Spatially constrained dictionary learning for source separation

Argheesh Bhanot

UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE MSII

Laboratoire ICube - UMR 7357

THÈSE présentée par :

Argheesh BHANOT

soutenue le 5 juillet 2021
pour obtenir le grade de : **Docteur de l'université de Strasbourg**
Discipline/ Spécialité : Signal, image, automatique, robotique

Apprentissage par dictionnaire pour la séparation de sources sous contraintes spatiales

THÈSE dirigée par :

M. HEITZ Fabrice	Professeur, Université de Strasbourg, Directeur de thèse
Mme HARSAN Laura	MCU-PH, Université de Strasbourg, Directrice de thèse

RAPPORTEURS :

M. FERRARI André	Professeur, Université de Nice-Sophia Antipolis
Mme ACHARD Sophie	Directrice de recherche, LJK, Université Grenoble Alpes

EXAMINATEURS :

M. BRIE David	Professeur, Université de Nancy
Mme MEILLIER Céline	Maître de Conférences, Université de Strasbourg, encadrante

It seems to me that the natural world is the greatest source of excitement; the greatest source of visual beauty, the greatest source of intellectual interest. It is the greatest source of so much in life that makes life worth living.

David Attenborough

Direction de la **recherche**
et de la **valorisation**
Collège **doctoral**
Université de Strasbourg

Déclaration sur l'honneur
Declaration of Honour

J'affirme être informé que le plagiat est une faute grave susceptible de mener à des sanctions administratives et disciplinaires pouvant aller jusqu'au renvoi de l'Université de Strasbourg et passible de poursuites devant les tribunaux de la République Française.

Je suis conscient(e) que l'absence de citation claire et transparente d'une source empruntée à un tiers (texte, idée, raisonnement ou autre création) est constitutive de plagiat.

Au vu de ce qui précède, **j'atteste sur l'honneur que le travail décrit dans mon manuscrit de thèse est un travail original et que je n'ai pas eu recours au plagiat ou à toute autre forme de fraude.**

I affirm that I am aware that plagiarism is a serious misconduct that may lead to administrative and disciplinary sanctions up to dismissal from the University of Strasbourg and liable to prosecution in the courts of the French Republic.

I am aware that the absence of a clear and transparent citation of a source borrowed from a third party (text, idea, reasoning or other creation) is constitutive of plagiarism.

In view of the foregoing, ***I hereby certify that the work described in my thesis manuscript is original work and that I have not resorted to plagiarism or any other form of fraud.***

Nom : Prénom : BHANOT Argheesh

Ecole doctorale :MSII

Laboratoire : ICube

Date : 5 juillet 2021

Signature :

Contents

List of Figures

List of Tables

Introduction

In the past few decades, data acquisition methods have made tremendous progress. The advancement of novel acquisition techniques has led to an increase in the complexity of data. The information from large amounts of complex data has helped us understand a myriad of phenomena. For example, such information has enabled us to analyse data beyond the visible spectral range to discover galaxies or black holes in astronomy. In the medical field, complex data such as functional Magnetic Resonance Imaging (fMRI) have enabled to understand brain activities; similarly, data from scintigraphy have helped in understanding the physical responses of the various body parts. On another front, in a more consumer-oriented market, visual images and speech signals can be used to identify individuals. In addition to data acquisition, advancement in the ability of magnetic devices to stock large data and improvements in calculation performances led us to find important information such data contain. Furthermore, novel experiments involve multi-modal data from different types of acquisition that can be used to provide more insights into the application in question.

We see that in the examples presented before, there are some *sources* that interest us. For instance in astronomy, it could be the galaxies, planets, or other celestial objects; in fMRI, it could be a region whose activity changes due to ageing or due to disease; in remote sensing domain, it could be the ores in a particular mine. To extract these sources from a mixture of signals/images, i.e. to *unmix* them to get related or specific information about them, source separation methods are necessary.

This work presents a source separation method for application to problems of source separation in spatially structured data: 2D or 3D images that contain temporal information (fMRI, scintigraphy) or light spectrum information (hyperspectral imaging) where an *a priori* information about the approximate spatial localisation of the sources is available. This information about the regions of interest as belonging to different modalities is not always precise, which therefore would require some kind of registration and maybe some slight adjustments to have an exact correspondence between the sources in the two modalities. In the absence of such exact information, an approximation or just partial information about the locations can then be used. Massive datasets with strong mixing (extremely high number of sources) can lead to indeterminate solutions; it is hence useful to incorporate such information even

though it is partial.

The unmixing problem involves simultaneous estimation of the precise locations of the sources, as well as their contributions within each pixel/voxel. The mixing could strongly depend on the application and the data at disposition. Generally, during the acquisition of such data, the compromise between spatial resolution and temporal/spectral resolution is often at the expense of spatial resolution due to the preference to the spectral or temporal resolution. This results in a potentially large mixing of sources in the same pixel/voxel with a lower spatial precision of the structures of the sources. In addition to this, the signals to unmix could present similarities and may add to the complexity of the unmixing model. This requires the addition of special constraints to counter the ill-posedness of the unmixing problem and thus increases the complexity of the algorithm.

Source separation methods must incorporate spatial information to estimate the contribution and signature of each source in the image. The extra knowledge related to the potential localisations of the sources is generally in the form of another observation modality (even manually defining the Regions of Interests (ROIs)) which requires setting up the preprocessing techniques adapted for registering the spatial information between them. Depending upon the application, this preprocessing step could be complicated.

In the approach developed during this PhD, the problem to unmix sources incorporating the approximate external spatial information is dealt with by introducing a spatial constraint based on an indicator function that allows unmixing strictly under the regions of interest. The proposed unmixing is done using a dictionary learning method for solving the constrained optimisation problem. The easy adaptability of the proposed algorithm to drop or add constraints on source properties made it possible to use it on varied applications.

A generic algorithm is proposed, and its application to different kinds of data and thus, different domains of application is described. As the problem of source separation is not the only objective of the thesis, approaches for change detection and detection of new sources are equally detailed for specific applications.

The first chapter focuses on the various methods of source separation considered as state-of-the-art. The cornerstones of blind source separation, i.e. Principal Component Analysis (PCA) and Independent Component Analysis (ICA), and the analysis of their applicability on 2D or 3D images with a temporal or spectral dimension, are presented. This follows up with the methods of dictionary learning with some classic priors that exploit spatial structures in the data. Geometrical-based approaches, which are known for their application in hyperspectral images, are also described. The formulation of the problems that interest us is incomplete without considering the Bayesian inference approach, which is also mentioned. Non-linear models and tensorial decomposition are introduced.

The second chapter is the core of this research work, where the proposed unmixing model has been detailed. The chapter also covers the evaluation and the

performances of our method on synthetic data. The chapter concludes with a discussion on the performance and perspectives to further improve the execution times of the method.

The third chapter contains the studies on fMRI data. One of the medical issues mainly dealt with during the PhD is the detection of changes in functional connectivity during neurodegenerative diseases (typically in Alzheimer's disease). The context behind the principal application of this PhD on fMRI data is a PhD supervised by two teams: IMAGeS (Images, Modélisation, Apprentissage, Géométrie et Statistique) team for the signal processing part and IMIS (Imagerie Multimodale Intégrative en Santé) for the acquisition of the data and medical application part. The data was acquired internally on the IRIS (Imagerie, Robotique et Innovation pour la Santé) platform of the ICube laboratory in the framework to study the neurodegenerative diseases such as Alzheimer's. The proposed hypotheses for the detection of cerebral networks, or the regions presenting similar activity in the brain, are presented. These cerebral networks are composed of different anatomical regions in the brain, and their cerebral activity is measured with fMRI data. Alzheimer's impacts the cerebral activity and thus the connectivity between the regions. This study permits to estimate the changes in the structures of cerebral networks and temporal activity at different stages during the development of Alzheimer; this further enables to look for changes in the correlation matrices formed by the signals estimated by our algorithm. The approach required the adaptation of a whole preprocessing pipeline for fMRI mouse data. The innovative nature of the approach on the biological side has been explained. An application on quasi-real data where the performance of the method was evaluated is presented. In addition to the application of the proposed algorithm on the fMRI unmixing problem, the chapter introduces statistical methods to evaluate differences between different groups of mice and change detection for a longitudinal study. The chapter concludes with the application of these techniques on real data after a general evaluation of the unmixing by looking at the connectivity of a group of control/healthy mouse brains and the analysis of results.

In order to demonstrate the generic nature of the algorithm, the application of the algorithm on different types of data where the local knowledge about the sources is different from the one in fMRI has been proposed in the fourth and last chapter. It was discovered that the unmixing algorithm with some changes allowed to enlarge the fields of applications to datasets other than fMRI. While in the fMRI data the unmixing consists in finding contributions (in the form of proportions) of the different anatomical regions in each voxel, the decomposition of the image sequences in scintigraphy consists in estimating the activity of the different organs that are superimposed (due to 2D acquisition of 3D structures). It is then necessary to modify the constraints of the sources in the mixture model. Moreover, the spatial localisation information is much more approximate since the ROIs are roughly delimited by a doctor. However, we see that sources estimates are comparable to the state-of-the-art methods developed for this application. Another targeted application is in

the field of astrophysics, where the estimation of spectra belonging to galaxies in hyperspectral images was performed. The mixing here is additive, i.e. the observed mixed signal is the sum of the signals from individual sources (galaxies). In this case, the localisation information is from Hubble Space Telescope dataset, and the unmixing is performed on the Multi Unit Spectroscopic Explorer (MUSE) Ultra Deep field 10 (UDF-10) dataset. The localisation information of different objects in the two datasets uses the same coordinate system; thus, no registration step is required. The results by the proposed method are compared to the analysis realised by the team behind the construction of the MUSE instrument and the production of this hyperspectral data on which several promising results have been published. Methods to refine the structures of galaxies and the estimation of galaxies not visible in the high resolution spatial images of Hubble are equally presented. For the astronomical case as well as the scintigraphy case, the constraints to be taken into account and the implementation details to unmix data are provided.

Liste of publications

Journal article

[1] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Spatially Constrained Online Dictionary Learning for Source Separation, IEEE Transactions on Image Processing, February 2021. DOI: 10.1109/TIP.2021.3058558

Conference paper

[2] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Online Dictionary Learning for Single-Subject fMRI Data Unmixing, EUSIPCO, A Coruña, Spain, September 2019. DOI: 10.23919/EUSIPCO.2019.8902991

[3] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf, Gretsi, Lille, France, August 2019.

Conference with abstracts

[4] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen Mouse Brain en IRMf chez la souris, SFRMBM, Strasbourg, France, March 2019.

Other communications

[5] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Online dictionary learning for single-subject fMRI data unmixing: présentation orale, Journée Data Science and Artificial Intelligence (DSAI), ICube, Strasbourg, 2019.

[6] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf; poster, 7èmes Journées scientifiques de la Fédération de la Médecine Translationnelle (FMTS) de Strasbourg, 2019.

[7] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen mouse brain en IRMf chez la souris. Journée poster des doctorants ICube, 2018.

Résumé en français

Au cours des dernières décennies, les méthodes d'acquisition de données ont fait d'énormes progrès. L'avancement de nouvelles techniques d'acquisition a conduit à une augmentation de la complexité des données. Les informations provenant de grandes quantités de données complexes nous ont aidés à comprendre une myriade de phénomènes. Par exemple, ces informations nous ont permis d'analyser des données au-delà de la gamme spectrale visible pour découvrir des galaxies ou des trous noirs en astronomie. Dans le domaine médical, des données complexes telles que l'imagerie par résonance magnétique fonctionnelle (IRMf) ont permis de comprendre les activités cérébrales ; de même, les données de la scintigraphie ont aidé à comprendre les réponses physiques des différentes parties du corps. Par ailleurs, dans un marché plus orienté vers le consommateur, les images visuelles et les signaux vocaux peuvent être utilisés pour identifier les individus. Outre l'acquisition de données, l'amélioration de la capacité des dispositifs magnétiques à stocker de grandes données et l'amélioration des performances de calcul nous ont permis d'extraire des informations de ces données.. En outre, de nouvelles expériences font intervenir des données multimodales provenant de différents types d'acquisition qui peuvent être utilisées pour fournir davantage d'informations sur l'application en question en combinant des mesures de différentes natures.

Dans les exemples précédemment présentés, de nombreuses sources d'intérêt nécessitent d'être étudiées. Par exemple, en astronomie, il peut s'agir des galaxies, des planètes ou d'autres objets célestes ; en IRMf, il peut s'agir d'une région dont l'activité change en raison du vieillissement ou d'une maladie. Les méthodes de séparation des sources sont nécessaires pour extraire ces sources d'un mélange de signaux et d'images et obtenir des informations connexes ou spécifiques à leur sujet.

Dans cette thèse, nous nous intéressons aux problèmes de séparation de sources dans des signaux spatialement structurés : des images 2D ou 3D qui contiennent des informations temporelles (IRMf, scintigraphie) ou des informations sur le spectre lumineux (imagerie hyperspectrale) où une information *a priori* sur la localisation spatiale approximative des sources est disponible. Cette information sur les régions d'intérêt extraite de modalités différentes n'est pas toujours précise, ce qui nécessite un recalage entre les images pour obtenir une correspondance exacte entre les sources dans les deux modalités. En l'absence de ces informations exactes, on peut alors

utiliser une approximation ou des informations partielles sur les emplacements. Les ensembles de données massifs avec un fort mélange (nombre extremement élevé de sources superposées) peuvent conduire à des solutions indéterminées ; il est donc utile d'incorporer ces informations, même si elles sont partielles.

Le problème du démélange implique l'estimation simultanée de l'emplacement précis des sources, ainsi que de leurs contributions dans chaque pixel/voxel. Le mélange peut dépendre fortement de l'application et des données à disposition. En général, lors de l'acquisition de ces données, le compromis entre la résolution spatiale et la résolution temporelle/spectrale se fait souvent au détriment de la résolution spatiale en raison de la préférence accordée à la résolution spectrale ou temporelle. Il en résulte un mélange potentiellement important de sources dans le même pixel/voxel avec une précision spatiale moindre des structures des sources. En outre, les signaux à démélanger peuvent présenter des similitudes et ajouter à la complexité du modèle de démélange. Cela nécessite l'ajout de contraintes spéciales pour contrer le caractère mal posé du problème de démélange et augmente ainsi la complexité de l'algorithme.

Les méthodes de séparation des sources doivent intégrer des informations spatiales pour estimer la contribution et la signature de chaque source dans l'image. La connaissance supplémentaire liée aux localisations potentielles des sources se présente généralement sous la forme d'une autre modalité d'observation (voire d'une définition manuelle des régions d'intérêt (ROI)) qui nécessite la mise en place de techniques de prétraitement adaptées pour recaler les informations spatiales entre elles. Selon l'application, cette étape de prétraitement peut être compliquée.

Dans l'approche développée au cours de cette thèse, le problème du démélange des sources incorporant l'information spatiale externe approximative est traité en introduisant une contrainte spatiale basée sur une fonction indicatrice qui permet un démélange strictement sous les régions d'intérêt. Le démélange proposé est effectué en utilisant une méthode d'apprentissage par dictionnaire pour résoudre le problème d'optimisation sous contrainte. La facilité d'adaptation de l'algorithme proposé pour supprimer ou ajouter des contraintes sur les propriétés des sources a permis de l'utiliser pour des applications variées.

Un algorithme générique est proposé, et son application à différents types de données et donc à différents domaines d'application est décrite. Le problème de la séparation des sources n'étant pas le seul objectif de la thèse, les approches de détection de changement et de détection de nouvelles sources sont également détaillées pour des applications spécifiques.

Modéle et méthode

Le modèle linéaire classique utilisé dans la séparation des sources peut être écrit comme suit :

$$\mathbf{Y} \simeq \mathbf{UA}, \tag{1}$$

où $\mathbf{Y} \in \mathbb{R}^{N \times P}$ sont des données observées. Dans la séparation spatio-spectrale, N peut être interprété comme la longueur spectrale et dans la séparation spatio-temporelle N est la longueur des signaux temporels. P est le nombre de voxels ou de pixels, selon l'ensemble de données. La matrice $\mathbf{U} \in \mathbb{R}^{N \times R}$ contient les signatures temporelles/spectrales où R est le nombre de sources. La matrice $\mathbf{A} \in \mathbb{R}^{R \times P}$, généralement appelée matrice de mélange ou d'abondance, contient la fraction des contributions des composants R à chaque voxel ou pixel.

Formulation d'optimisation sous contrainte

Étant donné le modèle d'observation (1), le problème de minimisation suivant :

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 \tag{2}$$

n'a pas de solution unique en raison de l'estimation conjointe de $\mathbf{A}$ et $\mathbf{U}$, et du caractère mal posé du problème. Afin de limiter le nombre de solutions, nous introduisons une contrainte spécifique sur la forme de la matrice $\mathbf{A}$ définie en fonction de certaines informations supplémentaires provenant de la segmentation haute résolution (HR) des sources ou de la connaissance de l'emplacement des sources : nous savons quelles ROIs peuvent contribuer à un voxel donné, c'est-à-dire présenter une proportion non nulle à ce voxel. Lorsque le nombre total de sources R est élevé, cette connaissance *a priori* permet de contraindre les solutions possibles du problème de minimisation. Certaines contraintes standard sur la matrice $\mathbf{A}$, telles que la *positivité* des coefficients de mélange, peuvent être ajoutées dans un terme de contrainte supplémentaire $g(\mathbf{A})$. De la même manière, des contraintes sur la matrice $\mathbf{U}$ peuvent être modélisées par une contrainte générique $h(\mathbf{U})$. Dans le cadre le plus général, le problème de démélange est reformulé comme suit :

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) + g(\mathbf{A}) + h(\mathbf{U}), \tag{3}$$

où le premier terme est le terme de fidélité aux données et le second terme $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ est la fonction indicatrice sur l'ensemble $M(\tilde{\mathbf{A}})$ de matrices ayant une structure similaire à une "matrice de structure" binaire donnée $\tilde{\mathbf{A}}$, c'est-à-dire $\mathbf{A} \in M(\tilde{\mathbf{A}})$ si et seulement si $\mathbf{A} \in \mathbb{R}^{R \times P}$ et ses coefficients $\mathbf{A}_{i,j} = 0$ si $\tilde{\mathbf{A}}_{i,j} = 0$. La matrice $\tilde{\mathbf{A}}$ est une matrice binaire, où l'élément $(\tilde{\mathbf{A}})_{r,i} = 1$ si, selon les connaissances a priori sur la localisation spatiale des sources, la r^{ime} région d'intérêt pourrait exister dans le i^{ime} voxel, et 0 sinon. Il en résulte que $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) = \infty$ si au moins un élément de $\mathbf{A}$ est non nul alors qu'il est nul dans $\tilde{\mathbf{A}}$, et 0 sinon.

L'estimation conjointe de $\mathbf{U}$ et $\mathbf{A}$ dans l'équation (3) est un problème typique de l'apprentissage par dictionnaire (DL). Mais, contrairement aux algorithmes DL classiques, le modèle proposé dans cette thèse n'a pas de terme de régularisation de la parcimonie sous la forme d'une pénalité ℓ_1 : ce sont les informations de localisation

des sources codées dans le terme structurel $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ qui imposent la décomposition de chaque voxel. Une manière classique de résoudre le problème d'estimation conjointe consiste à optimiser alternativement la fonction de coût eq. (3) selon $\mathbf{U}$ et $\mathbf{A}$ comme présenté dans l'algorithme

```
1 Initialisation de la matrice binaire Ã
2 Initialisation de la matrice U^(0), l = 0
3 while CRITÈRE D'ARRÊT ≠ VRAI do
4 |   Minimisation du problème (3) par rapport à A
5 |   Minimisation du problème (3) par rapport à U
6 end
7 return U^(l+1), A^(l+1)
```

Algorithm 1: Schéma d'optimisation alternatif de l'algorithme d'apprentissage par dictionnaire pour résoudre le problème générique (3).

Initialement, la méthode a été évaluée sur un exemple purement synthétique composé de données de dimension trois (deux dimensions pour l'image et une pour le temps). Dans ces données on trouve différents cas d'interactions entre les sources : a) sans superposition , b) partiellement superposées c) source superposée sur différentes sources. La disponibilité de la vérité terrain nous a permis de calculer des mesures quantitatives comme l'erreur quadratique moyenne normalisée, l'erreur moyenne absolue normalisée et la distance spectrale pour les décours temporels et les carte spatiales des sources estimées avec notre algorithme. Nous avons obtenu une bonne estimation en présence de bruit et lorsque l'hypothèse de pixel pur n'est pas respectée pour toutes les sources.

Les applications du modéle générique sur les données quasi-réelles et réelles sont présentés dans les prochaines sections.

IRMf

Étant données les bonnes performances obtenues sur les données purement synthétiques avec le modèle proposé, nous l'avons utilisé sur des images IRMf (dimensions 3D + temps). Un des objectifs de cette thèse était d'analyser les données longitudinales d'un modèle de souris Alzheimer. Les données sont constituées d'une image IRMf de repos et d'une image anatomique 3D acquise lors du même examen. L'image anatomique a une dimension $256 \times 256 \times 34$ et une résolution spatiale de $0.08299 \times 0.07812 \times 0.4\ mm$. L'IRMf est de taille $147 \times 87 \times 27 \times 500$ avec une résolution spatiale de $0.1445 \times 0.2299 \times 0.5\ mm$ et $2s$ pour la résolution temporelle. L'atlas *Allen Brain Atlas* [8] fournit une image 3D du cerveau (template) et une carte de segmentation qui permet d'identifier $R = 613$ structures anatomiques d'intérêt dans le cerveau de la souris. Ces deux jeux de données ont une résolution spatiale

de $25 \times 25 \times 25$ μm.

Nous avons commencé avec les données brutes et avant le démélange, les étapes classiques de pré-traitement des données IRMf (masquage du cerveau , *slice timing*, correction du mouvement au cours de l'examen si besoin) et de recalage sont effectuées. On choisit de recaler l'atlas avec les données IRM anatomiques qui sont bien mieux résolues spatialement que les données IRMf. On tire parti du fait que les données IRMf sont quasiment parfaitement recalées entre elles (la souris n'a pas bougé) et on effectue d'abord un recalage non rigide des données anatomiques sur les données IRMf pour corriger les distorsions géométriques parfois observées entre les deux modalités. Les données IRMf voient leur résolution spatiale artificiellement augmentée au préalable en subdivisant chaque voxel en $3 \times 6 \times 2$ voxels (la valeur du voxel original est dupliquée dans les $J = 36$ sous-voxels). Ceci permet d'atteindre une résolution spatiale comparable à celles de l'atlas et des données anatomiques, sauf dans la troisième dimension qui correspond à l'épaisseur des tranches lors de l'acquisition IRM (qui ne peut être réduite qu'au détriment de la résolution temporelle). L'atlas est ensuite recalé sur les données anatomiques par recalage non rigide. Les deux étapes de recalages successifs fournissent des champs de déformation que l'on peut appliquer en cascade aux régions de la carte de segmentation de l'atlas haute résolution afin d'obtenir les masques binaires des R régions projetées sur les données IRMf artificiellement augmentées. Ces régions de l'atlas projetées sur les données IRMf permettent de construire la matrice A_tilde de notre modéle de démélange et donc de définir la contrainte. Un pipeline de traitements complet, entièrement automatisé, adapté aux souris pour préparer les cerveaux des souris pour la méthode de démélange a ainsi été réalisé pendant la thèse.

Afin de tester les performances de la méthode proposée en conditions réalistes, on utilise tout d'abord un jeu de données IRMf réelles dans lequel des signaux temporels synthétiques connus sont introduits. Nous avons introduit les signaux dans l'image IRMf augmentée afin de simuler le mélange avec les données réelles en dégradant ensuite l'image jusqu'à atteindre la (basse) résolution de l'IRMf d'origine. La contrainte sur les abondances pour le modéle peut écrire comme $g(\mathbf{A}) = \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_S(\mathbf{A})$ pour respecter la positivité, la somme égale un des abondances, et la contrainte spatiale ($\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$). Dans le cas des données IRMf la contrainte $h(\mathbf{U}) = \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2$ ne contient que le terme de Tikhonov. Empiriquement l'algorithme propose une solution acceptable et stable. Dans la Fig. 1 on peut regarder les signaux temporels estimés pour les différentes régions dans lesquelles ont été introduits des signaux synthétiques. Ceci nous a permis d'utiliser notre algorithme pour l'analyse des souris dans une étude longitudinale.

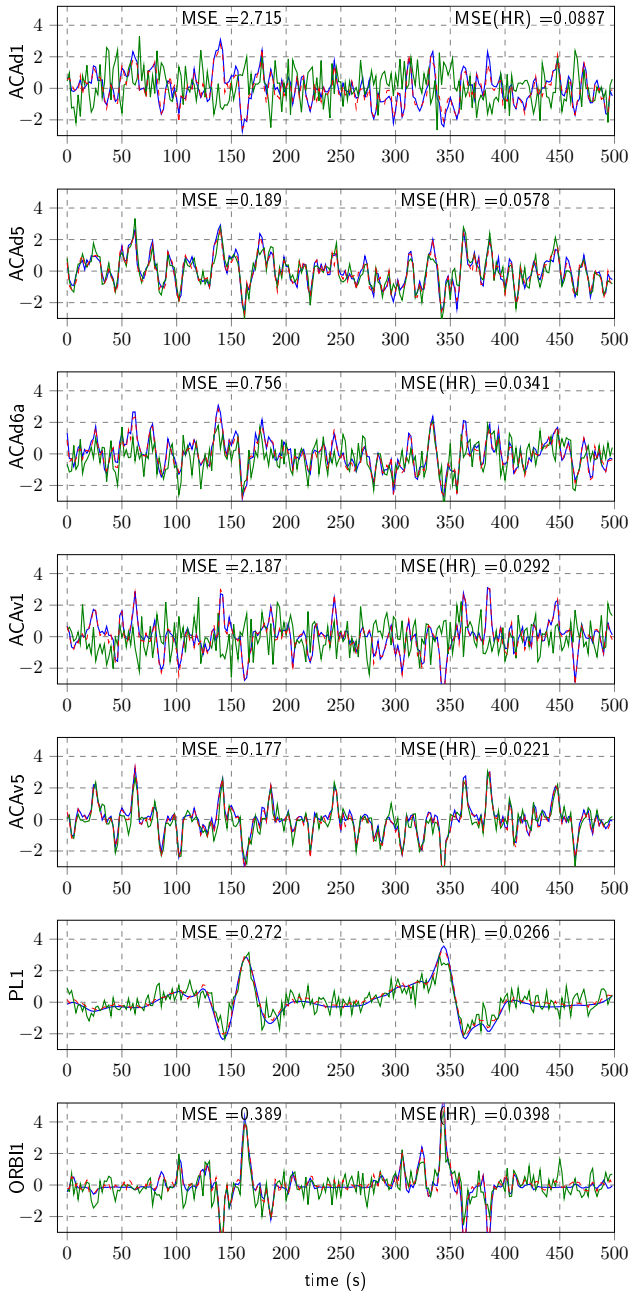


Figure 1: Données IRMf. Des échantillons correspondant aux 500 premières secondes des signaux synthétiques (en bleu) et leurs durées estimées correspondantes en utilisant les informations atlas des haute-résolution (en rouge pointillé) et sans utiliser les informations de l'atlas des haute-résolution (en vert). Les erreurs quadratiques moyennes (MSE pour mean square error en anglais) sont affichés au dessus des signaux.

Scintigraphie

Les algorithmes de démélange basés sur des régions spatiales d'intérêts sont également classiquement utilisés dans le domaine de la scintigraphie et nous avons testé notre algorithme sur ce type de données. Dans l'état de l'art, la méthode de référence RUDUR[9] calcule les sources avec des contraintes spatiales relaxées, c'est-à-dire que la contribution spatiale des sources peut dépasser les régions d'intérêt définies par l'expert. Nos contraintes spatiales, sous forme d'indicatrices sont des contraintes dures, *i.e.* les sources sont forcément estimées à l'intérieur des régions. Pour les données en scintigraphie la somme égale à un n'est pas importante sur les abondances. Aussi, la contrainte que les signaux temporels sont positifs doit être réspectée. Donc pour ce cas nous avons $g(\mathbf{A}) = \mathcal{I}_{\mathbb{R}^+}(\mathbf{A})$ et $h(\mathbf{U}) = \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{U})$ dans (3). Les résultats trouvés par notre méthode sont très proches des résultats trouvés par la méthode de réference [9], sur un cas quasi-réel de données scintigraphiques, montrant le large potentiel de notre modèle générique. Les résultats pour ces données sont montrés dans la Fig. 2.

