemble des classes des méthodes utilisées dans la littérature pour traiter les mélanges linéaires et par la suite nous évoquons les extensions proposées de certaines d'entre elles pour traiter les mélanges linéaires-quadratiques.

- **Chapitre 2 :** démélange spectral.
  Dans ce chapitre, nous présentons quelques généralités sur l'application de télédétection spatiale ainsi qu'un état de l'art sur les méthodes les plus utilisées pour démélanger les images issues de cette application, autrement dit, les méthodes de démélange spectral.

Deuxième partie : développement des nouvelles méthodes de SAS

- **Chapitre 3 :** une méthode d'ACPa pour le démélange hyperspectral des mélanges LQ.

  Dans ce chapitre, nous proposons une approche de démélange hyperspectral en utilisant un modèle linéaire-quadratique. En se basant sur la méthode d'Analyse en Composantes Parcimonieuses (ACPa), l'approche proposée nécessite l'existence d'une zone mono-source ou d'au moins deux pixels purs par matériau pur, dans la scène observée.

- **Chapitre 4 :** une méthode de NMF basée sur l'estimateur du maximum a posteriori (MAP) pour le démélange hyperspectral des mélanges LQ.

  Dans ce chapitre, nous proposons une autre méthode de démélange hyperspectral adaptée au modèle LQ. Celle-ci correspond à une méthode de Factorisation en Matrices Non-négatives (FMN), ou encore NMF (pour son appellation en anglais *Non-negative Matrix Factorization*), se basant sur l'estimateur du Maximum A Posteriori (MAP). L'idée principale consiste à prendre en compte les informations a priori sur les distributions des coefficients de mélange afin de mieux les estimer.

- **Chapitre 5 :** une méthode d'ACI basée sur les statistiques du second ordre pour les mélanges bilinéaires.

  Dans ce chapitre, nous proposons une nouvelle méthode de séparation de sources basée sur l'analyse en composantes indépendantes (ACI) en exploitant les Statistiques de Second Ordre (SSO) pour traiter un cas particulier du mélange linéaire-quadratique qui correspond au mélange bilinéaire.

Chapitre 1

Séparation Aveugle de Sources

Sommaire

1.1 Introduction . . . 5
1.2 Principe général de la séparation aveugle de sources . . . 5
1.3 Mélanges linéaires . . . 9
1.3.1 Mélanges linéaires instantanés . . . 9
1.3.2 Mélanges convolutifs . . . 10
1.3.3 Mélanges à atténuation et à décalage . . . 10
1.4 Méthodes de séparation linéaires instantanés . . . 11
1.4.1 Analyse en Composantes Indépendantes . . . 11
1.4.1.1 Maximisation de la non-gaussianité . . . 12
1.4.1.2 Minimisation de l'information mutuelle . . . 13
1.4.1.3 Maximisation de la vraisemblance . . . 13
1.4.1.4 Méthodes tensorielles basées sur les statistiques d'ordre supérieur . . . 14
1.4.1.5 Méthodes basées sur les Statistiques de Second Ordre (SSO) . . . 15
1.4.2 Analyse en Composantes Parcimonieuses . . . 18
1.4.2.1 Méthodes utilisant de fortes hypothèses de parcimonie . . . 19
1.4.2.2 Méthodes "quasi-non-parcimonieuses " . . . 19
1.4.2.3 Méthodes hybrides . . . 20
1.4.3 Factorisation en Matrices Non-négatives . . . 21
1.4.4 Approches bayésiennes . . . 25
1.5 Mélange linéaire-quadratique . . . 26
1.6 Méthodes linéaires-quadratiques . . . 29
1.6.1 Extension des méthodes d'ACI . . . 29
1.6.1.1 Méthodes d'ACI pour séparer les sources i.i.d . . . 29
1.6.1.2 Méthodes d'ACI pour séparer les sources non-i.i.d 31
1.6.2 Extension des méthodes NMF . . . 31
1.6.3 Extension des méthodes ACPa . . . 32
1.6.4 Approches bayésiennes . . . 33
1.7 Conclusions . . . 33

1.1 Introduction

Le premier chapitre de la thèse présente un état de l'art du problème de séparation aveugle de sources. Tout d'abord, nous commençons par introduire le principe général de la séparation aveugle de sources. Ensuite, nous donnons un tour d'horizon sur les méthodes de séparation de sources proposées pour les deux types de mélange, linéaire et non-linéaire, en se concentrant sur les méthodes dédiées aux mélanges linéaires-quadratiques. Pour les deux types de mélange, nous commençons par présenter la formulation mathématique du modèle de mélange, et par la suite, nous dressons un panorama des travaux effectués dans la littérature dans ce cadre.

1.2 Principe général de la séparation aveugle de sources

Le problème de la séparation aveugle de sources (SAS ou BSS selon son appellation en anglais *Blind Source Separation*) a été abordé au début des années 1980 par Christian Jutten, Jeanny Hérault et Bernard Ans dans le cadre de la modélisation neuronale (Hérault et al, 1985). Plus tard, le problème de SAS a reçu une attention considérable de la part de la communauté scientifique dans plusieurs domaines tels que le traitement de signal, les statistiques, les réseaux de neurones, etc. Au sens large, le problème de SAS consiste à récupérer un ensemble de *signaux sources* inconnus à partir d'un ensemble de *signaux observés* qui résultent d'un *mélange* inconnu de ces signaux sources (Mansour et al, 2000; Comon and Jutten, 2010; Deville, 2016). Le terme *aveugle* se réfère au fait que ni les signaux sources, ni le mélange sont connus.

Les applications des méthodes de SAS sont diverses, parmi lesquelles nous citons l'astrophysique (Berné, O. et al, 2007; Nuzillard, D. and Bijaoui, A., 2000), le biomédical (Kardec Barros et al, 1998; Deville, 2003; Kachenoura et al, 2007), la télédétection (Bayliss et al, 1997; Wang and Chang, 2006; Pauca et al, 2006; Sajda et al, 2003; Nascimento and Bioucas-Dias, 2005a) et la détection des composants chimiques (Bedoya, 2006; Duarte et al, 2014; Ando et al, 2015). Un aperçu détaillé des différentes applications des méthodes de SAS est présenté dans (Mansour et al, 2000; Mansour and Kawamoto, 2003; Hyvärinen et al, 2001; Cichocki and Amari, 2003; Jutten and Comon, 2007).

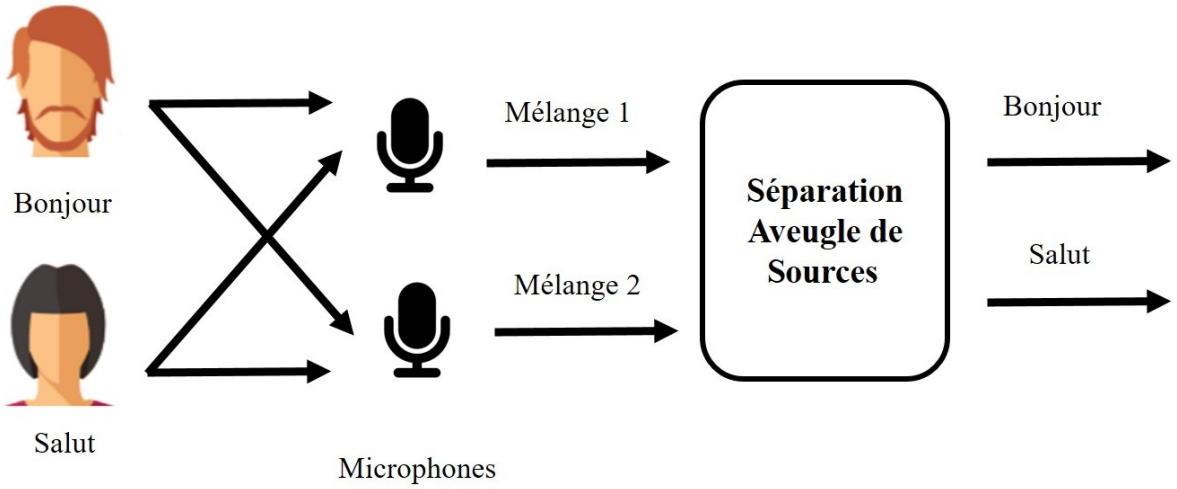


Figure 1.1: *Exemple de séparation de sources.*

Afin d'illustrer le principe de la SAS, prenons l'exemple classique du problème de la soirée cocktail (ou *Cocktail Party Problem* en anglais), où il y a un certain nombre de personnes qui sont en train de discuter dans une pièce. Dans ce cas, un être humain avec une audition en bonne santé peut faire la distinction entre les différentes voix parvenant mélangées à ses oreilles, et ainsi, il peut identifier une voix particulière et la comprendre. Il en est de même pour la SAS, dont le rôle est d'offrir une solution en séparant les signaux sources, correspondant aux signaux vocaux des invités, à partir des mélanges enregistrés par des microphones placés dans différents endroits. Le principe de cet exemple de séparation est illustré en Figure 1.1.

Un problème général de SAS peut être défini mathématiquement comme suit : nous avons un ensemble de K observations qui sont générées à partir d'un mélange de L signaux sources comme suit

$$\mathbf{x} = \mathbf{f}(\mathbf{s}), \tag{1.1}$$

où $\mathbf{x} = [x_1, \cdots, x_K]^T$ est le vecteur d'observations , $\mathbf{s} = [s_1, \cdots, s_L]^T$ est le vecteur des signaux sources, ^{T} désigne la transposée, et $\mathbf{f}$ désigne l'opérateur de mélange.

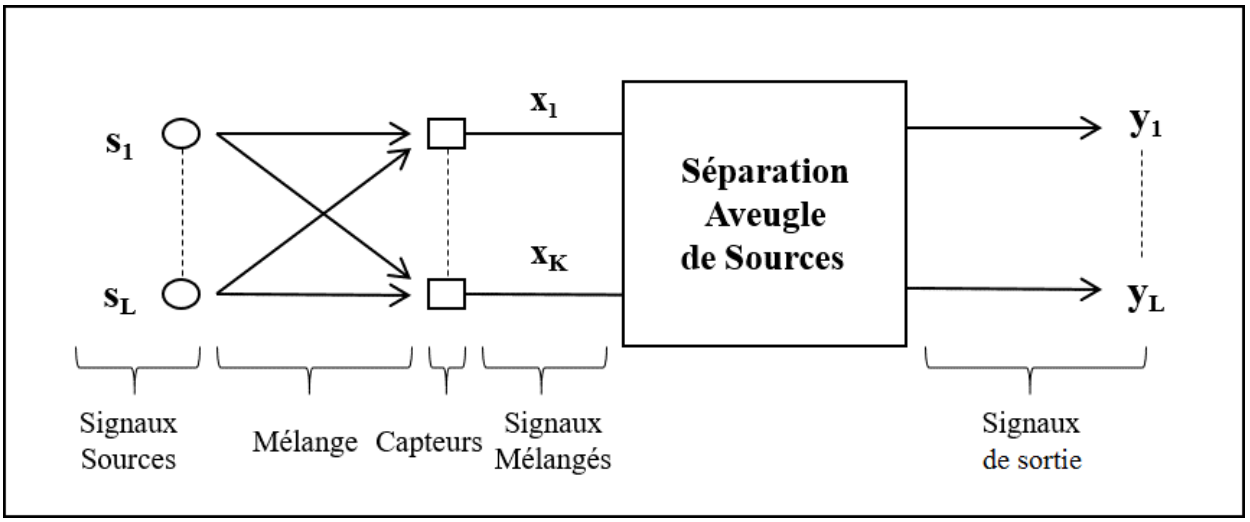


Figure 1.2: *Configuration générale de SAS*

Le but idéal de la SAS est de trouver la fonction inverse de $\mathbf{f}$, appelée le séparateur $\mathbf{g}$. Si cette fonction inverse existe, la séparation conduit à un vecteur de sortie $\mathbf{y} = [y_1, \cdots, y_L]^T$ qui contient les estimations des signaux sources $\mathbf{s}$ à certaines indéterminations près. Les sources estimées sont ainsi obtenues avec l'équation suivante :

$$\mathbf{y} = \mathbf{g}(\mathbf{x}). \tag{1.2}$$

La configuration générale du problème de SAS est illustrée dans la Figure 1.2.

Les méthodes de séparation de sources peuvent être classifiées principalement selon trois critères : le type de mélange, le nombre de signaux observés par rapport au nombre de signaux sources, et les propriétés des sources considérées.

Pour le premier critère, nous trouvons dans la littérature plusieurs types de mélange. Un premier type, concerne le mélange linéaire, c'est-à-dire le cas où les observations sont des mélanges linéaires des signaux sources. Les méthodes de SAS traitant ce type de mélange sont les plus étudiées en raison de la simplicité du modèle linéaire. Néanmoins, dans plusieurs applications, le modèle linéaire ne suffit pas pour décrire certains phénomènes physiques intervenant dans le mélange. Pour résoudre ce problème, il est nécessaire d'utiliser un modèle non-linéaire. La spécificité du modèle non-linéaire est sa capacité d'offrir une meilleure description de la réalité physique. L'extension des méthodes de SAS aux mélanges non-linéaires reste toujours moins étudiée en raison

de leur complexité. Pour simplifier leur utilisation, des modèles non-linéaires moins complexes ont été proposés dans la littérature. Nous citons, par exemple, le modèle post-non-linéaire où le mélange est formé d'une partie linéaire suivie d'une distorsion non-linéaire due aux capteurs. Un autre modèle est le modèle linéaire-quadratique qui contient à la fois des termes linéaires et des termes croisés résultant des interactions qui peuvent se produire entre les différentes sources. Nous nous intéressons dans cette thèse à ce dernier modèle qui sera traité en détail par la suite.

Une deuxième classification possible des méthodes de SAS est selon le rapport entre le nombre de signaux observés et le nombre de sources. Lorsque le nombre de signaux observés est supérieur au nombre de sources, nous parlons d'un mélange sur-determiné, et quand le nombre de signaux observés est inférieur au nombre de sources, le mélange devient sous-déterminé. Enfin, le mélange est considéré déterminé lorsque les nombres de sources et de signaux observés sont égaux.

Le troisième critère de classification concerne les propriétés vérifiées par les sources. En effet, la SAS n'est théoriquement possible que si certaines hypothèses sont vérifiées par les signaux sources. Dans la littérature, plusieurs propriétés ont été utilisées : signaux à valeurs réelles ou complexes, stationnaires, cyclo-stationnaires ou non-stationnaires, mutuellement indépendants ou non, indépendants et identiquement distribués (i.i.d.) ou non-i.i.d., positifs ou non, parcimonieux ou non, etc. La plupart des méthodes de séparation de sources ont été proposées pour les signaux sources i.i.d. et mutuellement indépendants.

Dans ce chapitre, nous considérons un classement selon le type de mélange étudié. Nous présentons, tout d'abord, les mélanges linéaires et les différentes méthodes utilisées en SAS pour les séparer, et par la suite, nous traitons en détail les mélanges linéaires-quadratiques et les méthodes élaborées dans ce cadre.

1.3 Mélanges linéaires

Dans la littérature, les mélanges linéaires ont largement été étudiés en séparation de sources. Ce type de mélange sera présenté dans cette partie d'une façon succincte. Tout d'abord, nous présentons les différents types de mélanges linéaires ainsi que leurs formulations mathématiques. Ensuite, nous exposons un panorama des méthodes de séparation de sources proposées pour ce type de mélange.
Les mélanges linéaires peuvent être répartis en trois sous-catégories présentées dans les sous-sections qui suivent.

1.3.1 Mélanges linéaires instantanés

Tout d'abord, nous commençons par décrire le modèle de mélange linéaire instantané. Dans ce cas, l'opérateur de mélange correspond à une somme pondérée et instantanée des signaux sources. Ainsi, l'équation de mélange pour chaque observation $\mathbf{x}_i$ à l'indice d'échantillon n (ou l'instant t dans le cas d'un signal temporel), en absence de bruit, peut s'exprimer comme suit

$$x_i(n) = \sum_{j=1}^{L} a_{ij} s_j(n), \quad \forall 1 \leq n \leq N, \ \forall 1 \leq i \leq K, \tag{1.3}$$

où $s_j(n)$ représente le $j^{ème}$ signal source à l'indice d'échantillon n, a_{ij} est un coefficient représentant la contribution de la source j sur le capteur i et N est le nombre d'échantillons.

Le modèle reliant les K observations $\mathbf{x}_i$ aux L signaux sources $\mathbf{s}_j$ peut s'écrire sous la forme matricielle suivante :

$$\mathbf{X} = \mathbf{AS}, \tag{1.4}$$

où $\mathbf{X} \in \mathbb{R}^{K \times N}$ est la matrice contenant les signaux observés, $\mathbf{A} \in \mathbb{R}^{K \times L}$ est la matrice contenant les coefficients de mélange a_{ij}, et $\mathbf{S} \in \mathbb{R}^{L \times N}$ est la matrice contenant les sources.

1.3.2 Mélanges convolutifs

Le modèle de mélange convolutif est une généralisation du modèle précédent. Dans ce cas-là, les mélanges observés résultent des versions décalées des sources originales. Par conséquent, en considérant tous les décalages possibles de la source $\mathbf{s}_j$ contribuant à l'échantillon d'indice n de l'observation $\mathbf{x}_i$, l'équation du mélange peut s'écrire sous la forme suivante

$$x_i(n) = \sum_{j=1}^{L} \sum_{\tau=-\infty}^{\infty} a_{ij}(\tau) s_j(n-\tau), \tag{1.5}$$

où τ représente tout décalage entier possible de la source $\mathbf{s}_j$ dans l'observation $\mathbf{x}_i$.

Une autre formulation du problème, pour le même type de mélange, peut être obtenue en considérant l'effet du filtrage de la source par la réponse impulsionnelle d'un filtre inconnu, ainsi nous obtenons l'équation suivante

$$x_i(n) = \sum_{j=1}^{L} a_{ij}(n) * s_j(n), \tag{1.6}$$

où $*$ représente l'opérateur de convolution et $a_{ij}(n)$ est la réponse impulsionnelle du filtre de mélange associé à la contribution de la source $\mathbf{s}_j$ dans l'observation $\mathbf{x}_i$.

1.3.3 Mélanges à atténuation et à décalage

Un cas particulier des mélanges convolutifs est le mélange à atténuation et à décalage appelé aussi mélange "anéchoïque". Chaque observation $\mathbf{x}_i$ résulte de la superposition de tous les signaux sources $\mathbf{s}_j$, $j = 1, \cdots, L$, éventuellement décalés et attenués. Les signaux observés sont ainsi définis comme suit

$$x_i(n) = \sum_{j=1}^{L} a_{ij} \mathbf{s}_j(n - \tau_{ij}), \tag{1.7}$$

où τ_{ij} est le décalage entre la source j et le capteur i.

Dans ce qui suit, nous présentons un aperçu des méthodes de SAS proposées dans la littérature pour la séparation des modèles de mélanges linéaires. Nous nous focalisons sur les approches les plus classiques associées aux mélanges linéaires instantanés.

1.4 Méthodes de séparation linéaires instantanés

Comme nous l'avons mentionné au début du chapitre, les méthodes de SAS visent à estimer un ensemble de signaux sources inconnus à partir des observations connues. Pour répondre à cet objectif, différentes méthodes ont été proposées. Ces méthodes se regroupent principalement en quatre classes dont nous donnerons un aperçu général ci-après : les méthodes basées sur l'Analyse en Composantes Indépendantes (ACI), les méthodes basées sur l'Analyse en Composantes Parcimonieuses (ACPa), les méthodes basées sur la Factorisation en Matrices Non-négatives (FMN), et enfin, les méthodes basées sur l'approche bayésienne.

1.4.1 Analyse en Composantes Indépendantes

Les méthodes basées sur l'Analyse en Composantes Indépendantes (ACI) sont les méthodes de séparation de sources les plus étudiées. Le point clé de ces méthodes consiste à appliquer des transformations aux observations de façon à obtenir en sortie des signaux statistiquement indépendants les uns des autres, et qui sont égaux aux signaux sources à certaines indéterminations près (permutations, facteur d'échelle).

Par définition, si $\mathbf{s} = [s_1, ..., s_L]^T$ est un vecteur aléatoire avec une densité de probabilité conjointe $p(\mathbf{s})$, alors $\mathbf{s}$ a des composantes indépendantes si sa densité de probabilité conjointe est égale au produit des densités marginales $p_i(s_i)$, comme suit

$$p(\mathbf{s}) = \prod_{i=1}^{L} p_i(s_i). \tag{1.8}$$

Il est à noter que pour réaliser la séparation des signaux sources en ACI, certaines méthodes supposent que les signaux sources sont i.i.d. avec au plus un seul signal

source ayant une distribution gaussienne. D'autres méthodes supposent que les signaux sources sont stationnaires et autocorrélés ou qu'ils sont non-stationnaires, et elles s'appliquent même lorsque les sources sont gaussiennes.

Plusieurs critères ont été utilisés pour la mesure de l'indépendance statistique entre les signaux sources. Dans ce qui suit, nous exposons brièvement les critères les plus connus.

1.4.1.1 Maximisation de la non-gaussianité

L'un des critères les plus utilisés en ACI est la maximisation de la non-gaussianité. En effet, une condition nécessaire pour séparer les signaux sources i.i.d. est la non-gaussianité de ceux-ci. Cette condition est basée sur le théorème de la limite centrale qui affirme que la somme des variables indépendantes tend vers une distribution gaussienne. Ainsi, on peut déduire que dans un problème de SAS, nous pouvons retrouver les composantes indépendantes en augmentant la non-gaussianité des sources estimées. Une mesure classique de la non-gaussianité est l'autocumulant d'ordre 4 normalisé, appelé encore le kurtosis normalisé, dont l'expression est définie comme suit (Mansour and Jutten, 1999)

$$Kurt(y) = \frac{E\{y^4\}}{E\{y^2\}^2} - 3, \tag{1.9}$$

où $E\{.\}$ est l'opérateur d'espérance mathématique et y est une variable aléatoire centrée.

En effet, le kurtosis normalisé mesure la différence entre la densité de probabilité d'une variable aléatoire et une densité gaussienne. Ainsi, la maximisation de la valeur absolue du kurtosis normalisé équivaut à une maximisation de la non-gaussianité de la variable estimée, et permet, par conséquent, d'extraire les sources indépendantes.

Dans cette catégorie de méthodes, plusieurs travaux ont été réalisés, parmi lesquels nous citons la plus connue qui est la méthode FastICA (Hyvärinen, 1999).

1.4.1.2 Minimisation de l'information mutuelle

Le critère d'information mutuelle dans l'ACI a été initialement introduit par P. Comon pour séparer les mélanges linéaires instantanés (Comon, 1994). Le critère d'information mutuelle permet de mesurer l'indépendance statistique des variables en évaluant la similitude entre la densité de probabilité conjointe et le produit des densités marginales des variables aléatoires. Cette similitude peut être mesurée par le calcul de la divergence de Kullback-Leibler comme suit (Comon and Jutten, 2010)

$$I(y_1, \cdots, y_K) = \int_{-\infty}^{+\infty} p(y_1, \cdots, y_K) \log \frac{p(y_1, \cdots, y_K)}{\prod_i p_i(y_i)} \, \mathrm{d}y_1, \cdots, \mathrm{d}y_K \tag{1.10}$$

où $y_1, \cdots, y_K$ sont des variables aléatoires, $p(y_1, \cdots, y_K)$ est la densité de probabilité conjointe de ces variables aléatoires, et $p_i(y_i)$ est la densité marginale de y_i.

Le critère d'information mutuelle est toujours positif et s'annule seulement lorsque les variables y_i sont indépendantes, autrement dit, lorsque ces variables n'ont aucune information commune.

Parmi les méthodes basées sur le critère de minimisation de l'information mutuelle, nous pouvons citer les travaux étudiés dans (Bell and Sejnowski, 1995; Amari et al, 1996; Pham, 1999; Erdogmus et al, 2004; Hild et al, 2001; Rhabi et al, 2004).

1.4.1.3 Maximisation de la vraisemblance

Le principe des méthodes d'ACI basées sur la maximisation de la vraisemblance consiste à trouver les paramètres de mélange qui maximisent la densité de probabilité des observations $p_{\mathbf{x}}(\mathbf{x})$. L'expression de la densité de probabilité des observations pour un modèle linéaire instantané peut être exprimée de la façon suivante

$$p_{\mathbf{x}}(x(n)) = |det(\mathbf{B})| \prod_{i=1}^{L} p_i(s_i(n)), \tag{1.11}$$

où $\mathbf{B}$ est l'inverse de la matrice de mélange $\mathbf{A}$ et $p_i(\mathbf{s}_i)$ est la densité de probabilité de la source $\mathbf{s}_i$.

En supposant que les sources sont i.i.d., et en disposant de N échantillons pour chaque signal observé, la vraisemblance est calculée de la manière suivante

$$V = \prod_{n=1}^{N} \left[|det(\mathbf{B})| \prod_{i=1}^{L} p_i(\mathbf{s}_i(n)) \right]. \tag{1.12}$$

Notons que la maximisation de la vraisemblance ou de la log-vraisemblance revient à la même chose. L'expression de la log-vraisemblance prend la forme suivante

$$L = \left[\sum_{n=1}^{N} \sum_{i=1}^{L} \log(p_i(\mathbf{s}_i(n))) \right] + N \log(|det(\mathbf{B})|). \tag{1.13}$$

L'inconvénient majeur de la méthode de maximisation de la vraisemblance en ACI est la non-connaissance des densités de probabilité des signaux sources. Pour pallier ce problème, il est nécessaire de supposer que ces densités de probabilité sont connues a priori ou qu'elles appartiennent à une famille de distributions.

Parmi les travaux réalisés dans ce cadre, nous mentionnons les méthodes étudiées dans (Cardoso, 1997; Gaeta and Lacoume, 1990; Guidara, 2009; Pham et al, 1992).

1.4.1.4 Méthodes tensorielles basées sur les statistiques d'ordre supérieur

Des méthodes basées sur les statistiques d'ordre supérieur, telles que les cumulants d'ordre supérieur à deux, ont été proposées pour séparer les signaux i.i.d. et non-gaussiens avec au plus un seul signal ayant une distribution gaussienne. L'idée est de séparer ces signaux en utilisant une fonction basée sur les cumulants d'ordre supérieur. En effet, les cumulants croisés entre les signaux sont nuls à tous les ordres lorsque ces signaux sont indépendants. D'une manière générale, il est difficile de déterminer les cumulants croisés à tous les ordres, c'est pourquoi les méthodes de SAS se limitent généralement à l'ordre quatre. L'utilisation des méthodes basées sur les cumulants s'effectue par une approche tensorielle qui emploie les tenseurs de cumulants d'ordre supérieur. Par définition, les tenseurs constituent une extension à N dimensions du concept de matrice. Par conséquent, le tenseur de cumulants étend le concept de matrice de covariance aux ordres supérieurs. La spécificité des tenseurs de cumulants est qu'ils sont diagonaux à tous les ordres lorsque les signaux sont indépendants. Le

tenseur de cumulants T d'ordre quatre correspond à un opérateur linéaire défini par les cumulants croisés tel que la transformation $T_{i,j}(M)$ de la matrice M est comme suit

$$T_{i,j}(M) = \sum m_{k,l} cum_4(y_i, y_j, y_k, y_l), \tag{1.14}$$

où $m_{k,l}$ est l'élément d'indice (k, l) de la matrice M, cum_4 désigne le cumulant d'ordre quatre, et y_i, y_j, y_k, y_l sont des variables aléatoires.

D'autres critères basés sur les cumulants d'ordre quatre pour des mélanges instantanés de deux sources, ont été étudiés dans (Mansour and Jutten, 1995).

Parmi les premières méthodes basées sur les cumulants d'ordre supérieur, nous citons celle proposée par Lacoume and Ruiz (1988) qui repose sur l'annulation des cumulants d'ordre deux et quatre. Par la suite, Comon (1994) a proposé une méthode appelée COM2 (pour *COntrast Maximization 2* en anglais) qui consiste à annuler les cumulants croisés d'ordre quatre. Une autre méthode appelée JADE (pour *Joint Approximate Diagonalization of Eigenmatrices* en anglais) a été proposée par Cardoso and Souloumiac (1993). Celle-ci se base sur la diagonalisation conjointe d'un ensemble de matrices de cumulants indépendantes. Une autre méthode, appelée STOTD (pour *Simultaneous Third Order Tensor Diagonalization* en anglais), a été proposée par Lathauwer (September 1997). En se basant sur le même principe que JADE, la méthode STOTD consiste à diagonaliser conjointement un ensemble de tenseurs d'ordre trois au carré. Moreau (2001) propose une extension de la méthode JADE, dénommée eJADE, utilisant des cumulants d'ordre supérieur ou égal à trois. Une méthode dénomée FOBIUM (pour *Fourth Order Blind Identification of Underdetermined Mixtures of sources* en anglais) a été proposée par Ferreol et al (2005) afin d'identifier les mélanges sous-déterminés.

1.4.1.5 Méthodes basées sur les Statistiques de Second Ordre (SSO)

Une autre classe de méthodes d'ACI est basée sur les Statistiques de Second Ordre (SSO) (ou SOS pour *Second Order Statistics* en anglais). L'objectif de ces méthodes est de rendre les signaux estimés indépendants en se limitant aux statistiques d'ordre deux. Il est à noter que les statistiques d'ordre deux ne sont pas suffisantes pour

effectuer la séparation des sources i.i.d. Il est donc nécessaire de supposer que les sources sont stationnaires et autocorrélées ou non-stationnaires.

Pour la première hypothèse, lorsque les sources sont stationnaires et autocorrélées, l'idée repose sur le fait que les matrices de corrélation des observations $R_{\mathbf{x}}(\tau)$ sont simultanément diagonalisables pour plusieurs décalages τ. Ces matrices de corrélation sont définies comme suit

$$R_{\mathbf{x}}(\tau) = E\{\mathbf{x}(n)\mathbf{x}(n-\tau)^T\} = \mathbf{A}R_{\mathbf{s}}(\tau)\mathbf{A}^T, \tag{1.15}$$

où $R_{\mathbf{s}}(\tau)$ sont les matrices de corrélation des sources dont l'expression est comme suit

$$R_{\mathbf{s}}(\tau) = E\{\mathbf{s}(n)\mathbf{s}(n-\tau)^T\}. \tag{1.16}$$

Notons que ces dernières matrices sont réelles et diagonales si les sources sont réelles, mutuellement non-corrélées, et de moyennes nulles.

Pour cette classe de méthodes, deux conditions doivent être vérifiées. La première consiste à supposer que les sources sont mutuellement non-corrélées et de moyennes nulles de sorte que

$$E\{s_i(n)s_j(n-\tau)\} = 0,\ i \neq j \tag{1.17}$$

et la deuxième consiste à considérer que les sources ont des statistiques distinctes aux décalages τ

$$E\{s_i(n)s_i(n-\tau)\} \neq E\{s_j(n)s_j(n-\tau)\},\ i \neq j. \tag{1.18}$$

La première condition permet aux matrices de corrélation des sources $R_{\mathbf{s}}(\tau)$ d'être diagonales, et la deuxième condition assure que ces matrices disposent des valeurs propres différentes.

Le point clé de ces méthodes consiste à utiliser une méthode de diagonalisation conjointe des matrices de corrélation pour estimer la matrice de mélange. En considérant un ensemble de matrices carrées $\mathbf{M} = \mathbf{M}_0, \mathbf{M}_1, \cdots, \mathbf{M}_k$, le principe de la diagonalisation conjointe est de déterminer une matrice $\mathbf{B}$ telle que les matrices $\mathbf{B}\mathbf{M}_j\mathbf{B}^T$, $j = 1, \cdots, k$, soient aussi proches que possible des matrices diagonales.

Une première méthode se basant sur les statistiques d'ordre deux est celle proposée par Tong et al (1990) sous l'appellation AMUSE (pour *Algorithm for Multiple Unknown*

Signal Extraction en anglais). Cette méthode vise à diagonaliser deux matrices de corrélation des observations. L'une est associée à un décalage nul tandis que l'autre est associée à un décalage non nul. Le principe général de la méthode AMUSE est décrit à travers les étapes suivantes. Tout d'abord, elle estime la matrice de corrélation des observations $R_{\mathbf{x}}(0)$. Ensuite, elle estime la matrice de blanchiment définie par

$$\mathbf{W} = \mathbf{D}^{-\frac{1}{2}}\mathbf{E}^T, \tag{1.19}$$

où $\mathbf{D}$ est la matrice diagonale contenant les valeurs propres de $R_{\mathbf{x}}(0)$ et $\mathbf{E}$ est la matrice contenant les vecteurs propres de $R_{\mathbf{x}}(0)$, et elle réalise le blanchiment des signaux observés en utilisant : $\mathbf{z}(n) = \mathbf{W}\mathbf{x}(n)$. Cette étape consiste à faire en sorte que les signaux observés soient décorrélés en vérifiant $R_z(0) = I$. Par la suite, elle détermine une matrice orthogonale $\mathbf{U}$ formée par les vecteurs propres de la matrice de corrélation des observations blanchies $R_{\mathbf{z}}(\tau)$ pour un décalage τ non-nul. Enfin, elle estime la matrice de mélange $\hat{\mathbf{A}} = \mathbf{W}^{\dagger}\mathbf{U}$, où $\dagger$ désigne la pseudo-inverse, et les signaux sources $\hat{\mathbf{s}}(n) = \mathbf{U}^T\mathbf{z}(n)$.

La méthode AMUSE se distingue notamment par sa simplicité, cependant, ses performances dépendent beaucoup du choix du décalage τ. Pour résoudre ce problème, une extension de AMUSE a été proposée par Belouchrani et al (1997) sous la dénomination SOBI (pour *Second Order Blind Identification* en anglais). L'idée de cette dernière consiste à diagonaliser simultanément plus de deux matrices de corrélation des observations à plus de deux décalages τ_i. La méthode SOBI commence par l'étape de blanchiment dans la méthode AMUSE, ensuite, elle estime la matrice $\mathbf{U}$ par la diagonalisation conjointe d'un ensemble de matrices de corrélation blanchies $R_z(\tau)$ à différents décalages τ_i non-nuls. Enfin, l'estimation des signaux sources et de la matrice de mélange s'effectue de la même manière que dans la méthode AMUSE.

D'autres méthodes utilisant la diagonalisation des matrices de corrélation ont été proposées. Celles-ci comprennent TDSEP (pour *Temporal Decorrelation SEParation* en anglais) (Ziehe and Müller, 1998) et RSOBI (pour *Robust SOBI* en anglais) (Belouchrani and Cichocki, 2000) qui est une version améliorée de l'algorithme SOBI.

En ce qui concerne la deuxième hypothèse, lorsqu'il s'agit de sources non-stationnaires, la décorrélation au décalage nul des signaux sorties entre elles est considérée suffisante pour réaliser la séparation. Dans ce cas, nous mentionnons la méthode proposée par Matsuoka et al (1995) où la séparation est réalisée en minimisant une fonction de contraste afin d'obtenir à tout instant t des sorties décorrélées. Une autre idée est de supposer que les sources non-stationnaires sont stationnaires par morceaux. Par conséquent, une diagonalisation conjointe des matrices de corrélation réalisées à partir de chaque morceau de données est capable de résoudre le problème. Dans ce cas, nous citons la méthode proposée dans (Souloumiac, 1995) où les signaux observés sont divisés en deux morceaux supposés stationnaires, et sa forme généralisée proposée par (Pham and Cardoso, 2001) qui réalise la séparation en considérant un nombre quelconque de morceaux. Aussi, nous mentionnons la méthode SEONS (pour *SEcond Order Nonstationary Source Separation* en anglais) proposée par Choi et al (2002) qui est considérée comme une extension de la méthode SOBI pour le traitement des signaux non-stationnaires.

1.4.2 Analyse en Composantes Parcimonieuses

Une deuxième classe de méthodes de séparation de sources est l'Analyse en Composantes Parcimonieuses "ACPa" (en anglais *Sparse Component Analysis* "SCA"). La parcimonie a été très largement utilisée dans le traitement des signaux audio, cependant, elle n'a été exploitée pour la SAS que durant la dernière décennie (Jourjine et al, 2000; Bofill and Zibulevsky, 2001; Donoho and Elad, 2002).

Par définition, un signal est dit *parcimonieux* dans un domaine de représentation ou d'analyse si la plupart de ses valeurs sont nulles dans ce domaine. Parmi ces domaines de représentation, nous mentionnons les domaines Spatial, Temporel, Fréquentiel, Temps-Fréquence (TF), etc. Il convient de noter qu'un signal peut être parcimonieux dans un certain domaine d'analyse sans l'être dans les autres.

En séparation de sources, la solution obtenue par les méthodes d'ACPa pour identifier la matrice de mélange $\mathbf{A}$ et l'ensemble des sources $\mathbf{S}$ est unique à certaines indéterminations près (Donoho and Elad, 2002). Cela est vrai même dans le cas où

les mélanges sont sous-déterminés, lorsque le nombre de sources est supérieur à celui des observations. Pour cette raison, l'ACPa est considérée comme une solution efficace pour faire face à des problèmes sous-déterminés (Bofill and Zibulevsky, 2001; Li et al, 2006).

Contrairement aux méthodes d'ACI, où les sources sont estimées directement à partir des mélanges en utilisant l'hypothèse d'indépendance, la plupart des méthodes basées sur la parcimonie effectuent la séparation des sources en deux étapes. Elles commencent, tout d'abord, par l'estimation de la matrice de mélange, et par la suite, elles estiment les signaux sources à partir de ce mélange.

Par ailleurs, les hypothèses de parcimonie sur lesquelles se basent les méthodes ACPa peuvent varier d'une méthode à une autre. Dans ce qui suit, nous présentons trois catégories de méthodes d'ACPa classées selon les hypothèses de parcimonie imposées.

1.4.2.1 Méthodes utilisant de fortes hypothèses de parcimonie

Ces méthodes supposent l'orthogonalité W-disjointe des sources (en anglais *W-Disjoint Orthogonality* "WDO"), autrement dit, elles considèrent qu'une seule source est active dans chaque point du domaine de représentation considéré. Parmi ces méthodes, nous citons la méthode DUET (pour son appellation en anglais *Degenerate Unmixing and Estimation Technique*) (Jourjine et al, 2000) qui permet la séparation des mélanges linéaires à atténuations et à décalages. Cette dernière a été généralisée par la suite pour la séparation des mélanges convolutifs dans (Melia and Rickard, 2005).

1.4.2.2 Méthodes "quasi-non-parcimonieuses"

Ces méthodes succitent le plus d'intérêt puisqu'elles exigent des hypothèses de parcimonie moins fortes que celles utilisées dans les méthodes basées sur l'orthogonalité W-disjointe des sources. En effet, la présence d'une seule source dans quelques petites régions ou zones du domaine d'analyse est considérée suffisante. Il est à noter que ces zones *mono-sources*, où une seule source est active, sont décelées en utilisant un certain critère de détection. Parmi ces méthodes, nous citons LI-TEMPROM (pour *Linear*

Instantaneous-TEMPoral Ratio Of Mixtures en anglais) et LI-TIFROM (pour *Linear-Instantaneous-TIme-Frequency Ratio Of Mixtures* en anglais) (Abrard and Deville, 2005) qui utilisent l'hypothèse de parcimonie dans les deux domaines temporel, et temps-fréquence, respectivement. L'idée de ces méthodes est d'utiliser la variance des rapports d'observations comme un critère pour détecter les zones mono-sources dans le domaine considéré. Si ce critère a une faible valeur dans une zone d'analyse, alors cette dernière est considérée comme une zone mono-source. Par la suite, à partir de chaque zone mono-source détectée, l'estimation des colonnes de la matrice de mélange est réalisée par un moyennage des rapports d'observations.

Dans le même contexte, d'autres méthodes nommées LI-TEMPCORR (pour *Linear Instantaneous-TEMPoral CORRelation* en anglais) et LI-TIFCORR (pour *Linear Instantaneous-TIme-Frequency CORRelation* en anglais) ont été proposées (Deville and Puigt, 2007). Dans ce cas, le critère utilisé pour détecter les zones mono-sources n'est plus la variance mais plutôt le coefficient d'intercorrélation entre les observations. Si dans une zone d'analyse, le module de ce critère a une valeur proche de 1, alors cette zone est considérée comme une zone mono-source. Dans ce cas, les colonnes de la matrice de mélange sont estimées par le calcul du rapport entre l'intercorrélation des deux observations et l'autocorrélation de l'une de ces observations.

Une autre méthode utilisant un autre critère de détection a été développée dans le cas du domaine Temps-Fréquence. Cette méthode, nommée LI-TIFCOHERE (pour *Linear Instantaneous-TIme-Frequency COHEREnce* en anglais) (Albouy, 2004), utilise la cohérence des observations comme un critère pour détecter les zones mono-sources. Comme dans LI-TEMPCORR et LI-TIFCORR, une zone d'analyse est considérée comme une zone mono-source lorsque le critère de détection a une valeur proche de 1 dans cette zone.

1.4.2.3 Méthodes hybrides

L'appellation hybride vient du fait que ces méthodes se basent à la fois sur les méthodes utilisant l'orthogonalité W-disjointe des sources et les méthodes quasi-parcimonieuses. Parmi ces méthodes, nous citons la méthode proposée par Arberet et al (2006) qui repose sur les deux méthodes DUET et LI-TIFROM. Cette approche est

basée sur un algorithme de regroupement appelé DEMIX (pour *Direction Estimation of Mixing matrIX* en anglais) afin d'estimer le nombre de sources et leurs directions. Pour distinguer les zones mono-sources de celles contenant plusieurs sources, un critère mesurant la fiabilité tel que celui utilisé dans LI-TIFROM est utilisé.

1.4.3 Factorisation en Matrices Non-négatives

La troisième classe de méthodes de séparation de sources est la FMN (pour Factorisation en Matrices Non-négatives), ou encore NMF (pour son appellation en anglais *Non-negative Matrix Factorization*). Cette dernière a été initialement étudiée au milieu des années 90 par Paatero and Tapper (1994) sous l'appellation FMP (Factorisation en Matrices Positives), et depuis lors, plusieurs travaux ont été investigués dans le domaine du traitement du signal et des images. L'avantage des méthodes NMF est qu'elles ne supposent ni l'indépendance statistique, ni la représentation parcimonieuse des sources, cependant, elles imposent la non-négativité des observations, des signaux sources, et des coefficients de mélange.

Les méthodes NMF ont été appliquées avec succès dans de nombreuses applications, où les données sont non-négatives, dont la classification des documents (Xu et al, 2003; Shahnaz et al, 2006), l'analyse spectrale des données (Pauca et al, 2006; Jia and Qian, 2009), le traitement d'images (Li et al, 2001; Gillis and Glineur, 2012). L'utilisation de la NMF a connu un essor, en particulier à la suite de l'article de Lee and Seung (1999). Cet article était fondé sur des principes physiologiques et cognitifs. En effet, les auteurs ont considéré que le cerveau humain est un objet complexe formé par la somme de ses composants. Ainsi, ils ont utilisé ce principe pour décomposer une image d'un objet en plusieurs images représentant les différentes parties qui composent cet objet.

En séparation de sources, le principe de base de la méthode NMF est le suivant : étant donné une matrice de données non-négatives $\mathbf{X}$, la NMF consiste à trouver une factorisation approchée de $\mathbf{X}$ en deux matrices non-négatives $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ comme suit

$$\mathbf{X} \approx \hat{\mathbf{A}}\hat{\mathbf{S}} \tag{1.20}$$

où $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ représentent les estimées des matrices $\mathbf{A}$ et $\mathbf{S}$, respectivement.

Afin de réaliser cette factorisation, l'utilisation d'une fonction de coût est nécessaire. Cette fonction de coût est en général un critère (distance, similarité, divergence, etc.) mesurant la différence entre la matrice des observations $\mathbf{X}$ et le produit entre les deux matrices $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$. Les mesures les plus couramment utilisées sont la norme de Frobenius, telle que

$$D(\mathbf{X} \parallel \hat{\mathbf{A}}\hat{\mathbf{S}}) = \frac{1}{2}\|\mathbf{X} - \hat{\mathbf{A}}\hat{\mathbf{S}}\|_F^2 = \frac{1}{2}\sum_{ij}(\mathbf{X}_{ij} - (\hat{\mathbf{A}}\hat{\mathbf{S}})_{ij})^2, \quad \hat{\mathbf{A}} \geqslant 0, \hat{\mathbf{S}} \geqslant 0, \tag{1.21}$$

ou encore, la divergence de Kullback-Leibler, définie par

$$KL(\mathbf{X} \parallel \hat{\mathbf{A}}\hat{\mathbf{S}}) = \sum_{ij} \mathbf{X}_{ij} \log(\frac{\mathbf{X}_{ij}}{(\hat{\mathbf{A}}\hat{\mathbf{S}})_{ij}}) - \mathbf{X}_{ij} + (\hat{\mathbf{A}}\hat{\mathbf{S}})_{ij}, \quad \hat{\mathbf{A}} \geqslant 0, \hat{\mathbf{S}} \geqslant 0. \tag{1.22}$$

Plusieurs autres mesures de divergence peuvent être utilisées et elles sont décrites dans (Cichocki et al, 2009).

La méthode NMF peut être formulée en tant qu'un problème d'optimisation qui minimise une fonction de coût, sous les contraintes de non-négativité, comme suit

$$\min_{\hat{\mathbf{A}},\hat{\mathbf{S}}} \; F(\mathbf{X}, \hat{\mathbf{A}}, \hat{\mathbf{S}}), \quad \hat{\mathbf{A}} \geqslant 0, \hat{\mathbf{S}} \geqslant 0, \tag{1.23}$$

où F représente le critère ou la fonction de coût à optimiser.

