



**HAL**  
open science

# Méthode géométrique de séparation de sources non-négatives : applications à l'imagerie dynamique TEP et à la spectrométrie de masse

Wendyam Ouedraogo

## ► To cite this version:

Wendyam Ouedraogo. Méthode géométrique de séparation de sources non-négatives : applications à l'imagerie dynamique TEP et à la spectrométrie de masse. Autre. Université de Grenoble; Université de Tunis El Manar, 2012. Français. NNT : 2012GRENT098 . tel-00859690

**HAL Id: tel-00859690**

**<https://theses.hal.science/tel-00859690>**

Submitted on 9 Sep 2013

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

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



## THÈSE

Pour obtenir le grade de

### DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Signal, Image, Parole, Télécoms**

Arrêté ministériel : 7 août 2006

Et de

### DOCTEUR DE L'UNIVERSITÉ DE TUNIS EL MANAR

Spécialité : **Télécommunications**

Présentée par

## Wendyam Serge Boris Ouedraogo

Thèse dirigée par **Christian JUTTEN** et **Meriem JAIDANE**

préparée au sein du **Laboratoire d'Outils pour l'Analyse de Données du CEA Saclay**

dans **l'École Doctorale Electronique, Electrotechnique, Automatique et Traitement du Signal**

et de **l'Unité de recherche Signaux et Systèmes de l'Ecole Nationale d'Ingénieurs de Tunis**

dans **l'École Doctorale Sciences et Techniques de l'Ingénieurs**

## **Méthode géométrique de séparation de sources non-négatives : Applications à l'imagerie dynamique TEP et à la spectrométrie de masse**

Thèse soutenue publiquement le **28 novembre 2012**, devant le jury composé de :

### **Monsieur, Pierre, COMON**

Directeur de Recherche CNRS, Président

### **Monsieur, Yannick, DEVILLE**

Pr., Université Paul Sabatier Toulouse 3, Rapporteur

### **Monsieur, Hichem, SNOUSSI**

Pr., Université de Technologie de Troyes, Rapporteur

### **Monsieur, Wady, NAANAA**

Pr., Université de Monastir, Examineur

### **Madame, Meriem, JAIDANE**

Pr., Université de Tunis El-Manar, Co-directrice

### **Monsieur, Christian, JUTTEN**

Pr., Université de Grenoble, Co-directeur

### **Monsieur, Antoine, SOULOUMIAC**

Dr., Ingénieur-chercheur CEA Saclay, Encadrant





A mon épouse et à mes parents  
Vous qui m'avez toujours soutenu malgré la distance  
Je vous dédie ce travail



# Remerciements

Avant tout propos, permettez-moi de remercier tous ceux qui ont contribué à l'aboutissement de mes travaux de thèse.

Je tiens d'abord à témoigner toute ma gratitude au Dr. Antoine SOULOUMIAC, ingénieur-chercheur au CEA LIST, qui a accepté de me prendre en thèse. Je voudrais te remercier Antoine, pour ta disponibilité constante et pour ton sens du travail bien fait. J'ai eu le plaisir d'apprendre avec toi à visualiser et à sentir les concepts et les équations mathématiques. J'espère avoir été un bon élève, et je ferai en sorte de te faire honneur.

Je voudrais remercier particulièrement Prof. Meriem JAIDANE, enseignante-chercheur à l'École Nationale d'Ingénieurs de Tunis. Md JAIDANE, mon aventure dans la recherche, et particulièrement en séparation de sources, a commencé avec vous. Je voudrais ici vous témoigner toute ma reconnaissance pour avoir cru en moi, et pour avoir accepté de co-diriger mes travaux de thèse. Votre rigueur dans l'organisation du travail et dans la rédaction des documents scientifiques est une chose que j'ai particulièrement noté et apprécié.

Je tiens à remercier Prof. Christian JUTTEN, enseignant-chercheur à l'Université Joseph Fourier de Grenoble, qui a co-dirigé mes travaux. Christian, je n'ai certes pas souvent été à Grenoble, mais tu as su intervenir quand il faut pour corriger la direction à suivre pour mes travaux. Je te remercie pour tes conseils avisés et ton sens du détail.

Permettez-moi de remercier aussi les membres du Département Capteurs Signal et Information (DCSI) du CEA LIST, département au sein duquel j'ai eu le plaisir d'effectuer mes travaux. Je remercie particulièrement Anthony LARUE, mon chef au Laboratoire d'Outils pour l'Analyse de Données (LOAD), qui n'a ménagé aucun effort pour me mettre dans les conditions optimales de travail.

Mes remerciements vont également à l'endroit des membres de l'Unité de recherche Signaux et Systèmes (U2S) de l'École Nationale d'Ingénieurs de Tunis, qui m'ont accueilli avec joie et m'ont mis dans les meilleures conditions de travail, lors de mes séjours de recherche à Tunis dans le cadre de ma co-tutelle de thèse.

Pour finir, je tiens à remercier mes rapporteurs Yannick DEVILLE et Hichem SNOUSSI, ainsi que les autres membres de mon jury de thèse, pour m'avoir fait l'honneur de participer à mon jury de thèse, et pour avoir contribué, à travers leurs commentaires pertinents, à l'amélioration de mon manuscrit.



# Résumé

Cette thèse traite du problème de séparation aveugle de sources non-négatives (c'est à dire des grandeurs positives ou nulles). La situation de séparation de mélanges linéaires instantanés de sources non-négatives se rencontre dans de nombreux problèmes de traitement de signal et d'images, comme la décomposition de signaux mesurés par un spectromètre (spectres de masse, spectres Raman, spectres infrarouges), la décomposition d'images (médicales, multi-spectrales ou hyperspectrales) ou encore l'estimation de l'activité d'un radionucléide. Dans ces problèmes, les grandeurs sont intrinsèquement non-négatives et cette propriété doit être préservée lors de leur estimation. En effet c'est cette propriété de non-négativité qui donne un sens physique aux composantes estimées.

La plupart des méthodes existantes de séparation de sources non-négatives requièrent de "fortes" hypothèses sur les sources (comme l'indépendance mutuelle, la dominance locale ou encore l'additivité totale des sources), qui ne sont pas toujours vérifiées en pratique.

Dans ce travail, nous proposons une nouvelle méthode de séparation de sources non-négatives fondée sur la répartition géométrique du nuage de points des signaux observés. Les coefficients de mélange et les sources sont estimées en cherchant le cône simplicial d'ouverture minimale contenant ce nuage de points. Cette méthode ne nécessite pas l'indépendance mutuelle des sources, ni même leur décorrélation ; elle ne requiert pas non plus la dominance locale des sources, ni leur additivité totale. Une seule condition est nécessaire et suffisante : l'orthant positif doit être l'unique cône simplicial d'ouverture minimale contenant le nuage de points des signaux sources. L'algorithme proposé est évalué avec succès dans deux situations de séparation de sources non-négatives de natures très différentes. Dans la première situation, nous effectuons la séparation de spectres de masse mesurés à la sortie d'un chromatographe en phase liquide haute précision, afin d'identifier et quantifier les différents métabolites (petites molécules) présents dans l'urine d'un rat traité au phénobarbital. Dans la deuxième situation, nous estimons les différents compartiments pharmacocinétiques du radio-traceur FluoroDeoxyGlucose marqué au fluor 18 ([ $^{18}\text{F}$ ]-FDG) dans le cerveau d'un patient humain, à partir d'une série d'images 3D TEP de cet organe. Parmi ces pharmacocinétiques, la fonction d'entrée artérielle présente un grand intérêt pour l'évaluation de l'efficacité d'un traitement anti-cancéreux en oncologie.





# Abstract

This thesis addresses the problem of non-negative blind source separation (i.e. positive or zero quantities). The situation of linear instantaneous mixtures of non-negative sources occurs in many problems of signal and image processing, such as decomposition of signals measured by a spectrometer (mass spectra, Raman spectra, infrared spectra), decomposition of images (medical, multi-spectral and hyperspectral) or estimation of the activity of a radionuclide. In these problems, the sources are inherently non-negative and this property should be preserved during their estimation, in order to get physical meaning components.

Most of existing non-negative blind source separation methods require “strong” assumptions on sources (such as mutual independence, local dominance or total additivity), which are not always satisfied in practice.

In this work, we propose a new geometrical method for separating non-negative sources. The mixing matrix and the sources are estimated by finding the minimum aperture simplicial cone containing the scatter plot of mixed data. The proposed method does not require the mutual independence of the sources, neither their decorrelation, nor their local dominance, or their total additivity. One condition is necessary and sufficient : the positive orthant must be the unique minimum aperture simplicial cone containing the scatter plot of the sources. The proposed algorithm is successfully evaluated in two different problems of non-negative sources separation. In the first situation, we perform the separation of mass spectra measured at the output of a liquid chromatograph to identify and quantify the different metabolites (small molecules) present in the urine of rats treated with phenobarbital . In the second situation, we estimate the different pharmacokinetics compartments of the radiotracer [18F]-FDG in human brain, from a set of 3D PET images of this organ, without blood sampling. Among these pharmacokinetics, the arterial input function is of great interest to evaluate the effectiveness of anti-cancer treatment in oncology.



# Table des matières

|  |              |
|--|--------------|
| <b>Remerciements</b>   | <b>v</b>     |
| <b>Résumé</b>  | <b>vii</b>   |
| <b>Abstract</b>  | <b>ix</b>    |
| <b>Abréviations</b>  | <b>xv</b>    |
| <b>Liste des figures</b>   | <b>xviii</b> |
| <b>Liste des tableaux</b>  | <b>xx</b>    |
| <b>Introduction générale</b>   | <b>2</b>     |
| <b>I État de l’art</b>   | <b>6</b>     |
| <b>1 Problématique et Applications</b>                                       | <b>8</b>     |
| 1.1 Introduction . . . . .   | 9            |
| 1.2 Séparation Aveugle de Sources (SAS) . . . . .                            | 9            |
| 1.2.1 Un seul problème théorique, plusieurs applications pratiques . . . . . | 9            |
| 1.2.2 Mise en équation du problème de SAS . . . . .                          | 10           |
| 1.2.3 Trois décennies de séparation de sources . . . . .                     | 13           |
| 1.3 Mélanges linéaires instantanés de sources non-négatives . . . . .        | 13           |
| 1.3.1 Équation de mélange . . . . .  | 13           |
| 1.3.2 Quelques situations de mélanges linéaires de sources non-négatives . . | 14           |
| 1.3.3 Séparation de spectres de masse en métabolomique . . . . .             | 14           |
| 1.3.4 Séparation d’images dynamiques TEP pour l’oncologie . . . . .          | 17           |
| 1.4 Hypothèses de travail . . . . .  | 19           |
| 1.5 Bruit en situation de sources non-négatives . . . . .                    | 20           |
| 1.6 Conclusion . . . . .   | 20           |

|           |  |           |
|-----------|--|-----------|
| <b>2</b>  | <b>Analyse critique des méthodes de SAS non-négatives</b>                                      | <b>22</b> |
| 2.1       | Introduction . . . . .   | 23        |
| 2.2       | Méthodes statistiques . . . . .  | 23        |
| 2.2.1     | Analyse en composantes indépendantes non-négatives . . . . .                                   | 23        |
| 2.2.2     | Méthode de maximum de vraisemblance . . . . .  | 25        |
| 2.2.3     | Inférence bayésienne . . . . .   | 27        |
| 2.3       | Méthodes déterministes . . . . .   | 27        |
| 2.3.1     | Factorisation en Matrices Non-négatives . . . . .  | 27        |
| 2.3.2     | Approches géométriques . . . . .   | 29        |
| 2.4       | Synthèse des méthodes de SAS non-négatives . . . . .   | 34        |
| 2.5       | Premiers résultats de séparation et premières conclusions . . . . .                            | 36        |
| 2.5.1     | Critères de performances . . . . .   | 37        |
| 2.5.2     | Génération de la matrice de mélange et du bruit . . . . .                                      | 37        |
| 2.5.3     | Résultats de séparation de sources indépendantes . . . . .                                     | 38        |
| 2.5.4     | Résultats de séparation de sources corrélées . . . . .   | 40        |
| 2.5.5     | Synthèse des résultats de séparation . . . . .   | 42        |
| 2.6       | Conclusion . . . . .   | 43        |
|           |  |           |
| <b>II</b> | <b>Développement d'une nouvelle méthode géométrique de séparation de sources non-négatives</b> | <b>45</b> |
|           |  |           |
| <b>3</b>  | <b>SCSA-UNS</b>  | <b>47</b> |
| 3.1       | Introduction . . . . .   | 48        |
| 3.2       | Synthèse de la méthode SCSA-UNS . . . . .  | 48        |
| 3.3       | Commentaires des résultats de séparation . . . . .   | 50        |
| 3.4       | Conclusions . . . . .  | 50        |
|           |  |           |
| <b>4</b>  | <b>Étude expérimentale de performance et de robustesse</b>                                     | <b>66</b> |
| 4.1       | Introduction . . . . .   | 67        |
| 4.2       | Initialisation de l'algorithme SCSA-UNS . . . . .  | 67        |
| 4.2.1     | Initialisation par l'orthant positif . . . . .   | 67        |
| 4.2.2     | Initialisation par VCA élargi . . . . .  | 70        |
| 4.3       | Étude expérimentale des performances de SCSA-UNS . . . . .                                     | 72        |
| 4.3.1     | Conditions de simulation . . . . .   | 72        |
| 4.3.2     | Performances en fonction du nombre d'échantillons . . . . .                                    | 72        |
| 4.3.3     | Performances en fonction de $n$ et $\tau$ . . . . .  | 74        |
| 4.4       | Robustesse de la méthode SCSA-UNS . . . . .  | 75        |
| 4.4.1     | Robustesse à la position du sommet du cône . . . . .   | 75        |

|  |   |            |
|--|---|------------|
| 4.4.2  | Robustesse à un bruit additif . . . . .   | 77         |
| 4.5  | Conclusion . . . . .  | 83         |
| <b>III Résultats de séparation sur données réelles LC-MS et TEP dynamiques</b> |   | <b>85</b>  |
| <b>5</b>   | <b>Résultats de séparation sur données LC-MS et TEP dynamiques</b>  | <b>87</b>  |
| 5.1  | Introduction . . . . .  | 88         |
| 5.2  | Métabolomique d'urine de rat traité au phénobarbital . . . . .  | 88         |
| 5.2.1  | Modèle de mélange des spectres de masse . . . . .   | 88         |
| 5.2.2  | Description des données LC-MS . . . . .   | 88         |
| 5.2.3  | Résultats de séparation des spectres de masse . . . . .   | 90         |
| 5.3  | Estimation des pharmacocinétiques du [18F]-FDG dans le cerveau . . . . .  | 97         |
| 5.3.1  | Modèle de mélange des images TEP . . . . .  | 97         |
| 5.3.2  | Résultats de séparation sur images TEP de fantômes . . . . .  | 97         |
| 5.3.3  | Résultats de séparation sur images TEP réelles de cerveau humain . . . . .  | 99         |
| 5.4  | Conclusion . . . . .  | 100        |
| <b>Conclusion générale et perspectives</b>                                     |   | <b>103</b> |
| <b>Annexes</b>   |   | <b>108</b> |
| <b>A</b>   | <b>Preuve de la non séparabilité par N-ICA de sources non-négatives, décor-<br/>réées et <i>well-grounded</i></b> | <b>108</b> |
| <b>B</b>   | <b>Quelques mesures de divergence utilisées en NMF</b>  | <b>112</b> |
| <b>C</b>   | <b>Liste des métabolites annotés dans l'urine de rat</b>  | <b>114</b> |
| <b>D</b>   | <b>Liste des publications</b>   | <b>117</b> |
| <b>Bibliographie</b>   |   | <b>129</b> |



# Abréviations

- CEA** : Commissariat à l'Énergie Atomique
- DRT** : Direction de la Recherche Technologique
- LIST** : Laboratoire d'Intégration des Systèmes et des Technologies
- DCSI** : Département Capteurs Signal et Information
- LOAD** : Laboratoire d'Outils pour l'Analyse de Données
- ENIT** : École Nationale d'Ingénieurs de Tunis
- U2S** : Unité de recherche Signaux et Systèmes
- UdG** : Université de Grenoble
- DSV** : Direction des Sciences du Vivant
- IBITEC-S** : Institut de Biologie et de Technologies de Saclay
- LEMM** : Laboratoire d'Étude du Métabolisme des Médicaments
- SPI** : Service de Pharmacologie et Immunoanalyse
- I<sup>2</sup>BM** : Institut d'Imagerie Biomédicale
- SHFJ** : Service Hospitalier Frédéric Joliot
- SAS** : Séparation Aveugle de Sources
- TEP** : Tomographie par Emissions de Positons
- LC-MS** : Liquid Chromatography - Mass Spectrometry
- FDG** : FluoroDéoxyGlucose
- AIF** : Arterial Input Function
- rAIF** : reference Arterial Input Function
- ACI** : Analyse en composantes indépendantes
- N-ICA** : Nonnegative Independent Component Analysis
- MV** : Maximum de Vraisemblance
- NMF** : Nonnegative Matrix Factorization
- BI-NSS** : Bayesian Inference - Non-negative Source Separation
- VCA** : Vertex Component Analysis
- SISAL** : Simplex Identification via Split Augmented Lagrangian
- MASC** : Minimum Aperture Simplicial Cone
- nLCA-IVM** : nonnegative Least-correlated Component Analysis based on Iterative Volume Maximization





# Table des figures

|     |  |    |
|-----|--|----|
| 1.1 | Procédure de mélange . . . . .   | 10 |
| 1.2 | Procédure de séparation . . . . .  | 12 |
| 1.3 | Spectre de masse de la nicotine, également appelée pyridine . . . . .  | 16 |
| 1.4 | Chaîne d'étude métabolomique au CEA-iBiTec-S/LEMM . . . . .  | 16 |
| 1.5 | Illustration du profil d'éluion et du spectre de masse de métabolite . . . . .   | 17 |
| 1.6 | Image TEP de cerveau humain . . . . .  | 18 |
| 1.7 | Modèle de mélange des images TEP . . . . .   | 19 |
| 2.1 | Diagrammes de dispersion de mélanges de deux sources uniformes et de mélanges de deux sources parcimonieuses . . . . .   | 30 |
| 2.2 | Diagrammes de dispersion de mélanges de sources à dominance locale . . . . .   | 31 |
| 2.3 | Nuage de points des signaux sources et observations, dans le cas de sources incluses dans un simplexe . . . . .  | 32 |
| 2.4 | Nuage de points des signaux sources et observations, dans le cas de sources vérifiant la contrainte d'additivité totale . . . . .  | 34 |
| 2.5 | Synthèse des méthodes de séparations de sources non-négatives . . . . .  | 35 |
| 2.6 | Sources corrélées : Faces humaines . . . . .   | 41 |
| 4.1 | Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange $\mathbf{A}_1$ et une initialisation par l'orthant positif . . . . . | 68 |
| 4.2 | Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange $\mathbf{A}_2$ et une initialisation par l'orthant positif . . . . . | 69 |
| 4.3 | Évolutions, en fonction du nombre de sources, du nombre moyen de blocages et du coût moyen de calcul, pour une initialisation par l'orthant positif . . . . .            | 70 |
| 4.4 | Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange $\mathbf{A}_1$ et une initialisation par VCA élargi . . . . .        | 71 |
| 4.5 | Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange $\mathbf{A}_2$ et une initialisation par VCA élargi . . . . .        | 71 |
| 4.6 | Évolutions en fonction du nombre de sources, du nombre moyen de blocages et du coût moyen de calcul, pour une initialisation par VCA élargi . . . . .                    | 72 |

|      |   |     |
|------|---|-----|
| 4.7  | Erreurs moyennes de séparation en fonction du nombre d'échantillons, et de la parcimonie des sources . . . . .  | 73  |
| 4.8  | Évolution en fonction du taux de parcimonie et du nombre de sources, du nombre minimum d'échantillons nécessaire pour une séparation exacte . . . . . | 74  |
| 4.9  | Performances en fonction du nombre de sources, pour $\tau = 0.01$ et $\tau = 0.1$ . . . . .   | 75  |
| 4.10 | Performances en fonction du nombre de sources, pour $\tau = 0.5$ et $\tau = 1$ . . . . .  | 76  |
| 4.11 | Effet d'un offset sur le nuage de points des signaux observations . . . . .   | 77  |
| 4.12 | Erreur de séparation en fonction de $\theta$ et de $\phi$ , pour $\rho = 10^{-3}$ . . . . .   | 78  |
| 4.13 | Erreur de séparation en fonction de $\theta$ et de $\phi$ , pour $\rho = 1$ . . . . .   | 78  |
| 4.14 | Nuage de points des signaux observations bruités . . . . .  | 79  |
| 4.15 | Effet du bruit sur le nuage de points des différentes classes . . . . .   | 80  |
| 4.16 | Illustration du cône engendrée par l'estimée de la matrice de mélange, pour $\lambda_b = 0.1$ . . . . .   | 81  |
| 4.17 | Variations des indices moyens de performance, en fonction du taux de parcimonie et du rapport signal sur bruit . . . . .                              | 82  |
| 4.18 | Variations des indices moyens de performance, en fonction du nombre de sources et du rapport signal sur bruit . . . . .                               | 82  |
| 5.1  | Profil d'élution et spectre de masse, attendus pour un métabolite . . . . .   | 89  |
| 5.2  | Chromatogramme total . . . . .  | 90  |
| 5.3  | Chromatogramme sur l'intervalle 12,4821 min - 21,5337 min . . . . .   | 91  |
| 5.4  | Profils d'élution des 6 sources estimées sur l'intervalle 12,4821 min -14,4178 min . . . . .  | 91  |
| 5.5  | Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 12,4821 min -14,4178 min . . . . .                            | 92  |
| 5.6  | Profils d'élution des 8 sources estimées sur l'intervalle 14,4178 min - 16,2441 min . . . . .   | 92  |
| 5.7  | Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 14,4178 min - 16,2441 min . . . . .                           | 93  |
| 5.8  | Profils d'élution des 10 sources estimées sur l'intervalle 16,2441 min - 18,6184 min . . . . .  | 94  |
| 5.9  | Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 16,2441 min - 18,6184 min . . . . .                           | 95  |
| 5.10 | Profils d'élution des 8 sources estimées sur l'intervalle 18,6184 min - 21,5337 min . . . . .   | 96  |
| 5.11 | Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 18,6184 min - 21,5337 min . . . . .                           | 96  |
| 5.12 | Coupes sagittale, coronale et axiale d'une image TEP de fantôme . . . . .   | 98  |
| 5.13 | Compartiments pharmacocinétiques estimés à partir d'images TEP de fantômes . . . . .  | 99  |
| 5.14 | Compartiments pharmacocinétiques estimés à partir d'images TEP du cerveau humain . . . . .  | 101 |
| C.1  | Liste des métabolites annotés, dans l'urine de rat traité au phénobarbital . . . . .  | 115 |



# Liste des tableaux

|      |  |     |
|------|--|-----|
| 2.1  | Listes des algorithmes utilisés pour les simulations . . . . .                               | 36  |
| 2.2  | Erreur de séparation $E_{sep}(dB)$ , pour des sources indépendantes et parcimonieuses .      | 38  |
| 2.3  | Erreur de corrélation $E_{corr}(dB)$ , pour des sources indépendantes et parcimonieuses .    | 38  |
| 2.4  | Négativité des sources estimées $E_{neg}$ , pour des sources indépendantes et parcimonieuses | 38  |
| 2.5  | Coût de calcul $T(s)$ , pour des sources indépendantes et parcimonieuses . . . . .           | 39  |
| 2.6  | Erreur de séparation $E_{sep}(dB)$ , pour des sources indépendantes et uniformes . . . .     | 39  |
| 2.7  | Erreur de corrélation $E_{corr}(dB)$ , pour des sources indépendantes et uniformes . . . .   | 39  |
| 2.8  | Négativité des sources estimées $E_{neg}$ , pour des sources indépendantes et uniformes .    | 40  |
| 2.9  | Coût de calcul $T(s)$ , pour des sources indépendantes et uniformes . . . . .                | 40  |
| 2.10 | Erreur de séparation $E_{sep}(dB)$ , pour des sources corrélées . . . . .                    | 41  |
| 2.11 | Erreur de corrélation $E_{corr}(dB)$ , pour des sources corrélées . . . . .                  | 41  |
| 2.12 | Négativité des sources estimées $E_{neg}$ , pour des sources corrélées . . . . .             | 41  |
| 2.13 | Coût de calcul $T(s)$ , pour des sources corrélées . . . . .                                 | 42  |
| 4.1  | Variations en fonction du seuil de bruit, des indices de performances : $RSB = 40dB$         | 81  |
| 5.1  | Durée d'échantillonnage des images TEP de fantômes . . . . .                                 | 98  |
| 5.2  | Durée d'acquisition des images TEP de cerveau humain . . . . .                               | 100 |



# Introduction générale

Cette thèse rentre dans le cadre du programme **TechnoThèse CEA - Tunisie 2009**, elle s'est déroulée dans le cadre d'une convention de cotutelle entre l'Université de Grenoble (UdG) et l'École Nationale d'Ingénieurs de Tunis (ENIT). Les travaux ont été effectuée alternativement au sein du Laboratoire d'Outils pour l'Analyse de Données (LOAD) du *CEA\DRT\LIST\DCSI* et de l'Unité de recherche Signaux et Systèmes (U2S) de l'ENIT, pour une durée cumulée de 27 mois au LOAD et de 9 mois à U2S. Les données traitées dans ce travail nous ont été fournies par le Laboratoire d'Étude du Métabolisme des Médicaments (LEMM) rattaché au *CEA\DSV\IBITEC - S\SPI* et par le Service Hospitalier Frédéric Joliot (SHFJ) rattaché au *CEA\DSV\I<sup>2</sup>BM*.

## Problématique

Analyser un phénomène complexe comme un mélange linéaire de phénomènes élémentaires est une idée très ancienne qui se justifie chaque fois que les intensités de ces phénomènes élémentaires sont suffisamment faibles pour qu'il y ait superposition sans interactions. Le concept de séparation aveugle de sources, introduit dans les années 80, a permis de développer de nouvelles méthodes pour effectuer cette analyse. De nos jours, les méthodes de séparation aveugle de sources sont utilisées dans plusieurs applications allant du traitement de signaux audio au traitement d'images, en passant par les télécommunications ou encore la chimométrie. Dans de nombreux problèmes, comme la séparation de spectres en chimie ou la séparation d'images (médicales ou hyperspectrales), les sources à séparer sont des grandeurs non-négatives, cette propriété doit être prise en compte lors de la séparation car c'est elle qui donne un sens physique aux sources estimées. Plusieurs algorithmes ont été proposés pour résoudre le problème de séparation aveugle de sources non-négatives. Mais bien qu'efficaces, la plupart de ces algorithmes requièrent de "fortes" hypothèses sur les sources qui ne sont pas toujours vérifiées en pratique (par exemple l'indépendance mutuelle ou la dominance locale des sources) et/ou présentent des problèmes de grande sensibilité au bruit ou de coût de calcul très élevé, comme nous le verrons par la suite.

Ce travail de thèse a pour ambition de développer un nouvel algorithme simple, efficace et robuste de séparation de sources non-négatives, cet algorithme devra être suffisamment générique pour pouvoir être utilisé dans différents problèmes de séparation de sources non-négatives. En particulier, deux situations de séparation de sources non-négatives seront traitées dans cette thèse : la séparation de spectres de masse enregistrés à la sortie d'un

chromatographe en phase liquide pour l'analyse métabolomique de fluides biologiques et la séparation d'images dynamiques de tomographie par émission de positons (TEP) pour l'estimation des pharmacocinétiques d'un radio traceur dans le cerveau humain.

## Contributions de la thèse

Les contributions de ce travail à la littérature de la séparation de sources se situent à trois niveaux :

D'abord, au niveau de l'analyse critique des méthodes existantes de séparation de sources non-négatives. Cette analyse permettra de dégager les forces et les faiblesses des principales méthodes de séparation de sources non-négatives et indiquera la direction suivie dans cette thèse pour le développement d'un nouvel algorithme.

Ensuite, au niveau du développement d'une nouvelle méthode géométrique de séparation de sources non-négatives. Cette partie, qui constitue la contribution majeure de notre travail, nous avons établi les fondements et réalisé l'implémentation algorithmique de la méthode proposée, ainsi que son analyse théorique et expérimentale. Cette méthode requiert peu d'hypothèses sur les sources et présente une meilleure robustesse au bruit que les autres méthodes géométriques.

Enfin, la dernière partie de ce travail est consacrée à l'évaluation de notre algorithme sur différents jeux de données simulées et réelles. D'une part, la séparation de spectres de masse a permis d'identifier les métabolites de différents fluides biologiques. D'autre part, la séparation d'images de tomographie par émission de positons a permis d'estimer, sans prélèvement invasif, les différentes pharmacocinétiques d'un radio-traceur dans le cerveau d'un patient humain, notamment la fonction d'entrée artérielle qui présente un grand intérêt en oncologie.

## Organisation du manuscrit

Ce mémoire est organisé en trois parties. La première partie présente le problème de séparation de sources non-négatives et fait une analyse critique des méthodes existantes afin de dégager leurs forces et leurs faiblesses. La deuxième partie est consacrée au développement d'une nouvelle méthode géométrique de séparation de sources non-négatives et à l'étude théorique et expérimentale de ses performances. La troisième partie présente les résultats de séparation sur différents jeux de spectres de masse et d'images TEP. Une conclusion générale et quatre annexes complètent ce mémoire. La première annexe donne la preuve de la non séparabilité de sources décorréelées et "*well-grounded*" (uniformément distribuées sur une sphère), la deuxième annexe présente quelques mesures de divergence utilisées en factorisation de matrice non-négatives. La troisième donne la liste des métabolites présents dans l'urine de rat traité au phénobarbital et annotés par les biologistes du LEMM, enfin la quatrième annexe dresse la liste des publications que nous avons effectué durant cette thèse.



## **Première partie : État de l'art**

Cette partie du mémoire est composée des chapitres 1 et 2. Dans le chapitre 1, nous présentons d'abord le problème de séparation aveugle de sources. Nous abordons ensuite le cas des sources non-négatives et nous décrivons les problématiques de séparation de spectres de masse en métabolomique et de séparation d'images dynamiques de tomographie par émission de positons en oncologie. Enfin, ce premier chapitre se termine par l'établissement des hypothèses de travail et par une discussion sur le modèle de bruit en situation de sources non-négatives. Le chapitre 2 fait une analyse critique des méthodes statistiques et des méthodes déterministes de séparation de sources non-négatives. Ces méthodes sont décrites, évaluées et comparées dans différents contextes. Cette analyse permettra d'indiquer la direction suivie dans cette thèse pour le développement d'un nouvel algorithme.

## **Deuxième partie : Développement d'une nouvelle méthode géométrique de séparation de sources non-négatives**

La deuxième partie de ce manuscrit est consacrée au développement d'une méthode géométrique de séparation de sources non-négatives, elle regroupe les chapitres 3 et 4. Le chapitre 3 est présenté sous forme d'un article scientifique soumis à la revue IEEE Transactions on Signal Processing le 8 octobre 2012. Nous y développons une nouvelle méthode géométrique pour la séparation de sources non-négatives basée sur les cônes simpliciaux. Nous établissons aussi des conditions nécessaires et des conditions suffisantes sur les sources pour garantir l'unicité de la solution du problème de séparation. Ce chapitre présente aussi des résultats de séparation sur des données synthétiques ainsi que sur des données réelles de spectres de masse et d'images dynamiques TEP. Dans le chapitre 4, nous effectuons une étude expérimentale de performance et nous évaluons la robustesse de la méthode développée. Pour ce faire, nous avons évalué l'influence de l'initialisation, du nombre de sources et du nombre d'échantillons sur les performances de notre méthode. La robustesse de la méthode proposée, au bruit et à une mauvaise estimation du nombre de sources, est aussi évaluée.

## **Troisième partie : Résultats de séparation sur données LC-MS et TEP dynamiques**

Cette dernière partie, composée du chapitre 5, présente les résultats de séparation, à l'aide de l'algorithme proposé, de spectres de masse et d'images TEP. D'une part, nous effectuons la séparation de spectres de masse issus de l'analyse, par spectrométrie de masse couplée à la chromatographie liquide, d'urine de rat, afin d'étudier le métabolisme de rat traité au phénobarbital. D'autre part, nous estimons les pharmacocinétiques du FluoroDeoxyGlucose marqué au fluor 18 ([<sup>18</sup>F]-FDG) dans le cerveau humain, en effectuant la séparation d'images dynamiques de tomographie par émission de positons de cet organe.



Première partie

État de l'art



# Chapitre 1

## Problématique et Applications

### Sommaire

---

|            |  |           |
|------------|--|-----------|
| <b>1.1</b> | <b>Introduction</b>  | <b>9</b>  |
| <b>1.2</b> | <b>Séparation Aveugle de Sources (SAS)</b>                         | <b>9</b>  |
| 1.2.1      | Un seul problème théorique, plusieurs applications pratiques       | 9         |
| 1.2.2      | Mise en équation du problème de SAS                                | 10        |
| 1.2.3      | Trois décennies de séparation de sources                           | 13        |
| <b>1.3</b> | <b>Mélanges linéaires instantanés de sources non-négatives</b>     | <b>13</b> |
| 1.3.1      | Équation de mélange  | 13        |
| 1.3.2      | Quelques situations de mélanges linéaires de sources non-négatives | 14        |
| 1.3.3      | Séparation de spectres de masse en métabolomique                   | 14        |
| 1.3.4      | Séparation d'images dynamiques TEP pour l'oncologie                | 17        |
| <b>1.4</b> | <b>Hypothèses de travail</b>                                       | <b>19</b> |
| <b>1.5</b> | <b>Bruit en situation de sources non-négatives</b>                 | <b>20</b> |
| <b>1.6</b> | <b>Conclusion</b>  | <b>20</b> |

---

## 1.1 Introduction

Nous introduisons dans ce chapitre la problématique de séparation aveugle de sources non-négatives. Dans un premier temps, nous présentons le problème de séparation aveugle de sources dans le cas général, en expliquant les procédures de mélange et de séparation. Dans un second temps, nous abordons le cas des sources non-négatives en décrivant deux applications de natures très différentes qui sont considérées dans ce travail. La première application considérée est la séparation de spectres de masses issus de l'analyse par spectrométrie de masse couplée à la chromatographie liquide (LC-MS) pour la métabolomique de fluide biologique et la deuxième application est la séparation d'images dynamiques de tomographie par émission de positons (TEP) pour l'oncologie. Pour finir, nous justifions les hypothèses qui seront faites sur les sources et sur le système de mélange et nous discutons le modèle de bruit.

## 1.2 Séparation Aveugle de Sources (SAS)

### 1.2.1 Un seul problème théorique, plusieurs applications pratiques

Introduit dans les années 80 par J. Héroult, C. Jutten et B. Ans pour une problématique de décodage du mouvement des vertébrés [HJA85], la séparation aveugle de sources (SAS) a été largement étudiée dans la littérature et se rencontre dans de nombreux problèmes de traitement du signal et des images. En partant d'un ensemble de "signaux complexes" mesurés par des capteurs (observations), la séparation aveugle de sources consiste à estimer les "signaux élémentaires" (sources) qui ont contribué à générer ces signaux complexes, ainsi que le système de mélange, sans aucune connaissance a priori sur les sources et sur le système de mélange. Les techniques de séparation aveugle de sources ont été utilisées dans de nombreux domaines :

- En traitement de signaux audio, pour séparer les différents signaux de paroles d'une conversation (problème de cocktail party) [CC97, MSAM06], pour séparer la voix du fond musical [LB97, Van03, OPBG07], pour séparer les sons issus de différents instruments musicaux [KP00, YW08] ou encore pour la transcription automatique de la musique [PAB<sup>+</sup>02, Ber09].
- En chimométrie, pour séparer et identifier les spectres des différents constituants d'une substance chimique [GPH04, Mou05, ZCHW08, SZC09].
- En électrophysiologie, pour supprimer les artéfacts des signaux électrocardiographiques (ECG) et électroencéphalographiques (EEG) [BMO98, MPC<sup>+</sup>07, KAS07] ou pour décomposer les signaux électromyographiques (EMG) [FLF<sup>+</sup>03, LYC<sup>+</sup>05].
- En traitement d'images médicales, pour séparer les images des différents constituants d'un organe [LLCL01, ZLY06, CAC07].
- En télécommunications, pour la radio surveillance [CT03, IIA07], la réduction d'interférences [XB99, BTMF04] ou encore pour la radio cognitive [ZXY09, LW11].
- En analyse d'images hyperspectrales, pour identifier les différents éléments présents dans une scène observée par un spectro-imageur [NBD05, DKS06, MQ07].

- En sismologie, pour estimer les différentes sources à l’origine d’un séisme [TML95, Kil03, ACM<sup>+</sup>03].

D’autres applications des méthodes de SAS telles que la finance, l’analyse de documents textes ou encore la gestion de machines tournantes sont présentées dans les livres consacrés à la séparation de sources, nous pouvons consulter pour cela [AH01, JC07, CZPA09, CJ10]. Le paragraphe 1.2.2 met en équation le problème de séparation aveugle de sources en expliquant les procédures de mélange et de séparation.

## 1.2.2 Mise en équation du problème de SAS

### 1.2.2.1 Le mélange

Notons par  $\mathbf{s}$  le vecteur des sources,  $\mathbf{s}(k) = [s_1(k) \ s_2(k) \ \cdots \ s_n(k)]^T$  où  $s_i(k)$  est le  $k$ ème échantillon de la source élémentaire  $i$ , et par  $\mathbf{x}$  le vecteur des observations,  $\mathbf{x}(k) = [x_1(k) \ x_2(k) \ \cdots \ x_m(k)]^T$  où  $x_j(k)$  est le  $k$ ème échantillon du signal issu du capteur  $j$  (observation  $j$ );  $n$  le nombre de sources et  $m$  le nombre d’observations. Suivant le cas,  $k$  peut désigner un indice temporel, un indice spatial ou un indice fréquentiel. La figure 1.1 illustre la procédure de mélange. La relation entre les sources et les observations (mélanges des sources)

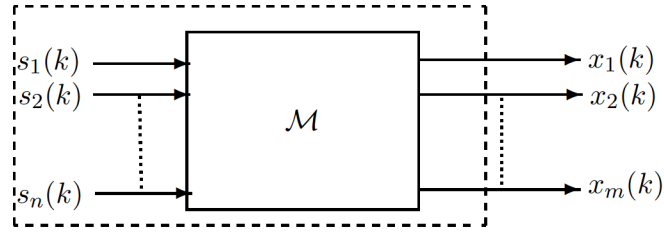


FIGURE 1.1 – Procédure de mélange

est exprimée par l’équation (1.1) où  $\mathcal{M}$  est le système de mélange :

$$\mathbf{x}(k) = \mathcal{M}[\mathbf{s}(k)] \quad (1.1)$$

Si l’on dispose de  $K$  échantillons de chaque source (et donc de chaque observation), alors on peut regrouper les échantillons de sources et d’observations respectivement dans les matrices  $\mathbf{S}$  et  $\mathbf{X}$ . L’équation de mélange peut se réécrire sous la forme matricielle donnée par (1.2) :

$$\mathbf{X} = \mathcal{M}[\mathbf{S}] \quad (1.2)$$

où  $\mathbf{X} \in \mathbb{R}^{m \times K}$  est la matrice des observations et  $\mathbf{S} \in \mathbb{R}^{n \times K}$  est la matrice des sources.

Le système de mélange  $\mathcal{M}$ , dépend de la physique des phénomènes que l’on mesure. Par exemple les signaux mesurés par des électrodes placés sur le cuir chevelu peuvent être modélisés comme des mélanges *linéaires instantanés* de différents signaux électrophysiologiques. Dans ce cas, le système de mélange est modélisé par une matrice  $\mathbf{A}$  et la relation entre les sources et les observations est donnée par l’équation (1.3) :

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{b}(k) \quad (1.3)$$

où  $\mathbf{A} \in \mathbb{R}^{m \times n}$  est la matrice de mélange et  $\mathbf{b}(k)$  est un vecteur de bruit additif qui exprime les erreurs de mesures et de modélisation.

Dans un autre contexte, les mélanges de signaux audio enregistrés par des microphones sont *linéaires convolutifs*, la relation entre les sources et les observation est donnée par l'équation (1.4) :

$$\mathbf{x}(k) = \mathbf{H}(k) * \mathbf{s}(k) + \mathbf{b}(k) \quad (1.4)$$

où  $\mathbf{H}(k)$  est le filtre de mélange et  $\mathbf{b}(k)$  est un vecteur de bruit additif qui exprime les erreurs de mesures et de modélisation.

En potentiométrie, les mesures des activités de différents ions à l'aide d'électrodes sélectives ioniques, ont été modélisées comme des mélanges *non linéaires* de ces activités. Le modèle de mesure est donnée par l'équation (1.5) [DJM09] :

$$x_i(k) = c_i + d_i \log \left( s_i(k) + \sum_{j \neq i} a_{ij} s_j(k)^{\frac{z_i}{z_j}} \right) \quad (1.5)$$

où  $c_i$ ,  $d_i$ ,  $z_i$  et  $z_j$  sont des paramètres du modèle et  $a_{ij}$  est la proportion de la source  $j$  dans l'observation  $i$ .

Les équations (1.3), (1.4) et (1.5) constituent trois modèles d'observations différents, de complexités différentes, associés à trois phénomènes physiques différents. Il est donc nécessaire de modéliser le plus fidèlement le système de mélange, qui exprime la relation entre les sources et les observations, pour pouvoir estimer correctement les sources initiales.

Par ailleurs, indépendamment de  $\mathcal{M}$ , on peut distinguer trois situations de mélange en fonction du nombre d'observations relativement au nombre de sources :

- Le mélange *sous-déterminé*, quand le nombre d'observations est inférieur au nombre de sources ( $m < n$ )
- Le mélange *déterminé*, quand le nombre d'observations est égal au nombre de sources ( $m = n$ )
- Le mélange *sur-déterminé*, quand le nombre d'observations est supérieur au nombre de sources ( $m > n$ )

Dans le problème de séparation *aveugle* de sources, les sources et le système de mélange *ne sont pas connus* a priori, on ne dispose que des observations.

### 1.2.2.2 La séparation

En partant uniquement des observations (mélanges des sources), la séparation aveugle de sources consiste à construire un système de séparation  $\mathcal{S}$ , qui permet d'estimer au mieux les sources initiales. La figure 1.2 illustre la procédure de séparation. Si l'on note par  $\hat{\mathbf{s}}$  le vecteur des sources estimées,  $\hat{\mathbf{s}}(k) = [\hat{s}_1(k) \hat{s}_2(k) \cdots \hat{s}_n(k)]^T$  où  $\hat{s}_i(k)$  est le  $k$ ième échantillon de la source estimée  $i$ , alors l'équation de séparation est donnée par (1.6) :

$$\hat{\mathbf{s}}(k) = \mathcal{S}[\mathbf{x}(k)] \quad (1.6)$$

Dans le cas matriciel, l'équation de séparation est donnée par (1.7) où  $\hat{\mathbf{S}} \in \mathbb{R}^{n \times K}$  est la matrice des sources estimées :

$$\hat{\mathbf{S}} = \mathcal{S}[\mathbf{X}] \quad (1.7)$$



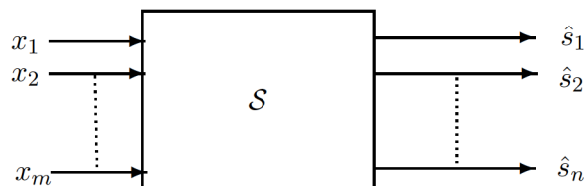


FIGURE 1.2 – Procédure de séparation

### Approches de séparation

Deux approches peuvent être considérées pour la séparation : *l'approche directe* et *l'approche indirecte* :

L'*approche directe* consiste à calculer d'abord  $\hat{\mathcal{M}}$ , estimation du système de mélange. Si le mélange est inversible (cas déterminé et cas sur-déterminé) alors les sources sont estimées par simple inversion de ce système. Si le mélange n'est pas inversible (cas sous-déterminé), la connaissance de  $\hat{\mathcal{M}}$  (ou même de  $\mathcal{M}$ ) ne suffit pas à retrouver les sources initiales. En effet, l'estimation des sources à partir des observations et de l'estimée du système de mélange est un problème mal posé car il consiste à résoudre un système d'équations avec plus d'inconnues que d'équations, il y a donc une infinité de solutions. Dans cette dernière situation, on peut réduire l'ensemble des solutions admissibles en prenant en compte des connaissances a priori sur les sources, nous reviendrons plus loin sur cet aspect.

L'*approche indirecte* consiste à estimer une inverse du système de mélange ( $\hat{\mathcal{M}}^{-1}$ ), à partir des observations. Ceci n'est possible que dans les cas déterminé et sur-déterminé. Les sources sont ensuite obtenues par application de ( $\hat{\mathcal{M}}^{-1}$ ) aux observations.

### Quid de l'aveugle ?

Le problème de *séparation aveugle de sources* (c'est à dire sans hypothèses sur les sources et sur le système de mélange) est un problème mal posé, et ce indépendamment du modèle de mélange et de l'approche que l'on adopte pour la séparation (approche directe ou approche indirecte). En effet, tout couple  $(\mathcal{W}, \mathbf{Y})$  tel que  $\mathcal{W}(\mathbf{Y}) = \mathbf{X}$  est une solution admissible du problème aveugle, et sans aucune information a priori sur les sources et/ou sur le système de mélange, il n'est pas justifié de privilégier une solution admissible par rapport à une autre. Pour résoudre ce problème de nombre infini de solutions, il est donc nécessaire de faire des hypothèses sur les sources et/ou sur le système de mélange afin de parvenir à une solution unique ou tout au moins de réduire le nombre de solutions admissibles. Ainsi, en fonction des informations a priori dont on dispose sur le problème, on peut favoriser un type de solutions en imposant des contraintes sur les sources estimées comme par exemple leur indépendance statistique, leur parcimonie, leur non-négativité ou encore leur cyclostationnarité, etc..., ces contraintes permettent d'estimer des grandeurs qui correspondent à la physique du problème

que l'on considère.

### Indéterminations

Malgré les hypothèses que l'on fait sur les sources et/ou sur le système de mélange, plusieurs travaux ont montré que les sources peuvent être au mieux estimées à une permutation et à une distorsion près. En effet, si l'on cherche une solution  $\hat{\mathbf{s}}$  du problème de séparation de sources dont les composantes vérifient une certaine propriété (indépendance mutuelle, parcimonie, non-négativité, etc...), alors pour toute matrice de permutation  $\Pi$ , les composantes de  $\Pi\hat{\mathbf{s}}$  vérifient les mêmes propriétés que celle de  $\hat{\mathbf{s}}$ ,  $\Pi\hat{\mathbf{s}}$  est donc aussi une solution admissible du problème de séparation de sources.

La distorsion sur les sources estimées peut être plus ou moins sévère suivant le modèle de mélange. Si le mélange est linéaire instantané, la distorsion se réduit à une simple matrice diagonale (facteurs d'échelles) [Com94]. Dans le cas d'un mélange linéaire convolutif, chaque source est estimée à un filtre scalaire près.

### 1.2.3 Trois décennies de séparation de sources

Dans l'absolue, nous pensons qu'il est impossible d'explorer toutes les facettes de la bibliographie de trois décennies de séparation de sources dans un paragraphe. Nous nous contentons donc de donner ici un aperçu rapide des grandes lignes vues sous notre angle.

Les premiers travaux sur la séparation de sources et la grande majorité des auteurs ont discuté la tâche de séparation de mélanges linéaires (instantanés ou convolutifs) de sources mutuellement indépendantes. Le concept d'analyse en composantes indépendantes (ACI) a ainsi été introduit dans les années 90 [Com94] et est aujourd'hui la méthode traditionnelle de séparation de sources. Quelques travaux ont abordé aussi la séparation de sources corrélées [XNN10]. Par la suite d'autres connaissances a priori sur les sources ont été considérées, conduisant à d'autres types de méthodes. Ainsi l'utilisation de la propriété de parcimonie des sources a conduit à l'introduction de l'analyse en composantes parcimonieuses dans les années 2000 [ZP01, LCiA03]. Le problème de séparation de mélanges non linéaires et post-non linéaires a été abordé dès la fin des années 90, mais comparativement au cas linéaire, ce problème a été peu étudié dans la littérature [HP99, TJ99, APJ01, DH07, DJM09].

Dans ce travail, nous nous intéressons à la séparation de mélange linéaires instantanés de sources non-négatives.

## 1.3 Mélanges linéaires instantanés de sources non-négatives

### 1.3.1 Équation de mélange

Le modèle de mélange linéaire instantané est donné par l'équation (1.8) :

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{b}(k) \quad (1.8)$$

où  $\mathbf{A} \in \mathbb{R}^{m \times n}$  est la matrice de mélange ( $\mathbf{A} = [a_{ij}]_{1 \leq i \leq m, 1 \leq j \leq n}$  où  $a_{ij}$  est la contribution de la source  $j$  à l'observation  $i$ ) et  $\mathbf{b}(k)$  est un vecteur de bruit additif qui exprime les erreurs de

mesure et de modélisation ( $\mathbf{b}(k) = [b_1(k) b_2(k) \cdots b_m(k)]^T$  et  $b_j(k)$  est le  $k$ ème échantillon du bruit sur l'observation  $j$ ). Ce modèle de mélange peut s'écrire aussi par l'équation matricielle (1.9) :

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{B} \quad (1.9)$$

où  $\mathbf{X} \in \mathbb{R}^{m \times K}$  est la matrice des observations,  $\mathbf{A} \in \mathbb{R}^{m \times n}$  est la matrice de mélange,  $\mathbf{S} \in \mathbb{R}^{n \times K}$  est la matrice des sources et  $\mathbf{B} \in \mathbb{R}^{m \times K}$  est la matrice de bruit.

### 1.3.2 Quelques situations de mélanges linéaires de sources non-négatives

Dans de nombreux problèmes de séparation de sources, les sources à estimer sont connues pour être des grandeurs non-négatives. Cette contrainte doit être prise en compte lors de la séparation car c'est elle qui donne un sens physique aux sources estimées. Par exemple, en séparation d'images (médicales, hyperspectrales, etc...) [ZLY06, CAC07, MQ07], les sources sont caractérisées par des grandeurs non-négatives (collections de pixels ou de voxels). De même en chimimétrie, la séparation de sources consiste le plus souvent à estimer les spectres (Raman, Infrarouges, spectres de masse) des différents constituants d'une substance chimique [Mou05, ZCHW08, SZC09], ces spectres sont également des grandeurs non-négatives. Le contexte de mélanges linéaires instantanés de sources non-négatives se rencontre aussi lorsqu'il s'agit de séparer les spectres de puissance de signaux audio [FBD09] ou sismiques [CCJ12].

La situation de séparation de mélanges linéaires instantanés de sources non-négatives se rencontre donc dans de nombreux problèmes, la séparation de spectres de masse en métabolomique et la séparation d'images dynamiques de Tomographie par Emission de Positons pour l'oncologie sont deux exemples concrets de ces problèmes.

### 1.3.3 Séparation de spectres de masse en métabolomique

#### 1.3.3.1 Métabolomique

La *métabolomique* est une science récente qui étudie le *métabolome* (ensemble des métabolites) des fluides biologiques [Por07, Tch11]. Un *métabolite* est une molécule organique de petite taille résultant du métabolisme (ensemble des transformations moléculaires et énergétiques se produisant à l'intérieur d'une cellule, d'un tissu ou d'un organisme vivant) [Har08, Mat11]. Parmi les métabolites, on peut citer les sucres, les acides aminés et les acides gras, les vitamines, les antibiotiques, etc... La métabolomique est utilisée dans de nombreux domaines comme la santé pour le diagnostic de maladie métabolique ou pour étudier l'effet d'un corps étranger sur le métabolisme d'un organisme vivant, la pharmacologie/toxicologie pour prédire les effets toxiques d'un médicament dans sa phase de développement, la nutrition et l'agroalimentaire pour l'identification de bio-marqueurs ou pour la traçabilité des organismes génétiquement modifiés (OGM) [Por07, Séb09, SP09].