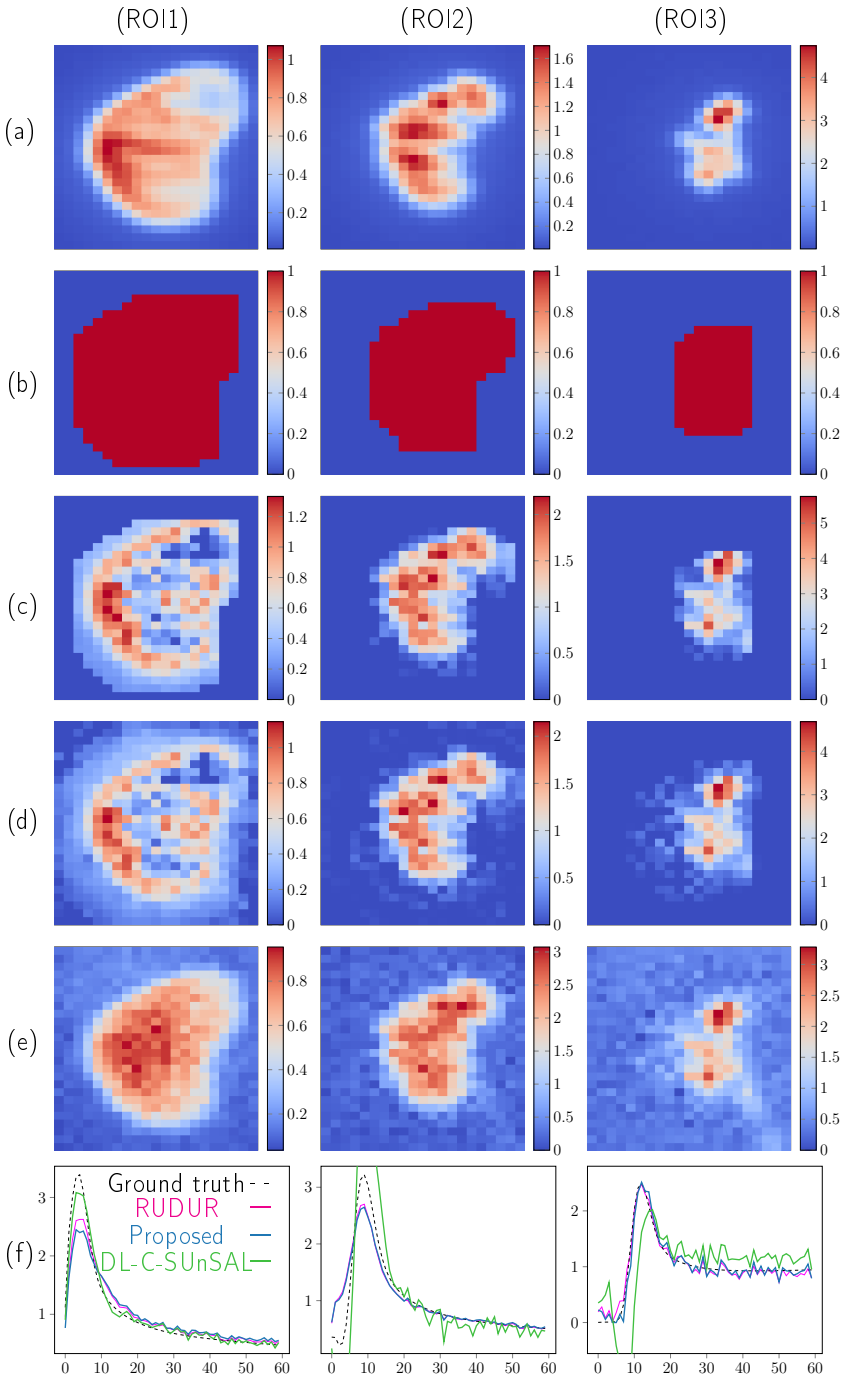


Figure 2: Données de scintigraphie. (a) vérité terrain pour les cartes spatiales, (b) regions initiales, (c) cartes spatiales estimées (d) cartes spatiales estimées par une méthode de l'état de l'art (RUDUR), (e) activités temporelles estimées par RUDUR et la méthode proposée.

Astronomie

Notre méthode a également été évaluée sur des observations astronomiques d'une portion du ciel observée par le téléscope spatial Hubble (HST) à haute résolution spatiale et par l'instrument MUSE au sol qui produit des images hyperspectrales de moins bonne résolution spatiale mais spectralement très résolues. Les régions d'intérêts ont été obtenues grâce à l'analyse préalable de l'image Hubble : [10] propose une carte de segmentation des galaxies observées. Le démélange spectral a ensuite été réalisé sur les données MUSE (Image 2D + spectres dans la troisième dimension) en utilisant la carte de segmentation de l'image Hubble pour exprimer la contrainte de localisation spatiale. Comme la résolution spatiale de l'image MUSE est faible, les spectres des galaxies dans l'image MUSE sont souvent noyés dans les spectres d'autres galaxies. En démélangeant une portion de l'image MUSE qui avait été analysée par le consortium qui a produit l'image, nous retrouvons, avec notre méthode, les mêmes estimations de spectres que dans leurs travaux, ce qui est une validation supplémentaire de notre approche. Le mélange des spectres de galaxies correspond à un mélange additif, donc nous enlevons la contrainte de somme égale à un. Les contraintes sont donnés par $g(\mathbf{A}) = \mathcal{I}_{\mathbb{R}^+}(\mathbf{A})$ et $h(\mathbf{U}) = \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{U})$. Les résultats sont montrés dans la Fig. 3. Également la méthode proposée a été lancée sur le champs entier des données MUSE à notre disposition (UDF-10).

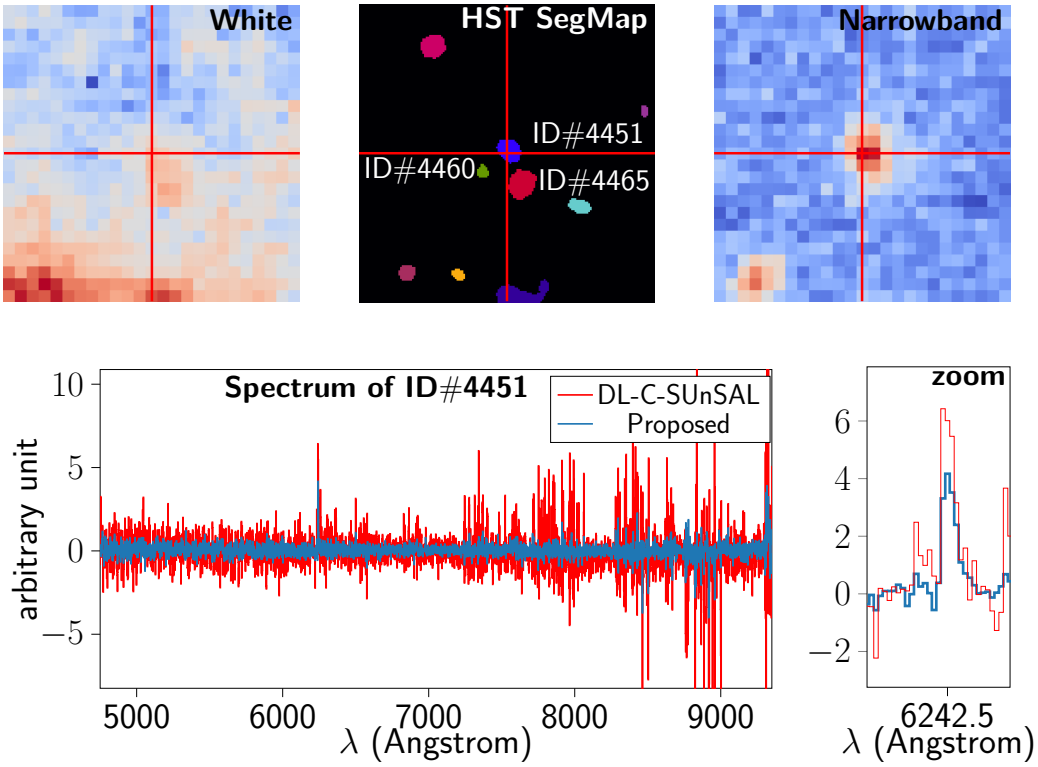


Figure 3: Données astronomiques. En haut, de gauche à droite, une image MUSE, une carte de segmentation de l'image HST Rafelski et une image à bande étroite centrée sur $\lambda = 6242.5$ Ang (position de la ligne d'émission dans le spectre estimé de la source ID#4451). La source centrale de Rafelski indiquée par la croix rouge est ID#4451. En bas, de gauche à droite : le spectre λ estimé pour ID#4451 sur toutes les de longueurs d'onde et zoom sur la ligne d'émission estimée à $\lambda = 6242,5$.

Conclusion

Le cadre de ce travail a évolué au fil du temps. Initialement, l'accent était mis sur le développement d'une méthode d'optimisation en tenant compte des aspects de l'IRMf. Dans le contexte de l'IRMf, l'application principale était de trouver la différence entre les souris AD et les souris WT. Selon l'hypothèse envisagée, chaque réseau fonctionnel est composé de régions anatomiques fines. On a envisagé d'utiliser un atlas avec une carte de segmentation pour la localisation de ces petites régions sur les images d'IRMf. Pour effectuer le démélange, des contraintes spéciales ont été ajoutées au problème d'optimisation pour faciliter le démélange. Pour les données d'IRMf, des études statistiques ont été nécessaires pour déterminer les différences entre les souris malades et les souris témoins.

Pour vérifier les différences entre les souris AD et WT, le test gaussien sur les données longitudinales a fourni des résultats intéressants. Les régions qui présentaient des changements de connectivité significativement différents pour les souris AD mais qui n'avaient pas subi de changements pour les souris WT ont été trouvées. Il a été observé que certaines de ces régions, comme l'entorhinal et les régions olfactives, sont également liées à la maladie d'Alzheimer chez l'homme.

Dans la phase ultérieure de ce travail, la généralisation du modèle de démélange nous a permis de l'appliquer à différents jeux de données tels que la scintigraphie et l'astronomie. L'algorithme d'apprentissage par dictionnaire avec des contraintes spatiales proposé a été comparé à l'état de l'art pour les différentes applications.

Contributions

Les contributions dans les différents domaines sont les suivantes :

- Par nature, les images IRMf ne contiennent aucune information sur la localisation des régions disponibles. Il peut donc être difficile d'analyser quantitativement les changements entre les souris témoins et les souris malades. Pour résoudre ce problème, un atlas haute résolution avec une carte de segmentation très détaillée (avec environ 600 régions anatomiques définies) a été recalé sur l'IRMf. Il convient de noter que les données de l'IRMf ont une résolution très faible et que la projection de l'atlas sur les données entraîne un chevauchement des régions. Cela nous a conduit à envisager une technique de démélange plus sophistiquée que la méthode des moindres carrés.

- Classiquement, une contrainte ℓ_1 est utilisée pour restreindre l'ensemble des solutions du fait du caractère mal posé du problème de démélange en induisant la parcimonie des sources dans le modèle. Pour la méthode de démélange proposée, la contrainte classique ℓ_1 pour les cartes spatiales/abondance a été remplacée par une contrainte spatiale sous la forme construite grâce à une connaissance a priori de la localisation potentielle des sources. Ceci permet un démélange respectant les limites des cartes spatiales initiales, alors que

la norme classique ℓ_1 ne contraint pas spatialement les sources. Les performances ont été validées sur un exemple synthétique avec différents cas de recouvrements que l'on peut trouver dans les données réelles. Une fois les erreurs évaluées pour les données synthétiques, l'algorithme a été appliqué à des données humaines réelles d'IRMf provenant d'un logiciel d'analyse d'IRMf populaire, SPM. Les résultats ont validé le modèle proposé pour les données d'IRMf. Enfin, la méthode a été appliquée à des données réelles de souris et interprétée par un neurobiologiste expert, avec des résultats compatibles avec ce qui était attendu pour les pathologies considérées dans ce modèle animal.

- L'hypothèse qui sous-tend l'utilisation d'un atlas de segmentation détaillé est que les réseaux fonctionnels sont constitués de petites régions anatomiques. Classiquement, les données d'IRMf sont recalées sur l'atlas et le démélange est effectué. Dans ce travail, l'atlas a été recalé sur les données d'IRMf. Un pipeline, différent de l'état de l'art, a été développé pour recaler la carte de segmentation hautement détaillée sur les données d'IRMf sans trop altérer les signaux temporels originaux. L'hypothèse et le pipeline présentés ont conduit au développement de codes adaptables à différents ensembles de données et utilisés en interne dans l'équipe.

  Afin de comparer la connectivité du cerveau, l'analyse des matrices de corrélation est l'approche privilégiée dans cette étude. Pour la validation, des signaux quasi-réels avec des corrélations connues ont été introduits dans les régions préfrontales, et le démélange a été effectué. Les signaux estimés et les matrices de corrélation ont confirmé la bonne performance de l'approche de démélange. Différents types de validations statistiques ont été effectués pour rechercher des changements longitudinaux sur les matrices de corrélation des données quasi-réelles et synthétiques. La validation des tests statistiques nous a permis d'effectuer les tests statistiques sur des données réelles.

- La généralisation du modèle a été détaillée pour adapter le modèle de démélange à contrainte spatiale à d'autres applications. Des résultats similaires à l'état de l'art ont été fournis pour l'application scintigraphique et astronomique. Le travail sur ces deux applications a nécessité une compréhension des différentes modalités impliquées. Les résultats prometteurs sur des données quasi-réelles en scintigraphie prouvent l'applicabilité de l'algorithme sur des données scintigraphiques réelles. Comme le problème proposé est adaptable, de l'information à priori sur les signaux temporels pourraient être ajoutés si nécessaire. Les données hyperspectrales pour l'application astronomique sont produites par l'instrument MUSE, et les connaissances externes sur l'emplacement des sources proviennent d'un catalogue développé à partir des observations du télescope Hubble en [10]. L'algorithme a pu fournir des résultats de démélange efficaces pour l'ensemble des données UDF-10 de MUSE, prouvant ainsi son évolutivité. Divers outils d'analyse des données ont été fournis avec un pipeline

permettant d'estimer les spectres de galaxies présentant des lignes d'émission faibles ou de retrouver des galaxies qui n'existent pas dans le catalogue initial mais qui sont confirmées par d'autres approches. La généralisation du modèle a donné lieu à une publication dans IEEE Transactions on Image Processing.

Perspectives

Les perspectives sont divisées en perspectives basées sur les modèles et en perspectives basées sur les applications.

Modèle

- Introduction d'un lissage spatial : Le modèle de démélange proposé dans le manuscrit effectue le démélange en considérant les pixels indépendamment et, de plus, les estimations d'abondance sont mises en parallèle selon les pixels dans l'implémentation de l'algorithme. Afin d'introduire plus de cohérence spatiale dans chaque carte d'abondance, une contrainte de lissage spatial dans les régions pourrait être introduite dans une future variante de l'algorithme. Cependant, en fonction de la nature de la contrainte ajoutée, cela augmenterait la complexité de l'algorithme. L'une des façons d'ajouter une telle contrainte est sous la forme d'une régularisation par variation totale (VT). Le problème d'optimisation à contrainte spatiale prenant en compte les pixels voisins pour les cas impliquant des images 2D avec une information temporelle ou spectrale a été donné.

- En ce qui concerne les aspects calculatoires de l'approche proposée dans le deuxième chapitre, certaines améliorations pourraient être apportées pour augmenter la vitesse de l'estimation. Une implémentation GPU parallélisée pourrait être mise en oeuvre pour accélérer les calculs dans l'algorithme d'optimisation. La projection sur les ensembles convexes est actuellement effectuée à l'aide de l'algorithme de Michelot [11]. L'une des améliorations possibles est l'implémentation de [12] pour une projection plus rapide sur les ensembles convexes, ce qui pourrait éventuellement réduire le temps de calcul.

Applications

- fMRI : Afin d'éviter de modifier les données IRMf, celles-ci ont été conservées dans leur résolution d'origine et un atlas a été projeté sur les données IRMf. Cette étape est différente des études de la littérature où les données IRMf sont enregistrées sur l'atlas, puis un algorithme de démélange est exécuté sur les données. Notre approche a impliqué un pipeline de prétraitement qui pourrait constituer la base d'un futur article avec l'ajout d'une analyse technique supplémentaire.

Quelques régions de l'hippocampe ont été divisées en plusieurs sous-régions sur la base de preuves d'activités fonctionnelles différentes à l'intérieur de ces régions. Cette analyse par noyaux à plus grande échelle pourrait également être effectuée pour d'autres régions pour lesquelles de telles preuves peuvent être faites. Une façon d'y parvenir est d'effectuer une ICA à l'intérieur d'une région particulière, puis de vérifier les cartes spatiales estimées. Dans un atlas de segmentation très détaillé, à sa résolution la plus fine, certaines régions anatomiques fines sont divisées en plusieurs couches fines. Ces couches se chevauchent à l'échelle de résolution inférieure de l'IRMf et dans ce cas, il est moins approprié de conserver cette segmentation fine. L'idée de fusionner les régions anatomiques pour rester à une échelle anatomique supérieure permettrait d'améliorer la difficulté du problème en réduisant les dimensions du problème d'estimation (par exemple, en améliorant le conditionnement de la matrice $\mathbf{A}$).

Au cours de ce travail, les applications IRMf ciblées concernaient les données de souris en état de repos. Bien qu'une application basée sur une tâche humaine soit présentée, de futures applications pourraient impliquer le démélange sur des données IRMf humaines au repos. D'autres ensembles de données impliquant un cryosonde ont un faible bruit, et l'approche que nous proposons pourrait être intéressante pour ces données. Un bruit plus faible entraînerait une meilleure estimation des signaux temporels et donc de meilleures matrices de corrélation, ce qui pourrait améliorer le regroupement des régions anatomiques en réseaux fonctionnels.

- Astronomie : Divers résultats corroborent que la méthode proposée peut être appliquée pour le démélange de données hyperspectrales lorsqu'une carte de segmentation des galaxies est disponible. Théoriquement, la méthode proposée devrait fournir des spectres mieux estimés que les spectres estimés par sélection manuelle par des experts lorsque les signaux sont contaminés par les signaux d'autres sources. L'analyse des résidus après démélange des sources connues grâce au catalogue [10] a permis de trouver une galaxie qui n'était pas présente dans la carte de segmentation. Cette galaxie a été également détectée par d'autres méthodes développées dans le cadre de l'analyse des données du projet MUSE. Il faut maintenant explorer l'ensemble du cube MUSE pour rechercher d'autres sources dans les résidus et relancer l'algorithme en ajoutant les nouvelles sources dans les cartes initiales afin d'affiner les résultats. Le pipeline permettant de trouver des galaxies avec des lignes d'émission faibles dans les résidus pourrait être utilisé pour valider d'autres galaxies après discussion avec les astronomes.

Liste de publications

Article de revue internationale avec comité de lecture

[1] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Spatially Constrained Online Dictionary Learning for Source Separation, IEEE Transactions on Image Processing, février 2021. DOI: 10.1109/TIP.2021.3058558

Articles de conférence avec comité de lecture

[2] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Online Dictionary Learning for Single-Subject fMRI Data Unmixing, EUSIPCO, A Coruña, Espagne, septembre 2019. DOI: 10.23919/EUSIPCO.2019.8902991
[3] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf, Gretsi, Lille, France, août 2019.

Conférence avec résumés

[4] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen Mouse Brain en IRMf chez la souris, SFRMBM, Strasbourg, France, Mars 2019.

Communications

[5] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Online dictionary learning for single-subject fMRI data unmixing: présentation orale, Journée Data Science and Artificial Intelligence (DSAI), ICube, Strasbourg, 2019.
[6] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf; poster, 7èmes Journées scientifiques de la Fédération de la Médecine Translationnelle (FMTS) de Strasbourg, 2019.
[7] A. Bhanot, C. Meillier, F. Heitz, L. Harsan, Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen mouse brain en IRMf chez la souris. Journée poster des doctorants ICube, 2018.

Chapter 1

Source separation state-of-the-art

Contents

1.1 Introduction to source separation

The issue of source separation, or unmixing, is well known to the signal and image processing community. It concerns a very large number of applications and can occur under different conditions of source mixing. A large part of the literature is devoted to blind source separation (BSS) [13]–[15]. BSS methods allow solving *cocktail party* problems for which P signals (or images) composed of a mixture of R sources are observed, without any *a priori* on the properties of the sources. If the number of sources involved in the mixtures is not known *a priori*, then it must also be estimated [16], [17].

The first record of a BSS method is in 1901, where an application of Principal component analysis (PCA) was presented by Pearson [18], later it was developed independently by Hotelling in [19] and started developing around the 1960s by Malinowski. In signal processing, the PCA or Karhunen-Loéve transform marks the beginning of source separation starting with [13], where they used a principal component model to explain the observed noisy data, $\mathbf{Y}$. It was followed by early BSS methods that mainly comprised Independent Component Analysis (ICA) [14], [15], [20] followed by sparse decomposition analysis [21], [22]. The first article on ICA

for the international community was published in the early 1990s [14] with an even earlier article in French from 1988 [23]. ICA and PCA require the specification of the number of components, whereas in sparse decomposition analysis, a prior is based on the idea that the number of involved sources in a given observation is low, which are generic priors under blind source separation [24]. Many variants of the ICA [14], [15], [25] approach have been proposed in the literature to solve BSS problems. All of them are based on the general principle of spatial independence of the sources, which makes it possible to estimate their temporal (or spectral) signatures.

For instance, for brain functional networks detection in functional Magnetic Resonance Imaging (fMRI) data, ICA is widely used to separate spatial sources by assuming the independence of the temporal signals associated with each spatial source, *i.e.* functional network. Spatial ICA has proven effective in [26], [27] for fMRI data, but the main drawback of ICA approach is the unknown number of sources which is set arbitrarily and may lead to a large number of nuisance sources that must be screened manually or by a semi-automatic method [28]. In neuroscience, SPM [29] has various ICA algorithms implemented for fMRI analysis. ICA is used in group studies, where the redundancy of the information from different subjects is useful. In the case of single-subject studies or longitudinal studies, change detection methods that do not solely depend on the independence of the sources are required.

In contrast to the BSS problem, many unmixing problems involve a dictionary of pre-defined bases such as Discrete Cosine Transform (DCT), wavelets, and curvelets [30], [31]. The dictionary columns or *atoms* may also be fixed; for example, in hyperspectral imaging for remote sensing, libraries of light spectra corresponding to the different materials that may be observed in the scene are available [32], [33], so that only the proportion of the different materials in each pixel is estimated. Between these two extreme cases, there are a large number of unmixing problems where some information on the form or location of the sources or the type of mixture is known [30], [31], [34]. As the problems for BSS methods are ill-posed, it is necessary to add constraints to reduce the size of the solution space, in other words, search space. *Sum-to-one* and *positivity* constraints on the coefficients of the mixing matrix are classic in signal and image processing [35]. For mixing matrices in remote sensing, these are known as the abundance sum-to-one-constraint (ASC) and abundance non-negativity constraint (ANC). In remote sensing applications, hyperspectral data linear unmixing is carried out by methods based on nonnegative matrix factorization [36], [37]. In recent years, sparse decomposition methods have been widely used to solve source separation problems [22], [33], [38], [39]. The sparsity constraint is another way to reduce the set of solutions [40]. It can be combined with the two latter constraints. The sparsity may concern the mixing itself, *i.e.* for a given observed signal, only a few number of sources is involved, or the decomposition of the sources on a dictionary (wavelet, discrete cosine transform, or custom atoms (containing sparse signals with most of the intensities with values equal to zero), known orthogonal atoms, or non-orthogonal possible in some cases) [30], [31], [41]. Dictionary

learning methods take into account the spatial sparsity of the sources in the form of ℓ_1 constraints on the mixing matrix in the minimisation problem. Recently for fMRI applications, where sources are functional networks, sparse analysis based on dictionary learning methods has proven to be promising [38], [39], [42], [43]. In the hyperspectral domain, provided a library of spectra, Constrained-Sparse Unmixing by variable Splitting and Augmented Lagrangian (C-SUnSAL) [44] is a classical algorithm for solving optimisation problem with the *sum-to-one*, *positivity* condition and an ℓ_1 constraint on the spatial maps matrix.

Classical linear model

The classical linear model used in source separation in the presence of an i.i.d. noise may be written as:

$$\mathbf{Y} \simeq \mathbf{UA}, \tag{1.1}$$

where $\mathbf{Y} \in \mathbb{R}^{N \times P}$ is the observed data. In spatio-spectral separation, N can be interpreted as the spectral length, and in spatio-temporal separation, N is the length of the temporal signals. P is the number of voxels or pixels, depending upon the dataset. Matrix $\mathbf{U} \in \mathbb{R}^{N \times R}$ contains the temporal/spectral signatures where R is the number of sources. Matrix $\mathbf{A} \in \mathbb{R}^{R \times P}$, usually called mixing or abundance matrix, codes the fraction of the R components contributions at each voxel or pixel. Notations and the model are graphically represented in Fig. 1.1. If $R < P$, the unmixing problem is overdetermined, and if $R > P$, then it is underdetermined.

In the vector form, for the observed signal for a given pixel p along with noise $\boldsymbol{\eta}_p$, the mixing model can be written as:

$$\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{u}_r \mathbf{a}_{r,p} + \|\boldsymbol{\eta}_p\|_2, \tag{1.2}$$

where $\mathbf{y}_p$ is the observed signal for the p^{th} voxel/pixel, r is the indice of the r^{th} region, $\mathbf{u}_r$ is the r^{th} column of $\mathbf{U}$ and $\mathbf{a}_{r,p}$ is the proportion of the r^{th} region in the p^{th} voxel.

The approach followed for the taxonomy of the state-of-the-art algorithms explained in this chapter is based on their utilisation in spatio-temporal or spatio-spectral unmixing. PCA and ICA lay the foundations for blind source separation and are explained at the beginning, followed by dictionary learning methods which recently are being widely used for fMRI data (our principal application). The geometric approaches are presented in the next section, followed by the Bayesian approach. The last section is dedicated to non-linear methods and tensor-based models.

1.2 PCA

Along with its popularity as a dimension reduction and a visualisation tool, PCA is an indispensable tool used to separate sources based on maximal variance. To

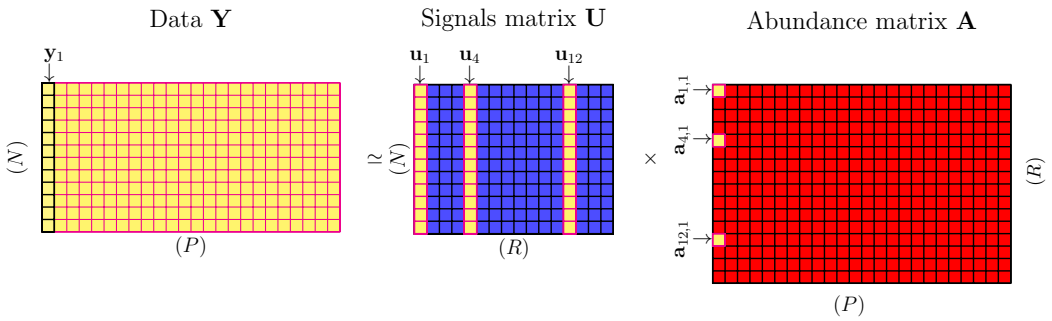


Figure 1.1: Matrix representation of the classic linear model. Here the column $\mathbf{y}_1$ is the signal of the first element of the image. This signal can be considered to be composed of $\mathbf{u}_1$, $\mathbf{u}_4$ and $\mathbf{u}_{12}$ with non-null abundances equal to $\mathbf{a}_{1,1}$, $\mathbf{a}_{4,1}$ and $\mathbf{a}_{12,1}$ respectively.

understand PCA for blind source separation, let us consider the $N \times P$ matrix $\mathbf{Y}$, where N is the number of observations of the P variables (i.e. each pixel/voxel is considered as a variable). The idea of PCA is to find a basis with dimensions less than the dimension of the variables P. Taking our model as the base, PCA analyses the variance-covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{P \times P}$, of the P variables observed in the matrix $\mathbf{Y} \in \mathbb{R}^{N \times P}$. This matrix is used to find a basis with the axes corresponding to different uncorrelated sources, in other words, orthogonal spatial maps. A linear combination of these maps explains the observations. Depending upon the application and requirements, the data may be standardised. In this case, PCA components are found using the correlation matrix instead of the variance-covariance matrix. The implementation of PCA remains the same whether we use the correlation matrix or the covariance matrix. So for convenience, let us consider that $\mathbf{\Sigma}$ is the variance-covariance matrix and that the columns of $\mathbf{Y}$ are demeaned/centered.