Par la suite, afin de déterminer les deux matrices $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ qui minimisent la fonction de coût, l'utilisation d'un algorithme itératif est nécessaire. Ce dernier doit permettre d'adapter chacune des matrices $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$, jusqu'à la convergence, tout en respectant les contraintes de non-négativité de celles-ci.

Il existe différents algorithmes itératifs permettant de parvenir à cette tâche, parmi lesquels nous citons :

- **Algorithme de gradient projeté :** une première solution pour minimiser la fonction de coût est d'utiliser un algorithme de gradient projeté. L'idée de cette méthode consiste à utiliser un algorithme de gradient avec la projection des mises à jour dans l'intervalle des solutions acceptables. Dans chaque itération de cet

algorithme, les règles de mise à jour suivantes sont utilisées

$$\hat{\mathbf{A}} \leftarrow [\hat{\mathbf{A}} - \mu_{\mathbf{A}} \frac{\partial D}{\partial \hat{\mathbf{A}}}]_+, \tag{1.24}$$

$$\hat{\mathbf{S}} \leftarrow [\hat{\mathbf{S}} - \mu_{\mathbf{S}} \frac{\partial D}{\partial \hat{\mathbf{S}}}]_+, \tag{1.25}$$

où $\mu_{\mathbf{A}}$ et $\mu_{\mathbf{S}}$ sont des pas d'apprentissage positifs. Les valeurs de ces derniers sont soit fixées par l'utilisateur, soit déterminées par un algorithme capable de les adapter à chaque itération. $[.]_+$ correspond à la projection dans l'intervalle $\mathbb{R}^{*+}$ afin de vérifier la non-négativité de $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ à chaque itération. Dans le cas général, les valeurs négatives sont remplacées par des valeurs négligeables.

Dans la littérature, nous pouvons trouver des méthodes basées sur le principe de gradient projeté dans (Lin, 2007; Zdunek and Cichocki, 2008).

- **Algorithme multiplicatif :** Lee et Seung ont proposé un algorithme multiplicatif pour minimiser la fonction de coût. Les règles de mise à jour des deux matrices $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ lorsque la fonction de coût est égale à la norme de Frobenius, présentée dans (1.21), sont comme suit (Lee and Seung, 2001)

$$\hat{\mathbf{A}} \leftarrow \hat{\mathbf{A}} \odot ((\mathbf{X}\hat{\mathbf{S}}^T) \oslash (\hat{\mathbf{A}}\hat{\mathbf{S}}\hat{\mathbf{S}}^T)), \tag{1.26}$$

$$\hat{\mathbf{S}} \leftarrow \hat{\mathbf{S}} \odot ((\hat{\mathbf{A}}^T\mathbf{X}) \oslash (\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}})), \tag{1.27}$$

où $\odot$ et $\oslash$ représentent la multiplication et la division élément par élément, respectivement.
Dans ce cas, la non-négativité de $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$ est garantie lorsque leurs valeurs initiales sont non-négatives.

Lorsque la fonction de coût est égale à la divergence de Kullback-Leibler, présentée dans (1.22), les règles multiplicatives sont définies comme suit (Lee and Seung, 2001)

$$\hat{\mathbf{A}} \leftarrow \hat{\mathbf{A}} \odot ((\mathbf{X} \oslash \hat{\mathbf{A}}\hat{\mathbf{S}})\hat{\mathbf{S}}^T) \oslash (\mathbf{E}\hat{\mathbf{S}}^T)), \tag{1.28}$$

$$\hat{\mathbf{S}} \leftarrow \hat{\mathbf{S}} \odot (\hat{\mathbf{A}}^T(\mathbf{X} \oslash \hat{\mathbf{A}}\hat{\mathbf{S}})) \oslash (\hat{\mathbf{A}}^T\mathbf{E})), \tag{1.29}$$

où $\mathbf{E}$ représente une matrice dont les éléments sont égaux à 1.

L'avantage de l'algorithme multiplicatif, en le comparant avec l'algorithme de gradient projeté, est qu'il n'y a pas de pas à fixer ou à calculer à chaque itération. En outre, la non-négativité est assurée par les règles de mise à jour sans avoir recours à la projection à chaque itération.

Les méthodes NMF basée sur l'algorithme multiplicatif et ses variantes sont citées dans (Lee and Seung, 2001; Virtanen, 2007; Févotte and Idier, 2011; Bertin et al, 2009).

- **Algorithme des moindres carrés alternés :** le premier algorithme des moindres carrés alternés a été proposé par Paatero and Tapper (1994). Le principe de cet algorithme consiste à calculer, tour à tour, une des deux matrices $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$, en minimisant un critère des moindres carrés avec contrainte de non-négativité. Ces deux étapes sont répétées jusqu'à la convergence de l'algorithme.

  Parmi les méthodes réalisées dans ce cadre, nous citons les travaux réalisés dans (Cichocki and Zdunek, 2007; Cichocki et al, 2007) .

La principale difficulté dans l'utilisation des méthodes NMF est due à la non-unicité de la factorisation de la matrice $\mathbf{X}$ en tenant compte seulement de la non-négativité de $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$. En effet, nous pouvons facilement démontrer l'existence d'un nombre infini de factorisations de la matrice $\mathbf{X}$. La démonstration peut être réalisée en considérant une matrice inversible $\mathbf{D}$, qui peut être non-négative, telle que $\mathbf{X}=\hat{\mathbf{A}}\mathbf{D}\mathbf{D}^{-1}\hat{\mathbf{S}}$ et que $\hat{\mathbf{A}}\mathbf{D} = \tilde{\mathbf{A}}$ et $\mathbf{D}^{-1}\hat{\mathbf{S}} = \tilde{\mathbf{S}}$ soient deux matrices non-négatives (Cichocki et al, 2009). Ceci montre qu'il existe plus qu'une factorisation matricielle permettant d'avoir une reconstruction exacte des données observées. Nous pouvons trouver plus de détails sur l'unicité de la factorisation matricielle dans les travaux réalisés par Moussaoui et al (2005) ou encore par Donoho and Stodden (2003).

Un autre problème rencontré en NMF est le fait que l'espace des solutions est non-convexe (Berry et al, 2007; Lin, 2007). En effet, la fonction de coût utilisée peut être convexe selon $\hat{\mathbf{A}}$ ou $\hat{\mathbf{S}}$ individuellement, mais pas toutes les deux au même temps. C'est pourquoi les méthodes développées peuvent avoir de nombreux minima locaux. Ainsi, la convergence dépend beaucoup des initialisations de l'algorithme, et par la

suite, un mauvais choix des initialisations peut entraîner la convergence vers un mauvais minimum local. Une solution possible pour résoudre ce problème est d'effectuer plusieurs initialisations et puis choisir celle qui donne la fonction de coût la plus faible (Cichocki et al, 2009). D'autres solutions consistent à déterminer l'initialisation convenable en utilisant des algorithmes de classification (Rezaei et al, 2011) ou des algorithmes génétiques (Rezaei and Boostani, 2010).

Les algorithmes NMF décrits précédemment peuvent être étendus afin de réduire le nombre de solutions possibles en imposant des contraintes supplémentaires en fonction des informations a priori disponibles sur les inconnues du problème. Ceci peut être réalisé en combinant la fonction de coût initiale avec des fonctions de coût supplémentaires caractérisant les contraintes requises, comme suit

$$J(\hat{\mathbf{A}}, \hat{\mathbf{S}}) = F(\mathbf{X} \parallel \hat{\mathbf{A}}\hat{\mathbf{S}}) + w_A f_A(\hat{\mathbf{A}}) + w_S f_S(\hat{\mathbf{S}}), \tag{1.30}$$

où J représente la nouvelle fonction de coût à optimiser, $f_A(\hat{\mathbf{A}})$ et $f_S(\hat{\mathbf{S}})$ sont des fonctions qui imposent les contraintes nécessaires sur $\hat{\mathbf{A}}$ et $\hat{\mathbf{S}}$, respectivement, et w_A et w_S précisent la pondération de chaque fonction de coût supplémentaire.

Des contraintes supplémentaires ajoutées aux algorithmes de NMF comprennent la parcimonie (Hoyer, 2002), la continuité temporelle (Virtanen, 2007), et la contrainte de lissage (Chen et al, 2006). Les études réalisées par Cichocki et al (2009) montrent de façon plus détaillée les différentes contraintes qui peuvent être imposées à la méthode NMF.

1.4.4 Approches bayésiennes

L'approche bayésienne pour la séparation de sources a été très largement employée (Roberts, 1998; Knuth, 1999; Mohammad-Djafari, 2001; Senecal and Amblard, 2000; Rowe, 2002; Ichir and Mohammad-Djafari, 2003). Parmi les travaux effectués pour l'appliquer à la séparation des sources non-négatives, nous pouvons citer (Roberts and Choudrey, 2003; Moussaoui et al, 2004; Abdallah and Plumbley, 2004; Duarte et al, 2009). L'inférence bayésienne offre un cadre théorique très puissant pour prendre en compte la non-négativité des données et surtout les connaissances a priori sur les coefficients de mélange et les signaux sources.

L'idée principale de l'approche bayésienne pour la séparation de sources consiste à utiliser non seulement la vraisemblance, mais aussi toute information a priori sur les signaux sources et les coefficients de mélange à travers l'assignation des lois de probabilité a priori $p(\mathbf{s})$ et $p(\mathbf{A})$. Selon le théorème de Bayes, la loi a posteriori des données est exprimée comme suit

$$p(\mathbf{A}, \mathbf{S}|\mathbf{X}) = \frac{p(\mathbf{X}|\mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S})}{p(\mathbf{X})} \propto p(\mathbf{X}|\mathbf{A}, \mathbf{S})p(\mathbf{A})p(\mathbf{S}) \tag{1.31}$$

où $\mathbf{A}$ et $\mathbf{S}$ sont ici considérées indépendantes.

À partir de cette densité de probabilité a posteriori, l'estimation de $\mathbf{A}$ et $\mathbf{S}$ peut être réalisée en utilisant divers estimateurs bayésiens tels que le Maximum A Posteriori (MAP) et la Moyenne a Posteriori (MP) (Robert, 2001). L'optimisation de ces critères est souvent réalisée par la méthode de Monte-Carlo par Chaînes de Markov (MCMC pour *Markov Chain Monte Carlo* en anglais) (Gilks et al, 1999).

Cependant, la tâche la plus délicate dans l'approche bayésienne est le choix des distributions de probabilité appropriées pour modéliser les informations disponibles a priori sur le bruit, les signaux sources, et les coefficients de mélange.

Pour plus de détails sur l'utilisation des approches bayésiennes en séparation de sources, nous citons les travaux réalisés dans (Snoussi et al, 2002; Stolyarov et al, 2005; Kuruoglu, 2010).

1.5 Mélange linéaire-quadratique

Comme il a été mentionné précédemment dans ce chapitre, la plupart des recherches traitant le problème de séparation de sources supposent que le modèle de mélange est linéaire instantané, où les signaux observés résultent des combinaisons linéaires des signaux sources. Néanmoins, pour certaines applications, le modèle de mélange linéaire n'est pas valable et doit être remplacé par un modèle non-linéaire. Dans ce contexte, le modèle de mélange linéaire-quadratique (LQ) a montré sa capacité d'offrir une meilleure description du processus de mélange et des interactions qui peuvent avoir lieu entre les signaux sources. En effet, les recherches menées dans le cadre du

modèle LQ ont démontré la pertinence de son utilisation dans diverses applications telles que la télédétection (Meganem et al, 2014a,b; Eches and Guillaume, 2014) , l'analyse de données issues des capteurs chimiques (Bedoya, 2006; Ando et al, 2015) et le traitement des documents scannés (Liu and Wang, 2013; Almeida and Almeida, 2012; Merrikh-Bayat et al, 2011; Duarte et al, 2011).

L'originalité du modèle LQ, en le comparant au modèle linéaire, est qu'il permet de décrire les interactions qui peuvent se produire entre les différentes sources. Cela est réalisé par l'ajout des termes croisés de second ordre dans le modèle linéaire. Dans ce cas-là, en considérant K observations résultant d'un mélange LQ de L sources, la relation entre les signaux observés et les signaux sources peut être caractérisée par l'équation suivante

$$x_i(n) = \sum_{j=1}^{L} a_j(i)s_j(n) + \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)s_j(n)s_k(n), \tag{1.32}$$

où $x_i(n)$ est le $i^{ème}$ signal observé à l'échantillon d'indice n, $s_j(n)$ est le $j^{ème}$ signal source inconnu, $a_j(i)$ est le coefficient linéaire associé au $j^{ème}$ signal source et au $i^{ème}$ signal observé, $a_{j,k}(i)$ est le coefficient quadratique associé au $i^{ème}$ signal observé et résultant de l'interaction entre la $j^{ème}$ et la $k^{ème}$ sources. Il est à noter que l'ensemble des *vraies sources* $s_j(n)$ et des *pseudo-sources* $s_j(n)s_k(n)$ est appelé *sources étendues*.

Il existe un cas particulier du modèle LQ, appelé le modèle bilinéaire (BL), où les coefficients quadratiques $a_{j,j}(i)$ du terme au carré $s_j(n)^2$ sont considérés nuls. Dans ce cas, le modèle (1.32) devient comme suit

$$x_i(n) = \sum_{j=1}^{L} a_j(i)s_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)s_j(n)s_k(n). \tag{1.33}$$

De même, une version quadratique du modèle LQ est obtenue lorsque tous les coefficients $a_j(i)$ sont nuls, telle que

$$x_i(n) = \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)s_j(n)s_k(n). \tag{1.34}$$

La réécriture du modèle de mélange LQ sous forme matricielle est comme suit

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), \tag{1.35}$$

où

- $\mathbf{x}(n) = [\ x_1(n) \cdots x_K(n)\ ]^T$ est le vecteur contenant les K signaux observés à l'échantillon d'indice n.

- $\mathbf{s}(n) = [\ s_1(n) \cdots s_L(n) \quad s_1(n)s_1(n) \quad s_1(n)s_2(n) \cdots s_L(n)s_L(n)\ ]^T$ est le vecteur contenant les vraies sources et pseudo-sources à l'échantillon d'indice n .

- $\mathbf{A} = \begin{pmatrix} a_1(1) & \cdots & a_L(1) & a_{1,1}(1) & \cdots & a_{L,L}(1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_1(K) & \cdots & a_L(K) & a_{1,1}(K) & \cdots & a_{L,L}(K) \end{pmatrix}$ est la matrice, de taille $K \times (\frac{L(L+3)}{2})$, contenant les coefficients linéaires et les coefficients quadratiques.

Dans le cadre de cette thèse, nous nous intéressons au développement de nouvelles méthodes de séparation de sources basées sur le modèle LQ. A cette fin, nous consacrons la partie suivante à un tour d'horizon sur les méthodes de séparation de sources proposées pour ce type de mélange.

1.6 Méthodes linéaires-quadratiques

Les méthodes de séparation de sources traitant le modèle LQ peuvent être classées principalement en quatre catégories décrites dans les sous-sections suivantes.

1.6.1 Extension des méthodes d'ACI

1.6.1.1 Méthodes d'ACI pour séparer les sources i.i.d

Une première classe de ces méthodes a été destinée aux mélanges linéaires-quadratiques déterminés par rapport aux vraies sources, c'est-à-dire, lorsque le nombre de vraies sources L est égal au nombre d'observations K. Dans ce cas, la séparation peut être réalisée en utilisant des réseaux récurrents non-linéaires.

Parmi les méthodes proposées, nous citons (Hosseini and Deville, 2003) où les auteurs envisagent une extension de la méthode proposée par Jutten and Herault (1991) pour traiter le modèle linéaire afin de séparer les sources réelles mélangées selon le modèle LQ. Le point clé de cette méthode consiste à utiliser une structure de séparation basée sur les réseaux récurrents et les algorithmes d'adaptation permettant d'obtenir des composantes indépendantes en sortie du réseau.

Une deuxième méthode décrite dans (Mokhtari et al, 2009) propose une extension de la méthode proposée par Comon (1994) pour traiter le modèle linéaire. Le principe de cette approche consiste à adapter la structure récurrente non-linéaire pour estimer les signaux sources, cependant, les coefficients de mélange sont déterminés à l'aide d'une méthode de minimisation de l'information mutuelle. Dans le même contexte, nous trouvons d'autres méthodes proposées dans (Hosseini and Deville, 2013; Chaouchi et al, 2009; Merrikh-Bayat et al, 2011) qui permettent de traiter le mélange LQ et plusieurs autres types de mélanges non-linéaires.

Une autre approche basée sur la maximisation de la vraisemblance des observations a été proposée dans (Hosseini and Deville, 2004, 2013; Chaouchi et al, 2010a). Le lien entre les méthodes de SAS basées sur la maximisation de vraisemblance et la minimisation de l'information mutuelle a été étudié dans (Deville et al, 2011) pour les mélanges non-linéaires, y compris les mélanges LQ.

Les méthodes de séparation basées sur la structure récurrente sont intéressantes puisqu'elles ne nécessitent que la connaissance de l'expression analytique du modèle de mélange contrairement aux méthodes basées sur le modèle inverse. Toutefois, les structures récurrentes non-linéaires peuvent avoir quelques limitations telles que :

- elles peuvent avoir des problèmes d'instabilité, ou elles peuvent même aboutir à un comportement chaotique (Hosseini and Deville, 2003; Deville and Hosseini, 2009) (aussi (Deville and Hosseini, 2007b, 2010) pour les réseaux étendus qui permettent de résoudre ces problèmes : ces réseaux ont été utilisés par Ando et al (2012) comme des outils originaux pour la résolution d'équations non-linéaires),
- elles peuvent avoir des points d'équilibre parasites,
- elles nécessitent un calcul itératif pour chaque vecteur de sortie.

Pour pallier ce problème, une solution simple peut être adoptée en augmentant le nombre d'observations pour atteindre le nombre de sources étendues. Ainsi, le modèle de mélange est considéré déterminé et linéaire par rapport à ces sources étendues et surdéterminé par rapport aux vraies sources. Parmi les travaux réalisés dans ce cadre, nous citons (Duarte et al, 2010) où les auteurs ont basé leur méthode sur la minimisation de l'information mutuelle des signaux sources et ils n'ont traité que le cas de mélanges de deux sources.

Une autre approche proposée pour estimer le modèle de mélange exploite les expressions de certains cumulants ou moments des signaux observés par rapport à ceux des signaux sources et aux coefficients de mélange. Ainsi, la résolution de ces équations pour des valeurs connues des cumulants ou des moments des signaux observés sert principalement à déterminer une estimation des valeurs de coefficients de mélange (à des indéterminations près). L'idée de cette approche a été appliquée dans (Chaouchi et al, 2010b) pour traiter les mélanges quadratiques.

Toujours dans le même contexte, une méthode d'identification aveugle du modèle LQ qui se base sur l'utilisation de la décomposition en valeurs singulières des matrices de cumulants a été proposée dans (Krob and Benidir, 1993). Dans cette dernière, les sources sont supposées indépendantes, circulaires et centrées.

1.6.1.2 Méthodes d'ACI pour séparer les sources non-i.i.d

D'autres méthodes de séparation de sources ont été élaborées en considérant que les signaux sources sont non-i.i.d et en exploitant leur autocorrélation et/ou leur non-stationnarité. Ces deux propriétés ont été utilisées pour l'extension de la méthode basée sur la vraisemblance proposée par Hosseini and Deville (2014).

Une méthode d'identification aveugle des mélanges LQ a été proposée par Abed-Meraim et al (1996). Dans ce travail, les sources sont supposées être auto-corrélées, circulaires, complexes et mutuellement indépendantes. Le principe de cette méthode est basé sur la diagonalisation conjointe d'un ensemble de matrices de corrélation des signaux observés. En exploitant le même principe, une méthode d'extraction des termes croisés de second ordre a été également proposée par Georgiev (2001). Cette méthode suppose que les sources sont non-corrélées avec des auto-corrélations distinctes.

1.6.2 Extension des méthodes NMF

En examinant la forme matricielle du modèle de mélange LQ dans (1.35), nous pouvons remarquer qu'il peut être exprimé comme étant un modèle linéaire par rapport aux sources étendues. Par conséquent, si les deux matrices $\mathbf{A}$ et $\mathbf{S}$ sont non-négatives, alors il est possible de développer des méthodes de séparation de sources pour les estimer conjointement par l'extension des méthodes NMF linéaires classiques. Cependant, dans ce cas, les algorithmes d'optimisation utilisés peuvent avoir des structures beaucoup plus complexes vu que le calcul des dérivées de la fonction de coût doit prendre en compte les coefficients quadratiques $a_{j,k}(i)$ ainsi que les termes croisés de second ordre $s_j(n)s_k(n)$.

Parmi les extensions des méthodes NMF proposées dans ce cadre, nous citons (Meganem et al, 2014a,b) où les auteurs proposent une variante de la méthode NMF pour séparer les mélanges linéaires-quadratiques. Cette méthode prend en compte le fait que les termes croisés de second ordre sont les produits de deux vraies sources présentes dans le mélange. En outre, à part la contrainte de non-négativité, les auteurs ont tenu compte d'autres contraintes liées aux coefficients de mélange (les coefficients

linéaires ont la somme égale à un et les coefficients quadratiques sont inférieur à 0.5). Dans ce travail, la fonction de coût à minimiser est définie comme suit

$$J = \frac{1}{2}\sum_{i,n}\left(x_i(n) - \sum_{j=1}^{L}\hat{a}_j(i)\hat{s}_j(n) - \sum_{j=1}^{L}\sum_{k=j}^{L}\hat{a}_{j,k}(i)\hat{s}_j(n)\hat{s}_k(n)\right)^2, \tag{1.36}$$

où "ˆ" se réfère aux valeurs estimées.

La minimisation de J a été réalisée en utilisant un algorithme de gradient projeté à pas fixe, sa variante améliorée en utilisant une mise à jour de Newton, et un algorithme multiplicatif.

Une approche similaire, dédiée aux mélanges bilinéaires, est décrite dans (Eches and Guillaume, 2014). En outre, les travaux réalisés par Liu and Wang (2013) traitent une configuration spécifique des mélanges LQ en considérant deux mélanges de deux sources originales, avec la même contribution quadratique pour les deux mélanges.

1.6.3 Extension des méthodes ACPa

Des méthodes exploitant la parcimonie ont été proposées pour séparer les mélanges linéaires-quadratiques. Dans ce cadre, nous citons les travaux réalisés dans (Deville and Hosseini, 2007a) où les auteurs proposent une extension aux mélanges bilinéaires de la méthode LI-TEMPCORR (Deville and Puigt, 2007) proposée initialement pour traiter les mélanges linéaires. L'hypothèse de parcimonie utilisée dans cette méthode ne nécessite que l'existence d'une zone temporelle où une seule source est active. En contrepartie, cette méthode exige que les sources soient centrées et linéairement indépendantes.

Dans le même contexte, une autre méthode basée sur la parcimonie a été développée par Duarte et al (2012) pour traiter les mélanges sur-déterminés. L'idée de cette méthode consiste à annuler, d'abord, la partie quadratique des mélanges LQ, puis à appliquer une méthode de SAS linéaire à la partie linéaire restante. Notons que, dans cette méthode, le cas d'un mélange de deux sources est considéré.

1.6.4 Approches bayésiennes

Il existe aussi des approches bayésiennes qui traitent le mélange LQ (Duarte et al, 2009, 2011). Ce type de méthode permet d'éviter les problèmes liés aux mélanges déterminés de vraies sources puisqu'elles n'utilisent pas explicitement un système de séparation. L'approche décrite dans (Duarte et al, 2009, 2011) utilise la méthode MCMC et considère que les sources sont statistiquement indépendantes et ont une structure temporelle. Cependant, l'utilisation des méthodes bayésiennes reste limitée à cause de la complexité de leur mise en œuvre et leur coût de calcul élevé.

Un panorama des méthodes linéaires-quadratiques ainsi que des méthodes post-non-linéaires est présenté dans (Deville and Duarte, 2015).

1.7 Conclusions

Ce premier chapitre a été consacré à la présentation du problème de SAS ainsi que des modèles de mélange et des méthodes les plus utilisées. Tout d'abord, nous avons donné un aperçu général sur le problème de la SAS. Ensuite, nous avons présenté une vue d'ensemble sur les modèles de mélange linéaires ainsi que les méthodes de SAS classiques proposées pour séparer les mélanges linéaires instantanés. En effet, le modèle linéaire instantané est le modèle le plus utilisé dans la SAS, c'est pourquoi nous trouvons une vaste production scientifique à ce sujet. Par la suite, nous avons présenté le modèle linéaire-quadratique qui permet l'ajout des termes croisés du second ordre par rapport au modèle linéaire. Enfin, nous avons présenté un panorama des méthodes de SAS proposées pour le traiter.

Dans le cadre de cette thèse, nous nous intéressons à la séparation des mélanges linéaires-quadratiques. Deux de nos méthodes de SAS proposées sont dédiées au démélange hyperspectral des images issues de la télédétection. Pour cela, nous avons consacré le chapitre suivant pour faire un tour d'horizon sur les différentes méthodes de démélange hyperspectral.

Chapitre 2

Démélange spectral

Sommaire

2.1 Introduction 35
2.2 Principe de la télédétection hyperspectrale 36
2.3 Modèle linéaire pour le démélange hyperspectral 39
2.4 Méthodes de démélange linéaire 41
2.4.1 Approches géométriques 42
2.4.1.1 Approches de pixels purs 42
2.4.1.2 Approches de Minimisation de Volume 43
2.4.2 Approches statistiques 44
2.4.3 Approches de régression parcimonieuse 45
2.5 Modèles non-linéaires pour le démélange hyperspectral . 46
2.5.1 Modèle de mélange intime 48
2.5.2 Modèles bilinéaires pour le démélange hyperspectral 49
2.5.3 Modèle linéaire-quadratique pour le démélange hyperspectral 52
2.5.4 Autres modèles basés sur des approximations physiques . . 52
2.6 Méthodes de démélange non-linéaires 53
2.6.1 Méthodes de démélange non-linéaires basées sur des modèles 54
2.6.1.1 Démélange non-linéaire supervisé 54
2.6.1.2 Démélange non-linéaire non-supervisé 56
2.6.2 Méthodes de démélange non-linéaires non-basées sur des modèles 57
2.7 Conclusions 59

2.1 Introduction

La télédétection spatiale permet d'étudier, à distance, la surface de la Terre grâce à des capteurs placés à bord d'un satellite ou d'un aéronef. Actuellement, de nombreux instruments de télédétection permettent d'obtenir des informations sur la surface de la Terre dans des centaines de bandes spectrales très étroites correspondant aux différentes longueurs d'onde. Ces instruments sont appelés des capteurs hyperspectraux. Ces derniers sont capables de produire des informations beaucoup plus détaillées que celles fournies par les capteurs classiques tels que les capteurs multispectraux, qui collectent des informations en utilisant seulement quelques bandes spectrales larges. L'information supplémentaire apportée par ces capteurs est utile dans différents domaines tels que l'agriculture, la géographie, la géologie, l'identification des minéraux, la planification urbaine, la surveillance environnementale, et la classification des données. Les images spatiales de télédétection sont acquises dans la plupart des cas sur des scènes hétérogènes composées de différents matériaux. Ceci entraîne des pixels mélangés dans l'image hyperspectrale et ainsi la présence de plusieurs signatures spectrales dans un seul pixel. Afin de mieux comprendre la scène, le traitement et l'analyse des données dans l'imagerie hyperspectrale nécessitent parfois d'estimer les spectres des matériaux purs et leurs proportions dans un pixel mélangé. Ce genre de traitement est ce qu'on appelle le démélange spectral. Le concept général du démélange spectral consiste à identifier et quantifier les matériaux inconnus dans une région de la Terre, ce qui est particulièrement utile dans diverses applications. Pour en citer quelques-unes, en agriculture, il permet de suivre l'évolution des produits agricoles (Asner and Lobell, 2000; Yang et al, 2010), il analyse également les composants biochimiques dans le processus de la végétation (Oppelt and Mauser, 2003), et en sciences de la mer, il permet d'analyser le mélange spectral des macroalgues trouvés dans la mer (Uhl et al, 2013).

Dans ce chapitre, nous abordons le problème de démélange hyperspectral. Pour cela, nous commençons, tout d'abord, par la présentation des concepts de base de l'imagerie hyperspectrale, et par la suite, nous présentons un état de l'art sur les méthodes de démélange pour les deux cas linéaire et non-linéaire.

2.2 Principe de la télédétection hyperspectrale

L'analyse spectroscopique se réfère à l'étude des interactions entre les matériaux et les rayonnements électromagnétiques en utilisant un spectromètre. La télédétection hyperspectrale combine l'imagerie et la spectroscopie dans un seul système. L'idée est d'équiper un satellite ou un aéronef d'un spectromètre pour capturer à distance les données. Ceci est fait en utilisant la télédétection passive (voir Figure 2.1), où l'énergie lumineuse du soleil touche un objet sur la terre, et par la suite, une partie de la lumière réfléchie est enregistrée par les capteurs hyperspectraux. Ces capteurs permettent d'enregistrer les données dans différentes bandes spectrales du visible à l'infrarouge.

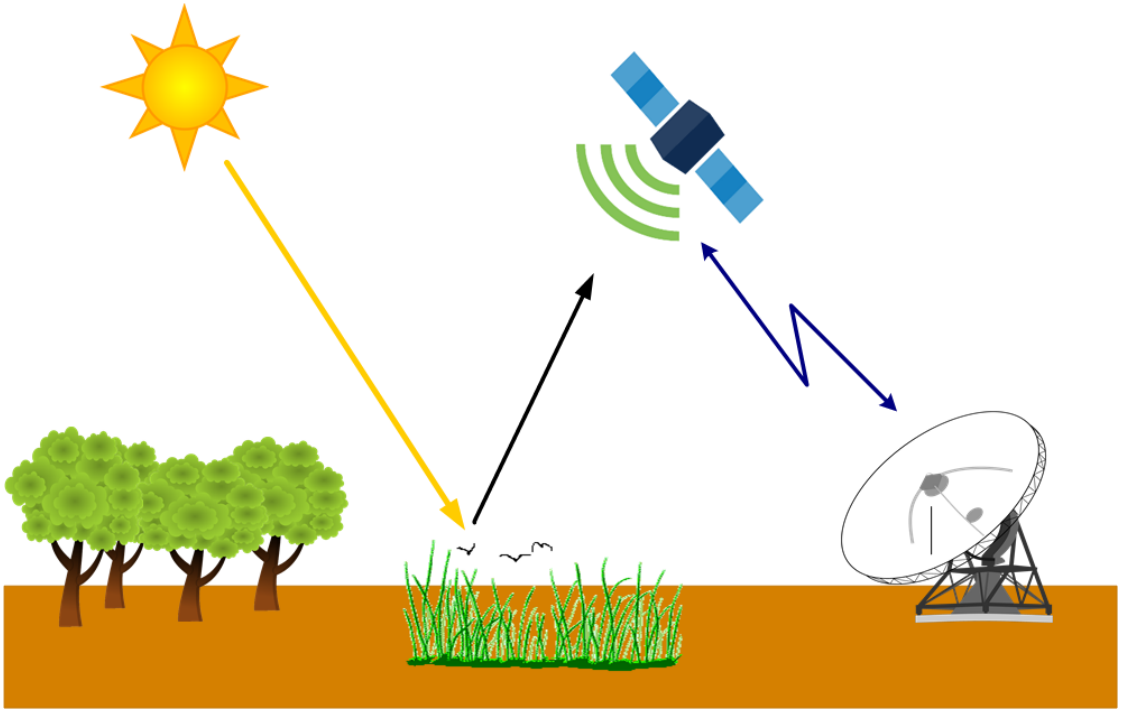


Figure 2.1: *La télédétection passive par des capteurs satellitaires.*

L'imagerie hyperspectrale est typiquement collectée et représentée sous la forme d'un cube de données ou d'une image en 3D avec des informations spatiales représentées dans le plan X-Y et des informations spectrales représentées dans la direction Z. Ainsi, chaque pixel de l'image est représenté par un spectre de réflectance qui correspond au rapport de l'énergie réfléchie sur l'énergie incidente pour chaque bande spectrale. Notons que la réflectance est une grandeur sans unité se situant dans l'intervalle $[0, 1]$.

La Figure 2.2 illustre un exemple de cube de données hyperspectrales formé par des images enregistrées dans différentes bandes spectrales.

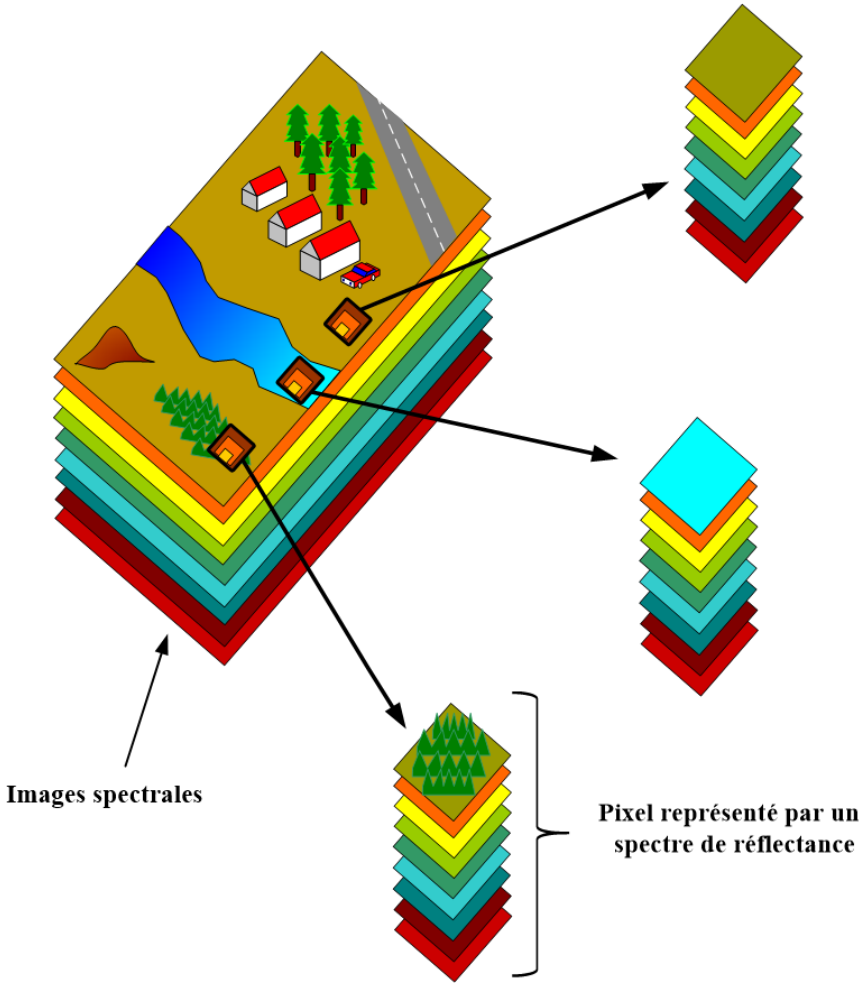


Figure 2.2: *Exemple d'une image hyperspectrale.*

Afin de bénéficier d'une résolution spectrale plus élevée, les capteurs hyperspectraux offrent une résolution spatiale relativement faible à cause des limitations matérielles et de la réalisation des compromis entre les résolutions spatiale et spectrale. Ceci entraîne des pixels mélangés dans l'image hyperspectrale, autrement dit, la présence de plusieurs signatures spectrales dans un seul pixel. La Figure 2.3 illustre le concept des pixels mélangés lorsque les données sont considérées comme des mélanges des signatures spectrales des matériaux purs trouvés dans la scène, dénommées aussi, les *endmembers*. Les pixels correspondant à une seule signature spectrale sont appelés pixels purs alors que les pixels mélangés sont appelés mixels. Les mixels sont généralement

trouvés dans tous les endroits de la scène à cause de la résolution spatiale faible de l'image hyperspectrale. Cependant, nous pouvons aussi trouver des pixels mélangés lorsque la résolution spatiale est élevée (Villa et al, 2011). Une partie d'une scène ayant une résolution spatiale élevée ($4m \times 4m$) est illustrée en Figure 2.3. Celle-ci montre un exemple de pixels qui contiennent à la fois du sol et de la végétation.

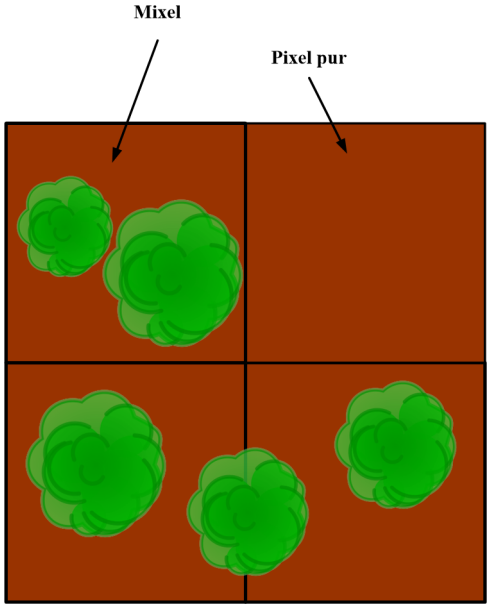


Figure 2.3: *Illustration du principe des mélanges dans l'imagerie hyperspectrale.*

Ainsi, nous pouvons considérer la réflectance observée dans un pixel comme un mélange de plusieurs réflectances de matériaux purs caractérisé par différents paramètres qui contribuent à sa formation. Ce phénomène de mélange peut être considéré comme étant un mélange linéaire ou non-linéaire de certaines signatures spectrales des matériaux constitutifs présents dans une scène. Le modèle linéaire ne prend pas en compte les interactions entre les matériaux purs. Le modèle non-linéaire, plus réaliste, vise à considérer les interactions entre les matériaux purs impliqués dans la scène.

Le modèle linéaire a été largement utilisé pour résoudre le problème de démélange hyperspectral (Shaw and Manolakis, 2002; Kerekes and Schott, 2006). Cela est dû à la simplicité de son interprétation physique. Cependant, un tel modèle n'est généralement valable que lorsque la scène observée est plate et éclairée de façon homogène. Néanmoins, dans le milieu urbain, les réflexions multiples de la lumière entre les différentes surfaces présentes dans la scène observée peuvent rendre le modèle de mélange beaucoup plus

complexe. Dans ce cas, le modèle de mélange linéaire n'est plus valable, et doit être remplacé par un modèle non-linéaire (Meganem et al, 2014a).

Dans les sections qui suivent, nous présentons les modèles linéaires et non-linéaires, ainsi que les méthodes utilisées pour résoudre le problème de démélange hyperspectral.

2.3 Modèle linéaire pour le démélange hyperspectral

La plupart des travaux de recherche en télédétection ont utilisé le modèle de mélange linéaire pour résoudre le problème de démélange hyperspectral (Bioucas-Dias et al, 2012). L'idée derrière le modèle linéaire est que chaque rayon lumineux incident interagit avec un seul endmember avant d'atteindre le capteur et que la probabilité d'interagir avec un endmember donné est proportionnelle à sa fraction d'abondance dans le pixel. Ce modèle est approprié lorsque la scène est plate et contient des régions spatialement séparées et formées de composants macroscopiquement purs. La Figure 2.4 présente un exemple illustratif des effets de réflexion de la lumière dans le cas du modèle linéaire.

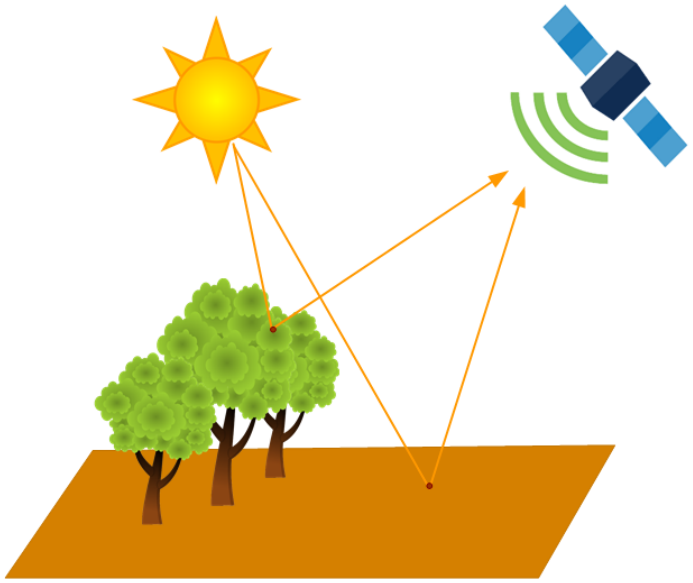


Figure 2.4: *Principe du modèle linéaire.*

Dans ce cas, le spectre d'un mixel peut être modélisé par un mélange linéaire de L spectres de matériaux purs, selon l'équation de mélange suivante

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \mathbf{b}_i, \tag{2.1}$$

où $\mathbf{x}_i = [x(1), \cdots, x(N)]^T$ est le spectre de réflectance associé au $i^{ème}$ pixel observé, $\mathbf{s}_j = [s_j(1), \cdots, s_j(N)]^T$ est le spectre du $j^{ème}$ matériau pur, N est le nombre de bandes spectrales, $a_j(i)$ est la fraction d'abondance du $j^{ème}$ matériau pur dans le pixel i considéré, et le vecteur $\mathbf{b_i}$ désigne un bruit additif, supposé gaussien, de moyenne nulle, et spatialement et spectralement i.i.d.. Notons qu'il existe différentes sources de bruit dans les données de télédétection. Parmi elles, nous citons le bruit du capteur causé par les limites de la qualité et le vieillissement des composants, le bruit du canal lors de la transmission des données, le bruit thermique causé par les circuits, etc. Cependant, en général, il est difficile de prévoir la quantité et le type de bruit dans les données réelles acquises par les capteurs hyperspectraux. Pour cette raison, dans la plupart des cas, on considère que le bruit est dû à l'appareil et aux imprécisions de modélisation.

Dans le cas du modèle linéaire, les données hyperspectrales imposent la contrainte de non-négativité des endmembers tandis que les fractions d'abondance sont contraintes par la non-négativité, ainsi que la somme égale à un dans chaque pixel de l'image. Par conséquent, les données hyperspectrales peuvent être délimitées par un simplexe dont les sommets sont les endmembers comme le montre la Figure 2.5.

Nous pouvons conclure à partir de la Figure 2.5 que le modèle de mélange linéaire permet de fournir une bonne représentation des données hyperspectrales en utilisant des concepts géométriques convexes (tels qu'un simplexe de dimension $(L-1)$ et de sommets les endmembers) (Boardman, 1995), ce qui explique pourquoi ce modèle est très répandu dans les travaux de recherche en télédétection (Bioucas-Dias et al, 2012; Ma et al, 2014).

La partie suivante est consacrée à la description des méthodes de démélange dédiées au modèle linéaire.

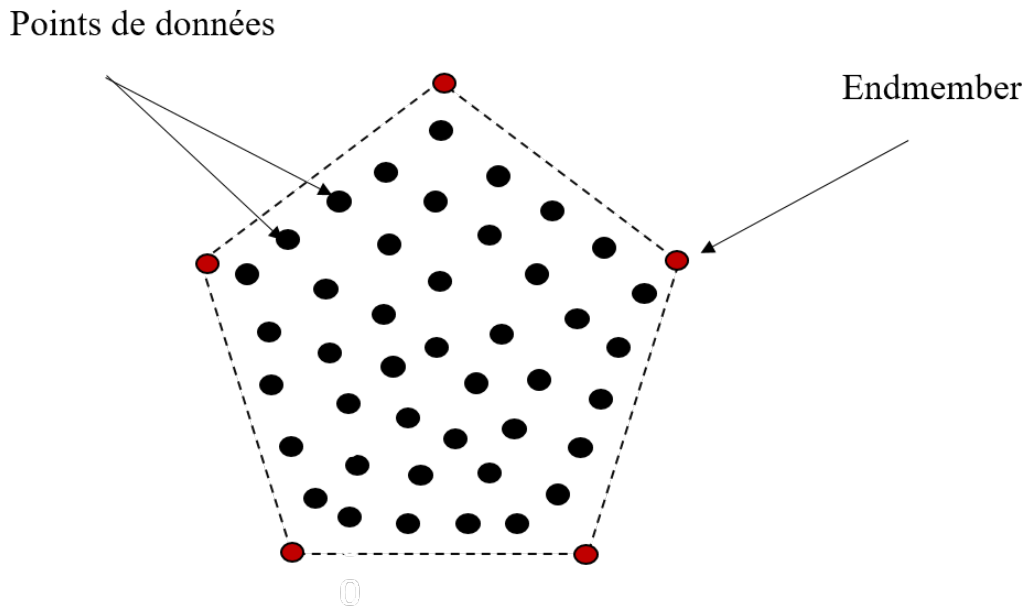


Figure 2.5: *Projection en dimension 2 d'un simplexe de dimension 4 contenant les données hyperspectrales.*

2.4 Méthodes de démélange linéaire

Les méthodes de démélange linéaire sont souvent classées en trois catégories : approches géométriques, statistiques, et de régression parcimonieuse (Bioucas-Dias et al, 2012).

- Les approches géométriques supposent que les vecteurs observés sont positionnés dans un simplexe ou un cône positif.
- Les approches statistiques utilisent des méthodes d'estimation des paramètres pour identifier les fractions d'abondance et les endmembers.
- Les méthodes de régression parcimonieuse supposent que les abondances sont parcimonieuses, et par la suite, elles formulent le problème de démélange comme étant une régression parcimonieuse.