L'une des techniques d'étude métabolomique est la spectrométrie de masse couplée à la chromatographie liquide (LC-MS pour Liquid Chromatography - Mass Spectrometry). Les paragraphes suivants décrivent ces deux techniques et leur combinaison.

### 1.3.3.2 Chromatographie liquide

La *Chromatographie* est une technique physique de séparation des différents constituants (appelés aussi *analytes*) d'un mélange homogène, grâce à leurs propriétés physico-chimiques (charge électrique, taille, forme, polarité, hydrophobicité, volatilité, ...) [Bri05, Cuq07, Gil08]. En chromatographie liquide (LC pour Liquid Chromatography), le mélange à séparer est entraîné par un liquide (appelé phase mobile) le long d'une colonne fixe (appelée phase stationnaire) [Bri05, Cuq07]. Chaque analyte du mélange parcourt la colonne à une vitesse qui lui est propre, en fonction de sa solubilité dans la phase mobile et de son affinité pour la phase fixe qui tend à le retenir. Les différents analytes sont ainsi séparés et élués<sup>1</sup> de la colonne chromatographique pendant un laps de temps appelé *profil d'éluion*. Le maximum du profil d'éluion est appelé *temps de rétention*. Un détecteur placé à la sortie de la colonne couplé à un enregistreur permet d'obtenir un tracé appelé *chromatogramme*, qui représente l'évolution de la concentration de chaque analyte en fonction du temps. Les analytes élués de la colonne chromatographique sont analysés par spectrométrie de masse afin de les identifier et de les quantifier.

### 1.3.3.3 Spectrométrie de masse

La *Spectrométrie de masse* (MS pour Mass Spectrometry) est une méthode d'analyse qui permet de déterminer la structure et la masse moléculaire d'un composé chimique [San05, Qui08]. A l'intérieur du *spectromètre de masse*, le composé à analyser est d'abord ionisé puis passe dans un analyseur qui trie les différents ions en fonction de leurs rapports masse sur charge ( $m/z$ ). Un détecteur placé à la suite de l'analyseur recueille les ions et produit un signal électrique proportionnel au nombre d'ions détectés. A la sortie du spectromètre de masse, un système informatique transforme le signal électrique en *spectre de masse*, formé d'une série de pics répartis sur une ligne de base. Chaque pic caractérise un ion particulier de la même molécule, de rapport masse sur charge donné, la hauteur de chaque pic est proportionnelle au nombre d'ions. La figure 1.3 illustre le spectre de masse de la pyridine. Nous pouvons remarquer que le caractère parcimonieux du spectre de masse. Lorsque la résolution du spectromètre de masse est suffisamment élevée, le spectre de masse d'un molécule peut permettre de déterminer sa formule brute.

### 1.3.3.4 Intérêt de la séparation de spectres de masse en métabolomique

Au Laboratoire d'Etude du Métabolisme des Médicaments (LEMM) de la Direction des Sciences du Vivant (DSV) du CEA, la chaîne d'étude métabolomique est formée du spectromètre de masse Orbitrap, couplé à un chromatographe en phase liquide ; la figure 1.4 illustre cette chaîne d'étude.

Cependant, les fluides biologiques analysés en métabolomique sont, dans la plupart des cas pratiques, des mélanges très complexes de plusieurs dizaines, voire centaines, de métabolites. Dans ces conditions, la colonne de chromatographie ne peut complètement séparer

---

1. l'éluion est l'action d'entraîner des espèces chimiques par un fluide, appelé éluant, en vue de les séparer

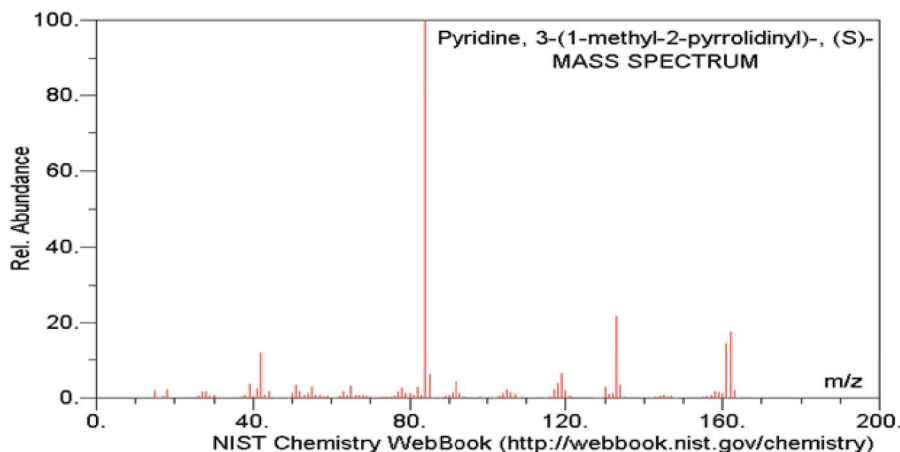


FIGURE 1.3 – Spectre de masse de la nicotine, également appelée pyridine

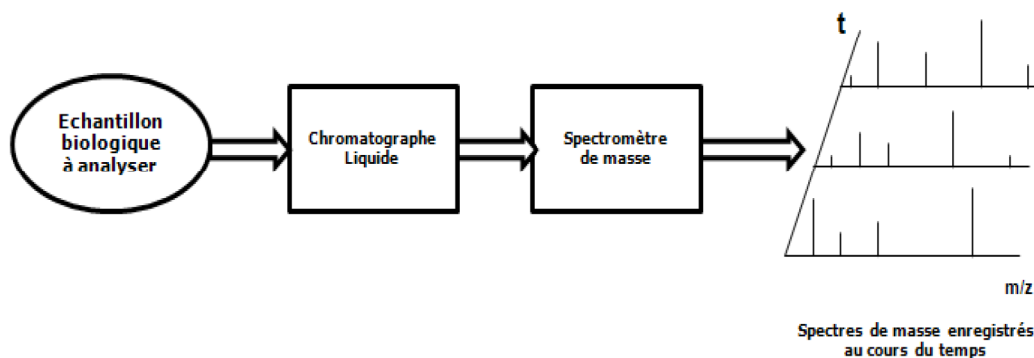


FIGURE 1.4 – Chaîne d'étude métabolomique au CEA-iBiTec-S/LEMM

les différents profils d'élutions; ces profils sont distincts mais ils se recouvrent parfois très fortement. Les spectres de masse mesurés à la sortie du chromatographe peuvent ainsi être modélisés comme des *mélanges linéaires instantanés* des spectres de masse des différents métabolites co-élus.

L'objectif est de séparer totalement les spectres de masse élémentaires de différents métabolites co-élus et d'estimer leurs profils d'élutions. La paire profil d'élution-spectre de masse, illustrée sur la figure 1.5, permet d'identifier et de quantifier un métabolite. Il s'agit donc d'un problème de séparation de mélanges linéaires instantanés de sources non-négatives. Les sources à séparer sont les spectres de masses des différents métabolites, la ligne  $\mathbf{s}_j$  de la matrice des sources  $\mathbf{S}$  contient le spectre de masse du métabolite  $j$  et la colonne  $\mathbf{a}_j$  de la matrice de mélange  $\mathbf{A}$  contient le profil d'élution du métabolite  $j$ . La ligne  $\mathbf{x}_i$  de la matrice des observations  $\mathbf{X}$  contient le spectre de masse mesuré par le spectromètre à l'instant  $i$ .

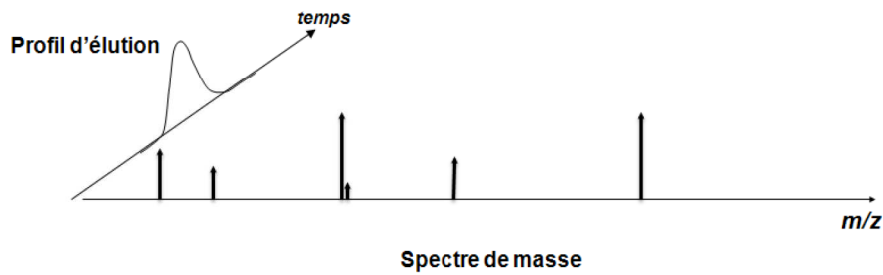


FIGURE 1.5 – Illustration du profil d'éluion et du spectre de masse de métabolite

### 1.3.4 Séparation d'images dynamiques de tomographie par émission de positons pour l'oncologie

#### 1.3.4.1 Tomographie par émission de positons

La tomographie par émission de positons (TEP) est une technique d'imagerie médicale qui permet de voir en trois dimensions l'activité métabolique d'un organe [dDMF<sup>+</sup>02, Che04]. Pour ce faire, un traceur radio-actif (molécule organique marquée avec un émetteur de positons comme le Carbone 11, le Fluor 18, l'Oxygène 15 ou encore de l'Azote 13) est introduit dans le sang. On peut noter que 90% des examens TEP utilisent comme radio-traceur le FluoroDéoxyGlucose (FDG) marqué au Fluor 18. L'annihilation d'un positon émis par le radio-traceur avec un électron entraîne l'émission de deux photons qui partent dans la même direction mais dans des sens contraires. La détection en coïncidence de ces photons par le colimateur TEP permet de localiser leur lieu d'émission. On peut ainsi déterminer la répartition du traceur dans l'organe.

#### 1.3.4.2 Oncologie et imagerie TEP

L'oncologie est la discipline de la médecine qui diagnostique et traite les tumeurs cancéreuses. En oncologie, l'examen TEP permet de détecter les cancers, de faire le bilan d'extension des cancers et d'évaluer l'efficacité des traitements anticancéreux.

Dans le cadre de l'évaluation de l'efficacité d'un traitement anticancéreux, lorsqu'un radio-traceur est injecté dans le sang, on s'intéresse à ses pharmacocinétiques qui indiquent l'évolution temporelle de la concentration du traceur dans les différents compartiments<sup>2</sup> (tissu, sang, etc...) de l'organe traité [TS08]. Parmi ces pharmacocinétiques, la cinétique artérielle des traceurs injectés et non métabolisés présente une importance particulière. En effet, elle permet de mesurer la quantité de traceur disponible pour les cellules de l'organe, d'où son nom de fonction d'entrée artérielle (AIF pour Arterial Input Function). Il n'existe aujourd'hui qu'une seule méthode de référence de cette mesure, celle invasive par prélèvements sanguins artériels. Cette méthode est de moins en moins utilisée chez l'homme en raison de la difficulté technique à poser un cathéter artériel et des risques encourus par le patient au cours de l'acte. En pré-clinique, elle est peu utilisée chez le rongeur à cause des difficultés

2. un compartiment est une partie de l'organe dans laquelle la concentration du traceur est homogène

techniques de l'acte et de la quantité de sang qu'il faut prélever à l'animal.

### 1.3.4.3 Intérêt de la séparation d'images dynamiques TEP en oncologie

L'objectif ici est d'estimer de manière non invasive les pharmacocinétiques du traceur FluoroDeoxyGlucose marqué au fluor 18 ([18F]-FDG) dans le cerveau humain à partir uniquement d'ensemble d'images TEP de cet organe reconstruites à différents instants. La figure 1.6 montre les coupes sagittale, coronale et axiale d'une image TEP de cerveau humain.

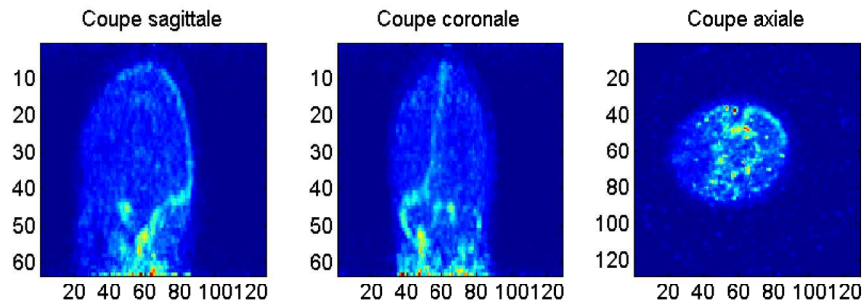


FIGURE 1.6 – Image TEP de cerveau humain

Si l'on suppose que la concentration du traceur radioactif d'un compartiment donné est homogène dans toute son extension spatiale - concentration homogène dans tout le réseau artériel visible sur les images par exemple - alors la contribution de ce compartiment aux différentes images TEP peut se factoriser comme le produit de la concentration du traceur (une fonction de temps exprimée en  $Bq/cm^3$ ) par la fraction volumique (sans dimension) de chaque voxel des images occupées par le compartiment en question (une fonction de l'espace). Cette hypothèse, et donc la factorisation temps/espace qui en découle, sont assez bien vérifiées à l'échelle d'un organe comme le cerveau où le temps d'homogénéisation d'un compartiment est très inférieur au temps de prise de chaque images. Ainsi, chaque image TEP reconstruite peut être modélisée comme un *mélange linéaire instantané* des images TEP des différents compartiments pharmacocinétiques du cerveau [TS08][ZFFM<sup>+</sup>09]. Le modèle de mélange des images TEP est illustré sur la figure 1.7

On cherche donc à estimer la distribution spatiale ainsi que la cinétique (évolution temporelle de la concentration) du traceur dans chaque compartiment. Il s'agit là aussi d'un problème de séparation de mélanges linéaires instantanés de sources non-négatives. Les sources à séparer sont les distributions spatiales du traceur dans les différents compartiments du cerveau, la ligne  $\mathbf{s}_j$  de la matrice des sources  $\mathbf{S}$  contient les échantillons de la distribution spatiale du traceur dans le compartiment  $j$ , obtenus par vectorisation de l'image 3D de ce compartiment. La colonne  $\mathbf{a}_j$  de la matrice de mélange  $\mathbf{A}$  contient l'évolution de la concentration du traceur dans le compartiment  $j$ . La ligne  $\mathbf{x}_i$  de la matrice des observations  $\mathbf{X}$  contient les échantillons de l'image TEP reconstruite à l'instant  $i$ , obtenus par vectorisation de l'image 3D.

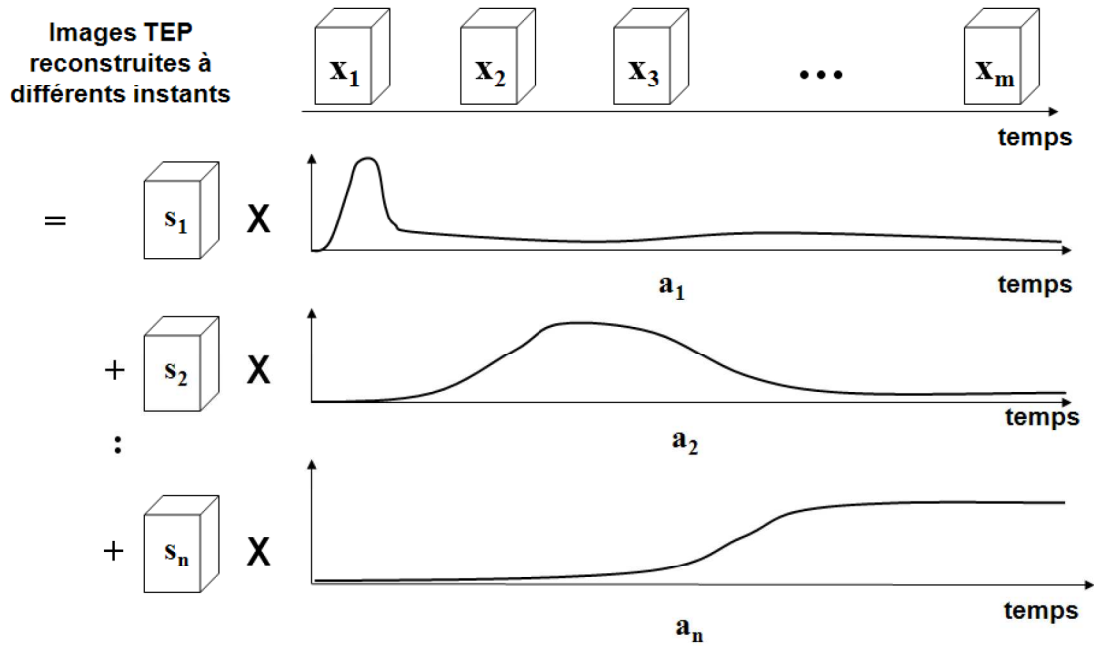


FIGURE 1.7 – Modèle de mélange des images TEP

## 1.4 Hypothèses de travail

La première hypothèse que nous ferons dans la suite de ce travail est la non-négativité des sources. Nous nous plaçons aussi dans la situation où le nombre d'observations est supérieur ou égal au nombre de sources et où la matrice de mélange est de rang plein. Ces trois hypothèses sont non seulement cohérentes avec les données que nous traitons (spectres de masse et images TEP), mais elles sont vérifiées dans les autres applications évoquées dans le paragraphe 1.3.2. Ces hypothèses sont mises en équations ci-dessous :

(H1) Les sources sont non-négatives :  $\forall 1 \leq i \leq n, \forall 1 \leq k \leq K, s_i(k) \geq 0$

(H2) Le mélange est déterminé ou surdéterminé :  $m \geq n$

(H3) La matrice de mélange est de rang plein :  $rg(\mathbf{A}) = n$

Il faut noter que dans les problèmes de séparation de spectres de masse pour la métabolomique et de séparation d'images dynamiques TEP pour l'oncologie, les matrices de mélange sont aussi à coefficients non-négatifs. Cette connaissance n'est cependant pas exploitée explicitement car, comme nous le verrons par la suite, elle n'est pas nécessaire pour le développement des méthodes que nous envisageons.

Par ailleurs, étant donné que le modèle de mélange est linéaire instantané, les sources peuvent au mieux être estimées à une matrice de permutation et à une matrice diagonale près (voir paragraphe 1.2.2.2). Dans le cas des sources non-négatives, la matrice diagonale est elle aussi à entrées non-négatives.



## 1.5 Bruit en situation de sources non-négatives

Lorsque les sources sont non-négatives, une attention particulière doit être accordée à la modélisation du terme de bruit  $\mathbf{B}$  dans l'équation (1.9). Lorsque la matrice de mélange est à coefficients quelconques (positifs et négatifs), le modèle générique de bruit, indépendant des données et gaussien, pourrait être envisagé. Par contre, lorsque la matrice de mélange est aussi à coefficients non-négatifs, ce modèle générique de bruit indépendant et gaussien peut ne pas être cohérent avec les données non-négatives fournies par les appareils de mesure. Dans ce cas, la question du modèle de bruit reste ouverte. Si l'on dispose de suffisamment de connaissances a priori, il est préférable d'adopter un modèle de bruit adapté à l'appareil de mesure. Par exemple, en spectrométrie de masse, il a été constaté que la puissance du bruit dépend de la puissance du signal non bruité [HFR<sup>+</sup>09][BMD<sup>+</sup>11], ce qui laisse penser que le bruit a une composante additive et une composante multiplicative. A défaut d'informations sur le bruit généré par l'appareil de mesure, nous pouvons considérer une distribution à support positif, reste à répondre à la question : quelle distribution choisir ? (Gamma, Exponentielle, Uniforme). Dans ce travail, nous avons adopté un bruit indépendant et gaussien, tronqué sur  $\mathbb{R}_+$ .

## 1.6 Conclusion

Dans ce premier chapitre, nous avons introduit la séparation aveugle de sources en décrivant les procédures de mélange et de séparation. Nous avons pu voir que le problème de séparation de mélanges linéaires instantanés de sources non-négatives se rencontre dans plusieurs problèmes, en particulier lorsqu'il s'agit de séparer des spectres de masse en métabolomique ou des images dynamiques TEP en l'oncologie. Nous avons établi dans ce chapitre aussi les hypothèses de travail et mener une première discussion du modèle de bruit en situation de sources non-négatives.

Le prochain chapitre fait une analyse critique des méthodes existantes de séparation de sources non-négatives afin de dégager leurs avantages et leurs inconvénients respectifs.



# Chapitre 2

## Analyse critique des méthodes de SAS non-négatives

### Sommaire

---

|            |  |           |
|------------|--|-----------|
| <b>2.1</b> | <b>Introduction</b>  | <b>23</b> |
| <b>2.2</b> | <b>Méthodes statistiques</b>                                     | <b>23</b> |
| 2.2.1      | Analyse en composantes indépendantes non-négatives               | 23        |
| 2.2.2      | Méthode de maximum de vraisemblance                              | 25        |
| 2.2.3      | Inférence bayésienne   | 27        |
| <b>2.3</b> | <b>Méthodes déterministes</b>                                    | <b>27</b> |
| 2.3.1      | Factorisation en Matrices Non-négatives                          | 27        |
| 2.3.2      | Approches géométriques   | 29        |
| <b>2.4</b> | <b>Synthèse des méthodes de SAS non-négatives</b>                | <b>34</b> |
| <b>2.5</b> | <b>Premiers résultats de séparation et premières conclusions</b> | <b>36</b> |
| 2.5.1      | Critères de performances   | 37        |
| 2.5.2      | Génération de la matrice de mélange et du bruit                  | 37        |
| 2.5.3      | Résultats de séparation de sources indépendantes                 | 38        |
| 2.5.4      | Résultats de séparation de sources corrélées                     | 40        |
| 2.5.5      | Synthèse des résultats de séparation                             | 42        |
| <b>2.6</b> | <b>Conclusion</b>  | <b>43</b> |

---

## 2.1 Introduction

Ce deuxième chapitre présente les méthodes existantes de séparation de sources non-négatives. Pour ce faire, nous distinguons deux grandes familles de méthodes : **les méthodes statistiques** et **les méthodes déterministes**. Les méthodes statistiques considèrent sources, et parfois aussi les coefficients de la matrice de mélange, comme des variables aléatoires. Elles utilisent les propriétés statistiques de ces variables aléatoires (notamment leur indépendance mutuelle) pour concevoir les algorithmes de séparation. Les méthodes déterministes, elles, utilisent les propriétés algébriques ou géométriques des sources pour effectuer la séparation. Nous adoptons dans ce chapitre une démarche critique, notamment sur les hypothèses faites par ces différentes méthodes, afin de dégager les forces et les faiblesses des méthodes existantes. Ainsi, après la présentation des principes de fonctionnement de ces méthodes, nous évaluons leurs capacités à séparer des sources indépendantes et des sources corrélées, en absence et en présence de bruit.

## 2.2 Méthodes statistiques

Dans cette première famille de méthodes, les sources (et parfois aussi les coefficients de la matrice de mélange) sont modélisées par des variables aléatoires. Ainsi  $s_i(k)$ , c'est l'échantillon  $k$  de la source  $i$ , est la  $k$ -ième réalisation de la variable aléatoire  $s_i$ . Dans certaines méthodes, le coefficient  $a_{lj}$  de la matrice de mélange est aussi la  $l$ -ième réalisation de la variable aléatoire  $a_j$ . Les approches statistiques de séparation de sources non-négatives sont fondées sur l'hypothèse de l'indépendance statistique des sources. Il convient de préciser qu'un ensemble de variables aléatoires  $s_1, s_2, \dots, s_n$  sont dites statistiquement indépendantes si :

$$Pr(s_1, s_2, \dots, s_n) = \prod_{i=1}^n Pr_i(s_i) \quad (2.1)$$

où  $Pr(s_1, s_2, \dots, s_n)$  est la densité de probabilité jointe des variables aléatoires  $s_1, s_2, \dots, s_n$  et  $Pr_i(s_i)$  est la densité de probabilité marginale de la variable aléatoire  $s_i$ . Parmi les méthodes statistiques, nous distinguons, l'analyse en composantes indépendantes non-négatives, la séparation de sources par maximum de vraisemblance et l'approche bayésienne.

### 2.2.1 Analyse en composantes indépendantes non-négatives

L'analyse en composantes indépendantes non-négatives (N-ICA pour Nonnegative Independent Component Analysis) a été introduite au début des années 2000 [Plu02], pour prendre en compte explicitement la contrainte de non-négativité des sources dans l'analyse en composantes indépendantes (ACI). N-ICA est une méthode indirecte de séparation. Plumbley a montré que la séparabilité des sources est garantie lorsque les conditions suivantes sont vérifiées [Plu02] :

- i. Les sources sont non-négatives
- ii. Les sources sont statistiquement indépendantes

iii. Les sources ont une probabilité non nulle d'être distribuées au voisinage de zero :

$$\forall \delta > 0, Pr(s_i < \delta) > 0, \forall 1 \leq i \leq n.$$

Les sources vérifiant l'hypothèse iii. sont dites “*well-grounded*”. A défaut d'un mot adapté en français, nous garderons ce terme anglo-saxon dans la suite de ce mémoire. Aucune contrainte de non-négativité n'est posée sur la matrice de mélange, supposée déterministe. Les sources sont estimées en effectuant d'abord un blanchiment des observations, suivi d'une rotation des observations blanchies de sorte à les rendre non-négatives [Plu02], cela est décrit étapes de blanchiment et de rotation ci-dessous.

## Blanchiment et réduction de dimension

La première étape consiste à blanchir les observations. Soit  $\mathbf{C}_\mathbf{X}$  la matrice de covariance des observations dont les coefficients sont calculés suivant l'équation (2.2) :

$$[\mathbf{C}_\mathbf{X}]_{ij} = \frac{1}{K} \sum_{k=1}^K [x_i(k) - \mu_{\mathbf{x}_i}] [x_j(k) - \mu_{\mathbf{x}_j}] \quad (2.2)$$

où  $\mu_{\mathbf{x}_l}$  est la moyenne de l'observation  $l$ . En effectuant la décomposition en éléments propres de  $\mathbf{C}_\mathbf{X}$ , on obtient :

$$\mathbf{C}_\mathbf{X} = \mathbf{E} \mathbf{D} \mathbf{E}^T \quad (2.3)$$

où  $\mathbf{D}$  est la matrice diagonale des valeurs propres de  $\mathbf{C}_\mathbf{X}$  et  $\mathbf{E}$  est la matrice orthogonale de ses vecteurs propres. Ainsi une matrice de blanchiment des observations,  $\mathbf{V}$ , peut être calculée par l'équation (2.4), où  $\mathbf{D}_n$  est la matrice diagonale des  $n$  plus grandes valeurs propres et  $\mathbf{E}_n$  est la matrice des  $n$  vecteurs propres correspondants.

$$\mathbf{V} = \mathbf{D}_n^{-\frac{1}{2}} \mathbf{E}_n^T \quad (2.4)$$

Les observations blanchies,  $\mathbf{Z}$ , sont obtenues par l'équation (2.5) :

$$\mathbf{Z} = \mathbf{V} \mathbf{X} \quad (2.5)$$

## Rotation

La deuxième étape de la méthode N-ICA consiste à chercher la matrice de rotation  $\mathbf{Q}$  qui rend les composantes de  $\mathbf{Z}$  les plus non-négatives possible. Ceci peut se faire en résolvant l'équation (2.6) [Plu03] :

$$\mathbf{Q} = \arg \min_{\mathbf{U}^T \mathbf{U} = \mathbf{I}_n} \mathcal{J}(\mathbf{U}) \quad (2.6)$$

où le critère  $\mathcal{J}(\mathbf{U})$  est donné par l'équation (2.7) :

$$\mathcal{J}(\mathbf{U}) = \frac{1}{2} \|\mathbf{Y}^-\|_F^2 \quad (2.7)$$

avec  $\mathbf{Y} = \mathbf{U} \mathbf{Z}$ ,  $\mathbf{Y}_{ij}^- = \min(\mathbf{Y}_{ij}, 0)$  et  $\|\cdot\|_F$  est la norme de Frobenius.

Les sources sont estimées par le filtrage spatial linéaire dont l'équation est donnée par (2.8) :

$$\hat{\mathbf{S}} = \mathbf{W}\mathbf{X} \text{ où } \mathbf{W} = \mathbf{Q}\mathbf{V} \quad (2.8)$$

Pour résoudre l'équation (2.6), nous pouvons consulter les algorithmes proposés dans [Plu03, YL05, Plu05] ainsi que dans [OSJ10, OJSJ11].

### Commentaires sur les hypothèses de la méthode N-ICA

Bien que supposant l'indépendance mutuelle des sources, nous remarquons que dans sa mise en œuvre, la méthode N-ICA se contente de décorreler les observations (statistiques de second ordre). Dans ces conditions, nous pouvons nous interroger sur la nécessité d'avoir l'indépendance mutuelle des sources, en d'autres termes, est-il toujours possible de séparer par la méthode N-ICA, des sources *non-négatives*, *décorrélées* et *well-grounded* ?

Pour tenter de répondre à cette question, nous considérons les trois sources telles que définies ci-dessous et dont la distribution jointe est uniformément répartie sur la sphère  $\mathcal{C}(\mathbf{c}, 1)$ , de centre  $\mathbf{c} = [1 \ 1 \ 1]^T$  et de rayon 1.

$$s_1 = 1 + \cos(\varphi) \quad (2.9)$$

$$s_2 = 1 + \sin(\theta) \sin(\varphi) \quad (2.10)$$

$$s_3 = 1 + \cos(\theta) \sin(\varphi) \quad (2.11)$$

$\theta$  est une variable aléatoire uniformément répartie sur  $[0, 2\pi]$ ,  $v$  est une variable aléatoire uniformément répartie sur  $[0, 1]$  et  $\varphi = \cos^{-1}(2v - 1)$ .

Il existe une infinité de bases dans lesquelles les trois sources ci-dessus sont *non-négatives*, *décorrélées* et *well-grounded*, leur mélange linéaire instantané n'est donc pas séparable par N-ICA (voir la preuve dans l'annexe A).

Ce résultat nous permet d'affirmer que la méthode N-ICA ne peut pas toujours séparer des sources non-négatives, uniquement décorréliées et *well-grounded*.

Il est donc nécessaire pour les sources non-négatives d'avoir une propriété qui soit à la fois plus forte que la décorrélation et le caractère *well-grounded* des lois marginales, mais plus faible que l'indépendance mutuelle. Nous conjecturons que pour des sources non-négatives soient séparables par N-ICA, il suffit qu'elles soient *décorrélées et well-grounded dans une unique base*.

### 2.2.2 Méthode de maximum de vraisemblance

La méthode de maximum de vraisemblance (MV) traite du problème de séparation de sources dans le cas déterminé ( $m = n$ ). Si le mélange est sur-déterminé ( $m > n$ ), alors on peut se ramener au cas déterminé en effectuant une première étape de blanchiment et de réduction de dimension comme décrite dans l'étape de blanchiment et réduction de dimensions de la méthode N-ICA. Dans l'approche de maximum de vraisemblance la matrice de mélange  $\mathbf{A}$  est un paramètre déterministe et les sources  $s_1, s_2, \dots, s_n$ , sont des variables aléatoires indépendantes, distribuées respectivement suivant les lois  $Pr_1(s_1), Pr_2(s_2), \dots, Pr_n(s_n)$ .

La matrice de mélange est estimée en résolvant l'équation (2.12) où  $\Pr(\mathbf{X}|\mathbf{A})$  est la vraisemblance des observations :

$$\hat{\mathbf{A}} = \arg \max_{\mathbf{A}} \Pr(\mathbf{X}|\mathbf{A}) \quad (2.12)$$

Dans le cas sans bruit, la vraisemblance des observations est donnée par l'équation (2.13) [PG97].

$$\Pr(\mathbf{X}|\mathbf{A}) = \frac{1}{|\det(\mathbf{A})|^K} \prod_{k=1}^K \prod_{i=1}^n \Pr_i([\mathbf{A}^{-1}\mathbf{x}(k)]_i) \quad (2.13)$$

En prenant l'opposé du logarithme de la vraisemblance, l'estimation de la matrice de mélange revient à résoudre l'équation (2.14) :

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \left[ K \log |\det(\mathbf{A})| - \sum_{k=1}^K \sum_{i=1}^n \log \Pr_i([\mathbf{A}^{-1}\mathbf{x}(k)]_i) \right] \quad (2.14)$$

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \left[ K \log |\det(\mathbf{A})| - \sum_{k=1}^K \sum_{i=1}^n \log \Pr_i(s_i(k)) \right] \quad (2.15)$$

En présence d'un bruit additif indépendant des sources, la vraisemblance d'un vecteur d'observations est donnée par l'équation (2.16) :

$$\Pr(\mathbf{x}(k)|\mathbf{A}, \mathbf{s}(k)) = \prod_{i=1}^n \Pr_{b_i}(b_i(k)) \quad (2.16)$$

où  $\Pr_{b_i}$  est la loi de probabilité du bruit sur le capteur  $i$ . On en déduit que :

$$\Pr(\mathbf{X}, \mathbf{S}|\mathbf{A}) = \Pr(\mathbf{B})\Pr(\mathbf{S}) = \prod_{k=1}^K \prod_{i=1}^n \Pr_{b_i}(b_i(k)) \Pr_i([\mathbf{A}^{-1}\mathbf{x}(k)]_i) \quad (2.17)$$

$$\Pr(\mathbf{X}, \mathbf{S}|\mathbf{A}) = \Pr(\mathbf{B})\Pr(\mathbf{S}) = \prod_{k=1}^K \prod_{i=1}^n \Pr_{b_i}(b_i(k)) \Pr_i(s_i) \quad (2.18)$$

La matrice de mélange est estimée en résolvant l'équation (2.19) :

$$\hat{\mathbf{A}} = \arg \max_{\mathbf{A}} \Pr(\mathbf{X}, \mathbf{S}|\mathbf{A}) \quad (2.19)$$

Dans la méthode de maximum de vraisemblance, l'hypothèse d'indépendance mutuelle des sources permet d'exprimer la vraisemblance des observations à l'aide du produit des lois marginales des sources, mais cette méthode n'optimise pas de critère d'indépendance. Il faut néanmoins fixer les densités de probabilités des sources et du bruit (équations (2.15) et (2.18)). Ceci nécessite une connaissance suffisamment précise des sources et du bruit pour choisir des distributions proches des distributions réelles. Dans le cadre de la séparation de sources non-négatives, nous pouvons choisir des densités de probabilité à supports positifs comme les lois Gamma, exponentielles, etc... Ce choix est discuté dans [Mou05].

### 2.2.3 Inférence bayésienne

Dans l'approche bayésienne, les sources, les coefficients de la matrice de mélange et le bruit sont des réalisations de variables aléatoires indépendantes. Suivant les connaissances a priori dont on dispose sur le problème, des densités de probabilité a priori  $Pr(\mathbf{S})$ ,  $Pr(\mathbf{A})$  et  $Pr(\mathbf{B})$  sont attribuées respectivement aux sources, à la matrice de mélange et au bruit. En utilisant la règle de *Bayes*, la loi a posteriori des données est exprimée par l'équation (2.20) :

$$Pr(\mathbf{A}, \mathbf{S}|\mathbf{X}) = \frac{Pr(\mathbf{X}|\mathbf{A}, \mathbf{S})Pr(\mathbf{A}, \mathbf{S})}{Pr(\mathbf{X})} \quad (2.20)$$

où  $Pr(\mathbf{X}|\mathbf{A}, \mathbf{S})$  est la vraisemblance des observations et  $Pr(\mathbf{X})$  est la densité de probabilité des observations. Par ailleurs  $Pr(\mathbf{X}|\mathbf{A}, \mathbf{S}) = Pr(\mathbf{B})$  et  $Pr(\mathbf{A}, \mathbf{S}) = Pr(\mathbf{A})Pr(\mathbf{S})$  (les sources et la matrice de mélange sont supposées statistiquement indépendantes).  $Pr(\mathbf{X})$  étant un terme de normalisation, on a ainsi :

$$Pr(\mathbf{A}, \mathbf{S}|\mathbf{X}) \propto Pr(\mathbf{B})Pr(\mathbf{A})Pr(\mathbf{S}) \quad (2.21)$$

Cette loi a posteriori permet de calculer des estimateurs des sources et de la matrice de mélange, soit directement à partir de la loi jointe, soit par marginalisation de cette dernière. Les hypothèses d'indépendance mutuelle des sources, des coefficients de la matrice de mélange et du bruit permettent d'exprimer la loi conjointe a posteriori  $Pr(\mathbf{A}, \mathbf{S}|\mathbf{X})$  à l'aide du produit des lois marginales a priori, cette méthode n'optimise pas non plus de critère d'indépendance. En pratique, il est difficile, voire impossible, de calculer une expression analytique de  $Pr(\mathbf{A}, \mathbf{S}|\mathbf{X})$ , les algorithmes bayésiens utilisent donc des procédures d'échantillonnage MCMC (Markov Chain Monte Carlo) pour générer des échantillons distribués suivant les lois a posteriori marginales.

L'approche bayésienne nécessite de fixer les lois a priori des sources, ainsi que celles des coefficients de la matrice mélange et du bruit. Pour garantir la non-négativité des sources et/ou des coefficients de la matrice de mélange, nous pouvons choisir des densités de probabilité à supports positifs (Gamma, exponentielle, Dirichlet, etc...). Contrairement à la méthode de maximum de vraisemblance, les choix des lois a priori peuvent être flexibles, les estimations des sources et de la matrice de mélange se faisant à partir des lois a posteriori.

Nous pouvons consulter [IM04, MBDC06, DMC<sup>+</sup>09] et [CJ10] au chapitre 12 pour plus de détails sur la séparation de sources non-négatives par inférence bayésienne.

## 2.3 Méthodes déterministes

Les méthodes déterministes ne modélisent pas de variables aléatoires, par conséquent, elles ne requièrent pas l'indépendance mutuelle des sources. Nous distinguons l'approche algébrique de factorisation en matrices non-négatives et les approches géométriques.

### 2.3.1 Factorisation en Matrices Non-négatives

Initialement introduite au milieu des années 90 par Paatero et Tapper [PP94] sous l'appellation PMF (pour désigner l'expression anglo-saxonne Positive Matrix Factorisation), la



factorisation en matrices non-négatives a été revisitée à la fin des années 90 par Lee et Seung [LS99] comme un moyen de décomposer une image sur une base de ses éléments élémentaires. Cette approche est désignée dans la littérature par le sigle NMF pour Nonnegative Matrix Factorization.

Sous les hypothèses de non-négativité des sources et des coefficients de la matrice de mélange ( $\mathbf{S} \geq 0$  et  $\mathbf{A} \geq 0$ ), la factorisation en matrices non-négatives cherche à décomposer la matrice non-négative des observations  $\mathbf{X}$  en produit de deux matrices non-négatives, l'une représentant les sources estimées  $\hat{\mathbf{S}}$  et l'autre la matrice de mélange estimée  $\hat{\mathbf{A}}$  telles que  $\hat{\mathbf{A}}\hat{\mathbf{S}} \approx \mathbf{X}$ . Les méthodes de NMF procèdent en deux étapes : la première étape consiste à définir une mesure de divergence (appelée aussi fonction de coût) entre les observations  $\mathbf{X}$  et le produit des matrices  $\mathbf{A}$  et  $\mathbf{S}$  (que l'on cherche). La deuxième étape consiste à chercher les matrices non-négatives  $\hat{\mathbf{A}}$  et  $\hat{\mathbf{S}}$  qui minimisent la mesure de divergence définie précédemment. Ces deux étapes sont résumées comme suit :

1. Définir une mesure de divergence  $\mathcal{D}$  entre  $\mathbf{X}$  et  $\mathbf{AS}$  :  $\mathcal{D}(\mathbf{X}, \mathbf{AS})$
2. Chercher les matrices optimales  $\hat{\mathbf{A}}$  et  $\hat{\mathbf{S}}$  à coefficients non-négatifs qui minimisent  $\mathcal{D}$

$$\left( \hat{\mathbf{A}}, \hat{\mathbf{S}} \right) = \arg \min_{\mathbf{A} \geq 0, \mathbf{S} \geq 0} \mathcal{D}(\mathbf{X}, \mathbf{AS}) \quad (2.22)$$

$\mathcal{D}$  est une application de  $\mathbb{R}_+^{m \times K} \times \mathbb{R}_+^{m \times K}$  dans  $\mathbb{R}_+$ , vérifiant la condition suivante :

$$\forall \mathbf{Y} \in \mathbb{R}_+^{m \times K}, \forall \mathbf{Z} \in \mathbb{R}_+^{m \times K}, \mathcal{D}(\mathbf{Y}, \mathbf{Z}) = 0 \iff \mathbf{Y} = \mathbf{Z} \quad (2.23)$$

Une première mesure de divergence naturelle qui peut être utilisée pour la NMF, est la *distance euclidienne* définie par l'équation (2.24) où  $\|\cdot\|_F$  est la norme de *Frobenius* :

$$\mathcal{D}_{EUC}(\mathbf{X}, \mathbf{AS}) = \|\mathbf{X} - \mathbf{AS}\|_F^2 \quad (2.24)$$

De nombreux algorithmes sont proposés dans la littérature pour optimiser cette fonction de coût, notamment la méthode de mise à jour multiplicative [LS01], la méthode de moindre carré alterné (ALS pour Alternating Least Square) avec contrainte de non-négativité [BJ97, ACD<sup>+</sup>06, CZ07], la méthode du gradient projeté [Lin07, ZC08] et la méthode Quasi Newton [ZC06]. Nous donnons dans l'annexe B d'autres mesures de divergence couramment utilisées en NMF.

Par ailleurs, les seules contraintes de non-négativité des sources et de la matrice de mélange ne suffisent pas à garantir l'unicité de la factorisation [DS04, MBI05, LCP<sup>+</sup>08]. En effet, si  $(\mathbf{W} \geq 0, \mathbf{Y} \geq 0)$  est une solution de la NMF alors  $(\mathbf{W}\mathbf{M}^{-1}, \mathbf{M}\mathbf{Y})$  est aussi une solution pour toute matrice inversible  $\mathbf{M}$  telle que  $\mathbf{W}\mathbf{M}^{-1} \geq 0$  et  $\mathbf{M}\mathbf{Y} \geq 0$ , dans ce cas  $\mathbf{M}$  est dite une indétermination de la NMF. Les indéterminations triviales de la NMF sont les matrices monomiales, en effet si  $\mathbf{M}$  est une matrice monomiale ( $\mathbf{M} = \mathbf{D}\mathbf{P}$  où  $\mathbf{D}$  est une matrice diagonale à coefficients positifs et  $\mathbf{P}$  est une matrice de permutation) alors  $\mathbf{M} \geq 0$  et  $\mathbf{M}^{-1} \geq 0$  [PC72]. Ceci correspond aux indéterminations canoniques de permutation et de facteur d'échelle intrinsèques à la séparation de sources. D'autres types d'indéterminations sont présentées dans [MBI05] et [Ber09].

Pour réduire le nombre de solutions admissibles, certaines méthodes NMF proposent de rajouter d'autres contraintes aux sources et/ou à la matrice de mélange. Ceci peut se matérialiser par le rajout de termes de régularisation à la mesure de divergence initiale, afin de prendre en compte des connaissances a priori sur le problème. L'équation à résoudre devient (2.25) :

$$\left(\hat{\mathbf{A}}, \hat{\mathbf{S}}\right) = \arg \min_{\mathbf{A} \geq 0, \mathbf{S} \geq 0} [\mathcal{D}(\mathbf{X}, \mathbf{A}\mathbf{S}) + \lambda_S \mathcal{R}(\mathbf{S}) + \lambda_A \mathcal{R}(\mathbf{A})] \quad (2.25)$$

où  $\mathcal{R}(\mathbf{S})$  et  $\mathcal{R}(\mathbf{A})$  sont des termes de régularisation des sources et de la matrice de mélange,  $\lambda_S$  et  $\lambda_A$  sont des coefficients positifs qui expriment les poids que l'on accorde aux termes de régularisation. Ainsi, en fonction de la physique du problème, on peut incorporer dans la NMF :

- une contrainte de *parcimonie* des sources en prenant  $\mathcal{R}(\mathbf{S})$  définie par l'équation (2.26) [Hoy02]. Cette NMF contrainte est aussi appelée Non-Negative Sparse Coding.

$$\mathcal{R}(\mathbf{S}) = \sum_{i=1}^n \sum_{k=1}^K s_i(k) \quad (2.26)$$

- une contrainte d'*orthogonalité* des sources, en prenant  $\mathcal{R}(\mathbf{S})$  définie par l'équation (2.27) [DLPP06], :

$$\mathcal{R}(\mathbf{S}) = Tr [\Lambda (\mathbf{S}\mathbf{S}^T - \mathbf{I}_n)] \quad (2.27)$$

où  $\Lambda$  est une matrice symétrique contenant les multiplicateurs de Lagrange

- une contrainte de *déterminant* sur la matrice de mélange en prenant  $\mathcal{R}(\mathbf{A})$  définie par l'équation (2.28) [SPL11], :

$$\mathcal{R}(\mathbf{A}) = \det(\mathbf{A}\mathbf{A}^T) \quad (2.28)$$

D'autres a priori comme l'harmonicité, la décorrélation ou encore la continuité des sources peuvent aussi être prises en compte. Nous pouvons consulter pour cela [CZPA09] et [CJ10] au chapitre 13.

## 2.3.2 Approches géométriques

Les approches géométriques constituent un autre point de vue sur le problème de séparation de sources non-négatives. Dans ces approches, chaque colonne de la matrice des observations  $\mathbf{X}$  est représentée par un point dans l'espace. L'ensemble des observations forment ainsi un nuage de points dans l'espace. Ces méthodes se fondent sur les propriétés géométriques de ce nuage de points pour concevoir les algorithmes de séparation. Il s'agit de méthodes directes qui cherchent d'abord à estimer la matrice de mélange.

Nous distinguons trois approches géométriques : celle fondée sur la classification des observations (Analyse en Composantes Parcimonieuses), celle fondée la détermination des points extrêmes du nuage des observations, et enfin celle fondée sur l'estimation du simplexe de volume minimal contenant le nuage des observations.

### 2.3.2.1 Analyse en composantes parcimonieuses

L'analyse en composantes parcimonieuses permet d'estimer la matrice de mélange dans le cas de sources parcimonieuses (pas forcément non-négatives). La figure 2.1 illustre les diagrammes de dispersion des observations dans le cas de deux mélanges de deux sources uniformes (2.1.a) et de deux mélanges de deux sources parcimonieuses (2.1.b).

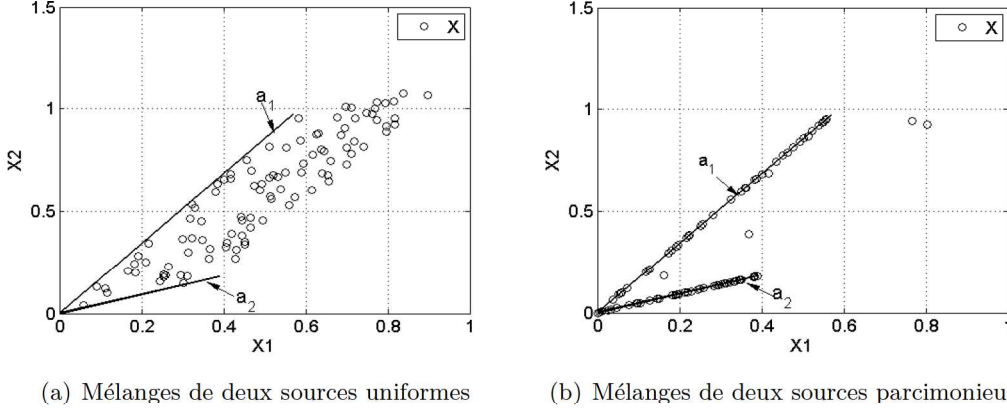


FIGURE 2.1 – Diagrammes de dispersion de mélanges de deux sources uniformes et de mélanges de deux sources parcimonieuses

Dans le cas des sources très fortement parcimonieuses, les observations sont principalement distribuées le long des axes formés par les colonnes de la matrice de mélange. On peut ainsi estimer la matrice de mélange en effectuant une classification des observations et en ajustant une droite à chaque classe. Cette idée est exploitée dans [BZMJM04, GS06, GTCB07, NMBZJ08]. L'analyse en composantes parcimonieuses permet de traiter le cas des mélanges sous-déterminés ( $m < n$ ) [ZP01, LCiA03].

### 2.3.2.2 Méthode des points extrêmes du nuage des observations

L'approche d'estimation de la matrice de mélange par détermination des points extrêmes du diagramme de dispersion des observations est fondée sur l'hypothèse fondamentale de la *dominance locale des sources*. Cette hypothèse signifie que pour chaque source, il existe au moins un indice d'observation pour lequel elle est seule active, c'est-à-dire non nulle (toutes les autres étant inactives, c'est-à-dire nulles) et elle se traduit par l'équation (2.29) :

$$\forall 1 \leq i \leq n, \exists k_i (1 \leq k_i \leq K) \text{ tel que } s_i(k_i) \neq 0 \text{ et } s_j(k_i) = 0 \text{ pour } j \neq i \quad (2.29)$$

Ainsi, la colonne  $k_i$  de la matrice des observations est proportionnelle à la  $i$ -ième colonne de la matrice de mélange. En effet l'observation  $x_j(k_i)$  est donnée par :

$$x_j(k_i) = \sum_{l=1}^n a_{jl} s_l(k_i) = a_{ji} s_i(k_i) \quad (2.30)$$

$$\mathbf{x}(k_i) = [x_1(k_i) \ x_2(k_i) \ \cdots \ x_m(k_i)]^T = [s_i(k_i) a_{1i} \ s_i(k_i) a_{2i} \ \cdots \ s_i(k_i) a_{mi}]^T = s_i(k_i) \mathbf{a}_i$$

où  $\mathbf{a}_i$  est la  $i$ -ième colonne de la matrice de mélange  $\mathbf{A}$ .

On peut donc ainsi estimer la matrice de mélange (à une matrice monomiale près) en déterminant les colonnes correspondant à la dominance locale dans les observations. Si l'on projette le nuage des observations sur un plan, alors les points de dominance locale correspondent aux points extrêmes du nouveau diagramme de dispersion comme cela est illustré sur la figure 2.2. Les données projetées sont obtenues par l'équation (2.31), et se répartissent sur le plan  $(\mathcal{P})_n$  défini par l'équation (2.32).

$$\tilde{x}_i(k) = x_i(k) / \sum_{l=1}^3 x_l(k) \quad (2.31)$$

$$(\mathcal{P})_n = \left\{ \mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n \mid \sum_{i=1}^n x_i = 1 \right\} \quad (2.32)$$

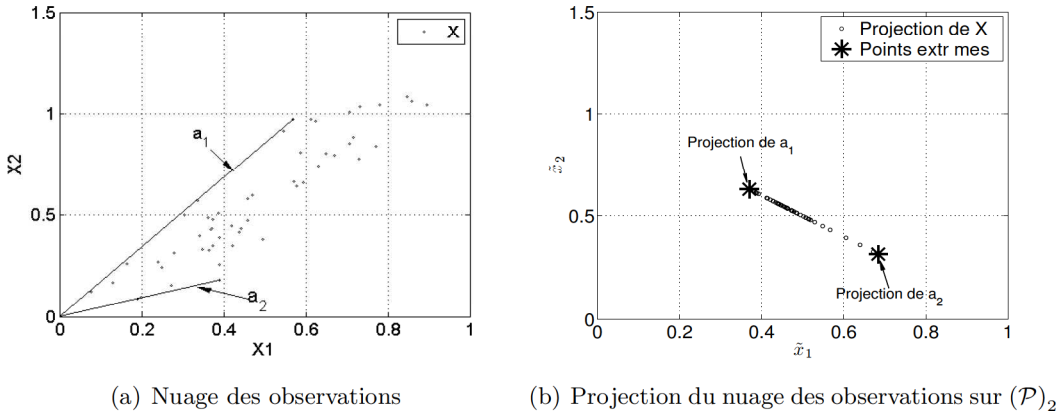


FIGURE 2.2 – Diagrammes de dispersion de mélanges de sources à dominance locale

Les points de dominance locale peuvent ainsi être identifiés en déterminant les sommets de l'enveloppe convexe du diagramme de dispersion des observations (projetées sur le plan). Des algorithmes exploitant cette approche sont mis en œuvre dans [IC99, NBD05, ZCHW08, CMCW08, LNN10].

On peut noter que si les sources sont très parcimonieuses, alors elles vérifient la contrainte de dominance locale. Le problème d'analyse en composantes parcimonieuses de sources non-négatives peut donc se résoudre par la méthode des points extrêmes.