The matrix $\mathbf{\Sigma}$ being a real and symmetric square matrix, can be diagonalised to an orthonormal basis. In PCA, during the diagonalisation of $\mathbf{\Sigma}$, we make sure to sort the eigenvalues in descending order so that the first principal component/vector explains the inertia of the data cloud (corresponding to the component with maximal variance), the second component explains the remaining inertia, and so on. In the end, among the P principal components, we would keep just the first R components, considering them as the sources and the others as noise.

The covariance matrix $\mathbf{\Sigma}$, can be written as:

$$\mathbf{\Sigma} = E(\mathbf{Y}^T\mathbf{Y}), \tag{1.3}$$

where $E(.)$ is the expectation and $(.)^T$ is the transpose of the matrix.

Eigenvalues and eigenvectors need to be found for (1.3), and the first component or the first eigenvector explains the maximum variance of the datacloud. This can

be done by a variety of methods, out of which the most popular is the singular value decomposition method (SVD). The SVD method is the general form of eigenvalue decomposition for any 2D matrix. If SVD decomposition of $\mathbf{Y} = \mathbf{QSV}^T$, then:

$$\mathbf{Y}^T\mathbf{Y} = \mathbf{VSQ}^T\mathbf{QSV}^T = \mathbf{VSSV}^T = \mathbf{V\Delta V}^T. \tag{1.4}$$

In the above equation $\mathbf{V}$ contains the right singular vectors, $\mathbf{S}$ is a diagonal singular matrix, $\mathbf{Q}$ contains the left singular vectors and $\mathbf{\Delta} = \mathbf{S}^2$. Optimisation problem for PCA can be explained in various forms, which have been detailed in [45]. In PCA terminology, $\mathbf{V}$ are the eigenvectors, principal axes or the principal plane direction coefficients. The projection of the data onto the principal axes provides the principal scores i.e., $\mathbf{YV}$. For our particular model, considering $\mathbf{A}$ as the sources: $\mathbf{V}$ can be replaced with $\mathbf{A}$ i.e., the spatial map matrix. In this case, each row of $\mathbf{A}$ is a PCA component.

1.3 ICA

Consider the same $N \times P$ matrix $\mathbf{Y}$, where N is the number of observations of the P variables (i.e., each pixel/voxel is considered as a variable). Independent component analysis (ICA) aims at decomposing $\mathbf{Y}$ as a mix of R independent sources $\mathbf{Y} = \mathbf{UA}$, where $\mathbf{A} \in \mathbb{R}^{R \times P}$ contains the independent spatial sources and $\mathbf{U} \in \mathbb{R}^{N \times R}$ stands for the mixing matrix. ICA provides an estimation of the sources $\mathbf{A} : \hat{\mathbf{A}} = \mathbf{WY}$, where $\mathbf{W}$ is a linear transform matrix that maximises the statistical independence of rows of $\hat{\mathbf{A}}$. This statistical independence can be measured through mutual information (from information theory), the non-Gaussianity, and the maximum likelihood. The idea is to use one of these measures as an objective function called the contrast function. The optimisation problem involving the contrast function aims to increase the difference between a Gaussian distribution and the independent sources; these sources found are called the independent components. Some well known ICA algorithms are FastICA [46], INFOMAX [47], JADE [48] and kernel ICA [49], each differing in the way the separation matrix is updated.

The uncorrelated and independent sources found in ICA are separated non-linearly using higher-order statistical moments and not just variance as in the case of PCA. Another notable difference between the new R basis formed by PCA from those in PCA is that in ICA, the basis is orthogonal, whereas in PCA, it is orthonormal.

Let us consider sources/components $\hat{\mathbf{A}}$, or the spatial maps matrix to be estimated, given by:

$$\hat{\mathbf{A}} = \mathbf{WY}. \tag{1.5}$$

In ICA, the idea is to estimate $\mathbf{W}$ i.e., the separation matrix, given just the data at disposition. Once $\mathbf{W}$ is known, the mixing matrix i.e. $\mathbf{U}$ in the case of spatial ICA can be estimated by $(\mathbf{W}^T\mathbf{W})^{-1}\mathbf{W}^T$ (comparing to our model $\mathbf{Y} \simeq \mathbf{U}\hat{\mathbf{A}}$). The columns of $\mathbf{U}$ correspond to spatially independent signals with the components being

the spatial maps. In this case, we are interested in spectro/temporal data that is spatially structured. This kind of ICA to find spatially independent signals can be noted as spatial-ICA (s-ICA). Spatial ICA has recently become popular in fMRI, where the interest is to find the different regions of the brain presenting different activations [50].

ICA can also be applied on Y^T(this is the classical implementation of ICA where the sources are temporal or spectral) and is called temporal-ICA (t-ICA) or even spectral ICA. To find the temporally independent sources or components for the $\mathbf{Y}$, ICA must be performed on $\mathbf{Y}^T$. In that case, the sources to be estimated i.e. $\mathbf{U}^T$ can be written as:

$$\mathbf{U}^T = \mathbf{W}\mathbf{Y}^T. \tag{1.6}$$

Implementation of ICA involves iterative updates for the separation matrix $\mathbf{W}$ until convergence. The rows of $\mathbf{W}$ represent the spatial maps, and the rows of $\mathbf{U}$ are the temporally independent signals.

Both t-ICA and s-ICA with results on data containing various combinations of temporal and spatial independence that could occur in task-based fMRI data are given in [50].

1.4 Dictionary learning

For ICA and PCA, the matrices $\mathbf{A}$ and $\mathbf{U}$ are estimated by imposing the constraints of statistical independence of sources in the former and the orthogonality and the maximal variance for the latter. Taking into consideration other types of constraints, the optimisation problem for $\mathbf{A}$ and $\mathbf{U}$ can be written in the following manner:

$$\min_{\mathbf{U},\mathbf{A}} \|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \lambda\Phi(\mathbf{U}, \mathbf{A}), \tag{1.7}$$

where $\Phi(\mathbf{U}, \mathbf{A})$ summarises the constraints on matrices $\mathbf{U}$ and $\mathbf{A}$. Changing $\Phi(\mathbf{U}, \mathbf{A})$ leads to different models. For example $\Phi(\mathbf{U}, \mathbf{A})$ can be separated as a sum of $\Phi_1(\mathbf{U})$ and $\Phi_2(\mathbf{A})$.

For the optimisation problem in 1.7, both $\mathbf{A}$ and $\mathbf{U}$ are unknown. A dominant approach in the literature for such dictionary learning problem is alternate optimisation where the matrices $\mathbf{A}$ and $\mathbf{U}$ are estimated alternatively. The initialisation is important in this case, and convergence of the algorithm is not assured in the presence of noise or unavailability of priors. Pseudocode for such alternating strategy is provided in Alg. 2.

Regularisations in dictionary learning

Some popular regularisations for dictionary learning have been explained as follows.

- There is a general interest for sparse models in dictionary learning to explain data as a linear combination of a few elements. For this purpose, some sparsity

```
Initialise A
while STOPPING CRITERIA ≠ TRUE do
    Estimate U by minimising the U sub-problem
    Estimate A by minimising the A sub-problem.
end
return A, U
```

Algorithm 2: Alternate optimisation pseudo-code to estimate $\mathbf{A}$ and $\mathbf{U}$. Role of $\mathbf{A}$ and $\mathbf{U}$ can be switched if initialising $\mathbf{U}$ is easier than initialising $\mathbf{A}$.

constraints are required in the optimisation problem. For sparse decomposition methods $\Phi(.) = \|.\|_x$, where x could be $\{0\}, \{1\}, \{2\}, \{0,1\}, \{1,0\}, \{1,1\}$, $\{1, \text{inf}\}$ etc. and $\|.\|$ is the norm of the rows/columns of matrix $\mathbf{A}$. The choice of 'x' depends on the amount of sparsity required. Mixed norms for sparsity have been well explained in [51].

- To introduce smoothing in the decomposed data, total variation techniques could be used. Total variation is high for a particular element if the neighbouring elements have dissimilar intensities and vice-versa. For total variation regularisation on 2D spatial maps, Φ can be written as

$$\Phi_{iso}(\mathbf{A}) = \sum_{k,l} \sqrt{|a_{k+1,l} - a_{k,l}|^2 + |a_{k,l+1} - a_{k,l}|^2} \tag{1.8}$$

  or for the anisotropic version, as:

$$\Phi_{aniso}(\mathbf{A}) = \sum_{k,l} |a_{k+1,l} - a_{k,l}| + |a_{k,l+1} - a_{k,l}|. \tag{1.9}$$

  In 1.8 and 1.9, k and l refer to the pixel indices in the actual 2D image. Total variation can similarly be used for signals smoothing and can also be extended to 3 dimensions for data consising of 3D spatial volumes.

- An important state-of-the-art regularisation on signals is $\Phi(\mathbf{U}) = \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2$, know as Tikhonov regularisation in the literature. Also popularly known as ridge regression, it is used to prevent multi-collinearity in linear regression. This regularisation also promotes smoothness in the temporal signals to be estimated.

Application-specific constraints can be also be defined in dictionary learning. For example, in remote sensing hyperspectral imaging, there exists a popular method called SUnSAL and its constrained version constrained-SUnSAL (C-SUnSAL) with the classical constraints of *sum-to-one* and *positivity* along with the $\ell 1$ sparsity of

the spatial maps. For fMRI data, a compressed online dictionary learning algorithm exists [39]. In scintigraphic imagery, RUDUR is a recent algorithm making use of the priors on regions of interest. These different constraints and methods interest us in processing data from different applications that we aim to study in this work.

- In the region-of-interest based algorithm RUDUR [9], $\Phi(\mathbf{A})$ contains the sparsity as well as a regularisation term that accounts for distances between between the ROIs. If i and j are two random pixels in the data, the distance of voxel i to the ROI r is calculated as :

$$D_{i,r} = \begin{cases} \min_{j/\tilde{\mathbf{A}}_{j,r}=1} dist_{euc}(i,j), \text{ for i and j not belonging to the same region,} \\ \text{otherwise 0.} \end{cases} \tag{1.10}$$

  Here $\tilde{\mathbf{A}}$ is the binary image of the initial ROIs and has the same dimension as that of $\mathbf{A}$, $\tilde{\mathbf{A}}_{j,r} = 1$ if an only if the pixel j belongs to the ROI r, and $dist_{euc}(i,j)$ is the euclidean distance between the pixels i and j. The distance $D_{i,r}$ is used to penalise the minimisation problem in RUDUR; the larger is the distance between the pixel i and some ROI r, the lower is its chance to belong to that particular ROI. In the absence of such a constraint all the pixels would be treated equally by the unmixing model.

- A well known algorithm in the hyperspectral imaging community for the estimation of abundance maps in hyperspectral images unmixing called SUnSAL [44] adds sparsity in the form of $\ell 1$ norm. If $\Phi(\mathbf{U},\mathbf{A})$ is equal to $\mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_S(\mathbf{A}) + \|\mathbf{A}\|_1$ the problem becomes similar to the one presented in the same work [44] by the name of C-SUnSAL.

1.5 Geometric approaches

Geometrical approaches are blind source separation methods popular in hyperspectral imaging; many of them have been summarised in [37]. These methods are based on the premise that data from a linear mixture of vectors lie in a simplex. All such methods involve some basic preprocessing techniques to project data (in our case $\mathbf{Y}$) in the form of data clouds in a reduced subspace of dimension $R-1$ i.e., the number of material (water, soil, buildings, etc.). Linear mixing subject to certain mathematical constraints forces the data clouds to be bounded in a simplex, or a cone in a $(R-1)$ dimension subspace. In the presence of constraints such as *sum-to-one* and *positivity* on the abundances, each linearly mixed vector $\mathbf{y}_p$, or the observed signal for a voxel/pixel, after dimensionality reduction can be written as:

$$\mathbf{y}_p = \sum_{r=1}^{R} \mathbf{u}_r a_{r,p} + \|\boldsymbol{\eta}_p\|_2, \text{ s.t.} \sum_{r=1}^{R} a_{r,p} = 1, a_{r,p} >= 0, \tag{1.11}$$

where r is the r^{th} region, p is the p^{th} voxel/pixel, $\mathbf{u}_r$ is the r^{th} column of $\mathbf{U}$ and $a_{r,p}$ is the proportion of the r^{th} region in the p^{th} voxel.

The endmembers i.e., the sources (columns of $\mathbf{U}$), with the abundance matrix $\mathbf{A}$ under the *positivity* and *sum-to-one* constraint are considered to form a simplex which must englobe the data cloud formed by $\mathbf{y}_p$'s. Simplex formed by a dataset in the reduced dimension subspace having three different kinds of sources has been illustrated in Fig. 1.2. Some geometrically driven algorithms based on the assumption of pure pixels to find the endmembers are PPI, NFINDR, and VCA. They have been described as follows :

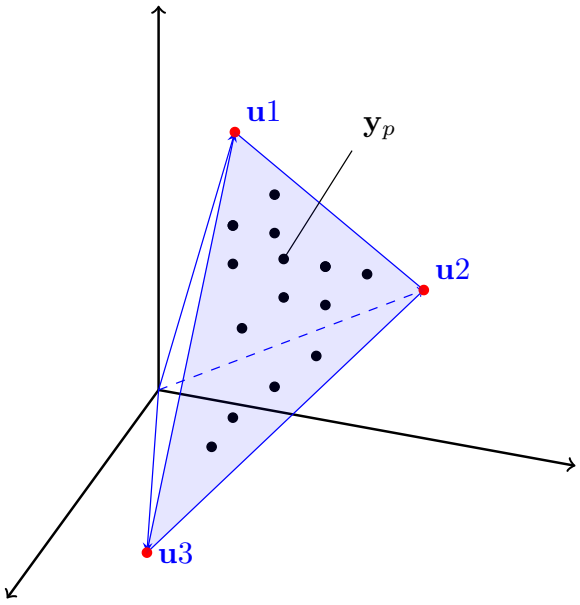


Figure 1.2: Data points $y_p \in \mathbb{R}^3$ are contained by the 2D-convex cone formed by the sources or the endmembers $\mathbf{u}_1$, $\mathbf{u}_2$ and $\mathbf{u}_3$

.

Pixel purity index(PPI)

To implement PPI, minimum noise fraction (MNF)[52] is a necessary preprocessing step that results in dimension reduction as well as whitening of data. Also, MNF arranges the components in terms of decreasing SNR of the images so one can select the components corresponding to the images. MNF gives the same result or order sequence for components as PCA. Once the data is reduced, the PPI [53], [54] method is implemented. This can be summarised as: counting the number of times the projection of a data point on a large number of generated random unit vectors results in it being found on the extremities of the vectors, and then declaring the most frequent points on the extremities as pure pixels. Here, the large set of random

unit vectors are called *skewers*.

N-FINDR

For N-FINDR [55], the simplex formed by the endmembers is assumed to be formed by pure pixels. Using this fact, the algorithm starts with an initial guess of the endmembers, then the volume formed by the endmembers is calculated. If there is a suitable pixel candidate, not contained in the existing volume, that could replace one of the present vertex of the simplex to result in a larger final volume, then the current vertex is replaced. It goes on iteratively, "inflating" the volume of the simplex at each iteration until no suitable candidates are left to replace the endmembers. In the absence of pure pixels, the algorithm may result in a mixed spectra for the estimated end members i.e., $\mathbf{u}_i$'s will have a higher correlation between them than it should be the case.

Vertex component analysis (VCA)

Like the other algorithms mentioned before, this one also exploits the fact that the vertices of the simplex are formed by the endmembers. Another property that VCA [56] exploits is that the affine transformation of a simplex results in a simplex. Making use of these properties, the cone formed by the already determined endmembers is projected to the orthogonal subspace in a repeated manner. As the transformations would result in a simplex, the convergence is achieved when all the members lie in the subspace simplex of dimension R, where R is the true number of endmembers.

Minimum volume approach

In minimum volume approaches, pure pixels are not necessary in the dataset but there should be at least $R-1$ members lying on each facet of the simplex for this algorithm to work. The optimisation problem can be cast as:

$$\min_{\mathbf{U},\mathbf{A}} \|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \lambda\|V(\mathbf{U})\|^2, s.t. \mathbf{A} >= 0, U >= 0, \mathbf{1}^T\mathbf{a} = \mathbf{1}_R^T. \tag{1.12}$$

Here, the volume $\|V(\mathbf{U})\|$ is proportional to the determinant of the simplex (formed by endmembers) in matrix form, i.e. $\det(\mathbf{U})$. MVSA [57] and SISAL [16] are two algorithms that follow the minimum volume approach but allow for violations of the *sum-to-one* constraint.

1.6 Bayesian inference

Probabilistic models present a different approach from other statistical source separation methods by defining explicit parametric forms of prior distributions to define

constraints. If $\mathbf{Y}$ is the observed data with $\mathbf{A}$ and $\mathbf{U}$ as the parameters to be estimated then the posterior probability in the Bayesian paradigm can be written as:

$$P(\mathbf{U}, \mathbf{A}|\mathbf{Y}) = \frac{P(\mathbf{Y}|\mathbf{U}, \mathbf{A})P(\mathbf{U}, \mathbf{A})}{P(\mathbf{Y})}. \tag{1.13}$$

In 1.13, P($\mathbf{Y}$|$\mathbf{U}$,$\mathbf{A}$) is the likelihood of $\mathbf{Y}$ given the distribution of $\mathbf{A}$ and $\mathbf{U}$, P($\mathbf{U}$) is the prior defining the probability distribution of the temporal/spectral signals and similarly P($\mathbf{A}$) for the abundances. P($\mathbf{Y}$) is the probability distribution of the data or the evidence. Estimation of posterior distributions of $\mathbf{U}$ and $\mathbf{A}$ would provide knowledge about the confidence intervals of parameters to be calculated which then can be used to infer the values of $\mathbf{A}$ and $\mathbf{U}$. The procedure for Bayesian inference requires defining some priors and then estimate the posteriors, but choosing priors can be complicated and may require application-specific expert knowledge or experience. For example, truncated normal priors with just the positive support could be introduced to force non-negative abundances in the estimation of $\mathbf{A}$. For the model given in 1.2 if $\|\boldsymbol{\eta}_p\|_2$ is an i.i.d. Gaussian white noise, then the likelihood is given by :

$$P(\mathbf{y}_p|\mathbf{U}, \mathbf{a}_p, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left[-\frac{\|\mathbf{y}_p - \mathbf{U}\mathbf{a}_p\|_2}{2\sigma^2}\right]. \tag{1.14}$$

σ^2 also to needs be estimated if not known a priori. As the noise is i.i.d, the model can be written as :

$$P(\mathbf{Y}|\mathbf{U}, \mathbf{A}, \sigma^2) = \Pi_{p=1}^{P} f(\mathbf{y}_p|\mathbf{U}, \mathbf{a}_p, \sigma^2). \tag{1.15}$$

The likelihoods presented in 1.14 and 1.15 for a linear hyperspectral model are similar to the problem solved in [58]. Interesting priors distributions for the signals $\mathbf{U}$ and the abundances $\mathbf{A}$, as well as the estimation of the posterior distribution using a hierarchical Bayes model, are detailed in [58].

Generally, with the addition of priors, the expressions for posterior distributions become complicated and require sampling strategies. Markov Chain Monte Carlo (MCMC) is the standard approach in the estimation of the posterior distribution parameters. Some popular automatic sampling methods in MCMC are RJMCMC (Reversible-Jump MCMC) [59], Gibbs sampler[60], Metropolis-Hastings [61], [62], hybrid Monte Carlo (or Hamiltonian Monte Carlo) [63], etc. If the posterior marginals are impossible to calculate, then variational Bayes approach may be used. A standard variation Bayes involves approximating the posterior distribution with another distribution easier to sample.

1.7 Other models and methods

Non-linear models

For most of the problems in blind source separation, linear assumptions hold true, and the optimisation of such provides good unmixing results. However, there could be cases where such an assumption is invalid, and a unique nonlinear mixing model may be required for each assumption. In hyperspectral imaging, for example, one of the primary reasons why non-linearity could be introduced is multiple reflections of the photons from different surfaces before reaching the camera sensor. There are various models in the literature that intend to solve nonlinear unmixing problems which have been well presented in [64]. A general optimisation problem for nonlinear mixture can be written as :

$$\min_{\Theta} \|\mathbf{Y} - \mathbf{\Theta}(\mathbf{U}, \mathbf{A})\|_F^2, \tag{1.16}$$

with $\mathbf{\Theta}(.)$, a non-linear function defining the relationships between the signals $\mathbf{U}$ and the abundances in $\mathbf{A}$. Many nonlinear approaches and methods for hyperspectral data are explained in [37]. A neural based approach to learn nonlinear mixtures has been detailed in [65].

Tensorial approach

Multi-dimensional data can be treated with multi-way analysis or in other words tensor decomposition. For example in hyperspectral, the third dimension could be angular data or time-series measurements of a particular area, for fMRI data with voxels in one dimension and the time-signals in the other: the third dimension could be the different subjects, etc. Various decompositions of a tensors are possible out of which Tucker [66] and Canonical Polyadic Decomposition (CPD) [67] are given below:

- Canonical Polyadic Tensor decomposition/Candecomp or Parafac [67]: An element $y_{i,j,k}$ of matrix $\mathbf{Y}$ can be denoted as:

$$y_{i,j,k} = \sum_{r=1}^{R} u_{ri} a_{rj} b_{rk}, \tag{1.17}$$

  where vectors $\mathbf{u}$, $\mathbf{a}$ and $\mathbf{b}$ are vectors of length R (R is the number of sources) and i,j and k are the indices of the elements in those vectors.

- Tucker decomposition : A three dimensional data $\mathbf{Y}$ measuring different properties along different dimensions can be viewed as a third order tensor $\mathcal{T}_{N,P,R} = \mathbf{G} \otimes \mathbf{U} \otimes \mathbf{A} \otimes \mathbf{B}$, where $\otimes$ is the tensor product. $\mathbf{U}$, $\mathbf{A}$ and $\mathbf{B}$ are the feature factors or feature matrices for the signals matrix, abundances matrix and some

other measured property respectively. An element $y_{i,j,k}$ of matrix $\mathbf{Y}$ can be expressed as:

$$y_{i,j,k} = \sum_{d=1}^{n} \sum_{e=1}^{p} \sum_{f=1}^{r} \mathbf{G}_{def} \mathbf{U}_{di} \mathbf{A}_{ej} \mathbf{B}_{fk}. \tag{1.18}$$

$\mathbf{G}_{d,e,f}$ is a diagonal tensor with dimensions $n \times p \times r$.

A graphical representation of Tucker decomposition to find features is given in Fig. 1.3. Tucker decomposition based non negative matrix factorisation involving *positivity* and *sparsity* constraints presented in [68] can be written as:

$$\min_{\mathbf{G},\mathbf{U},\mathbf{A},\mathbf{B} \geq 0} \frac{1}{2} \|\mathbf{Y} - \mathbf{G} \otimes \mathbf{U} \otimes \mathbf{A} \otimes \mathbf{B}\|_F^2. \tag{1.19}$$

For fMRI data, $\mathbf{U}$ and $\mathbf{A}$ can be seen as the signals and spatial maps matrices, respectively, and $\mathbf{B}$ contains the spatio-temporal information for each subject or may represent the spatio-temporal evolution with respect to new time series data.

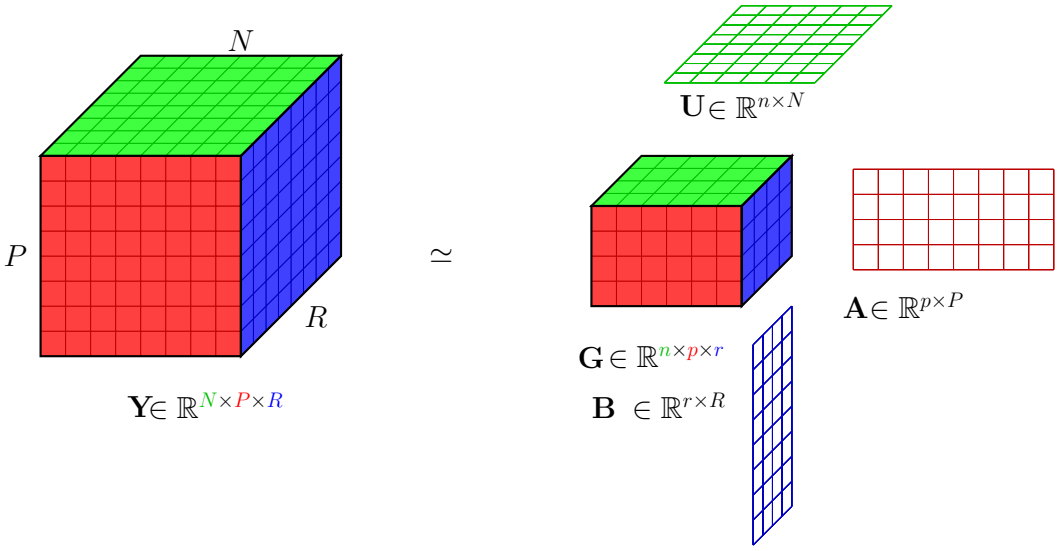


Figure 1.3: Tucker decomposition according to 1.19. The cube $\mathbf{Y}$ after reduction of dimension can be represented as $\mathbf{G}$; the dimension of N is reduced to n factors, the dimension R to r and the dimension P has been reduced to p. $\mathbf{U}$, $\mathbf{A}$ and $\mathbf{B}$ are the factor matrices.

Summary

Various methods that could be considered for source separation purposes have been explained in this chapter. Each would have certain advantages, as well as some general or data-specific disadvantages.

The classical data-driven methods of PCA, t-ICA, and s-ICA require the number of components to be estimated to be fixed. If a large number of components are specified, then many components would correspond to noise. On the contrary, a smaller number of specified components would lead to inability of these two methods to find all the components. For the main application of this work on fMRI, and also for the other applications on scintigraphy and astronomy, prior information about the localisation is available. The previous data-driven methods do not use these prior information. Similar is the case of geometric approaches, with the exception of the minimum volume approach, which by default do not exploit any a priori information. Another disadvantage of geometric methods is that they can be slower when the dimensions of the matrices are large. The utility of geometric methods is that they are blind methods and work well in the presence of similar spectra. Another interesting approach we came across is based on the Bayesian paradigm. For our problems, Bayesian approaches would require priors on the abundances as well as the nature of the timecourses or spectra. Priors of sum-to-one and positivity could still be defined on the abundances and timecourses, but the unavailability of a particular structure or form of the signals for our applications makes this approach impractical. The sampling methods required to estimate the posterior distributions are slow when dealing with large dimensions. Non-linear models could be interesting, but each application may require an individual model. It also requires an understanding of the physics behind the non-linearity assumption, which is beyond the scope of this thesis. The aim of this work is also to provide a generic adaptable algorithm for different kinds of spatio-temporal or spatio-spectral data, so non-linear models were not considered. Tensor decomposition models are novel and work in a similar way to NMFs but for the case of tensors. A necessary requirement for tensor decomposition is that it requires various instances of the data to be of the same dimension. It cannot be used for cases where data is multiway but have different dimensions, for example, data belonging to different modalities. This complex decomposition is not considered as a possible approach in this thesis.

From all the approaches we presented in this chapter, it can be noticed that the model-based approach of dictionary learning allows more straightforward additions of priors than the other methods in the literature. This becomes crucial when the compromise between spatial resolution and temporal/spectral resolution is present. This compromise is at the expense of spatial resolution, resulting in a potentially large mixing of sources in the same pixel/voxel. Source separation methods must incorporate spatial information to estimate the contribution and signature of each source in the image. We consider the particular case where the position of the sources is approximately known thanks to external information that may come from another

imaging modality or from *a priori* knowledge. This inspires us to use the dictionary learning approach for the optimisation problem treated in this thesis. It has been explained in the next chapter.

Chapter 2

Spatially constrained source separation

Contents

In this chapter, a source separation method that enables unmixing using spatial constraints is presented. The method proposed has the ability to use a high-resolution segmentation map associated with the data.

The chapter focuses on the problem of source separation in spatially structured data: 2D or 3D images that contain temporal information (fMRI, scintigraphy) or light spectrum information (hyperspectral imaging). We consider two categories of this kind of unmixing problem. In the first one, for a given pixel/voxel, different sources contribute to the mix in the sense that the spatial resolution is not fine enough to allow spatial separation of the sources. This is the case, for example, with unmixing problems in remote sensing [37] or fMRI applications [2], [39]. In this case,

the mixing matrix is a matrix of proportions where for a given pixel, the sum of the contributions of each source is equal to 1.