2.4.1 Approches géométriques

L'approche géométrique de démélange peut encore être classée en deux catégories : approches de Pixels Purs (PP), et approches de Minimisation du Volume (MV).

2.4.1.1 Approches de pixels purs

De nombreuses méthodes reposent sur l'hypothèse de PP, qui suppose que tous les matériaux présents dans l'image sont représentés par au moins un PP. Cette hypothèse est tout à fait raisonnable lorsque l'image a une bonne résolution spatiale.
Les méthodes d'extraction des endmembers qui reposent sur l'hypothèse de PP ont été largement utilisées en télédétection. Les méthodes les plus connues qui utilisent cette approche sont PPI (pour *Pixel Purity Index*) (Boardman, 1994), N-FINDR (Winter, 1999), et VCA (pour *Vertex Component Analysis*) (Nascimento and Bioucas-Dias, 2005b).
La méthode PPI commence par la réduction de la dimensionnalité des données en utilisant la méthode MNF (pour *Maximum Noise Fraction* en anglais) (Green et al, 1988). Par la suite, un grand nombre de vecteurs aléatoires de norme unité (*skewers*) est défini, et chaque spectre des données est projeté sur ces skewers. Pour chaque projection, les pixels ayant les valeurs extrêmes sont séléctionnés. A la fin, les pixels sont évalués et ceux qui ont le score le plus élevé sont considérés comme les pixels les plus purs dans les données. Le score est calculé par le nombre de fois où un pixel a été sélectionné comme point extrême. La méthode N-FINDR cherche à trouver l'ensemble des pixels qui maximisent le volume du simplexe contenant les données. Tout d'abord, une réduction de la dimension des données est réalisée. Ensuite, des pixels choisis au hasard sont considérés comme les endmembers et le volume formé par ces pixels est calculé. Par la suite, ce volume est recalculé en remplaçant l'un des endmembers par un autre pixel. Si le remplacement entraîne une augmentation du volume, le pixel est considéré comme un endmember. Cette procédure est répétée jusqu'à ce qu'il n'y ait plus de possibilité de remplacement des endmembers. La méthode VCA est une méthode itérative qui utilise le sous-espace défini par les endmembers déjà identifiés et elle projette les données disponibles sur une direction orthogonale à ce sous-espace.

L'extrémité de cette projection (le pixel correspondant à la valeur maximale) correspond au nouveau endmember.

2.4.1.2 Approches de Minimisation de Volume

L'approche de minimisation de volume suppose l'inexistence des pixels purs dans les données, ce qui rend le problème d'optimisation non-convexe, et ainsi, beaucoup plus difficile à résoudre.
Le principe de cette méthode consiste à trouver l'ensemble des endmembers qui minimisent le volume du simplexe défini par les données observées (Bioucas-Dias et al, 2012).

Parmi les méthodes les plus connues, nous trouvons la méthode MVC-NMF (pour *Minimum Volume Constrained Non-negative Matrix Factorization*) (Miao and Qi, 2007) qui considère que les données sont très mélangées. Par rapport à la méthode NMF originale, la méthode MVC-NMF introduit la contrainte de minimisation de volume dans le cadre d'une régularisation pour limiter l'espace des solutions. Cette méthode permet de respecter la contrainte de non-négativité, et la minimisation de la fonction de coût est effectuée en estimant alternativement les fractions d'abondance et les endmembers.

D'autres méthodes dénommées MVSA (pour *Minimum Volume Simplex Analysis*) (Li and Bioucas-Dias, 2008) et SISAL (pour *Simplex Identification via variable Splitting and Augmented Lagrangian*) (Bioucas-Dias, 2009) permettent de résoudre le problème d'optimisation en ajoutant des contraintes convexes. Ceci est réalisé en utilisant une séquence de sous-contraintes quadratiques pour la méthode MVSA et en utilisant des lagrangiens augmentés pour la méthode SISAL. Ces deux méthodes sont initialisées par les résultats de la méthode VCA, présentée dans la Sous-section 2.4.1.1, qui est basée sur la présence de pixels purs.

2.4.2 Approches statistiques

Lorsque les données sont fortement mélangées, le nombre de spectres sur les facettes du simplexe (les faces planes polygonales qui forment le simplexe) formé par les pixels de la scène diminue, et ceci entraînera la dégradation des résultats obtenus par les méthodes géométriques. Pour surmonter cette difficulté, les chercheurs se sont orientés vers les approches statistiques. Ces dernières sont puissantes mais de complexité calculatoire supérieure à celle des méthodes géométriques.

Lorsque les fractions d'abondance et les endmembers doivent être estimés, le démélange est considéré comme un problème de SAS. Dans ce contexte, l'ACI a été proposée pour le démélange des données hyperspectrales en considérant que les endmembers ou les fractions d'abondance sont mutuellement indépendants (Parra et al, 2007; Tu, 2000; Kosaka et al, 2005). Cependant, cette hypothèse n'est pas valide à cause des contraintes physiques imposées sur les données de télédétection (Nascimento and Bioucas-Dias, 2005a). En effet, Nascimento and Bioucas-Dias (2005a) ont conclu que l'ACI fournit des endmembers mal estimés à cause de la dépendance entre les données. De plus, la performance des ces méthodes a tendance à se dégrader avec la variabilité des endmembers et l'augmentation de leur nombre.

Les approches bayésiennes sont parmi les approches statistiques qui ont été appliquées pour le démélange hypespectral. Ces méthodes ont l'avantage de pouvoir exploiter les connaissances a priori sur les inconnues du problème. Pour ce faire, des distributions de probabilité a priori sont utilisées pour modéliser les informations disponibles sur les fractions d'abondance et les endmembers. Dans ce cas, les paramètres des distributions, ainsi que les inconnues du problème sont estimés conjointement. Cependant, un inconvénient majeur des méthodes bayésiennes est la non connaissance exacte des distributions de probabilité a priori en pratique.

Le modèle NCM (pour *Normal Compositional Model*), appelé aussi SMM (pour *Stochastic Mixing Model*), a été proposé par Eismann and Hardie (2004). Ce modèle considère que les pixels sont mélangés selon une combinaison linéaire des endmembers aléatoires. Dans (Eismann and Hardie, 2004), une méthode non-supervisée a été proposée pour estimer conjointement les endmembers et les fractions d'abondance.

Eches et al (2010b) utilisent une méthode semi-supervisée pour estimer, à partir d'une bibliothèque, les endmembers ainsi que leur nombre. Les fractions d'abondance sont estimées dans ce cas d'une manière parallèle.

Dans le même cadre, une approche bayésienne basée sur la méthode MCMC est proposée par Eches et al (2010a), où les valeurs moyennes des endmembers sont supposées connues. Dans (Zare et al, 2013), une approche similaire est adoptée, mais cette fois-ci les valeurs de covariance des endmembers sont supposées être connues. Dans (Nascimento and Bioucas-Dias, 2007, 2009a), les méthodes proposées utilisent des lois a priori qui prennent en considération la non-négativité et la somme égale à un des fractions d'abondance en les modélisant par des lois de Dirichlet.

2.4.3 Approches de régression parcimonieuse

Lorsqu'une bibliothèque spectrale est disponible, le démélange peut être réalisé en identifiant les endmembers à partir de cette bibliothèque. Celle-ci peut être construite par une collecte de données sur terrain ou obtenue directement auprès d'un laboratoire spécialisé. Ce type de démélange a été réalisé en utilisant une régression linéaire basée sur des termes régulateurs pour imposer la parcimonie. L'ensemble des endmembers est identifié en réduisant, de manière récursive, la bibliothèque spectrale.

Parmi les méthodes classiques utilisées pour résoudre ce problème, on trouve la méthode OMP (pour *Orthogonal Matching Pursuit* en anglais) (Pati et al, 1993) qui utilise la norme l_0 comme un terme régulateur de parcimonie, les méthodes BP (pour *Basis Pursuit* en anglais) et BPDN (pour *Basis Pursuit Denoising* en anglais) (Chen et al, 1998) qui utilisent la norme l_1.

Notons que les méthodes OMP et BPDN ne sont pas spécifiquement conçues pour le démélange hyperspectral, cependant, il existe de nombreuses méthodes de régression parcimonieuse qui sont proposées spécifiquement pour le démélange hyperspectral (Bioucas-Dias and Figueiredo, 2010; Iordache et al, 2011). Les méthodes de démélange hyperspectral les plus connues qui utilisent la régression parcimonieuse sont SUnSAL (pour *Sparse Unmixing by variable Splitting and Augmented Lagrangian* en anglais) (Bioucas-Dias and Figueiredo, 2010) et SUnSAL-TV (*TV* pour *Total Variation* en

anglais) (Iordache et al, 2011). SUnSAL et ses variantes sont basées sur la méthode des directions alternées ADA (en anglais ADMM pour *Alternating Direction Method of Multipliers*) (Eckstein and Bertsekas, 1992). Celle-ci permet de résoudre le problème des moindres carrés sous-contraintes qui est utilisé pour calculer les fractions d'abondance de spectres connus, le problème de BP sous-contraintes qui permet de trouver les mélanges parcimonieux de spectres à partir d'une bibliothèque spectrale, et le problème BPDN sous-contraintes qui est une généralisation de BP admettant les erreurs de modélisation (par exemple, le bruit dans les observations). SUnSAL-TV est similaire à SUnSAL, avec l'ajout d'un terme régulateur de variation totale (Rudin et al, 1992) qui permet d'intégrer l'homogénéité spatiale pour tenir compte de la possibilité que deux pixels voisins aient des fractions d'abondance similaires pour le même endmember dans le problème de démélange parcimonieux.

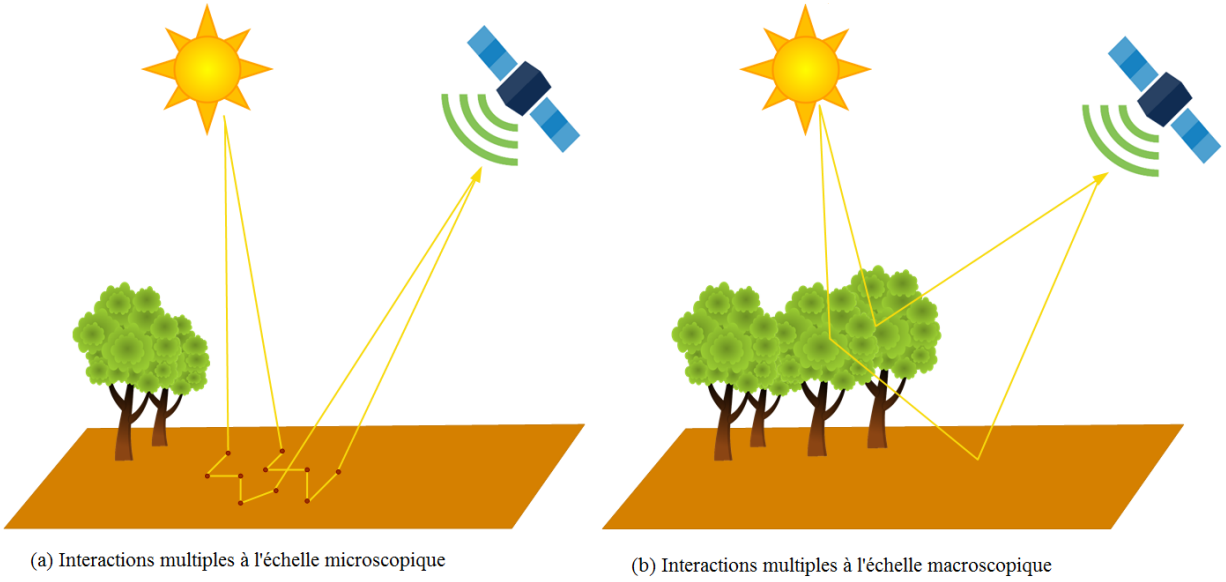
Une méthode NMF qui applique une régularisation de type parcimonie sur les fractions d'abondance a été proposée par Zymnis et al (2007). Dans cette méthode, la norme l_1 est utilisée comme un terme régulateur. Une autre méthode qui se base sur le même principe est $L_{1/2}$-NMF (Qian et al, 2011) qui vise à exploiter la parcimonie des fractions d'abondance en intégrant le terme régulateur $L_{1/2}$.

Pour avoir plus de détails sur les méthodes de démélange linéaires, il est possible de se référer aux revues (Keshava and Mustard, 2002; Bioucas-Dias et al, 2012).

2.5 Modèles non-linéaires pour le démélange hyperspectral

Le modèle linéaire est pertinent dans le cas où les différents composants présents dans le pixel occupent des régions spatialement séparées et aucune interaction entre ces composants n'est présente (Keshava and Mustard, 2002). Cependant, il est possible que les rayons lumineux subissent des réflexions multiples entre les différentes surfaces avant d'atteindre le capteur. Dans ce cas, des modèles de mélange non-linéaires doivent être considérés pour faire face à ces interactions multiples. Il existe deux scénarios justifiant l'utilisation des modèles de mélange non-linéaires pour décrire la complexité des interactions entre les matériaux présents dans la scène. Le premier scénario implique les réflexions multiples de la lumière sur différents matériaux purs à

l'échelle macroscopique. Un tel scénario se produit lorsqu'il existe des structures 3D dans la scène comme des bâtiments de différentes tailles ou des zones forestières. Pour le second scénario, le matériau réfléchissant est considéré comme un mélange intime de composants purs, qui se produit à des échelles microscopiques tel que le cas des grains de sable ayant des compositions différentes. Un exemple schématique de ces deux scénarios est présenté en Figure 2.6.



(a) Interactions multiples à l'échelle microscopique

(b) Interactions multiples à l'échelle macroscopique

Figure 2.6: *Illustration des réflexions multiples*

Cependant, une analyse appropriée de ces mélanges implique une modélisation très complexe, et il est généralement nécessaire de recourir à des modèles approximatifs pour rendre le problème traitable. Plusieurs modèles basés sur la théorie du transfert radiatif ont été proposés, tel que le modèle bidirectionnel proposé par Hapke (1981). Cependant, ce modèle utilise des approximations fortement non-linéaires ce qui le rend difficile à manipuler. En outre, les interactions multiples de la lumière entre les différents composants de la scène conduisent à des effets non-linéaires qui peuvent être pris en compte à l'aide des modèles bilinéaires (Altmann et al, 2011a; Nascimento and Bioucas-Dias, 2009b; Fan et al, 2009; Halimi et al, 2011a). Ces modèles généralisent le modèle linéaire standard en introduisant des termes d'interaction supplémentaires. Ils diffèrent les uns des autres par les contraintes d'additivité imposées aux coefficients

de mélange. Dans (Meganem et al, 2014a), les auteurs ont dérivé un modèle linéaire-quadratique en se basant sur la théorie du transfert radiatif appliquée à une scène urbaine. La principale différence entre le modèle linéaire-quadratique et le modèle bilinéaire est la considération des interactions entre les surfaces de même nature (par exemple sable-sable). L'avantage du modèle linéaire-quadratique est qu'il est basé sur des notions physiques ce qui permet de produire des résultats beaucoup plus précis. En outre, Altmann et al (2012) ont introduit un modèle pôlynomial post-non-linéaire qui a démontré sa capacité à décrire de nombreux effets non-linéaires, en particulier, dans les zones végétales.

Dans ce qui suit, nous fournirons une description détaillée des modèles non-linéaires approximatifs.

2.5.1 Modèle de mélange intime

La première hypothèse assurant la validité des mélanges linéaires est le déroulement du processus de mélange dans un échelle macroscopique. Cependant, dans certaines situations, les interactions peuvent se produire au niveau microscopique. Dans ce cas, les matériaux sont considérés intimement mélangés (Hapke, 2011). En se basant sur la théorie du transfert radiatif, les modèles de mélange intime ont été considérés pour décrire avec précision les interactions subies par la lumière lorsqu'elle rencontre une surface composée de particules, par exemple, dans le cas des scènes composées de sable ou des mélanges minéraux (Nash and Conel, 1974). Une illustration de ces interactions est représentée en Figure 2.6(a).

Les modèles les plus utilisés pour traiter les mélanges intimes sont ceux proposés par Hapke (2011). En effet, ces modèles utilisent des paramètres ayant une signification physique et ils sont capables de relier les mesures spectrales aux caractéristiques physiques des endmembers et leurs fractions d'abondance (associées aux fractions massiques des mélanges intimes). Dans (Hapke, 1981), l'auteur propose un modèle non-linéaire pour exprimer les réflectances mesurées en fonction de différents paramètres du problème, tels que les fractions de masse, les caractéristiques des particules (densité,

taille) et l'albédo de diffusion simple. Par définition, l'albédo correspond au rapport entre la lumière réfléchie par une surface donnée et la lumière incidente. En outre, dans (Draine, 1988) et (Shkuratov et al, 1999), les auteurs ont proposé des modèles de mélange non-linéaires approximatifs pertinents. Cependant, un inconvénient majeur de ces modèles est qu'ils dépendent fortement des paramètres externes liés à l'acquisition, par exemple, ils nécessitent la connaissance exacte de la position géométrique et l'orientation du capteur à l'égard de l'objet observé. Cette dépendance des paramètres externes rend l'étape d'estimation des fractions d'abondance très difficile à mettre en œuvre, surtout dans un scénario non-supervisé lorsque les endmemebers ne sont pas connus. Par conséquent, l'estimation des paramètres de mélange pour ce type de modèle est considérée difficile en pratique.

2.5.2 Modèles bilinéaires pour le démélange hyperspectral

La façon la plus simple de modéliser les réflexions multiples de la lumière est de ne considérer que les réflexions au second ordre entre les matériaux différents ou les interactions bilinéaires. Ce phénomène peut se produire lorsque la lumière diffusée par un matériau donné se réfléchit sur d'autres matériaux avant d'atteindre le capteur. C'est généralement le cas des scènes contenant des zones urbaines ou végétales où des interactions se produisent entre les différentes surfaces. Un exemple de ce genre de scène est montré dans la Figure 2.6 (b).
Mathématiquement, la plupart des modèles bilinéaires (Altmann et al, 2011a) proposés dans la littérature sont exprimés comme suit

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\mathbf{s}_j \odot \mathbf{s}_k + \mathbf{b}_i, \tag{2.2}$$

où $\odot$ est le produit de Hadamard terme à terme défini par

$$\mathbf{s}_j \odot \mathbf{s}_k = \begin{pmatrix} s_j(1) \\ \vdots \\ s_j(N) \end{pmatrix} \odot \begin{pmatrix} s_k(1) \\ \vdots \\ s_k(N) \end{pmatrix} = \begin{pmatrix} s_j(1)s_k(1) \\ \vdots \\ s_j(N)s_k(N) \end{pmatrix}. \tag{2.3}$$

Nous pouvons remarquer que le premier terme de l'équation (2.2) contient la partie linéaire du mélange alors que le deuxième terme présente les interactions non-linéaires qui se produisent entre les matériaux. Le coefficient $a_{j,k}(i)$ permet d'ajuster la quantité de non-linéarité entre les spectres $\mathbf{s}_j$ et $\mathbf{s}_k$ dans le $i^{ème}$ pixel.
Plusieurs modèles bilinéaires ont été proposés en imposant des contraintes sur les coefficients de mélange non-linéaires. Par exemple, Nascimento and Bioucas-Dias (2009b) ont proposé d'imposer les contraintes suivantes

$$\begin{cases} a_j(i) \geq 0 \quad \forall i \ \forall j \\ a_{j,k}(i) \geq 0 \ \ \forall i \ \forall j \neq k \\ \sum_{j=1}^{L} a_j(i) + \sum_{j=1}^{L-1} \sum_{k=j+1}^{L} a_{j,k}(i) = 1 \end{cases} \tag{2.4}$$

Ainsi, le modèle de Nascimento peut être assimilé à un modèle linéaire avec des endmembers virtuels supplémentaires en considérant le produit des spectres $\mathbf{s}_j \odot \mathbf{s}_k$ comme un endmember associé à l'abondance $a_{j,k}(i)$ au lieu de le considérer comme une composante de mélange non-linéaire. Ainsi, le modèle (2.2) en respectant (2.4) peut être re-exprimé comme suit

$$\mathbf{x}_i = \sum_{j=1}^{\tilde{L}} \tilde{a}_j(i)\tilde{\mathbf{s}}_j + \mathbf{b}_i, \tag{2.5}$$

avec

$$\begin{cases} \tilde{a}_j(i) = a_j(i), \tilde{\mathbf{s}}_j = \mathbf{s}_j \quad \forall j = 1, \cdots L \\ \tilde{a}_j(i) = a_{j,k}(i), \tilde{\mathbf{s}}_j = \mathbf{s}_j \odot \mathbf{s}_k \quad \forall j = L+1, \cdots \tilde{L} \end{cases} \tag{2.6}$$

et $\tilde{L} = L(L+1)/2$.
Notons que ce modèle se réduit au modèle de mélange linéaire lorsque les coefficients $\tilde{a}_j(i)$ sont nuls pour $j = L+1, \cdots \tilde{L}$.

Un deuxième modèle bilinéaire est celui proposé par Fan et al (2009). Celui-ci suppose que $a_{j,k}(i) = a_j(i)a_k(i)$, ce qui permet de réécrire le modèle bilinéaire (2.2) comme suit

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L-1} \sum_{k=j+1}^{L} a_j(i)a_k(i)\mathbf{s}_j \odot \mathbf{s}_k + \mathbf{b}_i, \tag{2.7}$$

en respectant les contraintes suivantes

$$\begin{cases} a_j(i) \geq 0 \quad \forall i \; \forall j \\ \sum_{j=1}^{L} a_j(i) = 1 \end{cases} \tag{2.8}$$

Ce modèle considère que si un matériau j est présent dans le mélange, alors la probabilité d'interaction avec lui doit être proportionnelle à son abondance $a_j(i)$. De plus, si un matériau j est absent dans un pixel donné, c'est à dire $a_j(i) = 0$, alors il ne devrait pas y avoir d'effets non-linéaires impliquant ce matériau dans le pixel, c'est-à-dire $a_{j,k}(i) = 0$. Contrairement au modèle de Nascimento, le modèle de Fan ne mène pas à un modèle linéaire.

Un autre modèle plus général est le modèle bilinéaire généralisé (GBM pour son appellation en anglais *Generalized Bilinear Model*) (Halimi et al, 2011a). Celui-ci propose d'ajouter un paramètre libre supplémentaire à chaque interaction bilinéaire tel que $a_{j,k}(i) = \gamma_{j,k}(i) a_j(i) a_k(i)$, avec $\gamma_{j,k}(i) \epsilon [0, 1]$. Le modèle résultant est formulé comme suit

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L-1} \sum_{k=j+1}^{L} \gamma_{j,k}(i) a_j(i) a_k(i) \mathbf{s}_j \odot \mathbf{s}_k + \mathbf{b}_i, \tag{2.9}$$

en respectant les conditions citées dans (2.8).

Le coefficient $\gamma_{j,k}(i)$ permet de contrôler l'interaction non-linéaire entre les deux spectres $\mathbf{s}_j$ et $\mathbf{s}_k$. Tout comme le modèle de Fan, si un endmember est absent dans le pixel alors il n'y a pas d'interaction non-linéaire avec cet endmember. Ce modèle a l'avantage de généraliser à la fois le modèle linéaire obtenu lorsque $\gamma_{j,k}(i) = 0$ et le modèle bilinéaire de Fan obtenu lorsque $\gamma_{j,k}(i) = 1$. Notons que tous ces modèles bilinéaires ne prennent en compte que des interactions entre les spectres différents $\mathbf{s}_j$ et $\mathbf{s}_k$ avec $j \neq k$ mais ils ne considèrent pas les interactions entre $\mathbf{s}_j$ et $\mathbf{s}_j$.

Notons aussi que les modèles bilinéaires présentés dans cette partie pour tenir compte de la non-linéarité ne considèrent que des effets de non-linéarité entre les matériaux présents dans le pixel considéré, cependant, ils ne tiennent pas en considération les interactions avec les matériaux présents dans les pixels voisins.

2.5.3 Modèle linéaire-quadratique pour le démélange hyperspectral

Les modèles bilinéaires ont été introduits pour modéliser les réflexions multiples entre les différents matériaux. Cependant, ces modèles ne tiennent pas compte des réflexions multiples impliquant une seule composante spectrale (par exemple sable-sable). Sur la base d'une analyse physique approfondie d'une scène urbaine, Meganem et al (2014a) ont proposé le modèle de mélange linéaire-quadratique (LQ) suivant

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)\mathbf{s}_j \odot \mathbf{s}_k + \ \mathbf{b}_i, \tag{2.10}$$

en tenant compte des contraintes suivantes

$$\begin{cases} s_j(n) \geq 0 & \\ a_j(i) \geq 0 & \\ \sum_{j=1}^{L} a_j(i) = 1 & \forall \\ a_{j,k}(i) \in [0, 0.5] & \end{cases} \begin{array}{l} 1 \leq i \leq K \\ 1 \leq j \leq L \\ j \leq k \leq L \\ 1 \leq n \leq N \end{array} \tag{2.11}$$

Ce modèle est similaire à la formulation générale des modèles bilinéaires (2.2), avec l'ajout des termes quadratiques $\mathbf{s}_j \odot \mathbf{s}_j$ dans la contribution non-linéaire. En outre, les études physiques établies dans ces travaux ont montré que les coefficients quadratiques $a_{j,k}(i)$ ne dépassent pas la valeur 0.5.

2.5.4 Autres modèles basés sur des approximations physiques

Altmann et al (2012) ont proposé un modèle approximatif capable de décrire une large classe de non-linéarités. La transformation proposée est un polynôme de second degré, conduisant au modèle Pôlynomial Post-Non-linéaire (PPN). Plus précisément, le $i^{ème}$ spectre du pixel observé est défini comme une transformation non-linéaire $g_i(.)$ d'un mélange linéaire des endmembers telle que

$$\mathbf{x}_i = g_i(\sum_{j=1}^{L} a_j(i)\mathbf{s}_j), \tag{2.12}$$

où la fonction non-linéaire $g_i(.)$ est une fonction pôlynomiale de second degré paramétrée par le paramètre de non-linéarité p_i, ce qui permet de réécrire le modèle comme suit

$$\mathbf{x}_i = \mathbf{S}\mathbf{a}_i + p_i(\mathbf{S}\mathbf{a}_i) \odot (\mathbf{S}\mathbf{a}_i), \tag{2.13}$$

où $\mathbf{S} = [\mathbf{s}_1, \cdots, \mathbf{s}_L]$, $\mathbf{a}_i = [a_1(i), \cdots, a_L(i)]^T$ et le paramètre p_i permet d'ajuster la quantité de non-linéarité dans le $i^{ème}$ pixel. L'avantage de ce modèle est sa capacité de se réduire au modèle linéaire ainsi qu'aux modèles bilinéaire et linéaire-quadratique.

Un autre modèle qui combine les deux mélanges Macroscopiques et Microscopiques (MM) est introduit dans (Close et al, 2012a), et il a la formulation suivante

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + a_{L+1}(i) R\Big(\sum_{j=1}^{L} f_j(i)\mathbf{w}_j\Big). \tag{2.14}$$

Ce modèle est composé de deux termes où le premier est similaire à celui rencontré dans le modèle linéaire et provient du processus de mélange macroscopique et le second est considéré comme un endmember additionnel pondéré par la fraction d'abondance a_{L+1}, qui modélise les mélanges intimes en utilisant les albédos de diffusion simple moyens (Hapke, 1981) exprimés en fonction des réflectances par l'intermédiaire d'une fonction de correspondance $R(\cdot)$.

Les méthodes de démélange non-linéaires sont présentées dans la partie suivante.

2.6 Méthodes de démélange non-linéaires

Nous pouvons distinguer deux catégories de méthodes de démélange non-linéaires selon qu'elles se basent sur un modèle de mélange, tel que les modèles présentés dans la section précédente, ou non. Dans la partie suivante, nous présentons quelques méthodes proposées pour ces deux catégories de méthodes.

2.6.1 Méthodes de démélange non-linéaires basées sur des modèles

Lorsque le modèle de mélange non-linéaire est défini, les méthodes de démélange non-linéaires peuvent être supervisées ou non-supervisées en fonction des connaissances a priori sur les endmembers. Dans le cas supervisé, le démélange non-linéaire peut être formulé comme un problème de régression non-linéaire sous-contraintes, et dans le cas non-supervisé, il peut être considéré comme un problème de SAS non-linéaire.

2.6.1.1 Démélange non-linéaire supervisé

Lorsque les endmembers sont connus à l'avance, la plupart des méthodes basées sur un modèle de mélange bilinéaire considèrent le démélange comme un problème de régression non-linéaire. Dans ce cas, le démélange du spectre observé $\mathbf{x}_i$ consiste à résoudre le problème de minimisation suivant

$$\theta_i = argmin\|\mathbf{x}_i - \phi(\mathbf{S}, \theta)\|, \tag{2.15}$$

où θ_i est un vecteur de paramètres contenant les fractions d'abondance ainsi que tous les autres paramètres de non-linéarité et $\phi(.,.)$ est une fonction non-linéaire. Ce problème est délicat à cause de la non-linéarité du modèle d'une part, et d'autre part, puisque certains paramètres doivent satisfaire les contraintes d'additivité et/ou de non-négativité.

Dans le cas du modèle bilinéaire de Nascimento, le problème de démélange peut être interprété comme un problème linéaire avec des endmembers virtuels supplémentaires qui correspondent aux termes croisés. Ainsi, l'estimation des inconnues peut être établie en utilisant une méthode des moindres carrés avec contraintes de non-négativité MCNN (Heinz and Chein-I-Chang, 2001). Cependant, l'estimation des termes croisés est un peu compliquée puisqu'ils devraient être identifiés et ajoutés au dictionnaire des signatures spectrales. Ainsi, si tous les termes croisés sont présents dans le mélange, la taille du dictionnaire va croître d'une façon importante.

Différentes méthodes ont été proposées pour résoudre le problème de démélange pour le modèle GBM. Halimi et al (2011a) ont proposé un algorithme bayésien basé

sur les méthodes MCMC. Le cadre bayésien a l'avantage de traiter facilement les contraintes imposées en définissant des distributions a priori pour les différents paramètres. Cependant, la méthode MCMC a l'inconvénient d'être gourmande en termes de calculs.

Une autre méthode applicable pour les deux modèles bilinéaires de Fan et GBM est proposée par Halimi et al (2011b) en se basant sur (Fan et al, 2009). Celle-ci vise à linéariser une fonction objective non-linéaire en utilisant une série de Taylor du premier ordre. Par la suite, l'ensemble des paramètres est estimé en utilisant un algorithme MCNN (Heinz and Chein-I-Chang, 2001). En outre, la méthode proposée par Halimi et al (2011b) consiste à utiliser un algorithme de descente en gradient, combiné à une méthode de recherche en ligne des pas de mise à jour du gradient. Dans (Yokoya et al, 2014), les auteurs ont proposé une méthode semi-NMF pour le démélange des mélanges GBM.

Concernant le modèle PPN, le démélange peut être réalisé en utilisant les algorithmes bayésiens proposés par Altmann et al (2012).

En ce qui concerne les mélanges intimes, l'approche présentée dans (Hapke, 1981) propose de convertir les mesures de réflectance en albédos de diffusion simple moyens. Étant donné que celui-ci suit un mélange linéaire, les fractions de masse associées à chaque endmember peuvent être estimées en utilisant un algorithme de démélange linéaire classique. Afin d'éviter la tâche de conversion des mesures de réflectance en albédos de diffusion simple moyens, une autre approche consiste à utiliser les réseaux de neurones pour l'apprentissage de la fonction non-linéaire du mélange intime. Dans (Guilfoyle et al, 2001), les auteurs ont utilisé un réseau de neurones à fonctions radiales pour le démélange intime. Par la suite, plusieurs améliorations ont été proposées par Altmann et al (2011b) pour réduire le calcul excessif dans l'étape d'apprentissage. D'autres algorithmes basés sur les réseaux de neurones ont été étudiés dans (Plaza et al, 2004, 2005).

2.6.1.2 Démélange non-linéaire non-supervisé

Les méthodes de démélange supervisées présentées dans la partie précédente supposent que les endmembers sont connus à l'avance. Cependant, dans certains cas, les endmembers doivent être aussi estimés. A cette fin, des méthodes non-linéaires et non-supervisées ont été proposées.

Dans (Gader et al, 2012), les auteurs ont proposé une méthode non-linéaire afin de traiter le modèle bilinéaire de Nascimento, en généralisant l'algorithme SPICE (pour *Sparsity Promoting Iterated Constrained Endmember detection* en anglais) (Zare and Gader, 2007) initialement proposé pour le traitement du modèle de mélange linéaire. Cette méthode alterne de manière itérative entre l'estimation des coefficients de mélange et des endmembers jusqu'à ce que la convergence soit obtenue. L'étape de mise à jour de chaque endmember est réalisée en tenant compte de tous les autres endmembers.

Des méthodes basées sur le principe de la NMF ont été proposées par Yokoya et al (2014) pour traiter le modèle bilinéaire GBM. Dans (Meganem et al, 2014b), des extensions de la méthode NMF ont été proposées pour traiter les mélanges linéaires-quadratiques. Aussi, dans (Eches and Guillaume, 2014), les auteurs ont proposé une méthode de démélange non-supervisée basée sur la méthode NMF pour traiter les mélanges bilinéaires de Fan.

Concernant le modèle PPN, une version non-supervisée de la méthode bayésienne basée sur la méthode MCMC initialement introduite dans (Altmann et al, 2012) a été étudiée dans (Altmann et al, 2014).

Dans (Heylen and Scheunders, 2012), les auteurs proposent une méthode géométrique basée sur le modèle GBM en calculant les distances géodésiques sur la variété [1] ("*manifold*" en anglais) décrite par le modèle GBM. L'idée consiste à identifier le simplexe de volume maximal contenu dans la variété définie par les pixels mélangés selon le modèle GBM.

[1] Par définition, une variété est un espace topologique qui ressemble localement à un espace euclidien n-dimensionnel près de chaque point. Par exemple, les variétés unidimensionnelles incluent les lignes et les cercles, et les variétés bidimensionnelles incluent les surfaces.

Dans (Close et al, 2012a,b), un algorithme non-supervisé a été introduit pour le démélange traitant le modèle MM.

2.6.2 Méthodes de démélange non-linéaires non-basées sur des modèles

Les modèles non-linéaires sous la forme générale ont l'avantage d'être très flexibles pour modéliser les différentes non-linéarités.

Dans ce cadre, des approches non-linéaires basées sur le concept du noyau ont prouvé leur efficacité dans la résolution des problèmes de démélange non-linéaires. L'algorithme KFCLS (pour *Kernel Fully Constrained Least Squares* en anglais) étudié dans (Broadwater et al, 2007; Broadwater and Banerjee, 2009) est une généralisation de l'algorithme FCLS (pour *Fully Constrained Least Squares* en anglais) (Heinz and Chein-I-Chang, 2001). Il a été obtenu principalement en remplaçant chaque produit des endmembers, dans la fonction de coût initial, par une fonction de noyau avec un paramètre réglable pour tenir compte de la non-linéarité. L'idée est de réaliser la correspondance entre le problème de mélange non-linéaire et le problème de mélange linéaire. Cela correspond à ce qui est réalisé pour les mélanges intimes où les mélanges non-linéaires de réflectances deviennent des mélanges linéaires d'albédos de diffusion simple moyens. Ainsi, l'algorithme KFCLS peut être considéré comme une fonction de distorsion non-linéaire appliquée aux endmembers indépendamment de leurs interactions. Cependant, la nature non-linéaire du mélange est causée également par les interactions non-linéaires entre les différents matériaux. Pour vaincre ce problème, dans (Chen et al, 2013), les auteurs ont proposé une méthode ayant une interprétation physique claire. Celle-ci est basée sur la reproduction des noyaux pour modéliser les non-linéarités résultant des interactions entre les endmembers.

Une autre méthode, introduite dans (Altmann et al, 2013), considère une approche de démélange non-supervisée basée sur le concept du noyau. Celle-ci consiste à réaliser une réduction de dimensionnalité non-linéaire en utilisant un modèle GPLVM (pour son appellation en anglais *Gaussian Process Latent Variable Model*). Dans ce cas, une forme particulière du noyau a été envisagée pour l'extension du modèle GBM. La

méthode bayésienne qui en résulte est totalement non-supervisée puisque les endmembers peuvent être conjointement estimés avec les fractions d'abondance.

Une autre stratégie pour traiter les non-linéarités dans les données hyperspectrales est d'effectuer la réduction de dimensionnalité en utilisant une approche géométrique basée sur les techniques d'apprentissage des variétés. L'idée de ces algorithmes consiste à projeter les données de dimension élevée dans une variété de faible dimension. Les méthodes les plus utilisées sont Isomap (pour *Isometric feature mapping* en anglais) (Tenenbaum et al, 2000) et LLE (pour *Locally Linear Embedding* en anglais) (Roweis and Saul, 2000) qui permettent d'appliquer des méthodes linéaires dans la variété considérée. Dans (Bachmann et al, 2005), les auteurs ont étudié le problème d'apprentissage de variété en se basant sur les algorithmes Isomap et LLE afin d'exploiter la structure non-linéaire des images hyperspectrales. En outre, dans (Heylen et al, 2011), les auteurs ont proposé une méthode géométrique de maximisation du volume basée sur les distances géodésiques afin de tenir compte de la non-linéarité des données. Cette méthode est une extension de la méthode N-FINDR initialement proposée pour le cas linéaire, en considérant une variété au lieu du simplexe. Cependant, un inconvénient majeur de la plupart des techniques de réduction de dimensionnalité non-linéaire est leur coût de calcul élevé qui les rendent peu pratiques dans les cas des scènes hyperspectrales de taille importante.

Des approches de démélange non-linéaires basées sur les réseaux de neurones ont été proposées pour établir une relation entrée-sortie entre les données observées et les coefficients de mélange par une méthode de type boîte noire. L'utilisation de ces approches peut rencontrer des difficultés à cause de la non-connaissance du processus et des paramètres de mélange. Dans (Plaza et al, 2004), les auteurs ont conçu une méthode hybride combinant un réseau de neurones de type perceptron multi-couches combiné à un réseau de neurones de Hopfield pour traiter des mélanges non-linéaires. Cependant, ces méthodes supervisées nécessitent l'apprentissage des réseaux de neurones, et ainsi leurs performances dépendent de la qualité des données d'apprentissage. De plus, pour chaque nouvel ensemble de spectres dans la scène, un nouveau réseau de neurones doit être entraîné avant que le démélange soit établi.

Pour des aperçus étendus et détaillés sur les méthodes de démélange non-linéaires, il est possible de se référer aux travaux réalisés dans (Dobigeon et al, 2014; Heylen et al, 2014).

2.7 Conclusions

Ce chapitre fournit un aperçu général du problème de démélange hypespectral qui permet d'identifier les signatures spectrales des composants purs présents dans la scène et leurs contributions correspondantes. Une étape essentielle dans la formulation du problème de démélange consiste à définir correctement le modèle de mélange. En effet, le modèle de mélange linéaire est le modèle le plus utilisé, pour cela, les méthodes linéaires ont été l'objet d'une vaste production scientifique. Cependant, dans certaines situations, le modèle linéaire n'est pas approprié et doit être remplacé par un modèle non-linéaire. Plusieurs modèles non-linéaires approximatifs ont montré leur capacité pour décrire la non-linéarité dans le mélange.
La résolution du problème de démélange spectral dans le cas non-supervisé peut être considérée comme un problème de séparation aveugle de sources. Dans le cadre de cette thèse, nous visons à proposer de nouvelles méthodes de SAS pour le démélange spectral du modèle de mélange linéaire-quadratique. La pertinence de ce modèle pour le démélange hyperspectral en zone urbaine a été montrée à l'aide d'études physique approfondies dans (Meganem et al, 2014a).

Chapitre 3

Une méthode d'ACPa pour le démélange hyperspectral des mélanges linéaires-quadratiques

Sommaire

3.1 Introduction 61
3.2 Modèle de mélange linéaire-quadratique 62
3.3 Méthode non supervisée de démélange hyperspectral . . . 63
3.3.1 Estimation du nombre de matériaux purs 63
3.3.2 Estimation des spectres des matériaux purs 65
3.3.2.1 Estimation par détection des zones mono-sources . 65
3.3.2.2 Estimation par détection des paires de pixels purs 68
3.3.3 Clustering 69
3.3.4 Estimation des coefficients d'abondance et quadratiques . . 70
3.4 Étude expérimentale 72
3.4.1 Critère de performance 72
3.4.2 Expériences réalisées et résultats 73
3.4.2.1 Expérience 1 74
3.4.2.2 Expérience 2 76
3.4.2.3 Expérience 3 78
3.5 Conclusions 84

3.1 Introduction

Dans ce chapitre, nous proposons une nouvelle méthode de séparation aveugle de sources, adaptée au modèle de mélange linéaire-quadratique pour le démélange des images hyperspectrales urbaines. L'objectif de cette méthode est d'extraire, à partir du spectre observé dans chaque pixel de l'image, l'ensemble des spectres de matériaux purs, appelés aussi *endmembers*, ainsi que leurs fractions d'abondance et les coefficients quadratiques associés aux produits des spectres de ces matériaux purs.

Étant basée sur la méthode d'Analyse en Composantes Parcimonieuses (ACPa), l'approche proposée nécessite l'existence d'une zone mono-source ou d'au moins deux pixels purs, par matériau pur dans la scène observée. Elle estime d'abord le nombre de matériaux purs présents dans l'image, puis extrait leurs spectres en détectant soit les zones mono-sources soit les pixels purs présents dans l'image. Enfin, elle estime les fractions d'abondance et les coefficients quadratiques dans chaque pixel en utilisant un algorithme des moindres carrés sous la contrainte de non-négativité.

Le chapitre est structuré comme suit. Nous commençons par un rappel sur le problème de démélange hyperspectral en utilisant le modèle linéaire-quadratique. Par la suite, nous fournissons une description détaillée de la méthode proposée. Enfin, nous présentons les résultats de simulation obtenus.

3.2 Modèle de mélange linéaire-quadratique

Rappelons que dans un problème de démélange hyperspectral, nous disposons de N images, observées dans N bandes spectrales distinctes. Chaque image est composée de K pixels, qui peuvent être des pixels purs contenant un seul matériau pur, ou des mixels issus d'un mélange de plusieurs matériaux purs. Dans (Meganem et al, 2014a), il a été montré que dans un milieu urbain, le spectre d'un mixel peut être modélisé par un mélange linéaire-quadratique de L spectres de matériaux purs, selon l'équation de mélange suivante :

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)\mathbf{s}_j \odot \mathbf{s}_k + \mathbf{b}_i, \tag{3.1}$$

avec

- $\mathbf{x}_i = [x_i(1), ..., x_i(N)]^T$ le spectre associé au pixel i,
- $\mathbf{s}_j = [s_j(1), ..., s_j(N)]^T$ le spectre du matériau pur j,
- $a_j(i)$ la fraction d'abondance du matériau pur j dans le pixel i,
- $a_{j,k}(i)$ le coefficient quadratique du matériau j sur le matériau k dans le pixel i,
- le symbole $\odot$ représente le produit de Hadamard terme à terme de deux vecteurs,
- $\mathbf{b}_i$ est un bruit additif présent sur le pixel i, généralement supposé gaussien, de moyenne nulle, spatialement et spectralement i.i.d.. Notons que l'effet du bruit n'a pas été pris en compte dans (Meganem et al, 2014a).

De plus, il a été montré dans (Meganem et al, 2014a) que les contraintes suivantes doivent être respectées :

$$\begin{cases} s_j(n) \geq 0 \\ a_j(i) \geq 0 \\ \sum_{j=1}^{L} a_j(i) = 1 \\ a_{j,k}(i) \in [0, 0.5] \end{cases} \quad \forall \begin{cases} 1 \leq i \leq K \\ 1 \leq j \leq k \leq L \\ 1 \leq n \leq N \end{cases} \tag{3.2}$$

Dans la suite, toutes les *vraies sources* $\mathbf{s}_j$ et les *pseudo-sources* $\mathbf{s}_j \odot \mathbf{s}_k$, $\forall j \in \{1, ..., L\}$, $\forall k \in \{j, ..., L\}$, sont appelées les *sources étendues*.

A titre d'exemple, dans le cas où deux matériaux purs sont présents dans un mixel, le modèle de mélange est comme suit

$$\mathbf{x}_i = a_1(i)\mathbf{s}_1 + a_2(i)\mathbf{s}_2 + a_{1,1}(i)\mathbf{s}_1 \odot \mathbf{s}_1 + a_{2,2}(i)\mathbf{s}_2 \odot \mathbf{s}_2 + a_{1,2}(i)\mathbf{s}_1 \odot \mathbf{s}_2 + \mathbf{b}_i \,. \quad (3.3)$$

Dans ce modèle, il y a 5 sources étendues : $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_1 \odot \mathbf{s}_1$, $\mathbf{s}_2 \odot \mathbf{s}_2$, $\mathbf{s}_1 \odot \mathbf{s}_2$, mais seulement 2 endmembers (vraies sources) : $\mathbf{s}_1$ et $\mathbf{s}_2$.

3.3 Méthode non supervisée de démélange hyperspectral

La méthode de démélange hyperspectral, proposée dans ce chapitre, vise à estimer les spectres $\mathbf{s}_j$ des matériaux purs présents dans les mixels, ainsi que les fractions d'abondance $a_j(i)$ et les coefficients quadratiques $a_{j,k}(i)$ associés à ces matériaux. Elle opère en quatre étapes présentées dans la Figure 3.1.

Les détails de chaque étape de la méthode sont mentionnés dans les sous-sections suivantes.