### 2.3.2.3 Méthode du Simplexe de Volume Minimal

Cette approche a été introduite à la fin des années 90 par Craig pour la résolution d'un problème de télédétection par imagerie multi-spectrale [Cra94]. Cette méthode opère dans le cas de mélange déterminé ( $m = n$ ). Si le mélange est sur-déterminé ( $m > n$ ), une première étape de réduction de dimensions doit être effectuée (comme présentée dans la méthode N-ICA) afin de se ramener au cas déterminé. La méthode du simplexe de volume minimal se

fonde sur l'hypothèse (2.33) :

$$\forall 1 \leq k \leq K, \sum_{i=1}^n s_i(k) \leq 1 \quad (2.33)$$

Géométriquement, l'hypothèse (2.33) implique que le diagramme de dispersion des sources est inclus dans un simplexe. Rappelons qu'un simplexe est l'extension en dimension quelconque d'un triangle. Pour  $n \geq 2$ , un  $n$ -simplexe a  $n + 1$  sommets,  $\frac{n(n+1)}{2}$  arêtes et  $C_{n+1}^3$  faces. Comme l'illustre la figure 2.3, la transformation linéaire d'un simplexe est aussi un simplexe. Il s'ensuit que le nuage des observations est inclus dans le  $n$ -simplexe dont les sommets sont les  $n$  points correspondant aux colonnes de la matrice de mélange et l'origine de l'espace.

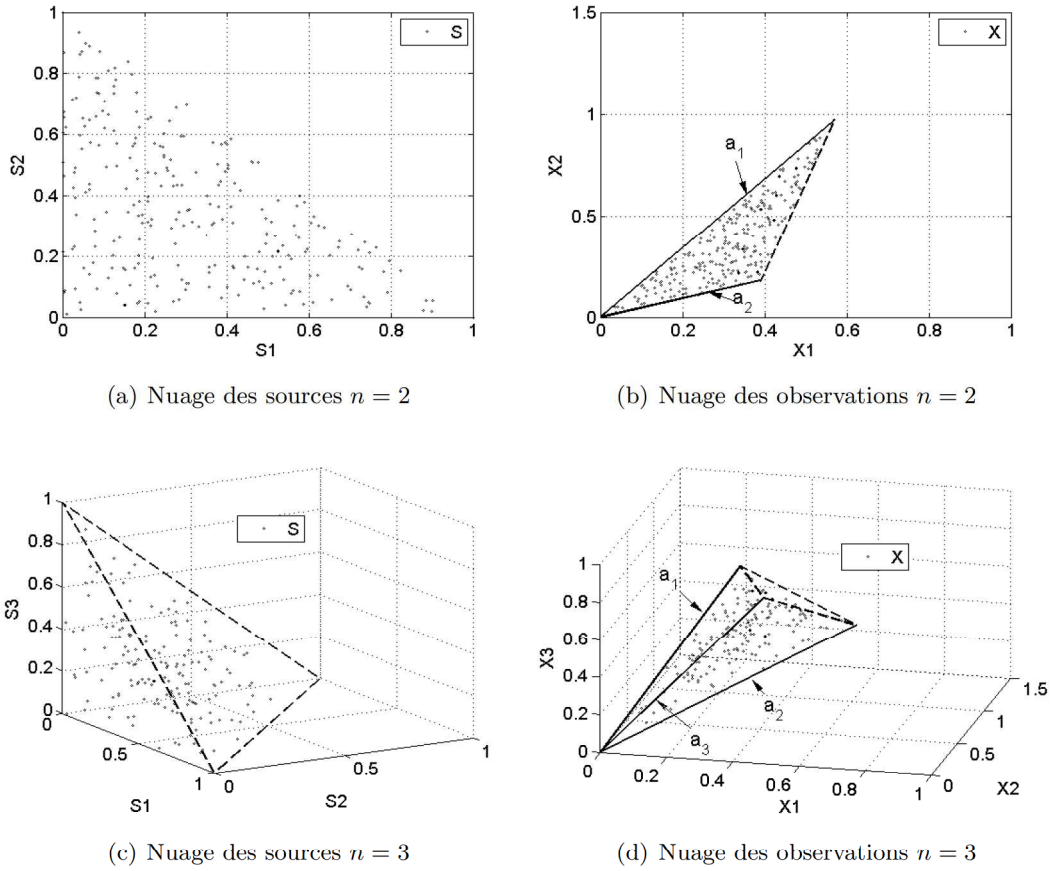


FIGURE 2.3 – Nuage de points des signaux sources et observations, dans le cas de sources incluses dans un simplexe

L'idée de cette approche est donc d'estimer la matrice de mélange en cherchant le simplexe de volume minimal contenant le diagramme de dispersion des observations. Le volume,  $\mathcal{V}(\mathbf{A})$ , d'un simplexe généré par une matrice  $\mathbf{A}$  est mesuré par la formule (2.34) [Cra94] :

$$\mathcal{V}(\mathbf{A}) = \frac{|\det(\mathbf{A})|}{n!} \quad (2.34)$$

L'estimation de la matrice de mélange est effectuée en résolvant l'équation (2.35) :

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}^{-1}\mathbf{X} \geq 0, \mathbf{1}_n^T \mathbf{A}^{-1}\mathbf{X} \leq \mathbf{1}_n^T} \mathcal{V}(\mathbf{A}) \quad (2.35)$$

Des algorithmes utilisant l'idée de simplexe de volume minimal ont aussi été développés pour la séparation d'images hyperspectrales [CWLO06, LBD08, CCMM09, BD09]. En imagerie hyperspectrale, l'hypothèse (2.33) est remplacée par l'hypothèse d'*additivité totale des sources* qui est donnée par l'équation (2.36) :

$$\forall 1 \leq k \leq K, \sum_{i=1}^n s_i(k) = 1 \quad (2.36)$$

Dans ce cas, les observations sont situées sur une face du simplexe engendré par la matrice de mélange comme l'illustre la figure 2.4. L'estimation de la matrice de mélange est toujours effectuée en résolvant l'équation (2.35), la contrainte  $\mathbf{1}_n^T \mathbf{A}^{-1} \mathbf{X} \leq \mathbf{1}_n^T$  étant remplacée par  $\mathbf{1}_n^T \mathbf{A}^{-1} \mathbf{X} = \mathbf{1}_n^T$ .

On peut lever les contraintes (2.33) ou (2.36) en projetant les observations sur le plan  $(\mathcal{P})_n$  [Cha09]. En effet, supposons que les sources ne satisfassent pas l'hypothèse d'additivité totale, alors :  $\exists 1 \leq k \leq K$  tel que  $\sum_{i=1}^n s_i(k) \neq 1$ . En absence de bruit, l'équation de mélange s'écrit :

$$x_j(k) = \sum_{l=1}^n a_{jl} s_l(k) \quad (2.37)$$

soit vectoriellement :

$$\mathbf{x}(k) = \mathbf{A} \mathbf{s}(k) \quad (2.38)$$

Supposons que  $\forall 1 \leq k \leq K, \sum_{i=1}^n x_i(k) \neq 0$  et  $\forall 1 \leq t \leq n, \sum_{j=1}^n a_{jt} \neq 0$ , et posons pour tout indice d'échantillon  $k$  ( $1 \leq k \leq K$ ) :

$$\tilde{\mathbf{x}}(k) = \frac{\mathbf{x}(k)}{\sum_{i=1}^n x_i(k)} \quad (2.39)$$

alors, pour tout indice d'observation  $j$  fixé :

$$\tilde{x}_j(k) = \frac{x_j(k)}{\sum_{i=1}^n x_i(k)} = \frac{\sum_{l=1}^n a_{jl} s_l(k)}{\sum_{i=1}^n x_i(k)} = \frac{\sum_{l=1}^n a_{jl} s_l(k)}{\sum_{i=1}^n x_i(k)} \frac{\sum_{j=1}^n a_{jt}}{\sum_{j=1}^n a_{jt}} = \sum_{l=1}^n \frac{a_{jl}}{\sum_{j=1}^n a_{jt}(k)} \frac{s_l(k) \sum_{j=1}^n a_{jt}}{\sum_{i=1}^n x_i(k)} \quad (2.40)$$

Si l'on pose :

$$\tilde{a}_{jl} = \frac{a_{jl}}{\sum_{j=1}^n a_{jt}} \text{ et } \tilde{s}_l(k) = s_l(k) \frac{\sum_{j=1}^n a_{jt}}{\sum_{i=1}^n x_i(k)} \quad (2.41)$$

alors

$$\tilde{\mathbf{x}}(k) = \tilde{\mathbf{A}} \tilde{\mathbf{s}}(k) \quad (2.42)$$

où  $\tilde{\mathbf{A}} = [\tilde{a}_{jl}]_{1 \leq j \leq n, 1 \leq l \leq n}$ . Nous avons alors,  $\forall 1 \leq k \leq K$  :

$$\sum_{l=1}^n \tilde{s}_l(k) = \sum_{l=1}^n s_l(k) \frac{\sum_{j=1}^n a_{jt}}{\sum_{i=1}^n x_i(k)} = \frac{\sum_{j=1}^n \sum_{l=1}^n a_{jl} s_l(k)}{\sum_{i=1}^n x_i(k)} = \frac{\sum_{j=1}^n x_j(k)}{\sum_{i=1}^n x_i(k)} = 1 \quad (2.43)$$

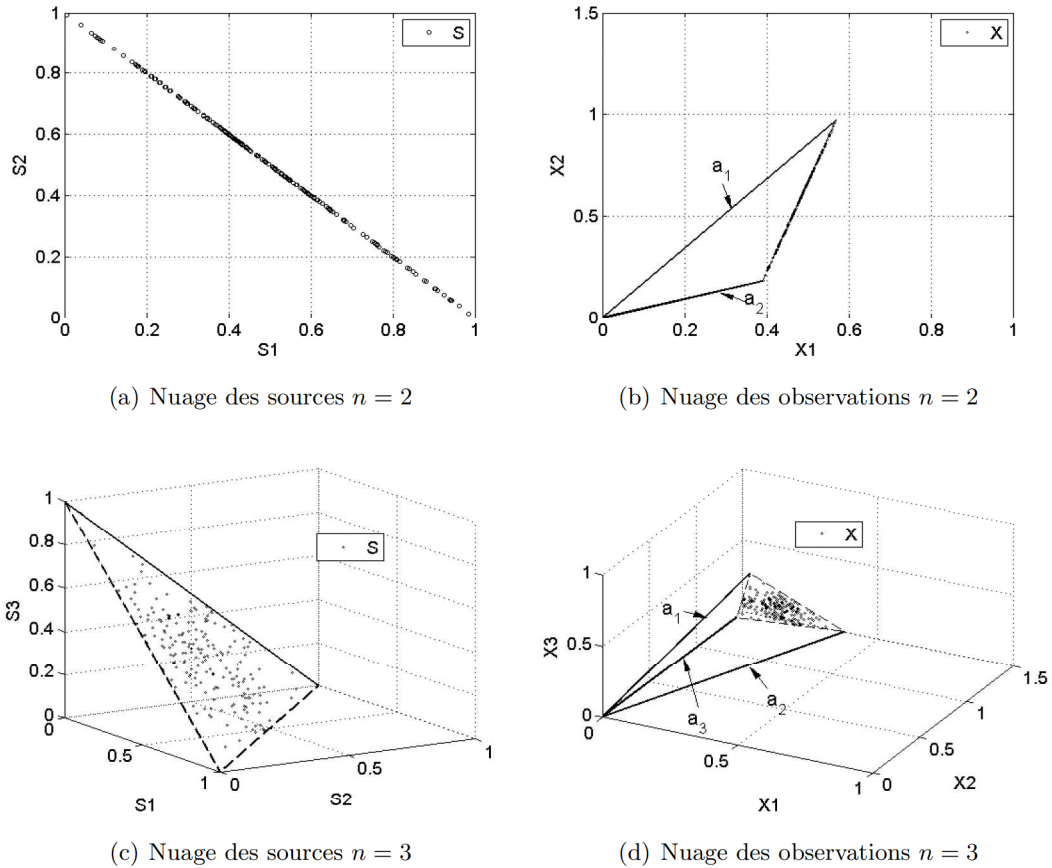


FIGURE 2.4 – Nuage de points des signaux sources et observations, dans le cas de sources vérifiant la contrainte d’additivité totale

## 2.4 Synthèse des méthodes de SAS non-négatives

Nous récapitulons dans le tableau de la figure 2.5 les différentes méthodes de séparation de sources non-négatives en précisant leurs hypothèses.

Au vu de la figure 2.5, il ressort que, sans prise en compte de connaissances a priori sur les sources et sur les coefficients de mélange, la factorisation en matrices non-négatives est la méthode qui requiert le moins d’hypothèses sur le problème de séparation de sources non-négatives. La méthode bayésienne est celle qui demande le plus d’information (assez flexibles cependant) sur les sources et sur la matrice de mélange. Nous pouvons noter aussi que toutes les méthodes supposent que la matrice de mélange est de rang plein, hypothèse est nécessaire pour assurer la séparabilité des sources.

| Approches statistiques                   |  | Approches déterministes  |   |  |   |
|--|--|--|---|--|---|
| Hypothèses sur les sources non-négatives | <ul style="list-style-type: none"> <li>▪ S aléatoire</li> <li>▪ A déterministe</li> </ul>  | <ul style="list-style-type: none"> <li>▪ S aléatoire</li> <li>▪ A aléatoire</li> </ul>   | <ul style="list-style-type: none"> <li>▪ S déterministe</li> <li>▪ A déterministe</li> </ul>  |  |   |
|  |  | <p>ACI Non-négatif</p> <p>Maximum de vraisemblance</p>   | <p>Méthode bayésienne</p>   | <p>Factorisation en matrices non-négatives sans contrainte</p> <p>Méthodes géométriques</p>                    |   |
|  |  |  |   | <p>Analyse en composantes parcimonieuses</p> <p>Points extrêmes</p> <p>Simplexe de volume minimal</p>          |   |
| Hypothèses sur la matrice de mélange     | <ul style="list-style-type: none"> <li>• Indépendance mutuelle</li> <li>• Well-grounded</li> </ul>                                 | <ul style="list-style-type: none"> <li>• Indépendance mutuelle</li> <li>• Densités de probabilités connues</li> </ul>              | <ul style="list-style-type: none"> <li>• Indépendance mutuelle et indépendantes des sources</li> <li>• Densités de probabilités a priori connues</li> </ul>   | <p>Parcimonie</p> <p>Dominance locale</p>  | <p>Additivité totale</p>  |
|  | <ul style="list-style-type: none"> <li>• Mélange déterminé ou sur-déterminé</li> <li>• Matrice de mélange de rang plein</li> </ul> | <ul style="list-style-type: none"> <li>• Mélange déterminé ou sur-déterminé</li> <li>• Matrice de mélange de rang plein</li> </ul> | <ul style="list-style-type: none"> <li>• Indépendance mutuelle et indépendantes des sources</li> <li>• Densités de probabilités a priori connues</li> <li>• Matrice de mélange de rang plein</li> </ul> | <ul style="list-style-type: none"> <li>• Non-négativité</li> <li>• Matrice de mélange de rang plein</li> </ul> | <p>Matrice de mélange de rang plein</p> <p>Matrice de mélange de rang plein</p> |

FIGURE 2.5 – Synthèse des méthodes de séparations de sources non-négatives



## 2.5 Premiers résultats de séparation et premières conclusions

L'objectif de ces premières simulations est de tester les performances des méthodes présentées précédemment, dans différentes situations de séparation de sources non-négatives. Nous évaluerons et comparerons, la capacité de ces algorithmes à séparer des sources indépendantes et des sources corrélées, sans et avec bruit, sans et avec dominance locale. Nous utiliserons pour cela les algorithmes indiqués dans le tableau 2.1 :

| Acronyme | Type de méthode  | Référence             |
|----------|--|-----------------------|
| N-ICA    | Analyse en composantes indépendantes non-négatives   | [Plu04]               |
| BI-NSS   | Bayésienne   | [DMC <sup>+</sup> 09] |
| multNMF  | NMF utilisant la distance euclidienne  | [LS01]                |
| detNMF   | NMF utilisant la distance euclidienne avec contrainte de déterminant minimal sur la matrice de mélange estimée | [SPL11]               |
| VCA      | Géométrie : points extrêmes du nuage   | [NBD05]               |
| SISAL    | Géométrie : simplexe de volume minimal   | [BD09]                |
| nLCA-IVM | Géométrie : simplexe de volume minimal   | [WCCW10]              |

TABLE 2.1 – Listes des algorithmes utilisés pour les simulations

L'algorithme BI-NSS [DMC<sup>+</sup>09] a été initialement proposé pour la séparation d'images hyperspectrale. Les auteurs ont cependant indiqué qu'il peut être utilisé à chaque fois que l'on cherche à séparer des sources non-négatives vérifiant l'hypothèse d'additivité totale, et nous avons montré dans le paragraphe 2.3.2.3 que l'on peut lever la contrainte d'additivité totale en normalisant les observations. Dans les simulations ci-dessous, la normalisation des observations a été effectuée pour les méthodes supposant l'additivité totale des sources.

### 2.5.1 Critères de performances

Nous utilisons quatre indices de performances, définis comme suit :

- *l'erreur de séparation*,  $E_{sep}$ , définie par l'équation (2.44) [MM94, ACY96], où  $\Delta = \hat{\mathbf{A}}^\dagger \mathbf{A}$  et  $\hat{\mathbf{A}}^\dagger$  est la pseudo-inverse de la matrice de mélange estimée  $\hat{\mathbf{A}}$ . Il s'agit d'un critère non aveugle qui mesure l'écart de la matrice  $\Delta$  à une matrice monomiale.  $E_{sep}$  s'apparente au pourcentage des autres sources présent dans l'estimation d'une source donnée.

$$E_{sep} = \frac{1}{2n(n-1)} \left[ \sum_{i=1}^n \left( \sum_{j=1}^n \frac{|\Delta_{ij}|}{\max_{1 \leq l \leq n} |\Delta_{il}|} - 1 \right) + \sum_{j=1}^n \left( \sum_{i=1}^n \frac{|\Delta_{ji}|}{\max_{1 \leq l \leq n} |\Delta_{li}|} - 1 \right) \right] \quad (2.44)$$

- *l'erreur de corrélation*,  $E_{corr}$ , définie par l'équation (2.45).  $E_{corr}$  est aussi un critère non aveugle, inversement proportionnelle à la corrélation entre les sources estimées et les sources réelles.

$$E_{corr} = -\log_{10} \left[ \frac{1}{n} \sum_{i=1}^n \frac{\hat{\mathbf{s}}_i \mathbf{s}_i^T}{\|\hat{\mathbf{s}}_i\|_2 \|\mathbf{s}_i\|_2} \right] \quad (2.45)$$

- *la négativité* des sources estimées,  $E_{neg}$ , définie par l'équation (2.46), il s'agit d'un critère aveugle quantifiant les valeurs négatives qui apparaissent dans les estimations des sources non-négatives. Avant de calculer cet indice, chaque source estimée est d'abord normalisée par son maximum.

$$E_{neg} = \left\| \hat{\mathbf{S}}^- \right\|_F^2 = \sum_{i=1}^n \sum_{k=1}^K [\hat{\mathbf{s}}_i(k)^-]^2 \quad (2.46)$$

où  $\hat{\mathbf{s}}_i(k)^- = \max [0, \hat{\mathbf{s}}_i(k)]$

- *le coût de calcul*,  $T$ , en terme de temps CPU nécessaire à la convergence pour chaque algorithme. Ce critère peut être mesuré aussi en aveugle.

Dans les simulations, les valeurs des critères  $E_{sep}$  et  $E_{corr}$  sont exprimées en décibels (dB), le critère  $E_{neg}$  est sans unité et le coût de calcul  $T$  est exprimé en secondes (s).

### 2.5.2 Génération de la matrice de mélange et du bruit

Dans cette première série de simulations, les coefficients des matrices de mélange sont générées aléatoirement suivant une distribution uniforme entre 0 et 1.

Les échantillons de bruit sont tirés aléatoirement suivant une distribution gaussienne de moyenne nulle et de variance unité tronquée sur  $\mathbb{R}_+$  afin de conserver la positivité du mélange bruité. Pour une observation  $i$  donnée, le niveau de bruit ou rapport signal sur bruit de l'observation (en abrégé RSB), est donné par l'équation (2.47) :

$$RSB_i = 10 \log_{10} \left( \sum_k \left[ \sum_{j=1}^n a_{ij} s_j(k) \right]^2 / \sum_{k=1}^K [b_i(k)]^2 \right) \quad (2.47)$$

Dans les simulations, nous fixons le même rapport signal sur bruit pour toutes les observations.

### 2.5.3 Résultats de séparation de sources indépendantes

Chaque source est générée suivant une *distribution uniforme parcimonieuse*. Cela signifie que les échantillons de chaque source ont un certain taux de parcimonie  $\tau$  et les échantillons non nuls sont tirés suivant une distribution uniforme entre 0 et 1.  $\tau$  est un nombre réel compris entre 0 et 1 qui indique la proportion d'échantillons non nuls de la source considérée. Un taux de parcimonie  $\tau = 0$  correspond à une source dont tous les échantillons sont nuls et un taux de parcimonie  $\tau = 1$  correspond à une source dont tous les échantillons sont non nuls.

Dans les simulations ci-dessous, nous fixons le nombre d'observations à  $m = 20$ , le nombre de sources à  $n = 4$  et le nombre d'échantillons  $K = 5000$

#### Sources parcimonieuses

Dans cette première simulation, nous fixons le taux de parcimonie  $\tau = 0.01$  pour toutes les sources. Ce taux donne 1% d'échantillons non nuls pour chaque source et correspond au taux de parcimonie des spectres de masse utilisés au chapitre 5. Nous pouvons remarquer que les sources vérifient ici l'hypothèse de dominance locale.

Les tableaux 2.2, 2.3, 2.4 et 2.5 montrent les indices moyens de performance des différentes méthodes pour différents niveaux de bruit, calculés à partir de 20 exécutions indépendantes Monte Carlo.

|              | N-ICA  | BI-NSS | multNMF | detNMF | VCA     | SISAL  | nLCA-IVM |
|--------------|--------|--------|---------|--------|---------|--------|----------|
| Sans bruit   | -23.08 | -24.2  | -9.66   | -10.25 | -140.49 | -14.19 | -90.43   |
| $RSB = 50dB$ | -19.99 | -7.43  | -15.11  | -6.99  | -2.95   | -2.31  | -3.85    |
| $RSB = 10dB$ | -18.76 | -3.36  | -10.17  | -9.72  | -8.51   | -2.83  | -3.36    |

TABLE 2.2 – Erreur de séparation  $E_{sep}(dB)$ , pour des sources indépendantes et parcimonieuses

|              | N-ICA  | BI-NSS | multNMF | detNMF | VCA     | SISAL  | nLCA-IVM |
|--------------|--------|--------|---------|--------|---------|--------|----------|
| Sans bruit   | -44.75 | -49.96 | -17.64  | -18.59 | -143.92 | -29.21 | -142.13  |
| $RSB = 50dB$ | -39.34 | -6.32  | -27.89  | -11.28 | -7.46   | -7.53  | -7.98    |
| $RSB = 10dB$ | -22.63 | -2.21  | -18.89  | -17.93 | -10.16  | -7.54  | -7.87    |

TABLE 2.3 – Erreur de corrélation  $E_{corr}(dB)$ , pour des sources indépendantes et parcimonieuses

|              | N-ICA           | BI-NSS | multNMF | detNMF | VCA   | SISAL           | nLCA-IVM |
|--------------|-----------------|--------|---------|--------|-------|-----------------|----------|
| Sans bruit   | $7, 10.10^{-3}$ | 0      | 0       | 0      | 0     | 0               | 0        |
| $RSB = 50dB$ | $2, 97.10^{-2}$ | 0      | 0       | 0      | 17.63 | $3, 74.10^{-8}$ | 0        |
| $RSB = 10dB$ | $1, 35.10^{-1}$ | 0      | 0       | 0      | 28.43 | $5, 18.10^{-4}$ | 0        |

TABLE 2.4 – Négativité des sources estimées  $E_{neg}$ , pour des sources indépendantes et parcimonieuses

**Commentaires des résultats de séparation de sources parcimonieuses :** D'abord en terme d'erreur de séparation  $E_{sep}$  et d'erreur de corrélation  $E_{corr}$ , toutes les méthodes

|              | N-ICA | BI-NSS | multNMF | detNMF | VCA  | SISAL | nLCA-IVM |
|--------------|-------|--------|---------|--------|------|-------|----------|
| Sans bruit   | 3.32  | 246.17 | 2.92    | 0.58   | 0.03 | 0.33  | 1.12     |
| $RSB = 50dB$ | 0.55  | 794.96 | 3.06    | 10.78  | 0.04 | 1.17  | 2.03     |
| $RSB = 10dB$ | 0.07  | 809.05 | 3.14    | 11.26  | 0.04 | 1.11  | 1.62     |

TABLE 2.5 – Coût de calcul  $T(s)$ , pour des sources indépendantes et parcimonieuses

présentent des performances satisfaisantes en absence de bruit, les méthodes géométriques donnant les meilleurs résultats. Les performances de la méthode N-ICA sont affectées par le blanchiment “imparfait” des observations dû au faible nombre d’échantillons non nuls, on note cependant que cette méthode résiste mieux au bruit. Les performances des méthodes géométriques se dégradent considérablement en présence de bruit (même à très fort rapport signal sur bruit).

Ensuite, nous pouvons remarquer que la méthode N-ICA et certaines méthodes géométriques présentent des valeurs négatives dans les sources estimées  $E_{neg}$ , le poids des valeurs négatives variant en sens opposé au RSB.

Enfin la méthode bayésienne possède un coût de calcul très élevé comparativement aux autres méthodes sans améliorer (dans cette simulation) les indices  $E_{sep}$  et  $E_{corr}$ .

### Sources uniformes

Dans cette deuxième simulation, nous fixons le taux de parcimonie à  $\tau = 1$  pour toutes les sources, ce qui à des sources uniformes 0 et 1. Il faut noter qu’ici, l’hypothèse de dominance locale des sources n’est pas vérifiée.

Les tableaux 2.6, 2.7, 2.8 et 2.9 montrent les indices moyens de performance des différentes méthodes pour différents niveaux de bruit, calculés à partir de 20 exécutions indépendantes Monte Carlo.

les indices de performance des différentes méthodes pour différents niveaux de bruit.

|              | N-ICA  | BI-NSS | multNMF | detNMF | VCA    | SISAL  | nLCA-IVM |
|--------------|--------|--------|---------|--------|--------|--------|----------|
| Sans bruit   | -20.44 | -16.15 | -4.47   | -5.62  | -10.18 | -22.49 | -28.55   |
| $RSB = 50dB$ | -21.11 | -15.23 | -3.83   | -4.04  | -9.16  | -22.92 | -24.66   |
| $RSB = 10dB$ | -12.36 | -9,63  | -3.85   | -3.16  | -4.52  | -12.57 | -8.09    |

TABLE 2.6 – Erreur de séparation  $E_{sep}(dB)$ , pour des sources indépendantes et uniformes

|              | N-ICA  | BI-NSS | multNMF | detNMF | VCA    | SISAL  | nLCA-IVM |
|--------------|--------|--------|---------|--------|--------|--------|----------|
| Sans bruit   | -39.49 | -32.08 | -9.98   | -12.36 | -17.37 | -45.98 | -55.75   |
| $RSB = 50dB$ | -41.3  | -29.42 | -9.25   | -9.97  | -14.97 | -45.11 | -41.23   |
| $RSB = 10dB$ | -11.94 | -12,32 | -7.91   | -6.91  | -4.89  | -11.93 | -9.25    |

TABLE 2.7 – Erreur de corrélation  $E_{corr}(dB)$ , pour des sources indépendantes et uniformes

|              | N-ICA          | BI-NSS | multNMF | detNMF | VCA    | SISAL          | nLCA-IVM |
|--------------|----------------|--------|---------|--------|--------|----------------|----------|
| Sans bruit   | $1,68.10^{-4}$ | 0      | 0       | 0      | 57     | $2,8.10^{-3}$  | 0        |
| $RSB = 50dB$ | $3,42.10^{-4}$ | 0      | 0       | 0      | 455.21 | $2,7.10^{-3}$  | 0        |
| $RSB = 10dB$ | $2,77.10^{-1}$ | 0      | 0       | 0      | 407.72 | $3,05.10^{-1}$ | 0        |

TABLE 2.8 – Négativité des sources estimées  $E_{neg}$ , pour des sources indépendantes et uniformes

|              | N-ICA | BI-NSS | multNMF | detNMF | VCA  | SISAL | nLCA-IVM |
|--------------|-------|--------|---------|--------|------|-------|----------|
| Sans bruit   | 0.16  | 751.95 | 2.96    | 10.84  | 0.03 | 1.14  | 3.32     |
| $RSB = 50dB$ | 0.19  | 747.92 | 3.25    | 11.05  | 0.03 | 1.22  | 2.66     |
| $RSB = 10dB$ | 0.28  | 759,78 | 2.92    | 10.84  | 0.03 | 1.09  | 1.76     |

TABLE 2.9 – Coût de calcul  $T(s)$ , pour des sources indépendantes et uniformes

**Commentaires des résultats de séparation de sources uniformes :** Les tableaux 2.6 et 2.7 nous indiquent que la méthode d'analyse en composantes indépendantes non-négatives, la méthode bayésienne ainsi que les méthodes géométriques basées sur l'estimation du simplexe de volume minimal contenant le nuage des observations (SISAL et nLCA-IVM) présentent des indices  $E_{sep}$  et  $E_{corr}$  satisfaisants et résistent assez bien au bruit comparativement au cas parcimonieux. En terme de négativité des sources estimées, nous pouvons faire sur le tableau 2.8 les mêmes remarques que dans le cas de sources parcimonieuses.

De même que dans la première simulation, la méthode bayésienne présente un coût de calcul extrêmement élevé par rapport aux autres méthodes sans améliorer significativement (dans cette situation) les indices  $E_{sep}$  et  $E_{corr}$ .

## 2.5.4 Résultats de séparation de sources corrélées

Pour cette simulation, nous considérons les sources représentées sur la figure 2.6 [CZ06]. Chaque image de taille  $350 \times 275$  pixels, est vectorisée pour former une ligne de la matrice des sources. Le nombre total de pixels est donc  $K = 96250$ . Ces sources ne vérifient pas l'hypothèse de dominance locale.

La matrice de corrélation empirique est estimée par l'équation (2.48) où  $\mu_{s_l}$  et  $\sigma_{s_l}$  sont respectivement la moyenne et la variance de la source  $l$ .

$$\left[\hat{C}_{\mathbf{S}}\right]_{ij} = \frac{\frac{1}{K} \sum_{k=1}^K [s_i(k) - \mu_{s_i}] [s_j(k) - \mu_{s_j}]}{\sigma_{s_i} \sigma_{s_j}} \quad (2.48)$$

Ce qui donne numériquement :

$$\hat{C}_{\mathbf{S}} = \begin{pmatrix} 1.00 & 0.91 & 0.84 & 0.84 \\ 0.91 & 1.00 & 0.84 & 0.84 \\ 0.84 & 0.84 & 1.00 & 0.93 \\ 0.84 & 0.84 & 0.93 & 1.00 \end{pmatrix} \quad (2.49)$$

Les sources sont très fortement corrélées. Les tableaux 2.10, 2.11, 2.12, 2.13 montrent les

indices moyens de performance des différentes méthodes pour différents niveaux de bruit, calculés à partir de 20 exécutions indépendantes Monte Carlo.



FIGURE 2.6 – Sources corrélées : Faces humaines

|              | N-ICA | BI-NSS | multNMF | detNMF | VCA    | SISAL  | nLCA-IVM |
|--------------|-------|--------|---------|--------|--------|--------|----------|
| Sans bruit   | -3.16 | -20.74 | -4.39   | -4.79  | -24.25 | -26.95 | -29.17   |
| $RSB = 50dB$ | -3.27 | -21.68 | -3.31   | -4.21  | -7.55  | -10.58 | -11.55   |
| $RSB = 10dB$ | -3.47 | -8.48  | -4.17   | -4.91  | -3.35  | -2.78  | -4.17    |

TABLE 2.10 – Erreur de séparation  $E_{sep}(dB)$ , pour des sources corrélées

|              | N-ICA | BI-NSS | multNMF | detNMF | VCA    | SISAL  | nLCA-IVM |
|--------------|-------|--------|---------|--------|--------|--------|----------|
| Sans bruit   | -8.19 | -46.31 | -19.12  | -18.89 | -49.65 | -55.67 | -59.57   |
| $RSB = 50dB$ | -8.36 | -52.02 | -15.84  | -19.88 | -15.06 | -27.93 | -27.14   |
| $RSB = 10dB$ | -7.44 | -12.07 | -19.39  | -17.26 | -0.64  | -17.75 | -18.17   |

TABLE 2.11 – Erreur de corrélation  $E_{corr}(dB)$ , pour des sources corrélées

|              | N-ICA        | BI-NSS | multNMF | detNMF | VCA             | SISAL           | nLCA-IVM |
|--------------|--------------|--------|---------|--------|-----------------|-----------------|----------|
| Sans bruit   | $1, 88.10^3$ | 0      | 0       | 0      | $3, 19.10^{-1}$ | $5, 5.10^{-4}$  | 0        |
| $RSB = 50dB$ | $1, 68.10^3$ | 0      | 0       | 0      | $2, 39.10^3$    | $5, 15.10^{-4}$ | 0        |
| $RSB = 10dB$ | $1, 33.10^3$ | 0      | 0       | 0      | $5, 26.10^3$    | $6.10^{-3}$     | 0        |

TABLE 2.12 – Négativité des sources estimées  $E_{neg}$ , pour des sources corrélées

**Commentaires sur les résultats de séparation :** Comme attendu, la méthode N-ICA (qui fait l'hypothèse de sources indépendantes) échoue à séparer des sources corrélées. Cependant bien que faisant aussi l'hypothèse d'indépendance des sources, la méthode BI-NSS présente de bons indices de performance  $E_{sep}$  et  $E_{corr}$ . Ce résultat nous permet de dire que l'hypothèse d'indépendance des sources supposée par cette méthode est une hypothèse

|              | N-ICA | BI-NSS   | multNMF | detNMF | VCA  | SISAL | nLCA-IVM |
|--------------|-------|----------|---------|--------|------|-------|----------|
| Sans bruit   | 8.68  | 13295.44 | 102     | 290.81 | 0.26 | 27.28 | 6.95     |
| $RSB = 50dB$ | 4.12  | 13380.77 | 104.44  | 302.88 | 0.28 | 29.54 | 4.64     |
| $RSB = 10dB$ | 2.82  | 62913,4  | 103.26  | 295.93 | 0.23 | 30.03 | 7.81     |

TABLE 2.13 – Coût de calcul  $T(s)$ , pour des sources corrélées

”technique” afin de pouvoir approximer la densité jointe par le produit des densités marginales mais n’est pas une condition sine qua non au fonctionnement de cette méthode. Par ailleurs, du fait du grand nombre d’échantillons, la méthode bayésienne présente un temps de calcul extrêmement élevé comparativement aux autres méthodes. Les méthodes géométriques présentent aussi de bonnes performances de séparation en absence de bruit mais ces performances se dégradent fortement avec le bruit.

### 2.5.5 Synthèse des résultats de séparation

A l’issu de ces premières simulations, nous tirons les conclusions suivantes :

- L’analyse en composantes indépendantes non-négatives, présente de bonnes performances de séparation dans le cas de sources indépendantes (autant pour les sources parcimonieuses que pour les sources non parcimonieuses) et une bonne résistance au bruit. Elle nécessite néanmoins d’avoir un nombre “suffisant” d’échantillons non nuls afin de minimiser l’erreur de blanchiment susceptible de dégrader les performances de séparation. De plus, cette méthode a l’inconvénient majeur de ne pas être adaptée aux sources corrélées et il y a apparition de valeurs négatives dans les sources estimées.
- L’algorithme bayésien utilisé dans les simulations a été initialement conçu pour la séparation d’images hyperspectrales. Les résultats de séparation montre que les choix des lois a priori des sources et des coefficients de la matrice de mélange peuvent être assez flexibles, cette méthode présente aussi une bonne résistance au bruit. D’autre part, les résultats de séparation sur des sources corrélées ont montré que l’hypothèse d’indépendance mutuelle des sources, des coefficients de la matrice de mélange et du bruit n’est pas indispensable au fonctionnement de la méthode. Il s’agit plutôt d’une hypothèse “technique”, qui permet d’exprimer la loi à posteriori conjointe à l’aide des lois marginales a priori. La méthode bayésienne présente néanmoins un coût de calcul extrêmement élevé, ce qui constitue une limite majeure à son utilisation dans des problèmes de grandes dimensions comme c’est le cas en séparation d’image TEP où le nombre d’échantillons peut atteindre plusieurs millions.
- Dans les simulations précédentes, les méthodes de factorisation en matrices non-négatives, sans et avec régularisation, ont en règle générale donné des résultats de séparation peu satisfaisants. Cela s’explique par le fait que d’un point de vue théorique, ces méthodes ne garantissent pas l’unicité de la solution de séparation et donc en pratique elles sont sensibles à l’initialisation.
- La méthode géométrique fondée sur la recherche des points extrêmes du nuage des observations (VCA) présente d’excellents indices de performances en absence de bruit

et lorsque l’hypothèse de dominance locale des sources est vérifiée. En revanche les performances de cette méthode se dégradent fortement en présence de bruit (en particulier lorsque les sources sont parcimonieuses) et/ou lorsque les sources ne vérifient pas la dominance locale, ce qui est précisément le cas des images TEP. Cette méthode possède aussi des valeurs négatives dans les sources estimées.

- Les méthodes géométriques fondées sur la recherche du simplexe de volume minimal contenant le nuage des observations (SISAL et nLCA-IVM) présentent aussi de très bons indices de performances en absence de bruit, autant pour les sources indépendantes que pour les sources corrélées. Mais comme pour VCA, les performances de ces méthodes se dégradent significativement avec le bruit (en particulier lorsque les sources sont parcimonieuses). Par ailleurs, ce type de méthodes nécessitent l’additivité totale des sources. Dans le cas où cette dernière hypothèse n’est pas vérifiée (comme par exemple c’est le cas dans les spectres de masse et dans les images TEP), il est possible de forcer cette contrainte en normalisant les observations. Cette opération peut néanmoins amplifier le bruit quand les sources ont de fortes dynamiques et dégrader les performances de séparations.

Il ressort de la synthèse des résultats de séparation que les méthodes géométriques fondées sur la recherche du simplexe de volume minimal contenant le nuage des observations présentent le plus d’avantages en termes de “faibles” hypothèses sur les sources (elles ne nécessitent pas l’indépendance ou la dominance locale des sources), d’efficacité de séparation et de coût de calcul. Ces méthodes sont néanmoins assez sensibles au bruit.

## 2.6 Conclusion

Dans ce deuxième chapitre, nous avons présenté l’état de l’art des méthodes de séparation de sources non-négatives, en distinguant les méthodes statistiques et les méthodes déterministes. Une série d’évaluations dans différentes situations de séparation de sources non-négatives nous a permis de mettre en exergue les avantages et les inconvénients de ces différentes méthodes.

Dans le prochain chapitre, nous développons, dans le même esprit que les méthodes géométriques fondées sur l’estimation du simplexe de volume minimal contenant les observations, une méthode géométrique fondée sur les cônes simpliciaux. La méthode proposée intègre une étape de prétraitement qui permet d’améliorer sa robustesse au bruit comparativement aux autres méthodes géométriques, en particulier lorsque les sources sont parcimonieuses.





## Deuxième partie

# Développement d'une nouvelle méthode géométrique de séparation de sources non-négatives



## Chapitre 3

# Séparation de sources non-négatives par estimation du cône simplicial d'ouverture minimale contenant le nuage de points des signaux observations

### Sommaire

---

|     |  |    |
|-----|--|----|
| 3.1 | Introduction . . . . .                             | 48 |
| 3.2 | Synthèse de la méthode SCSSA-UNS . . . . .         | 48 |
| 3.3 | Commentaires des résultats de séparation . . . . . | 50 |
| 3.4 | Conclusions . . . . .                              | 50 |

---

### 3.1 Introduction

Ce chapitre est présenté sous la forme d'un article scientifique soumis à la revue IEEE Transactions on Signal Processing le 8 octobre 2012. Nous y développons une méthode géométrique de séparation des sources non-négatives baptisée **SCSA-UNS** pour Simplicial Cone Shrinking Algorithm for Unmixing Non-negative Sources. La méthode développée estime la matrice de mélange et les sources en cherchant le cône simplicial d'ouverture minimale qui contient le nuage de points des signaux observés.

La première partie de l'article résume les méthodes existantes de séparation de sources non-négatives, ces méthodes ont déjà été présentées dans le chapitre 2 de mémoire. Nous détaillons ensuite la méthode SCSA-UNS, en établissant des conditions nécessaires et des conditions suffisantes sur les sources, pour garantir l'unicité de la solution. Les résultats de simulation montrent l'efficacité de la méthode proposée sur des données synthétiques indépendantes et corrélées. Nous avons aussi appliqué notre méthode à des données réelles de spectrométrie de masse mesurées à la sortie d'un chromatographe en phase liquide pour l'estimation des spectres de masse de composés chimiques purs, ainsi qu'à des images réelles de tomographie par émission de positons pour l'estimation des pharmacocinétiques d'un traceur radio-actif dans le cerveau.

### 3.2 Synthèse de la méthode SCSA-UNS

Pour commencer, nous montrons que si les sources  $\mathbf{S}$  sont non-négatives et le mélange surdéterminé alors, en absence de bruit, le problème d'estimation de la matrice de mélange et des sources ( $\hat{\mathbf{A}}$  et  $\hat{\mathbf{S}}$ ) à partir des observations  $\mathbf{X}$ , peut se ramener à rechercher une matrice carrée non-négative  $\mathbf{A}_r$  et l'estimée des sources  $\hat{\mathbf{S}}$  à partir d'une autre matrice non-négative  $\mathbf{X}_r$ . Ce résultat nous permet de nous restreindre à résoudre le problème de séparation de sources non-négatives dans le cas où la matrice de mélange est aussi non-négative ( $\mathbf{A} \geq 0$ ) et le mélange est déterminé ( $m = n$ ).

Nous définissons le cône simplicial engendré par une matrice  $\mathbf{V} \in \mathbb{R}^{n \times n}$  (pas forcément non-négative), noté  $\mathbf{Span}^+(\mathbf{V})$ , par :

$$\mathbf{Span}^+(\mathbf{V}) = \{\mathbf{z} \mid \mathbf{z} = \mathbf{V}\mathbf{y} \text{ avec } \mathbf{y} \in \mathbb{R}_+^n\} \quad (3.1)$$

On peut noter que l'orthant positif  $\mathbb{R}_+^n$  est le cône simplicial engendré par  $\mathbf{I}_n$ , la matrice identité de taille  $n \times n$  :  $\mathbb{R}_+^n = \mathbf{Span}^+(\mathbf{I}_n)$ .

**En absence de bruit :** Étant donné que les sources sont non-négatives, le nuage de points des signaux observés est contenu dans le cône simplicial engendré par les colonnes de la matrice de mélange :

$$\{\mathbf{X}(k), 1 \leq k \leq K\} \subseteq \mathbf{Span}^+(\mathbf{A}) \quad (3.2)$$

La matrice de mélange étant supposée aussi non-négative,  $\mathbf{Span}^+(\mathbf{A})$  est inclus dans l'orthant positif :

$$\mathbf{Span}^+(\mathbf{A}) \subseteq \mathbf{Span}^+(\mathbf{I}_n) \quad (3.3)$$

A partir des résultats (3.2) et (3.3), l'intuition de la méthode SCSA-UNS est d'estimer la matrice de mélange en cherchant un cône simplicial contenant le nuage de points des signaux observés et contenu dans l'orthant positif. Or, sans aucune contrainte supplémentaire, il existe une infinité de tels cônes (l'orthant positif en étant un exemple). Pour garantir de retrouver la vraie matrice de mélange, nous définissons *l'ouverture d'un cône simplicial* généré par les colonnes d'une matrice *non-négative*  $\mathbf{V}$  par l'équation (3.4) :

$$\mathcal{A}(\text{Span}^+(\mathbf{V})) = \frac{|\det(\mathbf{V})|}{\|\mathbf{v}_1\| \times \|\mathbf{v}_2\| \times \cdots \times \|\mathbf{v}_n\|} \quad (3.4)$$

où  $\mathbf{v}_i$  est la  $i$ -ième colonne de la matrice  $\mathbf{V}$ . Nous introduisons l'hypothèse suivante :

**(H)** :  $\text{Span}^+(\mathbf{A})$  est l'unique cône d'ouverture minimale contenant le nuage de points des signaux observations et contenu dans l'orthant positif.

Nous avons établi (voir article) des conditions nécessaires et des conditions suffisantes sur les sources pour satisfaire l'hypothèse **(H)**.

Sous cette hypothèse, nous avons montré que la matrice de mélange peut être estimée de façon unique en résolvant l'équation (3.5) :

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A} \geq 0, \mathbf{A}^{-1} \mathbf{x} \geq 0} \mathcal{A}(\text{Span}^+(\mathbf{A})) \quad (3.5)$$

SCSA-UNS est un algorithme de résolution de cette équation, le principe de la méthode consiste à partir d'un cône initial  $\text{Span}^+(\mathbf{A}_0)$  contenant la totalité du nuage de points des signaux observés et à construire une suite de cônes simpliciaux inclus les uns dans les autres et contenant chacun la totalité du nuage de points des signaux observés ; l'ouverture du cône courant est inférieure à l'ouverture du cône précédent. Si  $\text{Span}^+(\mathbf{W})$  est le cône simplicial courant contenant la totalité du nuage de points des signaux observés, le cône simplicial suivant,  $\text{Span}^+(\mathbf{U})$ , est obtenu en multipliant  $\mathbf{W}$  à droite par une matrice de la forme  $\mathbf{R}_i$  donnée ci-dessous, les coefficients  $r_{ji}$  sont calculés de sorte à conserver la totalité du nuage de points des signaux observations à l'intérieur de  $\text{Span}^+(\mathbf{U})$ .

$$\mathbf{U} = \mathbf{W}\mathbf{R}_i \quad (3.6)$$

$$\mathbf{R}_i = \begin{pmatrix} 1 & 0 & \cdots & 0 & r_{1i} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & r_{2i} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 & r_{(i-1)i} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & r_{(i+1)i} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{ni} & 0 & \cdots & 1 \end{pmatrix} \quad (3.7)$$

L'algorithme s'arrête lorsque le cône courant s'ajuste au nuage de points des signaux observés.

**En présence d'un bruit additif :** Nous proposons, de réduire d'abord l'effet du bruit en supprimant les points les plus affectés par le bruit, avant d'estimer la matrice de mélange et les sources. Ces points sont susceptibles de modifier significativement le nuage de points des signaux observés et donc de conduire à une mauvaise estimation de la matrice de mélange et des sources. Nous les identifions en sélectionnant toutes les colonnes de la matrice des observations de normes inférieures à un certain seuil que nous fixons en fonction du rapport signal à bruit.

### 3.3 Commentaires des résultats de séparation

**Sources synthétiques.** La méthode SCSA-UNS a d'abord été évaluée sur des sources synthétiques et comparée à d'autres méthodes de séparation de sources non-négatives. Les résultats de séparation sur des sources indépendantes ainsi que des sources corrélées montrent que la méthode développée permet d'estimer correctement la matrice de mélange et les sources dans les deux cas (sources indépendantes et sources corrélées) et présente une meilleure résistance au bruit que les autres méthodes géométriques, en particulier lorsque les sources sont parcimonieuses.

**Sources réelles.** Des évaluations ont été aussi effectuées sur des données réelles LC-MS et sur des images dynamiques TEP.

Les résultats de séparation sur un jeu de données réelles de spectrométrie de masse couplée à la chromatographie en phase liquide ont permis d'estimer les spectres de masse élémentaires de 11 composés chimiques commerciaux à partir des spectres de masse du mélange homogène de ces composés.

L'exécution de l'algorithme sur une série images de tomographie par émission de positons, a permis d'estimer les pharmacocinétiques du FluoroDéoxyGlucose (FDG) marqué au Fluor 18 dans le cerveau. En particulier, nous avons correctement estimé la cinétique artérielle du traceur uniquement à partir des images TEP. Ce résultat est confirmé en comparant la cinétique artérielle estimée par séparation de sources avec la cinétique artérielle de référence obtenue par prélèvements sanguins.

### 3.4 Conclusions

Nous avons développé dans ce chapitre 3 une méthode géométrique de séparation de sources non-négatives dénommée SCSA-UNS. Cette méthode estime la matrice de mélange et les sources en cherchant le cône simplicial d'ouverture minimale contenant le nuage de points des signaux observés. Elle ne requiert pas l'indépendance des sources, ni leur dominance locale, ni leur additivité totale. Les résultats de séparation sur des données synthétiques et sur des données réelles montrent l'efficacité de la méthode proposée.

Dans le prochain chapitre, nous donnons quelques détails de l'implantation pratique de la méthode SCSA-UNS et nous étudions expérimentalement ses performances et sa robustesse.

# Non-negative Blind Source Separation Algorithm based on Minimum Aperture Simplicial Cone

Wendyam Serge Boris Ouedraogo\*, Antoine Souloumiac, Meriem Jaidane, and Christian Jutten, *Fellow, IEEE*

**Abstract**—We address the problem of Blind Source Separation (BSS) when the hidden sources are Nonnegative (N-BSS). In this case, the scatter plot of the mixed data is contained in the simplicial cone generated by the columns of the mixing matrix. The proposed method, termed SCSA-UNS for Simplicial Cone Shrinking Algorithm for Unmixing Non-negative Sources, aims at estimating the mixing matrix and the sources by fitting a Minimum Aperture Simplicial Cone (MASC) to the cloud of mixed data points. SCSA-UNS is evaluated on both independent and correlated synthetic data and compared to other N-BSS methods. Simulations are also performed on real Liquid Chromatography-Mass Spectrum (LC-MS) data for the metabolomic analysis of a chemical sample, and on real dynamic Positron Emission Tomography (PET) images to study the pharmacokinetics of the [18F]-FDG (FluoroDeoxyGlucose) tracer in the brain.

**Index Terms**—Blind Source Separation, Non-negativity, Simplicial Cone, Aperture, LC-MS, dynamic PET imaging

## I. INTRODUCTION

THE well-known problem of Non-negative Blind Source Separation (N-BSS) occurs in many situations of signal and image processing. Solving the N-BSS problem consists in retrieving the hidden non-negative sources and their corresponding mixing profiles from the measured data (also denoted by observations), which are mixtures of the original sources. The use of N-BSS on actual data have continuously increased these last decades, including chemometrics [51] [45], biomedical image processing [27] [43], spectrometry and spectroscopy [5] [30], remote sensing [29] [18], speech and music power spectra decomposition [47] [15], to name a few. Interested reader can refer to [40] and [11] for more details. The batch noiseless linear instantaneous mixture model of the N-BSS problem is given by equation (1):

$$\mathbf{X} = \mathbf{A}\mathbf{S} \quad (1)$$

where  $\mathbf{X} \in \mathbb{R}_+^{m \times K}$  is the known non-negative matrix of observations,  $\mathbf{A} \in \mathbb{R}_+^{m \times n}$  is the unknown non-negative mixing matrix and  $\mathbf{S} \in \mathbb{R}_+^{n \times K}$  is the hidden non-negative sources matrix.  $m$ ,  $n$  and  $K$  are respectively the number of observations, the number of sources and the number of samples.  $x_i(k)$  is the  $k$ th sample of observation  $\mathbf{x}_i$ ,  $s_j(k)$  is the  $k$ th sample of source

$\mathbf{s}_j$ , and  $a_{ij}$  is the  $i$ th value of the profile of source  $j$ . The task of solving problem (1) consists in recovering the sources  $\mathbf{S}$  and their profiles  $\mathbf{A}$ , given only the observations  $\mathbf{X}$ , without any prior knowledge on  $\mathbf{S}$  and  $\mathbf{A}$  except their non-negativity. Several methods were proposed for solving problem (1), one can distinguish statistical approaches and deterministic ones.