In the second category of unmixing problems, the mixing is additive, the signals of the different sources are superimposed and their sum forms the observed mixing signal. Decomposition of scintigraphic image sequences into tissue images and their time-activity curves or unmixing of light sources in hyperspectral data in astronomy are examples of this second category of problems. In this case, the *sum-to-one* constraint is not relevant; the coefficients of the mixing matrix are the intensity of the contribution of each source in the mixture. Since the observed signals are observed in the form of images, the constraints that can be defined in the optimisation problem should be related to the location of the sources and not to their shape. We consider the case where no information on the temporal or spectral signature of the different sources or their dependence is available.

We propose a dictionary learning method that introduces sparsity constraints on the spatial localisation of sources from external knowledge. Additional constraints on the mixing matrix (positivity and *sum-to-one* constraints) can be added or removed depending on the application. To illustrate the potential of the approach, different applications have been considered, from fMRI data (in chapter 3) to scintigraphic data to astronomy (in chapter 4). All these data are of very different natures, as well as the *a priori* information available on the location of the sources. We thus show that our algorithm is adaptable to different types of data and different types of *a priori* knowledge on the location of sources. In the case of multimodal observations, information regarding the possible spatial location of sources is usually derived from a high spatial resolution image that does not provide the second dimension, namely temporal or spectral information. Unlike multimodal image fusion problems, such as pansharpening [69], our goal is not to produce a spatially and spectrally or temporally well-resolved image. We rather aim at exploiting segmentation information from a high spatial resolution image in order to improve the unmixing of spectral or temporal sources at a lower resolution image level. In some applications, such a segmentation map is not available, but approximate spatial location information can be provided by an expert who can define regions of interest (ROIs) (see for example, the unmixing method for the highly realistic simulated renography dataset in [9], [70]).

This PhD dealing with the primary application on rs-fMRI stimulated the research towards a model with the aforementioned qualities.

2.1 Proposed unmixing model and method

The classical linear model mentioned in the previous chapter as $\mathbf{Y} \simeq \mathbf{UA}$ in 1.1, needs to be suited/tailored for different applications with the addition of certain constraints and the following next subsections explaing the proposed approach.

2.1.1 Constrained optimisation formulation

Given the observation model (1.1), the following minimisation problem:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 \tag{2.1}$$

does not have a unique solution because of the joint estimation of $\mathbf{A}$ and $\mathbf{U}$, and the ill-posedness of the problem. In order to restrain the number of solutions, we introduce some standard constraints on matrix $\mathbf{A}$ such as the *positivity constraint* $\mathbf{A} \in \mathbb{R}^+$ and the *sum-to-one constraint* $\sum \mathbf{A}[.,i] = 1$, with $i = 1:P$, as these are the proportions voxel by voxel. The form of matrix $\mathbf{A}$ may also be constrained by some extra information from high-resolution (HR) segmentations of sources or source locations knowledge: we know which ROIs may contribute to a given voxel, i.e., present a non-zero proportion at this voxel. When the total number of sources R is high, this *a priori* knowledge allows to constrain the possible solutions of the minimisation problem. In the most general setting, the unmixing problem is recast as:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_S(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}), \tag{2.2}$$

where the first term is the data fidelity term, the second term is a Tikhonov regularisation controlled by parameter μ_σ set to 10^{-4} to prevent bad conditioning (see section 2.1.2). The third term is a positivity constraint where $\mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) = \infty$ if at least one of the elements of $\mathbf{A}$ is negative, and 0 otherwise. The fourth term in eq. (2.2) codes an optional *sum-to-one* constraint on each column of matrix $\mathbf{A}$, $\mathcal{I}_S(\mathbf{A}) = \infty$ if at least one column of $\mathbf{A}$ does not sum to one, and 0 otherwise. The last term $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ is the indicator function on the set $M(\tilde{\mathbf{A}})$ of matrices having a structure similar to a given binary "structure matrix" $\tilde{\mathbf{A}}$, *i.e.* $\mathbf{A} \in M(\tilde{\mathbf{A}})$ if and only if $\mathbf{A} \in \mathbb{R}^{R\times P}$ and coefficient $\mathbf{A}_{i,j} = 0$ if $\tilde{\mathbf{A}}_{i,j} = 0$. $\tilde{\mathbf{A}}$ is a binary matrix, where element $(\tilde{\mathbf{A}})_{r,i} = 1$ if, according to a priori knowledge about spatial localisation of the sources, the r^{th} region of interest could exist in the i^{th} voxel, and 0 otherwise. This results in $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) = \infty$ if at least one element of $\mathbf{A}$ is non-zero while it is zero in $\tilde{\mathbf{A}}$, and 0 otherwise.

Combining sets $\mathbb{R}^+ \cap S = S^+$, the optimisation problem can be rewritten as follows:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \tag{2.3}$$

Estimating jointly $\mathbf{U}$ and $\mathbf{A}$ in eq. (2.3) is a typical problem of dictionary learning (DL). But, unlike conventional DL algorithms, there is no sparsity regularisation term in the form of an ℓ_1 penalty: it is the sources localisation information coded in the structural term $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ which enforces the sparse decomposition of each voxel. A classical way to solve the joint estimation problem is to optimise alternatively the cost function eq. (2.2) along $\mathbf{U}$ and $\mathbf{A}$.

2.1.2 Estimation of the temporal / spectral signatures matrix U

Considering that $\mathbf{A}$ is fixed, problem (2.2) becomes:

$$\min_{\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2. \tag{2.4}$$

The Tikhonov regularisation term $\frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2$ is introduced to improve the conditioning of problem (2.4), μ_σ is set to 10^{-4} to prevent collinearity between columns of $\mathbf{U}$.

The solution of (**??**) is the ridge estimator defined by:

$$\hat{\mathbf{U}} = \mathbf{Y}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T + \mu_\sigma \mathbf{I}_R)^{-1}, \tag{2.5}$$

where $\mathbf{I}_R$ is the $R \times R$ identity matrix.

2.1.3 Estimation of the abundance / mixing matrix A

Consider that $\mathbf{U}$ is fixed, then problem eq. (2.2) becomes $\min_{\mathbf{A}} f(\mathbf{A})$, where:

$$f(\mathbf{A}) = \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \tag{2.6}$$

Note that this function is separable according to the pixels/voxels $i \in \{1, \ldots, P\}$, which leads to: $\min_{\mathbf{a}_i} f(\mathbf{a}_i)$, with:

$$f(\mathbf{a}_i) = \frac{1}{2}\|\mathbf{y}_i - \mathbf{U}\mathbf{a}_i\|_F^2 + \mathcal{I}_{S^+}(\mathbf{a}_i) + \mathcal{I}_{M(\tilde{\mathbf{a}}_i)}(\mathbf{a}_i), \tag{2.7}$$

where $\mathbf{a}_i$ is a column vector from the matrix $\mathbf{A}$ (and with an abuse of notation for $f(.)$). The set of all the vectors with a structure similar to $\mathbf{a}_i$ is given by $\tilde{\mathbf{a}}_i$, where $\tilde{\mathbf{a}}_i$ is a column of $\tilde{\mathbf{A}}$. The regularisation terms in eq. (2.7) can be summarised as:

$$g(\mathbf{a}_i) = \mathcal{I}_{M(\tilde{\mathbf{a}}) \cap S^+}(\mathbf{a}_i). \tag{2.8}$$

Note that the objective function eq. (2.7) is convex since the first term is convex and differentiable and $g(a_i)$ is convex but non differentiable. The proof of convexity for $g(\mathbf{a}_i)$ can be easily demonstrated.

2.1.4 Proof of Convexity of $\mathcal{I}_{M(\tilde{\mathbf{a}}_i) \cap S^+}(\mathbf{a}_i)$

The convexity of $\mathcal{I}_{M(\tilde{\mathbf{a}}_i) \cap S^+}(\mathbf{a}_i)$ can be proven if we can prove that the set $M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$ is convex. This can be easily verified by following the assumption that each voxel contains the contribution of at least one region of interest. By definition:

$$\begin{aligned}
&\mathcal{I}_{\mathbb{R}^+ \cap M(\tilde{\mathbf{a}}_i) \cap S}(\mathbf{a}_i) \\
&\qquad = +\infty \text{ if } \mathbf{a}_i \text{ has a non-zero value where } \tilde{\mathbf{a}}_i \text{ is 0} \\
&\qquad = +\infty \text{ if the sum of } \mathbf{a}_i \text{ is not equal to 1} \\
&\qquad = +\infty \text{ if at least one element in } \mathbf{a}_i \text{ is negative} \\
&\qquad = 0 \text{ otherwise}
\end{aligned}$$

where $(\tilde{\mathbf{a}}_i)$ is a binary vector where element $(\tilde{\mathbf{a}}_i)_r = 1$ if the r^{th} region of the segmentation map intersects the i^{th} voxel, and 0 otherwise. A set formed by the intersection $M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$ is convex if for all $t \in [0,1]$ and for all $\mathbf{v}, \mathbf{w} \in M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$, a line segment defined :

$$\mathbf{z} = t\mathbf{v} + (1-t)\mathbf{w} \tag{2.9}$$

lies in $M(\tilde{\mathbf{a}}_i) \cap S \cap \mathbb{R}^+$. Proof that $\mathbf{z} \in \mathbb{R}^+$ and $\mathbf{z} \in S$ is straightforward.

Proof that the line segment $\mathbf{z}$ belongs to subspace $M(\tilde{\mathbf{a}}_i)$ is detailed in the following.

Let $\mathbf{u} \in \mathbb{R}^+$. By introducing the complementary vector $(\mathbf{1} - (\tilde{\mathbf{a}}_i))$ of $\tilde{\mathbf{a}}_i$, where the notation $\mathbf{1}$ is a vector of ones of the same size as $\tilde{\mathbf{a}}_i$, we have the equivalence:

$$\mathbf{u} \in M(\tilde{\mathbf{a}}_i) \iff \mathbf{u}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \tag{2.10}$$

Implication $\mathbf{u} \in M(\tilde{\mathbf{a}}_i) \implies \mathbf{u}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0$ is straightforward (vector $\mathbf{u}$ lying in $M(\tilde{\mathbf{a}}_i)$ has its coefficients $(\mathbf{u})_r = 0$ when $1 - (\tilde{\mathbf{a}}_i)_r = 1$). It is easy to prove that the reverse is true provided that $\mathbf{u} \in \mathbb{R}^+$. As a consequence, vectors $\mathbf{v}$ and $\mathbf{w}$ lying in $M(\tilde{\mathbf{a}}_i)$ verify:

$$\mathbf{v}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \quad \text{and} \quad \mathbf{w}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0 \tag{2.11}$$

which yields:

$$\begin{aligned} & t\mathbf{v}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) + (1-t)\mathbf{w}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0, \\ \iff \quad & \mathbf{z}^T(\mathbf{1} - (\tilde{\mathbf{a}}_i)) = 0, \end{aligned}$$

which, since $\mathbf{z} \in \mathbb{R}^+$ and according to (2.10), implies that $\mathbf{z} \in M(\tilde{\mathbf{a}}_i)$.

Minimisation of the objective function given by eq. (2.7) belongs to the class of problems on which the proximal gradient methods can be applied. Different algorithms are available, for example, alternating direction method of multipliers (ADMM) [71], projected gradient, also known as iterative shrinkage-thresholding algorithm (ISTA) or FISTA (Fast ISTA) [72]. Algorithm FISTA was preferred for its rapid convergence: its implementation is given in algorithm 3.

In algorithm 3, $\nabla f(\mathbf{a}_i)$ is the gradient of $f(\mathbf{a}_i)$, given by $\mathbf{U}^T(\mathbf{U}\mathbf{a}_i - \mathbf{y}_i)$. The step size λ is set equal to the inverse of the Lipschitz constant of $\nabla f(\mathbf{a}_i)$ i.e. $1/L$, where $L = \|\mathbf{U}^T\mathbf{U}\|_F$. $t^{(k+1)}$ is an auxiliary variable that helps in the fast convergence of FISTA, $\boldsymbol{\omega}$ calculates intermediate values based on a special linear combination of the last two points, and *prox* refers to the proximal operator [72]. In our case, the proximal operator is just the projection of $\mathbf{a}_i$ in the positive orthant, with the vector normalised *to-sum-to-one*. This projection also forces the elements of abundance matrix $(\mathbf{A})_{r,i}$ to be non-zero only at positions where the region of interest r projects on pixel/voxel i ($\mathcal{I}_{M(\tilde{\mathbf{a}})}$ constraint). The proximal operator of the function g is:

$$prox_g(\mathbf{y}) = \underset{\mathbf{x} \in M(\tilde{\mathbf{a}}) \cap S^+}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{y}\|_2 = \mathcal{P}_{M(\tilde{\mathbf{a}}) \cap S^+}(\mathbf{y}), \tag{2.12}$$

Initialisation of $\mathbf{A}^{(0)}, l = 0$
while *STOPPING CRITERIA* $\neq$ *TRUE* **do**
  Minimisation problem w.r.t $\mathbf{U}$:
  $\mathbf{U}^{(l+1)} = \mathbf{Y}\mathbf{A}^{(l)^T}(\mathbf{A}^{(l)}\mathbf{A}^{(l)^T} + \mu_\sigma \mathbf{I}_R)^{-1}$
  Parallel minimisation w.r.t. the columns $\mathbf{a}_i$ of $\mathbf{A}$:
  for $\mathbf{a}_i$ *of* $\mathbf{A}$ **do**
    $\boldsymbol{\omega}_i^{(1)} = \mathbf{a}_i^{(l)}, \mathbf{b}_i^{(0)} = \mathbf{a}_i^{(l)}$
    for $k \leftarrow 1$ **to** *proxsteps* **do**
      $\mathbf{b}_i^{(k)} = prox_g(\boldsymbol{\omega}_i^{(k)} - \lambda \nabla f(\boldsymbol{\omega}_i^{(k)}))$ [POCS]
      $t^{(k+1)} = \frac{1+\sqrt{1+4\left(t^{(k)}\right)^2}}{2}$
      $\boldsymbol{\omega}_i^{(k+1)} = \mathbf{b}_i^{(k)} + \left(\frac{t^{(k)}-1}{t^{(k+1)}}\right)(\mathbf{b}_i^{(k)} - \mathbf{b}_i^{(k-1)})$
    end
    $\mathbf{a}_i^{(l+1)} = \mathbf{b}_i^{(proxsteps)}$
  end
  $l = l + 1$;
end
return $\mathbf{A}$, $\mathbf{U}$

Algorithm 3: Alternate optimisation algorithm to estimate $\mathbf{A}$ and $\mathbf{U}$ that combines three nested iterative algorithms. At each iteration l, the $\mathbf{A}$ and $\mathbf{U}$ matrices are updated. The estimation of $\mathbf{A}$ is pixel-parallelised i.e., for a given pixel i, the index k refers to the iterations of the FISTA algorithm. In each iteration k, the calculation of the $prox_g$ requires an iterative POCS algorithm detailed in 2.1.3.

where $\mathcal{P}$ is the projection operator on set $M(\tilde{\mathbf{a}}) \cap S^+$. The orthogonal projection of a vector $\mathbf{y} \in \mathbb{R}^R$ on $M(\tilde{\mathbf{a}}) \cap S^+$ is obtained using the projection onto convex sets (POCS) method [73]. POCS algorithm alternates projection onto the simplex $S^+ = \mathbb{R}^+ \cap S$ and projection onto the set $M(\tilde{\mathbf{a}})$ of vectors having the same structure as $\tilde{\mathbf{a}}_i$. Only a few iterations are required for convergence of the POCS algorithm. Various POCS algorithms exist in the literature, we are interested in the projection of the abundance vector on the positive orthant of the unit simplex. The simplest way is the use of euclidean projections but there are other algorithms optimised to perform it [11], [74]–[76]. Some classical algorithms have been detailed in [12] where the author proposes a projection method faster than the standard methods for projection on the ℓ_1 ball or the simplex.

For the POCS implementation, which is projection on the set $\mathcal{I}_{M(\tilde{\mathbf{a}}_i)\cap S^+}(\mathbf{a}_i)$, the implementation of the Michelot algorithm [11] code in $C++$ was used. Even though [12] performs better in theory, this was a technical choice because of the availability of an existing implementation that performed the projections utilising all the cores

of the CPU using the multiprocessing OpenMPI library in C++.

Convergence towards a global minimum of DL algorithms cannot be proven. In practice, a good initialisation of $\mathbf{A}$ and the presence of pure pixels (as in remote sensing applications) in each region guarantee a good joint estimation of $\mathbf{U}$ and $\mathbf{A}$. Previous work [2] we have demonstrated the importance of well-defining the spatial constraint on abundance $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ to ensure an acceptable estimate of abundances and spectral or temporal signatures.

2.2 Evaluation on synthetic dataset

In this section, we evaluate the unmixing performance of our algorithm on a synthetic dataset. Dataset I was created to show unmixing of signals/spectrum taking into account different situations that could occur in real applications such as fMRI or astronomical data unmixing.

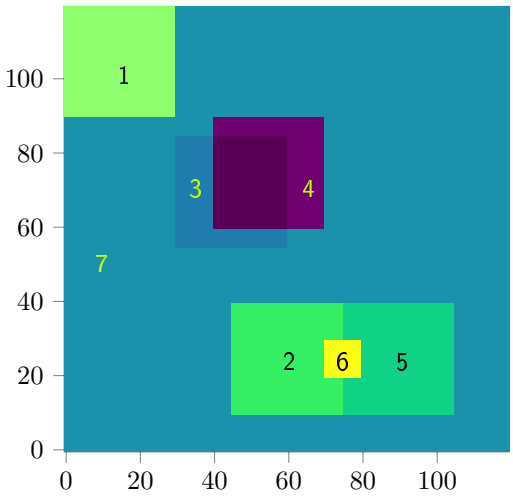
2.2.1 Data description

Unmixing algorithms are often sensitive to the assumption of pure pixels (i.e., each source or region has an abundance of 1 for at least one pixel of the image). To challenge this hypothesis, synthetic temporal data were simulated. Seven signals are mixed in a 120×120 pixels image. Ground truth signals and locations for the different regions are presented in Fig. 2.1.

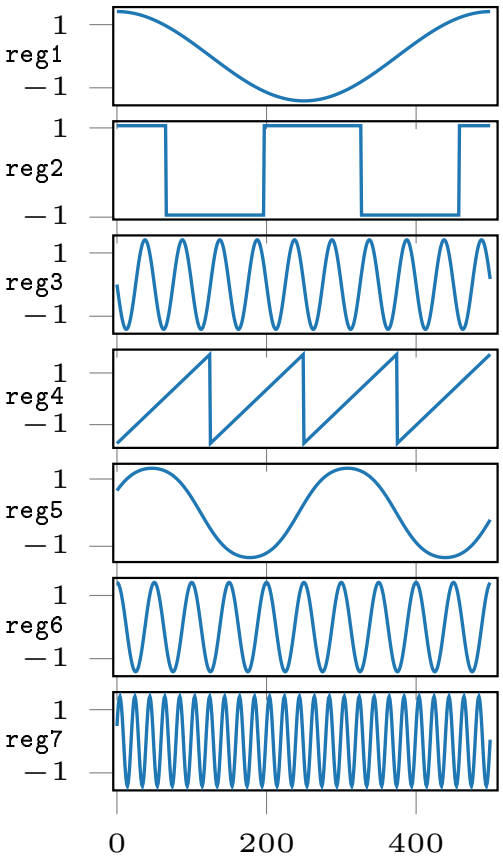
In Fig. 2.1a, we see a region 6 superimposed on two regions (2 and 5). Two other regions (3 and 4), partially covering each other, are included. Region 7 and region 1 are comprised of pixels not belonging to any other region. Data were generated for different SNRs ranging from -20dB to 20dB with a zero-mean Gaussian white noise.

2.2.2 Algorithm details

The ground truth is given by the localisation map in Fig. 2.1a. To initialise $\mathbf{A}^{(0)}$, each region was dilated with a 7 pixels square structuring element (shown in Fig. 2.2) and then the proportion for each region over each pixel was calculated, respecting the *sum-to-one* condition. The dilatation was done to introduce the uncertainty in the localisation of regions, as the localisation is seldom precise when dealing with real data. The algorithm used for unmixing is given in 3. 400 steps were adopted for FISTA, in combination with 50 steps of alternate optimisation. The weighting parameter in the Tikhonov regularisation was set to 10^{-4} as no more smoothing was required. A standard normalisation was applied to the data before processing: $\mathbf{y}_i = \frac{\mathbf{y}_i - \mu_i}{\sigma_{\mathbf{y}_i}}$, where μ_i is the mean of the temporal signal $\mathbf{y}_i$ of the i^{th} pixel and $\sigma_{\mathbf{y}_i}$ is the standard deviation of the timecourse of the i^{th} pixel.



(a) Localisation map for the 7 different sources. Region 6 is a difficult region, which overlaps region 2 and region 5. Region 7 and 1 do not intersect any other region. Regions 3 and 4 are partially overlapping.



(b) Nature of signals

Figure 2.1: Localisation map and temporal signatures used to build Dataset I.

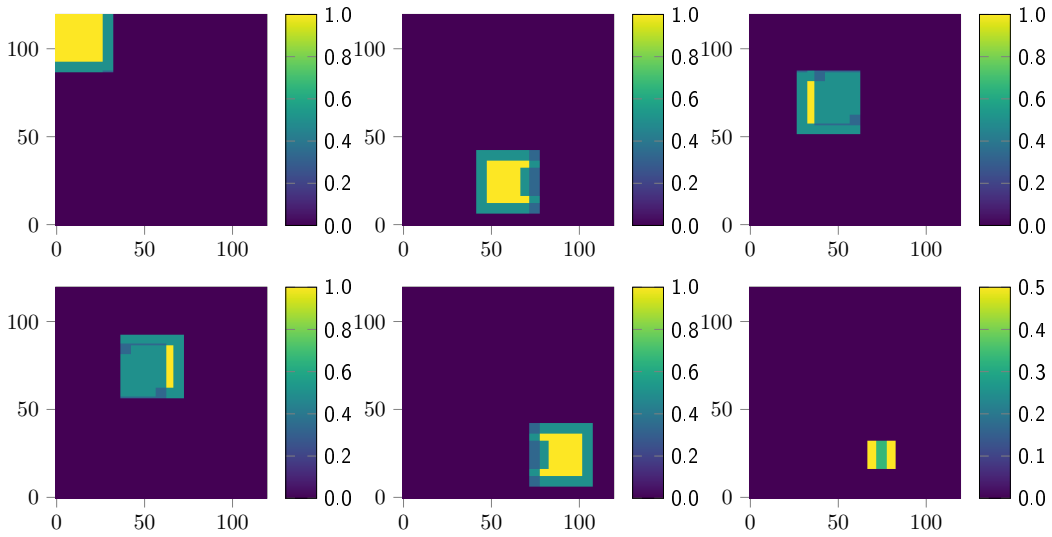


Figure 2.2: Initialisation maps of the abundances of the 6 regions.

2.2.3 Related works

To our knowledge, the optimisation problem eq. (2.3) is not solved in the state-of-the-art. The closest form to it consists in replacing the indicator on the support of matrix $\mathbf{A}$ by a sparsity constraint of type ℓ_1. Without the *sum-to-one* and *positivity* constraints, we would then have a classical problem of online dictionary learning where coefficients of the mixing matrix $\mathbf{A}$ and dictionary update are optimised alternatively until convergence to an acceptable solution. In presence of *sum-to-one* and *positivity* constraints, estimation of matrix $\mathbf{A}$ must be adapted. The constrained sparse unmixing by variable splitting and augmented Lagrangian method (C-SUnSAL) [44] is a possible candidate algorithm, widely used in the community of hyperspectral imaging, that contains the constraints of *sum-to-one* and *positivity* and an ℓ_1 constraint on the abundance map matrix. The optimisation problem then takes the following form, which is close to ours :

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2+\frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2+\mathcal{I}_{S^+}(\mathbf{A})+\|\mathbf{A}\|_1. \tag{2.13}$$

In the alternate optimisation scheme, the FISTA estimation of mixing matrix $\mathbf{A}$ is replaced by C-SUnSAL, while the estimation of $\mathbf{U}$ remains the same. For convenience purpose, let us call this algorithm DL-C-SUnSAL (Dictionary Learning-C-SUnSAL). As mentioned in the original paper [44], *sum-to-one* or *positivity* constraints can be dropped if necessary. In the following, we provide comparisons with this modified version of the optimisation problem solved by DL-C-SUnSAL (the code distributed by the authors of C-SUnSAL, with default parameters was used).

2.2.4 Results and discussion

We observe that the timecourses and the abundances for the seven regions are well estimated for different SNRs, even if the abundances are not perfectly initialised. The mean squared errors (MSE) for the estimation of the timecourses are displayed in table 2.1, The MSE in estimating the abundances are given in table 2.3. From table 2.1 and table 2.3 we see that for most of the regions the errors decrease as SNR increases. Another criteria to evaluate the timecourses is measuring the spectral angle distances (SAD). This criterion is useful when comparing the signals not having the same scale.

$$SAD(U_r^{GT}, U_r) = \cos^{-1} \frac{\sum_{n=1}^{N} U_{r,n}^{GT} U_{r,n}}{\sqrt{\sum_{n=1}^{N} U_{r,n}^{GT\,2}} \sqrt{\sum_{n=1}^{N} U_{r,n}^{2}}}, \tag{2.14}$$

where U_r^{GT} is the ground truth temporal signal for r^{th} region, U_r is the estimated signal for the r^{th} region, N is the length of the temporal signal. The SADs for the estimated timecourses given in table 2.2 follow a similar trend to MSE for an increase in SNR, proving the effectiveness of the unmixing method. The estimated timecourses for each region were normalised by standard deviation before calculating the SAD and MSE.

We also generated synthetic data where region 6 was completely included in region 5. In this case, due to noise, it was impossible to correctly estimate the timecourse of the region included in the other (and therefore its abundance). In practice, this case should not occur in our targeted applications. If such a case does occur, it could lead to poor estimation of the timecourses and thus the abundance maps.

Fig. 2.3 illustrates the convergence of the algorithm for different SNRs. Fig. 2.3 plots $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$, as a function of the number of steps in the alternate minimisation and η denotes the white Gaussian noise present in the dataset at different SNRs. In the very first steps of the alternate optimisation we see that the curves decrease sharply and ultimately settle around a particular value when convergence is achieved. In Fig. 2.3 the -20dB and -10dB curves converge to values less than 1 since some noise remains in the estimated timecourses. The 0^{th} iteration shows the ratio $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$ calculated with the initial enlarged $\mathbf{A}$ and $\mathbf{U}$, estimated with least squares.

The estimated abundance maps for each region have been shown in Fig. 2.4 to Fig. 2.8 for synthetics datasets with SNR -20dB to 20dB. For cases 10dB and 20dB, in Fig. 2.7 and Fig. 2.8 we can observe the high fidelity of the estimated maps to the ground truth. For the 0dB SNR case in Fig. 2.6 some noise remains after the estimation which aggravates in the estimated abundance maps for the -10dB SNR case Fig. 2.5 and completely dominates the region 6 in Fig. 2.4 for the -20dB SNR data.

The first 500 samples for the estimated timecourses for -20dB and -10dB SNR data can be seen in Fig. 2.9 and in Fig. 2.10 for 0dB and 10dB SNR data. The

	Method	-20dB	-10dB	0dB	10dB	20dB
reg1	Proposed	**9.5e-02**	**1.0e-02**	**1.1e-03**	**1.1e-04**	**1.1e-05**
	DL-C-SUnSAL	3.87e-01	6.91e-02	9.54e-03	1.23e-03	1.83e-04
reg2	Proposed	**1.3e-01**	**1.3e-02**	**1.2e-03**	**1.2e-04**	**1.3e-05**
	DL-C-SUnSAL	6.65e-01	1.40e-01	1.29e-02	1.85e-03	1.66e-03
reg3	Proposed	**2.2e-01**	**2.1e-02**	**2.4e-03**	**3.2e-03**	**4.0e-03**
	DL-C-SUnSAL	7.90e-01	1.04e-01	1.25e-02	5.58e-03	4.56e-03
reg4	Proposed	**2.3e-01**	**2.2e-02**	**2.5e-03**	**3.2e-03**	4.0e-03
	DL-C-SUnSAL	7.98e-01	1.06e-01	1.25e-02	4.66e-03	**3.81e-03**
reg5	Proposed	**1.3e-01**	**1.3e-02**	**1.3e-03**	**1.3e-04**	**1.3e-05**
	DL-C-SUnSAL	6.50e-01	1.29e-01	1.24e-02	1.81e-03	1.63e-03
reg6	Proposed	**1.4e+00**	**3.4e-01**	**4.2e-02**	**4.3e-03**	**6.3e-04**
	DL-C-SUnSAL	1.88e+00	1.30e+00	7.98e-02	5.75e-03	3.97e-03
reg7	Proposed	**9.6e-03**	**9.6e-04**	**9.6e-05**	**9.6e-06**	**1.0e-06**
	DL-C-SUnSAL	1.69e-01	5.03e-02	7.23e-03	9.40 e-04	1.04e-04

Table 2.1: Region wise mean squared errors for **U** for different SNRs. Best estimations for different cases have been highlighted in bold.