3.3.1 Estimation du nombre de matériaux purs

La première étape de cette méthode consiste à estimer le nombre de matériaux purs présents dans la scène observée. Dans le cas le plus simple, l'estimation de ce nombre peut être effectuée, d'une manière supervisée, par les experts du domaine lorsqu'il s'agit d'une image à haute résolution spatiale ou lorsqu'il y a des informations a priori sur les matériaux purs qui peuvent être présents dans certaines zones urbaines. Si ce n'est pas le cas, nous pouvons utiliser une méthode basée sur les valeurs propres de la matrice de covariance des données. Cette méthode nécessite la présence de tous les termes quadratiques, $\mathbf{s}_i \odot \mathbf{s}_j$ et $\mathbf{s}_i \odot \mathbf{s}_i$, dans le mélange.

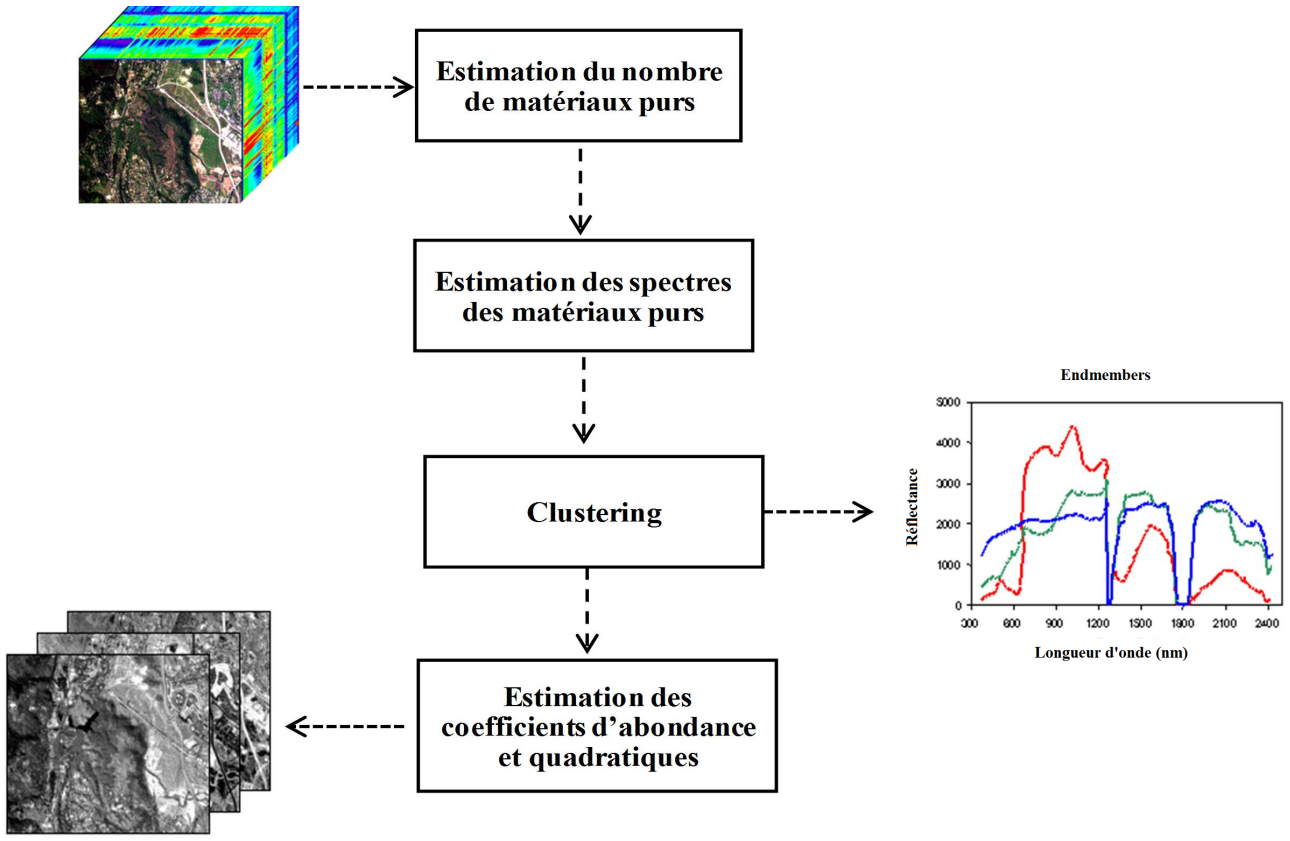


Figure 3.1: *L'organigramme de la méthode proposée*

Soit $\mathbf{X}$ une matrice d'observation de taille $N \times K$, chaque colonne de cette matrice correspondant au spectre observé dans un pixel de l'image, et soit $\mathbf{C}_X$ la matrice de covariance de $\mathbf{X}$, de taille $K \times K$, dont l'élément (i, j) représente la covariance des colonnes i et j de $\mathbf{X}$.

Dans le cas où les données sont non bruitées, le rang P de $\mathbf{C}_X$ correspond au nombre de sources étendues du modèle de mélange linéaire-quadratique, ce qui permet de déduire L, le nombre de matériaux purs présents dans l'image, grâce à une correspondance entre le nombre de sources et le nombre de sources étendues. En effet, la relation entre le nombre des sources étendues P et le nombre des vraies sources L s'écrit de la manière suivante

$$P = \frac{L(L+3)}{2} \ . \tag{3.4}$$

Ainsi, déterminer L revient à résoudre l'équation du second degré suivante

$$L^2 + 3L - 2P = 0 \ . \tag{3.5}$$

La résolution de (3.5), nous a permis de calculer L en fonction de P comme suit :

$$L = \frac{\sqrt{9+8P}-3}{2} \ . \tag{3.6}$$

Par exemple, dans le cas particulier du modèle (3.3), le rang de la matrice de covariance sera égal à 5 ce qui correspond à 5 sources étendues, et donc $L = \frac{\sqrt{9+8\times5}-3}{2} = 2$ matériaux purs.

Lorsque les données sont corrompues par un bruit i.i.d., la matrice de covariance a $K-P$ valeurs propres égales à la variance du bruit σ^2 et P valeurs propres, supérieures à σ^2, qui correspondent au signal. Pour obtenir une estimation robuste du nombre de composantes P, nous avons utilisé l'une des méthodes proposées dans (Luo and Zhang, 2000), et appelées EGM (*Eigenvalue Grads Methods*). Dans cette méthode, les K valeurs propres $\lambda(i)$ sont d'abord triées par ordre décroissant. Après cela, nous calculons le rapport logarithmique suivant :

$$\delta(i) = \ln\left(\frac{\lambda(i)}{\lambda(i+1)}\right) \qquad \forall 1 \leq i \leq (K-1). \tag{3.7}$$

Supposons que $\delta(i)$ est inférieur à un seuil donné pour $i > M$ et supérieur à ce seuil pour $i = M$. La valeur de M est alors considérée comme notre estimation du nombre de composantes P. La valeur de ce seuil est déterminée en utilisant des tests empiriques.

3.3.2 Estimation des spectres des matériaux purs

Dans cette étape, nous nous intéressons à l'estimation des spectres des matériaux purs présents dans la scène observée. Ceci a été réalisé en utilisant deux méthodes dont les détails sont fournis dans les paragraphes suivantes.

3.3.2.1 Estimation par détection des zones mono-sources

Cette première méthode, inspirée de (Deville and Hosseini, 2007b; Meganem et al, 2010; Karoui, 2012) est basée sur l'hypothèse d'existence des zones *mono-sources* dans la scène observée. Les définitions et hypothèses utilisées sont comme suit

- **Définition 1** : une *"zone d'analyse"* Ω est une petite zone spatiale formée de pixels adjacents.

- **Définition 2** : une zone d'analyse est appelée *mono-source* si elle ne contient qu'un seul matériau pur dont la fraction d'abondance vaut 1 alors que toutes les autres fractions d'abondance et tous les coefficients quadratiques sont nuls. En d'autres termes, dans cette zone, le seul matériau présent n'interagit pas avec les matériaux présents dans les pixels voisins.

- **Hypothèse 1** : dans la scène observée, il existe au moins une zone mono-source pour chaque matériau pur présent dans la scène.

L'hypothèse de parcimonie spatiale ci-dessus est réaliste pour des images à haute résolution spatiale, contenant des zones étendues pour les différentes classes d'occupation de sol d'intérêt (Karoui, 2012).

La détection des zones *mono-sources* est effectuée en balayant le domaine spatial par des *zones d'analyse* qui peuvent être adjacentes ou recouvrantes.
Notons par $\mathbf{x}_\Omega(k)$ le vecteur colonne contenant les observations associées à tous les pixels de la zone d'analyse Ω, dans la bande spectrale k. Dans chaque zone d'analyse, les coefficients d'intercorrélation $\rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q))$ entre les signaux observés dans les bandes spectrales p et q sont calculés de la manière suivante

$$\rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q)) = \frac{< \mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q) >}{\| \mathbf{x}_\Omega(p) \| \| \mathbf{x}_\Omega(q) \|}, \ \forall \ p, q = 1, \cdots N \ et \ \ p > q, \tag{3.8}$$

où $< ., . >$ désigne le produit scalaire et $\|.\|$ la norme vectorielle.

Si la zone d'analyse est une zone mono-source correspondant au matériau pur j, les vecteurs $\mathbf{x}_\Omega(p)$ et $\mathbf{x}_\Omega(q)$ sont deux vecteurs constants tels que

$$\mathbf{x}_\Omega(p) = s_j(p) \odot \overrightarrow{1} \ , \tag{3.9}$$

$$\mathbf{x}_\Omega(q) = s_j(q) \odot \overrightarrow{1} \ , \tag{3.10}$$

où $\overrightarrow{1}$ est un vecteur colonne de même dimension que le vecteur $\mathbf{x}_\Omega(p)$. Il s'en suit que

$$\mid \rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q)) \mid = 1, \ \forall \ p, q = 1, \cdots N \ et \ \ p > q, \tag{3.11}$$

dans les zones *mono-sources*.

Par ailleurs, comme il a été montré dans (Karoui, 2012) pour le cas des mélanges linéaires, si la zone d'analyse n'est pas mono-source, alors $\mid \rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q) \mid < 1$ sauf dans des cas très particuliers et donc peu probables, par exemple quand les fractions d'abondance de deux sources sont constants et non-nuls sur tous les pixels de la zone d'analyse. En se basant sur le même principe pour le cas des mélanges linéaires-quadratiques, nous considérons que si la zone d'analyse n'est pas mono-source, alors $\mid \rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q) \mid < 1$ sauf dans des cas très spécifiques et rares lorsque les coefficients d'abondance et les coefficients quadratiques des matériaux présents dans la zone d'analyse sont constants et non-nuls sur tous les pixels.

Pour obtenir un critère unique à partir de tous les coefficients d'intercorrélation correspondant aux différentes bandes spectrales p et q, nous choisissons le pire des cas, défini par le paramètre suivant

$$min\Big(\mid \rho(\mathbf{x}_\Omega(p), \mathbf{x}_\Omega(q)) \mid \Big), \ \forall \ p, q = 1, \cdots N \ et \ \ p > q. \tag{3.12}$$

Dans le cas idéal, la valeur de ce paramètre (3.12) doit être égale à 1 afin de considérer que la *zone d'analyse* étudiée est une zone *mono-source*. Cependant, en pratique, ce paramètre doit être supérieur à un seuil proche de 1, déterminé à l'aide de tests empiriques.

Après la détection des zones mono-sources dans tout le domaine spatial, il devient possible d'estimer les spectres des matériaux purs dans chacune de ces zones. Théoriquement, dans une zone mono-source, il n'existe qu'une seule fraction

d'abondance, égale à 1, alors que toutes les fraction d'abondance des autres matériaux et tous les coefficients quadratiques du mélange LQ sont nuls.

Ainsi, en utilisant l'équation (3.1) et le fait que la somme des fractions d'abondance est égale à un, chaque observation associée à chaque pixel de la zone *mono-source* détectée doit être égale à la valeur de la refléctance du matériau pur présent dans cette zone. Néanmoins, en pratique à cause du bruit ou d'autres sources d'erreur, nous ne trouvons pas exactement la même valeur dans tous les pixels de la zone mono-source, d'où l'idée de calculer la médiane des valeurs de tous les pixels présents dans la zone mono-source pour chaque bande spectrale, suggérée dans (Karoui, 2012). Cette idée permet d'obtenir une estimation plus précise de la valeur de la réflectance du matériau pur dans chaque bande spectrale. Autrement dit, chaque zone mono-source nous permet d'obtenir une estimation du vecteur $\mathbf{s}_j$ correspondant au spectre du matériau pur j de la manière suivante

$$\hat{\mathbf{s}}_j = \begin{pmatrix} médiane(\mathbf{x}_\Omega(1)) \\ médiane(\mathbf{x}_\Omega(2)) \\ \vdots \\ médiane(\mathbf{x}_\Omega(N)) \end{pmatrix}. \tag{3.13}$$

3.3.2.2 Estimation par détection des paires de pixels purs

La méthode présentée ci-dessus nécessite l'existence des zones mono-sources ((3×3) ou (4×4) au minimum pour avoir une estimation fiable des coefficients de corrélation). Dans la suite, nous proposons une autre méthode moins exigeante qui nécessite l'existence de 2 pixels purs par matériau pur dans l'image. Cette condition est généralement vérifiée lorsque la résolution spatiale de l'image n'est pas très faible. Sur la base de cette hypothèse, nous comparons, deux à deux, les spectres de tous les pixels afin d'identifier les pixels purs existants contenant le même matériau pur. La comparaison peut être effectuée en calculant les distances euclidiennes entre les

spectres des pixels, définie par :

$$D(i,j) = \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_F \quad \forall\ i,j = 1,\cdots K\ et\ j > i, \tag{3.14}$$

où $\left\| . \right\|_F$ représente la norme de Frobenius.

Pour des pixels purs composés du même matériau, cette distance est nulle dans le cas idéal, et suffisamment petite, et donc inférieure à un certain seuil, en présence d'un bruit faible.
Il est clair que cette condition est une condition nécessaire pour que les pixels détectés soient considérés comme pixels purs associés au même matériau. La suffisance de cette condition en tenant compte des contraintes physiques n'a pas été prise en compte dans cette thèse, mais mérite d'être étudiée ultérieurement.

3.3.3 Clustering

L'étape précédente nous fournit des estimations des spectres des matériaux purs présents dans l'image. Cependant, il est possible que l'image contienne plusieurs zones mono-sources (dans la première approche) ou plusieurs paires de pixels purs (dans la deuxième approche) contenant le même matériau pur. Ainsi, cette étape risque de nous fournir de nombreuses estimations des spectres de matériaux purs. Par conséquent, les spectres des matériaux purs détectés à l'étape précédente doivent être groupés, en fonction de leur similarité, dans L groupes, correspondant aux L matériaux purs recherchés. Pour y arriver, nous avons recours à une méthode de classification non supervisée ou "clustering". En effet, le clustering vise à regrouper un ensemble de données hétérogènes en différents groupes homogènes, appelés "*clusters*", en se basant sur le calcul d'une distance ou d'une mesure de similarité, ce qui nous permet de classer les spectres sélectionnés précédemment en L classes selon le critère de similarité mesuré.

Les algorithmes de clustering les plus utilisés sont l'algorithme k-means (Macqueen, 1967) et ses variantes telles que k-médianes (Dohan et al, 2015) et k-médoïdes (Kaufman and Rousseeuw, 1987). Dans cette étude, nous utilisons l'algorithme k-moyennes, qui s'est avéré assez performant pour déterminer les différentes classes des spectres

purs. Le principe de cet algorithme est le suivant : tout d'abord, il choisit k centres initiaux, où k représente le nombre de *clusters* et est fixé par l'utilisateur. Le paramètre k correspond, dans notre problème, à L, le nombre des matériaux purs estimé précédemment. Ensuite, chaque point des données, c-à-d dans notre cas chaque spectre de matériau pur estimé, est assigné au *cluster* dont le centre est le plus proche du point examiné. Par la suite, le centre de chaque *cluster* est recalculé par la moyenne des points appartenant au même *cluster* selon la nouvelle partition. Ces deux dernières étapes sont répétées jusqu'à la convergence de l'algorithme vers une partition stable, autrement dit, jusqu'à ce que les centres des *clusters* ne changent plus.

3.3.4 Estimation des coefficients d'abondance et quadratiques

Une fois que nous avons extrait les signatures spectrales des matériaux purs, il nous reste à estimer leurs fractions d'abondance et les coefficients quadratiques. Comme mentionné dans l'équation (3.2), les fractions d'abondance et les coefficients quadratiques sont non-négatifs et la somme des fractions d'abondance doit être égale à un dans chaque pixel. Donc, pour les extraire, nous avons besoin d'une méthode qui prend en compte la non-négativité des données. À cette fin, nous avons utilisé la méthode des Moindres Carrés avec contrainte de Non-Négativité (MCNN) (Lawson and Hanson, 1995) appliquée à chaque pixel i de l'image et nous l'avons adaptée afin de tenir compte de la somme égale à un des fractions d'abondance.

Étant donnés une matrice $\hat{\mathbf{B}}$ et un vecteur colonne $\mathbf{d}$, la méthode MCNN vise à trouver un vecteur non négatif $\hat{\mathbf{z}}$ de manière à minimiser

$$\left\| \hat{\mathbf{B}}\hat{\mathbf{z}} - \mathbf{d} \right\|_F . \tag{3.15}$$

Pour un modèle de mélange linéaire, les entrées de la méthode MCNN sont une matrice $\hat{\mathbf{B}}$, de taille $(N \times L)$, contenant les estimés des L spectres des matériaux purs $\hat{\mathbf{s}}_j$, et un vecteur colonne $\mathbf{d}$, de taille N, contenant le spectre observé $\mathbf{x}_i$ associé au pixel i. La sortie de la méthode est un vecteur colonne $\hat{\mathbf{z}}$ de taille L contenant les estimées des fractions d'abondance $a_j(i)$, $\forall j = 1, \cdots, L$, des L matériaux purs présents dans le pixel i.

Nous allons dans la suite appliquer cette méthode pour estimer les fractions d'abondance et les coefficients quadratiques du mélange LQ. Pour cela, nous réécrivons le modèle LQ présenté dans l'équation (3.1) en utilisant la notation matricielle suivante :

$$\mathbf{Bz} = \mathbf{d}, \tag{3.16}$$

où $\mathbf{B}$ est une matrice, de taille $(N \times P)$, contenant les vecteurs de l'ensemble des sources étendues, définie comme suit

$$\mathbf{B} = [\mathbf{s}_1 \;\; \mathbf{s}_2 \;\cdots\; \mathbf{s}_L \;\; \mathbf{s}_1 \odot \mathbf{s}_1 \;\; \mathbf{s}_1 \odot \mathbf{s}_2 \;\cdots\; \mathbf{s}_L \odot \mathbf{s}_L], \tag{3.17}$$

$\mathbf{z}$ est un vecteur colonne de P éléments, contenant tous les fractions d'abondance et les coefficients quadratiques présents dans le pixel i, défini comme :

$$\mathbf{z} = [a_1(i) \;\; a_2(i) \;\cdots\; a_L(i) \;\; a_{1,1}(i) \;\; a_{1,2}(i) \;\cdots\; a_{L,L}(i)]^T, \tag{3.18}$$

et $\mathbf{d}$ est un vecteur colonne de N éléments contenant le spectre observé $\mathbf{x}_i$ du pixel i.

En utilisant cette nouvelle formulation, et avec la connaissance de la matrice $\hat{\mathbf{B}}$ l'estimée de la matrice $\mathbf{B}$ (grâce aux estimations de $\mathbf{s}_1, \cdots, \mathbf{s}_L$ obtenues dans les étapes précédentes) et du vecteur $\mathbf{d}$, nous pouvons appliquer la méthode MCNN afin de minimiser la fonction de coût (3.15), et donc déterminer le vecteur $\hat{\mathbf{z}}$ contenant les estimés des fractions d'abondance $a_j(i)$ et des coefficients quadratiques $a_{j,k}(i)$ associés au pixel i.

D'un autre côté, la méthode MCNN doit être adaptée afin de respecter la contrainte de somme égale à un des fractions d'abondance, mentionnée dans l'équation (3.2). Pour cela, nous avons utilisé l'idée suggérée dans (Heinz and Chein-I-Chang, 2001) pour les mélanges linéaires et nous l'avons adaptée au mélanges linéaires-quadratiques dans notre algorithme. Nous avons procédé comme suit : nous avons ajouté une dernière ligne à la matrice $\hat{\mathbf{B}}$, dont les L premiers éléments sont égaux à une valeur constante μ et les autres éléments sont nuls, de cette manière

$$\hat{\mathbf{B}} = \begin{pmatrix} \hat{\mathbf{s}}_1 & \hat{\mathbf{s}}_2 & \cdots & \hat{\mathbf{s}}_L & \hat{\mathbf{s}}_1 \odot \hat{\mathbf{s}}_1 & \hat{\mathbf{s}}_1 \odot \hat{\mathbf{s}}_2 & \cdots & \hat{\mathbf{s}}_L \odot \hat{\mathbf{s}}_L \\ \mu & \mu & \cdots & \mu & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{3.19}$$

et nous avons ajouté, au vecteur colonne $\mathbf{d}$, un dernier élément égal à μ comme suit

$$\mathbf{d} = \begin{pmatrix} \mathbf{x}_i \\ \mu \end{pmatrix}. \tag{3.20}$$

En effectuant ces modifications sur la matrice $\hat{\mathbf{B}}$ et le vecteur $\mathbf{d}$, la minimisation de la fonction de coût (3.15) permet le respect de la contrainte de somme égale à un des fractions d'abondance. Les détails sur le choix de μ sont mentionnés dans (Heinz and Chein-I-Chang, 2001).

3.4 Étude expérimentale

Dans cette section, nous présentons les résultats de simulation obtenus par l'approche proposée. Elle est organisée comme suit : tout d'abord, nous présentons le critère de performance utilisé pour l'évaluation de la qualité des résultats obtenus, puis nous détaillons les expériences réalisées avec les différents types de données.

3.4.1 Critère de performance

Afin d'évaluer les performances de la méthode proposée, nous avons calculé l'Erreur Quadratique Moyenne Normalisée (EQMN). Comme indiqué par son appellation, ce critère permet de mesurer la moyenne du carré de l'erreur entre les valeurs réelles et les valeurs estimées, normalisée par la moyenne du carré des valeurs réelles. Les EQMN entre les valeurs réelles des spectres des matériaux purs, des fractions d'abondance et des coefficients quadratiques, et leurs valeurs estimées, sont données par

$$EQMN_{s_j} = \frac{\sum_{n=1}^{N}(s_j(n) - \hat{s}_j(n))^2}{\sum_{n=1}^{N} s_j(n)^2}, \qquad j = 1, \cdots, L \tag{3.21}$$

$$EQMN_{a_j} = \frac{\sum_{i=1}^{K}(a_j(i) - \hat{a}_j(i))^2}{\sum_{i=1}^{K} a_j(i)^2}, \qquad j = 1, \cdots, L \tag{3.22}$$

$$EQMN_{a_{j,k}} = \frac{\sum_{i=1}^{K}(a_{j,k}(i) - \hat{a}_{j,k}(i))^2}{\sum_{i=1}^{K} a_{j,k}(i)^2}, \qquad j, k = 1, \cdots, L \text{ et } k \geq j \tag{3.23}$$

Dans ces formules, " ˆ " se réfère aux valeurs estimées aprés suppression des indéterminations de permutation.

Afin de réduire la quantité des résultats rapportés, nous calculons l'EQMN moyenne sur toutes les sources de la manière suivante

$$EQMN_s = \frac{1}{L}\sum_{j=1}^{L} EQMN_{s_j}, \tag{3.24}$$

où L est le nombre de sources.

De même, nous calculons l'EQMN moyenne sur tous les coefficients de mélange (fractions d'abondance et coefficients quadratiques) :

$$EQMN_a = \frac{1}{P}(\sum_{j=1}^{L} EQMN_{a_j} + \sum_{j=1}^{L}\sum_{k=j}^{L} EQMN_{a_{j,k}}). \tag{3.25}$$

où P est le nombre total des fractions d'abondance et des coefficients quadratiques.

3.4.2 Expériences réalisées et résultats

Dans les expériences effectuées, nous avons considéré les données suivantes :

- **Expérience 1 :** les observations sont des mélanges artificiels et non-bruités générés à partir de spectres de réflectance réels et coefficients de mélange aléatoires mais réalistes.

- **Expérience 2 :** les observations sont des mélanges artificiels bruités générés à partir de spectres de réflectance réels et coefficients de mélange aléatoires mais réalistes.

- **Expérience 3 :** Les observations sont des mélanges artificiels, bruités et non bruités, générés à partir de spectres de réflectance réels et cartes de classification réelles.

Les détails des tests de simulation sont montrés dans les paragraphes suivants.

3.4.2.1 Expérience 1

Dans cette expérience, nous avons voulu tester notre méthode dans le cas le plus simple où les observations sont des mélanges artificiels de spectres réels sans présence de bruit dans les mélanges générés. Les spectres utilisés dans cette expérience ont été puisés d'une bibliothèque spectrale de neuf matériaux présents dans les zones urbaines, chacun mesuré pour 126 longueurs d'onde de 0,4 à 2,5 μm[1]. Ces matériaux sont : asphalte, béton, brique, gravier, trottoir, tuile, pelouse, terre et aluminium. Les spectres de ces matériaux sont présentés dans la Figure 3.2.

Selon l'équation (3.1), nous avons utilisé cet ensemble de spectres pour générer 126 images synthétiques chacune associée à une bande spectrale. Chaque image est formée de 50×50 pixels. Les fractions d'abondance et les coefficients quadratiques utilisés pour produire ces mélanges ont été créés par un processus aléatoire mais réaliste, en respectant les contraintes physiques citées dans l'équation (3.2). Plus précisément, les fractions d'abondance et les coefficients quadratiques sont générés par des valeurs aléatoires uniformément réparties dans les intervalles $[0, 1]$ et $[0, 0.5]$, respectivement, et la somme des fractions d'abondance sur chaque pixel est égale à 1.

Afin d'assurer la présence des pixels purs dans l'image, ces pixels ont été générés dans des positions aléatoires : la valeur de la fraction d'abondance d'un matériau pur a été mise à 1, et les autres fractions d'abondance et coefficients quadratiques ont été mis à zéro dans un tel pixel. Le pourcentage de présence des pixels purs par matériau est égal au moins à 6% et au plus à 10%.

[1]Base de données Memoires (http://www.onera.fr/data/memoires).

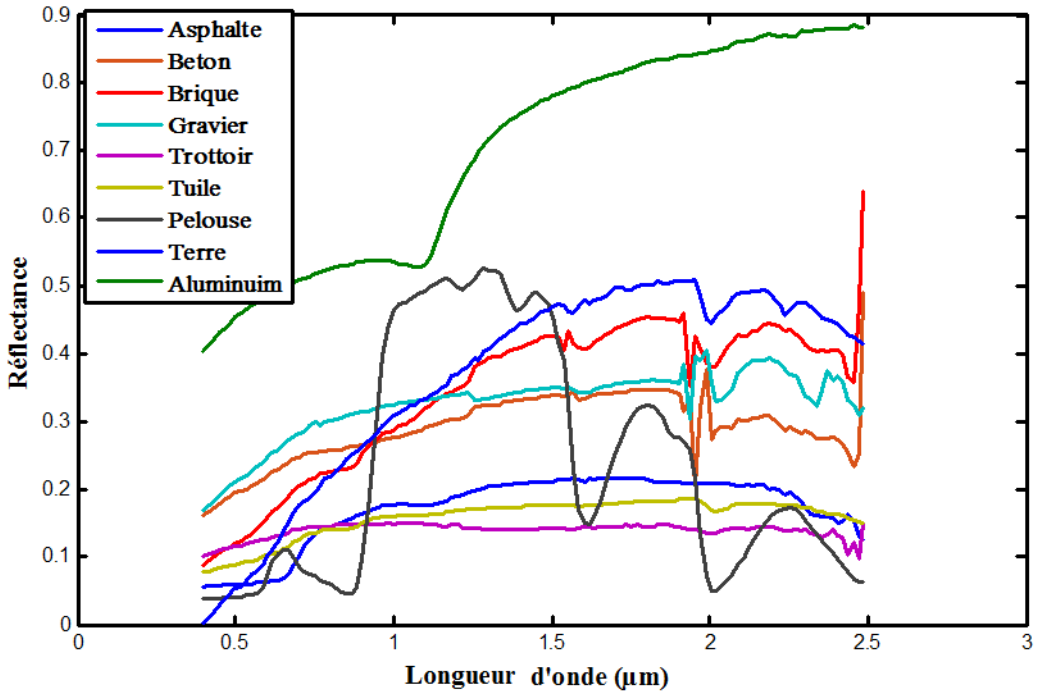


Figure 3.2: *Spectres des matériaux purs*

Tableau 3.1: *EQMN pour un mélange de 2 à 5 matériaux purs*

Moyenne EQMN \ Nombre de matériaux	**2**	**3**	**4**	**5**
Spectres des matériaux purs	$3.27\ 10^{-32}$	$7.98\ 10^{-32}$	$8.66\ 10^{-32}$	$9.95\ 10^{-32}$
Coefficients d'abondance/quadratiques	$3.95\ 10^{-26}$	$2.10\ 10^{-24}$	$2.71\ 10^{-13}$	$1.00\ 10^{-11}$

Dans cette expérience, nous avons effectué plusieurs tests en variant, à chaque fois, le nombre et la nature des matériaux purs mélangés. Dans chaque test, nous avons effectué une dizaine de simulations et, à chaque simulation, les fractions d'abondance et les coefficients quadratiques ont été modifiés et les spectres des matériaux purs ont été choisis aléatoirement parmi les 9 spectres présentés ci-dessus. L'estimation des spectres des matériaux purs a été réalisée selon la méthode détaillée dans la Section 3.3.2.2.

Le tableau 3.1 montre les valeurs moyennes des EQMN calculées pour 4 tests réalisés correspondant aux mélanges de 2 à 5 matériaux purs.

Nous pouvons remarquer que les valeurs obtenues pour l'EQMN des spectres de matériaux purs sont presque nulles (environ 10^{-32}). De même, les EQMN correspondant aux fractions d'abondance et aux coefficients quadratiques sont très petites (inférieures à 10^{-11}). Par conséquent, nous avons réussi à extraire les spectres de matériaux purs et leurs contributions relatives, ce qui confirme les bonnes performances de la méthode proposée lors de l'utilisation des images synthétiques non bruitées et contenant des pixels purs.

3.4.2.2 Expérience 2

Dans cette deuxième expérience, notre but est de tester notre méthode dans le cas où les observations sont des mélanges artificiels de spectres réels avec la présence de bruit dans les mélanges. Pour cela, nous avons d'abord utilisé la procédure décrite dans l'Expérience 1 pour la génération des images synthétiques. Nous avons ensuite ajouté à ces images un bruit gaussien et i.i.d..

Dans les premières simulations effectuées, notre objectif est l'évaluation de la robustesse de la méthode face à l'augmentation du taux de bruit dans les images générées. En ajoutant progressivement du bruit, nous avons calculé le Rapport Signal sur Bruit (RSB) de la manière suivante :

$$RSB = 10 \times \log_{10} \frac{\sum_{n=1}^{N} \sum_{i=1}^{K} x_i(n)^2}{\sum_{n=1}^{N} \sum_{i=1}^{K} b_i(n)^2} . \tag{3.26}$$

Pour chaque niveau de bruit, nous avons effectué une dizaine de simulations et, à chaque simulation, les fractions d'abondance et les coefficients quadratiques ont été modifiés et les spectres des matériaux purs ont été sélectionnés aléatoirement. Le nombre de matériaux purs utilisé dans chaque simulation a été fixé à 2.

Le Tableau 3.2 montre les valeurs moyennes des EQMN en variant le RSB de 80 dB à 30 dB.

Tableau 3.2: *EQMN pour un mélange de 2 matériaux purs avec un RSB variant de 30 à 80 dB*

RSB (dB) / Moyenne EQMN	80	60	50	40	30
Spectres des matériaux purs	$1.62\ 10^{-11}$	$1.07\ 10^{-9}$	$1.97\ 10^{-8}$	$2.58\ 10^{-7}$	$7.9\ 10^{-4}$
Coefficients d'abondance/quadratiques	$7.99\ 10^{-5}$	$1.77\ 10^{-4}$	$7.24\ 10^{-3}$	$2.54\ 10^{-2}$	$2.43\ 10^{-1}$

Tableau 3.3: *EQMN pour des mélanges de 2 à 5 matériaux purs avec RSB=50 dB*

Nombre de matériaux / Moyenne EQMN	2	3	4	5
Spectres des matériaux purs	$1.97\ 10^{-8}$	$2.79\ 10^{-8}$	$4.39\ 10^{-8}$	$6.31\ 10^{-8}$
Coefficients d'abondance/quadratiques	$7.24\ 10^{-3}$	$4.40\ 10^{-2}$	$5.74\ 10^{-2}$	$4.11\ 10^{-1}$

Nous pouvons remarquer que l'EQMN moyenne associée aux spectres des matériaux purs est faible jusqu'à un RSB égal à 40 dB ($\sim 10^{-7}$). Dans ce cas, l'EQMN moyenne associée aux coefficients d'abondance et quadratiques est acceptable ($\leq 2.54 \times 10^{-2}$ pour RSB=40 dB). A partir d'un $RSB = 30$ dB, les résultats se dégradent : on obtient une $EQMN$ moyenne égale à 7.9×10^{-4} pour les spectres et une $EQMN$ moyenne égale à 2.43×10^{-1} pour les fractions d'abondance et les coefficients quadratiques.

Par la suite, nous avons évalué les performances de la méthode proposée en fonction du nombre des matériaux purs présents dans l'image synthétique bruitée. Nous nous sommes limités au calcul des EQMN pour des mélanges contenant de 2 à 5 matériaux purs. Dans chaque test, nous avons effectué une dizaine de simulations et, à chaque simulation, les fractions d'abondance et les coefficients quadratiques ont été modifiés et les spectres des matériaux purs ont été choisis aléatoirement. Les résultats obtenus pour un RSB égal à 50 dB sont montrés dans le Tableau 3.3.

Nous pouvons remarquer que les valeurs obtenues pour l'EQMN associée aux spectres sont faibles ($\sim 10^{-8}$). Pour ce qui est des coefficients d'abondance et quadratiques, les résultats obtenus sont acceptables jusqu'à un nombre de matériaux purs égal à 4 ($\leq 5.74 \times 10^{-2}$). Par ailleurs, nous avons constaté que, dans toutes nos expériences, l'EQMN des spectres est plus faible que l'EQMN des coefficients d'abondance et quadratiques. Cela est dû au fait que cette dernière est calculée sur tout l'ensemble des

coefficients et pas seulement sur les fractions d'abondance. Ainsi, l'estimation moins bonne des coefficients quadratiques engendre la dégradation du résultat de l'estimation totale. Malgré ce fait, nous pouvons constater les bonnes performances de la méthode proposée lors de l'utilisation des données synthétiques en présence de bruit faible.

3.4.2.3 Expérience 3

Dans cette expérience, nous avons voulu tester notre méthode dans le cas d'une configuration plus réaliste que celles utilisées dans les expériences précédentes.
Ainsi, dans une première expérimentation, nous avons généré une image hyperspectrale synthétique à partir des cartes de classification d'occupation des sols d'une image satellitaire réelle. Cette dernière, présentée dans la Figure 3.3, correspond à une région de la ville de Toulouse.



Figure 3.3: *Image satellitaire d'une région de Toulouse*

Nous avons classifié cette image satellitaire, qui est de taille (500×500) pixels, à l'aide du logiciel de traitement d'images de télédétection ENVI (ENvironment for Visualizing Images) en tenant compte de la présence de quatre matériaux purs seulement qui sont : l'asphalte, la tuile, le gravier, et la végétation. A partir de cette classification d'occupation des sols, nous avons déduit quatre cartes d'abondance pour les quatre matériaux contenant chacune (50×50) pixels. Ceci a été réalisé en calculant les valeurs moyennes des pixels de la classification sur une fenêtre glissante de taille (10×10) pixels. Ces cartes d'abondances sont présentées dans la Figure 3.4. Ces cartes contiennent au moins 13% de pixels purs et au plus 15% de pixels purs.

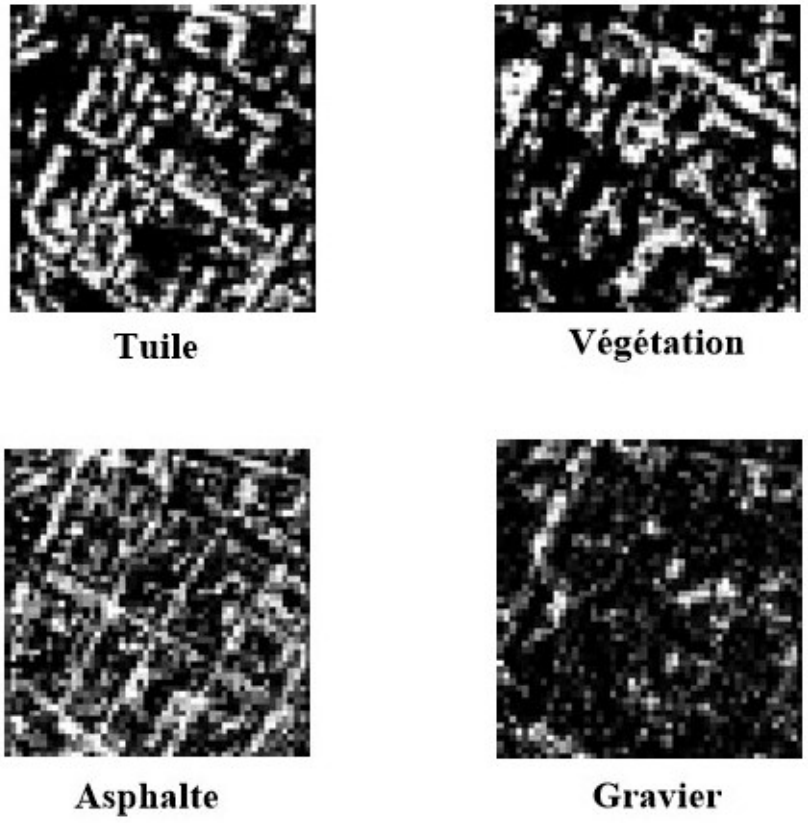


Figure 3.4: *Cartes d'abondances déduites à partir de l'image de Toulouse*

En outre, nous avons généré les coefficients quadratiques $a_{j,k}$ avec des valeurs aléatoires uniformément distribuées sur $[0, 0.5]$ afin d'aboutir à un mélange linéaire-quadratique. Les spectres de matériaux purs choisis ont été pris de la base spectrale utilisée dans l'Expérience 1. Pour la végétation, nous avons utilisé le spectre de la pelouse.
L'estimation des spectres des matériaux purs a été réalisée comme dans les expériences précédentes selon la méthode détaillée dans la Section 3.3.2.2.

Le Tableau 3.4 montre les résultats de simulations obtenus dans cette expérience lorsque l'image est non bruitée d'une part, et d'autre part en variant la moyenne du *RSB* de 80 dB à 30 dB.

Nous pouvons remarquer que l'EQMN associée aux spectres des matériaux purs est faible jusqu'à un *RSB* égal à 50 dB ($\sim 10^{-8}$). L'EQMN calculée pour les coefficients d'abondance et quadratiques est acceptable jusqu'à un *RSB* de 50 dB.
La Figure 3.5 montre les cartes d'abondances estimées dans le cas idéal sans présence de bruit.

Tableau 3.4: *EQMN pour l'**Expérience 3** avec un RSB variant de* 30 *à* 80 *dB*

RSB (dB) / EQMN	∞	**80**	**60**	**50**	**40**	**30**
Spectres des matériaux purs	$7.18\ 10^{-32}$	$2.57\ 10^{-11}$	$2.68\ 10^{-9}$	$2.27\ 10^{-8}$	$3.12\ 10^{-7}$	$2.54\ 10^{-6}$
Coefficients d'abondance/quadratiques	$5.31\ 10^{-25}$	$2.18\ 10^{-4}$	$1.64\ 10^{-2}$	$9.12\ 10^{-2}$	$3.52\ 10^{-1}$	$6.25\ 10^{-1}$

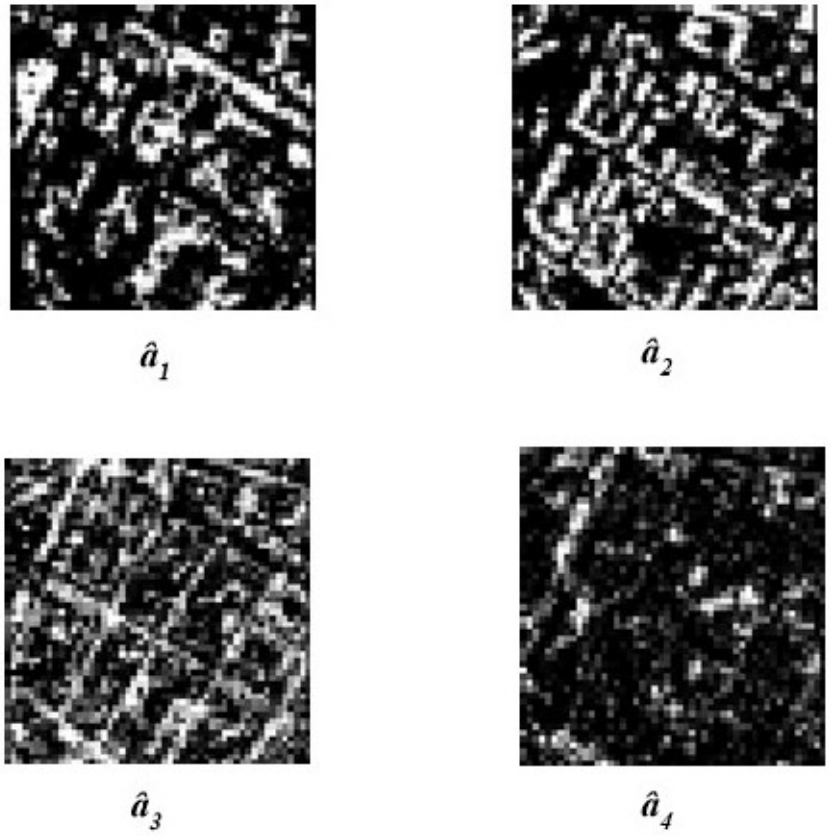


Figure 3.5: *Cartes d'abondances estimées*

Nous pouvons remarquer la bonne estimation des 4 cartes d'abondance en les comparant aux cartes d'abondances présentées dans la Figure 3.4, ce qui montre encore l'efficacité de la méthode proposée dans la séparation des mélanges linéaires-quadratiques avec la présence de pixels purs.

Une deuxième expérimentation a été réalisée en créant une image hyperspectrale à partir de 8 cartes d'abondance réalistes, de dimension (400×400) pixels, qui sont générées à partir d'une classification d'occupation de sols réelle (Karoui, 2012). L'ensemble des cartes est présenté dans la Figure 3.6. Ces cartes d'abondance contiennent au moins 1.5% et au plus 18% de pixels purs.

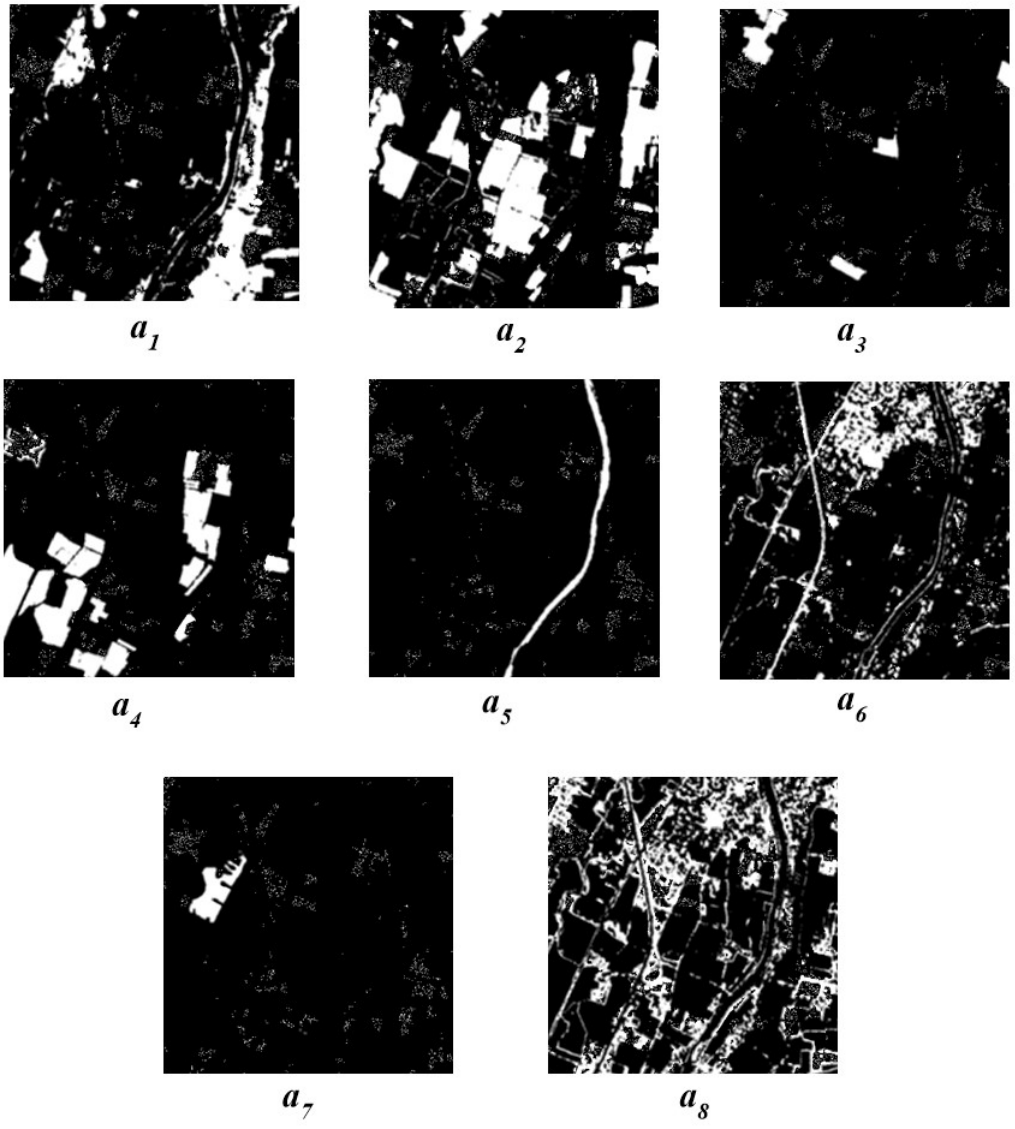


Figure 3.6: *Les 8 cartes d'abondance originales*

Comme dans les expériences précédentes, les coefficients quadratiques ont été générés avec des valeurs aléatoires uniformément distribuées sur l'intervalle $[0, 0.5]$ pour aboutir à un mélange linéaire-quadratique. Les 8 spectres de matériaux purs utilisés ont été pris de la base spectrale présentée en Figure 3.2. Selon l'équation (3.1),

nous avons utilisé cet ensemble de spectres, les cartes d'abondance et les coefficients quadratiques pour générer une image hyperspectrale non-bruitée.