The methods of the first class use the statistical properties of the sources to design the separating algorithms. They include Non-negative Independent Components Analysis (N-ICA) [37] and Bayesian Inference based Non-negative Source Separation (BI-NSS) [20] [30].

N-ICA assumes that the *non-negative* sources,  $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$  are random variables, which are mutually statistically *independent* and *well-grounded*<sup>1</sup> [36]. The sources and the mixing matrix are estimated by whitening the observations and by rotating the whitened data to make them non-negative [37] [39]. Unfortunately N-ICA is not suited to correlated sources. In Bayesian approach [20] [30], the sources and the mixing profiles (columns of the mixing matrix) are modeled by random variables. The idea is to assign a specific prior probability density to each variable, and to derive the joint posterior probability density. The sources and mixing profiles are obtained by posterior estimator (posterior mean or posterior maximum) using the Baye's rule and the Markov Chain Monte Carlo (MCMC) sampling. In the case of non-negative sources, BI-NSS uses non-negative prior probability density (such as Gamma density) for sources and profiles. This method can unfortunately be computationally complex and time-consuming especially for large scale data.

For solving the N-BSS problem, deterministic approaches require weak assumption on the sources distribution, in particular the independence of the sources is not necessary. These approaches include Non-negative Matrix Factorization (NMF) and geometrical methods.

Firstly introduced by Paatero and Tapper [34], NMF was popularized by Lee and Seung [25] and became the most used method for solving the N-BSS problem. In NMF,  $\mathbf{A}$  and  $\mathbf{S}$  are estimated by minimizing a divergence measure between the left and right parts of equation (1) under a non-negativity constraint on both the estimated mixing matrix and sources. Reference [12] gives more details on different divergence measures and optimization algorithms for NMF. Nevertheless, the non-negativity constraint alone (on  $\mathbf{S}$  and  $\mathbf{A}$ ) is not enough to guarantee the uniqueness of the factorization as shown in [14] [31] [23]. So, according to the prior knowledge about the

Wendyam Serge Boris Ouedraogo is with the GIPSA-lab/DIS, 11 rue des Mathématiques, BP 46, F-38402 Saint Martin d'Hères Cedex-France, (e-mail: wendyam.ouedraogo@gipsa-lab.grenoble-inp.fr).

Antoine Souloumiac is with CEA LIST/DCSI, Bat. 565, PC 192, F-91191 Gif-sur-Yvette Cedex France, (e-mail: antoine.souloumiac@cea.fr).

Meriem Jaidane is with ENIT/U2S, BP 37, Le Belvedere 1002 Tunis, Tunisie, (e-mail: meriem.jaidane@planet.tn).

Christian Jutten is with the GIPSA-lab/DIS, 11 rue des Mathématiques, BP 46, F-38402 Saint Martin d'Hères Cedex-France, and with Institut Universitaire de France, (e-mail: christian.jutten@gipsa-lab.grenoble-inp.fr).

<sup>1</sup>A *well-grounded* random variable  $s_i$  has a nonvanishing probability density function around zero, i.e.  $\forall \delta > 0, Pr(s_i < \delta) > 0$



sources (respectively the profiles), some additional constraints, such as the sparsity and/or the smoothness of the sources and/or the mixing profiles, were incorporated in NMF to improve the physical meaning of the estimated sources and/or mixing profiles, and reduce the number of possible solutions [17] [9]. A novel direction was introduced to tackle the non-uniqueness of NMF, by adding a minimum determinant constraint on the estimated mixing matrix [22] [44].

The geometrical methods are the other deterministic approaches for solving the Non-negative Blind Source Separation problem. The first geometrical method was introduced by Puntonet *et al.* [41] for unmixing two sources having bounded probability densities. The mixing matrix is estimated by finding the slopes of the parallelogram containing the scatter plot of mixed data. Babaie-Zadeh *et al.* [2] have proposed another geometrical method for separating sparse sources. The latter method, which can deal with more than two sources, estimates the mixing matrix by clustering the scatter plot of mixed data and fitting a line (hyper-plane) to each cluster.

Other geometric methods were proposed for sources satisfying certain properties. If for every source there is at least one instance where the underlined source is active and all the others are not (i.e.  $\forall 1 \leq i \leq n, \exists k_i (1 \leq k_i \leq K)$  such as  $s_i(k_i) \neq 0$  and  $s_j(k_i) = 0$  for  $j \neq i$ ), then the sources are termed *locally dominant*. In that case,  $\mathbf{A}$  can be estimated by finding the vertices of the convex hull, of the scatter plot of mixed data [24] [7] [32]. Unfortunately, the *local dominance* assumption (termed pure pixel assumption in hyperspectral unmixing) is seldom verified in practice.

If the sum on every column of the sources matrix equal to one (i.e.  $\forall 1 \leq k \leq K, \sum_{j=1}^n s_j(k) = 1$ ), then the sources are termed *full additive*, and one can estimate the mixing matrix by looking for the Minimum Volume Simplex (MVS) containing the scatter plot of mixed data [13] [28] [3] [6]. The MVS methods do not require the local dominance of the sources and, in noiseless case, one can relax the full additivity constraint by normalizing each column of the data matrix to a unit sum as shown in [8] and [16]. However, in the noisy case, enforcing this normalization may amplify the noise and/or lead to a bad estimation of the sources, especially if the number of sources is overestimated, as illustrated on Fig. 10.

This paper proposes a new geometrical method for solving the overdetermined N-BSS problem. The proposed method, denoted Simplicial Cone Shrinking Algorithm for Unmixing Non-negative Sources (SCSA-UNS), estimates the mixing matrix and the sources by finding the ‘‘Minimum Aperture Simplicial Cone’’ (MASC) containing the scatter plot of the mixed data. It neither require the independence of sources, nor their local dominance, or even their full additivity. The paper is organized as follows. In section II, we review the geometrical foundations of the N-BSS problem and we derive the Minimum Aperture Simplicial Cone based approach. Section III gives conditions on the sources for the uniqueness of N-BSS, and Section IV describes the proposed method. In Section V, we proposed a preprocessing for reducing the noise effect. Section VI and Section VII present simulation results on synthetic, and on real data. Finally Section VIII derives

conclusions.

## II. GEOMETRICAL FOUNDATIONS OF THE NON-NEGATIVE BLIND SOURCE SEPARATION PROBLEM

We restrict to the case where the mixture is determined (i.e.  $m = n$ ). If the mixture is overdetermined ( $m > n$ ), one can reduce to the previous case by a proper dimension reduction, as described in section IV-B.

### A. Useful definitions and concepts

In the following,  $\mathbb{R}_+^n$  denotes the set of non-negative  $n$ -dimensional vectors, and  $\mathbb{R}_+^{n \times n}$  denotes the set of  $n$ -size matrices whose all entries are non-negative.

Given a full rank matrix  $\mathbf{U} \in \mathbb{R}_+^{n \times n}$  ( $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ , where  $\mathbf{u}_i \in \mathbb{R}_+^n$  is the  $i$ th column of  $\mathbf{U}$ ), we introduce the following definitions:

*Definition 1: Simplicial Cone*

The *Simplicial Cone* generated by the columns of  $\mathbf{U}$ , denoted by  $\mathbf{Span}^+(\mathbf{U})$ , is defined as:

$$\mathbf{Span}^+(\mathbf{U}) = \{\mathbf{z} \mid \mathbf{z} = \mathbf{U}\mathbf{y} \text{ with } \mathbf{y} \in \mathbb{R}_+^n\} \quad (2)$$

*Definition 2: Edge Vector of a Simplicial Cone*

By abusing the notation, the  $i$ -th *Edge Vector* of the simplicial cone  $\mathbf{Span}^+(\mathbf{U})$ , denoted by  $\mathcal{E}_i(\mathbf{Span}^+(\mathbf{U}))$ , is defined as:

$$\mathcal{E}_i(\mathbf{Span}^+(\mathbf{U})) = \{\mathbf{z} \mid \mathbf{z} = \alpha \mathbf{u}_i \mid \alpha \geq 0\} \quad (3)$$

*Definition 3: Facet of a Simplicial Cone*

In the same way, the  $i$ -th *Facet* of the simplicial cone  $\mathbf{Span}^+(\mathbf{U})$ , denoted by  $\mathcal{F}_i(\mathbf{Span}^+(\mathbf{U}))$ , is defined as:

$$\mathcal{F}_i(\mathbf{Span}^+(\mathbf{U})) = \mathbf{Span}^+\{\mathbf{U} \setminus \{\mathbf{u}_i\}\} \quad (4)$$

*Definition 4: Vertex of a Simplicial Cone*

The *Vertex* of the simplicial cone  $\mathbf{Span}^+(\mathbf{U})$ , denoted by  $\mathcal{V}(\mathbf{Span}^+(\mathbf{U}))$ , is defined as:

$$\mathcal{V}(\mathbf{Span}^+(\mathbf{U})) = \bigcap_i \mathcal{E}_i(\mathbf{Span}^+(\mathbf{U})) \quad (5)$$

**Remarks :**

- For ease of the notations, we set  $\mathcal{E}_i(\mathbf{U}) = \mathcal{E}_i(\mathbf{Span}^+(\mathbf{U}))$ ,  $\mathcal{F}_i(\mathbf{U}) = \mathcal{F}_i(\mathbf{Span}^+(\mathbf{U}))$ , and  $\mathcal{V}(\mathbf{U}) = \mathcal{V}(\mathbf{Span}^+(\mathbf{U}))$
- The simplicial cone,  $\mathbf{Span}^+(\mathbf{U})$ , has  $n$  edge vectors,  $n$  facets and 1 vertex.
- The vertex of  $\mathbf{Span}^+(\mathbf{U})$  is located at the origin of  $\mathbb{R}^n$
- $\mathbb{R}_+^n$ , the *positive orthant*, is the simplicial cone generated by the identity matrix  $\mathbf{I}_n$ :  $\mathbb{R}_+^n = \mathbf{Span}^+(\mathbf{I}_n)$

Given the previous definitions, one can infer the following:

*Lemma 1:* The scatter plot of the sources  $\{\mathbf{s}(k), 1 \leq k \leq K\}$ , with  $\mathbf{s}(k) = [s_1(k) \ s_2(k) \ \dots \ s_n(k)]^T$ , is contained in the positive orthant,  $\mathbf{Span}^+(\mathbf{I}_n)$ :

$$\{\mathbf{s}(k), 1 \leq k \leq K\} \subseteq \mathbf{Span}^+(\mathbf{I}_n) \quad (6)$$

*Proof:* Follow the definition of a simplicial cone. ■

*Lemma 2:* The scatter plot of the mixed data  $\{\mathbf{x}(k), 1 \leq k \leq K\}$ , with  $\mathbf{x}(k) = [x_1(k) \ x_2(k) \ \dots \ x_n(k)]^T$ , is contained in  $\mathbf{Span}^+(\mathbf{A})$ , the simplicial cone generated by the columns of the mixing matrix:

$$\{\mathbf{x}(k), 1 \leq k \leq K\} \subseteq \mathbf{Span}^+(\mathbf{A}) \quad (7)$$

*Proof:* Follow the definition of a simplicial cone. Since  $\mathbf{A}$  is also non-negative, one can deduce that:

$$\text{Span}^+(\mathbf{A}) \subseteq \text{Span}^+(\mathbf{I}_n) \quad (8)$$

To illustrate Lemma 2, for  $m = n = 3$ , we consider the mixing matrix  $\mathbf{A}$  given below, and we generate data according to equation (1), where each source is generated following the uniform distribution between 0 and 1.

$$\mathbf{A} = \begin{pmatrix} 0.67 & 0.49 & 0.32 \\ 0.65 & 0.02 & 0.65 \\ 0.54 & 0.88 & 0.81 \end{pmatrix}$$

Fig. 1(a) shows the scatter plot of mixed data included in the simplicial cone generated by  $\mathbf{A}$ . For ease of visualization, we represent on Fig. 1(b) the projection of  $\text{Span}^+(\mathbf{A})$  and the mixed data, on the plane  $\mathcal{P}$  defined by:

$$\mathcal{P} = \left\{ \mathbf{z} = [z_1, z_2, z_3]^T \in \mathbb{R}^3 \mid \sum_{i=1}^3 z_i = 1 \right\} \quad (9)$$

The projected data  $\tilde{\mathbf{X}}$ , and cone  $\text{Span}^+(\tilde{\mathbf{A}})$  are obtained by:

$$\tilde{x}_i(k) = x_i(k) / \sum_{l=1}^3 x_l(k) \quad \text{and} \quad \tilde{a}_{ij} = a_{ij} / \sum_{l=1}^3 a_{il} \quad (10)$$

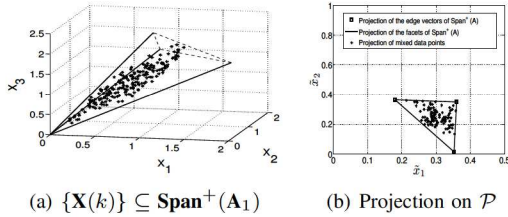


Fig. 1. Scatter plot of mixed data included in  $\text{Span}^+(\mathbf{A})$

*Theorem 1:* Given two full rank non-negative matrices  $\mathbf{U} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{V} \in \mathbb{R}_+^{n \times n}$ :

$$\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V}) \Leftrightarrow \mathbf{U} = \mathbf{V}\mathbf{M} \quad (11)$$

where  $\mathbf{M} \in \mathbb{R}_+^{n \times n}$  and  $rg(\mathbf{M}) = n$

*Proof:* Proof of Theorem 1 is given in Appendix A. ■

*Theorem 2:* Given two full rank non-negative matrices  $\mathbf{U} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{V} \in \mathbb{R}_+^{n \times n}$ :

$$\text{Span}^+(\mathbf{U}) = \text{Span}^+(\mathbf{V}) \Leftrightarrow \mathbf{U} = \mathbf{V}\mathbf{M} \quad (12)$$

where  $\mathbf{M} = \mathbf{D}\mathbf{P}$  and  $rg(\mathbf{M}) = n$

$\mathbf{D} \in \mathbb{R}_+^{n \times n}$  is a diagonal matrix, and  $\mathbf{P} \in \mathbb{R}_+^{n \times n}$  is a permutation one.  $\mathbf{M}$  is called a monomial matrix.

*Proof:* Proof of Theorem 2 is given in Appendix B. ■

### B. Minimum Aperture Simplicial Cone based N-BSS

According to equations (7) and (8), one can estimate the mixing matrix (and the sources) by looking for a simplicial cone containing the scatter plot of the mixed data and contained in the positive orthant. But without any additional constraint, there are infinite number of such cones [14] [23]

(the positive orthant itself is an example). By assuming that the sources satisfies the full additivity constraint (see Section I), some geometrical methods [13] [28] [3] [6], look for the Minimum Volume Simplex (MVS) containing the scatter plot of mixed data. In this paper, except their non-negativity, no other assumption is made on the sources. In particular, the sources and the mixed data do not have upper bounds, and the volume of a simplicial cone containing the scatter plot of mixed data is undefined. One can normalize each column of the data matrix to a unit sum, and consider the volume of the simplex containing the normalized data, however, as we mentioned in Section I above, enforcing this normalization may amplify the noise and/or lead to a poor estimation of the sources. Therefore, instead of using the volume, we characterize a simplicial cone by its *Aperture* defined below:

*Definition 5: Aperture of a Simplicial Cone*

The *Aperture* of the given simplicial cone,  $\text{Span}^+(\mathbf{U})$ , generated by a full rank matrix  $\mathbf{U}$ , denoted  $\mathcal{A}(\text{Span}^+(\mathbf{U}))$ , is defined by:

$$\mathcal{A}(\text{Span}^+(\mathbf{U})) = \frac{|\det(\mathbf{U})|}{\|\mathbf{u}_1\|_2 \times \|\mathbf{u}_2\|_2 \times \cdots \times \|\mathbf{u}_n\|_2} \quad (13)$$

**Remarks :**

- For ease of notation, we set:  $\mathcal{A}(\mathbf{U}) = \mathcal{A}(\text{Span}^+(\mathbf{U}))$
- For  $n = 2$ ,  $\mathcal{A}(\mathbf{U}) = \sin(\theta)$ , where  $\theta$  is the angle between the two column vectors of  $\mathbf{U}$ . For  $n > 2$ ,  $\mathcal{A}(\mathbf{U}) = \prod_{i=1}^{n-1} \sin(\theta_i)$ , where  $\theta_i$  is the principal angle between  $\mathcal{E}_{i+1}(\mathbf{U})$  and  $\text{Span}^+\{\mathbf{u}_1, \dots, \mathbf{u}_i\}$ .
- According to Hadamard's inequality which states that  $|\det(\mathbf{U})| \leq \prod_{i=1}^n \|\mathbf{u}_i\|_2$ , we have  $0 \leq \mathcal{A}(\mathbf{U}) \leq 1$ .
- $\mathcal{A}(\mathbf{I}_n) = 1$

*Theorem 3:* Given two full rank non-negative matrices  $\mathbf{U} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{V} \in \mathbb{R}_+^{n \times n}$ :

$$\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V}) \Rightarrow \mathcal{A}(\mathbf{U}) \leq \mathcal{A}(\mathbf{V}) \quad (14)$$

*Proof:* Proof of Theorem 3 is given in Appendix C. ■

*Theorem 4:* Given two full rank non-negative matrices  $\mathbf{U} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{V} \in \mathbb{R}_+^{n \times n}$ :

$$\text{Span}^+(\mathbf{U}) = \text{Span}^+(\mathbf{V}) \Rightarrow \mathcal{A}(\mathbf{U}) = \mathcal{A}(\mathbf{V}) \quad (15)$$

*Proof:* Proof of Theorem 4 directly comes from Theorem 2 and the definition of the *Aperture* of a simplicial cone. ■

*Definition 6: Minimum Aperture Simplicial Cone*

We state that  $\text{Span}^+(\mathbf{U})$  is a *Minimum Aperture Simplicial Cone* (MASC) containing the scatter plot of mixed data  $\{\mathbf{x}(k), 1 \leq k \leq K\}$ , if for any simplicial cone  $\text{Span}^+(\mathbf{V})$  containing the scatter plot of mixed data,  $\mathcal{A}(\mathbf{U}) \leq \mathcal{A}(\mathbf{V})$ .

By misnomer, we also state that  $\text{Span}^+(\mathbf{U})$  is a non-negative MASC containing the scatter plot of mixed data, if  $\text{Span}^+(\mathbf{U})$  is a MASC and  $\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{I}_n)$ .

Under certain conditions on the sources (which will be discussed in section III),  $\text{Span}^+(\mathbf{A})$  is the unique non-negative *Minimum Aperture Simplicial Cone* containing the scatter plot of the mixed data. In this case the mixing matrix can be estimated (up to positive scaling and permutation indeterminations) by finding this simplicial cone. In fact, if  $\text{Span}^+(\mathbf{W})$

is another non-negative MASC containing the scatter plot of mixed data, then  $\mathbf{Span}^+(\mathbf{W}) = \mathbf{Span}^+(\mathbf{A})$ , and therefore  $\mathbf{W} = \mathbf{ADP}$ , where  $\mathbf{D}$  is a diagonal matrix with non-negative entries, and  $\mathbf{P}$  is a permutation matrix (Theorem 2).

Before proposing an algorithm for estimating the mixing matrix, we give below conditions on the sources, under which  $\mathbf{Span}^+(\mathbf{A})$  is the unique non-negative *Minimum Aperture Simplicial Cone* containing the scatter plot of the mixed data.

### III. CONDITIONS ON THE SOURCES FOR UNIQUE N-BSS BY MINIMUM APERTURE SIMPLICIAL CONE

To ensure recovering of the true mixing matrix, and then the true sources in noiseless case,  $\mathbf{Span}^+(\mathbf{A})$  must be the unique non-negative MASC containing the scatter plot of mixed data. The following Theorem 5 transforms this condition into a condition on the sources.

**Theorem 5:**  $\mathbf{Span}^+(\mathbf{A})$  is the unique non-negative MASC containing the scatter plot of the mixed data if and only if  $\mathbf{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of the sources.

*Proof:* Proof of Theorem 5 is given in Appendix D. ■

#### A. Necessary condition for unique N-BSS by MASC

**Proposition 1:** If  $\mathbf{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of the sources, then there is at least one point of the cloud of sources on each facet of  $\mathbf{Span}^+(\mathbf{I}_n)$ , i.e.  $\forall 1 \leq i \leq n, \exists k_i$  such as  $s_i(k_i) = 0$

*Proof:* Assume that there is at least one facet of  $\mathbf{Span}^+(\mathbf{I}_n)$  where there is no point of the cloud of sources, i.e.  $\exists 1 \leq i \leq n \mid \forall 1 \leq k \leq K, s_i(k) > 0$  (Without loss of generality, set  $i = 1$  for the demonstration). Let's define the  $n$ -size square matrix  $\mathbf{U} = [u_{ij}]_{1 \leq i, j \leq n}$  as follow:

$$u_{ij} = \begin{cases} \min_{1 \leq k \leq K} \frac{s_1(k)}{s_2(k)}, s_2(k) \neq 0 & \text{if } i = 1 \text{ and } j = 2 \\ \delta_{ij} & \text{otherwise} \end{cases} \quad (16)$$

Since the sources are non-negative,  $u_{12} > 0$  and one can easily verify that  $\mathbf{U}^{-1}\mathbf{S} \geq 0$ . It comes that  $\mathbf{Span}^+(\mathbf{U})$  is a simplicial cone containing the scatter plot of the sources. Since  $u_{12} > 0$ , then  $\mathbf{Span}^+(\mathbf{U}) \subset \mathbf{Span}^+(\mathbf{I}_n)$ , hence  $\mathcal{A}(\mathbf{U}) < \mathcal{A}(\mathbf{I}_n)$ . Therefore  $\mathbf{Span}^+(\mathbf{I}_n)$  is not the unique non-negative MASC containing the scatter plot of sources. ■

#### B. Sufficient conditions for unique N-BSS by MASC

##### 1) Local dominance:

**Proposition 2:** If the sources are non-negative and *locally dominant* (see section I), then  $\mathbf{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of sources.

*Proof:* Since sources are non-negative and locally dominant, i.e.  $\forall 1 \leq i \leq n, \exists 1 \leq k_i \leq K \mid s_i(k_i) \neq 0$  and  $s_j(k_i) = 0$  for  $j \neq i$ , then the  $k_i$ th column of  $\mathbf{S}$  is  $\mathbf{s}(k_i) = s_i(k_i)\mathbf{e}_i$  where  $\mathbf{e}_i$  is the  $i$ th column of  $\mathbf{I}_n$ .

Let  $\mathbf{Span}^+(\mathbf{U})$  be a non-negative simplicial cone containing the scatter plot of sources, i.e.  $\forall 1 \leq i \leq n, \mathbf{s}(k_i) \in \mathbf{Span}^+(\mathbf{U})$  then  $\mathbf{e}_i \in \mathbf{Span}^+(\mathbf{U})$ , so  $\mathbf{Span}^+(\mathbf{I}_n) \subseteq \mathbf{Span}^+(\mathbf{U})$  and consequently  $\mathcal{A}(\mathbf{I}_n) \leq \mathcal{A}(\mathbf{U})$ . Therefore,  $\mathbf{Span}^+(\mathbf{I}_n)$  is a non-negative MASC containing the scatter plot of the sources.

If  $\mathbf{Span}^+(\mathbf{V})$  is another non-negative MASC containing the scatter plot of the sources, then  $\mathbf{V} \geq 0, \mathbf{V}^{-1}\mathbf{S} \geq 0$  and  $\mathcal{A}(\mathbf{V}) = \mathcal{A}(\mathbf{I}_n) = 1$ , which lead us to conclude that  $\mathbf{V} = \mathbf{I}_n\mathbf{M}$  where  $\mathbf{M}$  is a monomial matrix, and consequently  $\mathbf{Span}^+(\mathbf{I}_n)$  is the unique MASC containing the scatter plot of the sources. ■

##### 2) Independence and well-grounded:

**Proposition 3:** If the sources are *non-negative, independent and well-grounded*, then asymptotically,  $\mathbf{Span}^+(\mathbf{I}_n)$  tends to the unique non-negative MASC containing the scatter plot of the sources.

*Proof:* On the one hand the sources are well-grounded, then  $\forall 1 \leq i \leq n, \forall \delta > 0, Pr(\mathbf{s}_i < \delta) > 0$ . On the other hand, the sources are non-negative and independent, then for a fixed  $i$ :

$$Pr(\mathbf{s}_1 < \delta, \dots, \mathbf{s}_i > 0, \dots, \mathbf{s}_n < \delta) = Pr(\mathbf{s}_i > 0) \prod_{j=1, j \neq i}^n Pr(\mathbf{s}_j < \delta) \quad (17)$$

By making  $\delta \rightarrow 0$ , it follows that asymptotically, there will at least one index  $k_i$  such as  $s_i(k_i) > 0$  and  $s_j(k_i) \rightarrow 0$  for all  $j \neq i$ . Therefore, asymptotically, the sources tend to be locally dominant. Hence, asymptotically  $\mathbf{Span}^+(\mathbf{I}_n)$  tends to be the unique non-negative MASC containing the scatter plot of the source. ■

##### 3) Sufficiently spread condition:

**Proposition 4:** If for each facet of  $\mathbf{Span}^+(\mathbf{I}_n)$ , at least  $n - 1$  points of the scatter plot of the sources belong to underlined facet, and the vectors corresponding to these points are linearly independent, then  $\mathbf{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of the sources.

*Proof:* Let  $\mathbf{Span}^+(\mathbf{V})$  be a non-negative MASC containing the scatter plot of the sources, then  $\mathbf{V} \geq 0, \mathbf{V}^{-1}\mathbf{S} \geq 0$  and  $\mathcal{A}(\mathbf{V}) \leq 1$ . The task is to demonstrate that  $\mathbf{V}$  is a monomial matrix. For a fixed  $i$ , assume that  $\mathbf{s}(k_l^i), l = 1, 2, \dots, n - 1$  belongs to  $\mathcal{F}_i(\mathbf{Span}^+(\mathbf{I}_n))$ , then  $s_i(k_l^i) = 0$ , the submatrix  $\mathbf{S}^i = [\mathbf{s}(k_1^i) \ \mathbf{s}(k_2^i) \ \dots \ \mathbf{s}(k_{n-1}^i)]$  is of size  $n \times (n - 1)$ . Since  $\{\mathbf{s}(k_l^i), 1 \leq l \leq n - 1\} \subseteq \mathbf{Span}^+(\mathbf{V})$ , then there is a non-negative matrix  $\mathbf{Y} \geq 0$  of size  $n \times (n - 1)$ , such as  $\mathbf{S}^i = \mathbf{V}\mathbf{Y}$ . For  $1 \leq j \leq n$ , and  $1 \leq l \leq n - 1$ , the  $jl$ -th entry of  $\mathbf{S}^i$  is:

$$s_j(k_l^i) = \sum_{p=1}^n v_{jp}y_{pl} \quad (18)$$

where  $v_{jp}$  is the  $jp$ -th entry of  $\mathbf{V}$  and  $y_{pl}$  is the  $pl$ -th entry of  $\mathbf{Y}$ . For  $j = i$ , we get:

$$s_i(k_l^i) = 0 = \sum_{p=1}^n v_{ip}y_{pl} \quad (19)$$

Since  $v_{ip} \geq 0$  and  $z_{pl} \geq 0$ , then  $\forall 1 \leq p \leq n, v_{ip} = 0$  or  $y_{pl} = 0$ . For a fixed  $p$ , if  $v_{ip} > 0$  then  $\forall 1 \leq l \leq n - 1, y_{pl} = 0$ . It follow that each non zero  $v_{ip}$  involves that the corresponding row of  $\mathbf{Y}$  must have zeros entries. Furthermore, the vectors corresponding to the points  $\mathbf{S}(k_l^i), l = 1, 2, \dots, n - 1$  are linearly independent, then  $rg(\mathbf{S}^i) = n - 1 = \min(rg(\mathbf{V}), rg(\mathbf{Y}))$ , since  $\mathbf{Y}$  is size  $n \times (n - 1)$ , then there can be no more than one zeros row in  $\mathbf{Y}$ . Therefore, there is only one non zero entry on

the  $i$ th row of  $\mathbf{V}$ . By varying  $i$  from 1 to  $n$ , one can conclude that there is only one non zero entry on each row of  $\mathbf{V}$ . Since  $\mathbf{V}$  is non-singular, we deduce that  $\mathbf{V}$  is a monomial matrix. ■

Propositions 2, 3 and 4, give different sufficient conditions on the sources, under which  $\text{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing their scatter plot, and therefore  $\text{Span}^+(\mathbf{A})$  is the unique non-negative MASC containing the scatter plot of the mixed data. These conditions cannot all be deduced from each other, therefore we can state that each of this condition is sufficient but not necessary.

### C. Conjecture of necessary and sufficient condition on the sources for unique N-BSS by MASC

Based on the necessary condition, and the sufficient conditions described before, we conjecture that a necessary and sufficient condition on sources, under which  $\text{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of the sources can be established as follows:

- 1) There are at least one point of the scatter plot of sources belonging to each facet of the positive orthant, such as the vectors corresponding to these points are linearly independent:  $\forall 1 \leq i \leq n, \exists k_i \mid s_i(k_i) = 0$  and the  $n$ -size submatrix  $\mathbf{T} = [\mathbf{s}(k_1) \ \mathbf{s}(k_2) \ \cdots \ \mathbf{s}(k_n)]$  is full rank.
- 2) If  $\text{Span}^+(\mathbf{T}) \subset \text{Span}^+(\mathbf{I}_n)$ , then there is at least one point of the scatter plot of sources,  $\mathbf{s}(l)$  with  $1 \leq l \leq K$ , which lies outside any simplicial cone involving some or all of the vectors  $\mathbf{s}(k_1), \mathbf{s}(k_2), \dots, \mathbf{s}(k_n)$ .

Fig. 2 illustrates this intuitive condition on  $\mathcal{P}$ , in the case where  $m = n = 3$ . A proof of this conjecture is being studied.

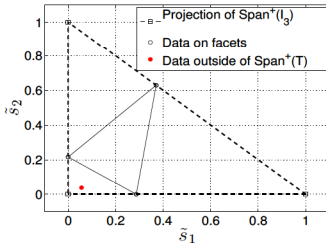


Fig. 2. Illustration of necessary and sufficient condition for  $n = 3$

## IV. SIMPLICIAL CONE SHRINKING ALGORITHM FOR UNMIXING NON-NEGATIVE SOURCES: SCSA-UNS

Assuming that  $\text{Span}^+(\mathbf{I}_n)$  is the unique non-negative MASC containing the scatter plot of the sources, and according to Theorem 5 the task of estimating the mixing matrix is reduced to finding the non-negative MASC containing the scatter plot of mixed data. This can be reduced to solving the following optimization problem (20), where  $\hat{\mathbf{A}}$  is the estimated mixing matrix and  $\mathcal{A}(\mathbf{A})$  is defined by equation (13):

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A} \geq 0, \mathbf{A}^{-1}\mathbf{X} \geq 0} \mathcal{A}(\mathbf{A}) \quad (20)$$

For solving (20), a possible direction is a gradient type method. However it would lead to non-convex optimization problem,

due to the denominator term of the criterion (13). Instead of this approach, we propose an iterative method which starts from an initial simplicial cone containing the scatter plot of the mixed data, and progressively decreases its aperture up to fit the scatter plot of the mixed data. The proposed algorithm, termed SCSA-UNS for Simplicial Cone Shrinking Algorithm for Unmixing Non-negative Sources, is described below.

### A. Proposed algorithm

#### 1) Finding a proper initial simplicial cone:

The first step of the proposed SCSA-UNS algorithm is finding an initial simplicial cone  $\text{Span}^+(\mathbf{U})$  containing the scatter plot of mixed data. Let  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ , for reason we will explain in Section IV-A2, we also want that  $\mathbf{u}_p^T \mathbf{u}_q \geq 0, \forall 1 \leq p, q \leq n$ . Since we first restrict to  $m = n$ , a trivial solution is  $\mathbf{U} = \mathbf{I}_n$ . One can also use an extended version of the solution of a local dominant based method, in this paper, we use extended VCA [32]. if  $\mathbf{V}$  is the mixing matrix estimated by VCA, then  $i$ -th column of  $\mathbf{U}$ ,  $\mathbf{u}_i$ , is computed by  $\mathbf{u}_i = \mathbf{v}_i - \alpha \bar{\mathbf{v}}$ , where  $\mathbf{v}_i$  is the  $i$ -th column of  $\mathbf{V}$  and  $\bar{\mathbf{v}}$  is the mean of the columns of  $\mathbf{V}$ . The coefficient  $\alpha$  should be adjusted in order to satisfy the required condition.

#### 2) Decreasing the aperture of the current Simplicial Cone:

We define the  $n$ -size matrices,  $\mathbf{R}_i$ , by equation (21) where  $r_{ii} = 1$  and  $r_{ji} \geq 0$  for  $1 \leq j \leq n, j \neq i$ .

$$\mathbf{R}_i = \begin{pmatrix} 1 & 0 & \cdots & 0 & r_{1i} & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & r_{2i} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & r_{(i-1)i} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & r_{(i+1)i} & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & r_{ni} & 0 & \cdots & 1 \end{pmatrix} \quad (21)$$

*Proposition 5:* Let  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ , the  $n$ -size matrix (where  $\mathbf{u}_p$  is the  $p$ -th column of  $\mathbf{U}$ ), and  $\text{Span}^+(\mathbf{U})$  the current simplicial cone containing the scatter plot of the mixed data, i.e  $\mathbf{U}^{-1}\mathbf{X} \geq 0$ . For a fixed  $i$ , let  $\mathbf{W} = \mathbf{U}\mathbf{R}_i$ . If  $\forall 1 \leq p, q \leq n, \mathbf{u}_p^T \mathbf{u}_q \geq 0$ , then:

- 1)  $\mathcal{A}(\mathbf{W}) \leq \mathcal{A}(\mathbf{U})$ .
- 2) The coefficients  $r_{ij}$  of  $\mathbf{R}_i$  can be computed such that  $\text{Span}^+(\mathbf{W})$  also contains the scatter plot of the mixed data, i.e  $\mathbf{W}^{-1}\mathbf{X} \geq 0$ .

*Proof:* Set  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n]$ , where  $\mathbf{w}_p$  is the  $p$ -th column of  $\mathbf{W}$ .

- i.  $|\det(\mathbf{W})| = |\det(\mathbf{U}\mathbf{R}_i)| = |\det(\mathbf{U})| |\det(\mathbf{R}_i)|$ , since  $\det(\mathbf{R}_i) = 1$  then  $|\det(\mathbf{W})| = |\det(\mathbf{U})|$
- ii.  $\forall l \neq i, \mathbf{w}_l = \mathbf{u}_l \Rightarrow \|\mathbf{w}_l\|_2 = \|\mathbf{u}_l\|_2$

iii.  $\mathbf{w}_i = \sum_{j=1}^n r_{ji} \mathbf{u}_j = \mathbf{u}_i + \sum_{j=1, j \neq i}^n r_{ji} \mathbf{u}_j$ , then

$$\mathbf{w}_i^T \mathbf{w}_i = \mathbf{u}_i^T \mathbf{u}_i + 2 \sum_{j=1, j \neq i}^n r_{ji} \mathbf{u}_i^T \mathbf{u}_j + \sum_{j=1, j \neq i}^n \sum_{p=1, p \neq i}^n r_{ji} r_{pi} \mathbf{u}_j^T \mathbf{u}_p \quad (22)$$

Since  $\forall 1 \leq p, q \leq n \mathbf{u}_p^T \mathbf{u}_q \geq 0$ , then  $\mathbf{w}_i^T \mathbf{w}_i \geq \mathbf{u}_i^T \mathbf{u}_i$ , and therefore  $\|\mathbf{w}_i\|_2 \geq \|\mathbf{u}_i\|_2$ .

According to the definition of the aperture of a simplicial cone given by equation (13), and considering i., ii. and iii., one can conclude that  $\mathcal{A}(\mathbf{W}) \leq \mathcal{A}(\mathbf{U})$ , this proves 1).

Let  $\mathbf{Y} = \mathbf{U}^{-1} \mathbf{X}$  and  $\mathbf{Z} = \mathbf{W}^{-1} \mathbf{X}$ , then  $\mathbf{Z} = [\mathbf{R}_i]^{-1} \mathbf{Y}$ . By computing  $[\mathbf{R}_i]^{-1}$ , one can easily verify that:  $\forall 1 \leq k \leq K$ ,  $z_i(k) = y_i(k)$  and  $z_j(k) = y_j(k) - r_{ji} y_i(k)$  for  $j \neq i$ .  $\text{Span}^+(\mathbf{W})$  contains the scatter plot of the mixed data, if  $\mathbf{Z} \geq 0$ , it is therefore enough to choose:

$$0 \leq r_{ji} \leq \min_{1 \leq k \leq K} \frac{y_j(k)}{y_i(k)}, y_i(k) \neq 0, \text{ for } j \neq i$$

In the algorithm, we set

$$r_{ji} = \frac{1}{2} \min_{1 \leq k \leq K} \frac{y_j(k)}{y_i(k)}, y_i(k) \neq 0 \quad (23)$$

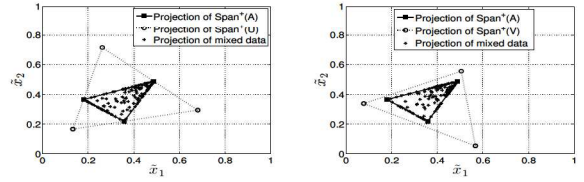
■

By starting from a proper initial simplicial cone containing the scatter plot of the mixed data (see Section IV-A1), say  $\text{Span}^+(\mathbf{U})$ , the proposed algorithm iteratively decreases  $\mathcal{A}(\mathbf{U})$  by performing several sweeps of  $n$  multiplications to the right of  $\mathbf{U}$  by  $\mathbf{R}_i$ ,  $i$  varying from 1 to  $n$ . At each iteration, the matrix  $\mathbf{R}_i$  is computed as described in Proposition 5, in order to decrease the aperture of the current simplicial cone, while keeping all the mixed data inside of the new simplicial cone. The algorithm stops when the current simplicial cone fits the scatter plot of the mixed data. In this case one cannot decrease anymore the aperture of the current simplicial cone, by the matrices  $\mathbf{R}_i$ , while keeping all the mixed data inside of the new cone.

### 3) Locking before convergence and proposed unlocking:

If  $\text{Span}^+(\mathbf{U})$  is the current estimated simplicial cone, the algorithm described above stops, if for any  $i$  between 1 and  $n$ , one cannot decrease anymore the aperture of the current simplicial cone, by the matrix  $\mathbf{R}_i$ , while keeping all the mixed data inside of the new cone. In this case the only matrix  $\mathbf{R}_i$  verifying, for  $\mathbf{W} = \mathbf{U} \mathbf{R}_i$ ,  $\mathcal{A}(\mathbf{W}) \leq \mathcal{A}(\mathbf{U})$  and  $\mathbf{W}^{-1} \mathbf{X} \geq 0$  is  $\mathbf{R}_i = \mathbf{I}_n$ . This situation often corresponds to the convergence of the current estimated mixing matrix  $\mathbf{U}$  to the true mixing matrix  $\mathbf{A}$ . However, this situation may also occur while  $\mathbf{U}$  has not converged yet to  $\mathbf{A}$ . The latter case, is called “locking before convergence”, and is illustrated on Fig.3(a) for  $n = 3$ , the data and the current simplicial cone are projected on the plane  $\mathcal{P}$ , as described in section II-A.

In case of locking before convergence, the proposed algorithm finds  $\mathbf{R}_i = \mathbf{I}_n$ ,  $\forall 1 \leq i \leq n$ , and therefore  $r_{ji} = \delta_{ji}$ . According to equation (23), it follows that,  $\forall 1 \leq i \leq n$ ,  $\exists 1 \leq k_i \leq K$  such as  $y_j(k_i) = 0$  and  $y_i(k_i) \neq 0$ . Since a locking before convergence causes zeros values in the current estimated sources  $\mathbf{Y} = \mathbf{U}^{-1} \mathbf{X}$ , the idea, for overcome this situation, is to modify the current simplicial cone  $\text{Span}^+(\mathbf{U})$



(a) Locking before convergence (b) Unlocking process

Fig. 3. Scatter plot of mixed data included in  $\text{Span}^+(\mathbf{A})$

to make the current estimated sources strictly positive, without increasing  $\mathcal{A}(\mathbf{U})$ . For this purpose we look for an “unlocking matrix”,  $\mathbf{Q}$ , that will rotate the current simplicial cone, such as:  $\mathbf{T} = \mathbf{U}^{-1} \mathbf{Q} \mathbf{X} > 0$ , and for  $\mathbf{V} = \mathbf{Q}^{-1} \mathbf{U}$   $\mathcal{A}(\mathbf{V}) = \mathcal{A}(\mathbf{U})$ . We introduce the criterion  $J$ , by equation (24), where  $t_{ik}$  is the  $ik$ -th entry of  $\mathbf{T}$ .

$$J(\mathbf{Q}) = \sum_{i=1}^n \sum_{k=1}^K t_{ik}^+ \text{ where } t_{ik}^+ = \begin{cases} 1 & \text{if } t_{ik} = 0 \\ 0 & \text{if } t_{ik} > 0 \\ +\infty & \text{if } t_{ik} < 0 \end{cases} \quad (24)$$

One may note that  $J(\mathbf{Q}) = 0$  if and only if  $\mathbf{T} > 0$ , and  $\mathcal{A}(\mathbf{Q}^{-1} \mathbf{U}) = \mathcal{A}(\mathbf{U})$  if  $\mathbf{Q}$  is an orthogonal matrix. A convenient unlocking matrix  $\mathbf{Q}$  can be computed by solving the optimization problem (25):

$$\mathbf{Q} = \arg \min_{\mathbf{O}^T \mathbf{O} = \mathbf{I}_n} J(\mathbf{O}) \quad (25)$$

To deal with problem (25) in one step by gradient like method, we slightly modify the original problem by :

- Regularizing the criterion  $J$  by  $J_\epsilon$  to avoid Dirac distributions when computing the gradient.

$$J_\epsilon(\mathbf{Q}) = \sum_{i=1}^n \sum_{k=1}^K \exp\left(-\frac{t_{ik}}{\epsilon}\right), \sigma > 0$$

$$J_\epsilon(\mathbf{Q}) \rightarrow J(\mathbf{Q}) \text{ when } \epsilon \rightarrow 0.$$

- Adding a penalty term  $J_{orth}(\mathbf{Q}) = \left\| \mathbf{Q}^T \mathbf{Q} - \mathbf{I}_n \right\|_F^2$  to  $J_\epsilon(\mathbf{Q})$ , to penalize its deviation to orthogonality.

The optimization problem becomes :

$$\mathbf{Q} = \arg \min_{\mathbf{O}} J_\sigma(\mathbf{O}) + \gamma J_{orth}(\mathbf{O}), \text{ with } \gamma \geq 0 \quad (26)$$

and can be solved by the iterative gradient algorithm (27):

$$\mathbf{Q}_{(p+1)} = \mathbf{Q}_p^{-\mu} \left[ -\frac{(\mathbf{U}^{-1})^T \mathbf{T}^{null} \mathbf{T}^T}{\sigma} + 4\gamma \mathbf{Q}_p \left( \mathbf{Q}_p^T \mathbf{Q}_p - \mathbf{I}_n \right) \right] \quad (27)$$

where  $\mathbf{T}_{ij}^{null} = \exp\left(-\frac{t_{ij}}{\sigma}\right)$ , and  $p$  denotes the iteration.

The unlocking process is illustrated on Fig.3(b), where  $\text{Span}^+(\mathbf{V})$ , is obtained by rotating  $\text{Span}^+(\mathbf{U})$ , with the unlocking matrix  $\mathbf{Q}$ .

4) Framework of the SCSA-UNS algorithm for noiseless determined mixture: The pseudo-code of the proposed SCSA-UNS algorithm in noiseless case is given by Algorithm 1.

**Algorithm 1** : SCSA-UNS

---

**Require:** Mixed data  $\mathbf{X}_{n \times K}$

**Ensure:** Estimated mixing matrix  $\hat{\mathbf{A}}_{n \times n}$ , and sources  $\hat{\mathbf{S}}_{n \times K}$

- 1: Initialization: Find a proper  $\mathbf{U}$  by extended VCA, as described in section IV-A1, and set  $\mathbf{Y} = \mathbf{U}^{-1}\mathbf{X}$
- 2: **repeat**
- 3:   **for**  $i = 1 \rightarrow n$  **do**
- 4:     Compute  $\mathbf{R}_i$  as described in Section IV-A2
- 5:      $\mathbf{U} \leftarrow \mathbf{U}\mathbf{R}_i$  and  $\mathbf{Y} \leftarrow [\mathbf{R}_i]^{-1}\mathbf{Y}$
- 6:   **end for**
- 7:   **if** a locking situation occurs **then**
- 8:     Compute  $\mathbf{Q}$  as described in Section IV-A3
- 9:      $\mathbf{U} \leftarrow \mathbf{Q}\mathbf{U}$  and  $\mathbf{Y} \leftarrow \mathbf{U}^{-1}\mathbf{Q}^{-1}\mathbf{X}$
- 10:   **end if**
- 11: **until**  $\mathbf{Q} = \mathbf{I}_n$
- 12:  $\hat{\mathbf{A}} \leftarrow \max(\mathbf{U}, 0)$
- 13:  $\hat{\mathbf{S}} \leftarrow \max(\mathbf{Y}, 0)$

---

5) *About the convergence of SCSA-UNS:*

The proposed algorithm, SCSA-UNS, is designed to monotonically decrease the criterion (i.e the aperture of the current simplicial cone). While the multiplications by the matrices  $\mathbf{R}_i$  do not increase the criterion, the unlockings may slightly increase the criterion, since the matrices  $\mathbf{Q}$  are not perfectly orthogonal (being computed by regularized gradient). Although, the monotonic decrease of the criterion is not guaranteed, we however noticed through simulations that the proposed algorithm does not diverge.

B. *Overdetermined mixture ( $m > n$ )*

In case of overdetermined mixture ( $m > n$ ), one must first perform a dimension reduction before running the SCSA-UNS algorithm. By computing the Singular Value Decomposition of  $\mathbf{X}$ , we get:

$$\mathbf{X}_{m \times K} \approx \mathbf{E}_{m \times n} \mathbf{F}_{n \times n} \mathbf{G}_{n \times K}^T \quad (28)$$

where  $\mathbf{F}$  is the diagonal matrix of the  $n$ -largest singular values,  $\mathbf{E}$  (respectively  $\mathbf{G}$ ) is the matrix of the corresponding left (respectively right) singular vectors. The second step consists in running the SCSA-UNS algorithm on the reduced data  $\mathbf{G}^T$ , to compute the MASC containing its scatter plot, say:  $\mathbf{G}^T = \mathbf{U}\mathbf{Y}$ . The mixing matrix and the source are then estimated by:

$$\hat{\mathbf{A}} = \max(\mathbf{E}\mathbf{F}\mathbf{U}, 0) \text{ and } \hat{\mathbf{S}} = \max(\mathbf{Y}, 0) \quad (29)$$

## V. NOISE REDUCTION

This Section considers the case where the observations are corrupted by additive noise. The mixed data are then obtained by the equation (30):

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{B} \quad (30)$$

Due to the additive noise,  $\mathbf{B}$ , the scatter plot of mixed data may expand outside the simplicial cone generated by the mixing

matrix. So running SCSA-UNS on the noisy data, without any pre-processing, will lead to bad a estimation of the mixing matrix (and the sources), as illustrated in Fig. 4, for  $m = n = 3$  sources. A pre-processing is necessary to reduce the noise effect, before estimating the mixing matrix.

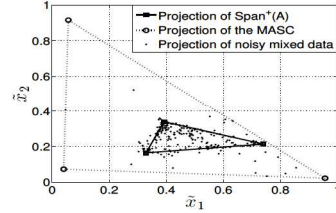


Fig. 4. Projection on  $\mathcal{P}$  of the scatter plot of the noisy mixed data,  $\text{Span}^+(\mathbf{A})$  and the MASC estimated by SCSA-UNS

Assuming the noise is independent and identically distributed, one can note that the noiseless mixed data points are not affected in the same way. In fact, a noisy mixed data point is given by  $\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{b}(k)$ , and let's define the "local signal to noise ratio" by  $\text{snr}(k) = -10 \log_{10} \frac{\|\mathbf{A}\mathbf{s}(k)\|_2}{\|\mathbf{b}(k)\|_2}$ . The small norms noiseless mixed data will be strongly affected by the noise, since their local signal to noise ratio will be very low. To reduce the noise effect, we propose to select and discard all the mixed data points whose norm is lower than a certain threshold  $\lambda_b$ , before running the SCSA-UNS algorithm. Finding a proper threshold  $\lambda_b$  is not trivial, and based on experimental result, we propose  $\lambda_b = \max_{1 \leq k \leq K} \|\mathbf{x}(k)\|_2$ . From the remaining mixed data, denoted  $\hat{\mathbf{X}}_r$  (where  $L$  is the number of remaining data, after removing the small norms data), one can estimate the mixing matrix,  $\hat{\mathbf{A}}_{m \times n}$ , using the SCSA-UNS algorithm for overdetermined mixture (see section IV-B). The sources are then estimated by:

$$\hat{\mathbf{S}} = \max\left(\hat{\mathbf{A}}^\dagger \mathbf{E} \mathbf{F} \mathbf{G}^T, 0\right) \quad (31)$$

where  $\hat{\mathbf{A}}^\dagger$  is the pseudoinverse of  $\hat{\mathbf{A}}$ ,  $\mathbf{F}$  is the diagonal matrix of the  $n$ -largest singular values of  $\mathbf{X}$ , and  $\mathbf{E}$  (respectively  $\mathbf{G}$ ) is the matrix of the corresponding left (respectively right) singular vectors.

## VI. EVALUATION ON SYNTHETIC DATA

In this section, the proposed method is evaluated on synthetic data, and compared to other Non-negative Blind Source Separation methods. The different methods are evaluated for both, mutually independent and mutually correlated sources.

A. *Performance indices and algorithms for comparison*

We consider two performance indices for comparison. The first one is the *mixing matrix estimation error*, also called separation error  $E_{sep}$ . Given the estimated mixing matrix  $\hat{\mathbf{A}}$  and its pseudoinverse  $\hat{\mathbf{A}}^\dagger$ ,  $E_{sep}$  is defined by equation (32) [1] where  $\Delta = \hat{\mathbf{A}}^\dagger \mathbf{A}$ .

$$E_{sep} = \frac{1}{2n(n-1)} \sum_{i=1}^n \left[ \sum_{j=1}^n \left( \frac{|\Delta_{ij}|}{\max_l |\Delta_{il}|} + \frac{|\Delta_{ji}|}{\max_l |\Delta_{li}|} \right) - 2 \right] \quad (32)$$

The smaller is  $E_{sep}$ , the better is the separation, and  $E_{sep}$  is zero for perfect estimation of the mixing matrix. We'll consider that the mixing matrix is properly estimated if  $E_{sep}(dB) < -10dB$ .

The second performance index is the *source estimation error*,  $E_{corr}$ , defined by equation (33):

$$E_{corr} = -\log_{10} \left[ \frac{1}{n} \sum_{i=1}^n \max_{1 \leq j \leq n} \frac{|\hat{\mathbf{s}}_i \mathbf{s}_j^T|}{\|\hat{\mathbf{s}}_i\| \|\mathbf{s}_j\|} \right] \quad (33)$$

where  $\hat{\mathbf{s}}_i$  is the estimate of the  $i$ th source and  $\mathbf{s}_j$  is the true  $j$ th source.  $E_{corr}$  is inversely proportional to the correlation between the true sources and the estimated ones. The smaller is  $E_{corr}$ , the better is the estimation of the sources, and  $E_{corr}$  is zero for perfect separation. We'll consider that the sources are properly estimated if  $E_{corr}(dB) < -20dB$ .