	Method	-20dB	-10dB	0dB	10dB	20dB
reg1	Proposed	**1.8e+01**	**5.8e+00**	**1.9e+00**	**5.9e-01**	**1.9e-01**
	DL-C-SUnSAL	3.62e+01	1.51e+01	5.60e+00	2.01e+00	7.76e-01
reg2	Proposed	**2.0e+01**	**6.5e+00**	**2.0e+00**	**6.4e-01**	**2.0e-01**
	DL-C-SUnSAL	4.81e+01	2.16e+01	6.50e+00	2.46e+00	2.34e+00
reg3	Proposed	**2.7e+01**	**8.3e+00**	**2.8e+00**	**3.2e+00**	**3.6e+00**
	DL-C-SUnSAL	5.28e+01	1.85e+01	6.41e+00	4.28e+00	3.87e+00
reg4	Proposed	**2.7e+01**	**8.6e+00**	**2.9e+00**	**3.2e+00**	3.6e+00
	DL-C-SUnSAL	5.31e+01	1.87e+01	6.40e+00	3.91e+00	**3.54e+00**
reg5	Proposed	**2.0e+01**	**6.5e+00**	**2.1e+00**	**6.5e-01**	**2.1e-01**
	DL-C-SUnSAL	4.76e+01	2.07e+01	6.38e+00	2.44e+00	2.32e+00
reg6	Proposed	**7.1e+01**	**3.4e+01**	**1.2e+01**	**3.7e+00**	**1.4e+00**
	DL-C-SUnSAL	8.66e+01	6.94e+01	1.62e+01	4.35e+00	3.61e+00
reg7	Proposed	**5.6e+00**	**1.8e+00**	**5.6e-01**	**1.8e-01**	**5.7e-02**
	DL-C-SUnSAL	2.37e+01	1.29e+01	4.87e+00	1.76e+00	5.85e-01

Table 2.2: Region wise spectral angle distances (in degrees) for U for different SNRs. Best estimations for different cases have been highlighted in bold.

estimations for timecourses for 20dB SNR data perfectly overlapped with the ground truth timecourses and have not been shown here. We can see that the quality of timecourse estimation deteriorates as the SNR decreases; much more than other regions because of the unavailability of pure pixels. The region 7 however has many pure pixels and the estimation is well even in the -20dB case.

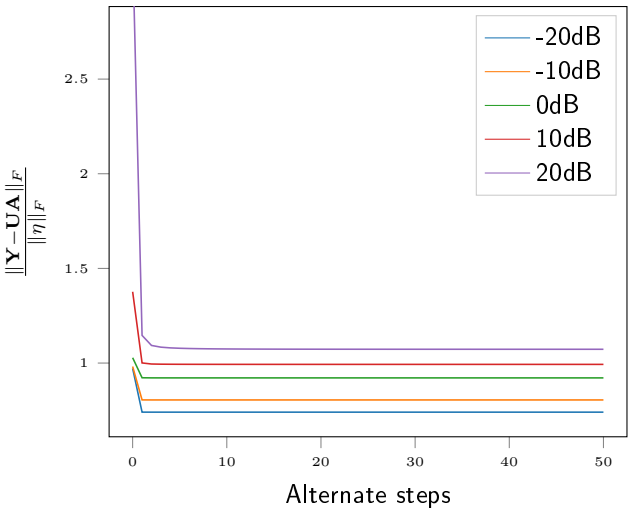


Figure 2.3: Convergence curves plotted for different values of SNR for Dataset I. The curves represent $\frac{\|\mathbf{Y}-\mathbf{UA}\|_F}{\|\eta\|_F}$ plotted against the number of iterations.

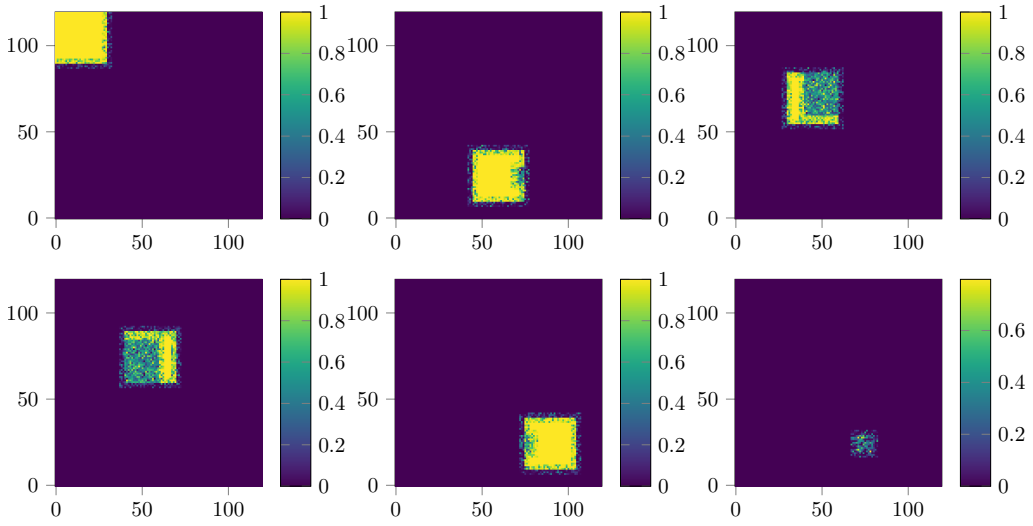


Figure 2.4: Estimated abundances at the end of 50 alternate optimisation steps for the -20dB SNR case

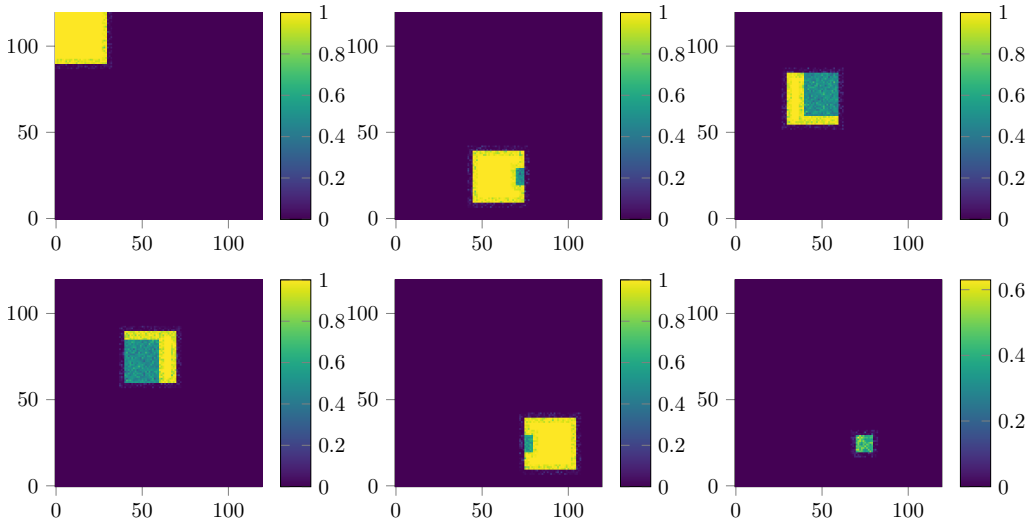


Figure 2.5: Estimated abundances at the end of 50 alternate optimisation steps for -10dB SNR case

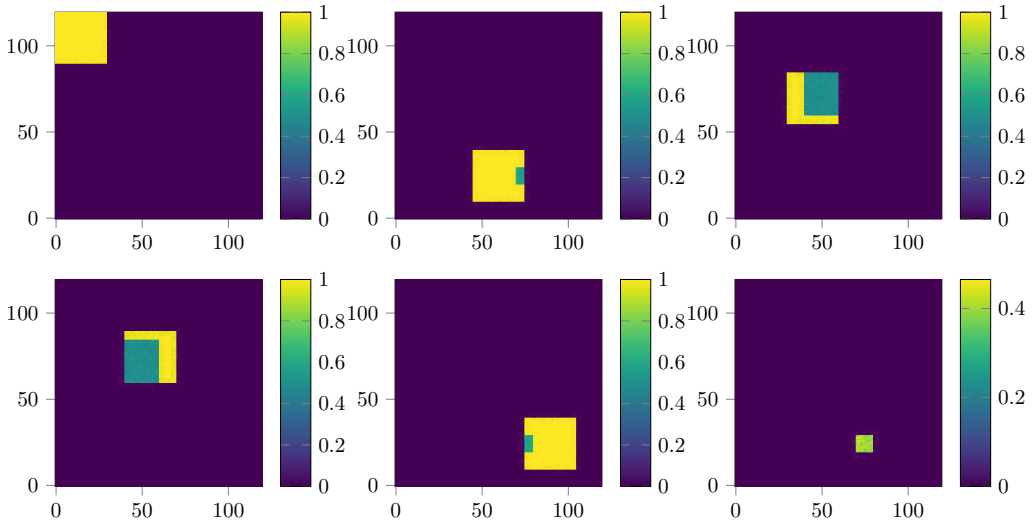


Figure 2.6: Estimated abundances at the end of 50 alternate optimisation steps for 0dB SNR case

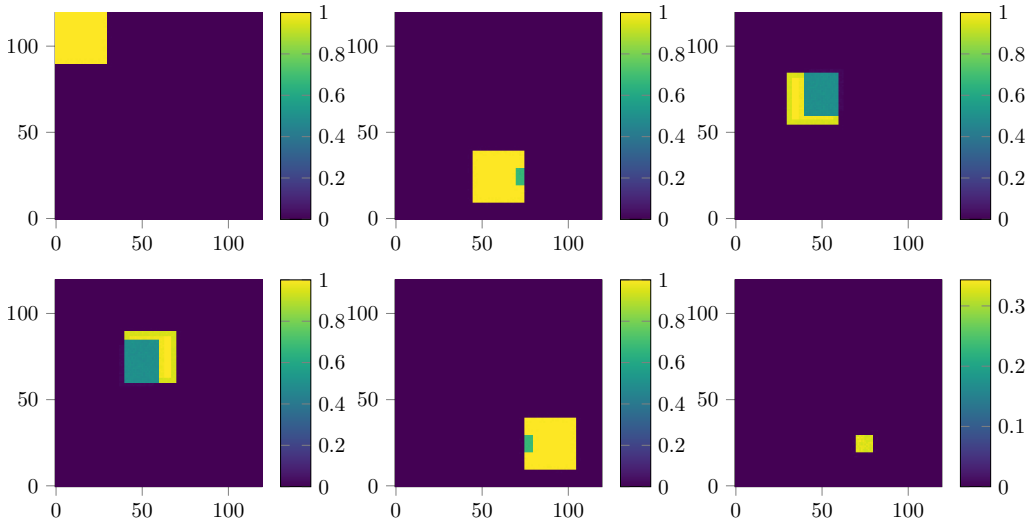


Figure 2.7: Estimated abundances at the end of 50 alternate optimisation steps have for 10dB SNR case

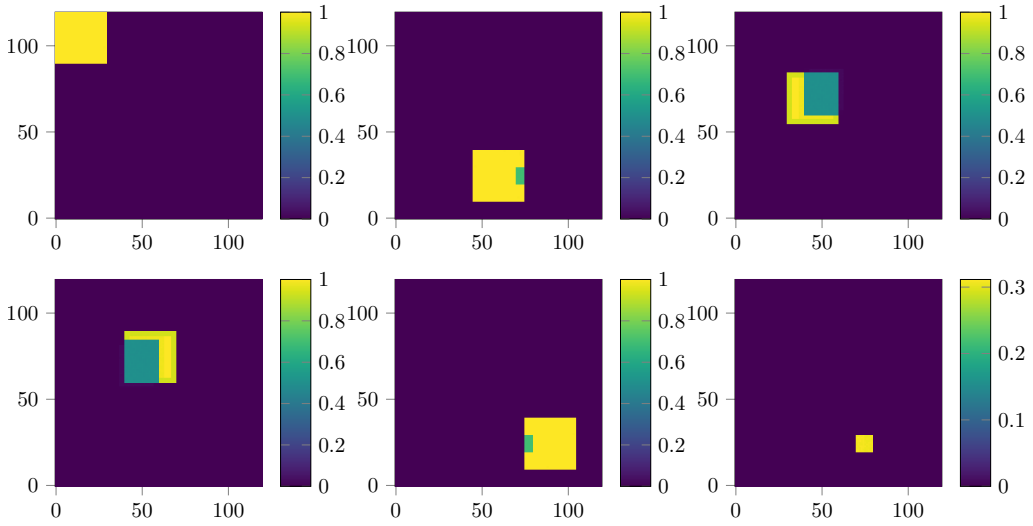


Figure 2.8: Estimated abundances at the end of 50 alternate optimisation steps for 20dB SNR case

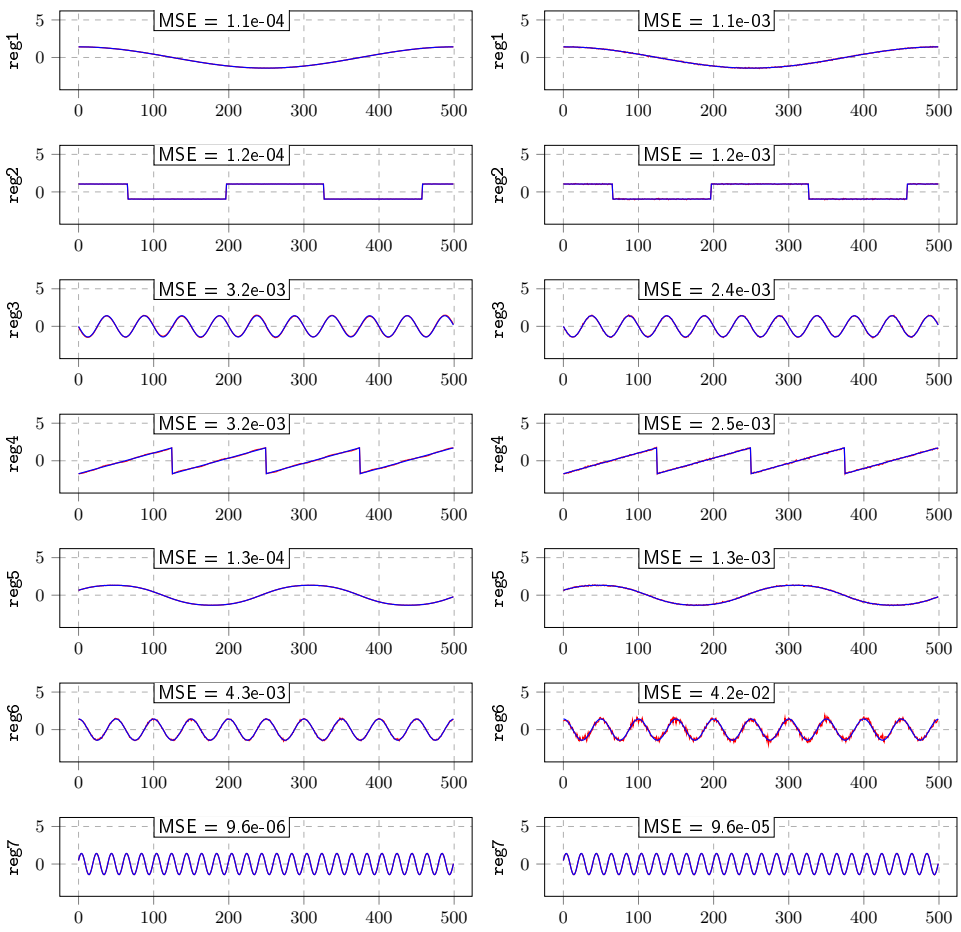


Figure 2.9: First 500 samples of ground truths(in blue) and signals estimated(in red) for 10dB SNR on the left and 0dB SNR on the right

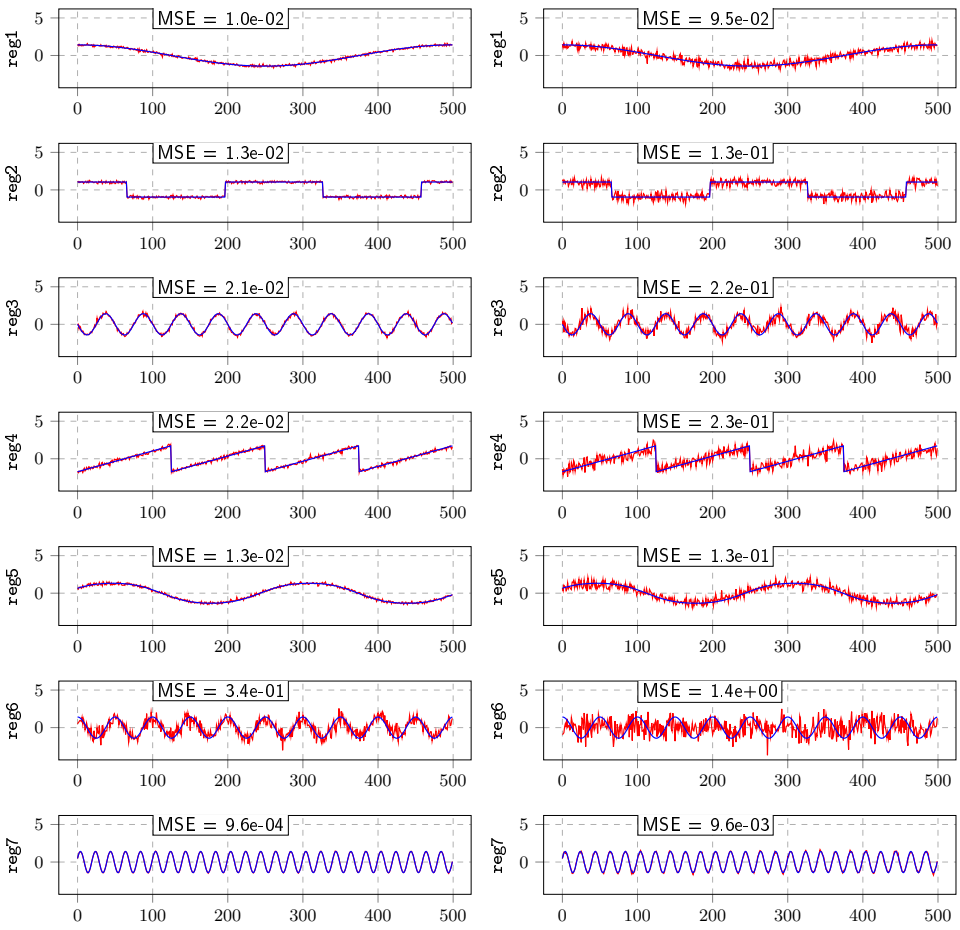


Figure 2.10: First 500 samples of ground truths (in blue) and signals estimated (in red) for -10dB SNR on the left and -20dB SNR on the right

	Method	-20dB	-10dB	0dB	10dB	20dB
reg1	Proposed	**5.8e-04**	**6.1e-05**	**4.7e-06**	**3.2e-07**	**2.8e-08**
	DL-C-SUnSAL	4.74e-02	2.05e-02	3.70e-03	5.99e-04	2.62e-04
reg2	Proposed	**2.2e-03**	**2.1e-04**	**3.8e-05**	**1.0e-04**	**1.3e-04**
	DL-C-SUnSAL	4.84e-02	2.40e-02	4.52e-03	7.79e-04	4.83e-04
reg3	Proposed	**3.2e-03**	**3.4e-04**	**4.9e-05**	**5.6e-05**	**7.0e-05**
	DL-C-SUnSAL	3.34e-02	1.57e-02	3.38e-03	1.72e-03	1.51e-03
reg4	Proposed	**3.4e-03**	**3.8e-04**	**5.4e-05**	**5.9e-05**	**7.1e-05**
	DL-C-SUnSAL	3.29e-02	1.54e-02	3.29e-03	1.61e-03	1.40e-03
reg5	Proposed	**2.1e-03**	**2.0e-04**	**4.6e-05**	**1.1e-04**	**1.3e-04**
	DL-C-SUnSAL	4.83e-02	2.39e-02	4.48e-03	7.67e-04	4.64e-04
reg6	Proposed	**9.2e-04**	**8.9e-05**	**7.0e-05**	**2.1e-04**	2.6e-04
	DL-C-SUnSAL	8.21e-03	2.91e-03	1.06e-03	2.72e-04	**1.70e-04**
reg7	Proposed	**5.4e-03**	**5.4e-04**	**6.3e-05**	**9.4e-05**	**1.2e-04**
	DL-C-SUnSAL	2.75e-01	1.06e-01	2.27e-02	2.93e-03	4.68e-04

Table 2.3: Region wise mean squared errors for A for different SNRs. Best estimations for different cases have been highlighted in bold.

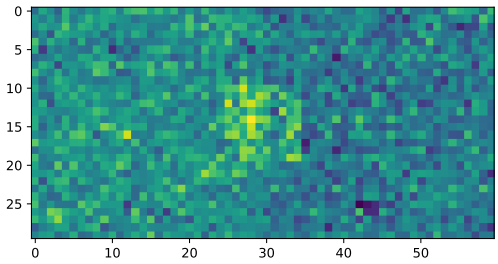


Figure 2.11: One of the 1000 frames for the 10dB case for the cropped regions

2.2.5 Geometric visualisation of convergence

Visually the convergence was evaluated on the cropped image of Fig. 2.2. The regions 2, 6 and 5 were cropped from the original synthetic data cube and one of the 1000 frames (length of the signal) has been represented in Fig. 2.11.

The algorithm proposed was run on the extracted cube with the same parameters as for the whole cube, i.e., 50 alternate optimisation steps with 400 steps of FISTA. The initialisation is the same as the abundance maps ground truth. PCA was used for the reduction of dimension of timecourses from 1000 to 2 to visualise the convergence in a 2D plane. In Fig. 2.13 we can see the representation of the initial timecourses (red), ground truth timecourses (green), and the estimations after each alternated step (in yellow) on a 2D plane. The data points, totalling 180 in number, are given in black and can be analysed as 4 cluster clouds A, B, C, and D.

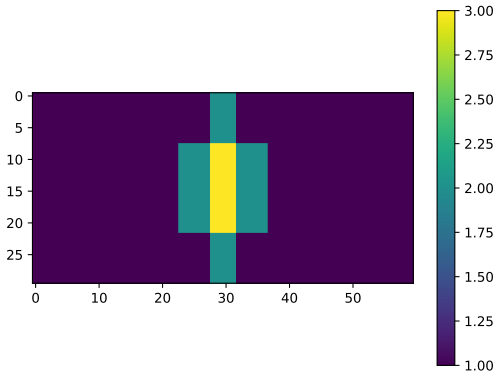


Figure 2.12: The amount of superposition for the three different regions can be understood with the help of the colour bar

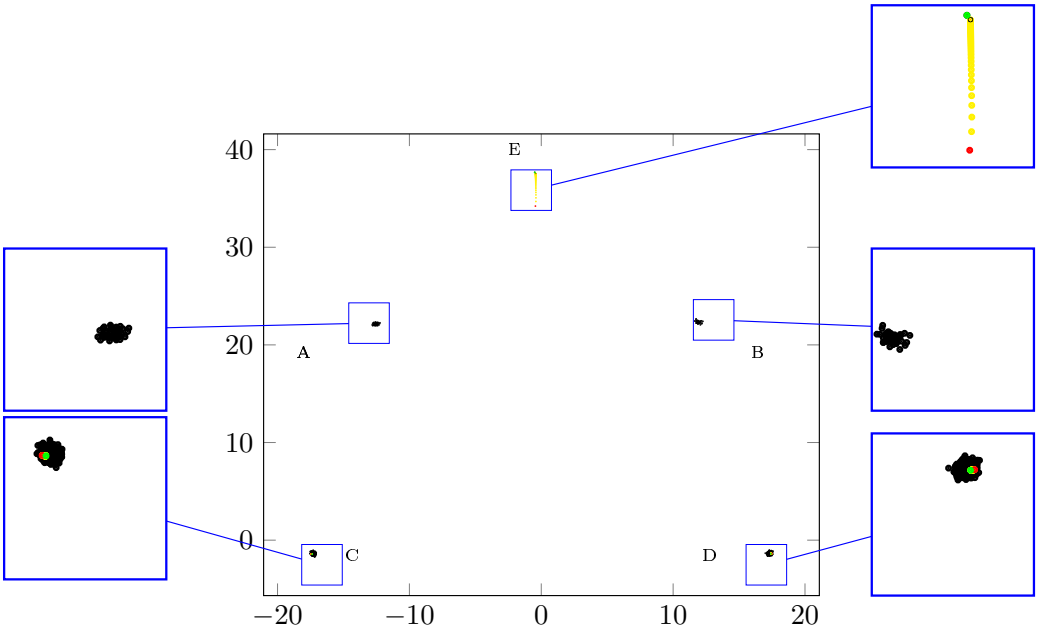


Figure 2.13: The initialisations can been seen in red, yellow data points represent the estimation after each alternated step. Green dot represents the ground truth.

In Fig. 2.13 the cloud C corresponds to the data points having timecourses similar to region 2, the region D corresponds to the data points having timecourses similar to those of region 5. The regions A and B with 100 points together correspond to the small region; we can observe that there are 50 points in the A cloud and an equal amount in B representing the data points from the two halves of the small region overlapping the larger regions. In Fig. 2.13 we can see the estimated members in yellow. The yellow dots represent the trajectories followed by the estimated members after each step of alternated optimisation. We see that they approach the ground truths represented by green dots. If the initialisation is perfect, then the least squares solution for the time courses is sufficient. This can be seen in Fig. 2.13 where the abundances of each of the region was already known.

A case for initialisation with dilated regions was done so that the check the convergence in real case scenarios where the abundance maps are not known initially. For this case the initialisation of $\mathbf{A}$ has been shown in Fig. 2.12. In Fig. 2.14 it can be observed that the initialisation for timecourses (in red) is further away from the ground truth (in green) than in the previous case Fig. 2.13; but after the alternate steps of our algorithm, it approaches the ground truth in the last iterations. The trajectories of the estimated timecourses (formed by yellow points) are comparatively longer than in the case when the abundances were precise.

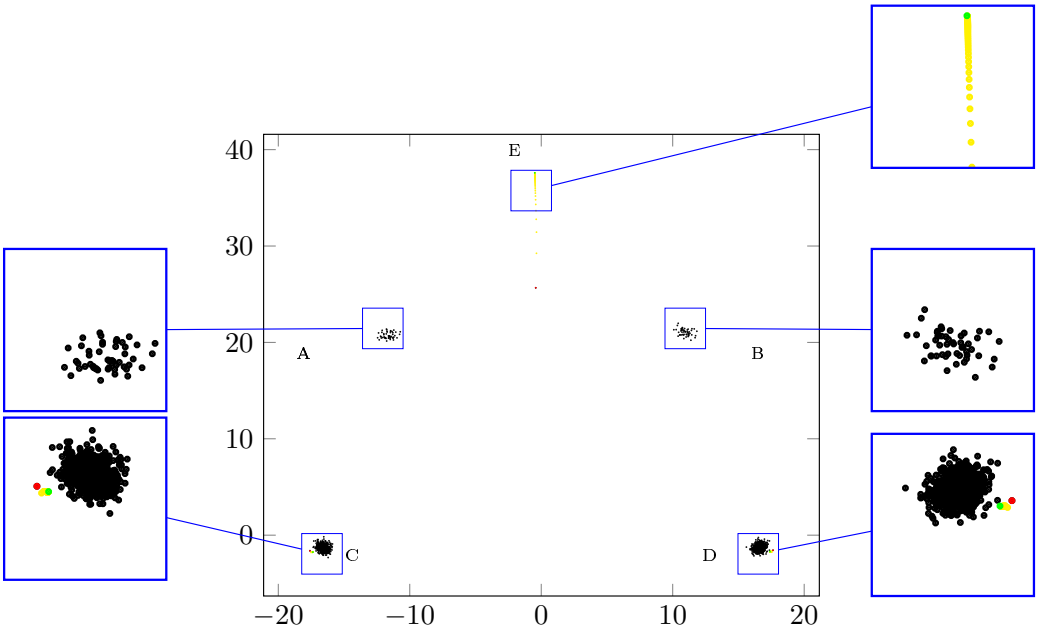


Figure 2.14: In this example the initialisations are further away from the ground truths because of less precise initialisation for the abundance map

2.3 Evaluation on a real dataset

The principal application in this Ph.D. is the unmixing of resting-state fMRI datasets in order to detect changes in cerebral activity induced by neurodegenerative diseases (explained in the next chapter). In order to evaluate the performance of the proposed algorithm on a real dataset related to the principal application of this Ph.D., the method was implemented on human block-based fMRI data (where the stimulus exists as blocks). The data treated is an example from SPM [29], where a person is made to listen bi-syllabic words binaurally. Due to the stimulus, the involved brain regions should have a temporal activity similar to the experimental paradigm. This data allowed us to compare regions showing correlated timecourses with the paradigm to ones obtained previously in the literature.