Vu que les cartes d'abondance contiennent des zones étendues où seulement un seul matériau pur est présent, nous avons utilisé la méthode basée sur la détection des zones *mono-sources* décrite en section 3.3.2.1 afin d'estimer les spectres des matériaux purs. Pour ce faire, nous avons analysé l'image sur des petites zones d'analyse adjacentes de taille (4×4) pixels. Les cartes d'abondance estimées, ainsi que les spectres réels et estimés sont présentés dans les Figures 3.7 et 3.8, respectivement.

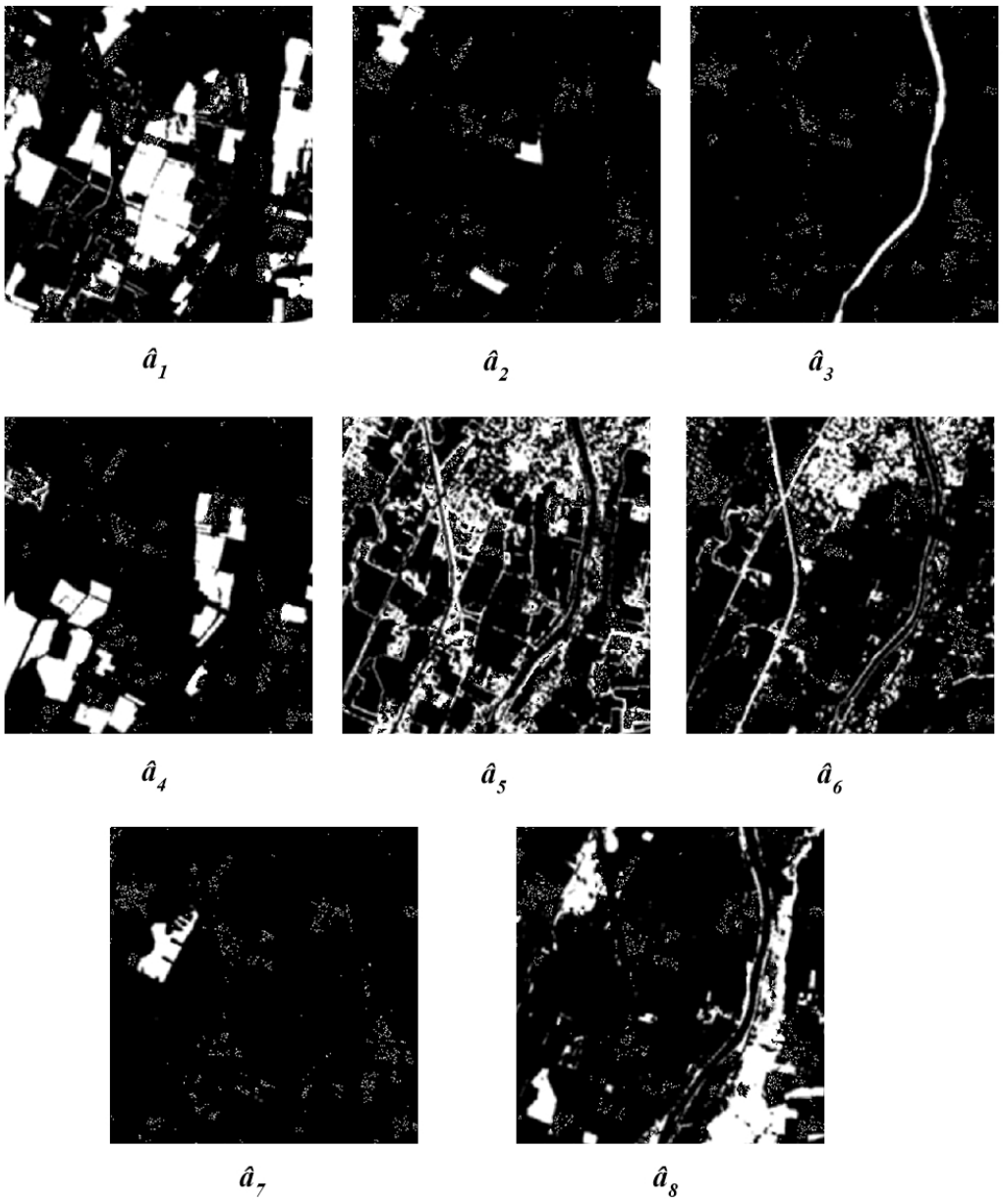


Figure 3.7: *Les 8 cartes d'abondance estimées*

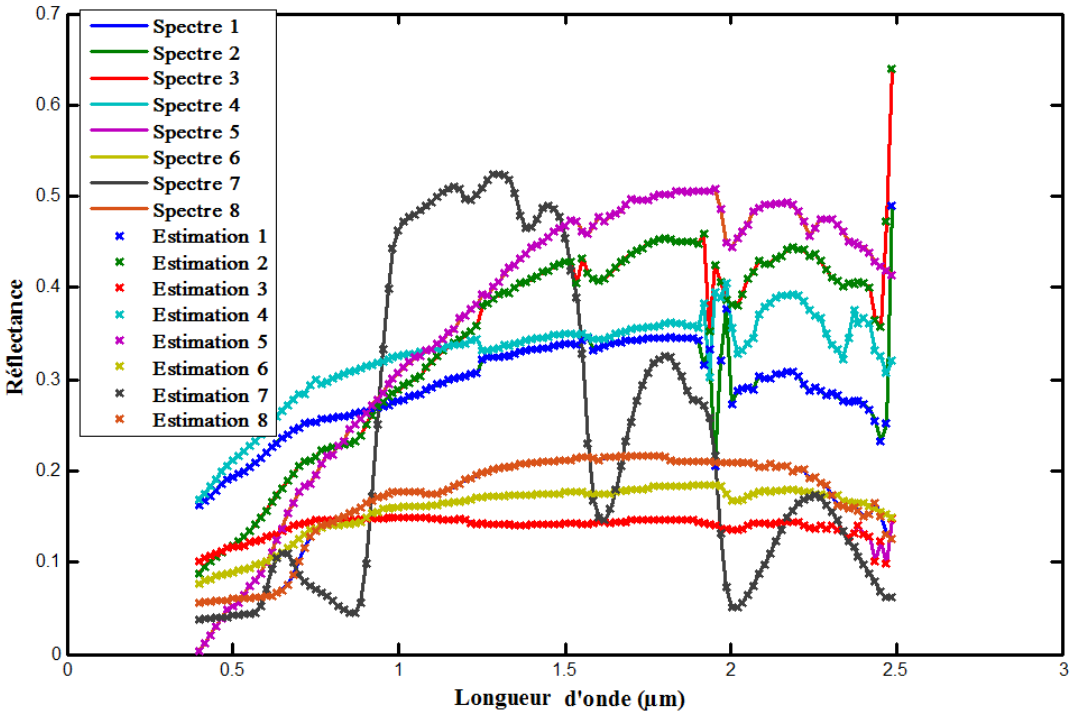


Figure 3.8: *Les 8 spectres estimés*

Nous pouvons remarquer que nous arrivons à estimer parfaitement les 8 spectres des matériaux purs : les courbes sont bien confondues et l'$EQMN_s$ est égale à 5.40×10^{-29}. De même, nous remarquons que nous avons bien estimé les 8 cartes d'abondance et l'$EQMN_a$ de l'ensemble des fractions d'abondance et coefficients quadratiques est égale à 1.22×10^{-7}. Ainsi, nous pouvons déduire les bonnes performances de la méthode proposée lors de la présence des zones mono-sources dans l'image hyperspectrale.

3.5 Conclusions

Dans ce chapitre, nous nous sommes intéressés au démélange non linéaire d'images hyperspectrales de télédétection urbaine en utilisant un modèle linéaire-quadratique. Nous avons proposé une approche qui identifie, d'abord, le nombre des matériaux purs dans l'image, puis estime les spectres des matériaux purs, en détectant soit des zones mono-sources soit des paires de pixels purs, et qui extrait finalement leurs fractions d'abondance et coefficients quadratiques. Les résultats expérimentaux ont montré que nous avons réussi à décomposer chaque spectre de pixel mélangé en plusieurs spectres de matériaux purs, et leurs contributions relatives dans le cas des images hyperspectrales non bruitées ou en présence d'un bruit faible.

Par ailleurs, l'estimation des spectres des pixels purs peut être effectuée d'une autre façon plus robuste pour éviter les fausses détections des pixels purs, dans le cas où deux mixels ont des spectres identiques. Ceci est possible en modifiant l'étape d'estimation des spectres des matériaux purs décrite dans la Section 3.3.2.2 par la comparaison de plus de deux spectres de pixels.

De plus, la détection des spectres des pixels purs peut être effectuée en calculant le coefficient d'intercorrélation entre les spectres des pixels au lieu de calculer la distance euclidienne surtout lorsqu'il s'agit d'une image bruitée vu que la corrélation est moins sensible au bruit.

L'approche proposée dans ce chapitre est efficace lors de la présence des zones mono-sources ou des pixels purs dans l'image, mais elle n'est plus applicable si aucune de ces hypothèses n'est vérifiée. Pour cela, une autre méthode qui ne se base pas sur l'hypothèse de la parcimonie est présentée dans le chapitre suivant.

Chapitre 4

Une méthode de NMF basée sur l'estimateur du maximum a posteriori (MAP) pour le démélange hyperspectral des mélanges linéaires-quadratiques

Sommaire

4.1 Introduction 86
4.2 Formulation du problème 87
4.2.1 Modèle Linéaire-Quadratique 87
4.2.2 Factorisation en Matrices Non-négatives (NMF) 88
4.3 Méthode proposée 89
4.3.1 Lois de probabilité a priori 89
4.3.1.1 Coefficients de mélange 89
4.3.1.2 Signaux sources 90
4.3.2 Vraisemblance 90
4.3.3 Estimateur MAP 91
4.3.4 Fonction de coût 91
4.3.5 Algorithmes d'optimisation 93
4.3.5.1 Algorithme du gradient projeté (Grd LQ-MAP) . 93
4.3.5.2 Algorithme multiplicatif projeté (Grd Mult-MAP) 95
4.3.5.3 Algorithme du gradient projeté modifié (Grd LQ-MAP-modif) 98
4.4 Résultats de simulations 101
4.4.1 Critère de performance 102
4.4.2 Tests expérimentaux 103
4.4.2.1 Test 1 103
4.4.2.2 Test 2 105
4.4.2.3 Test 3 106
4.5 Conclusions 109

4.1 Introduction

Dans le chapitre précédent, nous avons proposé une méthode ACPa pour le démélange des images hyperspectrales urbaines mélangées selon le modèle LQ. Cette méthode est basée sur l'hypothèse d'existence des zones mono-sources ou des paires de pixels purs par matériau dans le mélange. Néanmoins, cette hypothèse n'est pas toujours réaliste, notamment pour les images hyperspectrales à basse résolution spatiale.

Dans ce chapitre, nous proposons une autre méthode de séparation de sources adaptée au modèle LQ, qui n'exploite pas la parcimonie mais plutôt les informations a priori sur les inconnus du problème. Comme dans le chapitre précédent, notre but est de décomposer le spectre observé dans chaque pixel de l'image en un ensemble de spectres de matériaux purs, ainsi que leurs fractions d'abondance et leurs coefficients quadratiques. À cette fin, nous proposons une méthode de Factorisation en Matrices Non-négatives (NMF) se basant sur l'estimateur du Maximum A Posteriori (MAP). L'idée principale de cette méthode est de prendre en compte l'information préalable disponible sur la distribution des coefficients de mélange linéaires et quadratiques afin de mieux les estimer.

Ce chapitre est organisé comme suit. Nous commençons par un rappel sur le modèle de mélange LQ et la méthode NMF. Ensuite, nous présentons les lois de probabilité utilisées pour modéliser les paramètres de mélange, et nous en déduisons une fonction de coût adaptée à l'estimation MAP. Nous minimisons enfin cette fonction en utilisant différents algorithmes d'optimisation. Les résultats de simulation sur différents types de données, sont présentés à la fin du chapitre.

4.2 Formulation du problème

4.2.1 Modèle Linéaire-Quadratique

Rappelons que dans une image hyperspectrale urbaine acquise dans N bandes spectrales et composée par K pixels, il a été démontré (Meganem et al, 2014a) que le spectre de réflectance observé dans un pixel i peut être considéré comme un mélange LQ de L spectres de réflectance de matériaux purs, appelés encore "sources", comme indiqué dans l'équation suivante :

$$\mathbf{x}_i = \sum_{j=1}^{L} a_j(i)\mathbf{s}_j + \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)\mathbf{s}_j \odot \mathbf{s}_k + \mathbf{b}_i, \tag{4.1}$$

où $\mathbf{x}_i = [x_i(1), ..., x_i(N)]^T$ est le spectre associé au pixel observé i, $\mathbf{s}_j = [s_j(1), ..., s_j(N)]^T$ est le spectre du matériau pur j, $a_j(i)$ est la fraction d'abondance du matériau pur j dans le pixel i, $a_{j,k}(i)$ est le coefficient quadratique associé aux matériaux purs j et k dans le pixel i, le symbole $\odot$ représente le produit terme à terme de Hadamard, et $\mathbf{b}_i$ est un bruit additif, considéré gaussien, de moyenne nulle, i.i.d., et de variance σ_n^2. L'ensemble des vraies sources $\mathbf{s}_j$ et les pseudo-sources $\mathbf{s}_j \odot \mathbf{s}_k$ sont appelées sources étendues.

Il convient également de rappeler que les études physiques réalisées dans (Meganem et al, 2014a) ont montré que les contraintes physiques suivantes doivent être vérifiées :

$$\begin{cases} s_j(n) \geq 0 \\ a_j(i) \geq 0 \\ \sum_{j=1}^{L} a_j(i) = 1 \\ a_{j,k}(i) \in [0, 0.5] \end{cases} \forall \begin{cases} 1 \leq i \leq K \\ 1 \leq j \leq L \\ j \leq k \leq L \\ 1 \leq n \leq N \end{cases} \tag{4.2}$$

4.2.2 Factorisation en Matrices Non-négatives (NMF)

D'après (4.2), les sources et les coefficients de mélange sont non-négatifs. Ainsi, nous devons avoir recourt à une méthode de démélange qui permet de respecter cette propriété de non-négativité. Notre choix s'est porté vers la méthode de Factorisation en Matrices Non-négatives ou NMF qui permet de respecter cette propriété. Rappelons que le principe de la NMF proposée par Lee et Seung (Lee and Seung, 1999, 2001) est le suivant : étant donné une matrice de données non-négative $\mathbf{V}$, la NMF consiste à trouver une factorisation approchée de $\mathbf{V} \approx \mathbf{WH}$ en matrices non-négatives $\mathbf{W}$ et $\mathbf{H}$. Nous minimisons souvent un critère quadratique, défini par

$$f(\mathbf{W}, \mathbf{H}) = \|\mathbf{V} - \mathbf{WH}\|_F^2 \ . \tag{4.3}$$

La méthode NMF de base est seulement adaptée au modèle linéaire, ce qui n'est pas notre cas. Dans cette étude, nous cherchons à étendre cette méthode afin de prendre en compte les coefficients quadratiques, d'une part, et les informations a priori sur les paramètres de mélange d'autre part. Cela nécessite la réécriture du modèle de mélange linéaire-quadratique sous la forme matricielle suivante (Meganem et al, 2014b)

$$\mathbf{X} = \mathbf{AS} + \mathbf{B} \ , \tag{4.4}$$

où

- $\mathbf{X} = [\ \mathbf{x}_1 \cdots \mathbf{x}_K\ ]^T$ est la matrice contenant les K vecteurs d'observation.
- $\mathbf{A} = \begin{pmatrix} a_1(1) & \cdots & a_L(1) & a_{1,1}(1) & \cdots & a_{L,L}(1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_1(K) & \cdots & a_L(K) & a_{1,1}(K) & \cdots & a_{L,L}(K) \end{pmatrix}$ est la matrice contenant les fractions d'abondance et les coefficients quadratiques.
- $\mathbf{S} = [\ \mathbf{s}_1 \cdots \mathbf{s}_L\ \mathbf{s}_1 \odot \mathbf{s}_1 \cdots \mathbf{s}_L \odot \mathbf{s}_L\ ]^T$ est la matrice contenant les sources étendues, c'est-à-dire les vraies sources et leurs produits deux à deux.
- $\mathbf{B} = [\ \mathbf{b}_1 \cdots \mathbf{b}_K\ ]^T$ est la matrice contenant les K vecteurs de bruit.

Il est à noter que cette écriture matricielle permet de transformer le mélange LQ en un mélange linéaire des sources étendues.

Dans ce qui suit, nous expliquons comment les informations a priori disponibles sur les coefficients de mélange peuvent être utilisées afin d'améliorer la méthode NMF adaptée aux mélanges LQ (Meganem et al, 2014b) en déterminant l'estimateur MAP.

4.3 Méthode proposée

4.3.1 Lois de probabilité a priori

Dans certains cas, des informations sur les inconnues du problème peuvent être déterminées par les experts du domaine. L'utilisation de ces informations préalables est très avantageuse puisqu'elle permet de mieux estimer ces inconnues. L'exploitation de cette connaissance peut être réalisée en modélisant les lois de probabilité a priori des inconnues.
Dans cette partie, nous présentons les lois a priori que nous proposons pour modéliser les coefficients de mélange et les signaux sources.

4.3.1.1 Coefficients de mélange

En supposant que les coefficients quadratiques sont indépendants entre eux et qu'il sont aussi indépendants des fractions d'abondance, et en supposant que les paramètres de mélange dans un pixel sont indépendants de ceux des autres pixels, nous pouvons écrire

$$f(\mathbf{A}) = \prod_{i=1}^{K} \left[f(a_1(i), \cdots, a_L(i)) \prod_{j=1}^{L} \prod_{k=j}^{L} f(a_{j,k}(i)) \right], \tag{4.5}$$

où f représente la densité de probabilité.

D'après (4.2), les fractions d'abondance associées à chaque pixel doivent être à la fois non-négatives, et de somme égale à un. Elle peuvent donc être modélisées par une distribution de Dirichlet qui vérifie ces deux propriétés (Nascimento and Bioucas-Dias,

2007, 2012)

$$f(a_1(i),\cdots,a_L(i)) = \frac{\Gamma(\sum_{j=1}^{L}\theta_j)}{\prod_{j=1}^{L}\Gamma(\theta_j)}\prod_{j=1}^{L}a_j(i)^{\theta_j-1}, \tag{4.6}$$

où Γ est la fonction Gamma et θ_j sont les paramètres de Dirichlet.

Des études numériques dans un environnement urbain réalisées dans (Meganem et al, 2014a) ont montré que les coefficients quadratiques sont non-négatifs et qu'ils peuvent être modélisés par une densité de probabilité décroissante. En effet, dans la plupart des pixels, les termes quadratiques dans le mélange ont des valeurs faibles, et ils ne prennent leurs valeurs les plus élevées $(0,4-0,5)$ que dans quelques pixels. Ainsi, nous avons décidé de modéliser les coefficients quadratiques par une distribution demi-gaussienne, comme suit :

$$f(a_{j,k}(i)) = \begin{cases} \frac{2\vartheta_{j,k}}{\pi}\exp\left(\frac{-(a_{j,k}(i))^2\vartheta_{j,k}^2}{\pi}\right), \forall a_{j,k}(i)\geq 0 \\ 0, \text{ ailleurs} \end{cases} \tag{4.7}$$

où $\vartheta_{j,k} = \frac{\sqrt{\pi}}{\sigma_{j,k}\sqrt{2}}$ est le paramètre de cette densité.

4.3.1.2 Signaux sources

Comme aucune information a priori concernant les spectres des sources n'est disponible, nous affectons à chaque source une loi non-informative uniforme sur l'intervalle $[0,1]$.

4.3.2 Vraisemblance

La vraisemblance $f(\mathbf{X}|\mathbf{A},\mathbf{S})$ dépend de la distribution du bruit présent dans le mélange. Étant donné que le bruit dans (4.4) est considéré spatialement et spectralement i.i.d., gaussien, de moyenne nulle et de variance σ_n^2, alors la fonction $f(\mathbf{X}|\mathbf{A},\mathbf{S})$ et son logarithme sont présentés respectivement comme suit

$$f(\mathbf{X}|\mathbf{A},\mathbf{S}) = \prod_{i=1}^{K}\prod_{l=1}^{N}\frac{1}{\sqrt{2\pi\sigma_n^2}}\exp\left(\frac{-(X-AS)_{il}^2}{2\sigma_n^2}\right), \tag{4.8}$$

$$\log f(\mathbf{X}|\mathbf{A},\mathbf{S}) = C - \frac{1}{2\sigma_n^2}||\mathbf{X}-\mathbf{AS}||_F^2, \tag{4.9}$$

où $(X-AS)_{il} = x_{il} - \sum_{j=1}^{L} a_j(i)s_{jl} - \sum_{j=1}^{L}\sum_{k=j}^{L} a_{j,k}(i)s_{jl}s_{kl}$, et C est une constante.

4.3.3 Estimateur MAP

Afin d'estimer les inconnues du problème, nous avons choisi d'utiliser l'estimateur MAP. Comme son nom l'indique, le principe de ce dernier consiste à maximiser la densité de probabilité a posteriori $f(\mathbf{A},\mathbf{S}|\mathbf{X})$. L'avantage principal de l'estimateur MAP est qu'il permet d'inclure les informations a priori sur les inconnues dans le processus d'estimation. Ainsi, nous estimons les matrices inconnues $\mathbf{A}$ et $\mathbf{S}$ de la façon suivante

$$\{\hat{\mathbf{A}},\hat{\mathbf{S}}\} = \underset{\mathbf{A},\mathbf{S}}{\text{argmax}} f(\mathbf{A},\mathbf{S}|\mathbf{X}) = \underset{\mathbf{A},\mathbf{S}}{\text{argmax}} f(\mathbf{X}|\mathbf{A},\mathbf{S})f(\mathbf{A})f(\mathbf{S}). \tag{4.10}$$

Comme une loi non-informative uniforme a été affectée à chaque signal source, l'estimateur MAP s'écrit comme suit

$$\{\hat{\mathbf{A}},\hat{\mathbf{S}}\} = \underset{\mathbf{A},\mathbf{S}}{\text{argmax}} f(\mathbf{X}|\mathbf{A},\mathbf{S})f(\mathbf{A}) = \underset{\mathbf{A},\mathbf{S}}{\text{argmax}} \left[\log f(\mathbf{X}|\mathbf{A},\mathbf{S}) + \log f(\mathbf{A})\right]. \tag{4.11}$$

Dans ces formules et les formules qui suivent, "$\hat{}$" se réfère aux valeurs estimées.

4.3.4 Fonction de coût

En insérant (4.6) et (4.7) dans (4.5), nous obtenons

$$f(\mathbf{A}) = \prod_{i=1}^{K}\left[\left(\frac{\Gamma(\sum_{j=1}^{L}\theta_j)}{\prod_{j=1}^{L}\Gamma(\theta_j)}\prod_{j=1}^{L}a_j(i)^{\theta_j-1}\right)\left(\prod_{j=1}^{L}\prod_{k=j}^{L}\frac{2\vartheta_{j,k}}{\pi}\exp\left(\frac{-(a_{j,k}(i))^2\vartheta_{j,k}^2}{\pi}\right)\right)\right], \tag{4.12}$$

puis en calculant son logarithme,

$$\begin{aligned}\log(f(\mathbf{A})) &= \sum_{i=1}^{K}\Bigg[\log(\Gamma(\sum_{j=1}^{L}\theta_j)) - \sum_{j=1}^{L}\log(\Gamma(\theta_j)) + \sum_{j=1}^{L}(\theta_j-1)\log(a_j(i)) \\ &+ \sum_{j=1}^{L}\sum_{k=j}^{L}\left(\log(\frac{2\vartheta_{j,k}}{\pi}) - \frac{\vartheta_{j,k}^2}{\pi}a_{j,k}(i)^2\right)\Bigg], \end{aligned} \tag{4.13}$$

et en utilisant (4.9) et (4.11), l'estimateur MAP pour les inconnues $\mathbf{A}$ et $\mathbf{S}$ peut être écrit de la façon suivante

$$\begin{aligned}\{\hat{\mathbf{A}},\hat{\mathbf{S}}\} &= \underset{\mathbf{A},\mathbf{S}}{\text{argmax}}\ [\log f(\mathbf{X}|\mathbf{A},\mathbf{S}) + \log f(\mathbf{A})] \\ &= \underset{\mathbf{A},\mathbf{S}}{\text{argmax}}\Bigg[-\frac{1}{2\sigma_n^2}||\mathbf{X}-\mathbf{A}\mathbf{S}||_F^2 + \sum_{i=1}^{K}\Big[\log(\Gamma(\sum_{j=1}^{L}\theta_j)) - \sum_{j=1}^{L}\log(\Gamma(\theta_j)) \\ &+ \sum_{j=1}^{L}(\theta_j-1)\log(a_j(i)) + \sum_{j=1}^{L}\sum_{k=j}^{L}\left(\log(\frac{2\vartheta_{j,k}}{\pi}) - \frac{\vartheta_{j,k}^2}{\pi}a_{j,k}(i)^2\right)\Big] + C\Bigg], \end{aligned} \tag{4.14}$$

Ce problème de maximisation correspond à celui de la minimisation de la fonction de coût suivante

$$J = \frac{1}{2}||\mathbf{X}-\hat{\mathbf{A}}\hat{\mathbf{S}}||_F^2 - \eta R, \tag{4.15}$$

où R est un terme de régularisation lié à l'information a priori et défini par

$$\begin{aligned} R &= K\log\Gamma(\sum_{j=1}^{L}\hat{\theta}_j) - K\sum_{j=1}^{L}\log\Gamma(\hat{\theta}_j) + \sum_{j=1}^{L}\left[(\hat{\theta}_j-1)\sum_{i=1}^{K}\log\,\hat{a}_j(i)\right] \\ &+ \sum_{j=1}^{L}\sum_{k=j}^{L}\left[K\log(\frac{2\hat{\vartheta}_{j,k}}{\pi}) - \frac{\hat{\vartheta}_{j,k}^2}{\pi}\sum_{i=1}^{K}(\hat{a}_{j,k}(i))^2\right], \end{aligned} \tag{4.16}$$

et η est un paramètre de régularisation dont la valeur dépend de la variance du bruit σ_n^2. Dans la pratique, puisque σ_n^2 est inconnue, la valeur de η est sélectionnée manuellement.

4.3.5 Algorithmes d'optimisation

Dans cette partie, nous allons présenter les algorithmes d'optimisation utilisés afin de minimiser la fonction de coût J, définie dans l'équation (4.15).
Si les paramètres θ_j et $\vartheta_{j,k}$ des distributions a priori sont connus, nous pouvons facilement adapter la méthode LQ-NMF décrite dans (Meganem et al, 2014b) à la fonction de coût (4.15) juste en tenant compte du terme de régularisation R, défini par (4.16), lors du calcul du gradient par rapport aux éléments de la matrice $\hat{\mathbf{A}}$, c-à-d les paramètres $\hat{a}_j(i)$ et $\hat{a}_{j,k}(i)$. Cependant, en pratique, les paramètres des densités de probabilité ne sont pas exactement connus, bien que nous puissions avoir une idée sur leurs domaines de variation possibles. Par exemple, si nous savons que la scène observée est composée de pixels très mélangés, alors nous devons choisir des grandes valeurs pour les paramètres $\hat{\theta}_j$. En outre, puisque les valeurs des coefficients quadratiques $\hat{a}_{j,k}(i)$ sont réparties dans l'intervalle $[0, 0.5]$, les paramètres $\hat{\vartheta}_{j,k}$ devraient prendre des grandes valeurs (correspondant à des petites variances des densités demi-gaussiennes). Ainsi, la fonction de coût J doit être minimisée non seulement par rapport aux coefficients de mélange et les spectres des sources, mais aussi par rapport aux paramètres des distributions a priori. Notons que cet objectif peut être atteint en utilisant les méthodes de Monte Carlo par Chaînes de Markov (MCMC) (S. Moussaoui and Carteret, 2006; Dobigeon et al, 2008) au prix d'une augmentation significative du temps de calcul.
Les trois algorithmes d'optimisation que nous proposons dans les trois sous-sections suivantes sont beaucoup plus rapides que les algorithmes MCMC.

4.3.5.1 Algorithme du gradient projeté (Grd LQ-MAP)

La première méthode d'optimisation utilisée est l'algorithme du gradient projeté. Il s'agit d'un algorithme du gradient à pas fixe, qui projette, après chaque itération, les valeurs des paramètres de mélange, des spectres des sources et des paramètres des distributions a priori estimées dans leurs domaines de variation possibles. Les dérivées de J par rapport aux différentes inconnues sont énumérées ci-dessous.

Dans ces formules, ψ représente la fonction Digamma, c.à.d la dérivée logarithmique de la fonction Gamma. Pour calculer les dérivées de $\frac{1}{2}||\mathbf{X} - \hat{\mathbf{A}}\hat{\mathbf{S}}||_F^2$ par rapport aux paramètres de mélange et aux sources, nous utilisons les résultats fournis dans les

Equations (4) et (8) de Meganem et al (2014b).

$$\frac{\partial J}{\partial \hat{\theta}_\ell} = -\eta[K\psi(\sum_{j=1}^{L}\hat{\theta}_j) - K\psi(\hat{\theta}_\ell) + \sum_{i=1}^{K}\log\hat{a}_\ell(i)], \tag{4.17}$$

$$\frac{\partial J}{\partial \hat{\vartheta}_{m,n}} = -\eta[\frac{K}{\hat{\vartheta}_{m,n}} - \frac{2\hat{\vartheta}_{m,n}}{\pi}\sum_{i=1}^{K}(\hat{a}_{m,n}(i))^2], \tag{4.18}$$

$$\frac{\partial J}{\partial \hat{a}_m(\ell)} = [(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{m\ell} - \eta\frac{\hat{\theta}_m - 1}{\hat{a}_m(\ell)}, \tag{4.19}$$

$$\frac{\partial J}{\partial \hat{a}_{m,n}(\ell)} = [(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{(mn)\ell} + \eta\frac{2\hat{\vartheta}^2_{m,n}\hat{a}_{m,n}(\ell)}{\pi}, \tag{4.20}$$

$$\begin{aligned}\frac{\partial J}{\partial \hat{s}_{pn}} &= [\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{pn} + \sum_{j=1, j\neq p}^{L}\hat{s}_{jn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(jp)n} \\ &+ 2\hat{s}_{pn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(pp)n},\end{aligned} \tag{4.21}$$

où par exemple (jp) est l'indice de la colonne de la matrice $\hat{\mathbf{A}}$ correspondant aux coefficients des termes quadratiques $\hat{a}_{j,p}(n)$ et $\hat{s}_{pn}$ désigne l'élément (p,n) de la matrice $\hat{\mathbf{S}}$.

Dans chaque itération de l'algorithme du gradient projeté, les règles suivantes sont utilisées afin de mettre à jour les paramètres de mélange, les spectres des sources et les paramètres des distributions a priori :

$$\hat{a}_j(i) \leftarrow \left[\hat{a}_j(i) - \mu_a\frac{\partial J}{\partial \hat{a}_j(i)}\right]_P, \tag{4.22}$$

$$\hat{a}_{j,k}(i) \leftarrow \left[\hat{a}_{j,k}(i) - \mu_a\frac{\partial J}{\partial \hat{a}_{j,k}(i)}\right]_P, \tag{4.23}$$

$$\hat{s}_{ji} \leftarrow \left[\hat{s}_{ji} - \mu_s\frac{\partial J}{\partial \hat{s}_{ji}}\right]_P, \tag{4.24}$$

$$\hat{\theta}_j \leftarrow \left[\hat{\theta}_j - \mu_\theta\frac{\partial J}{\partial \hat{\theta}_j}\right]_P, \tag{4.25}$$

$$\hat{\vartheta}_{j,k} \leftarrow \left[\hat{\vartheta}_{j,k} - \mu_\vartheta\frac{\partial J}{\partial \hat{\vartheta}_{j,k}}\right]_P, \tag{4.26}$$

où μ_a, μ_s, μ_θ, μ_ϑ sont les pas d'apprentissage du gradient ayant des petites valeurs positives et fixes, et $[u]_P$ correspond à la projection de u dans l'intervalle P. En effet, si après la mise à jour du gradient, la valeur estimée est en dehors de P, elle est

remplacée par la borne la plus proche de P (Lin, 2007). Ces intervalles sont $[0, 1]$ pour $\hat{a}_j(i)$, $[0, 0.5]$ pour $\hat{a}_{j,k}(i)$, et $[0, \infty)$ pour $\hat{s}_{ji}$. Concernant les paramètres $\hat{\theta}_j$ et $\hat{\vartheta}_{j,k}$, si certaines informations sur leurs domaines de variation sont disponibles, elles peuvent être utilisées pour choisir les intervalles de projection, sinon ils sont projetés dans l'intervalle $[0, \infty)$.

Dans chaque itération de l'algorithme, l'équation (4.24) est utilisée pour mettre à jour les premières lignes de la matrice $\hat{\mathbf{S}}$ (correspondant aux vraies sources), puis les dernières lignes de $\hat{\mathbf{S}}$ correspondent aux produits deux à deux des premières lignes

$$\hat{s}_{(jp)n} \leftarrow \hat{s}_{jn} \times \hat{s}_{pn}. \tag{4.27}$$

Pour garantir que la somme des fractions d'abondance soit égale à 1, nous normalisons leurs valeurs estimées en les divisant par leur somme

$$\hat{a}_j(i) \leftarrow \frac{\hat{a}_j(i)}{\sum_{k=1}^{L} \hat{a}_k(i)}. \tag{4.28}$$

Le récapitulatif de cet algorithme est présenté dans Algorithme 1.

4.3.5.2 Algorithme multiplicatif projeté (Grd Mult-MAP)

L'un des algorithmes les plus utilisés pour la méthode NMF linéaire est l'algorithme multiplicatif. L'avantage de cet algorithme est que les valeurs des pas d'apprentissage sont auto-adaptatives à chaque itération. La méthode de base développée par Lee et Seung (Lee and Seung, 2001) pour le modèle linéaire consiste à réaliser, à chaque itération, les règles de mise à jour suivantes :

$$\hat{\mathbf{A}} \leftarrow \hat{\mathbf{A}} \odot ((\mathbf{X}\hat{\mathbf{S}}^T) \oslash (\hat{\mathbf{A}}\hat{\mathbf{S}}\hat{\mathbf{S}}^T + \varepsilon), \tag{4.29}$$

$$\hat{\mathbf{S}} \leftarrow \hat{\mathbf{S}} \odot ((\hat{\mathbf{A}}^T\mathbf{X}) \oslash (\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}} + \varepsilon), \tag{4.30}$$

où $\oslash$ représente la division élément par élément et ε est une petite valeur ajoutée au dénominateur pour éviter une division par zéro.

Algorithme 1 : Grd LQ-MAP

- Initialisation des paramètres $\hat{s}_{pn},\ \hat{a}_m(l),\ \hat{a}_{m,n}(\ell),\ \hat{\theta}_\ell$, et $\hat{\vartheta}_{m,n}$.

Faire

- Mise à jour des spectres des signaux sources :

$$\hat{s}_{pn} \leftarrow \Big[\hat{s}_{pn} - \mu_s\Big[\ [\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{pn} + \sum_{j=1, j\neq p}^{L} \hat{s}_{jn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(jp)n} + 2\hat{s}_{pn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(pp)n}\Big]\Big]_P$$

$$\hat{s}_{(jp)n} \leftarrow \hat{s}_{jn} \times \hat{s}_{pn}$$

- Mise à jour des fractions d'abondance :

$$\hat{a}_m(l) \leftarrow \Big[\hat{a}_m(l) - \mu_a\Big[\ [(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{m\ell} - \eta\tfrac{\hat{\theta}_m - 1}{\hat{a}_m(\ell)}\ \Big]\Big]_P$$

$$\hat{a}_m(l) \leftarrow \frac{\hat{a}_m(l)}{\sum_{k=1}^{L}\hat{a}_k(l)}$$

- Mise à jour des coefficients quadratiques :

$$\hat{a}_{m,n}(\ell) \leftarrow \Big[\hat{a}_{m,n}(\ell) - \mu_a\Big[\ [(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{(mn)\ell} + \eta\tfrac{2\hat{\vartheta}^2_{m,n}\hat{a}_{m,n}(\ell)}{\pi}\ \Big]\Big]_P$$

- Mise à jour des paramètres des distributions :

$$\hat{\theta}_\ell \leftarrow \Big[\hat{\theta}_\ell + \mu_\theta\ \eta[K\psi(\sum_{j=1}^{L}\hat{\theta}_j) - K\psi(\hat{\theta}_\ell) + \sum_{i=1}^{K}\log\hat{a}_\ell(i)\ ]\Big]_P$$

$$\hat{\vartheta}_{m,n} \leftarrow \Big[\hat{\vartheta}_{m,n} + \mu_\vartheta\eta[\frac{K}{\hat{\vartheta}_{m,n}} - \frac{2\hat{\vartheta}_{m,n}}{\pi}\sum_{i=1}^{K}(\hat{a}_{m,n}(i))^2]\Big]_P$$

Jusqu'à la convergence

Dans cette étude, nous modifions l'algorithme multiplicatif de base afin de traiter le modèle de mélange linéaire-quadratique, et tenir compte des informations a priori sur les coefficients de mélange. Pour la mise à jour des coefficients de mélange, nous avons adapté l'approche présentée dans (Cichocki et al, 2006) qui permet de garantir la non-négativité des données afin de vérifier les contraintes physiques imposées sur les données traitées. De cette façon, les pas d'apprentissage du gradient μ_{a_j} et $\mu_{a_{j,k}}$ sont définis comme suit

$$\mu_{a_j}(i) = \frac{\hat{a}_j(i)}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{ji}}, \qquad \mu_{a_{j,k}}(i) = \frac{\hat{a}_{j,k}(i)}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{(jk)i}}, \tag{4.31}$$

ce qui mène à avoir les règles de mise à jour suivantes

$$\hat{a}_j(i) \leftarrow \left[\hat{a}_j(i) \frac{[\mathbf{X}\hat{\mathbf{S}}^T]_{ji} + \eta \frac{\hat{\theta}_j - 1}{\hat{a}_j(i)}}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{ji} + \varepsilon} \right]_P, \tag{4.32}$$

$$\hat{a}_{j,k}(i) \leftarrow \left[\hat{a}_{j,k}(i) \frac{[\mathbf{X}\hat{\mathbf{S}}^T]_{(jk)i} - \eta \frac{2\hat{\vartheta}^2_{j,k}\hat{a}_{j,k}(i)}{\pi}}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{(jk)i} + \varepsilon} \right]_P. \tag{4.33}$$

Afin de garantir la somme égale à 1 des fractions d'abondance, nous normalisons les valeurs estimées selon l'équation (4.28).

En ce qui concerne les signaux sources, puisque le terme de régularisation R ne dépend pas de $\hat{\mathbf{S}}$, nous utilisons les calculs effectués dans (Meganem et al, 2014b). Ces calculs permettent de modifier la règle de mise à jour de l'algorithme multiplicatif linéaire pour prendre en compte les termes quadratiques. Ainsi, la mise à jour de $\hat{s}_{pn}$ a été réalisée comme suit

$$\hat{s}_{pn} \leftarrow \hat{s}_{pn} \frac{[\hat{\mathbf{A}}^T\mathbf{X}]_{pn} + 2\hat{s}_{pn}[\hat{\mathbf{A}}^T\mathbf{X}]_{(pp)n} + \sum_{j=1,j\neq p}^{L} \hat{s}_{jn}[\hat{\mathbf{A}}^T\mathbf{X}]_{(jp)n}}{[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{pn} + 2s_{pn}[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{(pp)n} + \sum_{j=1,j\neq p}^{L} \hat{s}_{jn}[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{(jp)n} + \epsilon}. \tag{4.34}$$

Par la suite, la mise à jour des termes quadratique $\hat{s}_{(jp)n}$ a été réalisée comme dans l'équation (4.27).

Afin de déterminer les paramètres des distributions $\hat{\theta}_j$ et $\hat{\vartheta}_{i,j}$, nous avons gardé les lois de mise à jour décrites auparavant dans les équations (4.25) et (4.26), respectivement.

Le récapitulatif de l'algorithme est présenté dans l'Algorithme 2.

Algorithme 2 : Grd Mult-MAP

- Initialisation des paramètres $\hat{s}_{pn}$, $\hat{a}_m(l)$, $\hat{a}_{m,n}(\ell)$, $\hat{\theta}_\ell$, et $\hat{\vartheta}_{m,n}$.

Faire

- Mise à jour des spectres des signaux sources :

$$\hat{s}_{pn} \leftarrow \left[\hat{s}_{pn}\frac{[\hat{\mathbf{A}}^T\mathbf{X}]_{pn}+2\hat{s}_{pn}[\hat{\mathbf{A}}^T\mathbf{X}]_{(pp)n}+\sum_{j=1,j\neq p}^{L}\hat{s}_{jn}[\hat{\mathbf{A}}^T\mathbf{X}]_{(jp)n}}{[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{pn}+2\hat{s}_{pn}[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{(pp)n}+\sum_{j=1,j\neq p}^{L}\hat{s}_{jn}[\hat{\mathbf{A}}^T\hat{\mathbf{A}}\hat{\mathbf{S}}]_{(jp)n}+\epsilon}\right]_P$$

$$\hat{s}_{(jp)n} \leftarrow \hat{s}_{jn} \times \hat{s}_{pn}$$

- Mise à jour des fractions d'abondance :

$$\hat{a}_m(l) \leftarrow \left[\ \hat{a}_m(l)\frac{[\mathbf{X}\hat{\mathbf{S}}^T]_{ml}+\eta\frac{\hat{\theta}_m-1}{\hat{a}_m(l)}}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{ml}+\varepsilon}\ \right]_P$$

$$\hat{a}_m(l) \leftarrow \frac{\hat{a}_m(l)}{\sum_{k=1}^{L}\hat{a}_k(l)}$$

- Mise à jour des coefficients quadratiques :

$$\hat{a}_{m,n}(l) \leftarrow \left[\ \hat{a}_{m,n}(l)\frac{[\mathbf{X}\hat{\mathbf{S}}^T]_{(mn)l}-\eta\frac{2\hat{\vartheta}^2_{m,n}\hat{a}_{m,n}(l)}{\pi}}{[\hat{\mathbf{A}}(\hat{\mathbf{S}}\hat{\mathbf{S}}^T)]_{(mn)l}+\varepsilon}\ \right]_P$$

- Mise à jour des paramètres des distributions :

$$\hat{\theta}_\ell \leftarrow \left[\hat{\theta}_\ell+\mu_\theta\ \eta[K\psi(\sum_{j=1}^{L}\hat{\theta}_j)-K\psi(\hat{\theta}_\ell)+\sum_{i=1}^{K}\log\hat{a}_\ell(i)\ ]\right]_P$$

$$\hat{\vartheta}_{m,n} \leftarrow \left[\hat{\vartheta}_{m,n}+\mu_\vartheta\eta[\frac{K}{\hat{\vartheta}_{m,n}}-\frac{2\hat{\vartheta}_{m,n}}{\pi}\sum_{i=1}^{K}(\hat{a}_{m,n}(i))^2]\right]_P$$

Jusqu'à la convergence

4.3.5.3 Algorithme du gradient projeté modifié (Grd LQ-MAP-modif)

Dans cette partie, nous décrivons les modifications apportées à l'algorithme du gradient projeté présenté dans l'Algorithme 1 afin de tenir compte de la contrainte de somme égale à 1 des fractions d'abondance directement dans le calcul des dérivées. En effet, lors du calcul du gradient dans la Section 4.3.5.1, nous avons tenu compte de la structure spécifique de la matrice $\hat{\mathbf{S}}$, c-à-d du fait que les dernières lignes de cette matrice sont les produits deux à deux de ses premières lignes (voir l'équation (4.27)). Néanmoins, nous n'avons pas utilisé la structure de la matrice $\hat{\mathbf{A}}$, due à la contrainte somme égale à un des fractions d'abondance, dans le calcul du gradient par rapport aux paramètres $\hat{a}_m(l)$ fourni dans l'équation (4.19). Pour tenir compte de cette structure,

nous réécrivons la dernière fraction d'abondance $\hat{a}_L$ de chaque pixel i en fonction de toutes les autres fractions d'abondance de ce pixel : $\hat{a}_L(i) = 1 - \sum_{j=1}^{L-1} \hat{a}_j(i)$.
Par conséquent, en tenant compte de cette nouvelle écriture, la fonction de coût J devient comme suit

$$J = \frac{1}{2}||\mathbf{X} - \hat{\mathbf{A}}\hat{\mathbf{S}}||_F^2 - \eta R^*, \tag{4.35}$$

où R^* est le nouveau terme de régularisation défini par

$$\begin{aligned} R^* &= K\log\Gamma(\sum_{j=1}^{L}\hat{\theta}_j) - K\sum_{j=1}^{L}\log\Gamma(\hat{\theta}_j) + \sum_{j=1}^{L-1}\left[(\hat{\theta}_j - 1)\sum_{i=1}^{K}\log\hat{a}_j(i)\right] \\ &+ (\hat{\theta}_L - 1)\sum_{i=1}^{K}\log\Big(1 - \sum_{j=1}^{L-1}\hat{a}_j(i)\Big) + \sum_{j=1}^{L}\sum_{k=j}^{L}\left[K\log(\frac{2\hat{\vartheta}_{j,k}}{\pi}) - \frac{\hat{\vartheta}_{j,k}^2}{\pi}\sum_{i=1}^{K}(\hat{a}_{j,k}(i))^2\right], \end{aligned} \tag{4.36}$$

Ainsi, le calcul du gradient par rapport aux fractions d'abondance $\hat{a}_m(l)$, pour $m = 1, ..., L-1$, est donné par

$$\frac{\partial J}{\partial a_m(l)} = [\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X}]_l[\hat{\mathbf{s}}_m - \hat{\mathbf{s}}_L]^T - \eta\Big[\frac{(\hat{\theta}_m - 1)}{\hat{a}_m(l)} - \frac{(\hat{\theta}_L - 1)}{1 - \sum_{j=1}^{L-1}\hat{a}_j(l)}\Big], \quad \forall\ 1 \leq m \leq (L-1) \tag{4.37}$$

où $[\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X}]_l$ correspond à la ligne l de la matrice $(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})$.