The mixed data were generated according to equation (30). The noise matrix,  $\mathbf{B}$ , entries are generated following the standard Normal distribution, and the Signal to Noise Ratio of the  $i$ -th observation,  $SNR_i$ , is calculated by:

$$SNR_i = 10 \log_{10} \left( \frac{\sum_k \left[ \sum_{j=1}^n a_{ij} s_j(k) \right]^2}{\sum_{k=1}^K [b_i(k)]^2} \right) \quad (34)$$

In all the simulations, the SNR are set to the same value for all the observations, and to keep the non-negativity of the noisy mixed data, we set the negative values to zero.

The proposed method, SCSA-UNS, is compared to six other N-BSS methods, namely SISAL [3], nLCA-IVM [48], VCA [32], N-ICA [38], MVC-NMF [22], and NMF [26]. For each method, the data are pre-normalized to satisfy its required conditions. For nLCA-IVM [48], each row of the data matrix is normalized to unit sum, and for SISAL [3] and VCA [32], each column of the data matrix is normalized to unit sum.

In all simulations, we set the SCSA-UNS parameters to  $\mu = 10^{-2}$ ,  $\epsilon = 10^{-3}$ ,  $\gamma = \frac{1}{4}$ . The parameters of the other methods follow their original paper.

### B. Independent sources

In the simulation, the sources are generated with different sparsity degrees,  $\tau$ , (i.e number of non-zeros elements). Each non-zeros entry of the sources matrix is generated following the uniform distribution between 0 and 1. Three sparsity degrees are considered,  $\tau = 90\%$  (i.e 10% of non-zeros entries),  $\tau = 50\%$  (i.e 50% of non-zeros entries) and  $\tau \rightarrow 0\%$  (i.e nearly 100% of non-zeros entries). The mixing matrix,  $\mathbf{A}$ , has a uniform random entries. We set the number of sources to  $n = 5$ , the number of observations to  $m = 20$ , and the number of samples is  $K = 10000$ .

Fig. 5, Fig. 6, and Fig. 7, depict the variation of the average over 50 independent Monte Carlo runs, of the performance indices versus the  $SNR$ , for the different sparsity degrees. These figures show, for all sparsity degree, the performance indices of the different methods improve when the SNR increases, N-ICA always presents the best performance indices, followed by SCSA-UNS, for  $\tau = 90\%$  and  $\tau = 50\%$ .

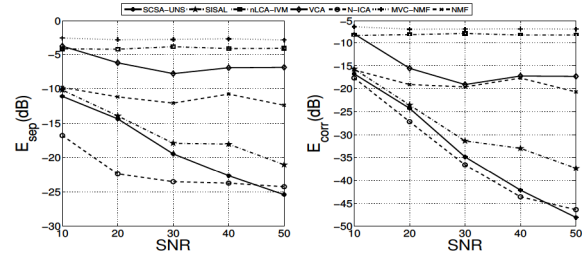


Fig. 5. Average indices vs SNR:  $n = 5$ ,  $m = 20$ ,  $K = 10000$ ,  $\tau = 90\%$

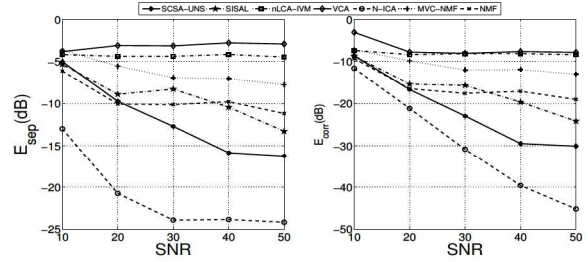


Fig. 6. Average indices vs SNR:  $n = 5$ ,  $m = 20$ ,  $K = 10000$ ,  $\tau = 50\%$

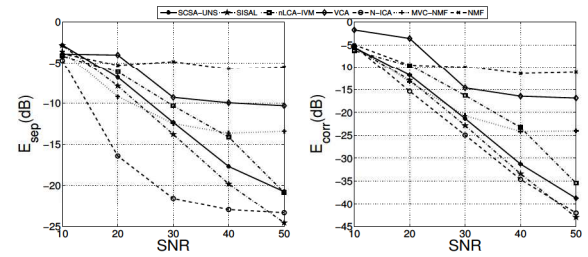


Fig. 7. Average indices vs SNR:  $n = 5$ ,  $m = 20$ ,  $K = 10000$ ,  $\tau \rightarrow 0\%$

### C. Correlated sources

Here, we evaluate the ability of the different methods to unmix mutually correlated sources. We consider four mutually correlated images of size  $350 \times 275$  shown in Fig. 8 [10]. Each image is reshaped to form one row of the source matrix.



Fig. 8. Four correlated images

The estimated empirical correlation matrix defined by equation (35) shows that the four images are highly correlated:

$$[\hat{\mathbf{C}}_s]_{ij} = \frac{\frac{1}{K} \sum_{k=1}^K [s_i(k) - \mu_{s_i}] [s_j(k) - \mu_{s_j}]}{\sigma_{s_i} \sigma_{s_j}} \quad (35)$$

$\mu_{s_l}$  and  $\sigma_{s_l}$  are the mean value and the variance of source  $l$ .

$$\hat{C}_s = \begin{pmatrix} 1.00 & 0.91 & 0.84 & 0.84 \\ 0.91 & 1.00 & 0.84 & 0.84 \\ 0.84 & 0.84 & 1.00 & 0.93 \\ 0.84 & 0.84 & 0.93 & 1.00 \end{pmatrix}$$

The number of sources is  $n = 4$ , the number of samples is  $K = 96250$ . The mixing matrix entries are randomly generated following a uniform distribution between 0 and 1, we set the number of observations to  $m = 20$ . Fig. 9 shows the variation of the average over 50 independent Monte Carlo runs, of the performance indices versus the  $SNR$ . One can

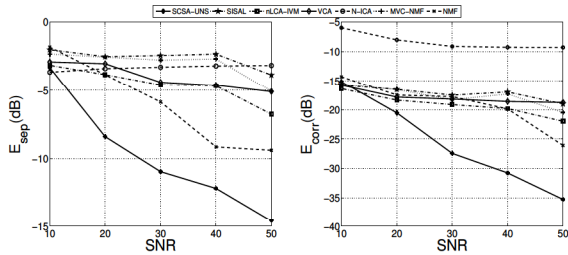


Fig. 9. Average indices vs SNR:  $n = 4$ ,  $m = 20$ ,  $K = 96250$

see on Fig.9 that N-ICA has the worst estimated sources, this result is not surprisingly since this method has been designed for independent sources, and then is not suited to mutually correlated ones. The proposed method, SCSA-UNS, presents the best performance indices.

#### Robustness to overestimation of the number of sources

One of the problem of sources separation in real blind context, is determining the proper number of sources. We compared here the performance of the different methods when the number of sources is overestimated: we assume that there are five sources. The mixing matrix estimation error is calculated by taking the columns of the estimated mixing matrix most correlated to the columns of the actual mixing matrix, and the sources estimation error is calculated by taking the estimated sources most correlated to actual ones. Fig 10 shows the sources estimated by the different method, visually one can see that only SCSA-UNS and nLCA-IVM properly estimate the four sources. SISAL fails due to the unit sum constraint on the estimated sources, and N-ICA cannot separate correlated sources. The performance indices recorded in Table I show that SCSA-UNS presents the best results.

|          | Noiseless case |                |
|----------|----------------|----------------|
|          | $E_{sep}(dB)$  | $E_{corr}(dB)$ |
| SCSA-UNS | -9.68          | -20.02         |
| SISAL    | -3.23          | -9.33          |
| nLCA-IVM | -4.45          | -17.44         |
| VCA      | -3.11          | -9.12          |
| N-ICA    | -2.61          | -8.26          |
| MVC-NMF  | -3.82          | -9.63          |
| NMF      | -6.91          | -13.26         |

TABLE I  
AVERAGE PERFORMANCE INDICES FOR CORRELATED SOURCES



Fig. 10. Correlated sources separation for  $SNR = 30dB$ : (row 1) original sources, and sources estimated by (row 2) SCSA-UNS, (row 3) SISAL, (row 4), nLCA-IVM, (row 5) VCA, (row 6) N-ICA, (row 7) MVC-NMF, (row 8) NMF

## VII. RESULTS ON ACTUAL DATA

This section presents separation results on two different data sets: images of positron emission tomography and Mass spectra. The indices  $E_{sep}$  and  $E_{corr}$ , used before to characterize the separation results on the synthetic data are no longer usable, because the original mixing matrices and sources are unknown. Results of SCSA-UNS are compared to those obtained by SISAL, N-ICA and MVC-NMF.

### A. Results on Dynamic Positron Emission Tomography images

In oncology, the effectiveness of an anticancer treatment is often achieved using Positron Emission Tomography (PET) images [33] [42]. But an additional measurement of the Arterial Input Function (AIF), which provides the tracer concentration available for organs, is necessary to evaluate quantitatively the tumor activity. The reference method for AIF estimation is the arterial blood sampling which is unfortunately too invasive for routine clinic use. We perform here, the separation of real Dynamic Positron Emission Tomography (PET) images, to study the pharmacokinetics of the [18F]-FDG (FluoroDeoxyGlucose) tracer on human brain. The main objective is to estimate the arterial pharmacokinetic (Arterial Input Function) using only the dynamic TEP images, without blood sampling [21] [46]. An arterial input function is obtained by blood sampling (reference Arterial Input Function rAIF) is



also considered to assess the accuracy of the AIF estimation by sources separation using PET images. The data were provided by the CEA/I2BM/SHFJ.

We have 19 human brain PET images recorded during 30mn. The original images are size  $128 \times 128 \times 63$ , but we only considered subimages of size  $63 \times 73 \times 58$  for only focusing on the voxels with activity. Each 3D PET image is reshaped to form one row of the observations matrix  $\mathbf{X}$ . The number of observations is  $m = 19$  and the number of samples is  $K = 266742$ . According to biological consideration, we set the number of sources to  $n = 3$ .

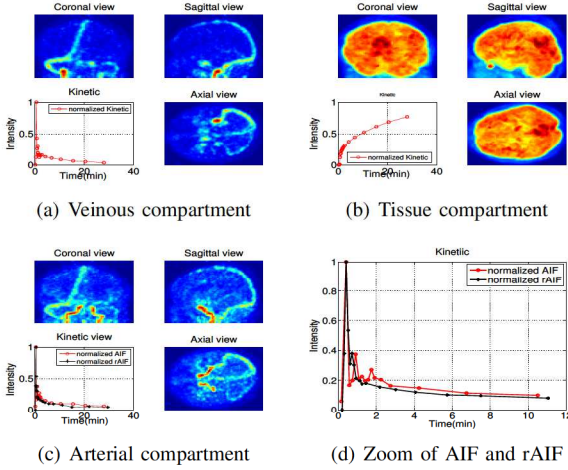


Fig. 11. Pharmacokinetic compartments estimated by SCSA-UNS

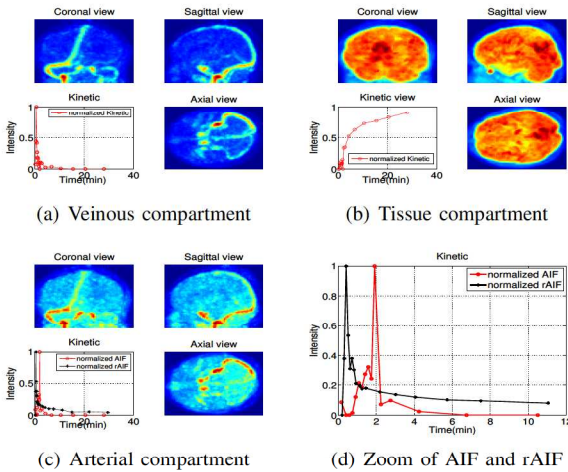


Fig. 12. Pharmacokinetic compartments estimated by SISAL

Figures 11, 12, 13, and 14, shows the pharmacokinetic compartments estimated respectively by SCSA-UNS, SISAL, N-ICA and MVC-NMF. Each subfigure a, b and c, represents, at lower left, the normalized kinetics (columns of the estimated mixing matrix) and the corresponding spatial distributions (rows of the estimated sources) according to the three views, coronal (upper left), sagittal (upper right) and axial (lower right). All the four methods properly estimate the Veinous and the Tissue compartments. SISAL and MVC-NMF fail to recover the Arterial compartment, and one can on Fig 12.d and

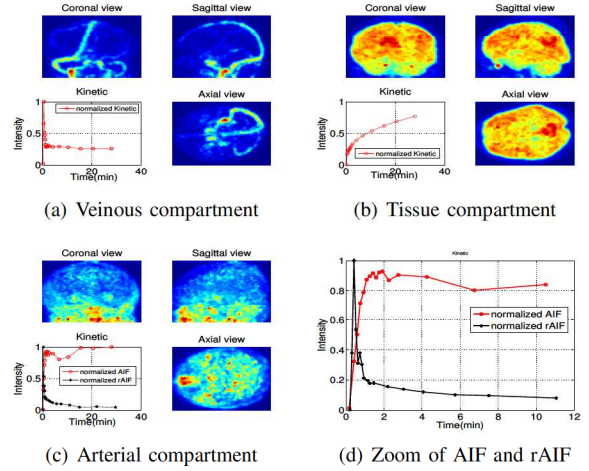


Fig. 13. Pharmacokinetic compartments estimated by N-ICA

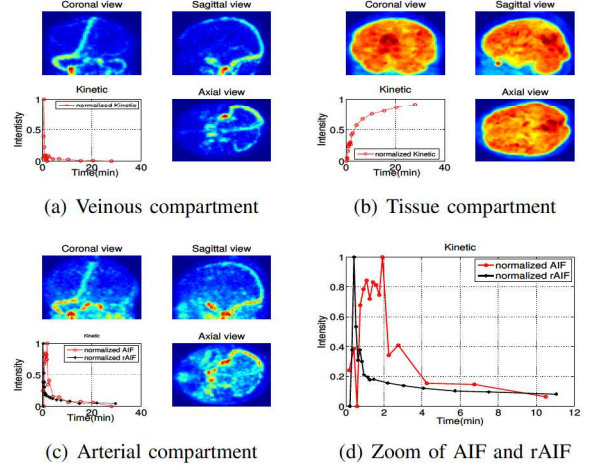


Fig. 14. Pharmacokinetic compartments estimated by MVC-NMF

Fig 14.d that the AIF estimated by these methods are far from the reference AIF, obtained by blood sampling. The third sources estimated by N-ICA suggests a nose compartment. Unlike the previous methods, SCSA-UNS found the cortid arteries (Fig. 11.c) [21], and Fig. 11.d shows that the AIF estimated by SCSA-UNS closely follows the reference AIF.

## B. Results on Mass spectra for Metabolomics

Liquid Chromatography - Mass Spectrometry (LC-MS) technique is used for identifying and quantifying small organic molecules (also called metabolites) of biological/chemical fluids [50]. The sample to analyse is first introduced into a liquid chromatograph which separates its different metabolites according to their physico-chemical properties. The metabolites thus separated, flow out of the chromatograph during different time intervals, called elution profiles. At the output of the chromatograph, a mass spectrometer measures the mass spectrum of the eluted metabolites. A metabolite is characterized by its elution profile and its mass spectrum. Real biological/chemical fluids (urine, blood, saliva) are complex mixtures of tens or even hundreds of metabolites, therefore the liquid chromatograph cannot completely separate the elution

profiles of the different metabolites which overlap. So, the mass spectra measured can be modeled as linear combinations of the mass spectra of the different metabolites co-eluted (i.e. metabolites whose elution profiles overlap). In this experiment, we seek to evaluate the efficiency of SCSA-UNS for separating the elementary elution profiles and mass spectra of different metabolites co-eluted, given the measured mass spectra [49]. The data is provided by CEA/DSV/LEMM.

Eleven commercial chemical compounds were first individually analyzed by the combination of HPLC and LTQ-Orbitrap (mass spectrometer) to build a reference database of their elution profiles and mass spectra. The eleven compounds are mixed and the resulting homogeneous solution is also analyzed by the combination of HPLC and LTQ-Orbitrap. The mass spectra were recorded during 20 min and stored in an observation matrix whose rows correspond to mass spectra measured at different times, and columns correspond to mass indices. The observation matrix is preprocessed by removing all the zeros columns and all the rows whose maximum value is smaller than  $10^4$  (threshold set by our partners biologists). The resulting number of observations is  $m = 1453$ , and the number of samples is  $K = 5638$ .

We performed several runs the SCSA-UNS algorithm by incrementing each time the number of sources, starting from  $n = 11$ . We noticed that all the components are found when  $n \geq 16$ . Figure 15 shows the elution profiles, and corresponding mass spectra estimated by SCSA-UNS. This figure should be compared with the initial components database, give at appendix E. The estimated components do not exactly fit to the components referenced in the database, but the greater part of the peaks are retrieved.

The other methods, namely SISAL, MVC-NMF and N-ICA present similar results, which are not shown here.

### VIII. CONCLUSION

In this paper, we propose a geometrical method for separating non-negative sources. The proposal, denoted SCSA-UNS, estimates the mixing matrix and the sources, by first reducing the dimension of the mixed data, followed by fitting a Minimum Aperture Simplicial Cone (MASC) to the scatter plot of the dimension reduced data. SCSA-UNS does not require the independence of sources, neither their local dominance, but the positive orthant must be the unique MASC containing the scatter plot of the sources, to ensure recovering the true mixing matrix and the true sources. In noisy case, the proposed method starts by discarding the points most corrupted by the noise, which can significantly expand the scatter plot of mixed data, before looking for the MASC containing the data. Simulation on synthetic data have showed that the proposed method performs good separation for both independent and mutually correlated sources. The proposal has also been successfully used to estimate the pharmacokinetic compartments of [18F]-FDG tracer on human brain (in particular to estimated the Arterial Input Function) and to separate the elementary mass spectra of different chemical compounds, from the mass spectra measured at the output of a liquid chromatograph.

Future works include improving the robustness of the proposed method to additive noise, and incorporating multiplica-

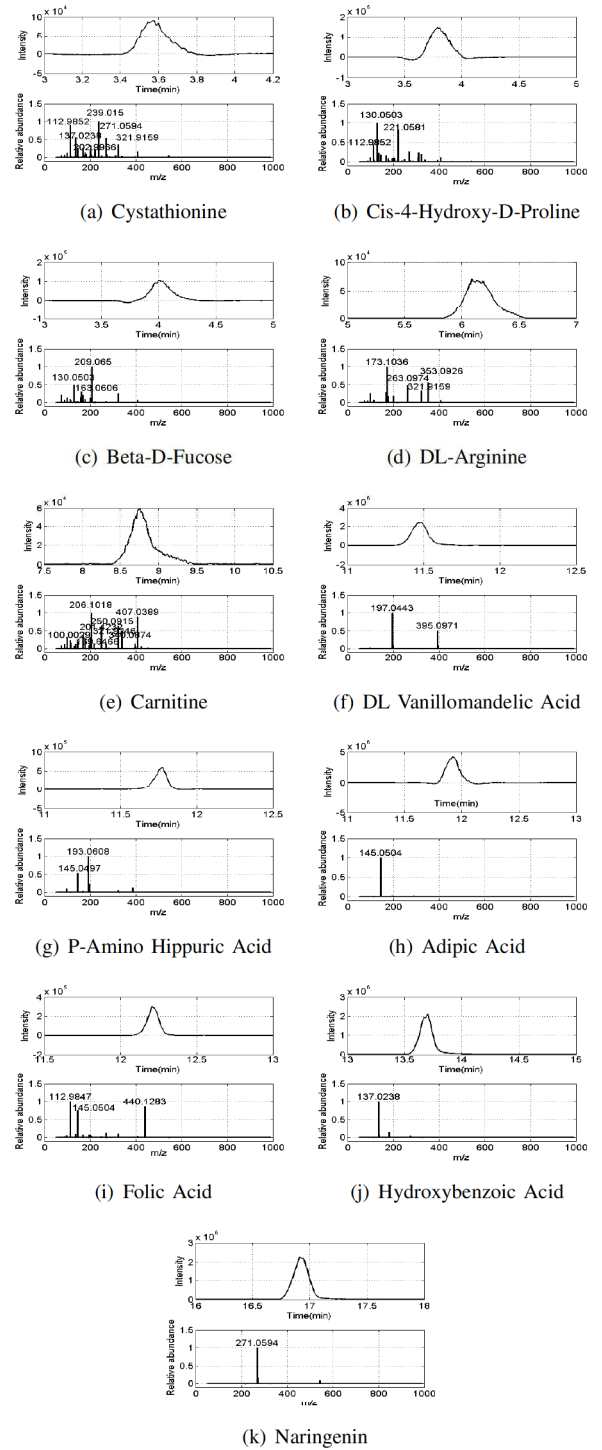


Fig. 15. 16 elution profiles and corresponding mass spectra estimated by SCSA-UNS from HPLC-LTQ Orbitrap data

tive noise. In fact, in certain application the noise seem to also have a multiplicative part [19] [4]. The proof of the necessary and sufficient condition on the sources, under which the simplicial cone generated by the mixing matrix is the unique non-negative MASC containing the scatter plot of the sources will also be investigated. Evaluation on other real data (such as hyperspectral images) will also performed.

## ACKNOWLEDGMENT

The authors would like to thank Régine Trébossen and Bertrand Tavitian of the CEA, I2BM, Service Hospitalier Frédéric Joliot, for the dynamic PET images. We also thank Christophe Junot of the CEA, DSV, Laboratoire d'Étude du Métabolisme des Médicaments (LEMM), for the LC-MS data.

## APPENDIX A

## PROOF OF THEOREM 1

$\mathbf{U} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{V} \in \mathbb{R}_+^{n \times n}$  are full rank matrices, let  $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$  and  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$  where  $\mathbf{u}_i$  (respectively  $\mathbf{v}_i$ ) the  $i$ -th column of  $\mathbf{U}$  (respectively  $\mathbf{V}$ ). If  $\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V})$  then  $\forall 1 \leq i \leq n, \mathbf{u}_i \in \text{Span}^+(\mathbf{V})$ , therefore there is  $\mathbf{m}_i \in \mathbb{R}_+^n$  such as  $\mathbf{u}_i = \mathbf{V}\mathbf{m}_i$ . Let  $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n]$ , then  $\mathbf{M} \in \mathbb{R}_+^{n \times n}$  and  $\mathbf{U} = \mathbf{V}\mathbf{M}$ . Since  $rg(\mathbf{U}) = rg(\mathbf{V}) = n$  then  $rg(\mathbf{M}) = n$ . Reciprocally, assume that there is a square full column rank non-negative matrix  $\mathbf{M} \in \mathbb{R}_+^{n \times n}$  such as  $\mathbf{U} = \mathbf{V}\mathbf{M}$ . For any  $\mathbf{z} \in \text{Span}^+(\mathbf{U})$ , there is  $\mathbf{y} \in \mathbb{R}_+^n$  such as  $\mathbf{z} = \mathbf{U}\mathbf{y}$ . Since  $\mathbf{U} = \mathbf{V}\mathbf{M}$ , then  $\mathbf{z} = \mathbf{V}\mathbf{M}\mathbf{y}$ . It follows that  $\mathbf{z} \in \text{Span}^+(\mathbf{V})$  (because  $\mathbf{M}\mathbf{y} \in \mathbb{R}_+^n$ ), and therefore  $\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V})$ .

## APPENDIX B

## PROOF OF THEOREM 2

Proof of the forward sense: Assume that  $\text{Span}^+(\mathbf{U}) = \text{Span}^+(\mathbf{V})$ , then  $\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V})$ , and there is a square full column rank non-negative matrix  $\mathbf{M} \geq 0$  such as  $\mathbf{U} = \mathbf{V}\mathbf{M}$  (due to Theorem 1), so  $\mathbf{V} = \mathbf{U}\mathbf{M}^{-1}$ . For any  $\mathbf{y} \geq 0$ ,  $\mathbf{V}\mathbf{y} \in \text{Span}^+(\mathbf{V})$ , then  $\mathbf{V}\mathbf{y} \in \text{Span}^+(\mathbf{U})$  (since  $\text{Span}^+(\mathbf{V}) = \text{Span}^+(\mathbf{U})$ ). Furthermore  $\mathbf{V}\mathbf{y} = \mathbf{U}\mathbf{M}^{-1}\mathbf{y} \Rightarrow \mathbf{U}\mathbf{M}^{-1}\mathbf{y} \in \text{Span}^+(\mathbf{U})$ , then  $\mathbf{M}^{-1}\mathbf{y} \geq 0$ , and therefore  $\mathbf{M}^{-1} \geq 0$ . Since  $\mathbf{M} \geq 0$  and  $\mathbf{M}^{-1} \geq 0$ , one can conclude that  $\mathbf{M}$  is a monomial matrix [35].

Proof of the reverse sense: Assume that  $\mathbf{U} = \mathbf{V}\mathbf{M}$  where  $\mathbf{M}$  is a monomial matrix. For any  $\mathbf{z} \in \text{Span}^+(\mathbf{U})$ , there is  $\mathbf{y} \geq 0$  such as  $\mathbf{z} = \mathbf{U}\mathbf{y}$ . Then  $\mathbf{z} = \mathbf{V}\mathbf{M}\mathbf{y}$ , since  $\mathbf{M}$  is a monomial matrix then  $\mathbf{M}\mathbf{y} \geq 0$ , therefore  $\mathbf{z} \in \text{Span}^+(\mathbf{V})$ , and one can conclude that  $\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V})$ . In the same way, and using the fact that  $\mathbf{M}$  is a monomial matrix, one can easily show that  $\text{Span}^+(\mathbf{V}) \subseteq \text{Span}^+(\mathbf{U})$ , which lead us to conclude that  $\text{Span}^+(\mathbf{U}) = \text{Span}^+(\mathbf{V})$ .

## APPENDIX C

## PROOF OF THEOREM 3

Let  $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n]$  and  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n]$  with  $\mathbf{g}_i = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|_2}$  and  $\mathbf{h}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|_2}$  then  $\text{Span}^+(\mathbf{G}) = \text{Span}^+(\mathbf{U})$  and  $\text{Span}^+(\mathbf{H}) = \text{Span}^+(\mathbf{V})$  (Theorem 2). It is enough to prove that  $\mathcal{A}(\mathbf{G}) \leq \mathcal{A}(\mathbf{H})$ .

$\text{Span}^+(\mathbf{U}) \subseteq \text{Span}^+(\mathbf{V}) \Rightarrow \text{Span}^+(\mathbf{G}) \subseteq \text{Span}^+(\mathbf{H})$  there is a full column rank non-negative matrix  $\mathbf{M}$  such as  $\mathbf{G} = \mathbf{H}\mathbf{M}$  (Theorem 1). Let  $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_n]$ , where  $\mathbf{m}_i$  is the  $i$ th column of  $\mathbf{M}$ . For fixed  $i$ :  $g_{li} = \sum_{r=1}^n h_{lr} m_{ri}$ , then

$$g_{li}^2 = \left( \sum_{r=1}^n h_{lr} m_{ri} \right)^2 \geq \sum_{r=1}^n h_{lr}^2 m_{ri}^2$$

$$\text{thus } \sum_{l=1}^n g_{li}^2 \geq \sum_{l=1}^n \left( \sum_{r=1}^n h_{lr}^2 m_{ri}^2 \right) = \sum_{r=1}^n m_{ri}^2 \left( \sum_{l=1}^n h_{lr}^2 \right).$$

Since  $\sum_{l=1}^n g_{li}^2 = 1$  and  $\sum_{l=1}^n h_{li}^2 = 1$ , then  $1 \geq \sum_{r=1}^n m_{ri}^2$ , so  $1 \geq \|\mathbf{m}_i\|_2$ . It follows that:

$$1 \geq \prod_{i=1}^n \|\mathbf{m}_i\|_2 \Rightarrow 1 \leq \frac{1}{\|\mathbf{m}_1\| \times \|\mathbf{m}_2\| \times \dots \times \|\mathbf{m}_n\|}$$

Furthermore, since  $|\det(\mathbf{G})| = |\det(\mathbf{H})| |\det(\mathbf{M})|$  then

$$|\det(\mathbf{G})| \leq |\det(\mathbf{H})| \frac{|\det(\mathbf{M})|}{\|\mathbf{m}_1\|_2 \times \|\mathbf{m}_2\|_2 \times \dots \times \|\mathbf{m}_n\|_2}$$

On the other hand,  $\frac{|\det(\mathbf{M})|}{\|\mathbf{m}_1\|_2 \times \|\mathbf{m}_2\|_2 \times \dots \times \|\mathbf{m}_n\|_2} \leq 1$  due to Hadamard's inequality, then  $|\det(\mathbf{G})| \leq |\det(\mathbf{H})|$ . Moreover, since  $\prod_{i=1}^n \|\mathbf{g}_i\|_2 = 1$  and  $\prod_{i=1}^n \|\mathbf{h}_i\|_2 = 1$ , then  $\mathcal{A}(\mathbf{G}) \leq \mathcal{A}(\mathbf{H})$ .

## APPENDIX D

## PROOF OF THEOREM 5

We proceed by contradiction.

Proof of the forward sense: Suppose that  $\text{Span}^+(\mathbf{I}_n)$  is not the unique non-negative MASC containing the scatter plot of sources, there is thus a non-negative matrix  $\mathbf{U}$  such as  $\mathbf{U}^{-1}\mathbf{S} \geq 0$  and  $\mathcal{A}(\mathbf{U}) \leq \mathcal{A}(\mathbf{I}_n)$ .

$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} \Rightarrow \mathbf{U}^{-1}\mathbf{A}^{-1}\mathbf{X} = (\mathbf{AU})^{-1}\mathbf{X} \geq 0$ , therefore  $\text{Span}^+(\mathbf{AU})$  is another simplicial cone containing the scatter plot of the mixed data. Since  $\mathbf{A} \geq 0$  and  $\mathbf{U} \geq 0$  then  $\text{Span}^+(\mathbf{AU}) \subseteq \text{Span}^+(\mathbf{I}_n)$ . Furthermore  $\mathcal{A}(\mathbf{AU}) \leq \mathcal{A}(\mathbf{A})$  (because  $\mathbf{U}$  is non-negative and due to Theorems 1 and 3), then  $\text{Span}^+(\mathbf{A})$  is not the unique non-negative MASC containing the scatter plot of the mixed data.

Proof of the reverse sense: Assume that  $\text{Span}^+(\mathbf{A})$  is not the unique non-negative MASC containing the scatter plot of the mixed data, there is thus another non-negative MASC, say  $\text{Span}^+(\mathbf{W})$ , containing the scatter plot of the mixed data (i.e  $\mathbf{X} = \mathbf{W}\mathbf{Y}$ , where  $\mathbf{Y} \geq 0$ ), and  $\mathcal{A}(\mathbf{W}) \leq \mathcal{A}(\mathbf{A})$ .

$\mathbf{S} = \mathbf{A}^{-1}\mathbf{X} \Rightarrow \mathbf{S} = \mathbf{A}^{-1}\mathbf{W}\mathbf{Y}$  with  $\mathbf{Y} \geq 0$ , then  $\text{Span}^+(\mathbf{A}^{-1}\mathbf{W})$  is another simplicial cone containing the scatter plot of the sources.  $\forall \mathbf{y} \geq 0$ ,  $\mathbf{W}\mathbf{y}$  is in the scatter plot of the mixed data, then  $\mathbf{A}^{-1}\mathbf{W}\mathbf{y} \geq 0$ , and it follow that  $\mathbf{A}^{-1}\mathbf{W} \geq 0$ . Since  $\mathcal{A}(\mathbf{A}^{-1}\mathbf{W}) \leq 1$  (due to Hadamard's inequality), and  $\mathcal{A}(\mathbf{I}_n) = 1$ , then  $\mathcal{A}(\mathbf{A}^{-1}\mathbf{W}) \leq \mathcal{A}(\mathbf{I}_n)$ , and one can conclude that  $\text{Span}^+(\mathbf{I}_n)$  is not the unique non-negative MASC containing the scatter plot of the sources.

## APPENDIX E

## DATABASE OF THE COMMERCIAL CHIMICAL COMPOUNDS

## REFERENCES

- [1] S. Amari, A. Cichocki, and H. Yang, "A New Learning Algorithm for Blind Signal Separation," in *Advances in Neural Information Processing Systems*. MIT Press, 1996, pp. 757–763.
- [2] M. Babaie-Zadeh, A. Mansour, C. Jutten, and F. Marvasti, "A Geometric Approach for Separating Several Speech Signals," in *Independent Component Analysis and Blind Signal Separation*, ser. Lecture Notes in Computer Science, 2004, vol. 3195, pp. 798–806.
- [3] J. Bioucas-Dias, "A Variable Splitting Augmented Lagrangian Approach to Linear Spectral Unmixing," in *First IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009, pp. 1–4.

| Chemical Compound       | Times (min) | ID                 | WVZ        |
|-------------------------|-------------|--------------------|------------|
| Cystathionine           | 3.75        | C4 H8 O2 N2 S      | 134.02612  |
|                         |             | C7 H13 O4 N2 S     | 221.059    |
| Cis-4-Hydroxy-D-Proline | 3.83        | C9 H8 N O3         | 130.05097  |
|                         |             | C6 H8 N Na2 O5     | 220.020338 |
| Beta-D-Fucose           | 4.01        | C3 H5 O3           | 89.02442   |
|                         |             | C6 H11 O5          | 163.0612   |
|                         |             | C7 H13 O7          | 209.06668  |
| DL-Arginine             | 6.23        | C8 H14 O N2        | 131.0826   |
|                         |             | C6 H13 O2 N4       | 173.1044   |
| Carnitine               | 8.87        | C6 H13 O2 N3       | 174.10143  |
|                         |             | C10 H18 C N Na3 O9 | 400.03687  |
| DL Vanillic Acid        | 11.48       | C9 H9 O5           | 197.04553  |
|                         |             | C18 H19 O10        | 385.09637  |
| P-Amino Hippuric Acid   | 11.8        | C8 H9 O N2         | 145.07204  |
|                         |             | C9 H9 O N2         | 153.06187  |
|                         |             | C15 H17 O6 N3 Na3  | 467.07414  |
| Adipic Acid             | 11.92       | C11 H20 O2         | 128.04342  |
|                         |             | C6 H9 O4           | 145.05063  |
| Folic Acid              | 12.22       | C13 H16 N O4       | 246.03309  |
|                         |             | C9 H6 O N3         | 128.03532  |
|                         |             | C14 H16 N O3       | 254.04127  |
| Hydroxybenzoic Acid     | 13.69       | C14 H16 N O4       | 254.04127  |
|                         |             | C19 H18 O N7       | 440.1324   |
| Naringenin              | 16.92       | C7 H5 O3           | 137.02442  |
|                         |             | C7 H5 C N Na O3    | 194.98304  |
|                         |             | C15 H11 O5         | 271.0612   |
|                         |             | C30 H23 O10        | 543.12967  |

Fig. 16. Database of the eleven commercial chemical compounds

- [4] F. P. Breitwieser, A. Müller, L. Dayon, T. Köcher, A. Hainard, P. Pichler, U. Schmidt-Erfurth, G. Superti-Furga, J.-C. Sanchez, K. Mechtler, and n. J. C. K. L. Bennett, "General Statistical Modeling of Data from Protein Relative Expression Isobaric Tags," *Journal of Proteome Research*, vol. 10, no. 6, pp. 2758–2766, 2011.
- [5] R. H. C. Gobinet, E. Perrin, "Application of Non-negative Matrix Factorization to Fluorescence Spectroscopy," in *12th European Signal Processing Conference*, 2004, pp. 1095–1098.
- [6] T. H. Chan, C. Y. Chi, Y. M. Huang, and W. K. Ma, "A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing," *IEEE Transactions on Signal Processing*, vol. 57, pp. 4418–4432, 2009.
- [7] T. H. Chan, W. K. Ma, C. Y. Chi, and Y. Wang, "A Convex Analysis Framework for Blind Separation of Non-Negative Sources," *IEEE Transactions on Signal Processing*, vol. 56, pp. 5120–5134, 2008.
- [8] T.-H. Chan, "Convex Analysis Based Non-negative Blind Source Separation for Biomedical and Hyperspectral Image Analysis," Ph.D. dissertation, Institute of Communications Engineering, NTHU, 2009.
- [9] Z. Chen and A. Cichocki, "Nonnegative matrix factorization with temporal smoothness and/or spatial decorrelation constraints," in *Laboratory for Advanced Brain Signal Processing, RIKEN, Tech. Rep.*, 2005.
- [10] A. Cichocki and R. Zdunek. NMFLAB MATLAB Toolbox for Non-Negative Matrix Factorization. [Online]. Available: <http://www.bsp.brain.riken.jp/ICALAB/nmflab.html>
- [11] A. Cichocki, R. Zdunek, A. H. Phan, and S. Amari, "8- Selected Applications," in *Nonnegative Matrix and Tensor Factorizations Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. John Wiley and Sons, Ltd, 2009.
- [12] —, *Nonnegative Matrix and Tensor Factorizations, Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*, Wiley, Ed. John Wiley and Sons, Ltd, 2009.
- [13] M. D. Craig, "Minimum-Volume Transforms for Remotely Sensed Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 542–552, 1994.
- [14] D. Donoho and V. Stodden, "When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts?" in *Advances in Neural Information Processing Systems 16*. MIT Press, 2004.
- [15] J.-L. Durrieu and J.-P. Thiran, "Sparse Non-negative Decomposition of Speech Power Spectra for Formant Tracking," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 5260 – 5263.
- [16] N. Gillis, "Sparse and Unique Nonnegative Matrix Factorization Through Data Preprocessing," *Journal of Machine Learning Research*, vol. 13, pp. 3349–3386, 2012.
- [17] P. O. Hoyer, "Non-negative Matrix Factorization with Sparseness Constraints," *Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, 2004.
- [18] A. Huck and M. Guillaume, "Robust Hyperspectral Data Unmixing with Spatial and Spectral Regularized NMF," in *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2010, pp. 1–4.
- [19] C. Hundertmark, R. Fischer, T. Reinl, S. May, F. Klawonn, and L. Jansch, "MS-specific noise model reveals the potential of iTRAQ in quantitative proteomics," *Bioinformatics*, vol. 25, no. 8, pp. 1004–1011, 2009.
- [20] M. M. Ichir and A. Mohammad-Djafari, "Bayesian Blind Source Separation of Positive Non Stationary Sources," in *American Institute of Physics Conference Series*, vol. 735, 2004, pp. 493–500.
- [21] K. Kim, H. W. M., Shidahara, J. Ahn, S. Choi, N. Kudomi, K. Hayashida, Y. Miyake, and H. Iida, "Noninvasive estimation of cerebral blood flow using image-derived carotid input function in H215O dynamic PET," in *IEEE Nuclear Science Symposium Conference Record*, vol. 3, 2001, pp. 1282–1285.
- [22] H. Q. L. Miao, "Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765–777, 2007.
- [23] H. Laurberg, M. G. Christensen, M. D. Plumbley, . K. Hansen, and S. H. Jensen, "Theorems on Positive Data: On the Uniqueness of NMF," *Computational Intelligence and Neuroscience*, vol. 2008, 2008.
- [24] C. Lazar, D. Nuzillard, and A. Nowé, "A New Geometrical BSS Approach for Non Negative Sources," in *Latent Variable Analysis and Signal Separation*, ser. Lecture Notes in Computer Science, 2010, vol. 6365, pp. 530–537.
- [25] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.
- [26] —, "Algorithms for Non-negative Matrix Factorization," *Neural Information Processing Systems*, vol. 13, pp. 556–562, 2001.
- [27] J. S. Lee, D. D. Lee, S. Choi, and D. S. Lee, "Application Of Non-Negative Matrix Factorization To Dynamic Positron Emission Tomography," in *3rd International Conference on Independent Component Analysis and Blind Signal Separation*, 2001, pp. 629–632.
- [28] J. Li and J. M. Bioucas-Dias, "Minimum Volume Simplex Analysis : A Fast Algorithm to Unmix Hyperspectral Data," in *IEEE International Symposium on Geoscience and Remote Sensing Symposium*, vol. 3, 2008, pp. III – 250–III – 253.
- [29] L. Miao and H. Qi, "A Constrained Non-Negative Matrix Factorization Approach to Unmix Highly Mixed Hyperspectral data," in *IEEE International Conference on Image Processing*, 2007, pp. 185–188.
- [30] S. Moussaoui, D. Brie, A. M. Djafari, and C. Carteret, "Separation of Non-Negative Mixture of Non-Negative Sources Using a Bayesian Approach and MCMC Sampling," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4133–4145, 2006.
- [31] S. Moussaoui, D. Brie, and J. Idier, "Non-Negative Source Separation: Range of Admissible Solutions and Conditions for the Uniqueness of the Solution," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2005, pp. 289–292.
- [32] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 898–910, 2005.
- [33] J. Okada, H. Oonishi, K. Yoshikawa, K. Imaseki, K. Uno, J. Itami, and N. Arimizu, "FDG-PET for the evaluation of tumor viability after anticancer therapy," *Annals of Nuclear Medicine*, vol. 8, no. 2, pp. 109–113, 1994.
- [34] P. Paatero and U. Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.
- [35] R. J. Plemmons and R. E. Cline, "The Generalized Inverse of a Nonnegative Matrix," in *American Mathematical Society*, vol. 31, no. 1, 1972, pp. 46–50.
- [36] M. Plumbley, "Conditions for Nonnegative Independent Component Analysis," *IEEE Signal Processing Letters*, vol. 9, pp. 177–180, 2002.
- [37] —, "Algorithms for Nonnegative Independent Component Analysis," *IEEE Transactions on Neural Networks*, vol. 14, no. 3, pp. 534–543, 2003.
- [38] —, "Optimization using Fourier Expansion over a Geodesic for Non-Negative ICA," *Lecture Notes in Computer Science*, vol. 3195, pp. 44–56, 2004.
- [39] —, "Geometrical methods for non-negative ICA: Manifolds, Lie groups and toral subalgebras," *Neurocomputing*, vol. 67, pp. 161–197, 2005.
- [40] M. Plumbley, A. Cichocki, and R. Bro, "Chapter 13 -Non-negative mixtures," in *Handbook of Blind Source Separation: Independent Component Analysis and Applications*, P. Comon and C. Jutten, Eds. Academic Press, 2010, pp. 515–547.
- [41] C. Puntonet, A. Mansour, and C. Jutten, "A Geometrical Algorithm for Blind Source Separation," in *GRETSI*, Juan-les-Pins, September 1995, pp. 273–276.

- [42] P. Rigo, P. Paulus, B. J. Kaschten, R. Hustinx, T. Bury, G. Jerusalem, T. Benoit, and J. Foidart-Willems, "Oncological applications of positron emission tomography with fluorine-18 fluorodeoxyglucose," *European Journal of Nuclear Medicine*, vol. 23, no. 23, pp. 1641–1674, 1996.
- [43] P. Sajda, S. Du, T. Brown, R. Stoyanova, D. Shungu, and L. P. X. Mao, "Nonnegative Matrix Factorization for Rapid Recovery of Constituent Spectra in Magnetic Resonance Chemical Shift Imaging of the Brain," *IEEE Transactions on Medical Imaging*, vol. 23, no. 12, pp. 1453–1465, 2004.
- [44] R. Schachtner, G. Pöppel, and E. Lang, "Towards unique solutions of non-negative matrix factorization problems by a determinant criterion," *Digital Signal Processing*, vol. 21, pp. 528–534, 2011.
- [45] X. Shao, Z. Zhichao, and W. Cai, "Extraction of chemical information from complex analytical signals by a non-negative independent component analysis," *Analyst*, vol. 134, pp. 2095–2099, 2009.
- [46] Y. Su, A. M. Arbelaez, T. Benzinger, A. Z. Snyder, A. G. Vlassenko, M. A. Mintun, and M. E. Raichle, "Noninvasive estimation of the arterial input function in positron emission tomography imaging of cerebral blood flow," *Journal of Cerebral Blood Flow & Metabolism*, vol. 33, pp. 115–121, 2012.
- [47] T. Virtanen, "Monaural Sound Source Separation by Nonnegative Matrix Factorization With Temporal Continuity and Sparseness Criteria," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 3, pp. 1066–1074, 2007.
- [48] F. Y. Wang, C. Y. Chi, T. H. Chan, and Y. Wang, "Nonnegative Least-Correlated Component Analysis for Separation of Dependent Sources by Volume Maximization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 875–888, 2010.
- [49] X.-C. Xiong, X. Fang, Z. Ouyang, Y. Jiang, Z.-J. Huang, and Y.-K. Zhang, "Feature Extraction Approach for Mass Spectrometry Imaging Data Using Non-negative Matrix Factorization," *Chinese Journal of Analytical Chemistry*, vol. 40, no. 5, p. 663–669, 2012.
- [50] B. Zhou, J. Xiao, L. Tuli, and H. Resson, "LC-MS-based metabolomics," *Molecular BioSystems*, vol. 8, no. 2, pp. 470–481, 2012.
- [51] Y. Zhu, T. H. Chan, E. P. Hoffman, and Y. Wang, "Gene Expression Dissection by Non-negative Well-Grouped Source Separation," in *IEEE Workshop on Machine Learning for Signal Processing*, 2008, pp. 255–260.

**Meriem Jaidane****Wendyam Serge Boris Ouedraogo****Christian Jutten****Antoine Souloumiac**



# Chapitre 4

## Étude expérimentale de performance et de robustesse

### Sommaire

---

|            |   |           |
|------------|---|-----------|
| <b>4.1</b> | <b>Introduction</b>                                     | <b>67</b> |
| <b>4.2</b> | <b>Initialisation de l'algorithme SCSA-UNS</b>          | <b>67</b> |
| 4.2.1      | Initialisation par l'orthant positif                    | 67        |
| 4.2.2      | Initialisation par VCA élargi                           | 70        |
| <b>4.3</b> | <b>Étude expérimentale des performances de SCSA-UNS</b> | <b>72</b> |
| 4.3.1      | Conditions de simulation                                | 72        |
| 4.3.2      | Performances en fonction du nombre d'échantillons       | 72        |
| 4.3.3      | Performances en fonction de $n$ et $\tau$               | 74        |
| <b>4.4</b> | <b>Robustesse de la méthode SCSA-UNS</b>                | <b>75</b> |
| 4.4.1      | Robustesse à la position du sommet du cône              | 75        |
| 4.4.2      | Robustesse à un bruit additif                           | 77        |
| <b>4.5</b> | <b>Conclusion</b>                                       | <b>83</b> |

---

## 4.1 Introduction

Nous nous intéressons dans ce chapitre aux performances de la méthode SCSA-UNS ainsi qu'à quelques détails de son implantation pratique. Pour ce faire, nous discutons d'abord, en absence de bruit, l'initialisation de l'algorithme ainsi que l'influence du nombre d'échantillons et du nombre de sources sur les indices de performances. Ces discussions nous permettront d'introduire une technique d'accélération de la méthode. Ensuite, nous discutons la robustesse de la méthode SCSA-UNS à une erreur concernant la position du sommet du cône formé par le nuage de points des signaux observations. Pour finir, nous revenons sur les fondements de la solution proposée en présence de bruit additif (solution évoquée au chapitre 3).