Various preprocessing steps given in chapter 30 of [29] are performed on fMRI data, with the only major difference being in the registration step. In the SPM example, this step involves registration of the fMRI data to MNI (Montreal Neurological Institute) space (a standard human brain volume in the literature). In our case the MNI volume [77] along with a MMP (multi-modal parcellation) segmentation map by [78] is registered to the fMRI data. The registration of the segmentation map to the fMRI data provides the locations of the regions on the fMRI data; this can be observed in Fig. 2.15. The anatomical regions registered to the fMRI data are used to initialise the abundance matrix. Direct application of least squares to estimate temporal signals will be less accurate because the registration of the anatomical regions to the fMRI involves overlaps of regions; thus, an unmixing algorithm is required. The proposed alternate optimisation algorithm that refines the spatial maps, and in turn, the temporal signals iteratively, is run on the data and provides the estimated temporal signals Fig. 2.17 for each of the anatomic regions and the estimated spatial maps.

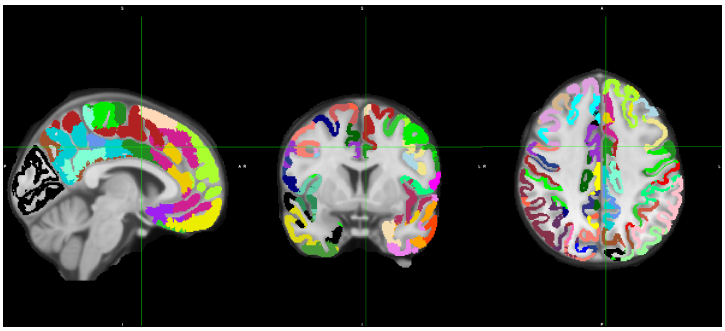


Figure 2.15: The MNI volume in grayscale with the MMP (multi-modal parcellation) segmentation map for humans superposed on the former.

We wish to verify that the signals estimated by our proposed method are linked

to the audio tasks as well as the regions highlighted in the SPM results, so we conduct further analysis. The block signal related to the experiment was convolved with the haemodynamic response function (h.r.f.) of the brain and can be observed as the paradigm signal in Fig. 2.17. The correlation coefficients of this paradigm signal to the estimated signals from different parts of the brain are calculated to find the regions showing similar activities. It should be noted that the labels for the left and the right hemispheres are different for the atlas we used. We observed that the estimated signals for three regions in the left brain and five regions in the right brain had higher correlation coefficients with the paradigm. These regions can be observed in Fig. 2.16 and the corresponding signals estimated by our algorithm are given in Fig. 2.17. The estimated signals for these eight regions correlated to the paradigm have a structure similar to the blocks in the paradigm signal. These regions correspond to the auditory cortex and can be considered to react to sounds as found in chapter 30 of SPM book [29]. In SPM analysis, each voxel of fMRI data is decomposed following the general linear model (GLM) defined as:

$$\mathbf{y} = \beta\mathbf{X} + \epsilon, \tag{2.15}$$

where β are the regression coefficients, the design matrix $\mathbf{X}$ is known and contains the regressors (stimulus convolved with h.r.f. as the first column and the offset in the second column, for example, as the regression signals) and ϵ is the model noise. In SPM [29], the values for β=[β_1, β_2] are calculated in the least squares sense. The coefficient β_1 corresponds to the paradigm, and β_2 corresponds to the regression coefficient for the offset.

In [29] the listening $>$ rest statistical t-test is performed considering all the voxels of the brain. The parameter for the statistics is the contrast coefficient; following the case of SPM, these can be written as $c^T\beta$. These contrast coefficients are supposed to have positive values for voxels affected by stimulus and this has been explained in [29]; a t-test is performed on the values of these coefficients $c^T B$ for all voxels. The estimated spatial maps after thresholding the t-test are shown in Fig. 2.18. It can be observed that the spatial maps estimated by the thresholding of the t-test superpose with the regions for which the signals are estimated by the proposed method.

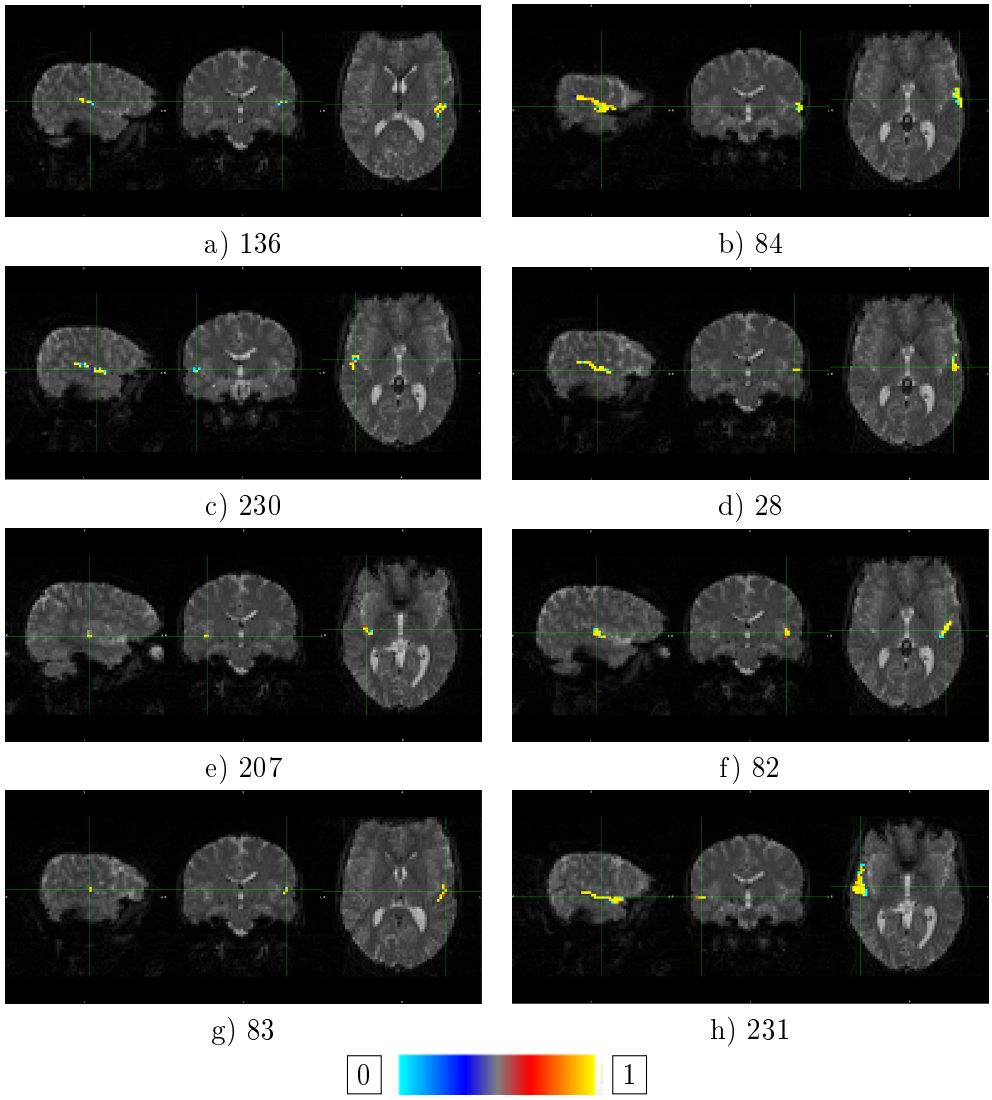


Figure 2.16: Three regions in the left brain and five in the right brain for the subject were identified to be linked with the audio stimulus. The spatial maps for the different regions have been shown in the figure. The colorbar varies from 0 to 1 for the abundance values in the spatial map.

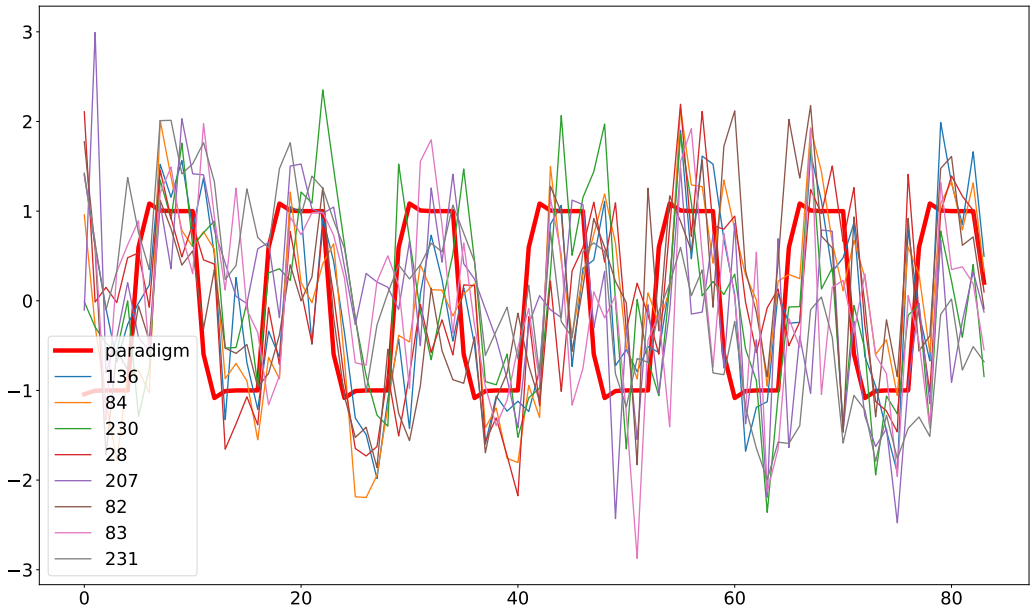


Figure 2.17: The paradigm and the estimated signals for different anatomical regions showing correlation with the paradigm.

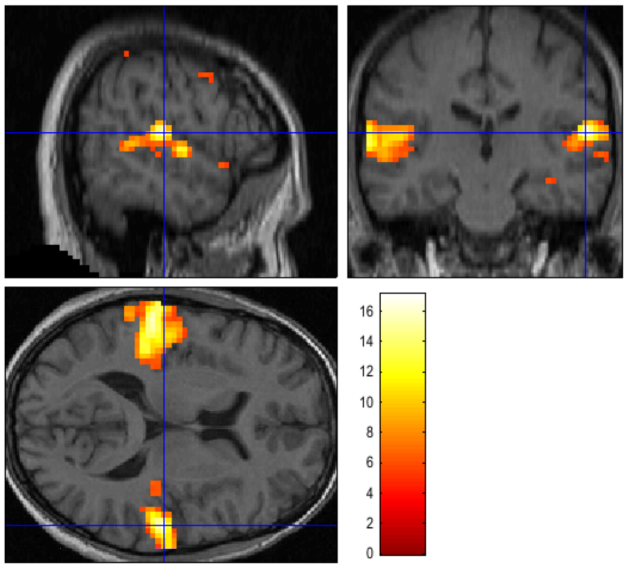


Figure 2.18: Reproduction of figure 30.19 from SPM book [29]. The figure shows the estimated maps by SPM with the statistical test listening > rest. Colorbar represents the t-test value of voxels above the threshold.

2.4 Discussion and conclusion

In this chapter, a method to unmix data consisting of an image and a temporal or spectral dimension is explained with the performances evaluated on synthetic data. From an algorithmic point of view, the strong points of the proposed approach are the small number of parameters to be set and its genericity. The algorithm has a single intrinsic parameter μ_σ used in the Tikhonov regularisation, it is to be set to a very low arbitrary value, as explained in section 2.1.2.

The convergence of the algorithm is empiric in nature and depends on a good initialisation. The algorithm is affected by the absence of pure pixels, which can be seen in the synthetic example given in section 2.2.

The algorithm is also tested on real block-based human fMRI data and shows promising results for this application. This usage has been exploited in the next chapter for resting-state mice fMRI data by following a similar but more elaborate pipeline.

Concerning the execution time of the algorithm, the main factors are the stopping criteria of the different nested iterative algorithms and the size of the images to be unmixed. Furthermore, the calculation time depends much more on the spatial dimension of the image than on the temporal/spectral dimension as the complexity lies in the estimation of $\mathbf{A}$, and to a lesser extent, on the number of regions.

It should be noted that the method does not take into consideration the morphology, the local structure or the texture of the sources, but only their approximate locations. The spatial constraints are classically expressed as an ℓ_1-penalty to promote sparsity of the mixture in each voxel. Problems with such constraints are generally solved by dictionary learning algorithms. The originality of the proposed approach lies in the replacement of this penalty by a constraint on localisation of the different regions of interests.

Chapter 3

Application to rs-fMRI

Contents

The study of cerebral anatomical and functional connectivity is one of the significant issues in neuroscience, intending to gain a better understanding of the functioning of the brain. The spread of neurodegenerative diseases like Alzheimer's disease and depression is difficult to follow over time. One way is to evaluate the

disease-induced changes in cerebral connectivity by comparing the functional activity of healthy individuals with the functional activity of a patient. The detection of connectivity changes in the individual human brain is made difficult by the lack of homogeneity of the population, with high inter-subject variability. A standard way to cope with this issue is to consider controlled animal models with induced disorders, for instance, here, mice belonging to the same genetic strain. Neurogenerative disorders are supposed to change the connectivity between the brain regions at rest in humans. These changes are observed in mice also. Although, the structural form of the mouse brain is significantly different from the human brain, the mechanism by which Alzheimer's disease affects a mouse brain is similar to that in the human brain. Homogeneity of data (same mouse model) and the short life cycle of mice is thus advantageous to learn about the spread of disease and changes in functional connectivity in mice brains. In our work, we have more specifically considered the single-subject analysis of resting-state fMRI (rs-fMRI) data in mouse models of Alzheimer's disease and depression. The advantage of single-subject evaluation is that it provides insights on the individual changes as the subject serves as its control.

3.1 Introduction to fMRI

To understand brain activity, non-invasive in vivo techniques are necessary. One such technique is functional Magnetic Resonance Imaging (fMRI). FMRI allows the observation of changes in cerebral activity by analysing the blood-oxygen-level-dependent (BOLD) signal [79]. BOLD signal measures the local changes in the quantity of oxygen carried by the haemoglobin. These changes are due to neural activity. As the neuronal activity requires consummation of a much larger oxygen share from the blood, to provide enough oxygen for the neuronal firing, the oxygenated blood flow locally increases at the particular area, which leads to local changes in the magnetic field. The magnetic field changes because of the difference in magnetic susceptibility between the oxygenated and the deoxygenated blood. Thus looking at BOLD signals is an indirect measure of the brain activity as neural processes lead to changes in the local magnetic susceptibility of the blood, which consequently is reflected in the images of the brain.

FMRI data are 4 dimensional images comprising the brain volume, with each voxel having a BOLD signal (timecourse) related to it. Two types of fMRI acquisitions exist: task-based fMRI and rs-fMRI. In task-based fMRI, the subject is asked to perform an activity, e.g., looking at different pictures, finger tapping, etc., whereas in a resting-state fMRI, the subject has to stay still. Task-based fMRI is difficult on animals and is not possible on mice, as they are difficult to train; and in rs-fMRI, the mice are lightly sedated. In recent years, rs-fMRI has become the prevailing method to study functional brain connectivity at rest [80]. At rest, only spontaneous activity of the brain is measured, and a set of anatomical regions with the same fluctuations are considered part of a common resting-state network.

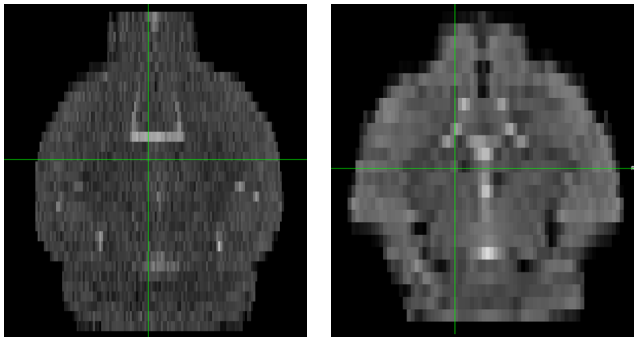
A certain number of resting-state networks have been examined in the mouse brain [81] and the human brain [80], [82]. Co-activation patterns are studied to determine the differences between healthy and pathological subjects using metrics such as correlation maps. Detecting precisely the different networks (localisation in the brain and temporal activity) is crucial for understanding a neurological disorder. The spatial resolution of the rs-fMRI is not high; thus, there is a strong probability that the functional activity of each brain region is not independently observed. The low-resolution will lead to a strong mix of functional activity on the neighbouring voxels between different functional regions. As there will be a mixture of functional activities from different regions on the same voxel, we need an unmixing model. There exist unmixing models for group studies in humans and mice, but they are not well adapted for a single-subject studies. The unmixing method proposed in chapter 2 can be used for single-subject as will be seen in the following.

3.2 Data description and material

For data acquisition of rs-fMRI and structural MRI of mice brains, a 7 Tesla scanner dedicated for small animals was installed at the IRIS (Imagerie, Robotique et Innovation en Santé) [83] platform of ICube. On this platform, a large number of data have been acquired during the past years. The data consist of 3D+t rs-fMRI and 3D anatomical imaged registered (acquired just after the fMRI scans) to the rs-fMRI imaged. The mice were scanned at five months, nine months, and thirteen months. The anatomical or the structural images have a dimension of $256 \times 256 \times 34$ and $0.08299 \times 0.07812 \times 0.4$ mm resolution. Functional images have a spatio-temporal dimension of $147 \times 87 \times 27 \times 500$ with $0.1445 \times 0.2299 \times 0.5$ mm spatial resolution and 2s for the temporal resolution. A slice from each is shown in Fig. 3.1. The structural MRI, having a better spatial resolution than fMRI, was used for the purpose of registration, explained in 3.4.1. The data were recorded and provided by Laetitia Diegorgis from IMIS (Imagerie Multimodale Intégrative en Santé) team, ICube for the study of neurodegenerative disorders (Alzheimer's) on a Thy-Tau22 transgenic mouse model of tauopathy. A setup with a head mount for the mice was installed to prevent head motion. Not avoiding head movements can lead to misinterpreted results or false activations at different sites in the brain [84].

3.3 Assumptions and proposed strategy for detection of cerebral networks

Independent component analysis (ICA) and dictionary learning (DL) methods are widely used to analyse resting-state functional Magnetic Resonance Imaging (rs-fMRI) in multi-subject studies. These methods aim at decomposing the multi-subject data into common spatial abundance maps and their related temporal sig-



(a) Structural (b) rs-fMRI

Figure 3.1: One of the axial plane images of the mice brain from structural and rs-fMRI image.

natures. In group studies, the interest is to find common resting-state functional networks; our goal here is different: detect changes in the same individual at different time points.

Detecting common networks through group analysis is clearly easier than single-subject analysis because of the redundancy of information present in group data. The redundant information present in the group studies enables to avoid the problem of low SNR. In our case, where few acquisitions or time data points are available, additionally, we have the problem of low Signal to Noise ratio (SNR). The poor resolution of rs-fMRI data is an additional source of difficulty, yielding noisy and blurry spatial maps. In the single-subject case, the rs-fMRI data of a unique subject must be decomposed according to model (1.1). The methods adopted for group studies often fail in this case because the problem is too ill-posed, requiring the use of additional prior information and the design of novel regularising constraints. The additional prior information is obtained by using the approximate locations of the regions using a highly resolved atlas with a detailed anatomical map in our case for fMRI data. The model using this information is presented in eq. (2.3). The Allen brain atlas volume is given in Fig. 3.2 and the regions of interests or the labels are given in Fig. 3.3.

The approach to detect the cerebral networks was thus different from the approach in the literature, which is mainly dedicated to detecting networks in groups rather than individuals. The classical estimation methods for cerebral networks based on ICA do not perform well for the single-subject cases because of the low SNR; therefore, we approach the problem differently. The proposed strategy is to unmix data not at the cerebral network scale but the finest resolution anatomical scale at our disposition. For mice, with the help of the exceptionally detailed segmentation of the brain, the hypothesis can be made that the functional networks

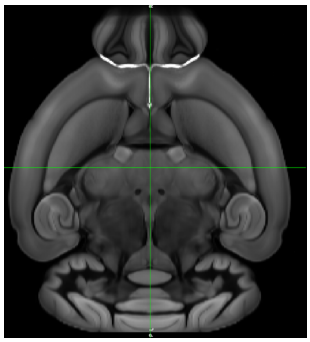


Figure 3.2: Atlas volume with very well defined anatomical regions

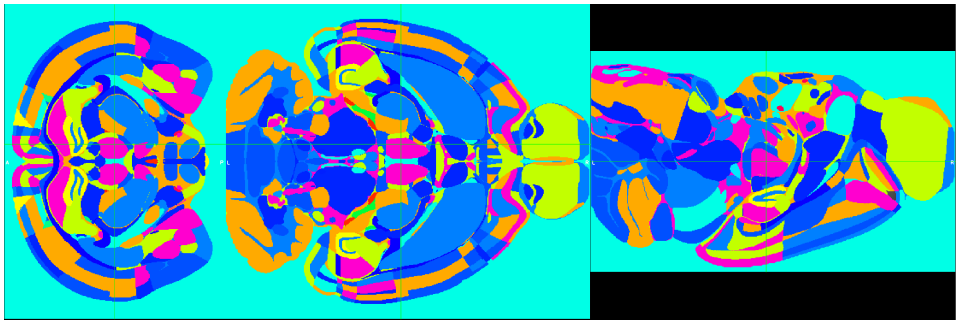


Figure 3.3: 3D representation of the segmentation map associated with the Allen Mouse Brain Atlas [8]. Each colour represents a label associated with an anatomical region.

at rest are in fact, a set of small anatomical regions. The functional networks will be formed *a posteriori* by studying the correlation coefficients between the different timecourses estimated for these anatomical regions and then grouping the anatomical regions presenting similarity in correlation coefficients.

3.4 Preprocessing pipeline for the mice fMRI data: tools and contributions

In this section, preprocessing carried out on the mice data is detailed. It is explained in two parts: spatial preprocessing and temporal preprocessing. Spatial preprocessing deals with the physical structures of the brain; this part contains details on masking, registration, and realigning of the mice brains. Temporal preprocessing involves cleaning of the timecourses belonging to the brain before the unmixing method proposed in the previous chapter 2 can be applied to the data. The structure of the preprocessing pipelines is globally the same and governed by the state-of-the-art; the

improvements made in the pipeline to adapt it to the data at disposition are detailed in this section.

3.4.1 Spatial preprocessing

Realign

This is the only part in spatial preprocessing which could modify the fMRI signals. Physiological noise due to subtle movements exists in our data. This can be caused by inadequate sedation, respiration, etc. Major movements are absent as the mice heads were fixed in a headgear during the acquisition. The images from the same examination were realigned using the classical SPM (Statistical Parameter Mapping) software [29]. The shifts induced by the realignment step are significantly less than the resolution of one voxel as the mice are anaesthetised.

Masking

In Fig. 3.1 the data is presented in its raw form, i.e., it contains other unnecessary parts acquired at the time of acquisition such as the skull, ears, eyes, etc., which should be stripped. Thus the anatomic or the structural images and the fMRI images acquired need to be masked. The masking of these two images is done by following two independent and different procedures. Unlike for human fMRI brains, there is no tool like the Brain Extraction Tool (BET) [85], [86] that exists for mice fMRI brains; thus, we decided to use the information from the Allen brain atlas to mask the structural image by registering the atlas to the structural image. The advantage of using the atlas image is that it contains just the brain without the parts surrounding it. It should be noted that the classical registration operations such as FLIRT and ANTs fail when the target image has a significantly lower resolution than the image to be registered. Even though the registration of the atlas to the fMRI data does not work well because of the significant difference in the resolution of the two images, it works well while registering the atlas to anatomic images. This is because of a smaller difference in the spatial resolution between the atlas and the anatomic image. The masking procedure can be followed visually in Fig. 3.4. In Fig. 3.4, in row (a), we can see the anatomical images for three different subjects. We can clearly see that in addition to the brain, there are some structures, the skull, and other unwanted regions present in the images. The masking of anatomical images (and fMRI) is impossible directly because of the absence of BET. To this end, to mask the anatomical images, we propose to register the Allen brain atlas to the structural brain image using the FLIRT toolbox [87]. This is advantageous because the Allen brain volume is already masked; thus, it allows information about the form of the brain and aids in refining the contours of the mask to be calculated. The registration works well in this case. One more advantage of using FLIRT is that it automatically realigns the brain images before performing the affine registration.

Row (b) of Fig. 3.4 shows the FLIRT registration from the atlas to the structural images. An affine registration is not enough for registration as can be seen in the lower regions for the three subjects. To refine this, we use a diffeomorphic registration using the ANTs toolbox [88] and the results are shown in Fig. 3.4 (c). We can observe that even the regions in the lower part are well registered to the anatomical image, which is not possible using just an affine registration. We use this registered image to mask the anatomical image. As the image contains just the atlas on the uniform black background, masking using the nilearn.masking python package is easy. The result is shown in row Fig. 3.4 (d). The masks are eventually refined by the binary closing morphological operation with a structural element of $3 \times 3 \times 3$ voxels.

Let us notice that the automatic registration of the atlas to structural images failed in cases where there was some high-intensity signal in close proximity to the skull due to gel. The gel used is a solution of 2 percent agar gel in NaCl solution. It is used to fill the gap between the top of the mouse head and the probe. This avoids the artefacts linked to the magnetic susceptibility difference between the air-tissue interface; this can introduce distortions of the signals due to the non-homogeneity of the local fields in the rs-fMRI.

In the case of calculating masks for fMRI images, the atlas was not used for masking because of a significant difference in spatial resolution of the fMRI and the atlas image; and as explained earlier, FLIRT followed ANTs registration is imprecise when registering the atlas to the fMRI. Usually, in the literature, it is the mean rs-fMRI image that is used to mask the images from the same subject. The problem of masking is similar to the case of masking the anatomical image; here, calculating the mean of the realigned fMRI volumes helps in increasing the SNR for the masking. The rs-fMRI image, in our case, was masked with the help of the masking function of nilearn.masking package in python [89] on the energy image of the fMRI instead of the mean image. The thresholding of the energy images heuristically provided better control over the calculation of the masks, and the same parameters were fixed for the whole data without any manual intervention. After calculating the masks using the energy images, holes were filled using morphological operators. The energy of the fMRI signal for each voxel can be written as :

$$e^{(i)}_{fMRI} = \sum_{n=1}^{500} |x^{(i)}_{fMRI}(n)|^2, \tag{3.1}$$

where x_{fMRI} is the raw fMRI image, n is the n^{th} timecourse sample and i is the i^{th} voxel of the brain.

In Fig. 3.5 we can visualise step by step how the brain was extracted, and a mask was made. In the first row of Fig. 3.5 the original fMRI data can be seen. In row (b) of Fig. 3.5 the realigned fMRI is observable. To find the mask, the energy of the fMRI signals was used, resulting in images given in Fig. 3.5 (c). The energy image is shown in row (d). The mask calculated on the Energy fMRI has been shown in Fig. 3.5 (e), and Fig. 3.5 row(f) shows the masked energy fMRI.