Après la mise à jour des $(L-1)$ premières fractions d'abondance avec la formule (4.22), la dernière fraction d'abondance est mise à jour de la manière suivante :

$$\hat{a}_L(l) \leftarrow [1 - \sum_{j=1}^{L-1}\hat{a}_j(l)]_P. \tag{4.38}$$

Les calculs de $\frac{\partial J}{\partial \hat{a}_{m,n}(l)}$, $\frac{\partial J}{\partial \hat{s}_{pn}}$, $\frac{\partial J}{\partial \hat{\vartheta}_{m,n}}$, et $\frac{\partial J}{\partial \hat{\theta}_\ell}$, restent inchangés. Et par la suite, leurs règles de mise à jour sont données par les équations (4.23)-(4.26) décrites auparavant. L'illustration de l'algorithme est dans l'Algorithme 3.

Notons que la méthode ci-dessus peut aussi être utilisée en absence des informations a priori sur les distributions des paramètres de mélange pour la prise en compte de la structure de la matrice $\mathbf{A}$ dans le calcul du gradient. Il suffit de poser $\eta = 0$ dans l'Algorithme 3. Nous obtenons ainsi une version améliorée de l'algorithme Grd-LQ

initial de Meganem et al (2014b) présentée dans l'Algorithme 4.

Algorithme 3 : Grd LQ-MAP-modif

- Initialisation des paramètres $\hat{s}_{pn},\ \hat{a}_m(l),\ \hat{a}_{m,n}(\ell),\ \hat{\theta}_\ell$, et $\hat{\vartheta}_{m,n}$.

Faire

- Mise à jour des spectres des signaux sources :

$$\hat{s}_{pn} \leftarrow \left[\hat{s}_{pn} - \mu_s\Big[\ [\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{pn} + \sum_{j=1, j\neq p}^{L} \hat{s}_{jn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(jp)n} + 2\hat{s}_{pn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(pp)n}\Big]\right]_P$$

$$\hat{s}_{(jp)n} \leftarrow \hat{s}_{jn} \times \hat{s}_{pn}$$

- Mise à jour des fractions d'abondance :

$$\hat{a}_m(l) \leftarrow \left[\hat{a}_m(l) - \mu_a\Big[[\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X}]_l[\hat{\mathbf{s}}_m - \hat{\mathbf{s}}_L]^T - \eta[\frac{(\hat{\theta}_m - 1)}{\hat{a}_m(l)} - \frac{(\hat{\theta}_L - 1)}{1 - \sum_{j=1}^{L-1}\hat{a}_j(l)}]\Big]\right]_P \quad \forall\ 1 \leq m \leq (L-1)$$

$$\hat{a}_L(l) \leftarrow [1 - \textstyle\sum_{m=1}^{L-1}\hat{a}_m(l)]_P$$

- Mise à jour des coefficients quadratiques :

$$\hat{a}_{m,n}(\ell) \leftarrow \left[\hat{a}_{m,n}(\ell) - \mu_a\Big[\ [(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{(mn)\ell} + \eta\frac{2\hat{\vartheta}^2_{m,n}\hat{a}_{m,n}(\ell)}{\pi}\ \Big]\right]_P$$

- Mise à jour des paramètres des distributions :

$$\hat{\theta}_\ell \leftarrow \Big[\hat{\theta}_\ell + \mu_\theta\ \eta[K\psi(\sum_{j=1}^{L}\hat{\theta}_j) - K\psi(\hat{\theta}_\ell) + \sum_{i=1}^{K}\log\hat{a}_\ell(i)\ ]\Big]_P$$

$$\hat{\vartheta}_{m,n} \leftarrow \Big[\hat{\vartheta}_{m,n} + \mu_\vartheta\eta[\frac{K}{\hat{\vartheta}_{m,n}} - \frac{2\hat{\vartheta}_{m,n}}{\pi}\sum_{i=1}^{K}(\hat{a}_{m,n}(i))^2]\Big]_P$$

Jusqu'à la convergence

Algorithme 4 : Grd LQ-modif

- Initialisation des paramètres $\hat{s}_{pn},\ \hat{a}_m(l),\ \hat{a}_{m,n}(\ell)$.

Faire

- Mise à jour des spectres des signaux sources :

$$\hat{s}_{pn} \leftarrow \left[\hat{s}_{pn} - \mu_s\Big[\ [\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{pn} + \sum_{j=1, j\neq p}^{L} \hat{s}_{jn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(jp)n} + 2\hat{s}_{pn}[\hat{\mathbf{A}}^T(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})]_{(pp)n}\Big]\right]_P$$

$$\hat{s}_{(jp)n} \leftarrow \hat{s}_{jn} \times \hat{s}_{pn}$$

- Mise à jour des fractions d'abondance :

$$\hat{a}_m(l) \leftarrow \left[\hat{a}_m(l) - \mu_a\Big[[\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X}]_l[\hat{\mathbf{s}}_m - \hat{\mathbf{s}}_L]^T\Big]\right]_P \quad \forall\ 1 \leq m \leq (L-1)$$

$$\hat{a}_L(l) \leftarrow [1 - \textstyle\sum_{m=1}^{L-1} \hat{a}_m(l)]_P$$

- Mise à jour des coefficients quadratiques :

$$\hat{a}_{m,n}(\ell) \leftarrow \left[\hat{a}_{m,n}(\ell) - \mu_a\Big[[(\hat{\mathbf{A}}\hat{\mathbf{S}} - \mathbf{X})\hat{\mathbf{S}}^T]_{(mn)\ell}\Big]\right]_P$$

Jusqu'à la convergence

4.4 Résultats de simulations

Dans cette partie, nous présentons les résultats des simulations effectuées afin d'évaluer les performances des méthodes proposées en les comparant à la méthode Grd-LQ présentée dans (Meganem et al, 2014b), qui n'exploite pas les informations a priori.

Nous nous contentons de tester les algorithmes Grd-LQ, Grd LQ-MAP, Grd Mult-MAP et Grd LQ-MAP-modif pour un cas particulier du modèle linéaire-quadratique, appelé modèle bilinéaire, où les coefficients quadratiques des termes au carré $a_{j,j}(i)$ sont considérés nuls. Autrement dit, nous supposons qu'il n'y a pas de réflexion entre les matériaux de même nature. L'utilisation du modèle bilinéaire est faite simplement pour simplifier le problème et pour réduire le temps de calcul, sinon le principe reste le même qu'en utilisant le modèle LQ.

Cette partie est organisée comme suit : tout d'abord, nous commençons par présenter les critères de performances utilisés afin d'évaluer les performances des différents algorithmes testés, puis nous détaillons les résultats expérimentaux obtenus avec les différents types de données.

4.4.1 Critère de performance

Afin d'évaluer et comparer les performances des algorithmes proposés, nous avons utilisé le Rapport Signal sur Interférence RSI (SIR pour l'appellation en anglais Signal-to-Interference Ratio). Pour chaque signal source, fraction d'abondance et coefficient quadratique, le RSI s'écrit comme suit

$$RSI_{s_j} = 10\log_{10}\frac{\sum_{n=1}^{N} s_j(n)^2}{\sum_{n=1}^{N}(s_j(n)-\hat{s}_j(n))^2}, \qquad j = 1,\cdots,L \tag{4.39}$$

$$RSI_{a_j} = 10\log_{10}\frac{\sum_{i=1}^{K} a_j(i)^2}{\sum_{i=1}^{K}(a_j(i)-\hat{a}_j(i))^2}, \qquad j = 1,\cdots,L \tag{4.40}$$

$$RSI_{a_{j,k}} = 10\log_{10}\frac{\sum_{i=1}^{K} a_{j,k}(i)^2}{\sum_{i=1}^{K}(a_{j,k}(i)-\hat{a_{j,k}}(i))^2}, \;\; j = 1,\cdots,L-1, \;\; k = j+1,\cdots,L \tag{4.41}$$

Dans ces formules, " ˆ " se réfère aux valeurs estimées en tenant compte des facteurs d'échelle et de permutations.

Afin de réduire la quantité des résultats rapportés, nous calculons plutôt le RSI moyen sur toutes les sources de la manière suivante

$$RSI_s = \frac{1}{L}\sum_{j=1}^{L} RSI_{s_j}, \tag{4.42}$$

avec L est le nombre de sources.

De même, nous calculons le RSI moyen sur tous les coefficients de mélange (fractions d'abondance et coefficients quadratiques) :

$$RSI_a = \frac{1}{P}(\sum_{j=1}^{L} RSI_{a_j} + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} RSI_{a_{j,k}}), \tag{4.43}$$

où P est le nombre total des fractions d'abondance et des coefficients quadratiques.

4.4.2 Tests expérimentaux

Dans les tests effectués, nous avons considéré les données suivantes :

- **Test 1 :** les observations sont des mélanges artificiels générés à partir de deux ou trois spectres synthétiques et coefficients de mélange aléatoires mais réalistes.

- **Test 2 :** les observations sont des mélanges artificiels générés à partir de deux ou trois spectres de réflectance réels et coefficients de mélange aléatoires mais réalistes.

- **Test 3 :** les observations sont des mélanges artificiels générés à partir de trois spectres de réflectance réels, cartes d'abondance réelles et coefficients quadratiques aléatoires mais réalistes.

Pour les différents tests effectués, les images sont formées par 10×10 pixels. Les détails sont montrés dans les paragraphes suivants.

4.4.2.1 Test 1

Dans cette expérience, nous avons testé les algorithmes développés pour des mélanges artificiels de deux et trois spectres synthétiques. Pour cela, nous avons généré les fractions d'abondance et les coefficients quadratiques en utilisant les distributions de Dirichlet et demi-gaussienne présentées dans les équation (4.6) et (4.7), respectivement. Concernant le choix des paramètres de distribution de Dirichlet, nous avons considéré $\theta = [70, 70]$ pour le cas d'un mélange de deux matériaux et $\theta = [50, 50, 50]$ pour le cas d'un mélange de trois matériaux. Ces valeurs permettent d'avoir une scène très mélangée sans la présence de pixels purs. Le paramètre des distributions demi-gaussienne $\vartheta_{j,k}$ a été choisi égal à 8.35 afin d'obtenir des valeurs de $a_{j,k}$ distribuées dans l'intervalle $[0, 0.5]$. Concernant les spectres synthétiques, nous avons généré deux et trois spectres synthétiques, selon le cas, formés de $(N = 126)$ valeurs uniformément réparties dans l'intervalle $[0, 1]$.

Nous avons effectué 100 simulations de Monte Carlo pour chaque algorithme testé. Dans chaque simulation, nous générons de nouvelles valeurs pour les fractions d'abondance, les coefficients quadratiques, et les spectres synthétiques. Les paramètres à estimer ont été initialisés comme suit : les fractions d'abondance avec des valeurs aléatoires dans l'intervalle [0,1], que nous normalisons par la suite selon l'équation (4.28) afin que leur somme soit égale à 1, de même les sources avec des valeurs aléatoires dans l'intervalle $[0, 1]$, les coefficients quadratiques avec des valeurs aléatoires réparties dans l'intervalle $[0, 0.5]$, les paramètres θ_j avec de grandes valeurs comprises entre 50 et 80, les paramètres $\vartheta_{j,k}$ avec des grandes valeurs satisfaisant la contrainte $a_{j,k} \in [0, 0.5]$. Les pas d'apprentissage du gradient ont été fixés à 0.0005 pour μ_a et μ_s, et 0.01 pour μ_θ et μ_ϑ. Les meilleurs résultats ont été obtenus lorsque la valeur du paramètre de régularisation η se situait entre 0.0001 et 0.001.

Dans le Tableau 4.1, nous présentons les résultats obtenus par le calcul des valeurs moyennes et des écarts-types de tous les RSI sur les 100 simulations Monte Carlo pour les différentes méthodes développées.

Tableau 4.1: *Résultats du Test 1*

Performances \ Algorithme		**Grd-LQ**	**Grd LQ-MAP**	**Grd LQ-MAP-modif**	**Grd Mult-MAP**
Cas de 2 matériaux	Moyenne/Ecart-type RSI_a (dB)	11.65/5.74	30.80/2.31	36.37/5.07	32.56/6.57
	Moyenne/Ecart-type RSI_s (dB)	12.03/4.42	29.58/3.79	38.54/5.77	35.20/5.10
Cas de 3 matériaux	Moyenne/Ecart-type RSI_a (dB)	5.38/1.45	26.70/1.25	28.70/ 2.45	31.30/12.98
	Moyenne/Ecart-type RSI_s (dB)	11.54/1.45	26.25/2.44	28.74/3.41	35.21/15.94

Les résultats des simulations sur des données synthétiques montrent que les performances s'améliorent pour les différents algorithmes proposés par rapport à l'algorithme Grd-LQ lorsque nous mélangeons 2 ou 3 matériaux. Ceci prouve l'intérêt de la prise en compte des informations a priori pour mieux estimer les paramètres de mélange. De plus, en comparant les résultats obtenus par les deux algorithmes Grd LQ-MAP et Grd LQ-MAP-modif, nous remarquons que ceux obtenus par Grd LQ-MAP-modif sont

meilleurs, ce qui montre que les performances d'estimation s'améliorent en ajoutant la contrainte de somme égale à un dans le calcul du gradient. En outre, Grd LQ-MAP-modif a les meilleures performances d'estimation en le comparant avec les trois autres algorithmes dans le cas d'un mélange de 2 matériaux alors que Grd Mult-MAP a les meilleures performances d'estimation dans le cas d'un mélange de 3 matériaux.

4.4.2.2 Test 2

Dans cette deuxième expérience, notre but est de tester les algorithmes dans le cas où les observations sont des mélanges artificiels de deux ou trois spectres réels rencontrés en milieu urbain. Pour ce faire, nous avons utilisé la configuration décrite dans la partie Test 1 pour générer les fractions d'abondance et les coefficients quadratiques, alors que pour les spectres, nous avons choisi trois matériaux parmi les spectres disponibles dans la bibliothèque spectrale présentée dans le chapitre précédent (voir Figure 3.2) [1]. Ces spectres de réflectance sont mesurés pour 126 longueurs d'onde variant de 0.4 μm à 2.5 μm. Les spectres des matériaux choisis dans ces tests sont ceux de béton, pelouse et sol nu. Ces spectres sont présentés dans la Figure 4.1.

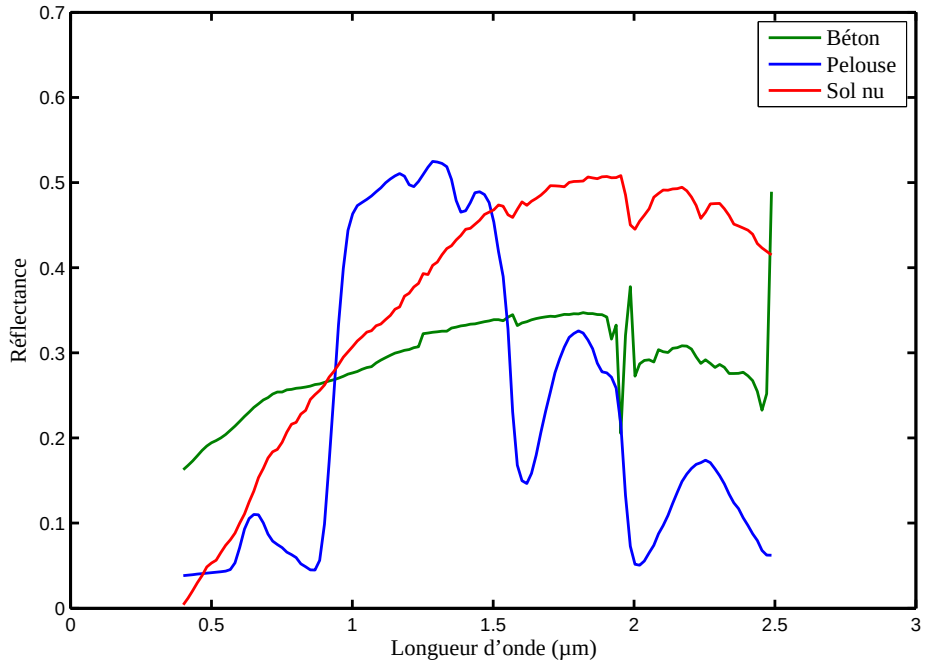


Figure 4.1: *Spectres utilisés dans le Test 2*

[1] (http://www.onera.fr/dota/memoires)

Nous avons effectué 100 simulations de Monte Carlo pour chaque algorithme testé. Dans chaque simulation, nous avons généré de nouvelles valeurs pour les paramètres de mélange (les fractions d'abondance et les coefficients quadratiques). Concernant les initialisations, elles ont été faites de la même façon que dans le Test 1. Le Tableau 5.2 affiche les résultats obtenus par le calcul de la valeur moyenne et l'écart-type de tous les RSI sur les 100 simulations Monte Carlo pour les différentes méthodes développées.

Tableau 4.2: *Résultats du Test 2*

Performances \ Algorithme		**Grd-LQ**	**Grd LQ-MAP**	**Grd LQ-MAP-modif**	**Grd Mult-MAP**
Cas de 2 matériaux	Moyenne/Ecart-type RSI_a	6.04/1.03	20.89/2.95	24.61/0.50	21.11/2.48
	Moyenne/Ecart-type RSI_s	14.73/0.36	20.85/3.12	24.87/1.67	22.07/3.62
Cas de 3 matériaux	Moyenne/Ecart-type RSI_a	2.95/0.89	15.34/2.38	13.60/1.09	22.28/4.66
	Moyenne/Ecart-type RSI_s	11.73/0.93	15.37/1.07	10.53/2.22	18.89/5.53

Nous pouvons remarquer que, dans la plupart des cas, les estimations des coefficients de mélange et des spectres réels par les algorithmes 1, 2 et 3 sont nettement meilleures que celles obtenues par l'algorithme Grd-LQ. Cela confirme que les informations a priori aident à mieux estimer les paramètres de mélange et les signaux sources. De même que dans le Test 1, nous constatons que l'algorithme Grd Mult-MAP a les meilleures performances en le comparant avec les trois autres algorithmes dans le cas d'un mélange de 3 matériaux alors que Grd LQ-MAP-modif donne les meilleurs résultats dans le cas d'un mélange de 2 matériaux. Nous remarquons dans ce test que les performances s'améliorent en ajoutant la contrainte de somme égale à un dans le calcul du gradient seulement dans le cas d'un mélange de 2 matériaux.

4.4.2.3 Test 3

Dans cette expérience, nous voulons utiliser une configuration plus réaliste que les configurations précédentes. Ainsi, nous avons généré une image hyperspectrale synthétique à partir d'une carte de classification d'occupation des sols d'une image satellitaire réelle de taille (180×180). Cette dernière, présentée dans la Figure 4.2, correspond à une partie de la commune de Tournefeuille située dans le département

de la Haute-Garonne.



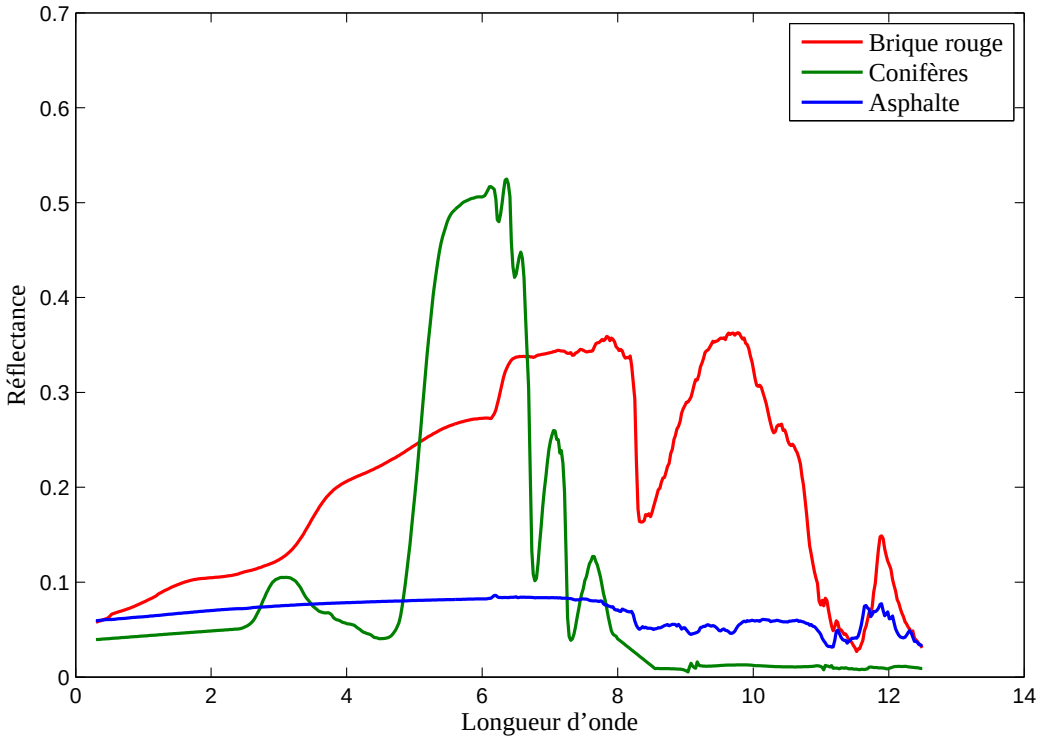
Figure 4.2: *Image satellitaire de Tournefeuille.*

La classification de l'image satellitaire a été réalisée à l'aide du logiciel de traitement d'images de télédétection ENVI en considérant la présence de trois grandes familles de matériaux seulement qui sont : le sol, la végétation et le mur. A partir de cette classification d'occupation des sols, nous avons déduit des cartes de fractions d'abondance réalistes des trois matériaux, contenant chacune (10×10) pixels. Ceci a été réalisé en calculant les valeurs moyennes des pixels de la classification sur une fenêtre glissante de taille (18×18) pixels. Afin d'obtenir un mélange linéaire-quadratique, nous avons généré les coefficients quadratiques $a_{j,k}$ selon la distribution demi-gaussienne présentée dans l'équation (4.7) avec un paramètre $\vartheta_{j,k} = 8.35$.
Dans ce test, nous avons utilisé trois spectres de matériaux provenant de la bibliothèque spectrale de Johns Hopkins University (JHU) [2]. Ces spectres sont mesurés pour 530 longueurs d'ondes variant de 0.3 μm à 12.5 μm et ils correspondent aux spectres de brique rouge pour le mur, asphalte de construction pour le sol, et conifères pour la végétation. Ils sont présentés dans la Figure 4.3.

Nous avons effectué 100 simulations de Monte Carlo pour chaque algorithme testé. Dans chaque simulation, nous avons initialisé les paramètres de la même manière que dans le Test 1.

[2](https://speclib.jpl.nasa.gov/documents/jhu_desc)

Figure 4.3: *Spectres utilisés dans le Test 3*



Dans le Tableau 4.3, nous présentons les valeurs moyennes et les écarts-types de tous les RSI obtenus pour les 100 simulations Monte Carlo pour les différentes méthodes proposées.

Tableau 4.3: *Résultats du Test 3*

RSI \ **Algorithme**	**Grd LQ-MAP**	**Grd LQ-MAP-modif**	**Grd Mult-MAP**
Moyenne/Ecart-type RSI_s	14.06/3.45	19.24/3.57	11.35/1.72
Moyenne/Ecart-type RSI_a	6.63/1.82	8.19/0.95	5.70/0.52

A partir de ces résultats de simulation, nous pouvons remarquer que les méthodes proposées permettent d'avoir une estimation acceptable des spectres même en se plaçant dans un contexte réaliste. Nous n'avons pas comparé nos méthodes avec l'algorithme Grd-LQ parce qu'il a divergé pour toutes les simulations.

4.5 Conclusions

Dans ce chapitre, nous avons proposé une nouvelle méthode de séparation de sources pour le démélange des images hyperspectrales urbaines mélangées selon le modèle linéaire-quadratique. Pour ce faire, nous avons modifié la méthode NMF en tenant compte des informations a priori sur les paramètres de mélange pour mieux les estimer. En se basant sur un critère de régularisation déduit de l'estimateur MAP, nous avons proposé différents algorithmes pour déterminer les paramètres inconnus du problème (les fractions d'abondance, les coefficients quadratiques, les signaux sources, et les paramètres des distributions). Les résultats des simulations sur différents types de données ont montré que les algorithmes développés ont permis une meilleure estimation en les comparants à l'algorithme Grd-LQ proposé dans (Meganem et al, 2014b). Ceci prouve l'intérêt de la prise en compte des informations a priori pour améliorer le résultat de séparation de sources.

Par ailleurs, notre méthode peut aussi être applicable pour le démélange d'autres mélanges LQ tel que le démélange des documents en recto-verso scannés où les paramètres des distributions a priori sont supposés approximativement connus (Duarte et al, 2011).

Comme travaux postérieurs, nous comptons étendre notre méthode par l'utilisation des méthodes MCMC pour l'estimation des inconnues du problème au lieu d'utiliser l'algorithme de gradient. De plus, nous allons travailler sur un algorithme du gradient à pas adaptatif afin de nous débarrasser du problème du choix du pas d'apprentissage de l'algorithme du gradient classique.

Chapitre 5

Une méthode d'ACI basée sur les Statistiques de Second Ordre pour les mélanges bilinéaires

Sommaire

5.1 Introduction 111
5.2 Modèle de mélange bilinéaire 112
5.3 Étude de la corrélation mutuelle des sources étendues . . 112
5.3.1 Cas de *vraies sources* de moyennes nulles 112
5.3.1.1 Corrélation entre $s_i(n)$ et $s_j(n)$ 113
5.3.1.2 Corrélation entre $s_i(n)$ et $s_j(n) \times s_k(n)$ 113
5.3.1.3 Corrélation entre $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ . . 114
5.3.2 Cas de *vraies sources* de moyennes non-nulles 115
5.3.2.1 Covariance entre $s_i(n)$ et $s_j(n)$ 115
5.3.2.2 Covariance entre $s_i(n)$ et $s_j(n) \times s_k(n)$ 115
5.3.2.3 Covariance entre $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ . . 116
5.4 Reformulation du modèle bilinéaire 117
5.5 Méthode proposée 120
5.5.1 Condition d'identifiabilité 120
5.5.2 Étapes de la méthode proposée 122
5.5.3 Identification des vraies sources centrées 123
5.5.4 Estimation des vrais coefficients de mélange 123
5.6 Résultats des simulations 125
5.6.1 Expérience 1 127
5.6.2 Expérience 2 128
5.6.3 Expérience 3 132
5.6.4 Expérience 4 133
5.7 Conclusions 135

5.1 Introduction

Dans ce chapitre, nous proposons une nouvelle méthode de séparation de sources adaptée aux mélanges linéaires-quadratiques et basée sur l'Analyse en Composantes Indépendantes (ACI). En effet, l'utilisation des méthodes d'ACI pour les mélanges Linéaires-Quadratiques (LQ) est généralement limitée par certaines propriétés des sources et/ou du mélange. Par exemple, certaines d'entre elles ne peuvent être utilisées que lorsque les sources sont à valeurs complexes et circulaires (Krob and Benidir, 1993; Abed-Meraim et al, 1996) ou binaires (Castella, 2008). D'autres méthodes sont adaptées uniquement à des mélanges déterminés (Deville and Hosseini, 2009; Hosseini and Deville, 2013; Almeida and Almeida, 2012), ce qui signifie généralement qu'elles n'exploitent pas une partie des informations utiles sur les observations lorsque les mélanges sont sur-déterminés. En outre, la plupart des méthodes d'ACI basées sur le modèle LQ existantes sont gourmandes en temps de calcul.

Dans ce travail, nous étudions un cas particulier du modèle LQ qui est le modèle bilinéaire dans une configuration sur-déterminée. Nous proposons une nouvelle méthode rapide de SAS, appelée B-SO-BSS (pour *Bilinear Second Order Blind Source Separation* en anglais), basée sur les Statistiques de Second Ordre (SSO) et la diagonalisation conjointe des matrices de corrélation des signaux observés centrés blanchis. Celle-ci est applicable dans le cas où les sources sont à valeurs réelles, auto-corrélées, mutuellement indépendantes et conjointement stationnaires au sens strict.

Le chapitre est structuré comme suit. Nous commençons par rappeler le modèle de mélange bilinéaire et étudier la corrélation entre les différentes *sources étendues*. Nous reformulons ensuite notre problème sur la base des résultats de cette étude afin de vérifier les conditions nécessaires pour l'utilisation des méthodes SSO, et par la suite, nous présentons les étapes de la méthode proposée. Enfin, nous exposons les résultats de simulation obtenus en utilisant des mélanges artificiels bruités et non-bruités de sources synthétiques et réelles.

5.2 Modèle de mélange bilinéaire

Nous rappelons que le modèle de mélange bilinéaire est un cas particulier du modèle LQ, où les coefficients quadratiques des termes au carré sont considérés nuls. En considérant K observations résultant d'un mélange bilinéaire de L sources, la relation entre les signaux observés et les signaux sources est comme suit

$$x_i(n) = \sum_{j=1}^{L} a_j(i)s_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)s_j(n)s_k(n), \tag{5.1}$$

où $x_i(n)$ est le $i^{ème}$ signal observé à l'indice d'échantillon n, $s_j(n)$ est le $j^{ème}$ signal source inconnu, $a_j(i)$ est le coefficient linéaire associé au $j^{ème}$ signal source et au $i^{ème}$ signal observé, $a_{j,k}(i)$ est le coefficient quadratique associé au $i^{ème}$ signal observé et résultant de l'interaction entre la $j^{ème}$ et la $k^{ème}$ sources.
Rappelons aussi que l'ensemble des *vraies sources* $s_j(n)$ et des *pseudo-sources* $s_j(n)s_k(n)$ est appelé *sources étendues*.

5.3 Étude de la corrélation mutuelle des sources étendues

Étant donné que le mélange est bilinéaire, il n'est pas formé seulement par les vraies sources comme dans le cas linéaire mais par l'ensemble des vraies sources et des pseudo-sources formant les sources étendues. En conséquence, il est nécessaire de vérifier, tout d'abord, les propriétés statistiques des sources étendues. Nous nous intéressons à deux scénarios : lorsque les vraies sources sont à moyennes nulles et lorsqu'elles ne le sont pas.

5.3.1 Cas de vraies sources de moyennes nulles

Dans cette partie, nous nous intéressons à étudier la corrélation entre les différentes sources étendues lorsque les vraies sources ont des moyennes nulles.
Tout d'abord, nous vérifions si les sources étendues sont à moyennes nulles dès lors que les vraies sources le sont.

En effet, les pseudo-sources sont définies par $s_j(n)s_k(n)$, avec $j \neq k$, et leurs facteurs $s_j(n)$ et $s_k(n)$ correspondent à des sources indépendantes et ainsi non-corrélées. Cela nous permet d'écrire

$$\begin{aligned} E\{s_j(n)s_k(n)\} &= E\{s_j(n)\}E\{s_k(n)\} \\ &= 0 \end{aligned} \tag{5.2}$$

où $E\{.\}$ représente l'espérance mathématique.
Ainsi, nous pouvons conclure que les pseudo-sources ont également des moyennes nulles, et par la suite, toutes les *sources étendues* ont des moyennes nulles.
Dans ce contexte, l'étude de la corrélation mutuelle entre toutes les sources étendues peut être réalisée en calculant leurs fonctions d'intercorrélation. Ces fonctions sont étudiées dans ce qui suit.

5.3.1.1 Corrélation entre $s_i(n)$ et $s_j(n)$

Lorsque deux *vraies sources* $s_i(n)$ et $s_j(n)$ sont indépendantes et de moyennes nulles, alors leur fonction d'intercorrélation est nulle pour tout décalage τ

$$R_{s_i,s_j}(\tau) = E\{s_i(n+\tau)s_j(n)\} = 0\ . \tag{5.3}$$

5.3.1.2 Corrélation entre $s_i(n)$ et $s_j(n) \times s_k(n)$

Si $(i \neq j)$ et $(i \neq k)$: étant donné que $s_i(n)$, $s_j(n)$ et $s_k(n)$ sont mutuellement indépendantes, nous pouvons écrire

$$\begin{aligned} R_{s_i,(s_j s_k)}(\tau) &= E\{s_i(n+\tau)(s_j(n)s_k(n))\} \\ &= E\{s_i(n+\tau)\}E\{s_j(n)\}E\{s_k(n)\} \\ &= 0\ . \end{aligned} \tag{5.4}$$

Ainsi, dans ce cas, $s_i(n)$ et $s_j(n) \times s_k(n)$ sont mutuellement non-corrélées.

Si $(i = j)$ ou $(i = k)$: le raisonnement est le même pour les deux cas. En

prenant par exemple $(i = j)$, nous obtenons

$$\begin{aligned} R_{s_i,(s_i s_k)}(\tau) &= E\{s_i(n+\tau)(s_i(n)s_k(n))\} \\ &= E\{s_i(n+\tau)s_i(n)\}E\{s_k(n)\} \\ &= 0\ . \end{aligned} \tag{5.5}$$

Nous remarquons que, dans ce cas aussi, $s_i(n)$ et $s_j(n) \times s_k(n)$ sont mutuellement non-corrélées.

5.3.1.3 Corrélation entre $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$

Si $(i \neq k)$, $(i \neq l)$, $(j \neq k)$ et $(j \neq l)$: l'indépendance entre $s_i(n)$, $s_j(n)$, $s_k(n)$ et $s_l(n)$ permet d'écrire

$$\begin{aligned} R_{(s_i s_j),(s_k s_l)}(\tau) &= E\{(s_i(n+\tau)s_j(n+\tau))(s_k(n)s_l(n))\} \\ &= E\{s_i(n+\tau)\}E\{s_j(n+\tau)\}E\{s_k(n)\}E\{s_l(n)\} \\ &= 0\ . \end{aligned} \tag{5.6}$$

Ainsi, $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ sont mutuellement non-corrélées.

Si $(i = (k \text{ ou } l))$ xor $(j = (k \text{ ou } l))$: le raisonnement est le même pour les différents cas. Par exemple, en considérant $(i = k)$ et donc $(j \neq (k \text{ et } l))$, l'indépendance entre les sources s_i, s_j et s_l nous permet d'écrire

$$\begin{aligned} R_{(s_i s_j),(s_i s_l)}(\tau) &= E\{(s_i(n+\tau)s_j(n+\tau))(s_i(n)s_l(n))\} \\ &= E\{s_i(n+\tau)s_i(n)\}E\{s_j(n+\tau)\}E\{s_l(n)\} \\ &= 0\ . \end{aligned} \tag{5.7}$$

Dans ce cas aussi, $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ sont mutuellement non-corrélées.

Ainsi, nous pouvons conclure que dans le cas où les *vraies sources* ont des moyennes nulles, toutes les *sources étendues* sont mutuellement non-corrélées.

5.3.2 Cas de vraies sources de moyennes non-nulles

Dans ce cas, en se basant sur la première ligne de l'équation (5.2), nous pouvons conclure que les *pseudo-sources* n'ont pas des moyennes nulles. En conséquence, la corrélation mutuelle entre les *sources étendues* doit être mesurée en utilisant leurs fonctions d'intercovariance.
En considérant que les *vraies sources* sont auto-corrélées de sorte que la fonction de covariance C_{s_i,s_i} soit non-nulle pour le décalage τ, nous présentons l'étude de la corrélation entre les différentes *sources étendues* dans ce qui suit.

5.3.2.1 Covariance entre $s_i(n)$ et $s_j(n)$

En considérant que les deux *vraies sources* $s_i(n)$ et $s_j(n)$ sont indépendantes, alors elles sont mutuellement non-corrélées et leur fonction d'intercovariance est nulle.

5.3.2.2 Covariance entre $s_i(n)$ et $s_j(n) \times s_k(n)$

Si $(i \neq j)$ et $(i \neq k)$: étant donné que $s_i(n)$, $s_j(n)$ et $s_k(n)$ sont mutuellement indépendantes, nous pouvons écrire

$$\begin{aligned} C_{s_i,(s_j s_k)}(\tau) &= E\{s_i(n+\tau)(s_j(n)s_k(n))\} - E\{s_i(n+\tau)\}E\{(s_j(n)s_k(n))\} \\ &= E\{s_i(n+\tau)\}E\{s_j(n)\}E\{s_k(n)\} - E\{s_i(n+\tau)\}E\{s_j(n)\}E\{s_k(n)\} \\ &= 0\ . \end{aligned} \tag{5.8}$$

Ainsi, dans ce cas, $s_i(n)$ et $s_j(n) \times s_k(n)$ sont mutuellement non-corrélées.

Si $(i = j)$ ou $(i = k)$: le raisonnement est similaire pour les deux cas. Par exemple, en considérant $(i = j)$, la fonction de covariance est exprimée comme suit

$$\begin{aligned} C_{s_i,(s_i s_k)}(\tau) &= E\{s_i(n+\tau)(s_i(n)s_k(n))\} - E\{s_i(n+\tau)\}E\{s_i(n)s_k(n)\} \\ &= E\{s_i(n+\tau)s_i(n)\}E\{s_k(n)\} - E\{s_i(n+\tau)\}E\{s_i(n)\}E\{s_k(n)\} \\ &= \Big(E\{s_i(n+\tau)s_i(n)\} - E\{s_i(n+\tau)\}E\{s_i(n)\}\Big)\ E\{s_k(n)\} \\ &= C_{s_i,s_i}(\tau)E\{s_k(n)\} \\ &\neq 0 \end{aligned} \tag{5.9}$$

Nous remarquons que, dans ce cas, la fonction de covariance est non-nulle, et ainsi, $s_i(n)$ et $s_j(n) \times s_k(n)$ sont mutuellement corrélées.

5.3.2.3 Covariance entre $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$

Si $(i \neq k)$, $(i \neq l)$, $(j \neq k)$ et $(j \neq l)$: l'indépendance entre $s_i(n)$, $s_j(n)$, $s_k(n)$ et $s_l(n)$ permet d'écrire

$$\begin{aligned} C_{(s_i s_j),(s_k s_l)}(\tau) &= E\{(s_i(n+\tau)s_j(n+\tau))(s_k(n)s_l(n))\} \\ &- E\{(s_i(n+\tau)s_j(n+\tau))\}E\{s_k(n)s_l(n)\} \\ &= E\{(s_i(n+\tau)s_j(n+\tau))\}E\{s_k(n)s_l(n)\} \\ &- E\{(s_i(n+\tau)s_j(n+\tau))\}E\{s_k(n)s_l(n)\} \\ &= 0\ . \end{aligned} \tag{5.10}$$

Ainsi, dans ce cas, $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ sont mutuellement non-corrélées.

Si $(i = (k \text{ ou } l))$ xor $(j = (k \text{ ou } l))$: le raisonnement est le même pour tous les cas. Par exemple, en considérant $(i = k)$ et donc $(j \neq (k \textit{ et } l))$, alors la fonction de covariance s'écrit

$$\begin{aligned} C_{(s_i s_j),(s_i s_l)}(\tau) &= E\{(s_i(n+\tau)s_j(n+\tau))(s_i(n)s_l(n))\} \\ &\quad -E\{s_i(n+\tau)s_j(n+\tau)\}E\{s_i(n)s_l(n)\} \\ &= \Big(E\{s_i(n+\tau)s_i(n)\}E\{s_j(n+\tau)\}E\{s_l(n)\}\Big) \\ &\quad -\Big(E\{s_i(n+\tau)\}E\{s_j(n+\tau)\}E\{s_i(n)\}E\{s_l(n)\}\Big) \\ &= C_{s_i,s_i}(\tau)E\{s_j(n+\tau)\}E\{s_l(n)\} \\ &\neq 0 \end{aligned} \tag{5.11}$$

Nous remarquons que, dans ce cas, $s_i(n) \times s_j(n)$ et $s_k(n) \times s_l(n)$ sont corrélées.

Nous pouvons conclure que dans le cas où les *vraies sources* ont des moyennes non-nulles, il existe des situations où certaines des *sources étendues* sont mutuellement corrélées. Pour cela, dans ce qui suit, nous avons l'intention de reformuler notre problème afin d'aboutir à un nouveau modèle bilinéaire exprimé en fonction des nouvelles sources étendues mutuellement non-corrélées.

5.4 Reformulation du modèle bilinéaire

Dans cette partie, nous proposons une méthode pour reformuler le modèle de mélange bilinéaire original afin d'avoir un nouveau modèle de mélange bilinéaire exprimé en fonction de nouvelles sources étendues mutuellement non-corrélées. L'utilisation de cette méthode n'est nécessaire que lorsque les vraies sources sont à moyennes non-nulles.

Étant donné que les *vraies sources* sont stationnaires au sens strict, alors les valeurs moyennes $E\{s_j(n)\}$ des *vraies sources* $s_j(n)$ ne dépendent pas de l'échantillon n. La valeur moyenne de s_j sera désignée par $\bar{s}_j$ par la suite. En conséquence, la version centrée de $s_j(n)$ peut être exprimée par $\widetilde{s}_j(n) = s_j(n) - \bar{s}_j$.

Le modèle bilinéaire (5.1) peut être reformulé comme suit

$$\begin{aligned} x_i(n) &= \sum_{j=1}^{L} a_j(i)(\widetilde{s}_j(n) + \bar{s}_j) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)(\widetilde{s}_j(n) + \bar{s}_j)(\widetilde{s}_k(n) + \bar{s}_k) \\ &= \sum_{j=1}^{L} a_j(i)\widetilde{s}_j(n) + \sum_{j=1}^{L} a_j(i)\bar{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\widetilde{s}_j(n)\widetilde{s}_k(n) \\ &\quad + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)(\bar{s}_k\widetilde{s}_j(n) + \bar{s}_j\widetilde{s}_k(n)) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k. \end{aligned} \tag{5.12}$$

Le quatrième terme du membre de droite de (5.12), noté F dans la suite, peut être réécrit comme suit

$$\begin{aligned} F &= \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\widetilde{s}_k(n) \\ &= \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n) + \sum_{k=1}^{L-1}\sum_{j=k+1}^{L} a_{k,j}(i)\bar{s}_k\widetilde{s}_j(n), \end{aligned} \tag{5.13}$$

où le dernier terme ci-dessus est obtenu juste en inversant les rôles des indices j et k. Nous introduisons alors les coefficients $a_{j,k}(i)$ avec $j > k$, définis par rapport aux vrais coefficients de (5.1), comme $a_{j,k}(i) = a_{k,j}(i)$.