Tout au long de ce chapitre et pour faciliter la visualisation en dimension  $n = 3$ , les cônes et les observations sont projetés sur l'hyperplan  $(\mathcal{P})_3$ , défini par :

$$(\mathcal{P})_3 = \left\{ \mathbf{y} = [y_1, y_2, y_3]^T \in \mathbb{R}^3 \mid \sum_{i=1}^3 y_i = 1 \right\} \quad (4.1)$$

Les observations projetées sont obtenues par l'équation (4.2)

$$\tilde{x}_i(k) = x_i(k) / \sum_{l=1}^3 x_l(k) \quad (4.2)$$

## 4.2 Initialisation de l'algorithme SCSA-UNS

Nous nous plaçons d'abord dans le cas sans bruit, nous supposons aussi que le mélange est déterminé ( $m = n$ ) et la matrice de mélange est non-négative ( $\mathbf{A} \geq 0$ ). Le principe de la méthode SCSA-UNS est de partir d'un cône simplicial initial,  $\text{Span}^+(\mathbf{A}_0)$ , contenant le nuage de points des signaux observés et de réduire itérativement l'ouverture de ce cône jusqu'à ce qu'il s'ajuste à ce nuage de points. La première étape de l'algorithme consiste donc à trouver un tel cône, c'est-à-dire trouver une matrice non-négative  $\mathbf{A}_0 \geq 0$  telle que :

$$(\mathbf{A}_0)^{-1} \mathbf{X} \geq 0 \quad (4.3)$$

### 4.2.1 Initialisation par l'orthant positif

Les observations étant aussi non-négatives ( $\mathbf{X} \geq 0$ ), nous pouvons de prime abord choisir  $\mathbf{A}_0 = \mathbf{I}_n$  [OSJJ12b, OSJJ12a]. Cependant, cette initialisation triviale, qui ne dépend pas du nuage de points des signaux observations, peut conduire aux *situations de blocage*, évoquées dans le chapitre 3. Pour illustrer nos propos, nous montrons sur les figures 4.1 et 4.2, l'évolution en fonction des itérations, des cônes courants estimés par l'algorithme SCSA-UNS et de leurs ouvertures, respectivement pour les matrices de mélange  $\mathbf{A}_1$  et  $\mathbf{A}_2$ . Pour ces deux situations, l'algorithme est initialisé par  $\mathbf{A}_0 = \mathbf{I}_3$ ; les paramètres sont fixés à  $\epsilon = 10^{-6}$ ,  $\gamma = \frac{1}{4}$  et  $\eta = 10^{-1}$ . Au cours des itérations, nous pouvons voir que la matrice de mélange  $\mathbf{A}_1$  est retrouvée sans blocage, alors que pour  $\mathbf{A}_2$  surviennent des situations de blocage.

$$\mathbf{A}_1 = \begin{pmatrix} 0.95 & 0.3 & 0.97 \\ 0.82 & 0.15 & 0.28 \\ 0.65 & 0.46 & 0.06 \end{pmatrix} \text{ et } \mathbf{A}_2 = \begin{pmatrix} 0.55 & 0.09 & 0.77 \\ 0.94 & 0.93 & 0.61 \\ 0.29 & 0.83 & 0.96 \end{pmatrix} \quad (4.4)$$



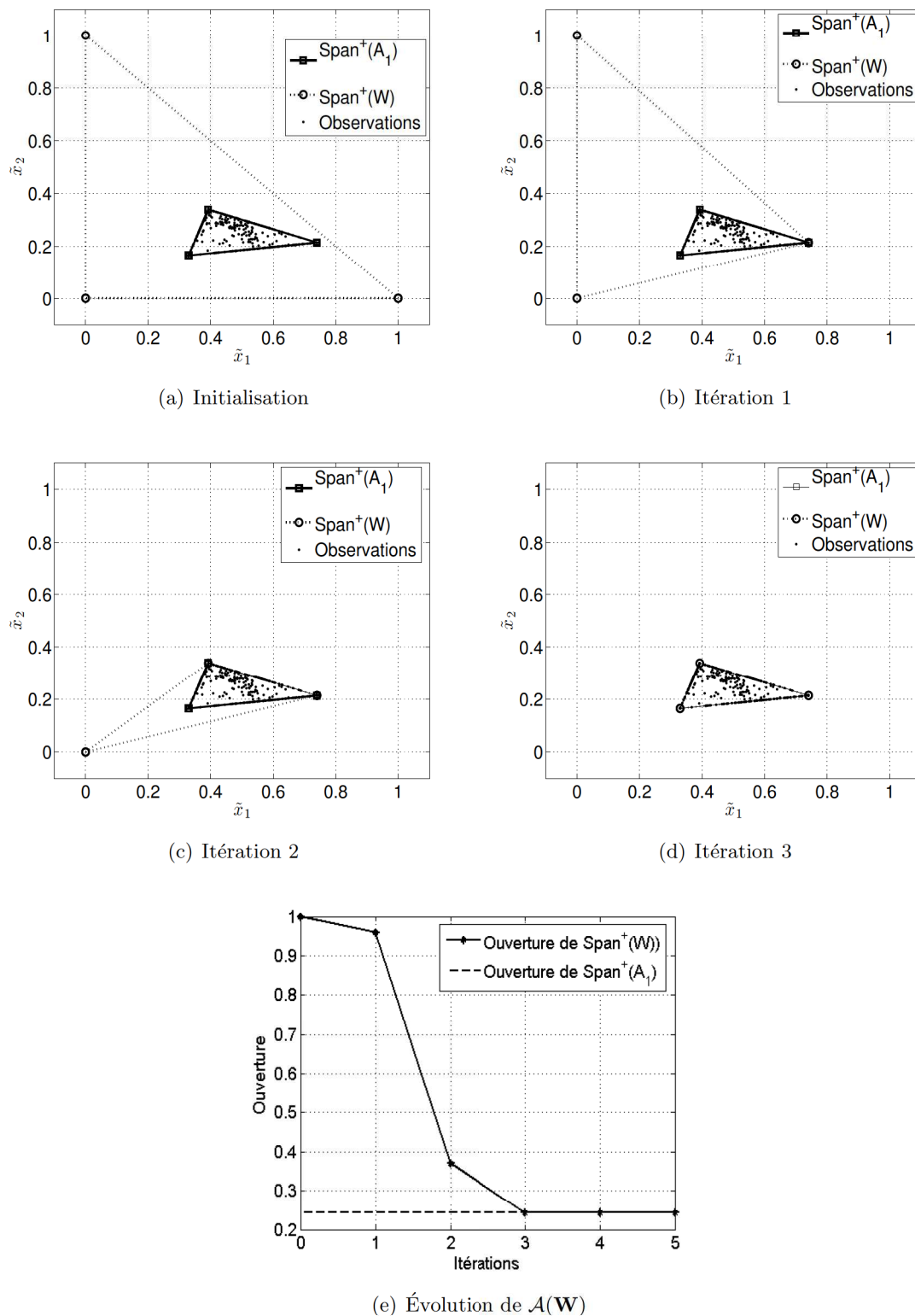


FIGURE 4.1 – Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange  $\mathbf{A}_1$  et une initialisation par l'orthant positif

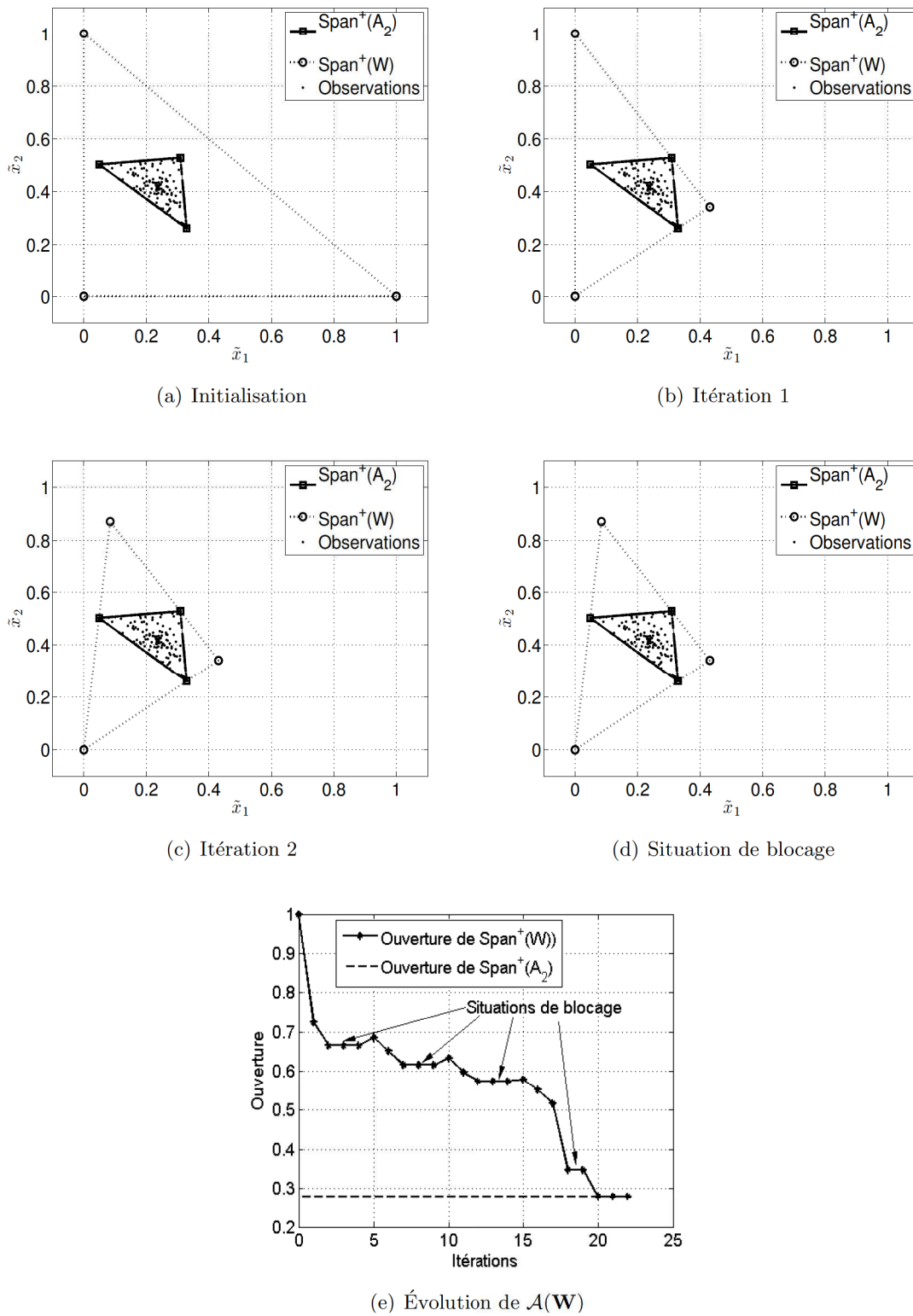


FIGURE 4.2 – Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange  $\mathbf{A}_2$  et une initialisation par l'orthant positif

### Nombre de blocages en fonction du nombre de sources

L’initialisation triviale par l’orthant positif peut conduire aux situations de blocages comme l’ont montré les deux exemples ci-dessus. Ces blocages ont pour effet d’augmenter le coût de calcul de l’algorithme; en effet, pour chaque blocage, un déblocage doit être effectué avant de poursuivre la réduction de l’ouverture du cône courant (voir chapitre 3). Afin d’appréhender la “gravité” de ce phénomène, nous évaluons ci-dessous (en absence de bruit), le nombre de blocages avant convergence, en fonction du nombre de sources. Pour ce faire, nous fixons  $m - n$  et  $K = 5000$ , les coefficients des matrices de mélange sont générés aléatoirement suivant une distribution uniforme entre 0 et 1. Les échantillons de chaque source sont générés aléatoirement suivant une distribution uniforme entre 0 et 1 avec 50% de coefficients non nuls. Les paramètres sont fixés à  $\epsilon = 10^{-6}$ ,  $\gamma = \frac{1}{4}$  et  $\eta = 10^{-1}$ . La figure 4.3 montre les évolutions, en fonction du nombre de sources  $n$ ; du nombre moyen de blocages et du coût moyen de calcul, pour 20 exécutions indépendantes Monte Carlo. Il ressort que le nombre de blocages et le coût de calcul augmentent significativement avec le nombre de sources. Il est donc nécessaire de trouver une meilleure initialisation afin de diminuer ces indicateurs.

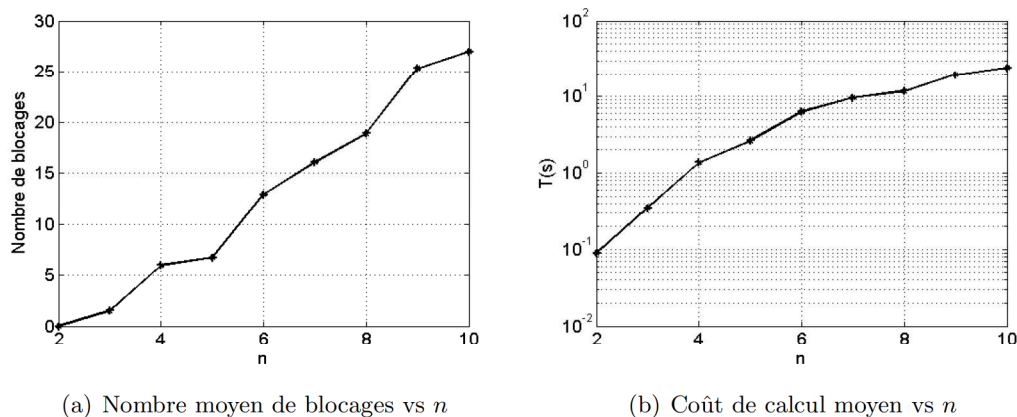


FIGURE 4.3 – Évolutions, en fonction du nombre de sources, du nombre moyen de blocages et du coût moyen de calcul, pour une initialisation par l’orthant positif

### 4.2.2 Initialisation par VCA élargi

Les méthodes géométriques fondées sur la dominance locale des sources permettent d’estimer les points extrêmes du nuage de points des signaux observés. En présence de bruit et/ou lorsque l’hypothèse de dominance locale des sources n’est pas satisfaite, ces points extrêmes ne permettent pas d’estimer correctement la matrice de mélange (voir chapitres 2 et 3), cependant ils fournissent en général un cône contenant une grande partie du nuage de points des signaux observations. L’idée est d’utiliser ce cône légèrement élargi comme initialisation de la méthode SCSA-UNS. Parmi les méthodes fondées sur la détermination des points extrêmes du nuage des observations, nous avons choisi l’algorithme VCA [NBD05], du

fait de son faible coût de calcul. Le cône obtenu par cet algorithme est élargi afin d'assurer qu'il contienne bien la totalité du nuage de points des signaux observations.

**Effet de l'initialisation par VCA élargi sur le nombre de blocages :** Les figures 4.4 et 4.5 montrent les cônes initiaux et les évolutions des ouvertures des cônes courants, respectivement pour les matrices de mélange  $\mathbf{A}_1$  et  $\mathbf{A}_2$ , lorsque l'algorithme est initialisé par VCA élargi. Nous pouvons voir que grâce à cette initialisation, les deux matrices de mélange  $\mathbf{A}_1$  et  $\mathbf{A}_2$  sont retrouvées sans blocages.

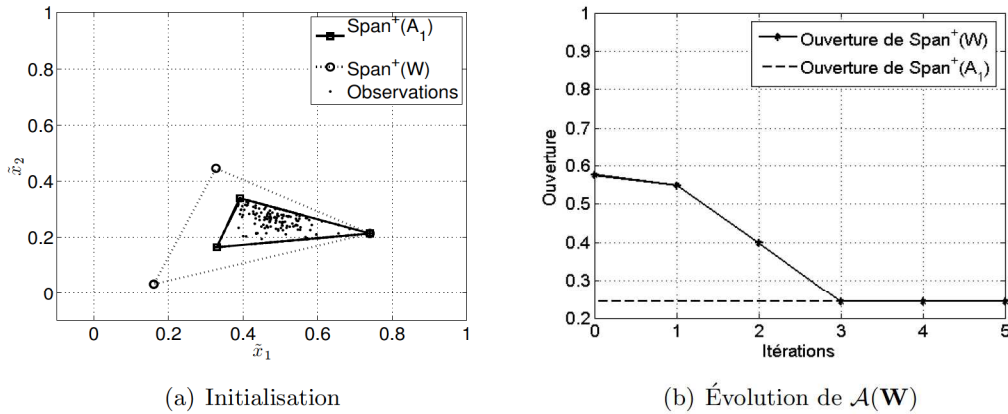


FIGURE 4.4 – Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange  $\mathbf{A}_1$  et une initialisation par VCA élargi

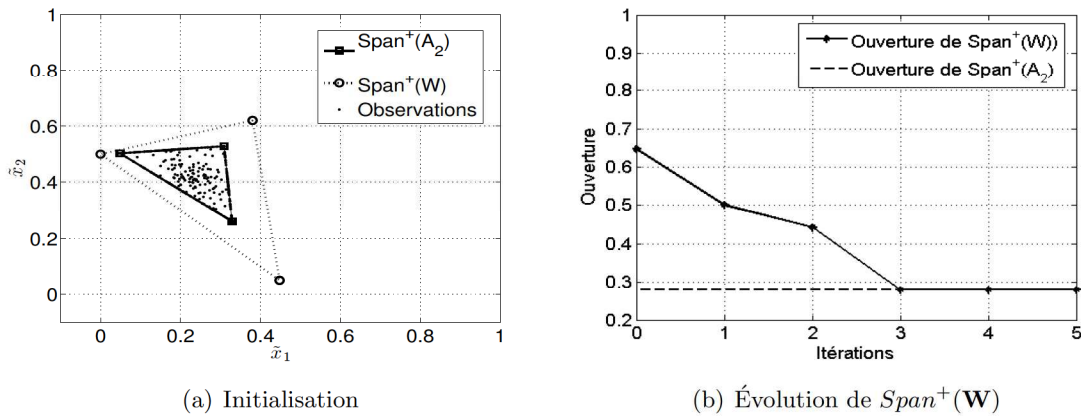


FIGURE 4.5 – Évolution en fonction des itérations du cône courant et de son ouverture pour la matrice de mélange  $\mathbf{A}_2$  et une initialisation par VCA élargi

Afin de quantifier l'apport de l'initialisation par VCA élargi, nous évaluons dans les mêmes conditions que le paragraphe 4.2.1, le nombre moyen de blocages et le coût moyen de calcul, en fonction du nombre de sources. La figure 4.6 montre les évolutions de ces deux indices obtenus respectivement, à partir d'initialisations par VCA élargi et par l'orthant positif. Nous voyons sur cette figure que l'initialisation par VCA élargi permet de réduire

considérablement le nombre moyen de blocages et le coût moyen de calcul. Ces améliorations sont d'autant plus significatives que le nombre de sources augmente.

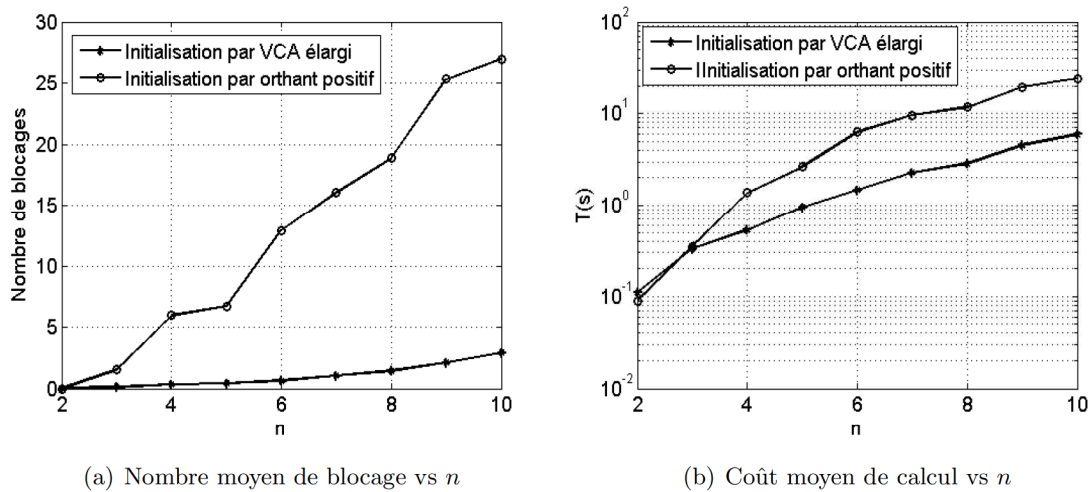


FIGURE 4.6 – Évolutions en fonction du nombre de sources, du nombre moyen de blocages et du coût moyen de calcul, pour une initialisation par VCA élargi

### 4.3 Étude expérimentale des performances de SCSA-UNS

Nous analysons dans cette section, en absence de bruit, les performances de l'algorithme SCSA-UNS en fonction du nombre d'échantillons, puis en fonction du nombre de sources. L'objectif est d'évaluer les performances de la méthode en grandes dimensions.

#### 4.3.1 Conditions de simulation

Nous nous plaçons dans le cas où le mélange est déterminé ( $m = n$ ) et où la matrice de mélange est aussi non-négative ( $\mathbf{A} \geq 0$ ). Les coefficients de la matrice sont générés aléatoirement suivant une distribution uniforme entre 0 et 1.

Les échantillons de chaque source sont générés suivant une distribution uniforme entre 0 et 1 avec une proportion  $\tau$  ( $0 \leq \tau \leq 1$ ) de coefficients non nuls,  $\tau$  est aussi appelé taux de parcimonie. Une valeur de  $\tau = 0$  correspond à une source dont tous les échantillons sont nuls et une valeur de  $\tau = 1$  correspond à une source dont tous les échantillons sont non nuls.

Nous fixons les paramètres de l'algorithme de déblocage à  $\epsilon = 10^{-6}$ ,  $\gamma = \frac{1}{4}$  et  $\eta = 10^{-1}$ .

Les figures ci-dessous montrent les évolutions des erreurs moyennes de séparation,  $E_{sep}(dB)$ , pour 20 exécutions indépendantes Monte Carlo.

#### 4.3.2 Performances en fonction du nombre d'échantillons

La figure 4.7 montre les évolutions, en fonction du nombre d'échantillons  $K$ , des erreurs moyennes de séparation  $E_{sep}(dB)$ , pour différents nombres de sources  $n$  et différents taux de parcimonie  $\tau$ . Cette figure indique que si le taux de parcimonie est inférieur à 1 ( $\tau < 1$ ) alors, au delà d'un certain nombre  $K_0$ , l'augmentation du nombre d'échantillons n'améliore plus

l'erreur de séparation. Ce résultat s'explique par le fait qu'un nombre fini de points sur les facettes d'un cône (au moins  $n - 1$  sur chaque facette) suffit à définir parfaitement ce cône. Par ailleurs, dans le nuage de points des signaux observations, seuls les points situés sur les facettes ou sur les arêtes du cône engendré par la matrice de mélange, donnent la forme de ce cône. Tous les autres points situés complètement à l'intérieur du cône engendré par la matrice de mélange n'apportent pas d'information sur la forme de ce cône, et par conséquent n'interviennent pas dans l'estimation de la matrice de mélange.

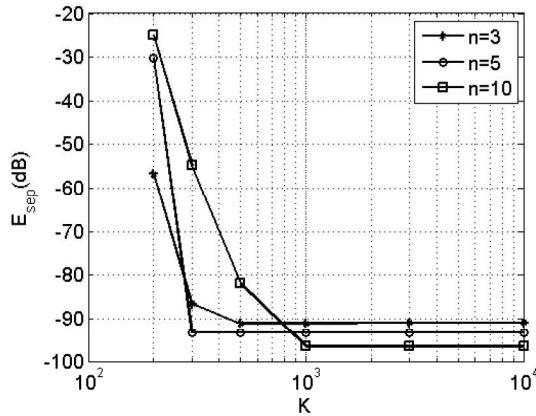
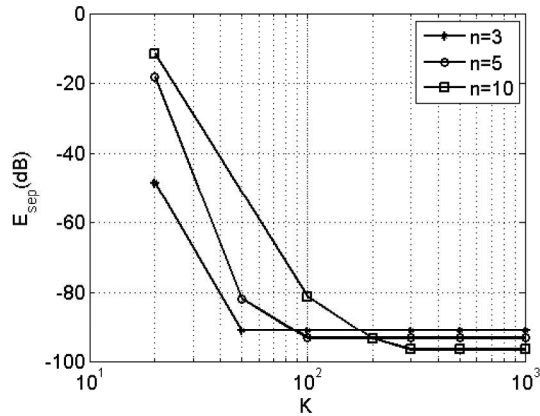
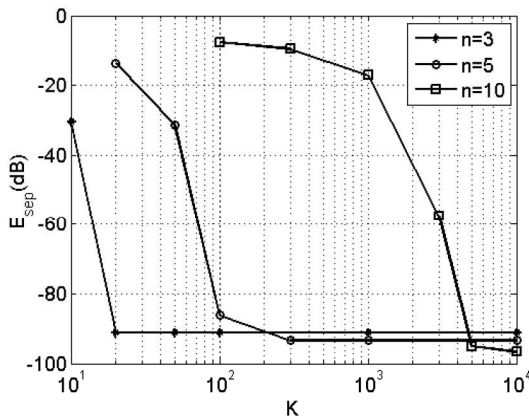
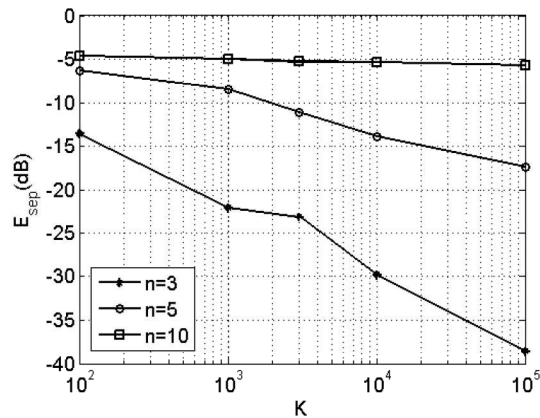
(a)  $E_{sep}(dB)$  vs  $K$  pour  $\tau = 0.01$ (b)  $E_{sep}(dB)$  vs  $K$  pour  $\tau = 0.1$ (c)  $E_{sep}(dB)$  vs  $K$  pour  $\tau = 0.5$ (d)  $E_{sep}(dB)$  vs  $K$  pour  $\tau = 1$ 

FIGURE 4.7 – Erreurs moyennes de séparation en fonction du nombre d'échantillons, et de la parcimonie des sources

Pour estimer correctement la matrice de mélange à l'aide de la méthode SCSA-UNS, il suffit que les facettes ou les arêtes du cône qu'elle engendre soient parfaitement identifiables à partir du nuage de points des signaux observations. Pour un nombre de sources  $n$  fixé, nous nous intéressons au nombre minimum d'échantillons,  $K_0^n$ , qu'il faut pour satisfaire cette condition. Théoriquement, si les sources sont localement dominantes, alors il suffit des  $n$  points de dominance locale pour estimer correctement la matrice de mélange, dans ce cas  $K_0^n = n$ . Si les sources ne sont pas localement dominantes, alors il peut suffire d'avoir  $n - 1$

points non alignés sur chaque facette du cône engendré par la matrice de mélange pour que ces facettes soit correctement identifiables selon [Hen97] et *Proposition 5* du chapitre 3. Dans ce cas, le nombre minimum d'échantillons est  $K_0^n = n \times (n - 1)$ .

Nous traçons sur la figure 4.8 l'évolution, en fonction du taux de parcimonie, du nombre minimum d'échantillons qu'il faut (trouvé expérimentalement), pour réaliser une séparation exacte. Nous voyons sur cette figure qu'en pratique,  $K_0^n > n(n - 1)$ , puisque tous les points ne sont pas sur les facettes. Cette figure nous montre aussi que pour le nombre de sources  $n$  fixé, il existe un taux de parcimonie  $\tau_n$  pour lequel le nombre minimum d'échantillons nécessaire pour une séparation exacte,  $K_0^n(\tau)$ , réalise son minimum.  $K_0^n(\tau)$  varie dans le même sens que  $\tau$  lorsque  $\tau > \tau_n$  et  $K_0^n(\tau)$  varie en sens contraire de  $\tau$  lorsque  $\tau < \tau_n$ .

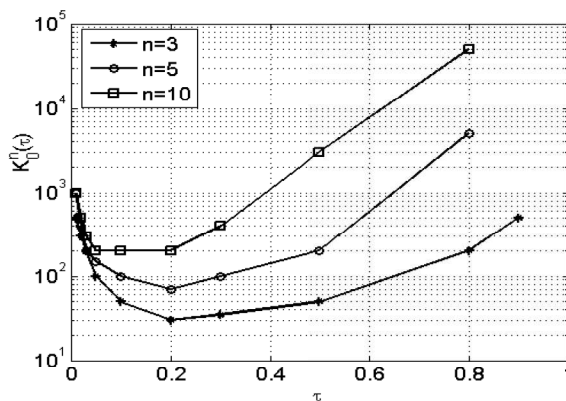


FIGURE 4.8 – Évolution en fonction du taux de parcimonie et du nombre de sources, du nombre minimum d'échantillons nécessaire pour une séparation exacte

### 4.3.3 Performances en fonction du nombre de sources et du taux de parcimonie, comparaison avec d'autres méthodes géométriques

Dans cette section, nous évaluons les performances en fonction du nombre de sources, pour différents taux de parcimonie. Ces performances sont comparées avec les méthodes géométriques SISAL [BD09] et nLCA-IVM [WCCW10], qui sont fondées sur la recherche du simplexe de volume minimal contenant le nuage de points des signaux observations. Les figures 4.9 et 4.10 montrent les évolutions en fonction du nombre de sources, des erreurs moyennes de séparation  $E_{sep}(dB)$  et des coûts moyens de calcul  $T(s)$ , pour différents taux de parcimonie. Le nombre d'échantillons est fixé à  $K = 10000$ .

Ces figures nous montrent que lorsque le taux de parcimonie est petit, la méthode SCSA-UNS présente les meilleurs indices de performance  $E_{sep}(dB)$  et  $T(s)$ ; ces indices restent stables lorsque le nombre de sources augmente. Par contre, lorsque le taux de parcimonie est grand, les performances de la méthode SCSA-UNS se dégradent au fur et à mesure que le nombre de sources augmente.

Nous voyons aussi sur les figures 4.9 et 4.10 qu'indépendamment du taux de parcimonie, lorsque le nombre de sources est supérieur à 10 ( $n > 10$ ), la méthode nLCA-IVM présente un coût de calcul extrêmement élevé (de l'ordre de quelques heures), ce qui la rend inutilisable

en grande dimension.

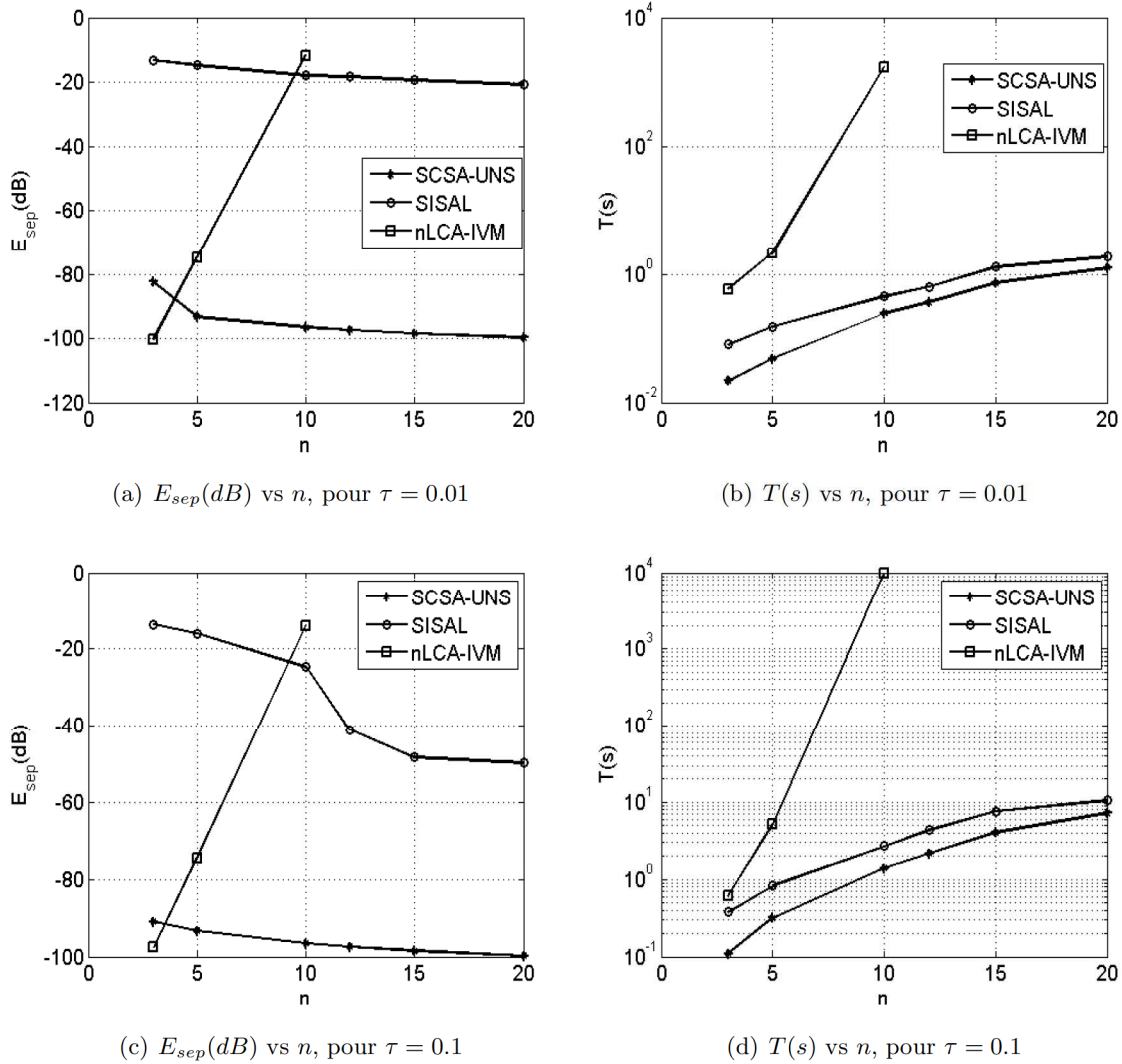


FIGURE 4.9 – Performances en fonction du nombre de sources, pour  $\tau = 0.01$  et  $\tau = 0.1$

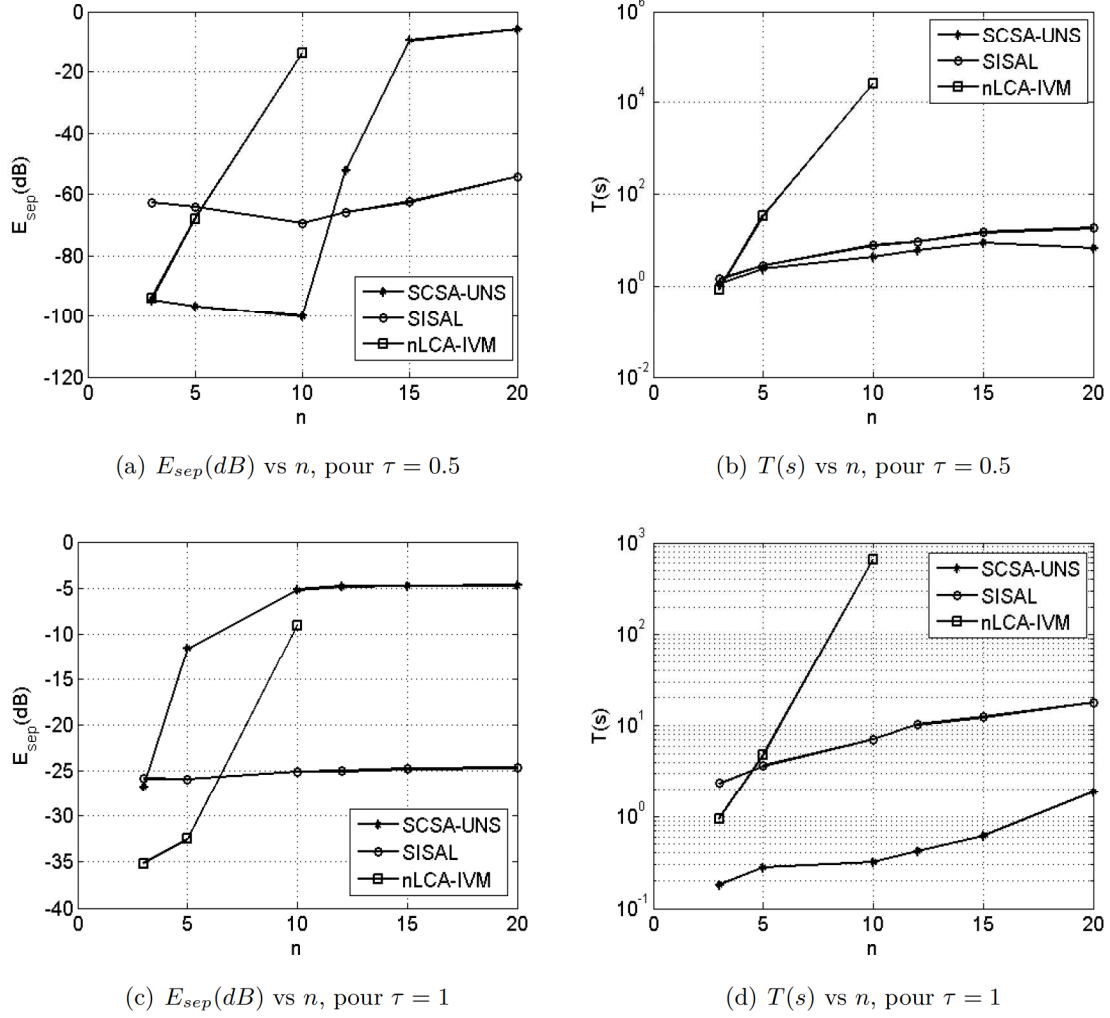
## 4.4 Robustesse de la méthode SCSA-UNS

Nous évaluons, dans cette section, la robustesse de la méthode SCSA-UNS à une erreur sur la position du sommet du cône simplicial engendré par le nuage de points des signaux observés, ainsi que la robustesse de cette méthode au bruit.

### 4.4.1 Robustesse à la position du sommet du cône

La méthode SCSA-UNS suppose que le sommet du cône formé par le nuage de points des signaux observés (cône engendré par la matrice de mélange) se situe à l'origine de l'espace. Cependant, la présence d'arrière-plans sur les observés (offsets engendrant des décalages par rapport à zéro) peut déplacer, par rapport à l'origine de l'espace, le sommet du cône



FIGURE 4.10 – Performances en fonction du nombre de sources, pour  $\tau = 0.5$  et  $\tau = 1$ 

formé par le nuage de points des signaux observations. Pour illustrer ce phénomène, nous considérons la matrice de mélange donnée par :

$$\mathbf{A} = \begin{pmatrix} 0.95 & 0.3 & 0.97 \\ 0.82 & 0.15 & 0.28 \\ 0.65 & 0.46 & 0.06 \end{pmatrix}$$

Les sources sont générées de la même manière que dans la section 4.3 avec  $\tau = 0.5$ , le nombre d'échantillons est  $K = 1000$ . Nous définissons le vecteur  $\mathbf{v}$  par :

$$\mathbf{v} = \begin{bmatrix} \rho \cos(\theta) \cos(\phi) \\ \rho \sin(\theta) \cos(\phi) \\ \rho \sin(\phi) \end{bmatrix} \quad (4.5)$$

avec  $\rho \in \mathbb{R}_+$ ,  $\theta \in [0 \frac{\pi}{2}]$  et  $\phi \in [0 \frac{\pi}{2}]$ .

Soient maintenant les observations données par l'équation (4.6),  $\forall 1 \leq k \leq K$  :

$$\mathbf{x}(k) = \mathbf{A}\mathbf{s}(k) + \mathbf{v} \quad (4.6)$$

La figure 4.11 illustre les projections sur  $(\mathcal{P})_3$  du cône engendré par  $\mathbf{A}$  et des nuages de points des signaux observations, lorsque  $\theta = \frac{\pi}{8}$  et  $\phi = \frac{\pi}{4}$ , respectivement pour  $\rho = 10^{-3}$  et  $\rho = 1$ .

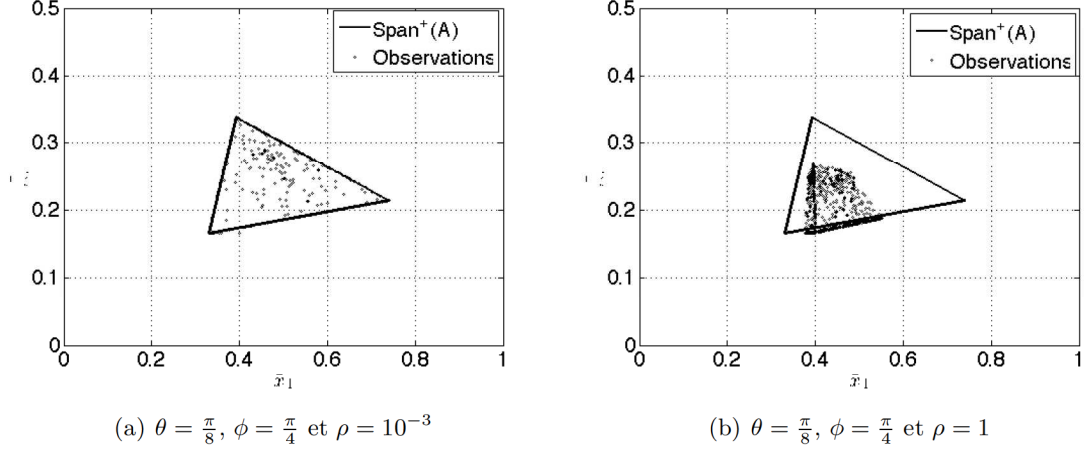


FIGURE 4.11 – Effet d’un offset sur le nuage de points des signaux observations

Dans le cas d’un “petit” déplacement du sommet par rapport à l’origine de l’espace (Figure 4.11.a), le cône formé par la matrice de mélange reste correctement identifiable à partir du nuage de points des observations. Par contre, un “grand” déplacement du sommet entraîne un décalage entre le cône formé par le nuage des observations et le cône engendré par la matrice de mélange (Figure 4.11.b).

Nous montrons sur les figures 4.12 et 4.13, la répartition de l’erreur de séparation  $E_{sep}(dB)$ , en fonction de  $\theta$  et de  $\phi$ , respectivement pour  $\rho = 10^{-3}$  et pour  $\rho = 1$ .

Ces deux figures nous permettent de dire qu’il existe un bassin en dehors duquel, le déplacement du sommet du cône formé par le nuage de points des observations, entraîne une mauvaise estimation de la matrice de mélange. La forme de ce bassin dépend de la matrice de mélange mais est difficilement caractérisable et/ou paramétrable.

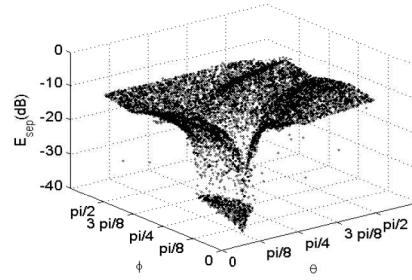
#### 4.4.2 Robustesse à un bruit additif

Nous traitons dans ce paragraphe de la robustesse de l’algorithme SCSA-UNS à un bruit additif. Pour ce faire, nous analysons d’abord “visuellement” l’effet d’un bruit additif sur le nuage de points des signaux observations non bruités, puis nous proposons une solution pour améliorer la robustesse au bruit de l’algorithme.

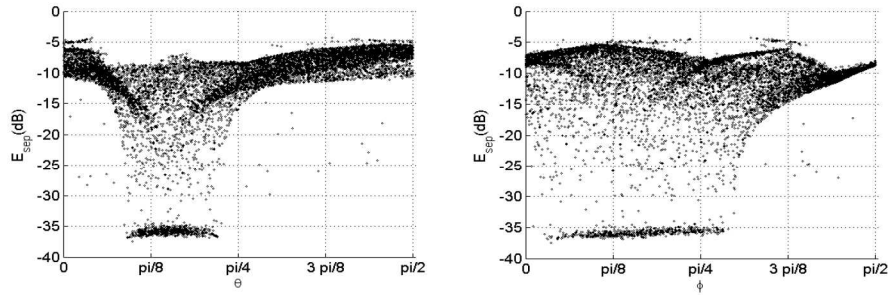
##### Analyse de l’effet d’un bruit additif sur le nuage des observations

Pour évaluer “visuellement” l’effet d’un bruit additif sur le nuage de points des signaux observations, nous nous plaçons dans la situation  $m = n = 3$ . La matrice de mélange vaut :

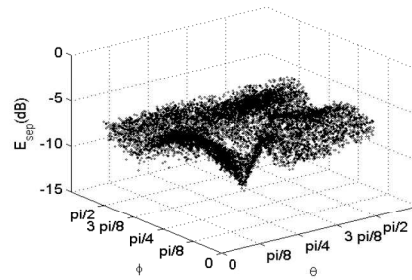
$$\mathbf{A} = \begin{pmatrix} 0.95 & 0.3 & 0.97 \\ 0.82 & 0.15 & 0.28 \\ 0.65 & 0.46 & 0.06 \end{pmatrix}$$



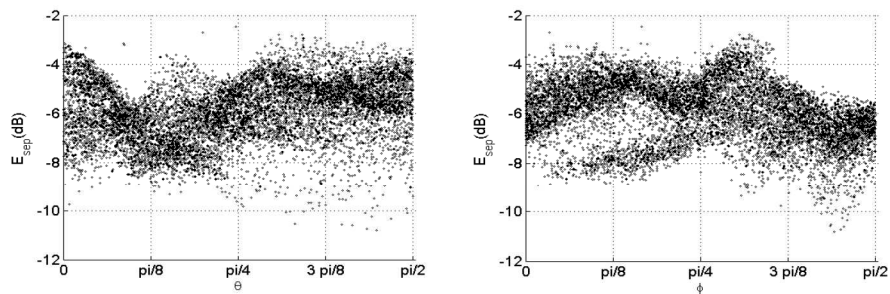
(a) Vue en trois dimensions



(b) Vue suivant les axes des abscisses et (c) Vue suivant les axes des ordonnées et des cotes

FIGURE 4.12 – Erreur de séparation en fonction de  $\theta$  et de  $\phi$ , pour  $\rho = 10^{-3}$ 

(a) Vue en trois dimensions



(b) Vue suivant les axes des abscisses et (c) Vue suivant les axes des ordonnées et des cotes

FIGURE 4.13 – Erreur de séparation en fonction de  $\theta$  et de  $\phi$ , pour  $\rho = 1$

Les échantillons de chaque source sont générés de la même manière que précédemment avec  $\tau = 0.5$  et  $K = 1000$ . Les échantillons de bruit sur chaque observation sont générés aléatoirement suivant une distribution gaussienne de moyenne nulle et de variance unité tronquée sur  $\mathbb{R}_+$ , avec un  $RSB = 40dB$ .

La figure 4.14 montre le nuage de points des signaux observations bruitées ainsi que le cône engendré par la matrice de mélange estimée à partir des observations bruitées. Cet exemple illustre à quel point, le bruit peut dégrader les performances de la méthode SCSA-UNS initiale, même à fort RSB ( $40dB$  dans cet exemple), ce qui peut sembler étonnant, mais que nous allons comprendre dans la suite.

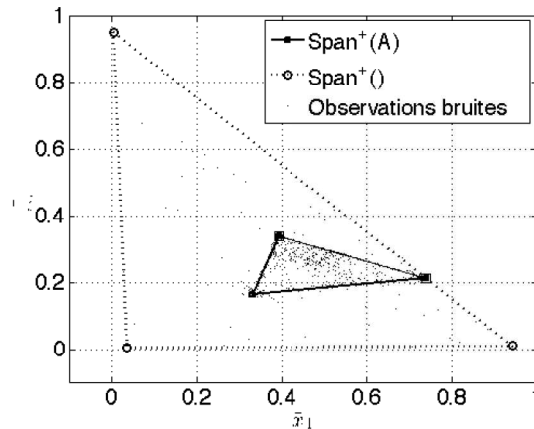


FIGURE 4.14 – Nuage de points des signaux observations bruités

Afin de mieux appréhender l'effet du bruit sur le nuage de points des signaux observations non bruités, nous analysons plus finement ce phénomène en divisant le nuage de points des signaux observations non bruités en quatre classes :

- Classe 1 : les points situés sur le sommet de  $Span^+(\mathbf{A})$
- Classe 2 : les points situés sur les arêtes de  $Span^+(\mathbf{A})$
- Classe 3 : les points situés sur les facettes de  $Span^+(\mathbf{A})$
- Classe 4 : les points situés à l'intérieur de  $Span^+(\mathbf{A})$

L'analyse a consisté à étudier séparément l'effet du bruit sur chaque classe. La figure 4.15 montre les nuages de points obtenus si l'on rajoute le bruit uniquement sur les observations de chaque classe, ainsi que le cône engendré par la matrice de mélange estimée. Il découle de cette analyse qu'en présence de bruit, les observations de la Classe 1 sont les plus susceptibles de modifier significativement le nuage de points des signaux observations et d'engendrer une mauvaise estimation de la matrice de mélange. Cette figure montre aussi que, dans cet exemple, la corruption par le bruit des points situés sur les facettes et à l'intérieur du cône engendré par la matrice de mélange n'a pas entraîné une grande dégradation de l'estimation de la matrice de mélange.

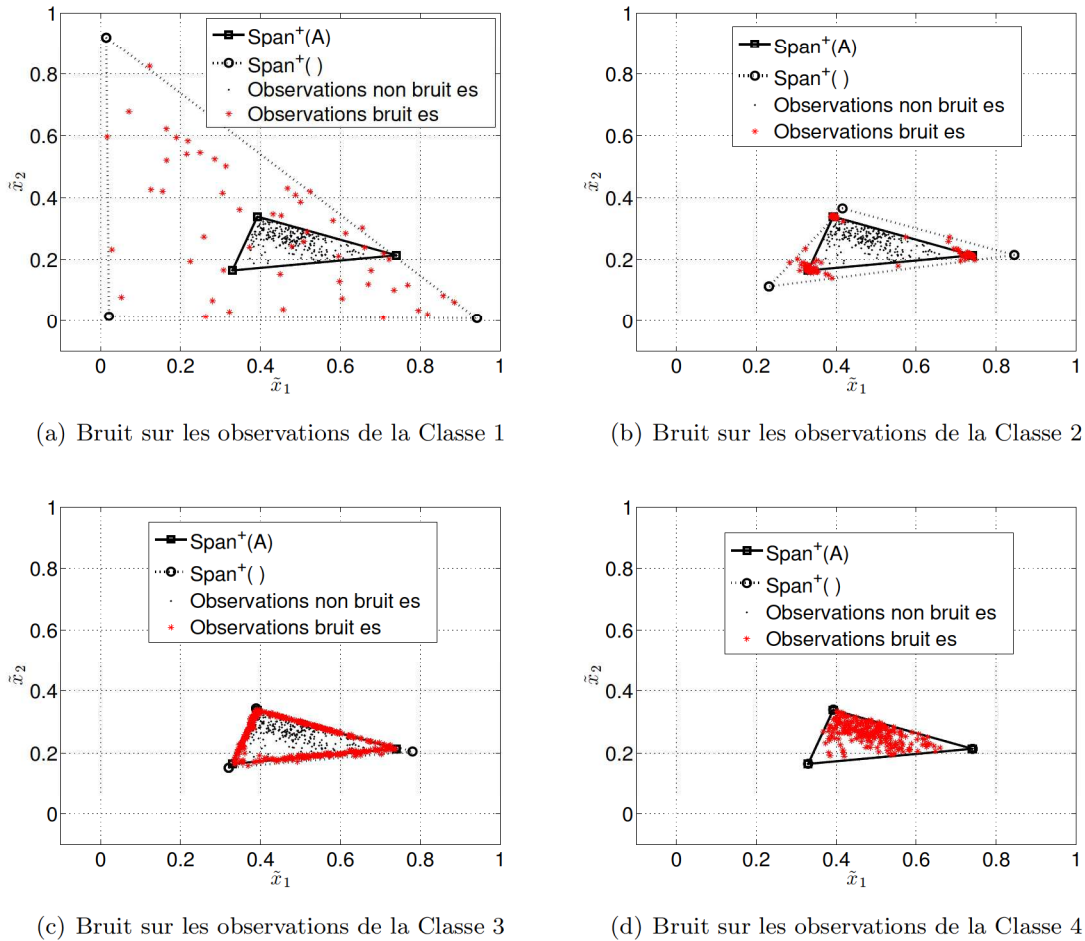
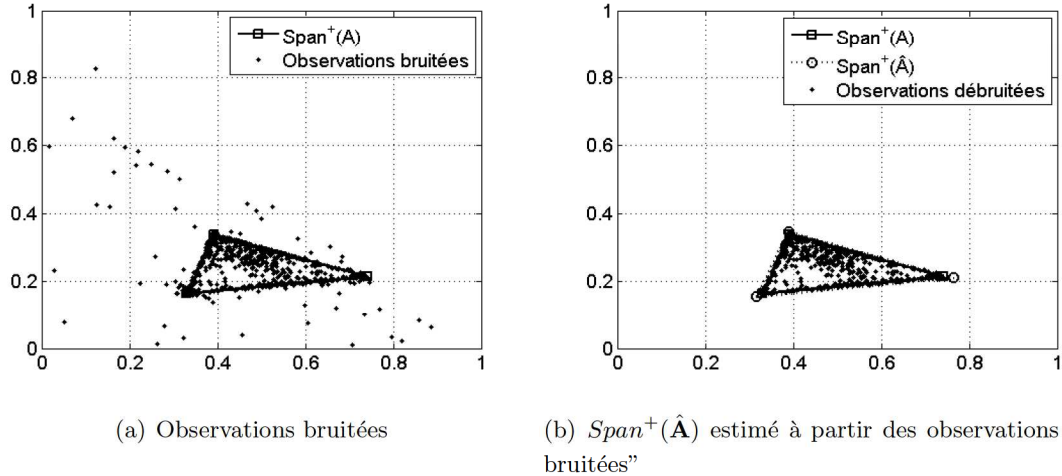


FIGURE 4.15 – Effet du bruit sur le nuage de points des différentes classes

### Approche intuitive de réduction de l'effet du bruit

L'analyse de l'effet du bruit sur le nuage de points des signaux observations a montré que les points de la Classe 1 sont les plus susceptibles de modifier significativement le nuage de points des signaux observations et d'engendrer une mauvaise estimation de la matrice de mélange. Or, en absence de bruit, ces points sont situés sur le sommet du cône engendré par la matrice de mélange, et ont donc la particularité d'avoir des coordonnées nulles et par conséquent des normes nulles. Nous en déduisons qu'en présence de bruit additif, si le rapport signal sur bruit est "suffisamment élevé", nous pouvons assimiler les points de la Classe 1 à ceux de "petites" normes. Ces points peuvent ainsi être identifiés en sélectionnant toutes les observations de normes inférieures à un certain seuil de bruit  $\lambda_b$ .

La figure 4.16 illustre le nuage de points des observations bruitées, ainsi que le cône engendré par l'estimée de la matrice de mélange, obtenu à partir des observations "débruitées" avec  $\lambda_b = 0.1$ . Comparativement à la figure 4.14, la figure 4.16 montre que la suppression des observations bruitées de faibles normes permet d'améliorer l'estimation de matrice de mélange.

FIGURE 4.16 – Illustration du cône engendré par l'estimée de la matrice de mélange, pour  $\lambda_b = 0.1$ 

### Influence du seuil de bruit $\lambda_b$ , sur les performances de séparation

Nous consignons dans le tableau 4.1, les variations des indices moyens de performances, pour 20 réalisations indépendantes Monte Carlo, le rapport signal sur bruit étant fixé à  $40dB$ . Ce tableau montre qu'une attention particulière doit être accordée au choix du seuil de bruit  $\lambda_b$ . Un seuil trop faible ne permet pas d'éliminer tous les points les plus bruités et un seuil trop fort peut entraîner l'élimination des observations donnant la forme du cône engendré par la matrice de mélange. Ces deux cas conduisent à une mauvaise estimation de la matrice de mélange et des sources.

|                    | $E_{sep}(dB)$ | $E_{corr}(dB)$ |
|--------------------|---------------|----------------|
| $\lambda_b = 0$    | -3.41         | -9.06          |
| $\lambda_b = 0.01$ | -7.45         | -15.35         |
| $\lambda_b = 0.1$  | -14.79        | -23.14         |
| $\lambda_b = 0.3$  | -12.33        | -10.65         |

TABLE 4.1 – Variations en fonction du seuil de bruit, des indices de performances :  $RSB = 40dB$ 

### Influence du taux de parcimonie, et du rapport signal sur bruit sur les performances de séparation

La méthode de débruitage proposée consiste à supprimer les observations de faible norme. Ces observations sont d'autant plus présentes que les sources sont parcimonieuses. La figure 4.17 montre les évolutions en fonction du  $RSB$  et du taux de parcimonie, des indices moyens de performance obtenus à partir de 20 exécutions indépendantes Monte Carlo. Nous fixons  $m = n = 3$  et  $K = 1000$ . Les sources et les coefficients de la matrice de mélange sont générés de la même manière que dans la section 4.3. Les échantillons de bruit sur chaque observation sont générés aléatoirement suivant une distribution gaussienne de moyenne nulle

et de variance unité tronquée sur  $\mathbb{R}_+$  et le seuil de bruit est fixé à  $\lambda_b = 0.1$ . Nous pouvons voir sur la figure 4.17 que les indices de performance sont d'autant meilleurs que le taux de parcimonie des sources est faible.