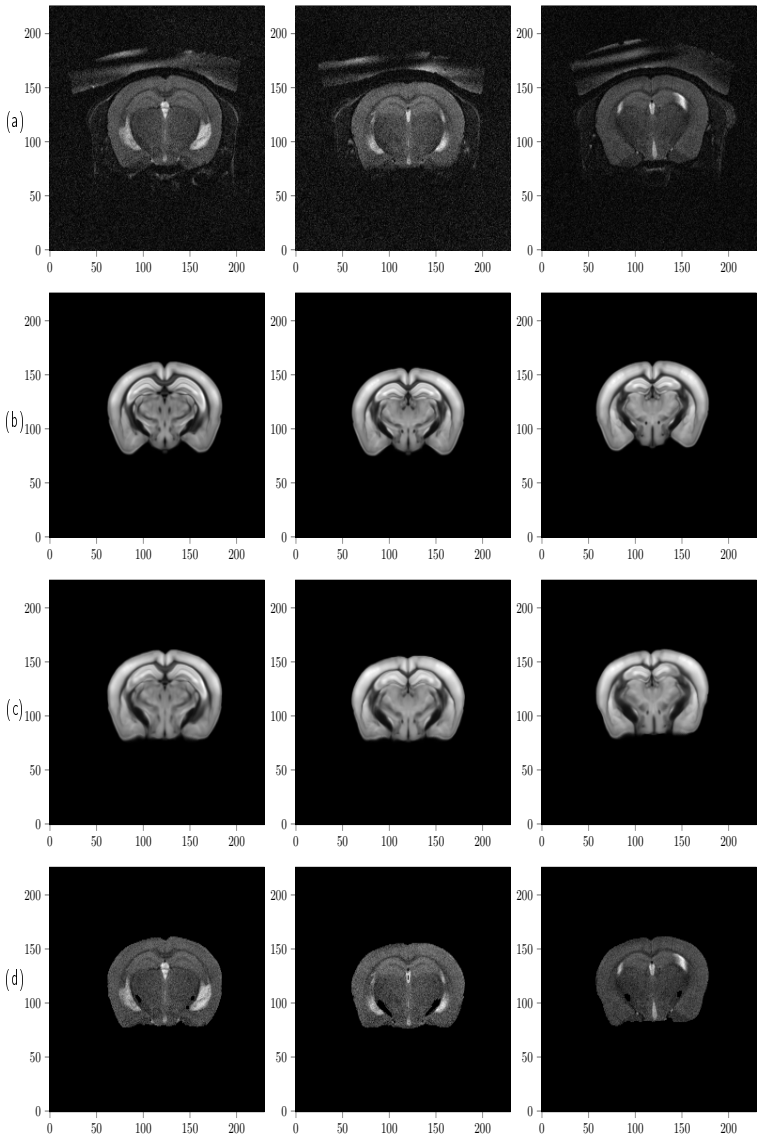


Figure 3.4: Preprocessing: masking of the anatomical image. (a) raw structural images for three different mice (b) atlas image registered to the structural image using FLIRT (c) results after ANTs diffeomorphic registration (d) masked structural image using the images provided in (c)

The masking method for fMRI was done on some human brains also and showed similar results to BET. Other notable methods for mouse brain masking are 3-D pulse-coupled neural networks (PCNN) [90] and Rapid Automatic Tissue Segmentation (RATS) [91], [92].

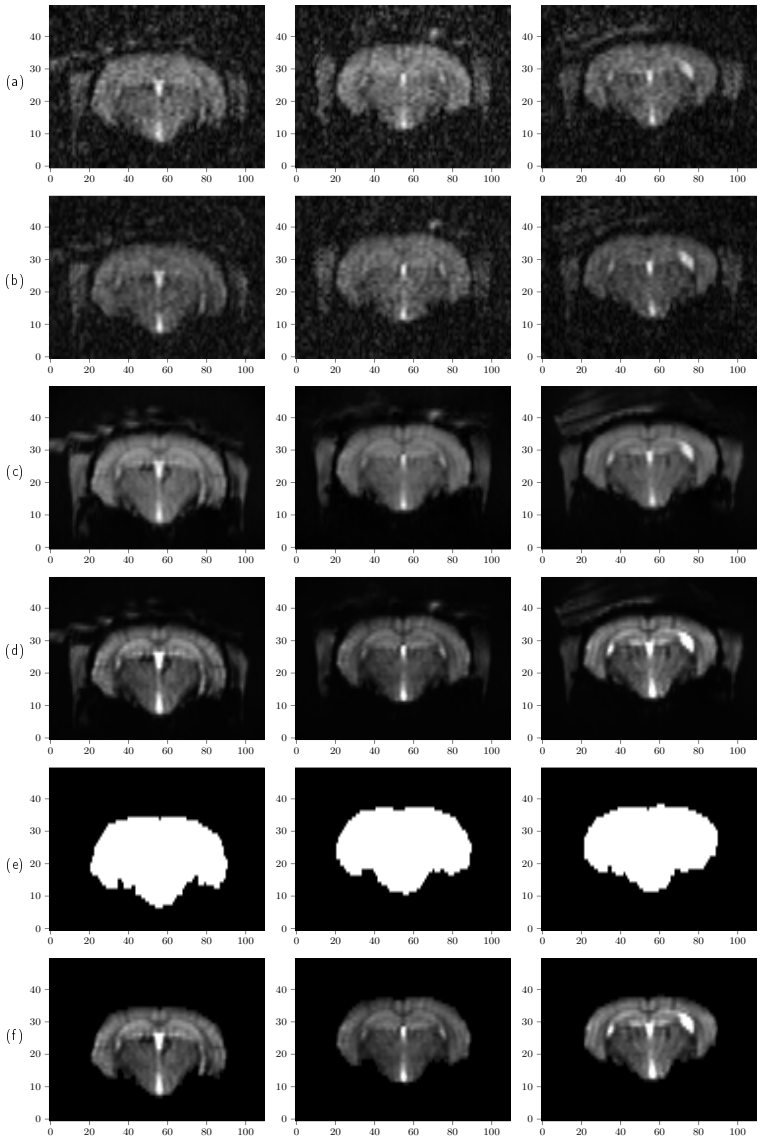


Figure 3.5: Masking of the fMRI data. In each row, we can see a coronal slice for three different mice for (a) raw rs-fMRI, (b) realigned rs-fMRI (c) mean rs-fMRI (d) energy image (e) fMRI mask (f) masked energy-fMRI that will be used for registration of the atlas to the rs-fMRI

Registration

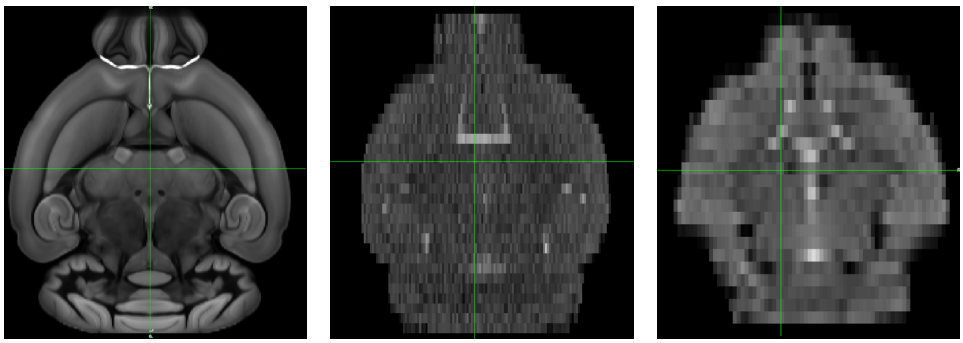
The classical registration operations such as FLIRT and ANTs fail when the target image has a significantly lower resolution than the image to be registered. The most

straightforward strategy is to increase the spatial resolution of the target image (in this case, rs-fMRI) artificially. The spatial resolution of rs-fMRI data is augmented by subdividing each original voxel into $3 \times 6 \times 2$ high-resolution voxels to reach the same order of resolution as the Allen brain atlas. In the same way, the spatial resolution of the energy rs-fMRI was increased. The registration of the anatomical image to the augmented energy rs-fMRI leads to an increase in its own resolution. In Fig. 3.7 the steps for registering the structural image to the augmented fMRI image are shown.

In Fig. 3.7 (a), we can see the masked structural images from three different mice. In Fig. 3.7 (b) the result of affine registration using FLIRT is shown. In Fig. 3.7 (c), the ANTs registration of the FLIRT image to the augmented fMRI can be visualised. In Fig. 3.7 (d) the augmented fMRI energy image (obtained by dividing the voxels into multiple parts without interpolation) is shown.

In Fig. 3.8, row (c) shows the structural image already registered to the fMRI image. The Allen brain atlas is registered to this image. In Fig. 3.8 (a), we can see the atlas after an affine registration to the structural image, which was already registered to the fMRI image. Fig. 3.8 (b) is the diffeomorphic registration of the affine-transformed atlas to the augmented structural image in the Fig. 3.8(c). In Fig. 3.8(d) the augmented fMRI energy image can be observed.

A slice from one of the axial planes of the atlas volume, the structural image, and the augmented-fMRI is shown in Fig. 3.6. The advantage of using an augmented fMRI image for registration is shown in section 3.5.



(a) atlas Volume (b) Structural (c) Augmented rs-fMRI

Figure 3.6: Axial images of mice brains from three different modalities that were used in the registration process. Atlas volume is having the highest resolution, followed by the structural and then the augmented fMRI image. The augmented fMRI image was formed by dividing the fMRI image voxels into multiple high-resolution voxels without using any interpolation

Due to the unavailability of ground truth for real data, it is hard to measure

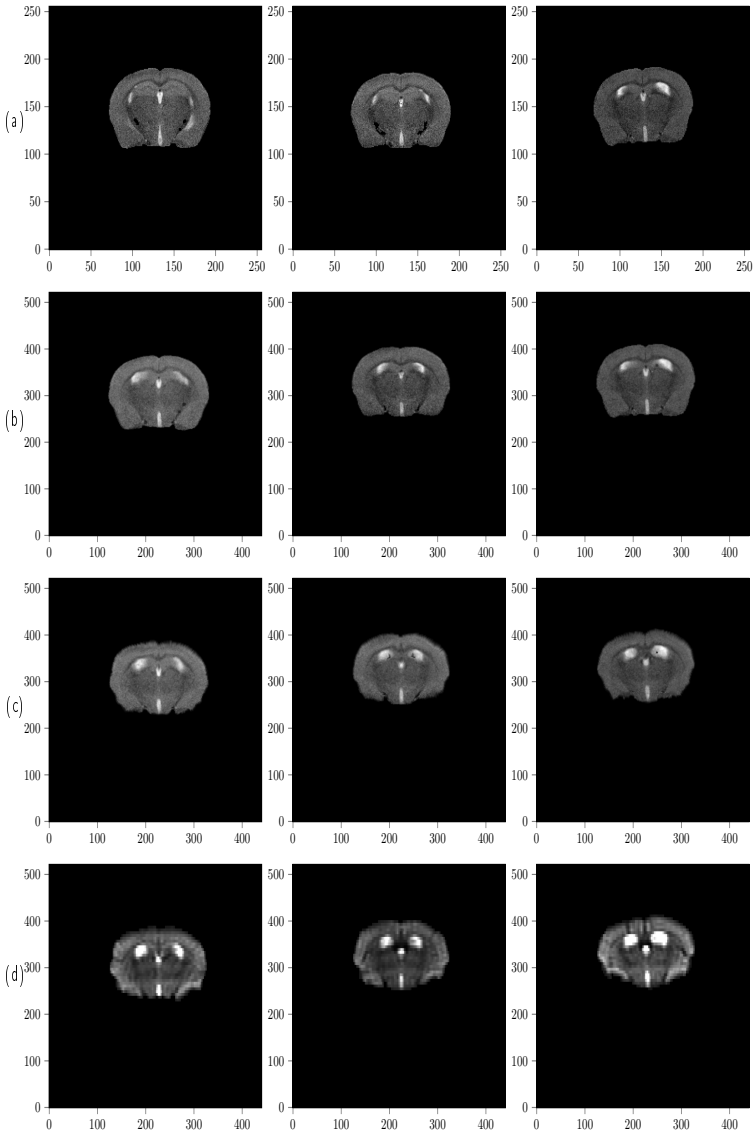


Figure 3.7: (a) masked structural images for three mice (b) results of FLIRT registration from structural to energy fMRI images (c) ANTs results (d) masked energy fMRI images

the effects of each pre-preprocessing step. To illustrate the influence of the different spatial preprocessing steps on the correlation matrices, we propose to compare the correlation matrices formed by the approximate estimation of the signals from the different regions before and after each step. Since the regions are projected on the rs-fMRI data, the mean timecourse for each region is calculated by calcu-

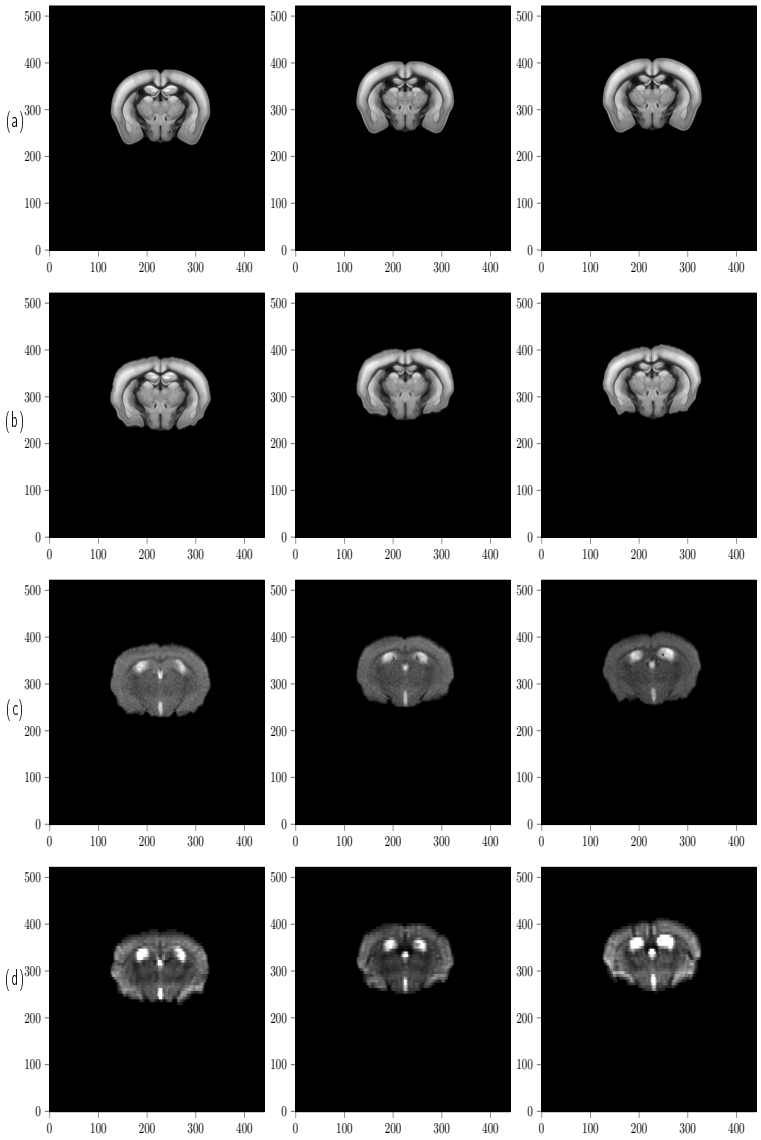


Figure 3.8: (a) Flirt registration of 112.the Allen brain atlas to the masked structural images (b) results of ANTs registration to the structural images (c) structural images to which the atlas is registered (d) energy fMRI images

lating the weighted means (proportions of the regions present on a voxel in the augmented fMRI) of the timecourses belonging to the voxels of a particular region. These timecourses can be used to evaluate the connectivity between the brain regions by calculating the Pearson correlation coefficients between different regions. The correlation matrix obtained by calculating the Pearson correlation for the mean

timecourses for the regions projected on realigned data can be seen in Fig. 3.9. With the atlas registered to the original data without realigning the brains, the correlation matrix is given in Fig. 3.10. Realignment was done to compensate for minor physiological movements which occur during the acquisition. It can be observed that the correlation matrix without the realignment step (Fig. 3.10) was not significantly different from the correlation matrix formed after realigning (Fig. 3.9) due the absence of significant movements.

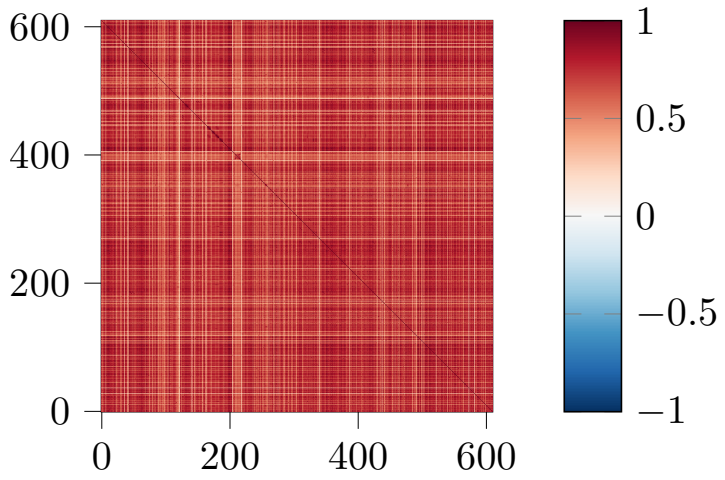


Figure 3.9: Correlation matrix (with 613 regions from the atlas projected on the fMRI images) after the realign step

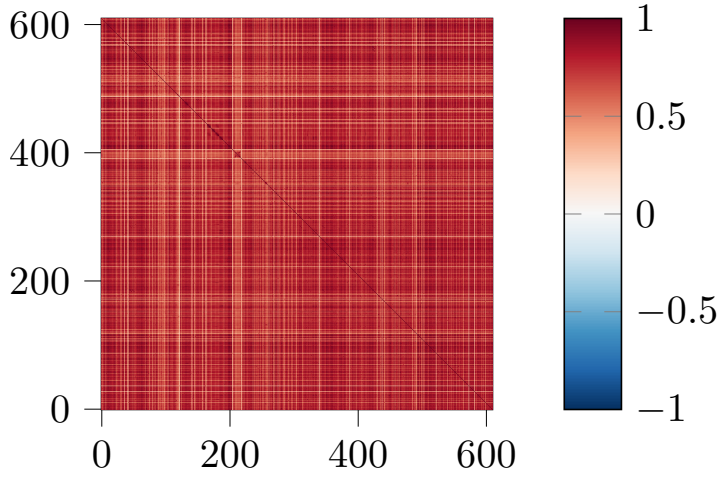


Figure 3.10: Correlation matrix (with 613 regions from the atlas projected on the fMRI images) for raw data

3.4.2 Temporal preprocessing

Slice timing

The use of slice timing in rs-fMRI is an open issue as neurobiologists differ in their opinion. The data were acquired with a repetition time (TR) of 2 seconds with an interleaved acquisition in the axial directions, i.e., in the z-direction, the two neighbouring voxels belonging to different layers were not acquired exactly at the same time. In addition to the temporal lag between the acquisition, if the mouse moves in the acquisition direction, the same layer will be acquired twice in the two successive layers. In this case, a slice timing would not be favourable. We performed slice timing by considering the fact that the mice didn't move much because of the fixed head mount during the acquisition and would not lead to signals being interpolated in an uncertain manner.

Slice timing is suggested to be performed before the realigning of the volumes in the case of limited head motion [93]. The correlation matrix after the slice-timing step has been shown in Fig. 3.11. We can observe that this step doesn't have a major influence on the correlation matrices.

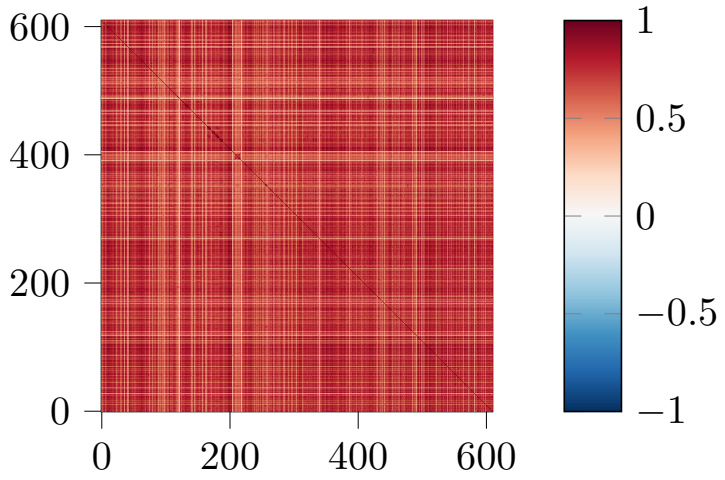


Figure 3.11: Correlation matrix for data after slice timing

Confounding signals

Ventricles were regressed by taking the mean of the ventricle signals in the brain and then regressing this signal. Physiological signals (cardiac and respiratory) are not available for the processed data set, so no regression of such signals was performed. The fMRI acquisition is perturbed by the unstable gradients in the few starting samples. So, we remove the first ten samples of the timecourse. The signals were filtered with a bandpass filter to contain frequencies solely between 0.01 to 0.1 Hz.

In Fig. 3.12 we can see the correlation matrix after the regression of the confounding signals. This time the resultant correlation matrix is significantly different from those in the previous steps. Predominantly, the changes are due to the regression of ventricles signals. The different regions of the brain contain fluid having a similar magnetic response to rs-fMRI as the fluid present in the ventricles.

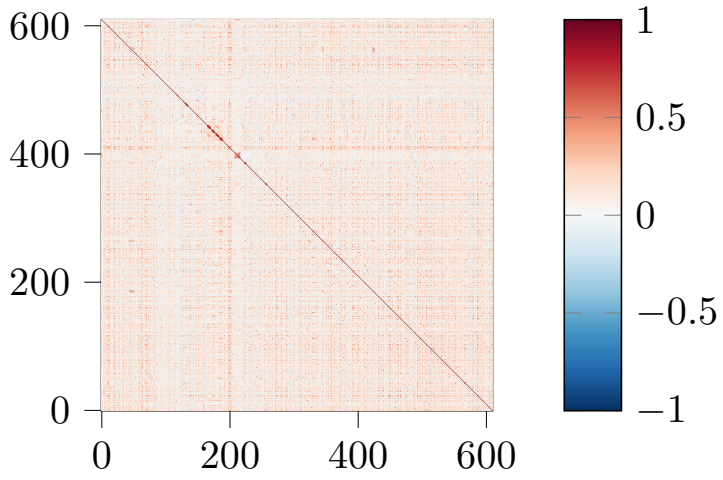


Figure 3.12: Correlation matrix for data after mean ventricles signal regression

Some of the tools were provided to the platform IRIS [83] and hospital and were used for an internal project at the laboratory with the purpose of creating quasi-real synthetic data to validate unmixing methods such as the one developed during this study.

3.5 Unmixing of the semi-real dataset

3.5.1 Validation dataset

For validating the contribution of the high-resolution (HR) segmentation map in unmixing the single-subject fMRI data, a set of synthetic temporal signatures are introduced in seven small regions arbitrarily chosen in the prefrontal cortex. A first synthetic signal is obtained by averaging the signals of the regions ACAd1 and ACAd5, which were already highly correlated in the real data. This signal is then modified to create signals with arbitrary high correlation or anti-correlation for the regions ACAd1, ACAd5, ACAd6a, ACAv1, ACAv5, ORBl1, and PL1 (see blue lines in plots of figure Fig. 3.18). These correlations don't have a physical significance; they are used as ground truth for the evaluation of the proposed algorithm performances. Their correlations are presented in table 3.1.

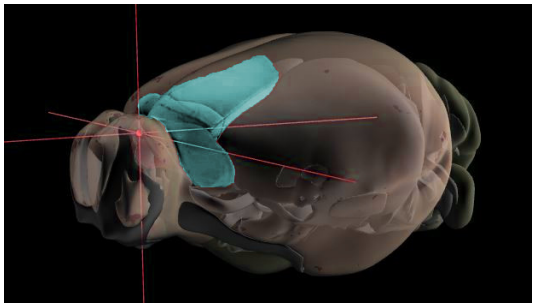


Figure 3.13: Localisation of the prefrontal cortex (in cyan) in the mouse brain where the quasi-real signals were added. The image is a screenshot from the Brain Explorer tool [94] of Allen Institute for Brain Science.

	ACAd1	ACAd5	ACAd6a	ACAv1	ACAv5	PL1	ORBl1
ACAd1	1.00	0.84	0.92	0.67	0.69	-0.45	-0.59
ACAd5	0.84	1.00	0.92	0.67	0.69	-0.47	-0.60
ACAd6a	0.92	0.92	1.00	0.66	0.67	-0.46	-0.58
ACAv1	0.67	0.67	0.66	1.00	0.88	-0.59	-0.82
ACAv5	0.69	0.69	0.67	0.88	1.00	-0.60	-0.82
PL1	-0.45	-0.47	-0.46	-0.59	-0.60	1.00	0.74
ORBl1	-0.59	-0.60	-0.58	-0.82	-0.82	0.74	1.00

Table 3.1: Correlation values between the seven synthetic temporal signatures introduced in the real data set.

Synthetic signals are introduced in the standardised artificially augmented fMRI data, which are then reduced to the initial low-resolution. These synthetic signals are thus mixed with the real signals in the voxels containing a portion of the seven selected regions. This has been visually explained in Fig. 3.14 for a case with two regions $r1$ and $r2$ in high-resolution fMRI data where the signals are added. The thick black lines correspond to the grid of low-resolution pixels, and the grey lines correspond to high-resolution voxels.

The classical fMRI preprocessing pipeline of slice timing and co-registration is applied on the rs-fMRI dataset. Also, the confounding signals are regressed before analysing the data. The next step consists of registering the spatially well-resolved Allen brain atlas template to the artificially augmented anatomical image (which is already perfectly aligned with the rs-fMRI data). The registration of the Allen brain atlas mouse template to the anatomical images provides the deformation field that is applied to the HR segmentation map to transport the different labelled regions on the augmented rs-fMRI data. The spatial resolution of rs-fMRI data is augmented by subdividing each original voxel into $3 \times 6 \times 2$ high-resolution voxels. The registration

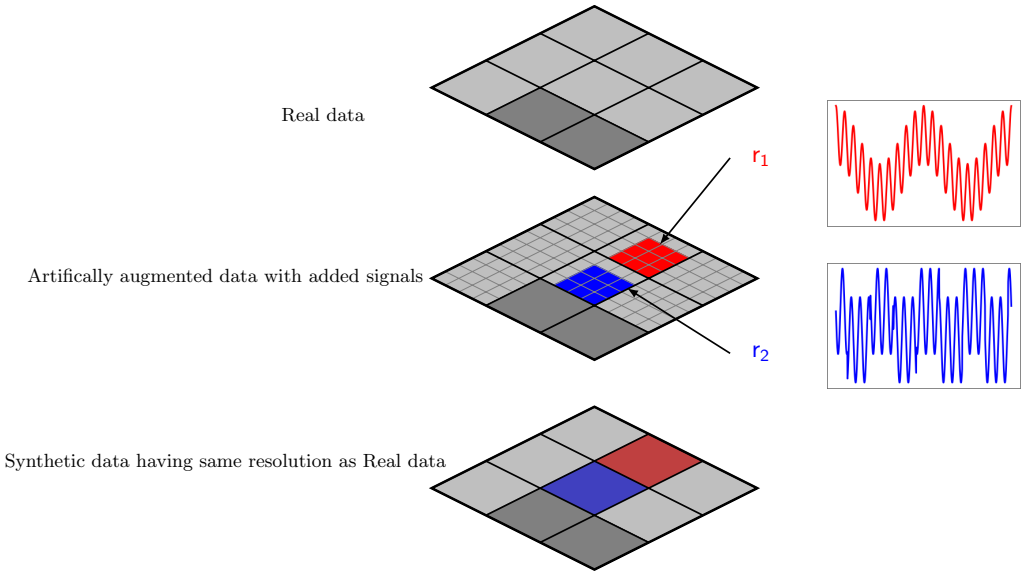


Figure 3.14: The semi-real data was created by augmenting the resolution and then adding signals at certain sub-voxels. Once in the original resolution, the synthetic signals get mixed with the original signals. $r1$ and $r2$ are the regions where the signals were added. The red timecourse is the synthetic signal added to the $r1$ and the blue timecourse is added to the region $r2$.

of the anatomical image to the augmented rs-fMRI leads to an increase in its own resolution.

Synthetic signals are introduced in the standardised artificially augmented fMRI data, which are then reduced to the initial low-resolution. These synthetic signals are thus mixed with the real signals in the voxels containing a portion of the seven selected regions. Since the atlas has a much higher spatial resolution than the fMRI or structural MRI data (up to a factor of 20 in one of the dimensions), the temporal signatures of the different anatomical regions are highly mixed within each low-resolution fMRI voxel. Let us note that the pure pixel assumption is not verified in the regions where the signals were added. The minimum and the maximum number of overlapping regions on the voxels of each region is given in table 3.2.