Ceci permet de déduire

$$\begin{aligned} F &= \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n) + \sum_{k=1}^{L-1}\sum_{j=k+1}^{L} a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n) \\ &= \left(\sum_{j=1}^{L-1}\sum_{k=j+1}^{L} \bigcup \sum_{k=1}^{L-1}\sum_{j=k+1}^{L}\right) a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n). \end{aligned} \tag{5.14}$$

La somme ci-dessus contient toutes les combinaisons possibles de $j \in [1, L]$, $k \in [1, L]$ telles que $j \neq k$. Elle peut alors être réécrite comme suit

$$F = \sum_{j=1}^{L} \sum_{k=1, k\neq j}^{L} a_{j,k}(i)\bar{s}_k\widetilde{s}_j(n). \tag{5.15}$$

En remplaçant (5.15) dans (5.12), nous obtenons

$$\begin{aligned} x_i(n) &= \sum_{j=1}^{L} \Big(a_j(i) + \sum_{k=1, k\neq j}^{L} a_{j,k}(i)\bar{s}_k\Big)\widetilde{s}_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\widetilde{s}_j(n)\widetilde{s}_k(n) \\ &\quad + \sum_{j=1}^{L} a_j(i)\bar{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k, \end{aligned} \tag{5.16}$$

ce qui permet de déduire

$$x_i(n) = \sum_{j=1}^{L} \widetilde{a}_j(i)\widetilde{s}_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\widetilde{s}_j(n)\widetilde{s}_k(n) + C_i, \tag{5.17}$$

où $\widetilde{a}_j(i)$ sont les coefficients linéaires dans le nouveau modèle de mélange bilinéaire, et ils sont définis par

$$\widetilde{a}_j(i) = a_j(i) + \sum_{k=1, k\neq j}^{L} a_{j,k}(i)\bar{s}_k, \tag{5.18}$$

et la constante C_i est définie par

$$C_i = \sum_{j=1}^{L} a_j(i)\bar{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k. \tag{5.19}$$

Étant donné que les *vraies sources centrées* $\widetilde{s}_j(n)$ et $\widetilde{s}_k(n)$ sont indépendantes et à moyennes nulles, alors à partir de (5.17), nous pouvons conclure que la valeur moyenne de $x_i(n)$ est égale à $\bar{x}_i = C_i$. En conséquence, la version centrée de $x_i(n)$ peut être exprimée comme suit

$$\begin{aligned} \widetilde{x}_i(n) &= x_i(n) - \bar{x}_i \\ &= \sum_{j=1}^{L} \widetilde{a}_j(i)\widetilde{s}_j(n) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\widetilde{s}_j(n)\widetilde{s}_k(n). \end{aligned} \tag{5.20}$$

Nous pouvons constater que les observations centrées forment un nouveau modèle de mélange bilinéaire exprimé en fonction des *vraies sources centrées*, bien que, dans ce cas, les coefficients de mélange de la partie linéaire ne soient pas les mêmes que ceux du mélange original.

En nous basant sur les résultats de l'étude effectuée dans la section 5.3.1, nous pouvons conclure que les nouvelles sources étendues $\widetilde{s}_j(n)$ et $\widetilde{s}_j(n)\widetilde{s}_k(n)$ sont toutes mutuellement non-corrélées.

Le nouveau modèle de mélange bilinéaire (5.20) peut être réécrit sous la forme matricielle suivante

$$\widetilde{\mathbf{x}}(n) = \widetilde{\mathbf{A}}\widetilde{\mathbf{s}}(\mathbf{n}), \tag{5.21}$$

où $\widetilde{\mathbf{x}}(n) = [\widetilde{x}_1(n), \cdots, \widetilde{x}_K(n)]^T$ est le vecteur contenant les K signaux observés centrés à l'échantillon n, $\widetilde{\mathbf{s}}(n) = [\widetilde{s}_1(n), \cdots, \widetilde{s}_L(n), \widetilde{s}_1(n)\widetilde{s}_2(n), \cdots, \widetilde{s}_{L-1}(n)\widetilde{s}_L(n)]^T$ est le vecteur contenant toutes les *sources centrées étendues* à l'échantillon n, et la matrice de mélange $\widetilde{\mathbf{A}}$, contient les coefficients de mélange $\widetilde{a}_j(i)$ et $a_{j,k}(i)$, tel que

$$\widetilde{\mathbf{A}} = \begin{pmatrix} \widetilde{a}_1(1) & \cdots & \widetilde{a}_L(1) & a_{1,2}(1) & \cdots & a_{L-1,L}(1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \widetilde{a}_1(K) & \cdots & \widetilde{a}_L(K) & a_{1,2}(K) & \cdots & a_{L-1,L}(K) \end{pmatrix}. \tag{5.22}$$

Il convient de noter que le mélange bilinéaire peut être considéré comme un mélange linéaire de $L(L+1)/2$ *sources centrées étendues*. Dans cette étude, nous supposons que $K \geq L(L+1)/2$ pour que le mélange linéaire reformulé ne soit pas sous-déterminé.

Dans ce qui suit, nous présentons la méthode que nous proposons pour estimer les vraies sources centrées et extraire les vrais coefficients de mélange. Nous considérons le cas où les *vraies sources* sont à moyennes nulles comme un cas particulier de la méthode proposée en considérant $\widetilde{\mathbf{s}} = \mathbf{s}$, $\widetilde{\mathbf{A}} = \mathbf{A}$, et $\widetilde{\mathbf{x}} = \mathbf{x}$.

5.5 Méthode proposée

D'après les résultats précédents, les observations centrées $\widetilde{x}_i(n)$ forment un mélange linéaire déterminé (ou surdéterminé) d'un ensemble de signaux sources non-corrélés, c'est-à-dire, les *sources centrées étendues*, liées aux vraies sources centrées. Par ailleurs, nous considérons dans la suite le cas où ces signaux sources modifiés sont autocorrélés et avec des fonctions d'autocorrélation différentes. Cette nouvelle configuration du mélange satisfait aux hypothèses proposées dans la littérature pour appliquer les méthodes de SAS linéaires basées sur les SSO, telles que la méthode AMUSE (pour *Algorithm for Multiple Unknown Signal Extraction* en anglais) (Tong et al, 1990) ou encore sa version améliorée, la méthode SOBI (pour *Second Order Blind Identification* en anglais) (Belouchrani et al, 1997). Cela nous permet de développer des versions étendues de ces méthodes classiques initialement proposées pour les mélanges linéaires, afin de traiter les mélanges bilinéaires. En particulier, nous proposons une nouvelle extension de la méthode SOBI, appelée B-SO-BSS (pour *Bilinear Second-Order Blind Source Separation* en anglais).

Dans ce qui suit, nous vérifions d'abord la condition d'identifiabilité du mélange, et nous présentons ensuite les différentes étapes effectuées pour estimer les vraies sources centrées et extraire les coefficients de mélange.

5.5.1 Condition d'identifiabilité

Une étape nécessaire de l'étude de la méthode proposée est de vérifier la condition d'identifiabilité du mélange. Comme la méthode SOBI, la méthode proposée utilise plusieurs matrices de corrélation des observations centrées blanchies pour différents décalages non-nuls $\tau_i \in \{\tau_1, ..., \tau_m\}$.

Pour un décalage τ_i, la matrice de corrélation des observations centrées blanchies $\mathbf{z}(n) = \mathbf{W}\widetilde{\mathbf{x}}(n)$(où $\mathbf{W}$ est une matrice de blanchiment) est exprimée comme suit

$$R_{\mathbf{z}}(\tau_i) = \mathbf{U} R_{\widetilde{\mathbf{s}}}(\tau_i) \mathbf{U}^T, \tag{5.23}$$

où $\mathbf{U}$ représente une matrice orthogonale, et $R_{\widetilde{\mathbf{s}}}(\tau_i)$ représente la matrice de corrélation des *sources centrées étendues* associée au décalage τ_i. La matrice $R_{\widetilde{\mathbf{s}}}(\tau_i)$ est une matrice diagonale puisque les *sources centrées étendues* sont mutuellement non-corrélées.

Considérons le théorème suivant :

Théorème : soient $\tau_i = \{\tau_1, \cdots, \tau_m\}$, m décalages non-nuls et $\mathbf{V}$ une matrice orthogonale, tels que :

$$\forall 1 \leq i \leq m \;\; \mathbf{V}^T R_{\mathbf{z}}(\tau_i)\mathbf{V} = diag[d_1(i), \cdots, d_{L(L+1)/2}(i)] \tag{5.24}$$

$$\forall 1 \leq j \neq k \leq L(L+1)/2, \;\; \exists i, 1 \leq i \leq m \;\; d_j(i) \neq d_k(i). \tag{5.25}$$

Dans ce cas, $\mathbf{U}$ et $\mathbf{V}$ sont *essentiellement* égales, c'est-à-dire qu'elles sont égales à une multiplication par une matrice $\mathbf{P}$ près, telle que $\mathbf{U} = \mathbf{VP}$, où $\mathbf{P}$ contient un élément non-nul dans chaque ligne et chaque colonne, dont la valeur est égale à ± 1.
Ce théorème permet de fournir une condition d'unicité pour la matrice $\mathbf{U}$ et, par conséquent, pour la matrice de mélange $\widetilde{\mathbf{A}}$.

Notons que la matrice de mélange ne peut être identifiée que lorsque les *sources centrées étendues* ont des spectres normalisés distincts, c'est-à-dire quand les versions normalisées des éléments diagonaux de $R_{\widetilde{\mathbf{s}}}(\tau_i)$ sont distinctes au moins pour un décalage τ_i non-nul. Dans ce cas, il est possible de trouver un ou plusieurs décalages τ_i non-nuls satisfaisant la condition du théorème (Belouchrani et al, 1997).

Il convient de noter que lorsque l'une des *vraies sources centrées* $\widetilde{s}_i(n)$ n'est pas auto-corrélée, alors toutes les *pseudo-sources* $\widetilde{s}_i(n)\widetilde{s}_j(n)$ avec $i \neq j$ ne sont pas non plus auto-corrélées, de sorte que toutes ces *sources centrées étendues* ont des spectres normalisés identiques (et constants dans ce cas). Ainsi, une condition nécessaire pour l'identifiabilité du mélange est que toutes les *vraies sources centrées* doivent être auto-corrélées. Notons que, dans la méthode classique SOBI linéaire, au plus une source peut être temporellement non-corrélée.

5.5.2 Étapes de la méthode proposée

La méthode (B-SO-BSS) proposée permet d'estimer les *vraies sources* centrées à des indéterminations d'échelle et de permutation près.

Les étapes de l'algorithme sont comme suit :

Algorithme 5 : B-SO-BSS

1. Estimer la matrice de corrélation pour le décalage nul $R_{\widetilde{\mathbf{x}}}(0)$ des signaux observés centrés $\widetilde{\mathbf{x}}(n)$.

2. Calculer la matrice de blanchiment $\mathbf{W} = \mathbf{D}^{-1/2}\mathbf{E}^T$ où $\mathbf{D}$ et $\mathbf{E}$ sont, respectivement, la matrice diagonale contenant les valeurs propres de l'estimée de $R_{\widetilde{\mathbf{x}}}(0)$ sur sa diagonale, et la matrice des vecteurs propres de norme unitaire de l'estimée de $R_{\widetilde{\mathbf{x}}}(0)$.

3. Blanchir les signaux observés centrés $\widetilde{\mathbf{x}}(n)$: $\mathbf{z}(n) = \mathbf{W}\widetilde{\mathbf{x}}(n)$.

4. Estimer les matrices de corrélation $R_{\mathbf{z}}(\tau_i)$ de $\mathbf{z}$, où $\tau_i = \{\tau_1, \cdots, \tau_m\}$ sont m décalages choisis.

5. Effectuer la diagonalisation conjointe des estimations des matrices $R_{\mathbf{z}}(\tau_i)$ pour obtenir une estimation $\hat{\mathbf{U}}$ d'une matrice orthogonale $\mathbf{U}$ telle que $R_{\mathbf{z}}(\tau_i) = \mathbf{U}R_{\widetilde{\mathbf{s}}}(\tau_i)\mathbf{U}^T$ où $\tau_i = \{\tau_1, \cdots, \tau_m\}$. Rappelons que la diagonalisation conjointe est un processus itératif dont le but est de déterminer une matrice $\mathbf{U}$ de telle sorte que les matrices $\mathbf{U}R_{\mathbf{z}}(\tau_i)\mathbf{U}^T$ soient les plus proches possible de matrices diagonales.

6. Calculer $\hat{\mathbf{s}}(n) = \hat{\mathbf{U}}^T\mathbf{z}(n)$ qui permet d'obtenir une estimation du vecteur des *sources centrées étendues* $\widetilde{\mathbf{s}}(n)$, à des indéterminations de permutation et d'échelle près.

7. Calculer $\hat{\mathbf{A}} = \mathbf{W}^{\dagger}\hat{\mathbf{U}}$ (où $\dagger$ représente la pseudo-inverse) qui permet d'avoir une estimation de la matrice $\widetilde{\mathbf{A}}$, à des indéterminations de permutation et d'échelle près.

8. Identifier les L *vraies sources centrées* estimées parmi toutes les *sources centrées étendues* estimées. Cette étape est détaillée dans ce qui suit dans la Sous-section 5.5.3.

5.5.3 Identification des vraies sources centrées

Les premières étapes de la méthode proposée fournissent un ensemble $\hat{\mathbf{s}}(n)$ de signaux qui correspondent aux estimations des $L(L+1)/2$ *sources centrées étendues* non-ordonnées à des indéterminations de permutation et d'échelle près. En considérant, par exemple, que $\hat{s}_j(n)$ et $\hat{s}_k(n)$ correspondent à l'estimation de deux *vraies sources centrées* et $\hat{s}_i(n)$ correspond à leur produit, alors, $\hat{s}_i(n)$ est idéalement proportionnel à $\hat{s}_j(n)\times\hat{s}_k(n)$. En conséquence, la valeur absolue du coefficient de corrélation entre $\hat{s}_i(n)$ et $\hat{s}_j(n) \times \hat{s}_k(n)$ doit être proche de la valeur 1. Ainsi, en calculant les coefficients de corrélation pour tous les triplets possibles $\{i, j, k\}, i \neq j \neq k$, il est possible d'identifier les *vraies sources centrées* estimées à partir des *sources centrées étendues* estimées. Notons que, dans le cas particulier où $L = 2$, il est également possible d'identifier les *vraies sources centrées* estimées à partir de l'estimation des 3 *sources centrées étendues* en utilisant un critère de mesure de l'indépendance statistique, comme l'information mutuelle. En effet, $\widetilde{s}_1$ et $\widetilde{s}_2$ sont mutuellement indépendantes mais $\widetilde{s}_1 \times \widetilde{s}_2$ n'est pas indépendant de $\widetilde{s}_1$ et $\widetilde{s}_2$.

5.5.4 Estimation des vrais coefficients de mélange

Une étape supplémentaire peut être ajoutée à la méthode proposée pour estimer les vrais coefficients de mélange. En effet, de nombreuses applications de SAS visent à estimer uniquement les *vraies sources* centrées. Cependant, dans certaines applications (comme dans la télédétection), il est nécessaire d'estimer les coefficients de mélange.

Dans le cas où les *vraies sources* ont des moyennes non-nulles, notre méthode (Algorithme 5) donne une estimation de la matrice $\widetilde{\mathbf{A}}$ (à une permutation et une matrice diagonale près) qui est constituée par

- des colonnes contenant les coefficients quadratiques $a_{j,k}(i)$, comme dans la matrice originale $\mathbf{A}$,
- des colonnes contenant les coefficients linéaires modifiés $\widetilde{a}_j(i)$, qui sont différents des coefficients $a_j(i)$, et qui sont définis par (5.18).

Dans la suite, nous proposons une méthode pour estimer la matrice $\mathbf{A}$ (aux indéterminations classiques près) à partir de l'estimation de la matrice $\widetilde{\mathbf{A}}$.

En effet, l'équation (5.18) permet de déduire

$$a_j(i) = \widetilde{a}_j(i) - \sum_{k=1,k\neq j}^{L} a_{j,k}(i)\bar{s}_k. \tag{5.26}$$

Ensuite, en insérant (5.26) dans (5.19), nous obtenons

$$\begin{aligned}
C_i &= \bar{x}_i \\
&= \sum_{j=1}^{L}[\widetilde{a}_j(i) - \sum_{k=1,k\neq j}^{L} a_{j,k}(i)\bar{s}_k]\bar{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k \\
&= \sum_{j=1}^{L}\widetilde{a}_j(i)\bar{s}_j - \sum_{j=1}^{L}\sum_{k=1,k\neq j}^{L} a_{j,k}(i)\bar{s}_k\bar{s}_j + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k \\
&= \sum_{j=1}^{L}\widetilde{a}_j(i)\bar{s}_j - 2\sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k \\
&= \sum_{j=1}^{L}\widetilde{a}_j(i)\bar{s}_j - \sum_{j=1}^{L-1}\sum_{k=j+1}^{L} a_{j,k}(i)\bar{s}_j\bar{s}_k \\
&= \sum_{j=1}^{L}\frac{\widetilde{a}_j(i)}{d_j}(d_j\bar{s}_j) + \sum_{j=1}^{L-1}\sum_{k=j+1}^{L}\frac{a_{j,k}(i)}{d_{j,k}}(-d_{j,k}\bar{s}_j\bar{s}_k),
\end{aligned} \tag{5.27}$$

où d_j et $d_{j,k}$ sont des facteurs d'échelle inconnus, représentant l'indétermination du facteur d'échelle pour les *sources centrées étendues* et les colonnes de la matrice $\widetilde{\mathbf{A}}$ estimées par notre Algorithme 5.

Le résultat ci-dessus peut être exprimé sous la forme matricielle suivante

$$\mathbf{c} = \widetilde{\mathbf{A}}_1\mathbf{e}, \tag{5.28}$$

où $\mathbf{c} = [C_1, \cdots, C_K]^T$, $\mathbf{e} = [d_1\bar{s}_1, \cdots, d_L\bar{s}_L, (-d_{1,2}\bar{s}_1\bar{s}_2), \cdots, (-d_{L-1,L}\bar{s}_{L-1}\bar{s}_L)]^T$, et $\widetilde{\mathbf{A}}_1$ est le résultat de la division des colonnes de la matrice $\widetilde{\mathbf{A}}$ par les facteurs d'échelle inconnus d_j et $d_{j,k}$. Autrement dit, $\widetilde{\mathbf{A}}_1$ correspond à la matrice $\hat{\mathbf{A}}$ fournie par notre Algorithme 5 à des erreurs d'estimation près. Notons que les colonnes de la matrice estimée $\hat{\mathbf{A}}$ peuvent également être permutées. Cependant, en utilisant la méthode appliquée dans l'étape 8 de l'Algorithme 5, nous pouvons identifier les colonnes contenant les coefficients linéaires et les colonnes contenant les coefficients quadratiques.

Il est alors possible de réarranger les colonnes de la matrice estimée de sorte que les

L premières colonnes correspondent à la partie linéaire (leur ordre n'est pas vraiment important) et que les autres colonnes soient correctement associées à ces L premières colonnes.
Nous précisons que $\mathbf{c}$ peut être obtenu en estimant les moyennes d'observations. Ainsi, $\mathbf{e}$ peut être déduit comme suit

$$\mathbf{e} = \widetilde{\mathbf{A}}_1^{\dagger}\mathbf{c}, \tag{5.29}$$

où † représente la pseudo-inverse.

En outre, (5.26) peut être réécrite sous la forme

$$a_j(i)\bar{s}_j = \frac{\widetilde{a}_j(i)}{d_j}(d_j\bar{s}_j) + \sum_{k=1,k\neq j}^{L} \frac{a_{j,k}(i)}{d_{j,k}}(-d_{j,k}\bar{s}_j\bar{s}_k). \tag{5.30}$$

Comme mentionné ci-dessus, $\frac{\widetilde{a}_j(i)}{d_j}$ et $\frac{a_{j,k}(i)}{d_{j,k}}$ $(\forall i = 1, \cdots, K)$ correspondent aux colonnes de $\widetilde{\mathbf{A}}_1$, estimées en utilisant l'étape 7 de l'Algorithme 5, et $(d_j\bar{s}_j)$ et $(-d_{j,k}\bar{s}_j\bar{s}_k)$ correspondent aux éléments de $\mathbf{e}$, estimés en utilisant (5.29). Par conséquent, les coefficients $a_j(i)$ $\forall i = 1, \cdots, K$ peuvent être estimés à des facteurs d'échelle inconnus $\bar{s}_j$ près selon l'équation (5.30). En d'autres termes, cette approche permet d'estimer les colonnes de la vraie matrice $\mathbf{A}$ contenant les coefficients linéaires à des facteurs d'échelle près. Notons que les colonnes de cette matrice contenant les coefficients quadratiques sont estimés directement par l'étape 7 de l'Algorithme 5 (à des facteurs d'échelle près).

5.6 Résultats des simulations

Dans cette section, nous présentons et discutons les résultats obtenus par la méthode B-SO-BSS proposée dans deux configurations différentes : en utilisant un seul décalage $(\tau_i = 1)$, et en utilisant 4 décalages $(\tau_i = \{1, 2, 3, 4\})$. Nous ajoutons aux étapes de l'Algorithme 5 l'étape supplémentaire décrite en Sous-section 5.5.4, permettant l'estimation des vrais coefficients du mélange. Nous présentons les résultats obtenus lorsque les sources n'ont pas des moyennes nulles, puisque nous avons trouvé les mêmes performances pour les deux cas lorsque les moyennes sont nulles et lorsqu'elles ne le sont pas.

Dans les simulations effectuées, les données traitées sont non-négatives. Ainsi, il est possible de comparer les résultats obtenus à ceux obtenus par l'algorithme *NMF-Grd-LQ* présenté dans (Meganem et al, 2014b). Rappelons que cette méthode est une généralisation de la NMF aux mélanges LQ, et qu'elle exploite donc la non-négativité des données impliquées dans le mélange. Notons que la version initiale de l'algorithme *NMF-Grd-LQ* a été développée pour le démélange de pixels en observation de la Terre et qu'elle suppose que la somme des coefficients linéaires est égale à 1 et que les coefficients de mélange quadratiques sont inférieurs à 0.5. Cependant, nos simulations ne concernent pas l'observation de la Terre, et par conséquent, la méthode NMF-Grd-LQ a été modifiée pour ne pas imposer ces contraintes. Afin d'évaluer et comparer les performances des méthodes développées, nous avons calculé les deux critères de RSI et d'EQMN décrits dans les chapitres précédents dans les équations (3.24), (3.25), (4.42) et (4.43).

Dans les simulations effectuées, nous avons considéré les données suivantes :

- **Expérience 1 :** les observations sont des mélanges artificiels générés à partir de deux signaux sources synthétiques et coefficients de mélange aléatoires.
- **Expérience 2 :** les observations sont des mélanges artificiels générés à partir de deux signaux sources réels et coefficients de mélange aléatoires.
- **Expérience 3 :** les observations sont des mélanges artificiels *bruités* générés à partir de deux signaux sources synthétiques et coefficients de mélange aléatoires.
- **Expérience 4 :** les observations sont des mélanges artificiels *bruités* générés à partir de deux signaux sources réels et coefficients de mélange aléatoires.

5.6.1 Expérience 1

Nous avons considéré des mélanges artificiels de signaux sources synthétiques. Les coefficients de mélange ont été générés d'une façon aléatoire avec des valeurs uniformément réparties entre 0 et 1. La génération des signaux sources a été réalisée comme suit : dans un premier temps, nous avons généré deux signaux i.i.d. $e_1(n)$ et $e_2(n)$ uniformément distribués sur $[0, 1]$, par la suite, nous les avons filtrés en utilisant deux filtres auto-régressifs de premier ordre afin qu'ils soient auto-corrélés selon le modèle $s_i(n) = e_i(n) + \rho_i s_i(n-1)$. Les paramètres choisis sont $\rho_1 = 0.7$ et $\rho_2 = 0.5$. Les tests ont été répétés en utilisant différents nombres d'échantillons N : 10000, 1000 et 100. Enfin, nous avons généré trois observations en utilisant le modèle bilinéaire.

Pour chaque test, nous avons effectué 100 simulations de Monte Carlo et, dans chaque simulation, nous avons modifié les réalisations des signaux sources et les coefficients du mélange.

Les valeurs moyennes des RSI et des EQMN des signaux sources et de tous les coefficients du mélange, ainsi que le temps CPU [1], calculés pour 100 simulations, sont présentés dans le Tableau 5.1.
Nous pouvons remarquer que notre méthode conduit aux meilleurs résultats d'estimation des sources et des coefficients du mélange. En terme de temps d'exécution, la méthode proposée est beaucoup plus rapide que NMF-Grd-LQ.
De plus, pour évaluer la qualité de l'estimation des sources, les vraies sources et leurs estimées obtenues par les différentes méthodes sont comparées en Figure 5.1. L'exemple présenté en Figure 5.1 correspond à des sources de $N = 100$ échantillons, mélangées par une matrice aléatoire $\mathbf{A_{Exp1}}$ donnée par

$$\mathbf{A_{Exp1}} = \begin{pmatrix} 0.0075 & 0.1461 & 0.3237 \\ 0.1221 & 0.4249 & 0.0723 \\ 0.4813 & 0.2845 & 0.5262 \end{pmatrix}. \tag{5.31}$$

Nous remarquons à partir de la Figure 5.1 que les deux sources synthétiques sont beaucoup mieux estimées par notre méthode que par la méthode NMF-Grd-LQ .

[1] Le calcul a été réalisé avec Matlab, sur un ordinateur avec un un processeur intel Core i7, une fréquence de 2,7 GHz et 16 Go de RAM.

Tableau 5.1: *Résultats des simulations en utilisant (B-SO-BSS) avec un et quatre décalages, et en utilisant NMF-Grd-LQ dans le cas de deux signaux synthétiques.*

Méthodes \ Nombre d'échantillons		**100**	**1000**	**10000**
B-So-BSS $\tau_i = 1$	RSI_s (dB)	9.06	20.18	30.15
	$EQMN_s$	0.33	0.022	0.0022
	RSI_a (dB)	8.71	11.25	19.38
	$EQMN_a$	0.36	0.33	0.08
	$Temps\ CPU\ (sec)$	$6.4\ 10^{-4}$	$8.8\ 10^{-4}$	$2.1\ 10^{-3}$
B-So-BSS $\tau_i = 1, 2, 3, 4$	RSI_s (dB)	8.05	19.54	29.80
	$EQMN_s$	0.36	0.028	0.0024
	RSI_a (dB)	8.73	11.15	18.89
	$EQMN_a$	0.39	0.32	0.08
	$Temps\ CPU\ (sec)$	$2.4\ 10^{-3}$	$2.8\ 10^{-3}$	$8.7\ 10^{-3}$
NMF-Grd-LQ	RSI_s (dB)	7.65	12.27	14.87
	$EQMN_s$	0.37	0.069	0.06
	RSI_a (dB)	7.21	7.45	11.09
	$EQMN_a$	0.64	0.70	0.18
	$Temps\ CPU\ (sec)$	4.94	19.87	127.77

5.6.2 Expérience 2

Dans cette expérience, nous avons généré des mélanges artificiels de deux sources réelles. Ces sources, présentées en Figure 5.2, correspondent aux activités des ions Na^+ et K^+ (qui peuvent être considérées comme des concentrations efficaces des ions) (Duarte et al, 2014). Ces deux signaux sources contiennent chacun 41 échantillons, et ils ne prennent que des valeurs non-négatives.

Comme dans l'Expérience 1, les coefficients de mélange ont été générés par des valeurs aléatoires uniformément réparties entre 0 et 1. Par la suite, nous avons généré trois signaux observés en utilisant le modèle bilinéaire (5.1). Notons que le modèle de mélange des concentrations des espèces chimiques correspond au modèle Nicolsky-Eisenman (Duarte et al, 2014) et il est généralement approché par un modèle linéaire-quadratique. Cependant, dans cette expérience, nous mélangeons les sources chimiques selon le modèle bilinéaire car elles vérifient les hypothèses de notre méthode.

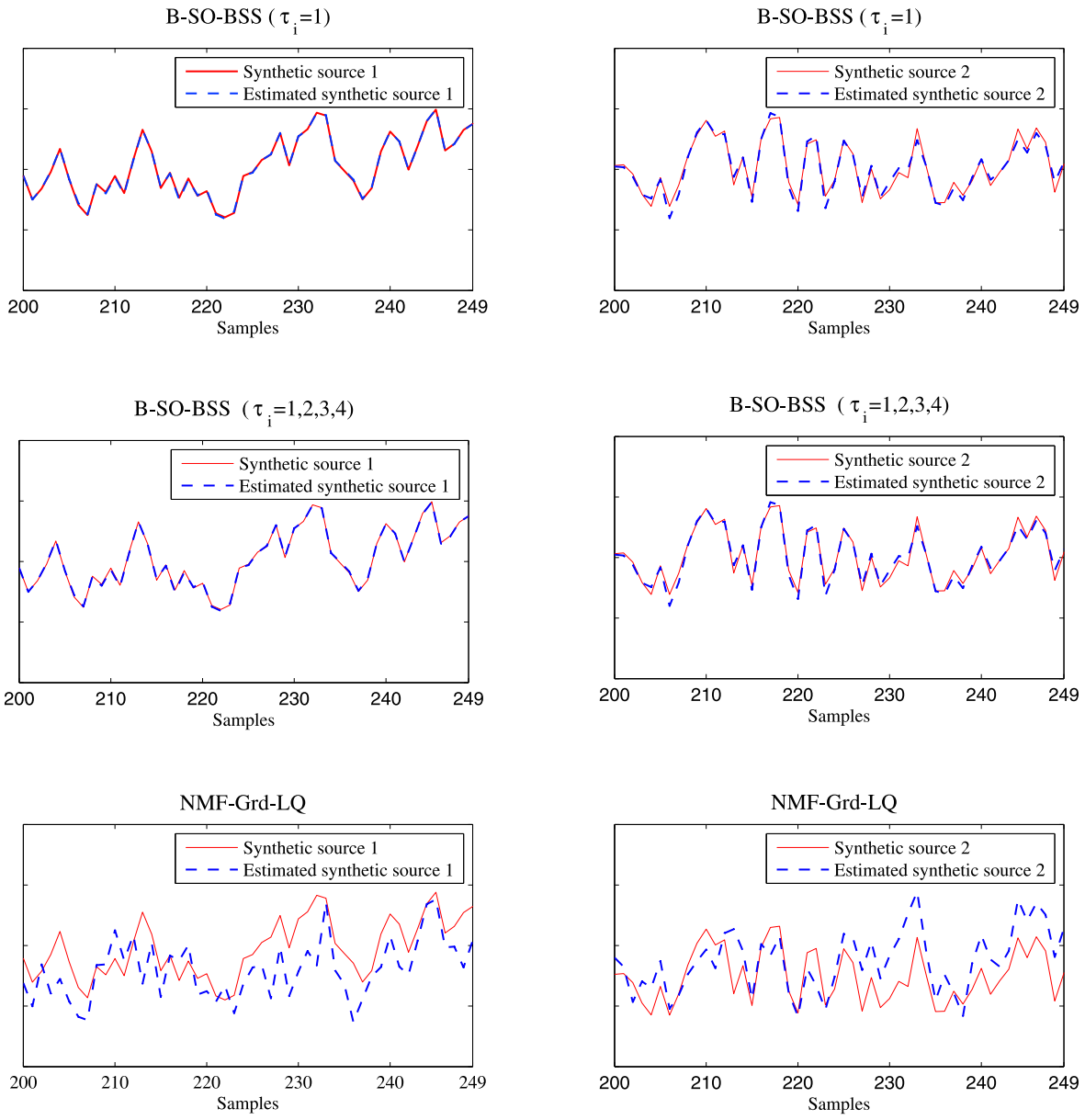


Figure 5.1: *Comparaison entre les vraies sources synthétiques et leurs estimées, dans le cas d'un mélange artificiel non-bruité, en utilisant 3 méthodes différentes : B-SO-BSS ($\tau_i = 1$) (haut), B-SO-BSS ($\tau_i = 1, 2, 3, 4$) (milieu) et NMF-Grd-LQ (en bas). Seule une partie des signaux correspondant à $n \in [200, 249]$ est présentée ici.*

Nous avons effectué 100 simulations Monte Carlo pour notre méthode et la méthode NMF-Grd-LQ. À chaque simulation, nous avons modifié seulement les coefficients de mélange. Les valeurs moyennes des RSI et des EQMN des sources et de tous les coefficients de mélange, ainsi que le temps CPU, calculés à partir des 100 simulations, sont présentés dans le Tableau 5.2.

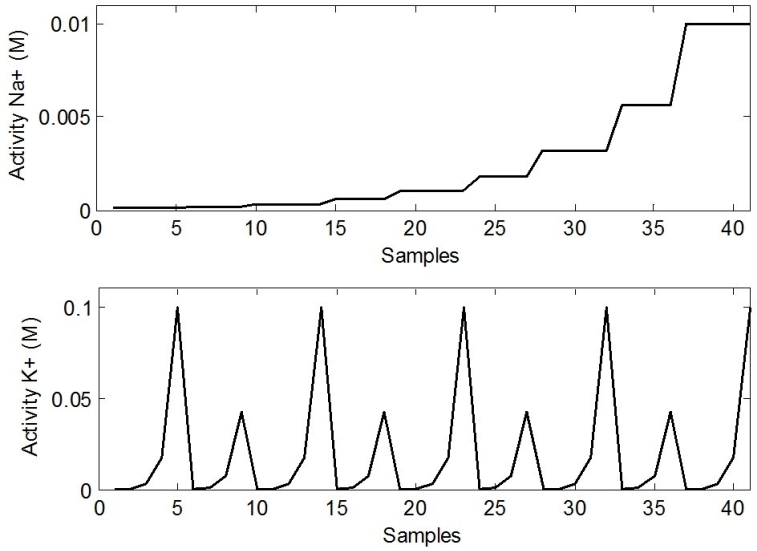


Figure 5.2: *Activités des ions Na+ et K+.*

Tableau 5.2: *Résultats des simulations en utilisant (B-SO-BSS) avec un et quatre décalages, et en utilisant NMF-Grd-LQ dans le cas de deux signaux chimiques.*

Méthodes / **Performances**	**B-So-BSS** $\tau_i = 1$	**B-So-BSS** $\tau_i = 1, 2, 3, 4$	**NMF-Grd-LQ**
RSI_s (dB)	13.68	15.99	7.48
$EQMN_s$	0.072	0.068	0.51
RSI_a (dB)	17.58	18.52	6.80
$EQMN_a$	0.31	0.31	0.41
$Temps\ CPU\ (sec)$	3.9 10^{-4}	8.5 10^{-4}	3.11

Nous pouvons remarquer que notre méthode permet d'avoir des meilleurs résultats par rapport à la méthode NMF-Grd-LQ. En terme de temps d'exécution, la méthode proposée est beaucoup plus rapide que NMF-Grd-LQ.

Nous montrons en Figure 5.3 un exemple de l'Expérience 2 où la matrice de mélange aléatoire $\mathbf{A_{Exp2}}$ est comme suit

$$\mathbf{A_{Exp2}} = \begin{pmatrix} 0.1549 & 0.1405 & 0.3916 \\ 0.5258 & 0.2041 & 0.9370 \\ 0.2047 & 0.5108 & 0.4310 \end{pmatrix}. \tag{5.32}$$

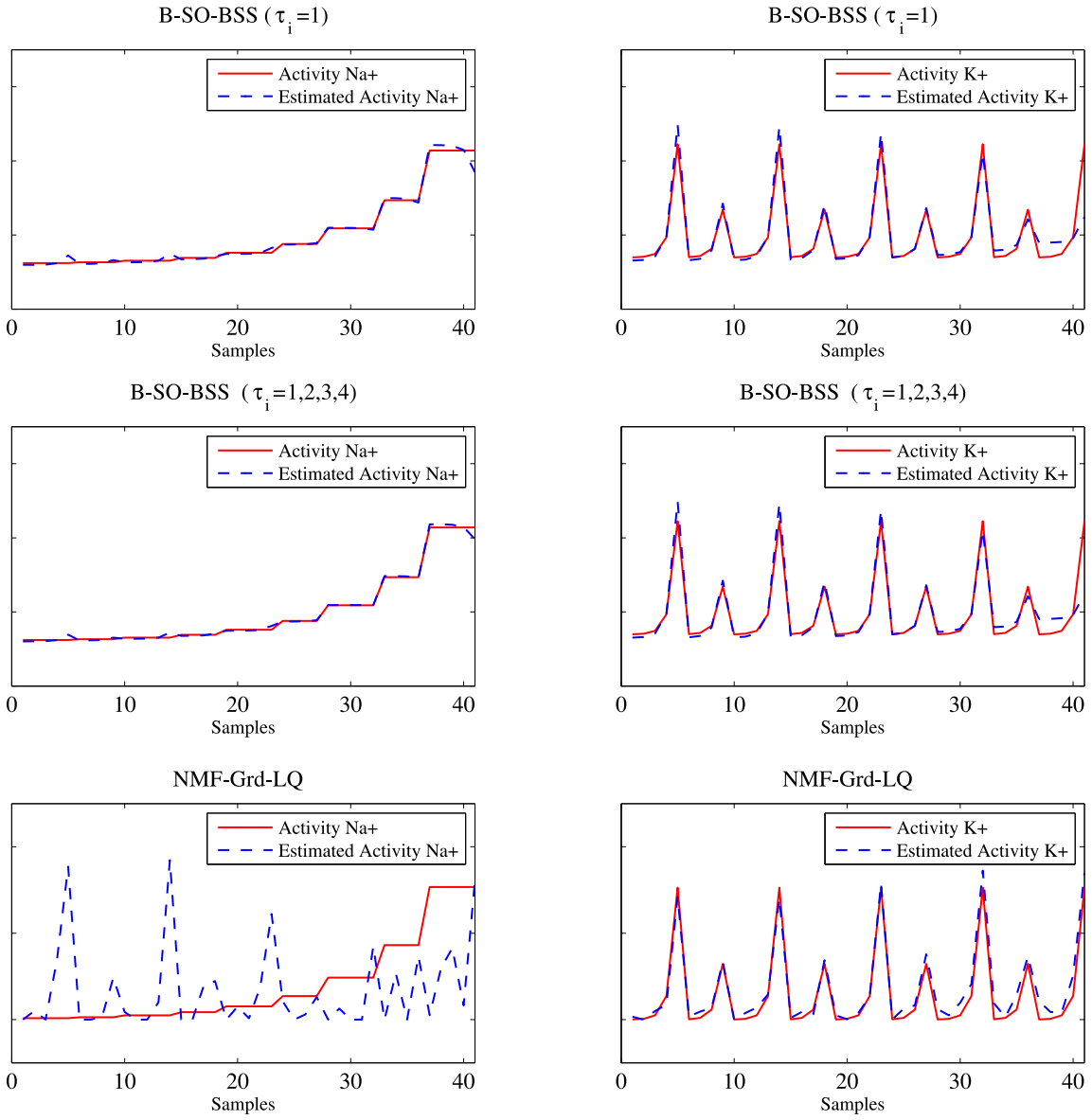


Figure 5.3: *Comparaison entre les sources chimiques et leurs estimées, dans le cas d'un mélange artificiel non-bruité, en utilisant 3 méthodes différentes : B-SO-BSS ($\tau_i = 1$) (haut), B-SO-BSS ($\tau_i = 1, 2, 3, 4$) (milieu) et NMF-Grd-LQ (en bas).*

Nous pouvons remarquer que les deux sources chimiques sont correctement estimées par notre méthode, tandis qu'une seule source est bien estimée par la méthode NMF-Grd-LQ.

5.6.3 Expérience 3

Le but de la troisième expérience est d'évaluer la robustesse de notre méthode face au bruit. Les mélanges ont été, d'abord, générés de la façon décrite dans l'Expérience 1 en considérant $N = 10000$. Par la suite, un bruit i.i.d, gaussien et de moyenne nulle a été ajouté aux signaux observés afin d'obtenir des mélanges bruités. Les valeurs du RSB (*Rapport Signal sur Bruit*) ont été variées de 60 dB à 30 dB.
Pour chaque niveau de bruit, nous avons effectué 100 simulations Monte Carlo et, à chaque simulation, nous avons modifié les signaux source et les paramètres du mélange. Les moyennes des RSI et des EQMN des sources et de tous les paramètres de mélange obtenus sont présentés dans le Tableau 5.3. Les variations du RSI_s en fonction du RSB sont présentées en Figure 5.4.

Tableau 5.3: *Résultats des simulations obtenus pour l'Expérience 3 en utilisant B-SO-BSS avec un et quatre décalages lorsque les mélanges sont bruités.*

Méthodes \ RSB (dB)		**60**	**50**	**40**	**30**
B-SO-BSS $\tau_i = 1$	RSI_s (dB)	25.46	20.49	13.60	6.13
	$EQMN_s$	0.03	0.08	0.21	0.43
	RSI_a (dB)	17.80	14.35	8.99	7.15
	$EQMN_a$	0.11	0.21	0.51	0.62
B-SO-BSS $\tau_i = 1, 2, 3, 4$	RSI_s (dB)	25.22	20.48	13.66	6.47
	$EQMN_s$	0.03	0.08	0.21	0.42
	RSI_a (dB)	18.17	14.52	9.21	7.29
	$EQMN_a$	0.10	0.21	0.51	0.61
NMF-Grd-LQ	RSI_s (dB)	14.82	14.61	13.93	12.77
	$EQMN_s$	0.06	0.06	0.06	0.07
	RSI_a (dB)	11.00	10.98	10.62	10.60
	$EQMN_a$	0.18	0.19	0.19	0.20

D'après la Figure 5.4, nous pouvons noter que les RSI_s obtenus par la méthode B-SO-BSS sont acceptables pour un RSB supérieur ou égal à 40 dB. En comparant ces résultats avec ceux obtenus par la méthode NMF-Grd-LQ, nous remarquons que pour des valeurs élevées de RSB, la méthode B-SO-BSS donne les meilleurs résultats, cependant, la méthode NMF-Grd-LQ semble plus efficace pour des faibles valeurs de RSB. En effet, nous pouvons remarquer que les RSI_s obtenus par NMF-Grd-LQ restent acceptables jusqu'à RSB = 30 dB.

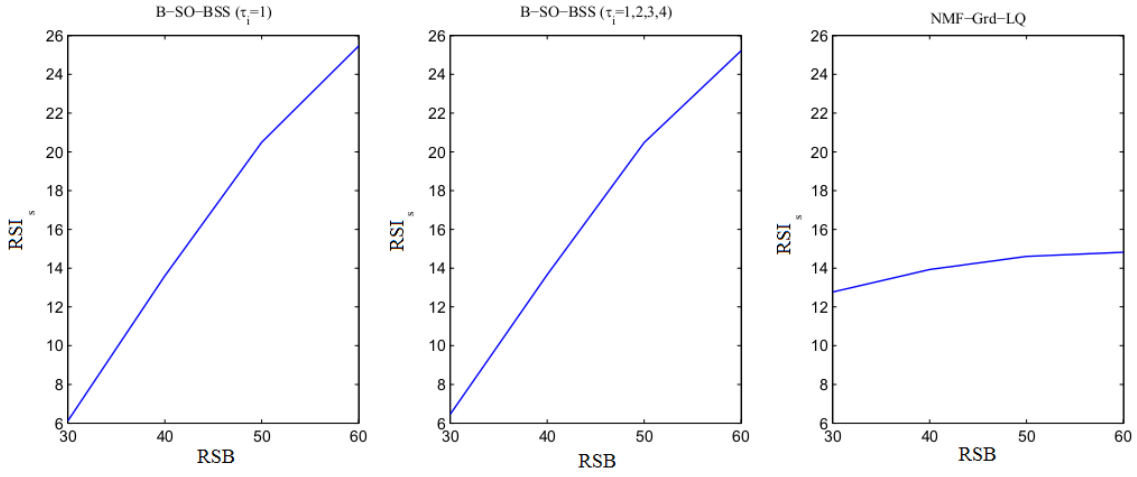


Figure 5.4: *RSI_s en fonction du RSB dans le cas de l'Expérience 3.*

5.6.4 Expérience 4

Nous avons ajouté un bruit i.i.d, gaussien et de moyenne nulle aux signaux observés générés dans l'Expérience 2. Comme dans l'Expérience 3, les valeurs du RSB sont variées de 60 dB à 30 dB.

Pour chaque niveau de bruit, nous avons effectué 100 simulations Monte Carlo, et à chaque simulation, nous avons modifié les coefficients du mélange.

Les moyennes des RSI et des EQMN des sources et de tous les paramètres du mélange obtenus sont présentées dans le Tableau 5.4. La représentation des RSI_s en fonction des valeurs du RSB est fournie en Figure 5.5.

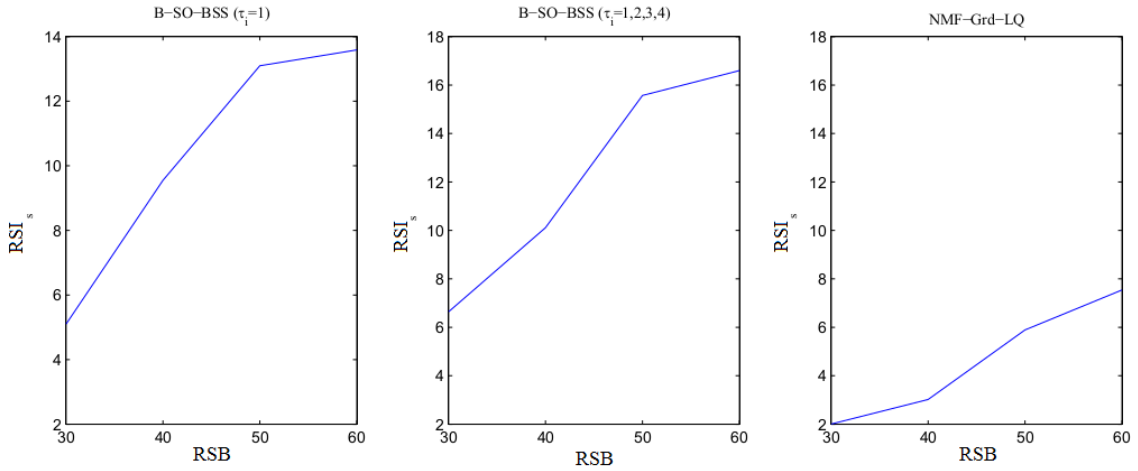


Figure 5.5: *RSI_s en fonction du RSB dans le cas de l'Expérience 4.*

Tableau 5.4: *Résultats des simulations obtenus pour l'Expérience 4 en utilisant B-SO-BSS avec un et quatre décalages lorsque les mélanges sont bruités.*

Méthodes \ RSB (dB)		**60**	**50**	**40**	**30**
B-SO-BSS $\tau_i = 1$	RSI_s (dB)	13.59	13.10	9.55	5.10
	$EQMN_s$	0.2881	0.33	0.45	0.61
	RSI_a (dB)	20.69	19.83	17.93	14.87
	$EQMN_a$	0.38	0.51	0.70	0.80
B-SO-BSS $\tau_i = 1, 2, 3, 4$	RSI_s (dB)	16.60	15.57	10.11	6.64
	$EQMN_s$	0.24	0.36	0.66	0.74
	RSI_a (dB)	20.72	20.20	16.66	14.47
	$EQMN_a$	0.34	0.42	0.73	0.56
NMF-Grd-LQ	RSI_s (dB)	7.54	5.89	3.02	2.01
	$EQMN_s$	0.51	0.54	0.66	0.74
	RSI_a (dB)	6.43	6.43	6.37	6.36
	$EQMN_a$	0.41	0.41	0.41	0.41

Dans le cas de l'Expérience 4, les résultats obtenus par la méthode B-SO-BSS sont acceptables pour un RSB supérieur ou égal à 50 dB. Au contraire, pour toutes les valeurs de RSB considérées, la méthode NMF-Grd-LQ ne parvient pas à fournir des estimations suffisamment précises.