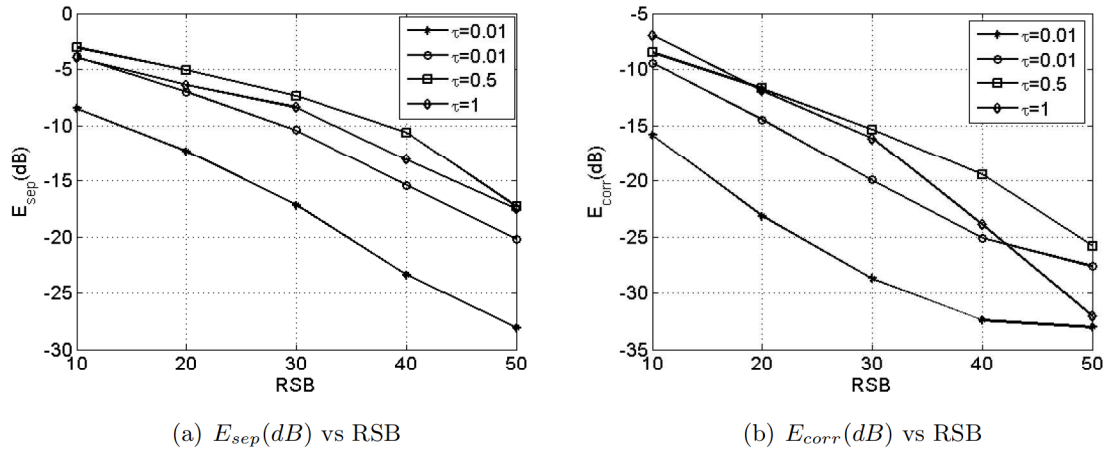


FIGURE 4.17 – Variations des indices moyens de performance, en fonction du taux de parcimonie et du rapport signal sur bruit

Afin d'évaluer l'efficacité de la méthode proposée pour traiter le bruit, nous traçons sur la figure 4.18, les évolutions des indices moyens de performance  $E_{sep}(dB)$  et  $E_{corr}(dB)$ , en fonction du nombre de sources, pour différents rapports signal sur bruit. Le taux de parcimonie est fixé à  $\tau = 0.1$ . Le mélange est déterminé ( $m = n$ ) et le nombre d'échantillons est fixé à  $K = 10000$ .

Cette figure montre que lorsque le  $RSB$  est élevé, La méthode SCSA-UNS donne des performances de séparation satisfaisantes, ces performances se dégradent lorsque le  $RSB$  diminue et/ou lorsque le nombre de sources augmente.

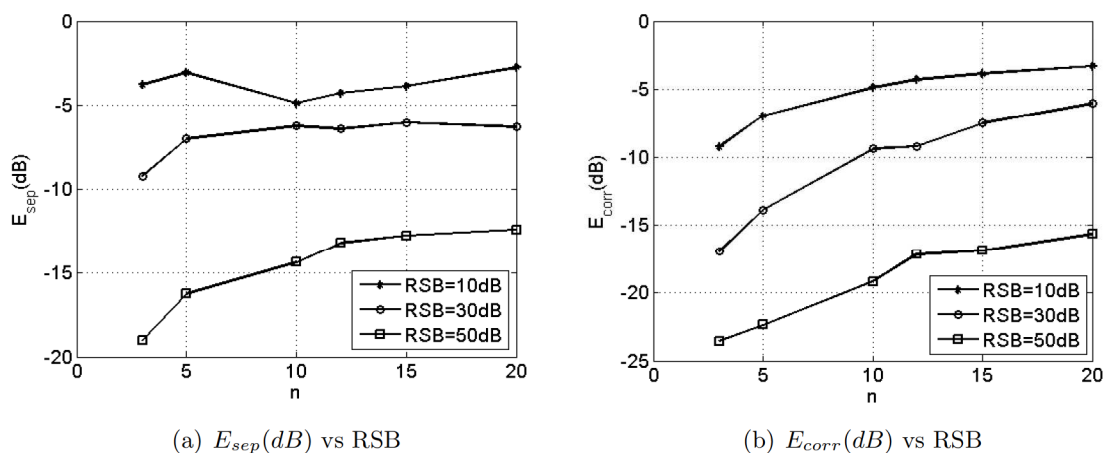


FIGURE 4.18 – Variations des indices moyens de performance, en fonction du nombre de sources et du rapport signal sur bruit

## 4.5 Conclusion

Dans ce chapitre, nous avons évalué les performances et la robustesse de l'algorithme SCSSA-UNS, après avoir discuté son initialisation. En absence de bruit, nous avons établi expérimentalement, dans un premier temps, que pour un nombre de sources et un taux de parcimonie donnés, il existe un nombre au-delà duquel l'augmentation du nombre d'échantillons n'améliore plus les performances de séparation. Nous avons montré aussi que les performances de séparation sont d'autant meilleures que les sources sont parcimonieuses, ce résultat est assez stable, vis-à-vis du nombre de sources.

Toujours en absence de bruit, nous avons aussi établi expérimentalement qu'il existe un bassin en dehors duquel, le déplacement du sommet du cône formé par le nuage de points des signaux observations entraîne un décalage entre ce dernier cône et le cône engendré par la matrice de mélange, ce qui entraîne une mauvaise estimation de la matrice de mélange. La forme de ce bassin dépend de la matrice de mélange mais est difficilement caractérisable et/ou paramétrable.

Pour finir nous avons réalisé une analyse qualitative de l'effet d'un bruit additif sur le nuage de points des signaux observations non bruitées. Cette analyse nous a permis de proposer une approche intuitive de réduction de l'effet du bruit qui consiste à éliminer du nuage de points des observations, les observations de normes inférieures à un certain seuil. Le choix du seuil est un facteur déterminant de l'efficacité de la méthode proposée et mérite une attention particulière. La méthode de réduction de l'effet du bruit est d'autant plus efficace que les sources sont parcimonieuses et le nombre de sources est faible.

Dans le dernier chapitre qui suit, nous évaluons l'algorithme SCSSA-UNS sur des données réelles de spectres de masse et d'imagerie de tomographie par émission de positons.





## Troisième partie

# Résultats de séparation sur données réelles LC-MS et TEP dynamiques



## Chapitre 5

# Résultats de séparation sur données LC-MS et TEP dynamiques

### Sommaire

---

|            |   |            |
|------------|---|------------|
| <b>5.1</b> | <b>Introduction</b>   | <b>88</b>  |
| <b>5.2</b> | <b>Métabolomique d'urine de rat traité au phénobarbital</b>           | <b>88</b>  |
| 5.2.1      | Modèle de mélange des spectres de masse                               | 88         |
| 5.2.2      | Description des données LC-MS   | 88         |
| 5.2.3      | Résultats de séparation des spectres de masse                         | 90         |
| <b>5.3</b> | <b>Estimation des pharmacocinétiques du [18F]-FDG dans le cerveau</b> | <b>97</b>  |
| 5.3.1      | Modèle de mélange des images TEP                                      | 97         |
| 5.3.2      | Résultats de séparation sur images TEP de fantômes                    | 97         |
| 5.3.3      | Résultats de séparation sur images TEP réelles de cerveau humain      | 99         |
| <b>5.4</b> | <b>Conclusion</b>   | <b>100</b> |

---

## 5.1 Introduction

Ce chapitre présente les résultats de séparation, à l'aide de l'algorithme SCSA-UNS, de spectres de masse et d'images dynamiques de tomographie par émission de positons (TEP). Dans la première situation, la séparation est effectuée sur les spectres de masse issus de l'analyse LC-MS d'urine de rat, afin d'étudier le métabolisme de rat traité au phénobarbital. Dans la deuxième situation, nous estimons les pharmacocinétiques du traceur FluoroDeoxy-Glucose marqué au fluor 18 ([18F]-FDG) dans le cerveau de patient humain, en effectuant la séparation d'images dynamiques TEP de cet organe.

## 5.2 Métabolomique d'urine de rat traité au phénobarbital

### 5.2.1 Modèle de mélange des spectres de masse

Rappelons que la *métabolomique* est la science qui étudie (identifie et quantifie) l'ensemble des métabolites (petites molécules) d'un fluide biologique. L'une des techniques d'étude métabolomique est la chromatographie en phase liquide couplée à la spectrométrie de masse. Dans ce cas, un spectromètre mesure différents spectres de masse à la sortie d'un chromatographe. Un spectre de masse mesuré à un instant donné est modélisé comme étant le mélange linéaire instantané des spectres de masse des différents métabolites co-élus (c'est-à-dire les métabolites sortant de la colonne chromatographique à cet instant). Les coefficients de mélange sont les évolutions en fonction du temps des concentrations des différents métabolites, appelés aussi profils d'élution<sup>1</sup>. Le modèle de mélange est ainsi donné par l'équation (5.1) :

$$x_t(m/z) = \sum_{i=1}^n a_i(t)s_i(m/z) \quad (5.1)$$

où  $x_t(m/z)$  est l'échantillon du spectre de masse mesuré à l'instant  $t$  et à l'indice de masse  $m/z$ ,  $a_i(t)$  est la concentration du métabolite  $i$  à l'instant  $t$  et  $s_i(m/z)$  est l'échantillon à l'indice de masse  $m/z$  du spectre de masse du métabolite  $i$ . Un métabolite est caractérisé par son profil d'élution et son spectre de masse, constitué de pics issus de l'ionisation du métabolite, ainsi que l'ionisation de ses fragments<sup>2</sup>, de ses adduits<sup>3</sup>, de ses polymères<sup>4</sup>, de ses isotopes<sup>5</sup>, etc... La figure 5.1 illustre le profil d'élution et le spectre de masse attendu pour un métabolite.

### 5.2.2 Description des données LC-MS

Les données sont fournies par le Laboratoire d'Étude du Métabolisme des Médicaments (LEMM) de la Direction des Sciences du Vivant (DSV) du Commissariat à l'Énergie Ato-

---

1. l'élution est l'action d'entraîner des espèces chimiques par un fluide, appelé éluant, en vue de les séparer.  
 2. Les fragments sont obtenus par fragmentation (cassure) d'une molécule.  
 3. Les adduits sont obtenus par liaison de la molécule avec d'autres atomes ou groupe d'atomes.  
 4. Les polymères sont constitués de plusieurs exemplaires de la même molécule liés entre eux.  
 5. Les isotopes sont obtenus lorsque plusieurs mêmes molécules ont des atomes isotopes (c'est-à-dire des atomes ayant le même nombre de charge mais des nombres de masse différents), comme par exemple le carbone 12 et 13, l'oxygène 16 et 18, l'azote 14 et 15, etc...

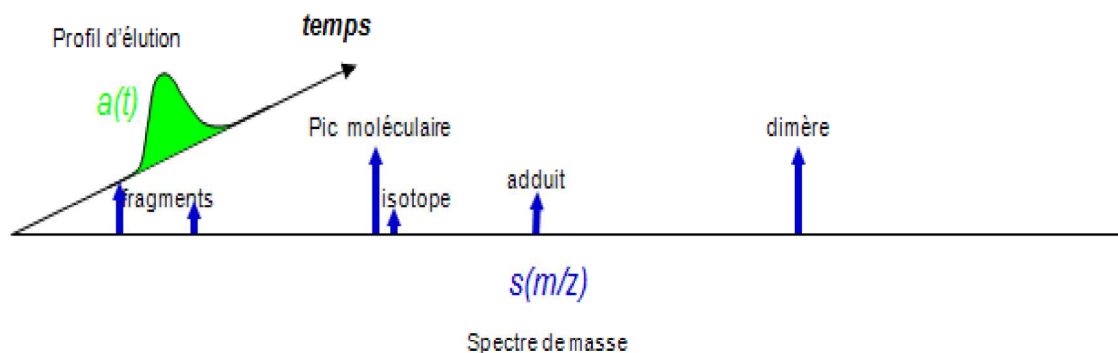


FIGURE 5.1 – Profil d'élution et spectre de masse, attendus pour un métabolite

mique et aux Énergies Alternatives (CEA-EA).

On analyse ici l'urine d'un rat traité au phénobarbital (médicament utilisé pour le traitement de l'épilepsie et des troubles du sommeil [MhdCSD08]). On s'intéresse à l'effet du médicament sur le métabolisme de l'animal.

Les mesures sont effectuées par un spectromètre LTQ-Orbitrap, à la sortie d'un chromatographe en phase liquide haute performance (HPLC). La résolution du spectromètre Orbitrap est de 5ppm<sup>6</sup> avec une précision de 4 chiffres après la virgule. La grille de masse est comprise entre 74.9998 et 1005.9999. Au total 6274 spectres ont été mesurés pendant une heure ; l'Orbitrap fournit à chaque instant de mesure les listes des masses non-nulles et des intensités correspondantes. En réalité chaque spectre de masse mesuré contient virtuellement 400000 échantillons dont en moyenne seuls 4000 échantillons sont non-nuls (les indices de masse et les intensités de ces échantillons sont fournis par l'Orbitrap). Le taux de parcimonie des spectres de masse est donc de l'ordre de  $\tau = 0.01$ .

### Prétraitements des spectres de masse mesurés

**Sélection des indices de masse :** Ce prétraitement a consisté à supprimer tous les indices de masse pour lesquels :

- Les intensités sont nulles sur tous les spectres mesurés
- Le maximum des intensités, sur tous les spectres mesurés, est inférieur à 10000 (seuil conseillé par les biologistes)
- Les intensités sont quasi-constantes sur tous les spectres mesurés

Ce premier prétraitement permet de réduire le nombre d'échantillons par spectres de 400000 à 36878. La nouvelle grille de masse est comprise 75,0276 et 995,5791.

**Sélection de la zone d'analyse** Pour chaque spectre de masse mesuré, le courant total est obtenu en additionnant les intensités de ses différents pics selon la dimension "masse". La figure 5.2 montre le chromatogramme total, représentant l'évolution du courant total en fonction du temps.

6. 1ppm = 1 part per million = 0.0001%

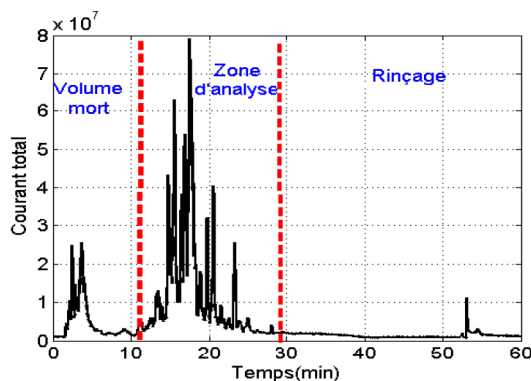


FIGURE 5.2 – Chromatogramme total

Trois zones peuvent être distinguées sur le chromatogramme total :

- La première zone, avant 10 min, correspond à l'élution des constituants polaires de l'urine de rat. Les constituants polaires interagissent peu avec la colonne chromatographique qui est apolaire dans le cadre de l'expérience qui nous intéresse ici. Les profils d'élution des différents métabolites sont donc peu séparés. Cette zone n'est pas encore utilisée par les biologistes, son exploitation constitue un défi futur.
- La deuxième zone, située entre 10 min et 30 min, correspond à l'élution de tous les métabolites apolaires qui interagissent avec la colonne chromatographique. L'analyse métabolomique est effectuée dans cette zone.
- La troisième zone, au delà de 30 min, où le courant total est quasiment nul. Tous les métabolites ont été élués du chromatographe. On note cependant une brève activité au environ de 55 min, cette activité correspond à un bruit résiduel.

Nous nous restreignons, pour la séparation, à l'intervalle 12,4821 min - 21,1015 min. La figure 5.3 montre le chromatogramme de la zone sélectionnée. 956 spectres de masse sont mesurés sur cet intervalle. Par ailleurs, un métabolite donné est élué en une seule fois du chromatographe pendant un laps de temps de l'ordre de 30 secondes. Nous pouvons ainsi effectuer la séparation de sources par fenêtre temporelle. Prendre des fenêtres temporelles permet de diminuer le volume des données sans rompre le modèle de mélange linéaire. Les spectres de masse à analyser sont donc subdivisés en 4 groupes correspondant à 4 fenêtres temporelles : [12,4821 min - 14,4178 min], [14,4178 min - 16,2441 min], [16,2441 min - 18,6184 min] et [18,6184 min - 21,1015 min].

### 5.2.3 Résultats de séparation des spectres de masse

A l'aide d'outils comme XCMS [SWO<sup>+</sup>06] ou MZmine [PCVBO10], la méthode traditionnelle, utilisée par les biologistes du LEMM pour l'analyse métabolomique, consiste à effectuer une recherche de pics dans le plan masse - temps de rétention, puis à grouper les pics trouvés suivant le temps de rétention. Cependant, ces méthodes de groupement a posteriori -fondée sur des calculs de corrélation temporelle- sont souvent prises en défaut. Les pics d'un même métabolites ne sont donc pas toujours identifiés comme provenant de la même molécule, en particulier en cas de chevauchement des profils d'élution, ce qui complique l'interprétation

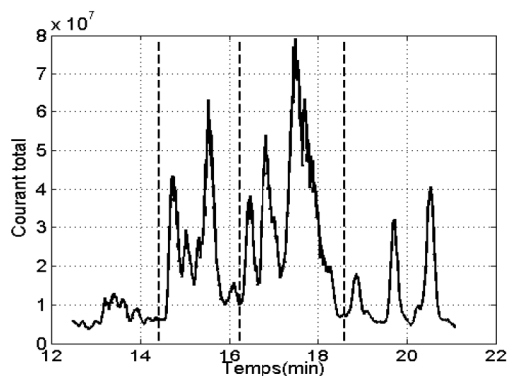


FIGURE 5.3 – Chromatogramme sur l'intervalle 12,4821 min - 21,5337 min

des données LC-MS. L'avantage de la séparation de sources c'est qu'il opère ce groupement a priori, au sein même du modèle de mélange linéaire instantané et fournit des données structurées par métabolites. Nous montrons sur les figures ci-dessous, les résultats de séparation obtenus par la méthode SCSSA-UNS. Ces résultats sont confrontés à ceux obtenus par les biologistes du LEMM et annotés dans l'annexe C.

Les paramètres de l'algorithme sont fixés à  $\epsilon = 10^{-6}$ ,  $\gamma = \frac{1}{4}$  et  $\eta = 10^{-1}$ . Sur chaque fenêtre temporelle, le seuil de bruit  $\lambda_b$  et le nombre de sources  $n$  sont optimisés manuellement de sorte à retrouver au moins tous les métabolites annotés par les biologistes dans cette fenêtre.

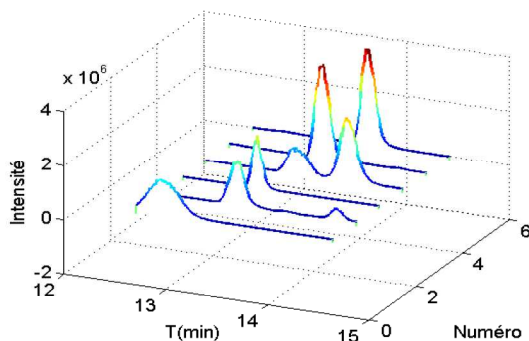


FIGURE 5.4 – Profils d'élution des 6 sources estimées sur l'intervalle 12,4821 min - 14,4178 min

### Résultats de séparation sur la fenêtre 12,4821 min - 14,4178 min

Le nombre de spectres mesurés dans cet intervalle est  $m = 206$ . Nous fixons  $n = 6$  et  $\lambda_b = 0.3$ . La figure 5.4 montre les profils d'élution des 6 sources estimées et la figure 5.5 montre les profils d'élution des deux métabolites annotés par les biologistes, complétés par les spectres de masse correspondants. Nous avons en outre extrait 4 autres métabolites qui n'avaient pas été identifiés par les biologistes, les profils d'élution de ces métabolites sont présentés sur la figure 5.4.



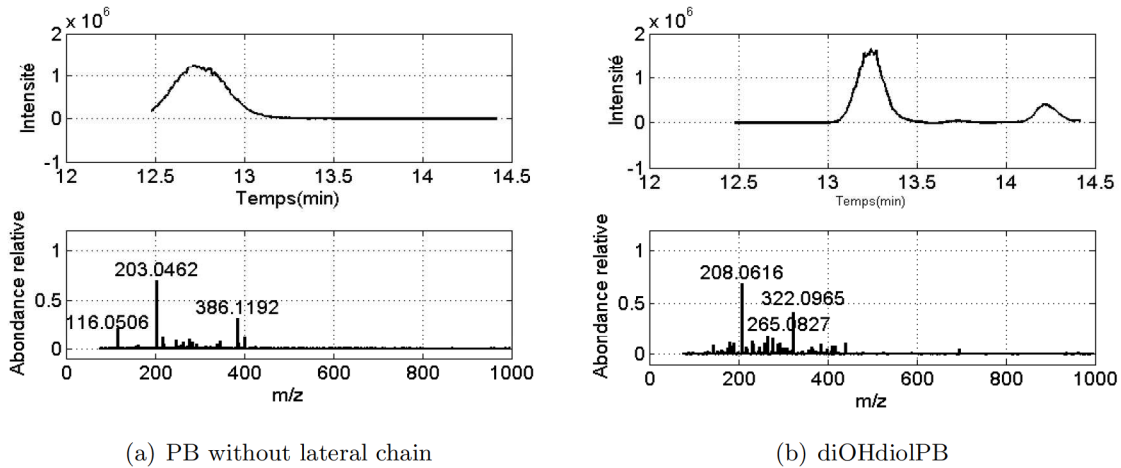


FIGURE 5.5 – Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 12,4821 min -14,4178 min

### Résultats de séparation sur la fenêtre 14,4178 min - 16,2441 min

La fenêtre temporelle 14,4178 min - 16,2441 min comprend 208 spectres. En optimisant le nombre de sources et le seuil de bruit, nous avons trouvé que les valeurs optimales sont respectivement  $n = 8$  et  $\lambda_b = 0.1$ . La figure 5.6 montre les profils d'élution des 8 sources estimées et la figure 5.7 montre les profils d'élution et les spectres de masse des 5 métabolites annotés par les biologistes. Nous pouvons voir que les 5 métabolites annotés, sur cet intervalle, sont correctement retrouvés ; nous avons en plus mis en évidence 3 autres métabolites, dont les profils d'élution sont présentés sur la figure 5.6, qui n'avaient été identifiés par les biologistes.

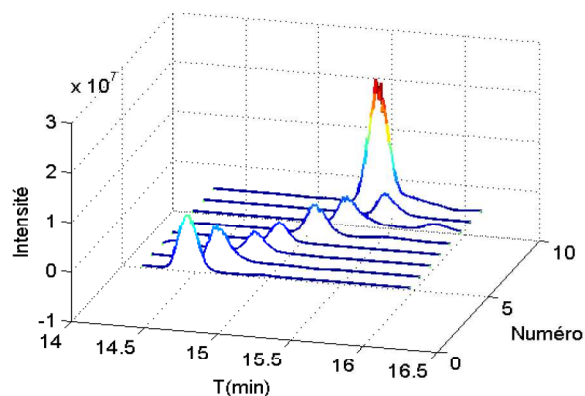
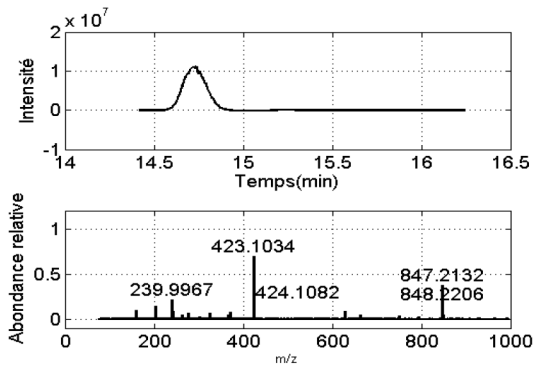


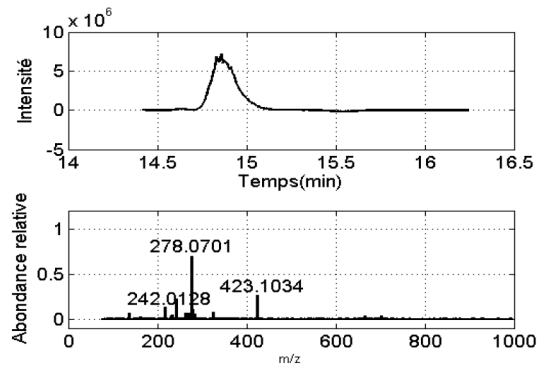
FIGURE 5.6 – Profils d'élution des 8 sources estimées sur l'intervalle 14,4178 min - 16,2441 min

### Résultats de séparation sur la fenêtre 16,2441 min - 18,6184 min

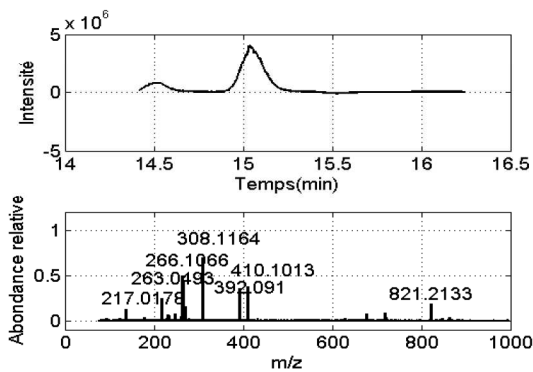
Dans cette fenêtre temporelle, 263 spectres de masse ont été mesurés. En suivant le même raisonnement que précédemment, nous avons trouvé que les paramètres optimaux sont  $n = 10$



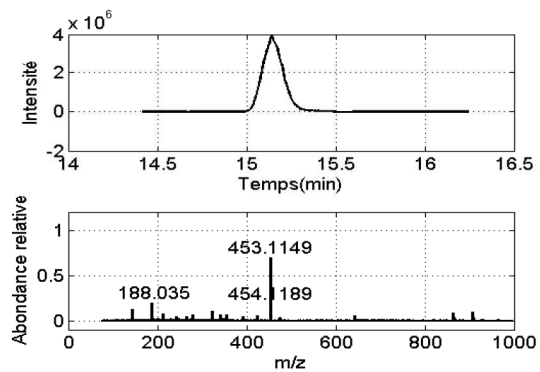
(a) OH-PB-O-Glucuronide



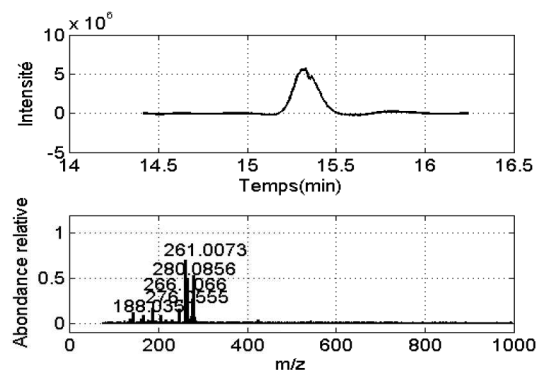
(b) PB-O-Glucuro-N-acetylhexopyranosylamine



(c) PB-mercaptopuric-acid



(d) O-Methyl-PB-O-Glucuronide



(e) diOH-PB

FIGURE 5.7 – Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 14,4178 min - 16,2441 min

et  $\lambda_b = 0.1$ . La figure 5.8 montre les profils d'élution des 10 sources estimées. Les spectres de masse et les profils d'élution des 6 métabolites annotés par les biologistes sont présentés sur la figure 5.9. Comme sur les fenêtres temporelles précédentes, nous avons extrait en plus 4 métabolites qui n'avaient pas été identifiés par les biologistes. Par ailleurs, nous voyons sur la figure 5.9.e que les métabolites PB-N-Glucuronide-1 et PB-N-Glucuronide-2 qui ont des spectres de masse identiques (voir annexe C), ont été regroupés dans la même composante, dont le profil d'élution présente une double bosse, correspondant aux temps de rétention de ces métabolites.

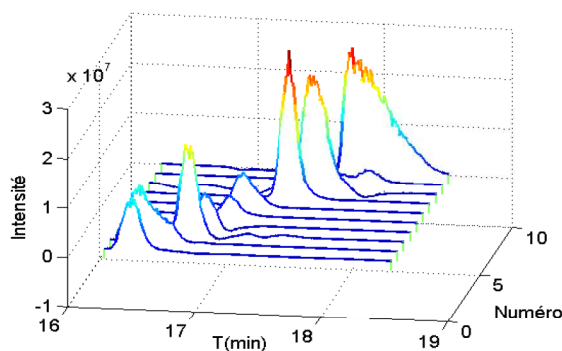


FIGURE 5.8 – Profils d'élution des 10 sources estimées sur l'intervalle 16,2441 min - 18,6184 min

### Résultats de séparation sur la fenêtre 18,6184 min - 21,5337 min

282 spectres de masse sont mesurés dans l'intervalle 18,6184 min - 21,5337 min. Nous fixons le nombre de sources à  $n = 8$  et le seuil de bruit à  $\lambda_b = 0.2$ . La figure 5.10 montre les profils d'élution des 8 sources estimées. Sur cette fenêtre temporelle aussi, les 2 métabolites annotés sont bien retrouvés, leurs profils d'élution et leurs spectres de masse sont présentés sur la figure 5.11. Nous avons également extrait 6 composantes qui n'avaient pas encore été identifiées.

### Synthèse des résultats de séparation sur données LC-MS

La séparation par SCSA-UNS des spectres de masse, issus de l'analyse LC-MS d'urine de rat traité au phénobarbital, a permis de retrouver tous les métabolites annotés par les biologistes. De plus, nous avons mis en évidence d'autres métabolites qui n'avaient pas encore été identifiés par les biologistes, ces profils et spectres sont en cours d'analyse. Par ailleurs, contrairement à ce qui était attendu, le profil d'élution de certains métabolites (annotés ou pas) présentent deux bosses plus ou moins éloignées dans le temps. Une double bosse correspond à la présence d'isomères, c'est-à-dire des molécules ayant la même formule brute (donc le même spectre de masse), mais des formules développées différentes. La formule développée exprime l'occupation de l'espace par la molécule, elle constitue donc une propriété physique de la molécule, qui agit sur son temps de rétention dans la colonne chromatographique. Deux isomères du même métabolite auront donc tendance à sortir du chromatographe pendant des

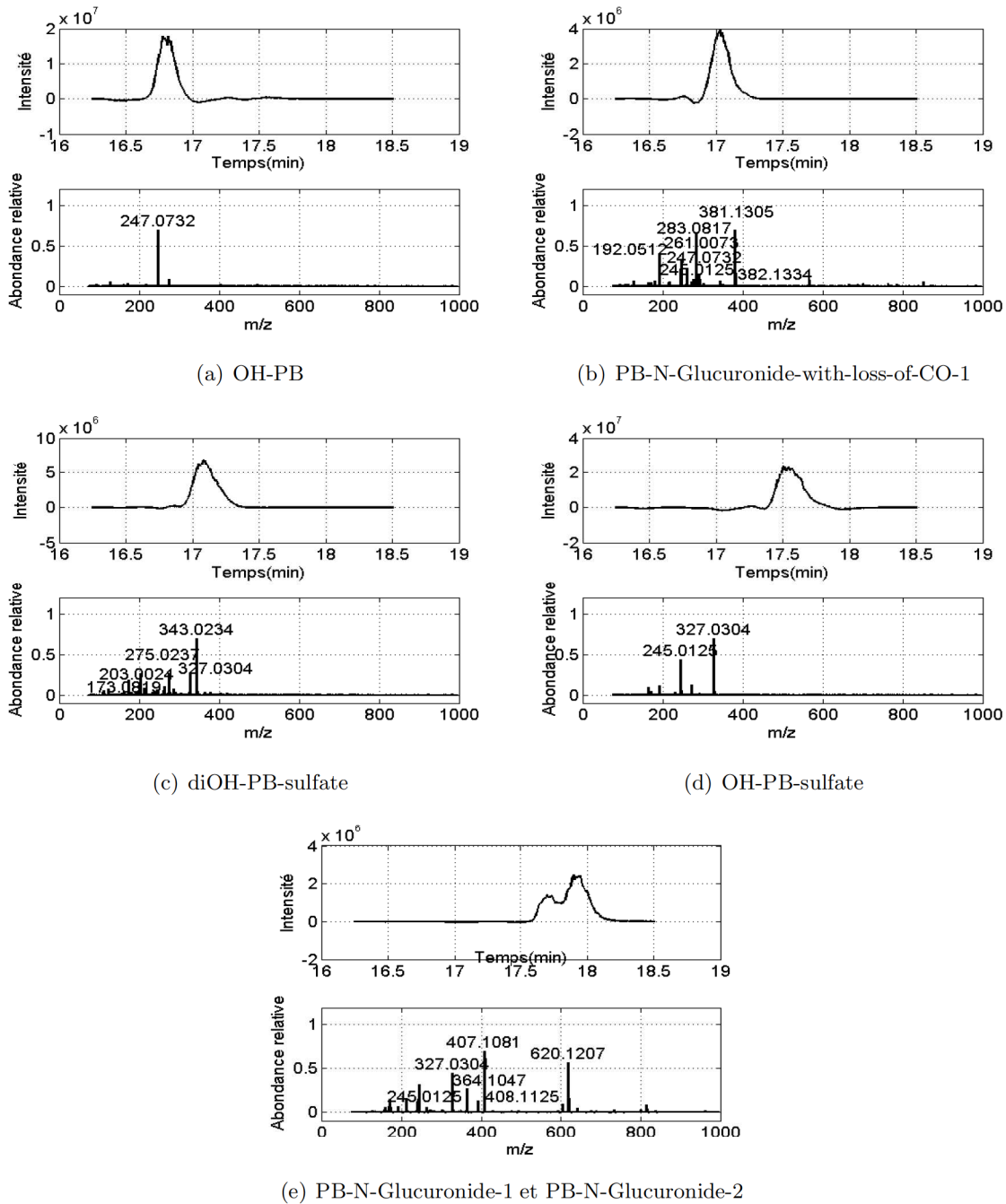


FIGURE 5.9 – Profils d'élution et spectres de masse estimés des métabolites annotés sur l'intervalle 16,2441 min - 18,6184 min

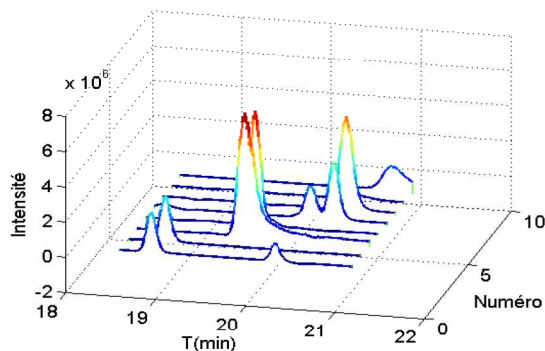


FIGURE 5.10 – Profils d'élu-tion des 8 sources estimées sur l'intervalle 18,6184 min - 21,5337 min

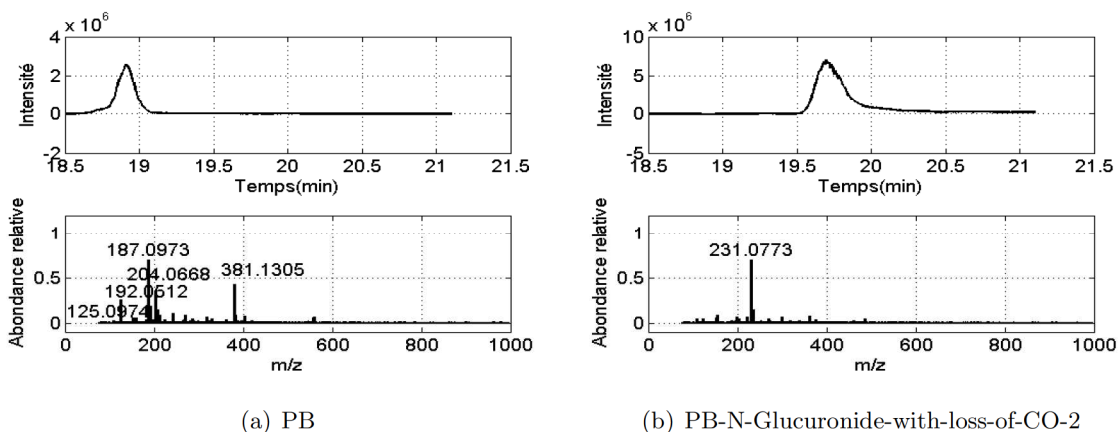


FIGURE 5.11 – Profils d'élu-tion et spectres de masse estimés des métabolites annotés sur l'intervalle 18,6184 min - 21,5337 min

laps de temps différents. Les isomères sont donc logiquement regroupés dans une même composante par la méthode de séparation de sources SCSA-UNS.

Du point de vue séparation de sources, nous avons remarqué que, outre le seuil de bruit, les résultats de séparation sont assez sensibles au nombre de sources. En effet, certains métabolites, annotés par les biologistes, ont des faibles concentrations, ce qui nous a amené, sur chaque fenêtre temporelle, à optimiser le nombre de sources de sorte à retrouver tous les métabolites annotés.

Pour finir, nous pouvons noter la présence de faibles valeurs négatives dans les profils d'élu-tion de certains métabolites. Nous avons attribué cela à une réduction insuffisante du bruit. Des travaux récents [HFR<sup>+</sup>09][BMD<sup>+</sup>11] ont montré qu'en spectrométrie de masse, le bruit semble avoir une composante additive et une composante multiplicative, or la méthode de débruitage que nous avons proposé ne traite que le bruit additif.

## 5.3 Estimation des pharmacocinétiques du [18F]-FDG dans le cerveau

Nous évaluons dans cette section, la capacité de notre méthode à estimer les pharmacocinétiques du traceur [18F]-FDG dans le cerveau humain, à partir d'ensembles d'images dynamiques TEP de cet organe. En particulier, nous cherchons à estimer la fonction d'entrée artérielle (AIF), du fait de son utilité en oncologie (voir chapitre 1). Les données nous sont fournies par le Service Hospitalier Frédéric Joliot (SHFJ), de l'Institut d'Imagerie Biomédicale (I<sup>2</sup>BM), de la Direction des Sciences du Vivant (DSV) du Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA-EA). Deux types de données sont traitées : des images synthétiques TEP de fantômes de cerveau humain et des images réelles TEP de cerveaux humains.

### 5.3.1 Modèle de mélange des images TEP

Nous considérons que le cerveau est formé de différents compartiments pharmacocinétiques. Rappelons qu'un compartiment pharmacocinétique du cerveau (comme par exemple la matière blanche, la matière grise, le réseau artériel, le réseau veineux, ...) est une partie du cerveau dans laquelle la concentration du traceur peut être considérée comme homogène. La contribution d'un compartiment à une image TEP du cerveau reconstruite à un instant  $t$ , peut se factoriser comme le produit de la concentration du traceur dans ce compartiment à cet instant, par la fraction volumique de chaque voxel de l'image occupée par le compartiment. Ainsi, chaque image TEP reconstruite peut être modélisée comme le mélange linéaire instantané des images TEP des différents compartiments pharmacocinétiques [TS08][ZFFM<sup>+</sup>09]. Le modèle de mélange des images TEP est ainsi donné par l'équation (5.2) :

$$X(x, y, z, t) = \sum_{i=1}^n a_i(t) S_i(x, y, z) \quad (5.2)$$

où  $X(x, y, z, t)$  est le voxel à la position  $(x, y, z)$  de l'image TEP reconstruite à l'instant  $t$ ,  $a_i(t)$  est la concentration du traceur dans le compartiment  $i$  et à l'instant  $t$ , et  $S_i(x, y, z)$  est la fraction volumique du voxel de position  $(x, y, z)$  occupée par le compartiment  $i$ . Nous cherchons à estimer la distribution spatiale ainsi que la cinétique (évolution temporelle de la concentration) du traceur dans chaque compartiment.

### 5.3.2 Résultats de séparation sur images TEP de fantômes

Pour cette première évaluation 26 images 3D TEP de fantôme, de taille  $128 \times 128 \times 63$ , ont été générées à l'aide d'un simulateur numérique [ZFFM<sup>+</sup>09]. Ce fantôme (non-réaliste) comporte plusieurs tissus cérébraux (matière blanche, matière grise, cervelet, etc...), deux courts cylindres (de diamètre 8mm et de hauteur 25mm) représentant deux morceaux d'artères carotides, ainsi que le crâne ; ce fantôme ne comporte pas de compartiment veineux. Les images dynamiques TEP ont été calculées par combinaisons linéaires des images des

différents compartiments du fantôme, pondérés par leurs cinétiques associées. Plus précisément, pour tenir compte de l'influence de la reconstruction, les sinogrammes des différents compartiments ont été combinés linéairement et bruités, puis les images ont été reconstruites par un algorithme standard<sup>7</sup> de reconstruction TEP. Les durées d'échantillonnage des images suivent celles couramment utilisées pour les études cliniques et sont indiquées dans le tableau 5.1 ; la durée totale de l'expérience est de 74 minutes.

| Numéro de l'image | Durée(s) |
|-------------------|----------|
| 1                 | 20       |
| 2 → 11            | 10       |
| 12 → 13           | 30       |
| 14 → 15           | 150      |
| 16 → 19           | 300      |
| 20                | 1200     |
| 21                | 120      |
| 22                | 300      |
| 23                | 120      |
| 24                | 600      |
| 25                | 120      |
| 26                | 300      |

TABLE 5.1 – Durée d'échantillonnage des images TEP de fantômes

La figure 5.12 montre les coupes sagittale, coronale et axiale d'une image sur lesquelles on distingue la forme du cerveau.

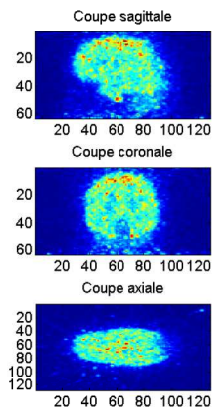


FIGURE 5.12 – Coupes sagittale, coronale et axiale d'une image TEP de fantôme

Les images 3D initiales sont réduites suivant les 3 axes pour ne considérer que les voxels présentant une activité, on obtient ainsi des images de tailles  $61 \times 81 \times 60$ . L'objectif principal étant de séparer le compartiment artériel des autres compartiments du cerveau, nous avons

7. OSEM : Ordered Set Exception Maximization

fait varier le nombre de sources à partir de  $n = 2$ , et nous avons optimisé manuellement le seuil de bruit  $\lambda_b$ . Les meilleurs résultats sont obtenus pour  $n = 3$  et  $\lambda_b = 0.3$ . La figure 5.13 montre les vues suivant les coupes (sagittale, coronale et axiale) des compartiments pharmacocinétiques estimés avec l'algorithme SCSA-UNS. Nous pouvons constater sur cette figure que trois différents compartiments pharmacocinétiques ont été nettement séparés. Ces compartiments correspondent aux deux tronçons de carotides (figure 5.13.a), au cervelet (figure 5.13.c) et au reste du tissu cérébral. De plus, la figure 5.13.b montre que la cinétique artérielle estimée (AIF) par SCSA-UNS suit correctement la cinétique artérielle utilisée pour la génération des images TEP (rAIF) ; ce qui était l'objectif principal de la séparation.

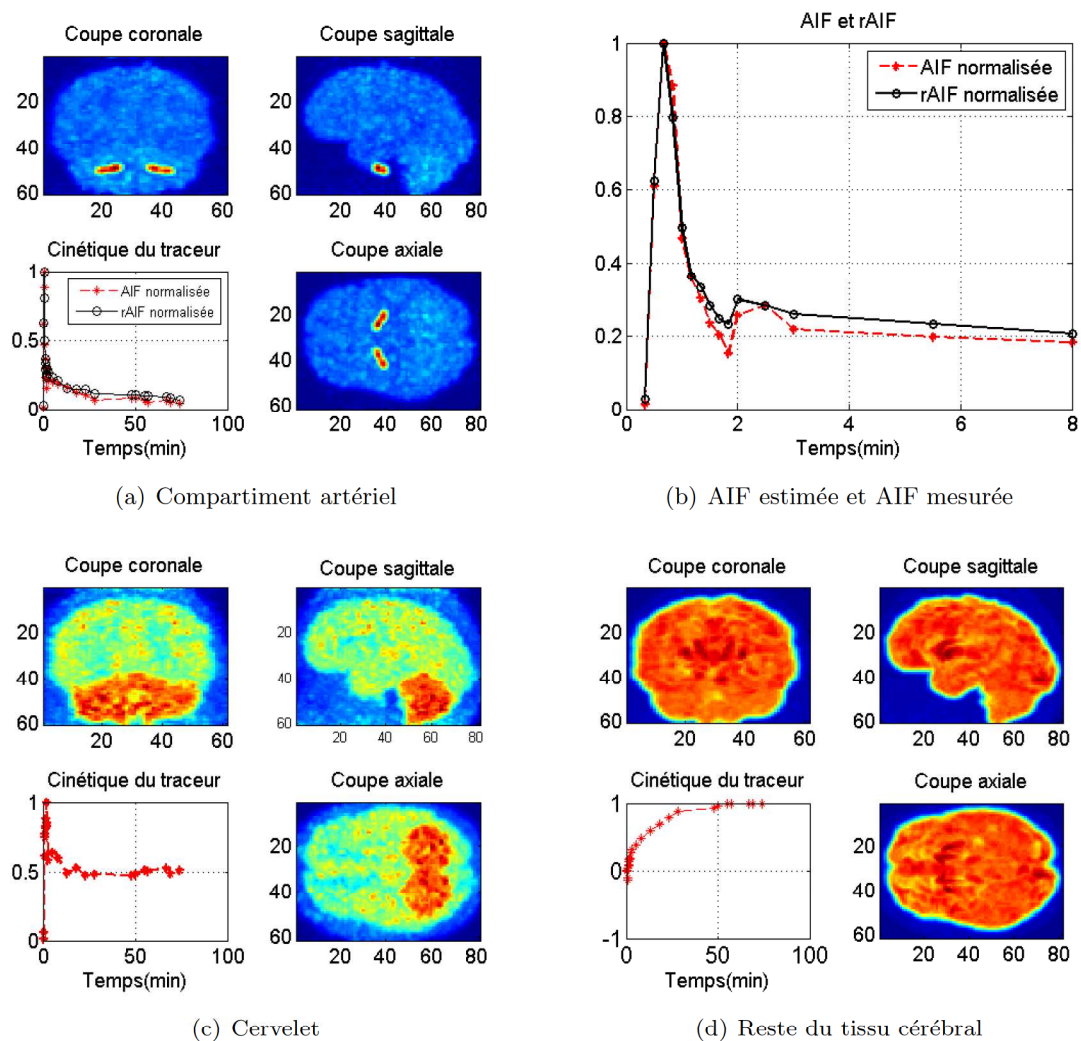


FIGURE 5.13 – Compartiments pharmacocinétiques estimés à partir d'images TEP de fantômes

### 5.3.3 Résultats de séparation sur images TEP réelles de cerveau humain

Dans cette expérience 21 images TEP de tailles  $128 \times 128 \times 63$  sont enregistrées pendant 71 min par un scanner TEP *Siemens ECAT HR plus*. Les 19 premières images ont été enregistrées dans les 25 minutes après l'injection du traceur. Les deux dernières images



ont été enregistrées, respectivement 53 et 71 minutes après l'injection du traceur. Les durées d'acquisition des différentes images sont consignées dans le tableau 5.2. Il faut préciser que

| Numéro de l'image | Durée(s) |
|-------------------|----------|
| 1                 | 20       |
| 2 → 11            | 10       |
| 12 → 13           | 30       |
| 14 → 15           | 150      |
| 16 → 21           | 300      |

TABLE 5.2 – Durée d'acquisition des images TEP de cerveau humain

pendant les acquisitions des images, une série de prélèvement sanguin artériel a été effectuée à l'aide d'un catheter. Ces prélèvements ont servi à construire une référence de la cinétique artérielle du traceur (rAIF).

Pour effectuer la séparation, nous ne considérons que les 19 premières images car les activités ne varient plus sur les 2 dernières images. Comme dans le cas des images de fantôme, les images initiales sont recadrées, pour ne considérer que les voxels présentant une activité. Les images ainsi réduites sont de tailles  $61 \times 67 \times 54$ . L'objectif ici aussi est de séparer le compartiment artériel des autres compartiments. Le nombre de sources et le seuil de bruit optimaux que nous avons trouvés sont respectivement  $n = 3$  et  $\lambda_b = 0.3$ . La figure 5.14 montre les vues suivant les trois coupes (sagittale, coronale et axiale) des compartiments pharmacocinétiques estimés par l'algorithme SCSA-UNS. La figure 5.14 montre que les différents compartiments pharmacocinétiques sont correctement séparés et identifiables. De plus la figure 5.14.b montre que la cinétique artérielle estimée par SCSA-UNS (AIF) suit correctement l'allure de la cinétique artérielle obtenue par prélèvement sanguin (rAIF). On constate cependant un décalage temporel entre les deux cinétiques artérielles (estimée par SCSA et obtenue par prélèvements sanguins), dû à des erreurs de datation des images TEP et/ou des instants des prélèvements sanguins. Il faut noter par ailleurs, la difficulté à synchroniser une image acquise sur 10 secondes, 2 minutes ou 5 minutes et un prélèvement artériel qui ne dure que quelques secondes.