En conclusion, pour les deux Expériences 3 et 4, la présence du bruit dans les signaux observés diminue les performances de séparation de notre méthode. Néanmoins, les performance restent acceptables pour des valeurs de RSB relativement élevées.

5.7 Conclusions

Dans ce chapitre, nous avons proposé une nouvelle méthode de SAS. Celle-ci représente une extension des méthodes SSO classiques proposées pour séparer les mélanges linéaires, afin de séparer les sources mélangées selon le modèle bilinéaire. D'abord, nous avons étudié les propriétés statistiques des différentes *sources étendues* lorsque les *vraies sources* ont des moyennes nulles et lorsqu'elles n'en ont pas. Ensuite, nous avons reformulé notre problème pour aboutir à des mélanges bilinéaires modifiés exprimés en fonction des sources centrées qui permettent de vérifier les conditions d'identifiabilité nécessaires pour l'utilisation des méthodes SSO. Nous avons présenté également les différentes étapes effectuées afin d'estimer les *vraies sources centrées* et les vrais coefficients de mélange. Enfin, nous avons présenté les résultats expérimentaux obtenus par la méthode proposée. Dans un premier temps, nous avons évalué les performances de séparation de la méthode lorsqu'elle est appliquée à des mélanges artificiels non-bruités de sources synthétiques ou chimiques. Ainsi, nous avons clairement remarqué l'efficacité de notre méthode comparée à la méthode NMF-Grd-LQ. Dans un deuxième temps, nous avons évalué la robustesse de la méthode face au bruit et, comme prévu, nous avons remarqué que la présence de bruit dans les mélanges générés diminue l'efficacité de notre méthode. Cependant, les performances restent acceptables pour des valeurs de RSB relativement élevées.

Comme travaux postérieurs à court terme, nous visons à étendre notre méthode proposée pour traiter le modèle de mélange linéaire-quadratique qui permet de bien décrire le mélange des concentrations des espèces chimiques.

Conclusion Générale

Dans cette thèse, nous nous sommes intéressés au problème de la séparation aveugle des mélanges linéaires-quadratiques/bilinéaires des sources. Les travaux effectués durant la thèse ont permis le développement de trois nouvelles méthodes qui exploitent différentes propriétés des signaux sources étudiés : parcimonie, positivité, lois de probabilités a priori, indépendance. Les deux premières méthodes sont particulièrement adaptées au démélange hyperspectral des images issues de la télédetection spatiale, tandis que la troisième peut être utilisée pour traiter différents types de données issues d'un mélange de sources indépendantes.

La première partie de la thèse est répartie en deux chapitres qui donnent un aperçu sur la SAS et le démélange spectral. Dans le premier chapitre, nous avons présenté le problème général de SAS, le modèle linéaire qui est le modèle de mélange le plus étudié, ainsi que les méthodes de SAS les plus classiques. Dans ce même chapitre, nous avons présenté le modèle linéaire-quadratique et les méthodes existantes pour le traiter. Dans le second chapitre, nous avons introduit quelques concepts de base sur la télédétection spatiale puis, nous avons donné un panorama des méthodes les plus populaires pour le démélange hyperspectral des deux modèles linéaire et non-linéaire.

La deuxième partie de ce document présente les apports méthodologiques de cette thèse. Chacun des chapitres de cette partie a été consacré à la présentation théorique et expérimentale d'une nouvelle méthode de SAS. Dans le troisième chapitre, nous avons développé une méthode de séparation de sources adaptée au modèle de mélange linéaire-quadratique pour le démélange des images hyperspectrales urbaines. L'objectif de cette méthode est d'extraire, à partir du spectre observé dans chaque pixel de l'image, tous les spectres des matériaux purs, ainsi que leurs fractions d'abondance et les coefficients quadratiques associés aux produits des spectres de ces matériaux purs. Dans cette approche, nous avons exploité la parcimonie des sources. Pour

ce faire, l'approche proposée nécessite l'existence d'une zone mono-source ou d'au moins deux pixels purs par matériau pur, dans la scène observée. Les étapes de cette approche consistent à estimer, d'abord, le nombre de matériaux purs présents dans l'image, puis à extraire les spectres des matériaux purs en détectant soit les zones mono-sources soit les pixels purs présents dans l'image, et enfin, à estimer leurs contributions dans chaque pixel en utilisant un algorithme des moindres carrés sous la contrainte de non-négativité. Les résultats des simulations sur différents types de données ont montré l'efficacité de la méthode proposée. Cette première méthode est basée sur l'hypothèse d'existence des zones mono-sources ou des paires de pixels purs par matériau dans le mélange, ce qui n'est pas toujours vérifié surtout dans le cas d'images hyperspectrales à faible résolution spatiale. Ainsi, pour pallier ce problème, nous avons proposé une autre méthode de séparation de sources adaptée au modèle LQ. Cette méthode, présentée dans le quatrième chapitre, correspond à une méthode de NMF se basant sur l'estimateur MAP. Le point clé de cette méthode consiste à prendre en compte l'information a priori disponible sur les distributions des coefficients de mélange afin de mieux les estimer. Cela a été réalisé par l'attribution de distributions de probabilité réalistes aux paramètres du mélange. En se basant sur un critère régularisé déduit de l'estimateur MAP, nous avons proposé différents algorithmes pour déterminer les paramètres inconnus du problème (les fractions d'abondance, les coefficients quadratiques, les signaux sources et les paramètres des distributions). Les résultats des simulations sur différents types de données ont montré la pertinence des algorithmes développés. Enfin, dans le dernier chapitre, nous avons proposé une nouvelle méthode de SAS basée sur l'ACI et les SSO pour séparer les mélanges bilinéaires. Cette méthode peut être appliquée à différents problèmes où les sources sont supposées à valeurs réelles, stochastiques, auto-corrélées, mutuellement indépendantes, et conjointement stationnaires. Nous avons étudié au début les propriétés statistiques des différentes sources étendues lorsque les vraies sources ont des moyennes nulles et lorsqu'elles n'en ont pas. À partir des résultats de cette étude, nous avons reformulé notre problème afin d'aboutir à des mélanges bilinéaires modifiés exprimés en fonction des sources centrées vérifiant les conditions nécessaires pour l'utilisation des méthodes SSO. Ensuite, nous avons présenté les différentes étapes effectuées pour séparer les vraies sources centrées et estimer les vrais paramètres de mélange. Les

résultats expérimentaux, obtenus avec des mélanges artificiels de sources synthétiques et réelles, ont confirmé l'efficacité de la méthode proposée.

Les travaux effectués dans cette thèse ouvrent la voie à des perspectives de recherches pour l'amélioration de chaque méthode proposée, telles que :

Dans le Chapitre 3 :

- Il serait intéressant d'améliorer l'étape d'estimation des spectres des matériaux purs en comparant un ensemble de spectres de pixels à la fois (trois ou quatre spectres) au lieu de deux spectres seulement afin d'éviter les fausses détections.
- Une deuxième amélioration peut être réalisée dans l'étape de détection des spectres des pixels purs en calculant le coefficient d'intercorrélation entre les spectres des pixels au lieu de calculer la distance euclidienne puisque la corrélation est moins sensible au bruit.

Dans le Chapitre 4 :

- Nous envisageons d'appliquer un algorithme du gradient à pas adaptatif afin de nous débarrasser du problème du choix du pas d'apprentissage de l'algorithme du gradient classique.
- Nous pensons aussi à étendre notre méthode par l'utilisation des méthodes MCMC pour l'estimation des inconnues du problème au lieu de l'algorithme du gradient.
- Nous comptons adapter la méthode proposée à d'autres applications faisant intervenir un modèle de mélange LQ, telles que le démélange des documents en recto-verso scannés où les paramètres des distributions a priori sont supposés approximativement connus.

Dans le Chapitre 5 :

- Nous visons à étendre notre méthode proposée pour traiter les mélanges linéaires-quadratiques qui permettent de bien décrire le mélange des concentrations des espèces chimiques.

Liste des publications

Publications dans des revues internationales :

1. L. JARBOUI, Y. DEVILLE, S. HOSSEINI, R. GUIDARA, A. BEN HAMIDA, and L. T. DUARTE, "A Second-Order Blind Source Separation Method for Bilinear Mixtures", Multidimensional Systems and Signal Processing Journal, pp. 1–20, DOI:10.1007/s11045-017-0493-9, Mai 2017.

Publications dans des conférences internationales :

1. L. JARBOUI, S. HOSSEINI, R. GUIDARA, Y. DEVILLE and A. BEN HAMIDA, "A MAP-based NMF approach to hyperspectral image unmixing using a linear-quadratic model" , Proceedings of ICASSP 2016, pp. 3356–3360, 20–25 March 2016, Shanghai, China.
2. L. JARBOUI, S. HOSSEINI, R. GUIDARA, Y. DEVILLE and A. BEN HAMIDA, "Séparation aveugle des mélanges linéaires-quadratiques pour les images hyperspectrales urbaines", Traitement et Analyse de l'Information Méthodes et Applications (TAIMA), 11-16 Mai 2015, Hammamet, Tunisia.
3. L. JARBOUI, S. HOSSEINI, Y. DEVILLE, R. GUIDARA, and A. BEN HAMIDA "A New Unsupervised Method for Hyperspectral Image Unmixing Using a Linear-quadratic Model", International Conference On Advanced Technologies For Signal & Image Processing (ATSIP) , 17-19 Mars 2014, Sousse, Tunisia.

Présentations dans des ateliers :

1. L. JARBOUI, S. HOSSEINI, Y. DEVILLE, R. GUIDARA, and A. BEN HAMIDA, "Méthodes avancées de séparation de sources : Application en imagerie spectrale ", Spring school on Advanced Technologies for signal & Image Processing (ATSIP), 21-24 March 2015, Monastir, Tunisia.

Bibliographie

Abdallah SA, Plumbley MD (2004) Polyphonic music transcription by non-negative sparse coding of power spectra. In: in Proc. Int. Conf. Music Information Retrieval (ISMIR), Barcelona, Spain, pp 318–328

Abed-Meraim K, Belouchrani A, Hua Y (1996) Blind identification of a linear-quadratic mixture of independent components based on joint diagonalization procedure. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96), Atlanta, GA, USA, pp 2718–2721

Abrard F, Deville Y (2005) A time-frequency blind signal separation method applicable to underdetermined mixtures of dependent sources. Signal Processing 85(7):1389–1403

Albouy B (2004) Méthodes de séparation aveugle de sources applicables à des signaux de parole. Thèse de doctorat, Université Paul Sabatier, Toulouse III

Almeida MSC, Almeida LB (2012) Nonlinear separation of show-through image mixtures using a physical model trained with ICA. Signal Processing 92(4):872–884

Altmann Y, Dobigeon N, Tourneret JY (2011a) Bilinear models for nonlinear unmixing of hyperspectral images. In: 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp 1–4

Altmann Y, Dobigeon N, Tourneret JY, McLaughlin S (2011b) Nonlinear unmixing of hyperspectral images using radial basis functions and orthogonal least squares. In: 2011 IEEE International Geoscience and Remote Sensing Symposium, pp 1151–1154

Altmann Y, Halimi A, Dobigeon N, Tourneret JY (2012) Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery. IEEE Transactions on Image Processing 21(6):3017–3025

Altmann Y, Dobigeon N, McLaughlin S, Tourneret JY (2013) Nonlinear spectral unmixing of hyperspectral images using gaussian processes. IEEE Transactions on Signal Processing 61(10):2442–2453

Altmann Y, Dobigeon N, Tourneret JY (2014) Unsupervised post-nonlinear unmixing of hyperspectral images using a hamiltonian monte carlo algorithm. IEEE Transactions on Image Processing 23(6):2663–2675

Amari S, Cichocki A, Yang HH (1996) A new learning algorithm for blind signal separation. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, MIT Press, pp 757–763

Ando R, Duarte L, Soriano D, Attux R, Suyama R, Deville Y, Jutten C (2012) Recurrent source separation structures as iterative methods for solving nonlinear equation systems. In: Proceedings of XXX Simposio Brasileiro de Telecomunicacoes, SBrT 2012, Brasilia, Brazil

Ando RA, Duarte LT, Jutten C, Attux R (2015) A blind source separation method for chemical sensor arrays based on a second order mixing model. In: 23rd European Signal Processing Conference (EUSIPCO-2015), Nice, France, pp 933–937

Arberet S, Gribonval R, Bimbot F (2006) A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture. In: Rosca J, Erdogmus D, Príncipe JC, Haykin S (eds) Independent Component Analysis and Blind Signal Separation: 6th International Conference, ICA 2006, Charleston, SC, USA, March 5-8, 2006. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 536–543

Asner GP, Lobell DB (2000) A biogeophysical approach for automated {SWIR} unmixing of soils and vegetation. Remote Sensing of Environment 74(1):99 – 112

Bachmann CM, Ainsworth TL, Fusina RA (2005) Exploiting manifold geometry in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 43(3):441–454

Bayliss J, Gualtieri JA, Cromp RF (1997) Analyzing hyperspectral data with independent component analysis. In: Proc. SPIE AIPR Workshop, SPIE, vol 9, pp 133–143

Bedoya G (2006) Non-linear blind signal separation for chemical solid-state sensor arrays. Ph.D. thesis, Universitat Politecnica de Cataluny

Bell AJ, Sejnowski TJ (1995) An information-maximization approach to blind separation and blind deconvolution. Neural Computation 7(6):1129–1159

Belouchrani A, Cichocki A (2000) Robust whitening procedure in blind source separation context. Electronics Letters 36(24):2050–2051

Belouchrani A, Abed-Meraim K, Cardoso JF, Moulines E (1997) A blind source separation technique using second-order statistics. IEEE Transactions on Signal Processing 45(2):434–444

Berné, O, Joblin, C, Deville, Y, Smith, J D, Rapacioli, M, Bernard, J P, Thomas, J, Reach, W, Abergel, A (2007) Analysis of the emission of very small dust particles from spitzer spectro-imagery data using blind signal separation methods. Astronomy and Astrophysics 469(2):575–586

Berry MW, Browne M, Langville AN, Pauca VP, Plemmons RJ (2007) Algorithms and applications for approximate nonnegative matrix factorization. Computational Statistics and Data Analysis 52(1):155 – 173

Bertin N, Badeau R, Vincent E (2009) Fast bayesian nmf algorithms enforcing harmonicity and temporal continuity in polyphonic music transcription. In: 2009 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Mohonk Mountain House New Paltz, NY, USA, pp 29–32

Bioucas-Dias JM (2009) A variable splitting augmented lagrangian approach to linear spectral unmixing. In: 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp 1–4

Bioucas-Dias JM, Figueiredo MAT (2010) Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In: 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp 1–4

Bioucas-Dias JM, Plaza A, Dobigeon N, Parente M, Du Q, Gader P, Chanussot J (2012) Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 5(2):354–379

Boardman JW (1994) Geometric mixture analysis of imaging spectrometry data. In: Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation., International, vol 4, pp 2369–2371 vol.4

Boardman JW (1995) Analysis, understanding, and visualization of hyperspectral data as convex sets in n space. vol 2480, pp 14–22

Bofill P, Zibulevsky M (2001) Underdetermined blind source separation using sparse representations. Signal Processing 81(11):2353–2362

Broadwater J, Banerjee A (2009) A comparison of kernel functions for intimate mixture models. In: 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp 1–4

Broadwater J, Chellappa R, Banerjee A, Burlina P (2007) Kernel fully constrained least squares abundance estimates. In: 2007 IEEE International Geoscience and Remote Sensing Symposium, pp 4041–4044

Cardoso JF (1997) Infomax and maximum likelihood for blind source separation. IEEE Signal Processing Letters 4(4):112–114

Cardoso JF, Souloumiac A (1993) Blind beamforming for non-gaussian signals. IEE Proceedings F - Radar and Signal Processing 140(6):362–370

Castella M (2008) Inversion of polynomial systems and separation of nonlinear mixtures of finite-alphabet sources. IEEE Transactions on Signal Processing 56(8):3905–3917

Chaouchi C, Deville Y, Hosseini S (2009) Nonlinear source separation: a quadratic recurrent inversion structure. In: Proceedings of ECMS 2009, Arrasate-Mondragon, Spain, p 91–98

Chaouchi C, Deville Y, Hosseini S (2010a) Nonlinear source separation: a maximum likelihood approach for quadratic mixtures. In: Proceedings of MaxEnt 2010, Chamonix, France

Chaouchi C, Hosseini S, Deville Y (2010b) Cumulant-based estimation of quadratic mixture parameters for blind source separation. In: Proceedings of EUSIPCO 2010, Aalborg, Denmark, pp 1826–1830

Chen J, Richard C, Honeine P (2013) Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model. IEEE Transactions on Signal Processing 61(2):480–492

Chen S, Donoho D, Saunders M (1998) Atomic decomposition by basis pursuit. SIAM Journal on Scientific Computing 20(1):33–61

Chen Z, Cichocki A, Rutkowski T (2006) Constrained non-negative matrix factorization method for eeg analysis in early detection of alzheimer's disease. In: Proceedings of 2006 IEEE Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 893–896

Choi S, Cichocki A, Belouchrani A (2002) Second order nonstationary source separation. Journal of VLSI signal processing systems for signal, image and video technology 32(1):93–104

Cichocki A, Amari SI (2003) Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. Wiley, New York

Cichocki A, Zdunek R (2007) Regularized alternating least squares algorithms for non-negative matrix/tensor factorization. In: Advances in Neural Networks – ISNN 2007: 4th International Symposium on Neural Networks, ISNN 2007, Nanjing, China, June 3-7, 2007, Proceedings, Part III, Springer Berlin Heidelberg, pp 793–802

Cichocki A, Amari S, Zdunek R, Kompass R, Hori G, He Z (2006) Extended smart algorithms for non-negative matrix factorization. In: Proceedings of the 8th International Conference on Artificial Intelligence and Soft Computing, Springer-Verlag, Berlin, Heidelberg, ICAISC'06, pp 548–562

Cichocki A, Zdunek R, Amari Si (2007) Hierarchical als algorithms for nonnegative matrix and 3d tensor factorization. In: Independent Component Analysis and Signal Separation: 7th International Conference, ICA 2007, London, UK, September 9-12, 2007. Proceedings, Springer Berlin Heidelberg, pp 169–176

Cichocki A, Zdunek R, Phan AH, Amari S (2009) Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation. John Wiley and Sons, Ltd

Close R, Gader P, Wilson J, Zare A (2012a) Using physics-based macroscopic and microscopic mixture models for hyperspectral pixel unmixing. vol 8390, pp 83,901L–83,901L–13

Close R, Gader P, Zare A, Wilson J, Dranishnikov D (2012b) Endmember extraction using the physics-based multi-mixture pixel model. In: Proc. SPIE 8515, Imaging Spectrometry XVII, vol 8515

Comon P (1994) Independent component analysis, a new concept? Signal Process 36(3):287–314

Comon P, Jutten C (2010) Handbook of Blind Source Separation. Independent Component Analysis and Applications. Academic Press, Oxford

Deville Y (2003) Panorama des applications biomédicales des méthodes de séparation aveugle de sources. In: GRETSI, Paris, France, pp 8–11

Deville Y (2016) Blind Source Separation and Blind Mixture Identification Methods. John Wiley

Deville Y, Duarte LT (2015) An overview of blind source separation methods for linear-quadratic and post-nonlinear mixtures. In: Proceedings of the 12th International Conference on Latent Variable Analysis and Signal Separation (LVA/ICA 2015), Springer International Publishing Switzerland, LNCS 9237, Liberec, Czech Republic, pp 155–167

Deville Y, Hosseini S (2007a) Blind identification and separation methods for linear-quadratic mixtures and/or linearly independent non-stationary signals. In: Proceedings of the International Symposium on Signal Processing and Its Applications (ISSPA-2007), Sharjah, United Arab Emirates, pp 1–4

Deville Y, Hosseini S (2007b) Stable higher-order recurrent neural network structures for nonlinear blind source separation. In: Independent Component Analysis and Signal Separation: 7th International Conference, ICA 2007, London, UK, September 9-12, 2007. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 161–168

Deville Y, Hosseini S (2009) Recurrent networks for separating extractable-target nonlinear mixtures. part i: Non-blind configurations. Signal Processing 89(4):378 – 393

Deville Y, Hosseini S (2010) Blind operation of a recurrent neural network for linear-quadratic source separation: Fixed points, stabilization and adaptation scheme. In: Latent Variable Analysis and Signal Separation: 9th International Conference, LVA/ICA 2010, St. Malo, France, September 27-30, 2010. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 237–244

Deville Y, Puigt M (2007) Temporal and time-frequency correlation-based blind source separation methods. part i: Determined and underdetermined linear instantaneous mixtures. Signal Processing 87(3):374 – 407

Deville Y, Hosseini S, Deville A (2011) Effect of indirect dependencies on maximum likelihood and information theoretic blind source separation for nonlinear mixtures. Signal Processing 91(4):793 – 800

Dobigeon N, Tourneret JY, Chang CI (2008) Semi-supervised linear spectral unmixing using a hierarchical bayesian model for hyperspectral imagery. IEEE Transactions on Signal Processing 56(7):2684–2695

Dobigeon N, Tourneret JY, Richard C, Bermudez JCM, McLaughlin S, Hero AO (2014) Nonlinear unmixing of hyperspectral images: Models and algorithms. IEEE Signal Processing Magazine 31(1):82–94

Dohan D, Karp S, Matejek B (2015) K-median algorithms: Theory in practice

Donoho D, Stodden V (2003) When does non-negative matrix factorization give correct decomposition into parts? In: Neural Information Processing Systems, MIT Press, p 2004

Donoho DL, Elad M (2002) Optimally sparse representation in general (non-orthogonal) dictionaries via l^1 minimization. In: PROC. NATL ACAD. SCI. USA 100 2197–202

Draine B (1988) The discrete-dipole approximation and its application to interstellar graphite grains. Astrophysics Journal 333:848–72

Duarte L, Ando RA, Attux R, Deville Y, Jutten C (2012) Separation of sparse signals in overdetermined linear-quadratic mixtures. In: Latent Variable Analysis and Signal Separation: 10th International Conference, LVA/ICA 2012, Tel Aviv, Israel, March 12-15, 2012. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 239–246

Duarte LT, Jutten C, Moussaoui S (2009) Bayesian source separation of linear-quadratic and linear mixtures through a mcmc method. In: 2009 IEEE International Workshop on Machine Learning for Signal Processing, pp 1–6

Duarte LT, Suyama R, Attux R, Deville Y, Romano JMT, Jutten C (2010) Blind source separation of overdetermined linear-quadratic mixtures. In: Latent Variable Analysis and Signal Separation: 9th International Conference, LVA/ICA 2010, St. Malo, France, September 27-30, 2010. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 263–270

Duarte LT, Tomazeli L, Jutten C, Moussaoui S (2011) Bayesian source separation of linear and linear-quadratic mixtures using truncated priors. Journal of Signal Processing Systems 65(3):311–323

Duarte LT, Romano JMT, Jutten C, Chumbimuni-Torres KY, Kubota LT (2014) Application of blind source separation methods to ion-selective electrode arrays in flow-injection analysis. IEEE Sensors Journal 14(7):2228–2229

Eches O, Guillaume M (2014) A bilinear-bilinear nonnegative matrix factorization method for hyperspectral unmixing. IEEE Geoscience and Remote Sensing Letters 11(4):778–782

Eches O, Dobigeon N, Mailhes C, Tourneret JY (2010a) Bayesian estimation of linear mixtures using the normal compositional model. application to hyperspectral imagery. IEEE Transactions on Image Processing 19(6):1403–1413

Eches O, Dobigeon N, Tourneret JY (2010b) Estimating the number of endmembers in hyperspectral images using the normal compositional model and a hierarchical bayesian algorithm. IEEE Journal of Selected Topics in Signal Processing 4(3):582–591

Eckstein J, Bertsekas DP (1992) On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators. Mathematical Programming 55(1):293–318

Eismann MT, Hardie RC (2004) Stochastic spectral unmixing with enhanced endmember class separation. Appl Opt 43(36):6596–6608

Erdogmus D, Hild KE, Rao YN, Principe JC (2004) Minimax mutual information approach for independent component analysis. Neural Computation 16(6):1235–1252

Fan W, Hu B, Miller J, Li M (2009) Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated forest hyperspectral data. International Journal of Remote Sensing 30(11):2951–2962

Ferreol A, Albera L, Chevalier P (2005) Fourth-order blind identification of underdetermined mixtures of sources (fobium). IEEE Transactions on Signal Processing 53(5):1640–1653

Févotte C, Idier J (2011) Algorithms for nonnegative matrix factorization with the beta-divergence. Neural Computation 23(9):2421–2456

Gader P, Dranishnikov D, Zare A, Chanussot J (2012) A sparsity promoting bilinear unmixing model. In: 2012 4th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp 1–4

Gaeta M, Lacoume JL (1990) Source separation without prior knowledge: The maximum likelihood solution. In: Proceedings of the European Signal Processing Conference, EUSIPCO, pp 621–624

Georgiev P (2001) Blind source separation of bilinearly mixed signals. In: Proceedings of ICA, San Diego, USA, p 328–331

Gilks W, Richardson S, Spiegehalter D (1999) Markov Chain Monte Carlo in Practice. London, U.K.: Chapman Hall

Gillis N, Glineur F (2012) A multilevel approach for nonnegative matrix factorization. Journal of Computational and Applied Mathematics 236(7):1708 – 1723

Green AA, Berman M, Switzer P, Craig MD (1988) A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing 26(1):65–74

Guidara R (2009) Méthodes markoviennes pour la separation aveugle de signaux et images. Thèse de doctorat, Université Paul Sabatier, Toulouse III

Guilfoyle KJ, Althouse ML, Chang CI (2001) A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks. IEEE Transactions on Geoscience and Remote Sensing 39(10):2314–2318

Halimi A, Altmann Y, Dobigeon N, Tourneret JY (2011a) Nonlinear unmixing of hyperspectral images using a generalized bilinear model. IEEE Transactions on Geoscience and Remote Sensing 49(11):4153–4162

Halimi A, Altmann Y, Dobigeon N, Tourneret JY (2011b) Unmixing hyperspectral images using the generalized bilinear model. In: 2011 IEEE International Geoscience and Remote Sensing Symposium, pp 1886–1889

Hapke B (1981) Bidirectional reflectance spectroscopy: 1. theory. Journal of Geophysical Research: Solid Earth 86(B4):3039–3054

Hapke B (2011) Theory of Reflectance and Emittance Spectroscopy:, 2nd edn. Cambridge, DOI 10.1017/CBO9781139025683

Heinz DC, Chein-I-Chang (2001) Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 39(3):529–545

Heylen R, Scheunders P (2012) Calculation of geodesic distances in nonlinear mixing models: Application to the generalized bilinear model. IEEE Geoscience and Remote Sensing Letters 9(4):644–648

Heylen R, Burazerovic D, Scheunders P (2011) Non-linear spectral unmixing by geodesic simplex volume maximization. IEEE Journal of Selected Topics in Signal Processing 5(3):534–542

Heylen R, Parente M, Gader P (2014) A review of nonlinear hyperspectral unmixing methods. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 7(6):1844–1868

Hild KE, Erdogmus D, Principe J (2001) Blind source separation using renyi's mutual information. IEEE Signal Processing Letters 8(6):174–176

Hosseini S, Deville Y (2003) Blind separation of linear-quadratic mixtures of real sources using a recurrent structure. In: Artificial Neural Nets Problem Solving Methods: 7th International Work-Conference on Artificial and Natural Neural Networks, IWANN2003 Maó, Menorca, Spain, June 3–6, 2003 Proceedings, Part II, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 241–248

Hosseini S, Deville Y (2004) Blind maximum likelihood separation of a linear-quadratic mixture. In: Independent Component Analysis and Blind Signal Separation: Fifth International Conference, ICA 2004, Granada, Spain, September 22-24, 2004. Proceedings, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 694–701

Hosseini S, Deville Y (2013) Recurrent networks for separating extractable-target nonlinear mixtures. part ii. blind configurations. Signal Processing 93(4):671 – 683

Hosseini S, Deville Y (2014) Blind separation of parametric nonlinear mixtures of possibly autocorrelated and non-stationary sources. IEEE Transactions on Signal Processing 62(24):6521–6533

Hoyer PO (2002) Non-negative sparse coding. In: Neural Networks For Signal Processing Xii Proc. IEEE Workshop On Neural Networks For Signal Processing, pp 557–565

Hérault J, Jutten C, Ans B (1985) Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In: GRETSI, Paris, France, pp 1014–1022

Hyvärinen A (1999) Fast and robust fixed-point algorithms for independent component analysis. IEEE Transactions on Neural Networks 10(3):626 – 634

Hyvärinen A, Karhunen J, Oja E (2001) Independent Component Analysis. Wiley, New York

Ichir M, Mohammad-Djafari A (2003) Wavelet domain blind image separation. vol 5207, pp 361–370

Iordache MD, Bioucas-Dias JM, Plaza A (2011) Total variation regulatization in sparse hyperspectral unmixing. In: 2011 3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), pp 1–4

Jia S, Qian Y (2009) Constrained nonnegative matrix factorization for hyperspectral unmixing. Geoscience Remote Sensing 47(1):161–173

Jourjine A, Rickard S, Yilmaz O (2000) Blind separation of disjoint orthogonal signals: demixing n sources from 2 mixtures. In: 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100), vol 5, pp 2985–2988

Jutten C, Comon P (2007) Séparation de sources, tome 2, Collection Traité IC2, Information-Commande-Communication. Hermès-Lavoisier, Paris, France

Jutten C, Herault J (1991) Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. Signal Processing 24(1):1–10

Kachenoura A, Albera L, Senhadji L (2007) Séparation aveugle de sources en ingénierie biomédicale. IRBM 28(1):20 – 34

Kardec Barros A, Mansour A, Ohnishi N (1998) Removing artifacts from ecg signals using independent components analysis. NeuroComputing 22:173–186

Karoui SM (2012) Méthodes de séparation aveugle de sources et application à la télédétection spatiale. Thèse de doctorat, Université Paul Sabatier, Toulouse III

Kaufman L, Rousseeuw P (1987) Clustering by Means of Medoids. Reports of the Faculty of Mathematics and Informatics, Faculty of Mathematics and Informatics

Kerekes JP, Schott JR (2006) Hyperspectral Imaging Systems, John Wiley and Sons, Inc., pp 17–45

Keshava N, Mustard JF (2002) Spectral unmixing. IEEE Signal Processing Magazine 19(1):44–57

Knuth K (1999) A bayesian approach for source separation. In: Proceedings of Int. Conf. Independent Component Analysis Blind Signal Separation (ICA), p 283–288

Kosaka N, Uto K, Kosugi Y (2005) Ica-aided mixed-pixel analysis of hyperspectral data in agricultural land. IEEE Geoscience and Remote Sensing Letters 2(2):220–224

Krob M, Benidir M (1993) Blind identification of a linear-quadratic model using higher-order statistics. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP-93), vol 4, pp 440–443

Kuruoglu EE (2010) Bayesian source separation for cosmology. IEEE Signal Processing Magazine 27(1):43–54

Lacoume JJ, Ruiz P (1988) Sources indentification: a solution based on the cumulants. In: Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling, pp 199–203

Lathauwer LD (September 1997) Signal processing based on multilinear algebra. Thèse de doctorat, Université catholique de Leuven, Belgique

Lawson C, Hanson R (1995) Solving Least Squares Problems. Society for Industrial and Applied Mathematics, DOI 10.1137/1.9781611971217

Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. Nature 401(6755):788–791

Lee DD, Seung HS (2001) Algorithms for non-negative matrix factorization. In: NIPS, MIT Press, pp 556–562

Li J, Bioucas-Dias JM (2008) Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In: IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium, vol 3, pp III – 250–III – 253

Li SZ, Hou XW, Zhang HJ, Cheng QS (2001) Learning spatially localized, parts-based representation. In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol 1, pp I–207–I–212 vol.1

Li Y, Amari S, Cichocki A, Ho DWC, Xie S (2006) Underdetermined blind source separation based on sparse representation. IEEE Transactions on Signal Processing 54(2):423–437

Lin CJ (2007) Projected gradient methods for nonnegative matrix factorization. Neural Comput 19(10):2756–2779

Liu Q, Wang W (2013) Show-through removal for scanned images using non-linear NMF with adaptive smoothing. In: Proceedings of ChinaSIP, Beijing, pp 650–654

Luo J, Zhang Z (2000) Using eigenvalue grads method to estimate the number of signal source. Beijing, vol 1, pp 223–225

Ma WK, Bioucas-Dias JM, Chan TH, Gillis N, Gader P, Plaza AJ, Ambikapathi A, Chi CY (2014) A signal processing perspective on hyperspectral unmixing: Insights from remote sensing. IEEE Signal Processing Magazine 31(1):67–81

Macqueen J (1967) Some methods for classification and analysis of multivariate observations. In: In 5-th Berkeley Symposium on Mathematical Statistics and Probability, pp 281–297

Mansour A, Jutten C (1995) Fourth-order criteria for blind sources separation. IEEE Transaction on Signal Processing 43(8):2022–2025

Mansour A, Jutten C (1999) What should we say about the kurtosis? IEEE Signal Processing Letters 6(12):321–322

Mansour A, Kawamoto M (2003) Ica papers classified according to their applications and performances. IEICE Trans on Fundamentals of Electronics, Communications and Computer Sciences E86-A(3):620–633

Mansour A, Barros AK, Ohnishi N (2000) Blind separation of sources: Methods, assumptions and applications. EICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences E83-A(8):1498–1512

Matsuoka K, Ohoya M, Kawamoto M (1995) A neural net for blind separation of nonstationary signals. Neural Networks 8(3):411 – 419

Meganem I, Deville Y, Puigt M (2010) Blind separation methods based on correlation for sparse possibly-correlated images. In: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, pp 1334–1337

Meganem I, Déliot P, Briottet X, Deville Y, Hosseini S (2014a) Linear-quadratic mixing model for reflectances in urban environments. IEEE Transactions on Geoscience and Remote Sensing 52(1):544–558

Meganem I, Deville Y, Hosseini S, Déliot P, Briottet X (2014b) Linear-quadratic blind source separation using NMF to unmix urban hyperspectral images. IEEE Transactions on Signal Processing 62(7):1822–1833

Melia T, Rickard S (2005) Extending the duet blind source separation technique. In: Actes Structure et Parcimonie pour la Représentation Adaptative de Signaux (SPARS'05))

Merrikh-Bayat F, Babaie-Zadeh M, Jutten C (2011) Linear-quadratic blind source separating structure for removing show-through in scanned documents. International Journal on Document Analysis and Recognition (IJDAR) 14(4):319–333

Miao L, Qi H (2007) Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing 45(3):765–777

Mohammad-Djafari A (2001) A bayesian approach to source separation. AIP Conference Proceedings 567(1):221–244

Mokhtari F, Babaie-Zadeh M, Jutten C (2009) Blind separation of bilinear mixtures using mutual information minimization. In: 2009 IEEE International Workshop on Machine Learning for Signal Processing, pp 1–6

Moreau E (2001) A generalization of joint-diagonalization criteria for source separation. IEEE Transactions on Signal Processing 49(3):530–541

Moussaoui S, Brie D, Caspary O, Mohammad-Djafari A (2004) A bayesian method for positive source separation. In: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol 5, pp V–485–8 vol.5

Moussaoui S, Brie D, Idier J (2005) Non-negative source separation: range of admissible solutions and conditions for the uniqueness of the solution. In: Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005., vol 5, pp 289–292

Nascimento JMP, Bioucas-Dias JM (2005a) Does independent component analysis play a role in unmixing hyperspectral data? IEEE Transactions on Geoscience and Remote Sensing 43(1):175–187

Nascimento JMP, Bioucas-Dias JM (2005b) Vertex component analysis: a fast algorithm to unmix hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing 43(4):898–910

Nascimento JMP, Bioucas-Dias JM (2007) Hyperspectral unmixing algorithm via dependent component analysis. In: 2007 IEEE International Geoscience and Remote Sensing Symposium, pp 4033–4036

Nascimento JMP, Bioucas-Dias JM (2009a) Learning dependent sources using mixtures of dirichlet: Applications on hyperspectral unmixing. In: 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp 1–5

Nascimento JMP, Bioucas-Dias JM (2009b) Nonlinear mixture model for hyperspectral unmixing. vol 7477, pp 74,770I–74,770I–8

Nascimento JMP, Bioucas-Dias JM (2012) Hyperspectral unmixing based on mixtures of dirichlet components. IEEE Transactions On Geoscience And Remote Sensing 50(3):863–878

Nash DB, Conel JE (1974) Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite. Journal of Geophysical Research 79(11):1615–1621

Nuzillard, D, Bijaoui, A (2000) Blind source separation and analysis of multispectral astronomical images. Astron Astrophys Suppl Ser 147(1):129–138

Oppelt N, Mauser W (2003) Hyperspectral remote sensing for the monitoring of plant parameters of maize (zea mays) 4879:13–24

Paatero P, Tapper U (1994) Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. Environmetrics 5(2):111–126

Parra L, Mueller KR, Spence C, Ziehe A, Sajda P (2007) Unmixing hyperspectral data. Adv Neural Inf Process Systems 12:942–948

Pati YC, Rezaiifar R, Krishnaprasad PS (1993) Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition. In: Proceedings of 27th Asilomar Conference on Signals, Systems and Computers, pp 40–44 vol.1

Pauca V, Piper J, Plemmons R (2006) Nonnegative matrix factorization for spectral data analysis. Linear Algebra and its Applications 416(1):29 – 47

Pham DT (1999) Mutual information approach to blind separation of stationary sources. IEEE Transactions on Information Theory 48:215–220

Pham DT, Cardoso JF (2001) Blind separation of instantaneous mixtures of nonstationary sources. IEEE Transactions on Signal Processing 49(9):1837–1847

Pham DT, Garat P, Jutten C (1992) Source separation without prior knowledge: The maximum likelihood solution. In: Signal Processing VI, Proceedings of the 6th European Signal Processing Conference, EUSIPCO, Brussels, Belgium, pp 771–774

Plaza J, Martinez P, Perez R, A P (2004) Nonlinear neural network mixture models for fractional abundance estimation in aviris hyperspectral images. In: Proc. XIII NASA/Jet propulsion laboratory airborne earth science workshop, Pasadena, CA, USA, p 1–12

Plaza J, Plaza A, Perez R, Martinez P (2005) Automated generation of semi-labeled training samples for nonlinear neural network-based abundance estimation in hyperspectral data. In: Proceedings. 2005 IEEE International Geoscience and Remote Sensing Symposium, 2005. IGARSS '05., vol 2, pp 1261–1264

Qian Y, Jia S, Zhou J, Robles-Kelly A (2011) Hyperspectral unmixing via $l_{1/2}$ sparsity-constrained nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing 49(11):4282–4297

Rezaei M, Boostani R (2010) Using genetic algorithm to enhance nonnegative matrix factorization initialization. In: 2010 6th Iranian Conference on Machine Vision and Image Processing, pp 1–5

Rezaei M, Boostani R, Rezaei M (2011) An efficient initialization method for nonnegative matrix factorization. Journal of Applied Sciences 11:354 – 359

Rhabi ME, Gelle G, Fenniri H, Delaunay G (2004) A penalized mutual information criterion for blind separation of convolutive mixtures. Signal Processing 84(10):1979 – 1984

Robert C (2001) The Bayesian Choice. 2nd ed. New York: Springer-Verlag

Roberts S, Choudrey R (2003) Data decomposition using independent component analysis with prior constraints. Pattern Recognition 36(8):1813–1825

Roberts SJ (1998) Independent component analysis: source assessment and separation, a bayesian approach. IEE Proceedings - Vision, Image and Signal Processing 145(3):149–154

Rowe D (2002) Multivariate Bayesian Statistics: Models for Source Separation and Signal Unmixing. Chapman and Hall/CRC

Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323–2326

Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena 60(1):259 – 268

S Moussaoui AMD D Brie, Carteret C (2006) Separation of non-negative mixture of non-negative sources using a bayesian approach and mcmc sampling. IEEE Transactions on Signal Processing 54:4133–4145

Sajda P, Du S, Parra LC (2003) Recovery of constituent spectra using non-negative matrix factorization. vol 5207, pp 321–331

Senecal S, Amblard PO (2000) Bayesian separation of discrete sources via gibbs sampling. In: Proceedings of the Second International workshop on Independent Component Analysis and Signal Separation : ICA'2000, pp 561–566

Shahnaz F, Berry MW, Pauca VP, Plemmons RJ (2006) Document clustering using nonnegative matrix factorization. Information Processing and Management 42(2):373 – 386

Shaw G, Manolakis D (2002) Signal processing for hyperspectral image exploitation. IEEE Signal Processing Magazine 19(1):12–16

Shkuratov Y, Starukhina L, Hoffmann H, Arnold G (1999) A model of spectral albedo of particulate surfaces: Implications for optical properties of the moon. Icarus 137(2):235 – 246

Snoussi H, Patanchon G, Macias-Perez JF, Mohammad-Djafari A, Delabrouille J (2002) Bayesian blind component separation for cosmic microwave background observations. AIP Conference Proceedings 617(1):125–140

Souloumiac A (1995) Blind source detection and separation using second order nonstationarity. In: 1995 International Conference on Acoustics, Speech, and Signal Processing, vol 3, pp 1912–1915 vol.3

Stolyarov V, Hobson MP, Lasenby AN, Barreiro RB (2005) All-sky component separation in the presence of anisotropic noise and dust temperature variations. Monthly Notices of the Royal Astronomical Society 357:145–155

Tenenbaum JB, de Silva V, Langford JC (2000) A global geometric framework for nonlinear dimensionality reduction. Science 290:2319–2323

Tong L, Soon VC, Huang YF, Liu R (1990) Amuse: a new blind identification algorithm. In: Proceedings of the International Symposium on Circuits and Systems (ISCAS-1990), vol 3, pp 1784–1787

Tu TM (2000) Unsupervised signature extraction and separation in hyperspectral images: a noise-adjusted fast independent component analysis approach. Optical Engineering 39(4):897–906

Uhl F, Oppelt N, Bartsch I (2013) Spectral mixture of intertidal marine macroalgae around the island of helgoland (germany, north sea). Aquatic Botany 111:112 – 124

Villa A, Chanussot J, Benediktsson JA, Jutten C (2011) Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution. IEEE Journal of Selected Topics in Signal Processing 5(3):521–533

Virtanen T (2007) Monaural sound source separation by nonnegative matrix factorization with temporal continuity and sparseness criteria. Trans Audio, Speech and Lang Proc 15(3):1066–1074

Wang J, Chang CI (2006) Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing 44(9):2601–2616

Winter ME (1999) N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. vol 3753, pp 266–275

Xu W, Liu X, Gong Y (2003) Document clustering based on non-negative matrix factorization. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, ACM, New York, NY, USA, SIGIR '03, pp 267–273

Yang C, Everitt JH, Du Q (2010) Applying linear spectral unmixing to airborne hyperspectral imagery for mapping yield variability in grain sorghum and cotton fields 4(1):041,887

Yokoya N, Chanussot J, Iwasaki A (2014) Nonlinear unmixing of hyperspectral data using semi-nonnegative matrix factorization. IEEE Transactions on Geoscience and Remote Sensing 52(2):1430–1437

Zare A, Gader P (2007) Sparsity promoting iterated constrained endmember detection in hyperspectral imagery. IEEE Geoscience and Remote Sensing Letters 4(3):446–450

Zare A, Gader P, Casella G (2013) Sampling piecewise convex unmixing and endmember extraction. IEEE Transactions on Geoscience and Remote Sensing 51(3):1655–1665

Zdunek R, Cichocki A (2008) Fast nonnegative matrix factorization algorithms using projected gradient approaches for large-scale problems. Intell Neuroscience 2008:3:1–3:13

Ziehe A, Müller KR (1998) Tdsep - an efficient algorithm for blind separation using time structure. In: in Proc. International Conference on Artificial Neural Networks, pp 675–680

Zymnis A, Kim SJ, Skaf J, Parente M, Boyd S (2007) Hyperspectral image unmixing via alternating projected subgradients. In: 2007 Conference Record of the Forty-First Asilomar Conference on Signals, Systems and Computers, pp 1164–1168