## 5.4 Conclusion

Nous avons présenté dans ce chapitre les résultats de séparation par l'algorithme SCSA-UNS de différents jeux de données (LC-MS et TEP). D'une part l'application de l'algorithme à des spectres de masse issus de l'analyse LC-MS d'urine de rat traité au phénobarbital a permis de retrouver tous les métabolites annotés par les biologistes. Nous avons en outre mis en évidence d'autres métabolites qui n'avaient pas encore été identifiés. D'autre part, la séparation d'images dynamiques TEP a permis de séparer correctement les différents compartiments pharmacocinétiques du  $[18F]$ -FDG à la fois sur fantôme et sur patient humain. En particulier, nous avons montré que la cinétique artérielle estimée par SCSA-UNS suit correctement la cinétique artérielle obtenue par prélèvements sanguins. Ces résultats illustrent

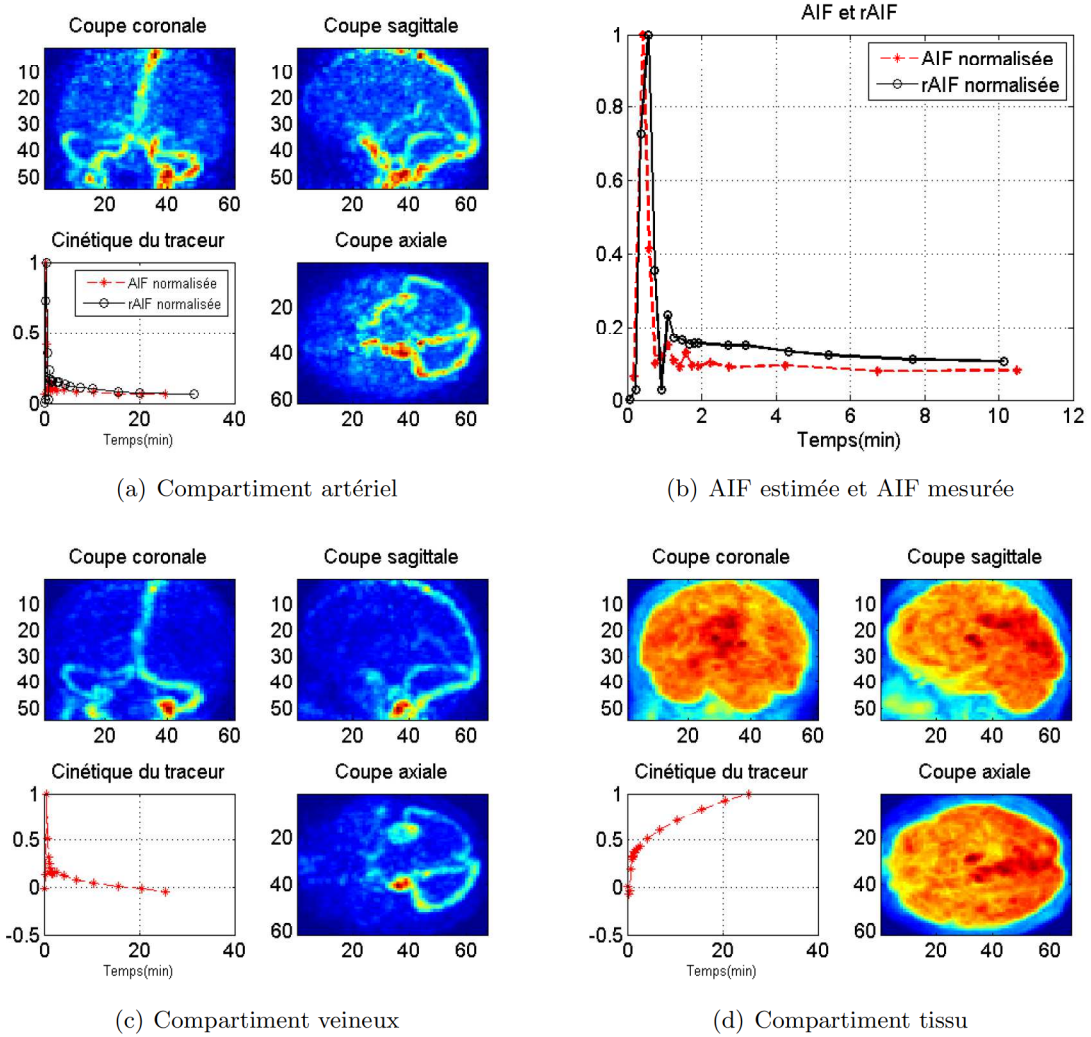


FIGURE 5.14 – Compartiments pharmacocinétiques estimés à partir d’images TEP du cerveau humain

l’efficacité de l’algorithme SCSA-UNS et sa capacité à traiter des problèmes réels de natures très différentes.



# Conclusion générale et perspectives

Ce travail de thèse a permis de développer une nouvelle méthode géométrique de séparation de sources non-négatives. La méthode développée est utilisable dans différents problèmes, en particulier pour la séparation de spectres de masse en vue de l'étude métabolomique de fluide biologique, et pour la séparation des compartiments pharmacocinétiques d'un radio-traceur dans un organe.

La première partie de notre travail a consisté à présenter l'état de l'art des méthodes existantes de séparation de sources non-négatives, à travers une analyse critique de ces méthodes. Nous avons ainsi pu établir que, parmi les approches statistiques, la méthode d'analyse en composantes indépendantes non-négatives n'est pas adaptée aux sources corrélées et les sources estimées présentent des valeurs négatives non négligeables. La méthode bayésienne nécessite de choisir des lois a priori pour tous les paramètres du modèle de mélange (sources, coefficients de mélange et bruit), ce choix est néanmoins assez flexible. Cette approche possède cependant un coût de calcul extrêmement élevé qui constitue une limite majeure à son utilisation dans des problèmes de grandes dimensions. Parmi les approches déterministes, les méthodes de factorisation en matrices non-négatives (sans ou avec régularisations) ne garantissent en général pas l'unicité de la solution de séparation ; de ce fait, leurs performances sont assez sensibles à l'initialisation. En ce qui concerne les méthodes géométriques, nous avons établi que de manière générale, leurs performances se dégradent fortement en présence de bruit, en particulier lorsque les sources sont parcimonieuses. De plus les méthodes géométriques fondées sur la recherche des points extrêmes du nuage de points des signaux observations, font l'hypothèse de la dominance locale des sources, hypothèse qui est rarement vérifiée en pratique. Quant aux méthodes géométriques fondées sur la recherche du simplexe de volume minimal contenant le nuage de points des signaux observations, elles font l'hypothèse de l'additivité totale des sources, hypothèse qui peut être levée en normalisant les observations ; mais avec le risque d'amplifier le bruit, surtout quand les sources ont de fortes dynamiques.

A la synthèse des résultats notre analyse, il est ressorti que les méthodes géométriques fondées sur la recherche du simplexe de volume minimal contenant le nuage des observations présentent l'avantage de ne nécessiter que de très "faibles" hypothèses sur les signaux sources (ni l'indépendance, ni la dominance locale des sources ne sont nécessaires), de réaliser efficacement la séparation pour un faible coût de calcul. Ce qui nous a motivé proposer une nouvelle méthode géométrique de séparation de sources non-négative fondée sur les cônes

simpliciaux, cette méthode a fait l'objet de la deuxième partie.

Dans la deuxième partie de ce mémoire, nous avons proposé une nouvelle méthode géométrique de séparation de sources non-négatives fondée sur les cônes simpliciaux, baptisée SCSA-UNS pour *Simplicial Cone Shrinking Algorithm for Unmixing Non-negative Sources*. En absence de bruit, la méthode SCSA-UNS estime la matrice de mélange et les sources en cherchant le cône simplicial d'ouverture minimale qui contient le nuage de points des signaux observations. La recherche de ce cône se fait par alternance de réduction de l'ouverture d'un cône initial contenant la totalité du nuage de points des signaux observations et de déblocage. En présence d'un bruit additif, la méthode commence par une étape de débruitage qui consiste à éliminer dans les observations, celles de norme inférieure à un certain seuil à fixer. La méthode SCSA-UNS ne nécessite pas l'indépendance mutuelle des sources, ni même leur décorrélation ; elle ne requiert pas non plus la dominance locale des sources, ni leur additivité totale. Une seule condition est nécessaire et suffisante : l'orthant positif doit être l'unique cône simplicial d'ouverture minimale contenant le nuage de points des signaux sources. Nous avons établi des conditions nécessaires et des conditions suffisantes pour vérifier cette contrainte. Nous avons également conjecturé une condition nécessaire et suffisante, qui mérite d'être démontrée. Du point de vue performances de séparation, nous avons montré, à travers des simulations sur données synthétiques, que l'algorithme SCSA-UNS présente de meilleures performances de séparation comparativement aux autres méthodes existantes, en particulier lorsque les sources sont parcimonieuses et/ou le bruit est faible. Cependant, une erreur sur la position du sommet cône formé le nuage de points des signaux observations peut entraîner une dégradation significative des performances de séparation. La méthode SCSA-UNS admet néanmoins un vaste champ d'applications.

C'est l'objet du troisième volet de notre travail qui évalue la méthode SCSA-UNS dans différentes situations réelles de séparation de sources non-négatives.

Dans la première situation, nous avons effectué la séparation de spectres de masse, mesurés à la sortie d'un chromatographe en phase liquide, en vue de l'étude métabolomique de fluides biologiques. Deux jeux de données ont été traités. Dans le premier cas, il s'est agi d'un mélange synthétique de composés chimiques réels, et dans le second cas, il s'est agi d'urine de rat traité au phénobarbital. Pour ces deux jeux de données, nous avons pu correctement extraire, à l'aide de la méthode SCSA-UNS, tous les métabolites annotés par les biologistes. De plus, nous avons mis en évidence d'autres métabolites qui n'avaient pas encore été identifiés. Dans la deuxième situation, nous avons correctement séparé, avec notre algorithme, les différents compartiments pharmacocinétiques du radio-traceur [18F]-FDG, dans le cerveau humain à partir d'images dynamiques de tomographie par émission de positons (TEP) de cet organe. En particulier, nous avons correctement estimé la cinétique artérielle du traceur à partir uniquement des images TEP, ce qui permet d'éviter les prélèvements sanguins artériels, très invasifs et potentiellement dangereux. Cette cinétique présente un grand intérêt en oncologie, pour l'évaluation de l'efficacité d'un traitement anti-cancéreux.

Certains points de ce travail peuvent être approfondis afin d'améliorer l'efficacité de l'algorithme SC-SA-UNS et/ou le rendre totalement autonome. Ces points sont présentés dans les perspectives ci-dessous.

## Perspectives

Une première perspective, pour automatiser complètement la méthode SC-SA-UNS, serait de chercher une méthode de détermination automatique du seuil de bruit. En effet, les performances de l'algorithme en présence de bruit additif dépendent de ce seuil. Les grandeurs qui peuvent être considérées pour la détermination de ce seuil sont le niveau de bruit et le taux de parcimonie des sources.

Un deuxième axe d'amélioration de la méthode SC-SA-UNS, serait de mettre en place une méthode d'estimation du sommet du cône formé par le nuage de points des signaux observations. Ceci, afin de minimiser les erreurs d'estimation de la matrice de mélange, liée à une erreur concernant la position du sommet du cône formé par les observations. Une méthode efficace et robuste d'estimation du sommet du cône permettrait de considérer le cas, plus général, de sources non plus positives mais distribuées sur une demi-droite. En effet, on pourra toujours se ramener aux cas de sources non-négatives en ramenant le sommet estimé à l'origine de l'espace.

Pour finir, il faudrait effectuer une modélisation plus réaliste du bruit, dans le cas de sources et de mélanges non-négatifs. En effet, dans certaines applications, comme en spectrométrie de masse, le bruit possède une composante multiplicative, en plus de sa composante additive. Cet aspect doit être pris en compte afin d'avoir un modèle de données qui colle plus avec la réalité des mesures.



# Annexes



## Annexe A

# Preuve de la non séparabilité par N-ICA de sources non-négatives, décorrélées et *well-grounded*

Nous considérons les sources  $s_1$ ,  $s_2$  et  $s_3$  définies par :

$$s_1 = 1 + \cos(\varphi) \tag{A.1}$$

$$s_2 = 1 + \sin(\theta) \sin(\varphi) \tag{A.2}$$

$$s_3 = 1 + \cos(\theta) \sin(\varphi) \tag{A.3}$$

où  $\theta$  est uniformément repartie sur  $[0, 2\pi]$ ,  $v$  est uniformément repartie sur  $[0, 1]$  et  $\varphi = \cos^{-1}(2v - 1)$ .

La distribution jointe des sources est uniformément répartie sur la sphère  $\mathcal{C}(\mathbf{c}, 1)$  de centre  $\mathbf{c} = [1 \ 1 \ 1]^T$  et de rayon 1. Leurs valeurs moyennes  $\mu_{s_1}$ ,  $\mu_{s_2}$  et  $\mu_{s_3}$  sont données par :

$$\mu_{s_1} = 1 + \int_0^1 (2v - 1) dv = \int_0^1 2v dv = 1 \tag{A.4}$$

$$\mu_{s_2} = 1 + \int_0^{2\pi} \int_0^1 \sin(\theta) \sin(\cos^{-1}(2v - 1)) \frac{d\theta}{2\pi} dv \tag{A.5}$$

$$= 1 + \frac{1}{2\pi} \int_0^{2\pi} \sin(\theta) d\theta \int_0^1 \sin(\cos^{-1}(2v - 1)) dv = 1 \tag{A.6}$$

$$\mu_{s_3} = 1 + \int_0^{2\pi} \int_0^1 \cos(\theta) \sin(\cos^{-1}(2v - 1)) \frac{d\theta}{2\pi} dv \tag{A.7}$$

$$= 1 + \frac{1}{2\pi} \int_0^{2\pi} \cos(\theta) d\theta \int_0^1 \sin(\cos^{-1}(2v - 1)) dv = 1 \tag{A.8}$$

Les covariances des variables  $s_1$ ,  $s_1$  et  $s_1$  sont données par :

$$E[(s_1 - \mu_{s_1})(s_2 - \mu_{s_2})] = \int_0^{2\pi} \int_0^1 \sin(\theta) (2v - 1) \sin(\cos^{-1}(2v - 1)) \frac{d\theta}{2\pi} dv \quad (\text{A.9})$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \sin(\theta) d\theta \int_0^1 (2v - 1) \sin(\cos^{-1}(2v - 1)) dv \quad (\text{A.10})$$

$$= 0 \quad (\text{A.11})$$

$$E[(s_1 - \mu_{s_1})(s_3 - \mu_{s_3})] = \int_0^{2\pi} \int_0^1 \cos(\theta) (2v - 1) \sin(\cos^{-1}(2v - 1)) \frac{d\theta}{2\pi} dv \quad (\text{A.12})$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \cos(\theta) d\theta \int_0^1 (2v - 1) \sin(\cos^{-1}(2v - 1)) dv \quad (\text{A.13})$$

$$= 0 \quad (\text{A.14})$$

$$E[(s_2 - \mu_{s_2})(s_3 - \mu_{s_3})] = \int_0^{2\pi} \int_0^1 \sin(\theta) \cos(\theta) [\sin(\cos^{-1}(2v - 1))]^2 \frac{d\theta}{2\pi} dv \quad (\text{A.15})$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \sin(\theta) \cos(\theta) d\theta \int_0^1 [\sin(\cos^{-1}(2v - 1))]^2 dv \quad (\text{A.16})$$

$$= 0 \quad (\text{A.17})$$

Les variables  $s_1$ ,  $s_2$ ,  $s_3$  sont donc décorrélées. Elle ne sont cependant pas indépendantes car :

$$[s_1 - 1]^2 + [s_2 - 1]^2 + [s_3 - 1]^2 = 1 \quad (\text{A.18})$$

Un point  $\mathbf{s} = [s_1 \ s_2 \ s_3]^T$  est située sur  $\mathcal{C}(\mathbf{c}, 1)$  si :

$$\|\mathbf{s} - \mathbf{c}\|^2 = 1 \quad (\text{A.19})$$

Soit  $\mathbf{R}(\phi)$  la rotation d'angle  $\phi$  autour de l'axe dirigé par le vecteur  $\mathbf{u} = [\frac{1}{\sqrt{3}} \ \frac{1}{\sqrt{3}} \ \frac{1}{\sqrt{3}}]^T$ , et dont la matrice est donnée par :

$$\mathbf{R}(\phi) = \frac{1}{3} \begin{pmatrix} 1 + 2 \cos(\phi) & 1 - \cos(\phi) - \sqrt{3} \sin(\phi) & 1 - \cos(\phi) + \sqrt{3} \sin(\phi) \\ 1 - \cos(\phi) + \sqrt{3} \sin(\phi) & 1 + 2 \cos(\phi) & 1 - \cos(\phi) - \sqrt{3} \sin(\phi) \\ 1 - \cos(\phi) - \sqrt{3} \sin(\phi) & 1 - \cos(\phi) + \sqrt{3} \sin(\phi) & 1 + 2 \cos(\phi) \end{pmatrix} \quad (\text{A.20})$$

Si  $\mathbf{s} \in \mathcal{C}(\mathbf{c}, 1)$  alors  $\forall 0 \leq \phi \leq 2\pi$ ,  $\mathbf{R}(\phi)\mathbf{s} \in \mathcal{C}(\mathbf{c}, 1)$ . En effet :

$$\|\mathbf{R}(\phi)\mathbf{s} - \mathbf{c}\|^2 = [\mathbf{R}(\phi)\mathbf{s} - \mathbf{c}]^T [\mathbf{R}(\phi)\mathbf{s} - \mathbf{c}] \quad (\text{A.21})$$

$$= [\mathbf{s}^T \mathbf{R}(\phi)^T - \mathbf{c}^T] [\mathbf{R}(\phi)\mathbf{s} - \mathbf{c}] \quad (\text{A.22})$$

$$= \mathbf{s}^T \mathbf{R}(\phi)^T \mathbf{R}(\phi)\mathbf{s} - 2\mathbf{c}^T \mathbf{R}(\phi)\mathbf{s} + \mathbf{c}^T \mathbf{c} \text{ car } \mathbf{c}^T \mathbf{R}(\phi) = \mathbf{c}^T \quad (\text{A.23})$$

$$= \mathbf{s}^T \mathbf{s} - 2\mathbf{c}^T \mathbf{s} + \mathbf{c}^T \mathbf{c} \quad (\text{A.24})$$

$$= \|\mathbf{s} - \mathbf{c}\|^2 = 1 \quad (\text{A.25})$$

La distribution des sources est donc invariant par  $\mathbf{R}(\phi)$ ,  $\forall 0 \leq \phi \leq 2\pi$ . Il s'en suit qu'il existe une infinité de bases dans lesquelles les sources  $s_1$ ,  $s_2$ ,  $s_3$  sont non-négatives, décorrelées et *well-grounded*. Leur mélange linéaire instantané n'est donc pas séparable par N-ICA.



## Annexe B

# Quelques mesures de divergence utilisées en NMF

Nous donnons dans cette annexe quelques mesures de divergence couramment utilisées dans les méthodes de factorisation en matrices non-négatives. Ces fonctions de coût sont étudiées plus en détails dans [CZPA09] et [Ber09], et des liens ont été établis entre elles.

### Distance Euclidienne

$$\mathcal{D}_{EUC}(\mathbf{X}, \mathbf{AS}) = \|\mathbf{X} - \mathbf{AS}\|_F^2 = \sum_{i=1}^m \sum_{k=1}^K \left[ x_i(k) - \sum_{j=1}^n a_{ij} s_j(k) \right]^2 \quad (\text{B.1})$$

### Divergence de Kullback-Leibler

$$\mathcal{D}_{KL}(\mathbf{X}, \mathbf{AS}) = \sum_{i=1}^m \sum_{k=1}^K \left[ x_i(k) \ln \left( \frac{x_i(k)}{\sum_{j=1}^n a_{ij} s_j(k)} \right) - x_i(k) + \sum_{j=1}^n a_{ij} s_j(k) \right] \quad (\text{B.2})$$

### Divergence d'Itakura-Saito

$$\mathcal{D}_{IS}(\mathbf{X}, \mathbf{AS}) = \sum_{i=1}^m \sum_{k=1}^K \left[ \ln \left( \frac{\sum_{j=1}^n a_{ij} s_j(k)}{x_i(k)} \right) + \frac{x_i(k)}{\sum_{j=1}^n a_{ij} s_j(k)} - 1 \right] \quad (\text{B.3})$$

### Beta divergence

$$\mathcal{D}_\beta(\mathbf{X}, \mathbf{AS}) = \sum_{i=1}^m \sum_{k=1}^K \left[ x_i(k) \frac{[x_i(k)]^\beta - [\sum_{j=1}^n a_{ij} s_j(k)]^\beta}{\beta} - \frac{[x_i(k)]^{\beta+1} - [\sum_{j=1}^n a_{ij} s_j(k)]^{\beta+1}}{\beta + 1} \right] \quad (\text{B.4})$$

### Alpha divergence

$$\mathcal{D}_\alpha(\mathbf{X}, \mathbf{AS}) = \frac{1}{\alpha(\alpha - 1)} \sum_{i=1}^m \sum_{k=1}^K \left[ [x_i(k)]^\alpha \left[ \sum_{j=1}^n a_{ij} s_j(k) \right]^{1-\alpha} - \alpha x_i(k) + (\alpha - 1) \left[ \sum_{j=1}^n a_{ij} s_j(k) \right] \right] \quad (\text{B.5})$$



## Annexe C

# Liste des métabolites annotés dans l'urine de rat

| Métabolite                              | Tr (min) | ID                                       | m/z      |
|---|----------|--|----------|
| PB_without_lateral_chain                | 12.66    | PB_without_lateral_chain                 | 203,0463 |
|   |          | PB_without_lateral_chain_13C             | 204,0491 |
|   |          | PB_without_lateral_chain_frag1           | 116,0512 |
| diOHdiolPB                              | 13.2     | diOHdiolPB                               | 265,083  |
|   |          | dihydrodiolPB_13C                        | 266,086  |
| OH-PB-O-Glucuronide                     | 14.65    | OH-PB-O-Glucuronide                      | 423,1047 |
|   |          | OH-PB-O-Glucuronide_13C                  | 424,1072 |
|   |          | OH-PB-O-Glucuronide_(2M-H)               | 847,2163 |
|   |          | OH-PB-O-Glucuronide_(2M-H)_13C           | 848,2195 |
|   |          | OH-PB-O-Glucuronide_(2M-H)_18O           | 849,2219 |
| PB-O-Glucuro-N-acetylhexopyranosylamine | 14.85    | PB-O-Glucuro-N-acetylhexopyranosylamine  | 626,1837 |
|   |          | PB-O-Gluc-N-acetylhexopyranosylamine_13C | 627,1874 |
| PB-mercapturic_acid                     | 15.01    | PB-mercapturic_acid                      | 410,1029 |
|   |          | PB-mercapturic_acid_(2M-H)               | 821,2128 |
|   |          | PB_mercapturic_acid_frag1                | 392,0912 |
|   |          | PB_mercapturic_acid_frag2                | 263,0497 |
| O-Methyl-PB-O-Glucuronide               | 15.1     | O-Methyl-PB-O-Glucuronide                | 453,1152 |
|   |          | O-methyl-PB-O-glucuronide_13C            | 454,1182 |
| diOH-PB                                 | 15.45    | diOH-PB                                  | 263,0675 |
| OH-PB                                   | 16.75    | OH-PB                                    | 247,0724 |
|   |          | OH-PB_13C                                | 248,0751 |
| OH-PB-O-Glucuronide_with_loss_of_CO     | 16.85    | OH-PB-O-Glucuronide_with_loss_of_CO      | 395,1097 |
|   |          | OH-PB-O-Gluc_with_loss_of_CO_13C         | 396,1127 |
| PB-N-Glucuronide_with_loss_of_CO_1      | 16.87    | PB-N-Glucuronide_with_loss_of_CO_1       | 381,1304 |
|   |          | PB-N-Gluc_with_a_loss_of_CO_1_13C        | 382,1333 |
|   |          | PB-N-Gluc_with_a_loss_of_CO_1_frag1      | 192,0515 |
| diOH-PB-sulfate                         | 16.99    | diOH-PB-sulfate                          | 343,0239 |
|   |          | diOH-PB-sulfate_frag2                    | 234,0282 |
| OH-PB-sulfate                           | 17.34    | OH-PB-sulfate                            | 327,0295 |
|   |          | OH-PB-sulfate_frag1                      | 247,0722 |
|   |          | OH-PB-sulfate_frag2                      | 204,0664 |
| PB-N-Glucuronide_1                      | 17.7     | PB-N-Glucuronide_1                       | 407,1095 |
|   |          | PB-N-Glucuronide_1_13C                   | 408,113  |
|   |          | PB-N-Glucuronide_1_frag1                 | 364,1049 |
|   |          | PB-N-Glucuronide_1_frag2                 | 174,0413 |
| PB-N-Glucuronide_2                      | 17.89    | PB-N-Glucuronide_2                       | 407,1095 |
|   |          | PB-N-Glucuronide_2_13C                   | 408,113  |
|   |          | PB-N-Glucuronide_2_frag1                 | 364,1049 |
|   |          | PB-N-Glucuronide_2_frag2                 | 174,0413 |
| PB-N-Glucuronide_with_loss_of_CO_2      | 18.95    | PB-N-Glucuronide_with_loss_of_CO_2       | 381,1304 |
|   |          | PB-N-Gluc_with_loss_of_CO_2_13C          | 382,1333 |
|   |          | PB-N-Gluc-with_loss_of_CO_2_(2M-H)       | 763,2697 |
|   |          | PB-N-Gluc-with_loss_of_CO-2_frag1        | 192,0515 |
| PB                                      | 19.69    | PB                                       | 231,0776 |
|   |          | PB_13C                                   | 232,0803 |
|   |          | PB_(2M-2H+Na)                            | 485,1442 |
|   |          | PB_(3M-3H+2Na)                           | 739,2118 |

FIGURE C.1 – Liste des métabolites annotés, dans l'urine de rat traité au phénobarbital





# Annexe D

## Liste des publications

### Journal

1. W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, C. Jutten, “Non-negative Blind Source Separation Algorithm based on Minimum Aperture Simplicial Cone”. soumis à IEEE Transactions on Signal Processing

### Communications

1. W. S. B. Ouedraogo, A. Souloumiac, and C. Jutten. “Non-negative Independent Component Analysis Algorithm based on 2D Givens Rotations and a Newton Optimization”. In Proc. of the 9th international conference on Latent Variable Analysis and Signal Separation, pages 522-529, 2010.
2. W. S. B. Ouedraogo, M. Jaidane, A. Souloumiac, and C. Jutten, “Regularized Gradient Algorithm for Non-negative Independent Component Analysis”, In Proc. of the IEEE International Conference on Accoustic, Speech and Signal Processing, pp. 2524 - 2527, 2011.
3. W.S.B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten. “Geometrical Method Using Simplicial Cones for Overdetermined Nonnegative Blind Source Separation : Application to Real PET Images”. In Proc. of the 10th international conference on Latent Variable Analysis and Signal Separation, pages 494-501, 2012.
4. W. S. B. Ouedraogo, A. Souloumiac, M. Jaidane, C. Jutten, “Simplicial Cone Shrinkng Algorithm for Unmixing Nonnegative Sources”, In Proc. of the IEEE International Conference on Accoustic, Speech and Signal Processing, pp. 2405-2408, 2012.

### Brevet

1. W. S. B. Ouedraogo, A. Souloumiac, “Procédé de Séparation Aveugle de Sources positives”, Brevet national, accepté à l’INPI en avril 2012.



# Bibliographie

- [ACD<sup>+</sup>06] R. Albright, J. Cox, D. Duling, A.N. Langville, and C.D. Meyer. Algorithms, Initializations, and Convergence for the Nonnegative Matrix Factorization. Technical report, NCSU, 2006.
- [ACM<sup>+</sup>03] F. Acernese, A. Ciaramella, S. De Martino, R. De Rosa, and M. Falanga and R. Tagliaferri. Neural Networks for Blind Source Separation of Stromboli Explosion Quakes. *IEEE Transactions on Neural Networks*, 14(1) :167–175, 2003.
- [ACY96] S. Amari, A. Cichocki, and H.H. Yang. A new learning algorithm for blind signal separation. In *Advances in Neural Information Processing Systems*, pages 757–763. MIT Press, 1996.
- [AH01] E. Oja A. Hyvärinen, J. Karhunen. *Independent Component Analysis*. John Wiley and Sons, 2001.
- [APJ01] S. Achard, D.-T. Pham, and C. Jutten. Blind Source Separation in Post Nonlinear Mixtures. In Te-Won Lee, Tzyy-Ping Jung, Scott Makeig, and Terrence J. Sejnowski, editors, *Third International Conference on Independent Component Analysis and Blind Signal Separation, ICA 2001, December, 2001*, pages 295–300, San-Diego, Etats-Unis, 2001.
- [BD09] J. Bioucas-Dias. A Variable Splitting Augmented Lagrangian Approach to Linear Spectral Unmixing. In *First IEEE GRSS Workshop on Hyperspectral Image and Signal Processing*, pages 1–4, 2009.
- [Ber09] Nancy Bertin. *Les factorisations en matrices non-négatives. Approches contraintes et probabilistes, application à la transcription automatique de musique polyphonique*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, 2009.
- [BJ97] R. Bro and S. De Jong. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics*, 11(5) :393–401, 1997.
- [BMD<sup>+</sup>11] F.P. Breitwieser, A. Müller, L. Dayon, T. Köcher, A. Hainard, P. Pichler, U. Schmidt-Erfurth, G. Superti-Furga, J.C. Sanchez, K. Mechtler, K. L. Bennett, and J. Colinge. General Statistical Modeling of Data from Protein Re-

- lative Expression Isobaric Tags. *Journal of Proteome Research*, 10(6) :2758–2766, 2011.
- [BMO98] A. K. Barros, A. Mansour, and N. Ohnishi. Removing artifacts from electrocardiographic signals using independent components analysis. *Neurocomputing*, 22 :173–186, 1998.
- [Bri05] T. Briere. Cours de chromatographie liquide, 2005. <http://departements.univ-reunion.fr/chimie/briere/SITE-NEUF/Chromato/chro-web-index.html>.
- [BTMF04] R. Boloix-Tortosa and Juan J. Murillo-Fuentes. Blind Source Separation in the Adaptive Reduction of Inter-channel Interference for OFDM. *Lecture Notes in Computer Science*, 3195 :1142–1149, 2004.
- [BZMJM04] M. Babaie-Zadeh, A. Mansour, C. Jutten, and F. Marvasti. A Geometric Approach for Separating Several Speech Signals. In Carlos Puntonet and Alberto Prieto, editors, *Independent Component Analysis and Blind Signal Separation*, volume 3195 of *Lecture Notes in Computer Science*, pages 798–806. Springer Berlin / Heidelberg, 2004.
- [CAC07] N. Correa, T. Adali, and V.D. Calhoun. Performance of Blind Source Separation Algorithms for fMRI Analysis using a Group ICA Method. *Magnetic Resonance Imaging*, 25(5) :684–694, 2007.
- [CC97] S. Choi and A. Cichocki. Adaptive Blind Separation of Speech Signals : Cocktail Party Problem. In *In Proc. Int. Conf. Speech Processing*, pages 617–622, 1997.
- [CCJ12] G. Carbras, R. Carniel, and J. Jones. Non-negative Matrix Factorization : an application to ErtaÁle volcano, Ethiopia. *Bollettino di Geofisica Teorica ed Applicata*, 53(2) :231–242, 2012.
- [CCMM09] T. H. Chan, C. Y. Chi, Y. M. Huang, and W. K. Ma. A Convex Analysis-Based Minimum-Volume Enclosing Simplex Algorithm for Hyperspectral Unmixing. *IEEE Transactions on Signal Processing*, 57 :4418–4432, 2009.
- [Cha09] Tsung-Han Chan. *Convex Analysis Based Non-negative Blind Source Separation for Biomedical and Hyperspectral Image Analysis*. PhD thesis, Institute of Communications Engineering, National Tsing Hua University, 2009.
- [Che04] G.M. Chevalier. Tomographie par émission de positrons (TEP), Oct. 2004. Texte offert à titre de renseignement par l’Association canadienne des radiologistes (CAR). <http://www.radiologyinfo.ca/utilisateur/documents/version-pdf/fr/mn-tep.pdf>.
- [CJ10] Pierre Comon and Christian Jutten, editors. *Handbook of Blind Source Separation : Independent Component Analysis and Applications*. Academic Press, 1st edition, 2010.

- [CMCW08] T. H. Chan, W. K. Ma, C. Y. Chi, and Y. Wang. A Convex Analysis Framework for Blind Separation of Non-Negative Sources. *IEEE Transactions on Signal Processing*, 56 :5120–5134, 2008.
- [Com94] P. Comon. Independent Component Analysis, A new concept ? *Signal Processing*, 36 :287–314, 1994.
- [Cra94] M. D. Craig. Minimum-Volume Transforms for Remotely Sensed Data. *IEEE Transactions on Geoscience and Remote Sensing*, 32 :542–552, 1994.
- [CT03] E. Carlos and J.I. Takada. ICA Based Blind Source Separation Applied to Radio Surveillance. *IEICE Transactions on Communications E*, 86 :3491–3497, 2003.
- [Cuq07] J.L. Cuq. Chromatographie Liquide, 2007. <http://diffusiondessavoirs.uomlr.fr/balado/wp-content/uploads/2008/01/chromato-liquide-2007.pdf>.
- [CWLO06] C.I. Chang, C.C. Wu, W.M. Liu, and Y.C. Ouyang. A New Growing Method for Simplex-Based Endmember Extraction Algorithm. *Geoscience and Remote Sensing, IEEE Transactions on*, 44(10) :2804–2819, oct. 2006.
- [CZ06] A. Cichocki and R. Zdunek. NMFLAB-MATLAB Toolbox for Non-Negative Matrix Factorization, 2006. <http://www.bsp.brain.riken.jp/ICALAB/nmflab.html>.
- [CZ07] A. Cichocki and R. Zdunek. Regularized Alternating Least Squares Algorithms for Non-negative Matrix/Tensor Factorization. In Liu D ; Fei S ; Hou ZG ; Zhang HG ; Sun CY, editor, *Advances in Neural Networks - ISNN 2007, Pt 3, Proceedings*, volume 4493 of *Lecture Notes in Computer Science*, pages 793–802. Springer-Verlag, 2007.
- [CZPA09] A. Cichocki, R. Zdunek, A. H. Phan, and S. Amari. *Nonnegative Matrix and Tensor Factorizations, Applications to Exploratory Multi-way Data Analysis and Blind Source Separation*. John Wiley and Sons, Ltd, 2009.
- [dDMF<sup>+</sup>02] O. de Dreuille, P. Maszelin, H. Foehrenbach, G. Bonardel, and J.F. Gaillard. Principe et technique de la tomographie par émission de positons (tep). *Encyclopédie Médico-Chirurgicale*, 35-310-A-10 :1–22, 2002.
- [DH07] Y. Deville and S. Hosseini. Stable higher-order recurrent neural network structures for nonlinear blind source separation. In *Proceedings of the 7th international conference on Independent component analysis and signal separation, ICA'07*, pages 161–168. Springer-Verlag, 2007.
- [DJM09] L. T. Duarte, C. Jutten, and S. Moussaoui. A Bayesian Nonlinear Source Separation Method for Smart Ion-Selective Electrode Arrays. *IEEE Sensors Journal*, 9(12) :1763–1771, December 2009.

- [DKS06] Q. Du, I. Kopriva, and H. Szu. Independent-component analysis for hyperspectral remote sensing imagery classification. *Optical Engineering*, 45(1) :1–13, January 2006.
- [DLPP06] C. Ding, T. Li, W. Peng, and H. Park. Orthogonal Nonnegative Matrix t-Factorizations for Clustering. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 126–135. ACM, 2006.
- [DMC<sup>+</sup>09] N. Dobigeon, S. Moussaoui, M. Coulon, J.Y. Tourneret, and A.O. Hero. Joint Bayesian Endmember Extraction and Linear Unmixing for Hyperspectral Imagery. *IEEE Transactions on Signal Processing*, 57(11) :4355–4368, 2009.
- [DS04] D. Donoho and V. Stodden. When Does Non-Negative Matrix Factorization Give a Correct Decomposition into Parts? *MIT Press*, 2004.
- [FBD09] C. Févotte, N. Bertin, and J. L. Durrieu. Nonnegative Matrix Factorization with the Itakura-Saito Divergence : With Application to Music Analysis. *Neural Computation*, 21(1) :793–830, 2009.
- [FLF<sup>+</sup>03] D. Farina, F. Lebrun, C. Févotte, C. Doncarli, and R. Merletti. Blind source separation of linear mixtures of non-stationary surface EMG signals. In *Proc. 19e colloque GRETSI sur le Traitement du Signal et des Images*, pages 942–945, Paris, France, Sep. 2003.
- [Gil08] S. Gillet. La chromatographie et l'électrophorèse, 2008. <http://perso.latrimu.com/shagar/steve/pdf/ch42BBM.pdf>.
- [GPH04] C. Gobinet, E. Perrin, and R. Huez. Application of Non-negative Matrix Factorization to Fluorescence Spectroscopy. In *12th European Signal Processing Conference. EUSIPCO'04*, 2004.
- [GS06] R. Gribonval and S. Lesage. A survey of Sparse Component Analysis for blind source separation : principles, perspectives, and new challenges. In *European Symposium on Artificial Neural Networks (ESANN)*, pages 323–330, 2006.
- [GTCB07] P. Georgiev, F. Theis, A. Cichocki, and H. Bakardjian. Sparse Component Analysis : a New Tool for Data Mining. In PanosM. Pardalos, VladimirL. Boginski, and Alkis Vazacopoulos, editors, *Data Mining in Biomedicine*, volume 7 of *Springer Optimization and Its Applications*, pages 91–116. Springer US, 2007.
- [Har08] E.D. Harris. Biochemical facts behind the definition and properties of metabolites, 2008. <http://www.fda.gov/ohrms/dockets/ac/03/briefing/3942b1-08-harris-20paper.pdf> : Consulté le 24 mai 2012.

- [Hen97] R. C. Henry. History and fundamentals of multivariate air quality receptor models. *Chemometrics and Intelligent Laboratory Systems*, 37 :37–42, 1997.
- [HFR<sup>+</sup>09] C. Hundertmark, R. Fischer, T. Reinl, S. May, F. Klawonn, and L. Jansch. MS-specific noise model reveals the potential of iTRAQ in quantitative proteomics. *Bioinformatics*, 25(8) :1004–1011, Apr. 2009.
- [HJA85] J. Hérault, C. Jutten, and B. Ans. Detection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Xème colloque GRETSI*, pages 1017–1022, Nice, France, 20-24 Mai 1985.
- [Hoy02] P. O. Hoyer. Non-Negative Sparse Coding. In *In Proc. IEEE Workshop on Neural Networks for Signal Processing*, pages 557–565, 2002.
- [HP99] A. Hyvärinen and P. Pajunen. Nonlinear Independent Component Analysis : Existence and Uniqueness Results. *Neural Networks*, 12 :429–439, 1999.
- [IC99] A. Ifarraguerri and C.I. Chang. Multispectral and hyperspectral image analysis with convex cones. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2) :756–770, 1999.
- [IIA07] K. Ichige, M. Imai, and H. Arai. Fastica-Based Blind Signal Separation and its Application to Radio Surveillance. In *14th Workshop on Statistical Signal Processing*, pages 546 – 550. IEEE, 2007.
- [IM04] M. M. Ichir and A. Mohammad-Djafari. Bayesian Blind Source Separation of Positive Non Stationary Sources. In R. Fischer, R. Preuss, & U. V. Toussaint, editor, *American Institute of Physics Conference Series*, volume 735 of *American Institute of Physics Conference Series*, pages 493–500, November 2004.
- [JC07] C. Jutten and P. Comon, editors. *Séparation de sources 2 : au-delà de l’aveugle et applications*, volume 2. *Traité IC2 - Traitement du signal et de l’image*, hermès - lavoisier edition, Mars 2007.
- [KAS07] A. Kachenoura, L. Albera, and L. Senhadji. Blind source separation in biomedical engineering. *ITBM-RBM*, 28 :20–34, 2007.
- [Kil03] Yalcin Kilic. Source Separation of Seismological Signals using Blind Deconvolution and Independent Component Analysis Techniques. Master’s thesis, Bogazici University, 2003.
- [KP00] J. Klingseisen and M.D. Plumbley. Towards Musical Instrument Separation Using Multiple-Cause Neural Networks. In *in Proceedings of the International Workshop on Independent Component Analysis And Blind Signal Separation*, pages 447–452, 2000.



- [LB97] T.W. Lee and A.J. Bell. Blind source separation of real world signals. In *International Conference on Neural Networks*, volume 4, pages 2129–2134. IEEE, 1997.
- [LBD08] J. Li and J.M. Bioucas-Dias. Minimum Volume Simplex Analysis : A Fast Algorithm to Unmix Hyperspectral Data. In *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, volume 3, pages III–250 –III–253, july 2008.
- [LCiA03] Y. Li, A. Cichocki, and S. i. Amari. Sparse Component Analysis for Blind Source Separation with Less Sensors than Sources. In *4th International Symposium on Independent Component Analysis and Blind Source Separation*, Nara, Japan, April 2003.
- [LCP<sup>+</sup>08] H. Laurberg, M.G. Christensen, M.D. Plumbley, L.K. Hansen, and S.H. Jensen. Theorems on Positive Data : On the Uniqueness of NMF. *Computational Intelligence and Neuroscience*, 2008, 2008.
- [Lin07] C. J. Lin. Projected Gradient Methods for Non-negative Matrix Factorization. *Neural Computation*, 19(10) :2756–2779, 2007.
- [LLCL01] J. S. Lee, D. D. Lee, S. Choi, and D. S. Lee. Application Of Non-Negative Matrix Factorization To Dynamic Positron Emission Tomography. In *In 3rd International Conference on Independent Component Analysis and Blind Signal Separation*, pages 629–632, 2001.
- [LNN10] C. Lazar, D. Nuzillard, and A. Nowé. A New Geometrical BSS Approach for Non Negative Sources. In *Latent Variable Analysis and Signal Separation*, volume 6365 of *Lecture Notes in Computer Science*, pages 530–537. 2010.
- [LS99] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Letter to nature*, 401 :788–791, 1999.
- [LS01] D. D. Lee and H. S. Seung. Algorithms for Non-negative Matrix Factorization. *Neural Information Processing Systems*, 13 :556–562, 2001.
- [LW11] C.H. Lee and W. Wolf. Blind Signal Separation for Cognitive Radio. *Journal of Signal Processing Systems*, 63(1) :67–81, 2011.
- [LYC<sup>+</sup>05] Q. Li, J.H. Yang, X. Chen, Z. Liang, and Y.X.Ren. The decomposition of surface EMG signals based on blind source separation of convolved mixtures. In *27th Annual International Conference of the Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005*, volume 6, pages 5912–5915. IEEE, 2005.
- [Mat11] R. Matthews. Definition of metabolites, August 2011. <http://www.livestrong.com/article/510498-definition-of-metabolites/> : Consulté le 24 mai 2012.

- [MBDC06] S. Moussaoui, D. Brie, A. M. Djafari, and C. Carteret. Separation of Non-Negative Mixture of Non-Negative Sources Using a Bayesian Approach and MCMC Sampling. *IEEE Transactions on Signal Processing*, 54(11) :4133–4145, 2006.
- [MBI05] S. Moussaoui, D. Brie, and J. Idier. Non-negative Source Separation : Range of Admissible Solutions and Conditions for the Uniqueness of the Solutions. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 5 :289–292, 2005.
- [MhdCSD08] C. M. Monzón, M. R. Delfino (h), M. d. C. Sarno, and M. R. Delfino. Phenobarbital in Pharmaceutical Tablets by modified Conductimetric Analysis. *The Journal of the Argentine Chemical Society*, 96 :101–110, 2008.
- [MM94] E. Moreau and O. Macchi. A one stage self-adaptive algorithm for source separation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pages 49–52, 1994.
- [Mou05] Said Moussaoui. *Séparation de sources non-négatives. Application au traitement des signaux de spectroscopie*. PhD thesis, Université Henri Poincaré, Nancy 1, 2005.
- [MPC<sup>+</sup>07] D. Mantini, M.G. Perrucci, S. Cugini, A. Ferretti, G.L. Romani, and C. Del Gratta. Complete artifact removal for EEG recorded during continuous fMRI using independent component analysis. *NeuroImage*, 34 :598–607, 2007.
- [MQ07] L. Miao and H. Qi. A constrained Non-Negative Matrix Factorization Approach to Unmix Highly Mixed Hyperspectral Data. In *IEEE International Conference on Image Processing*, pages 185–188, 2007.
- [MSAM06] R. Mukai, H. Sawada, S. Araki, and S. Makino. Frequency-Domain Blind Source Separation of Many Speech Signals Using Near Field and Far Field Models. *EURASIP Journal on Applied Signal Processing*, 2006 :1–13, 2006.
- [NBD05] J. M. P. Nascimento and J. M. Bioucas-Dias. Vertex Component Analysis : A Fast Algorithm to Unmix Hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43 :898–910, 2005.
- [NMBZJ08] F. M. Naini, G. H. Mohimani, M. Babaie-Zadeh, and C. Jutten. Estimating the mixing matrix in Sparse Component Analysis (SCA) based partial k-dimensional subspace clustering. *Neurocomputing*, 71 :2330–2343, 2008.
- [OJSJ11] W. S. B. Ouedraogo, M. Jaidane, A. Souloumiac, and C. Jutten. Regularized Gradient algorithm for Non-Negative Independent Component Analysis. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 2524–2527, may 2011.

- [OPBG07] A. Ozerov, P. Philippe, F. Bimbot, and R. Gribonval. Adaptation of Bayesian Models for Single-Channel Source Separation and its Application to Voice/Music Separation in Popular Songs. *IEEE Transactions on Audio, Speech, and Language processing*, 15(5) :1564–1578, July 2007.
- [OSJ10] W. S. B. Ouedraogo, A. Souloumiac, and C. Jutten. Non-negative Independent Component Analysis Algorithm based on 2D Givens Rotations and a Newton Optimization. *Lecture Note in Computer Sciences*, 6365 :522–529, 2010.
- [OSJJ12a] W.S.B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten. Geometrical Method Using Simplicial Cones for Overdetermined Nonnegative Blind Source Separation : Application to Real PET Images. In *Proceedings of the 10th international conference on Latent Variable Analysis and Signal Separation*, pages 494–501, 2012.
- [OSJJ12b] W.S.B. Ouedraogo, A. Souloumiac, M. Jaidane, and C. Jutten. Simplicial Cone Shrinking Algorithm for Unmixing Nonnegative Sources. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 2405–2048, march 2012.
- [PAB<sup>+</sup>02] M.D. Plumbley, S.A. Abdallah, J.P. Bello, M.E. Davies, G. Monti, and M.B. Sandler. Automatic music transcription and audio source separation. *Cybernetics and Systems*, 33(6) :603–627, 2002.
- [PC72] R. J. Plemmons and R. E. Cline. The Generalized Inverse of a Nonnegative Matrix. In *American Mathematical Society*, volume 31, pages 46–50, 1972.
- [PCVBO10] T. Pluskal, S. Castillo, A. Villar-Briones, and M. Oresic. Mzmine 2 : Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC*, 11(395) :1–11, 2010.
- [PG97] D. T. Pham and P. Garat. Blind Separation of Mixture of Independent Sources Through a Maximum Likelihood Approach. In *Proc. EUSIPCO*, pages 771–774, 1997.
- [Plu02] M. Plumbley. Conditions for Nonnegative Independent Component Analysis. *IEEE Signal Processing Letters*, 9 :177–180, 2002.
- [Plu03] M. Plumbley. Algorithms for Nonnegative Independent Component Analysis. *IEEE Transactions on Neural Networks*, 14(3) :534–543, 2003.
- [Plu04] M. Plumbley. Optimization using Fourier Expansion over a Geodesic for Non-Negative ICA. *Lecture Notes in Computer Science*, 3195 :44–56, 2004.
- [Plu05] M. Plumbley. Geometrical methods for non-negative ICA : Manifolds, Lie groups and toral subalgebras. *Neurocomputing*, 67 :161–197, 2005.

- [Por07] J. C. Portais. Métabolomique, 2007. <http://metasys.insa-toulouse.fr//docens/cours-M2Pro-JCP07-1.pdf> : Consulté le 24 mai 2012.
- [PP94] U. Tapper P. Paatero. Positive matrix factorization : A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5(2) :111–126, 1994.
- [Qui08] E. Quivet. Chimie analytique instrumentale notions de spectrométrie de masse, 2008. <http://gsite.univ-provence.fr/gsite/Local/lcp-ira/dir/Etienne/Cours-20MasS.pdf>.
- [San05] S. Sanglier. Cours spectrométrie de Masse : Introduction, Oct. 2005. <http://www-esbs.u-strasbg.fr/notesdecours/3eme-annee/gap/MmeSanglier05-10-05/Intro-SS.pdf>.
- [SP09] J.L. Sébédio and A. Paris. La métabolomique : un nouvel outil en nutrition et toxicologie. *Lettre Scientifique de l'Institut Français de la Nutrition*, 139 :1–8, 2009.
- [SPL11] R. Schachtner, G. Pöppel, and E.W. Lang. Towards unique solutions of non-negative matrix factorization problems by a determinant criterion. *Digital Signal Processing*, 21 :528–534, 2011.
- [SWO<sup>+</sup>06] C. A. Smith, E. J. Want, G. O'Maille, R. Abagyan, and G. Siuzdak. XCMS : Processing mass spectrometry data for metabolite profiling using Nonlinear peak alignment, matching, and identification. *Analytical Chemistry*, 78(3) :779–787, 2006.
- [SZC09] X. Shao, Z. Zhichao, and W. Cai. Extraction of chemical information from complex analytical signals by a non-negative independent component analysis. *Analyst*, 134 :2095–2099, 2009.
- [Séb09] J.L. Sébédio. Nutrition et Métabolomique. *Cholé-Doc*, 114, 2009.
- [Tch11] G. Tcherkez. Initiation à la métabolomique. [http://www.pmm.u-psud.fr/IMG/pdf/Initiation\\_metabolomique\\_P6.pdf](http://www.pmm.u-psud.fr/IMG/pdf/Initiation_metabolomique_P6.pdf), 2011. En ligne, Consulté le 24 mai 2012.
- [TJ99] A. Taleb and C. Jutten. Source separation in post-nonlinear mixtures. *IEEE Transactions on Signal Processing*, 47(10) :2807–2820, October 1999.
- [TML95] N. Thirion, J. Mars, and J.L. Lacoume. Séparation aveugle de signaux large bande : Un nouveau challenge en prospection sismique. In *GRETSI*, pages 1335–1338, Juan-les-Pins, Sep. 1995.
- [TS08] R. Trébossen and A. Souloumiac. Analyse des séries dynamiques TEP : Extraction des pharmacocinétiques des traceurs. Technical report, CEA, 2008.

- [Van03] P. Vanroose. Blind Source Separation of Speech and Background Music for Improved Speech Recognition. In *In Proceedings of the 24th Symposium on Information Theory*, pages 103–108, 2003.
- [WCCW10] F. Y. Wang, C. Y. Chi, T. H. Chan, and Y. Wang. Nonnegative Least-Correlated Component Analysis for Separation of Dependent Sources by Volume Maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 :875–888, 2010.
- [XB99] J. Xavier and V. Barroso. Blind Source Separation, ISI Cancellation and Carrier Phase Recovery in SDMA Systems for Mobile Communications. *Wireless Personal Communications*, 10(1) :53–76, 1999.
- [XNN10] Y. Xiang, S.N. Kui Ng, and V.K. Nguyen. Blind separation of mutually correlated sources using precoders. *IEEE Transactions on Neural Networks*, 21(1) :82–90, 2010.
- [YL05] H. Zheng Y. Li. Improvement for Nonnegative PCA Algorithm for Independent Component Analysis. *Neural Networks and Brain, 2005. ICNN&B '05. International Conference on*, 3 :2000–2002, Octobre 2005.
- [YW08] K. Youssef and P.Y. Woo. Instrument sound separation in songs. In *International Conference on Electro/Information Technology*. IEEE, 2008.
- [ZC06] R. Zdunek and A. Cichocki. Non-negative matrix factorization with quasi-newton optimization. In *In Eighth International Conference on Artificial Intelligence and Soft Computing, ICAISC*, pages 870–879. Springer, 2006.
- [ZC08] R. Zdunek and A. Cichocki. Fast Nonnegative Matrix Factorization Algorithms Using Projected Gradient Approaches for Large-Scale Problems. *Journal of Computational Intelligence and Neuroscience*, 2008 :1–13, 2008.
- [ZCHW08] Y. Zhu, T. H. Chan, E. P. Hoffman, and Y. Wang. Gene Expression Dissection by Non-negative Well-Grounded Source Separation. In *IEEE Workshop on Machine Learning for Signal Processing*, pages 255–260, 2008.
- [ZFFM<sup>+</sup>09] P. Zanotti-Fregonara, E.M. Fadailli, R. Maroy, C. Comtat, A. Souloumiac, S. Jan, M.-J. Ribeiro, V. Gaura, A. Bar-Hen, and R. Trébossen. Comparison of eight methods for the estimation of the image-derived input function in dynamic [18F]-FDG PET human brain studies. *Journal of Cerebral Blood Flow and Metabolism*, 29 :1825–1835, 2009.
- [ZLY06] N. Zhang, J. Lu, and T. Yahagi. Nonlinear Blind Source Separation Method for X-Ray Image Separation. *IEICE Transactions*, 89-A(4) :924–931, 2006.
- [ZP01] M. Zibulevsky and B.A. Pearlmutter. Blind source separation by sparse decomposition in a signal dictionary. *Neural Computation*, 13(4) :863–882, 2001.

- [ZXY09] Y. Zheng, X. Xie, and L. Yang. Cooperative Spectrum Sensing Based on Blind Source Separation for Cognitive Radio. In *First International Conference on Future Information Networks*, pages 398–402, 2009.