	ACAd1	ACAd5	ACAd6a	ACAv1	ACAv5	PL1	ORBl1
Min	4	2	4	4	6	4	2
Max	9	12	18	12	17	10	10

Table 3.2: Minimum and maximum number of regions overlapping on the voxels where signals were added.

3.5.2 Algorithm details

Finally, dictionary learning is performed at the (low) resolution of the initial fMRI data $\mathbf{Y} \in \mathbb{R}^{N \times P}$, where $N = 490$ temporal samples and $P = 21024$ voxels, after extracting the brain. The initial abundance matrix $A^{(0)} \in \mathbb{R}^{R \times P}$ is constructed as follows. Let's say that each voxel $i \in \{1, \ldots, \mathrm{P}\}$ was subdivided into J high-resolution voxels during the artificial augmentation step.

For each voxel i of $\mathbf{Y}$ and all regions $r \in \{1, \ldots, R\}$, the element $(A^{(0)})_{r,i}$ will contain the proportion of high-resolution voxels in voxel i, occupied by region r. This can be understood in figure Fig. 3.15. The black margins in the figure belong to the mesh for low-resolution voxels, and the green mesh shows the demarcation between the high-resolution voxels. The hypothesis made for the initialisation is that the functional contribution of a region in a timecourse belonging to a low-resolution voxel is equal to the proportion of the low-resolution voxel spatially occupied by the region after the registration of the atlas. This hypothesis has no biological foundation, but allows to obtain a better initialisation of the abundance matrix based on the only criteria at our disposal (spatial information). This has been shown to be a far better initialisation than a random initialisation for the abundance matrix.

If region r is not transported to the low-resolution voxel i then $(A^{(0)})_{r,i} = 0$. Matrix $\tilde{A}$ which supports the spatial constraint $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ in eq. (2.3) is defined as: $(\tilde{A})_{r,i} = 1$ when $(A^{(0)})_{r,i} > 0$ and 0 elsewhere. This has been explained visually in Fig. 3.16.

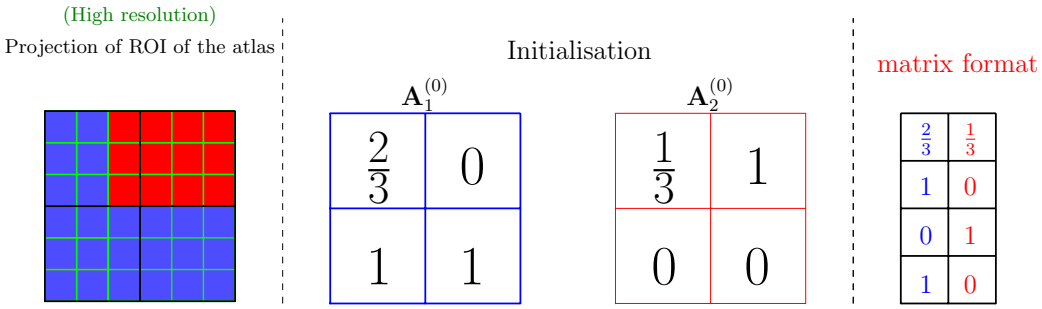


Figure 3.15: Initialisation of $A^{(0)}$ has been shown graphically. In the left-most image, two different regions (red and blue) are projected on the low-resolution fMRI. The black mesh in the first image is for the low-resolution voxels and the green mesh is for the high-resolution voxels. As two-thirds of the top-left low-resolution pixel is occupied by the blue region, $A_1^{(0)}$ is equal to $\frac{2}{3}$. One-third of the same voxel is occupied by the red region so $A_2^{(0)}$ is equal to $\frac{1}{3}$. In the right-most image the matrix form for the image is shown.

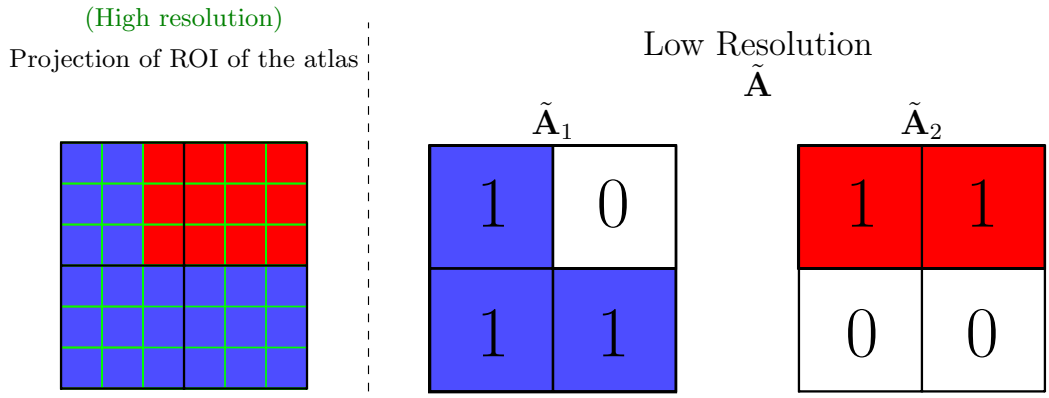


Figure 3.16: We observe that $(\tilde{A})_{1,i} = 1$ where $(A^{(0)})_{1,i} > 0$ and $(\tilde{A})_{2,i} = 1$ where $(A^{(0)})_{2,i} > 0$.

3.5.3 Results and discussion

The proposed DL method is applied to the semi-synthetic validation data set. Empirically, the algorithm converges to an acceptable solution for $\mathbf{A}$ and $\mathbf{U}$ after 500 iterations, see Fig. 3.17, corresponding to a gain on the optimization $\kappa_l < 10^{-3}$ (eq. (4.10)). For the estimation of $\mathbf{A}$, the FISTA algorithm requires a stopping criterion or a maximum number of iterations. In our implementation, FISTA is stopped when $\|\mathbf{a}_i^{(l-1)} - \mathbf{a}_i^{(l)}\|_2 < 10^{-8}$ or $l > 100$.

Fig. 3.18 shows the estimated timecourses in dashed red against the ground truth signals in blue. The mean squared error (MSE) of the estimated timecourses are given on the plots. Despite the strong mixing in the voxels of the seven regions, our algorithm provides a very good estimate of the synthetic timecourses introduced in the data. In neurosciences applications, these timecourses are then used to build functional brain networks.

To highlight the crucial contribution of a well-registered high-resolution segmentation map, we have applied the standard ANTS registration algorithm to the validation dataset without handling the augmentation of resolution. The Allen brain atlas template is thus directly registered on the low-resolution anatomical image using ANTS (initialisation of $\mathbf{A}^{(0)}$ is straightforward in this case). Figure 3.19b shows the correlation matrix obtained in this case, after 500 iterations. The inaccurate initial projection of the different anatomical structures on the low-resolution fMRI data yields a poor initialisation $\mathbf{A}^{(0)}$ for the abundance matrix. This results in a correlation matrix in figure 3.19b where the estimated correlations are far away from the ground truth.

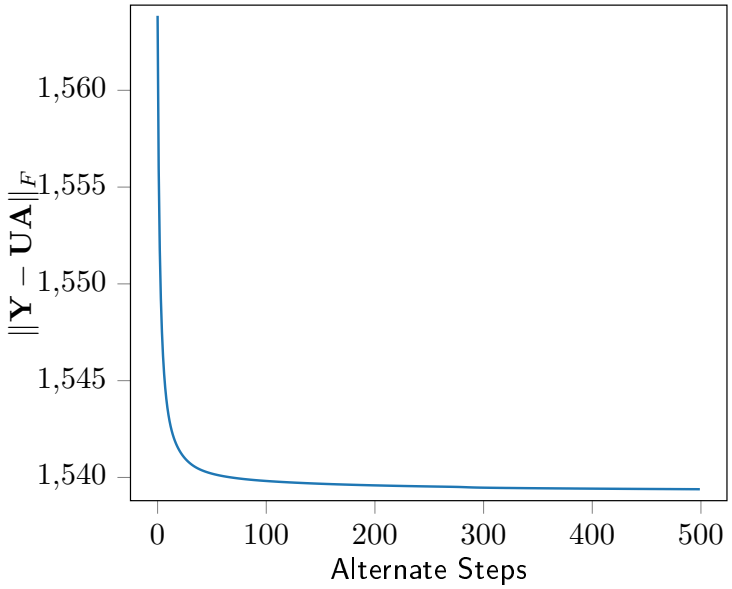


Figure 3.17: Rs-fMRI data. The figure shows that $\|\mathbf{Y} - \mathbf{UA}\|_F$ converges; as is seen by the flattening of the curve in the last iterations.

3.5.4 Precision in the estimation of A

Validation of the estimation of $\mathbf{A}$ on such an example is difficult, considering that no ground truth is available. Recently functional connectivity has been proven to be non-stationary [95], consequently inside the same voxel, the contribution of the different regions can change with time (*i.e.* the abundances). These changes would contribute to non-stationary changes in the brain signals. These other signals mix with the stationary signals added, so stationarity does not exist in the quasi-real dataset. Even if we know that the connectivity is non-stationary, we can assume that these non-stationarities are present in a uniform manner at the start and at the end of the acquisition. So, to validate the convergence, we divided the time-course matrix into two parts with an overlap of 5 samples (having almost half of the timecourse) and ran our method of alternate optimisation on each part individually and then with the whole timecourse matrix. It is possible that by performing the unmixing over two time intervals of the same dataset, the estimated abundances are not identical for the same given region in the same voxel. We can only assess the consistency of the estimate on the regions where a synthetic signal has been introduced because these signals and the proportions are stationary. Then, the mean square differences/deviations (MSD) between the abundance vectors for each case for the regions ACAd1, ACAd5, ACAd6a, ACAv1, ACAv5, ORBl1, and PL1 were calculated. Abundances were also calculated for two randomly chosen regions RR1

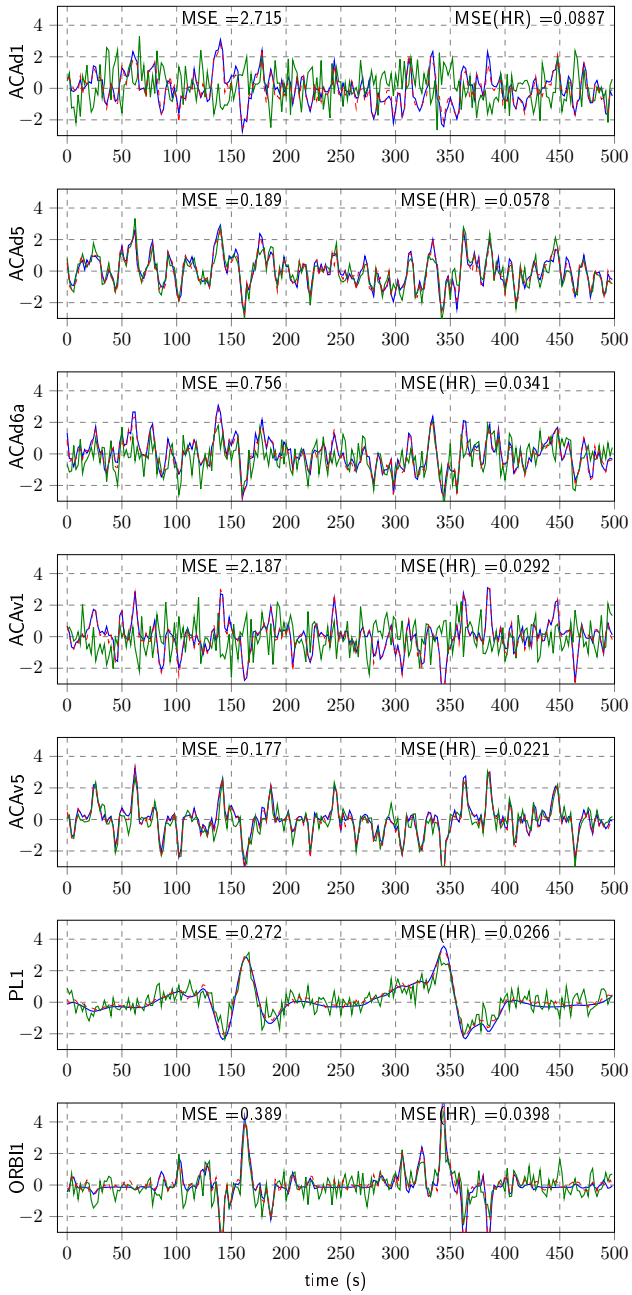
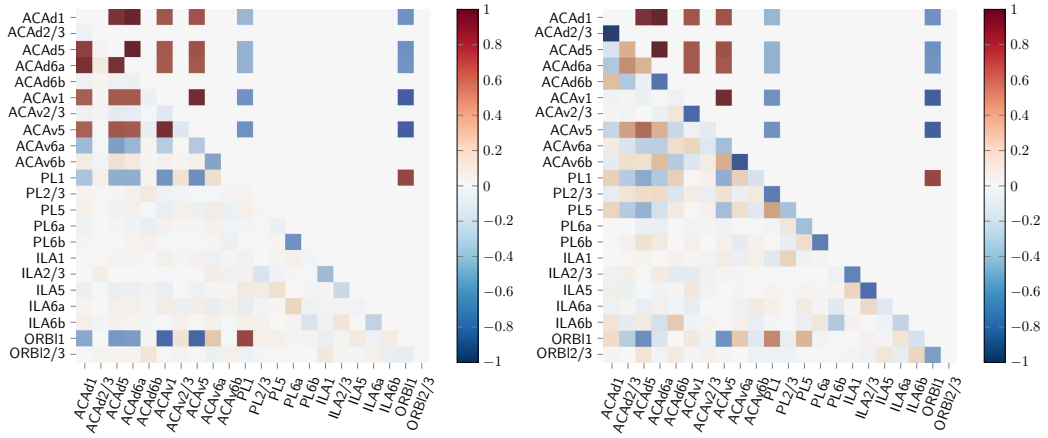


Figure 3.18: The plot shows samples corresponding to the first 500 seconds of the synthetic signals (in blue) and their corresponding estimated timecourses using HR altas information (in dashed red) and without using HR atlas information (in green).



(a) Correlation matrix from signal estimated guided by HR segmentation map

(b) Correlation matrix obtained without HR registration

Figure 3.19: Correlations in the prefrontal cortex. The lower triangular matrix contains estimated correlation and the upper triangular matrix contains the true ones for the seven synthetic signatures. Diagonal elements are set to zero.

and RR2. In table 3.3, $MSD_{1,full}$ refers to the MSD between abundances of the first half and the whole signal, $MSD_{2,full}$ refers to the MSD between abundances of the second half and the whole signal and $MSD_{1,2}$ refers to the MSD of the abundances between the first half and the second half. We observe that the differences between $MSD_{1,full}$, $MSD_{2,full}$ and $MSD_{1,2}$ are really low, suggesting a good precision in the estimation of $\mathbf{A}$ for the three different cases. The signals estimated were also stable; the differences occurring most probably due to the dynamics present in the brain.

	ACAd1	ACAd5	ACAd6a	ACAv1	ACAv5	PL1	ORBl1	RR1	RR2
$MSD_{1,2}$	0.030974	0.034711	0.033012	0.018284	0.042602	0.056220	0.011396	0.002691	0.128761
$MSD_{1,full}$	0.020485	0.013207	0.013604	0.007573	0.018063	0.006392	0.015329	0.000945	0.071989
$MSD_{2,full}$	0.005803	0.013927	0.007179	0.005692	0.013945	0.024941	0.012262	0.002643	0.045751

Table 3.3: Mean square deviations between the abundance vectors for each of the seven regions where the synthetic signals were added. RR1 and RR2 are two randomly selected regions.

3.6 Change detection

For medical studies and comparisons between the functional connectivity of healthy and pathological individuals, correlation matrices are used to infer the differences. Once the correlation matrices are obtained for the individuals, such as the one shown in 3.19a for the synthetic case, various statistical tests can be performed depending upon the study in question. If the study is a longitudinal analysis of an individual, the simplest of the tests could be a Gaussian test on the difference of pairs of regions using coefficient matrices calculated at two different time-points for a mouse. If it is a group study then to find the difference between two different groups, i.e., wild-type and Alzheimer's or depression, a Student's t-test or a permutations test could be performed (Appendix A). One of the interests in this thesis is to find the correlations coefficients that differ in Alzheimer's disease mice from those in normal or wild-type (WT) mice.

3.6.1 Longitudinal change detection

In WT mice, the longitudinal changes are principally due to ageing, whereas for pathological mice, there are changes due to ageing as well as the progression of the disease. To detect longitudinal changes due to diseases, our idea is to detect the pairs of regions which would suffer connectivity changes due to ageing in WT mice and then to compare these changes to the changes observed in pathological mice. To demonstrate the use of Gaussian statistical test to look for changes, the data given in section 3.5 was used to generate two timepoints. Two datasets were generated from the example given in section 3.5. For the first dataset, we ran the alternated optimisation method considering the first 230 time samples of the real fMRI dataset; this can be considered as timepoint A or the case without any changes. To construct the second dataset, we ran the proposed method considering the last 230 time samples of the quasi-real fMRI dataset; this can be considered as timepoint B. As the mouse used at timepoint A and B is the same, so there won't be many changes in most of the regions, except for the seven regions where the signals were added. There is no overlap between the time samples of the two datasets. The two datasets are considered to be i.i.d. considering the dynamic connectivity [95] present in the rs-fMRI datasets. Visually we can observe that there is a difference between the correlation matrices at timepoint A (see Fig. 3.20) and B (see Fig. 3.21) because of the dynamic connectivity present. The regions shown in the figures here are the sub-regions belonging to the prefrontal cortex. The name of the regions, in their order of occurrence in the correlation matrices Fig. 3.20 and Fig. 3.21 are given in B.1.1.

For Gaussian statistical tests, the correlation coefficients ρ are z-transformed to Gaussian values with the Fisher transformed values given here:

$$z = \frac{1}{2} ln \left(\frac{1+\rho}{1-\rho} \right) \tag{3.2}$$

This is done to transform the distributions of correlations coefficients (generally skewed) into normal distributions.

If the true correlation coefficients at time-point A and B are given by ρ_A and ρ_B respectively, then the Fisher transformed correlation coefficients at time-point A and B can be considered to follow a normal distribution given by

$$z_A \sim \mathcal{N}\left(\frac{1}{2}ln\left(\frac{1+\rho_A}{1-\rho_A}\right), \sigma_A^2 = \frac{1}{N_A - 3}\right) \tag{3.3}$$

and

$$z_B \sim \mathcal{N}\left(\frac{1}{2}ln\left(\frac{1+\rho_B}{1-\rho_B}\right), \sigma_B^2 = \frac{1}{N_B - 3}\right). \tag{3.4}$$

Here $N_A = N_B = 230$ are the length of the timecourses. We perform a Gaussian statistical test, and we define, in the absence of any changes in the two time-points, the following null hypothesis:

$$\mathcal{H}_0 : \rho_A = \rho_B \tag{3.5}$$

Following, the distribution of the difference $D = Z_A - Z_B$ follows under the null hypothesis:

$$D \mid \mathcal{H}_0 \sim \mathcal{N}\left(0, \sigma_A^2 + \sigma_B^2\right) \tag{3.6}$$

and under the alternative:

$$D \mid \mathcal{H}_1 \sim \mathcal{N}\left(\theta, \sigma_A^2 + \sigma_B^2\right), \theta \neq 0. \tag{3.7}$$

The changes between the two time-points were obtained using the Gaussian statistical test explained before. A two-tailed test is performed on the data. The p-values [refer to section A.1] for our problem are calculated using the formula:

$$p_{x_i} = 2 \times (1 - \Phi(D \mid \mathcal{H}_0)), \tag{3.8}$$

where Φ is the cumulative distribution function of the standard Gaussian distribution.

The expected p-values are supposed to follow a uniform distribution under $\mathcal{H}_0$ and this is reflected in the histogram given in Fig. 3.22. The peak on the left-most likely corresponds to p-values close to zero i.e. test under $\mathcal{H}_1$. The plot for p-values arranged in an increasing order has been given in Fig. 3.23, we see that the curve for p-values is flatter in the beginning because of the low p-values. The zoomed-in section with the smallest 100 p-values for the plot of p-values is given in Fig. 3.24; we see that there are many p-values close to zero qualifying them as discoveries. To threshold p-values or in turn reject the $\mathcal{H}_0$'s associated, Benjamini Hochberg [96] procedure was followed. This enables control of FDR (false discovery rate) and is explained in A.2.3. The q-value was set to 0.01 to control the FDR. The FDR line ($\frac{q_i}{N}$), where i

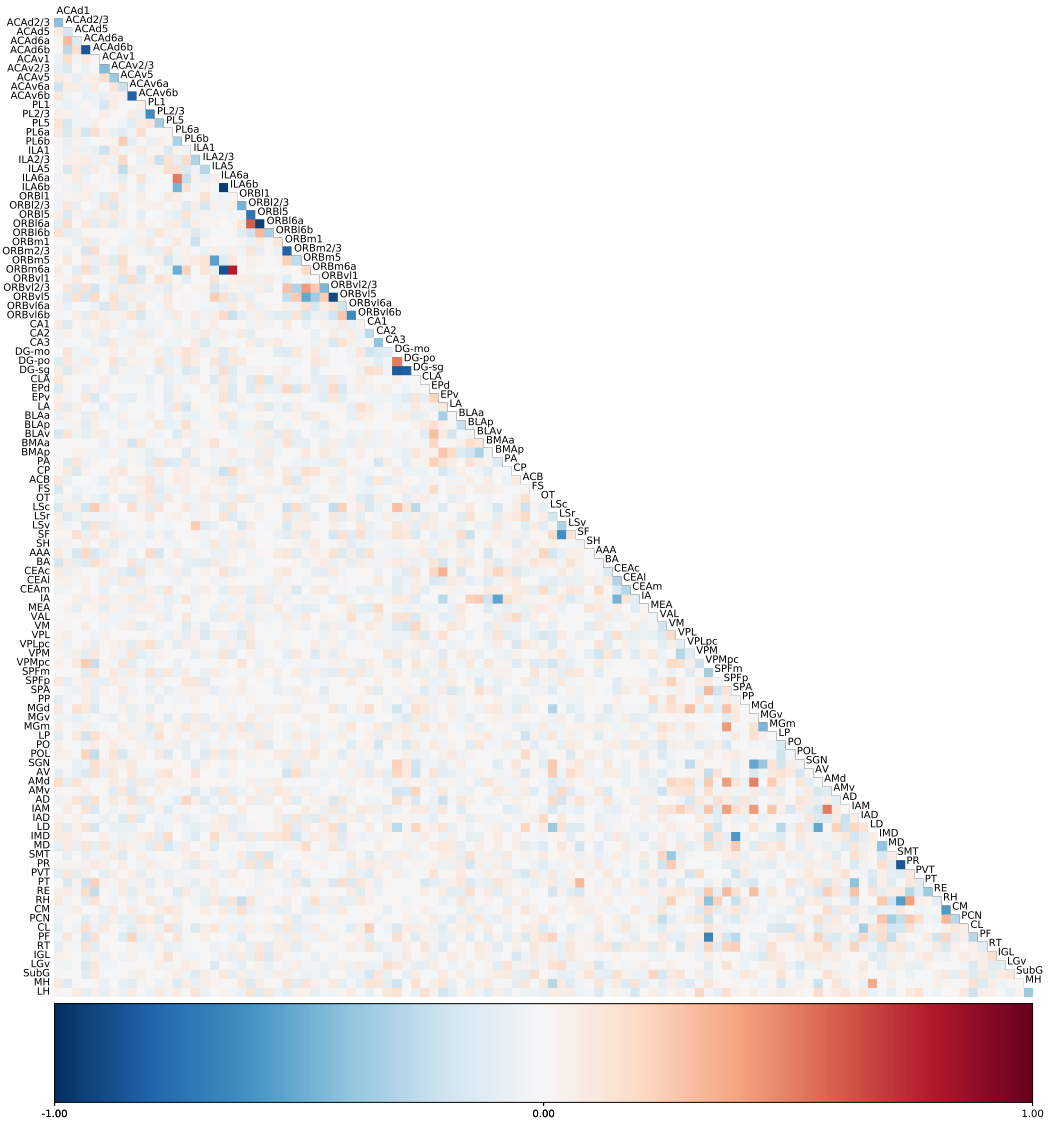


Figure 3.20: The correlation matrix considering all the regions under the prefrontal cortex at the considered timepoint A.

is the index corresponding to the sorted p values, in the figure can be seen in red. The values lying below the threshold are the changes observed for the pair of regions in the two timepoints. In Fig. 3.25 we can see the pairs that can be considered to have changed from timepoint A to timepoint B. We can observe that the majority

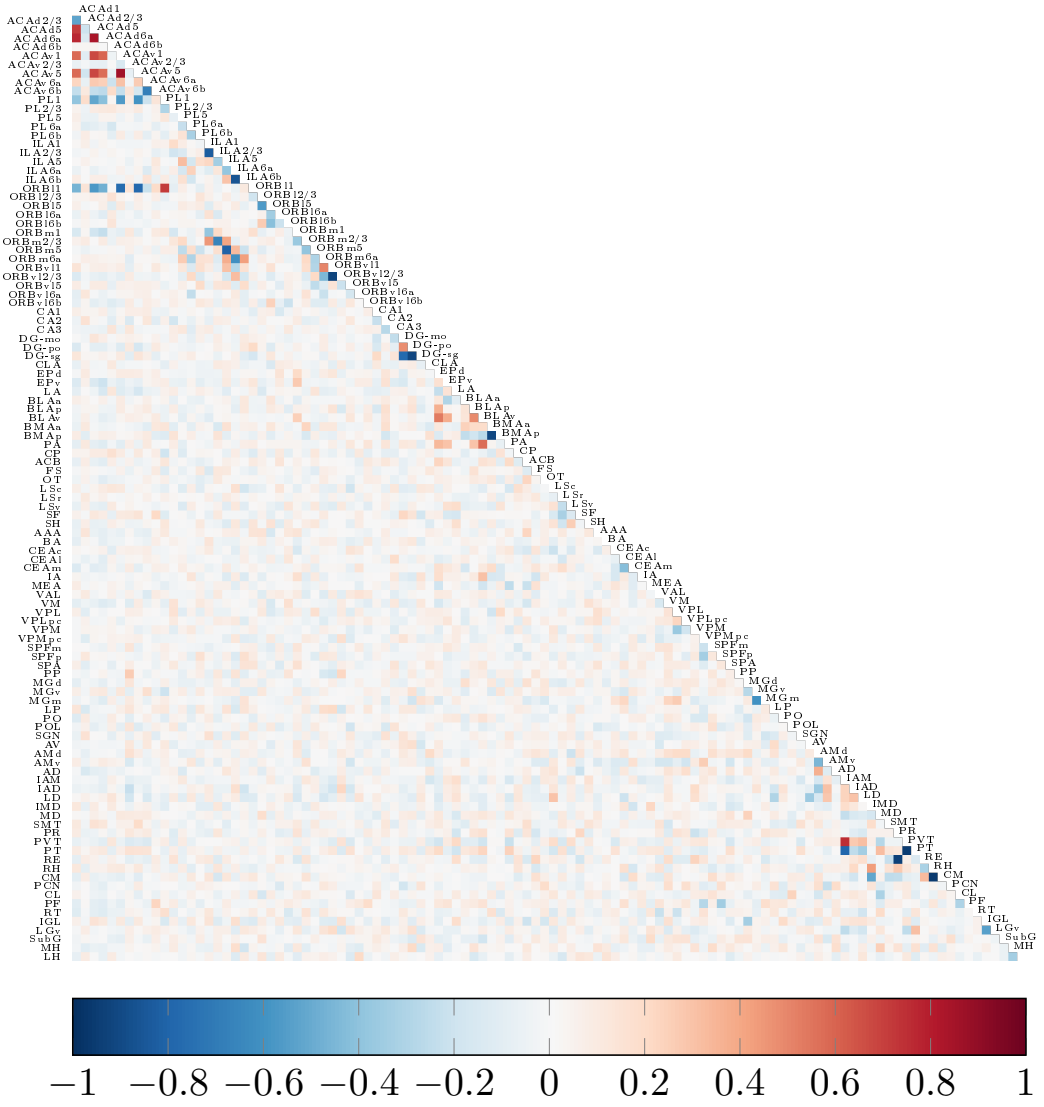


Figure 3.21: The correlation matrix considering all the regions under the prefrontal cortex at the considered timepoint B.

of the changes are in the regions where the synthetic signals were added. A few detections were made in the regions where the signals were not added, for example, the pairs ACAd6a-ACAd6b, ILA1-ILA2/3, ORBl5-ORBl6a are detected. This could be explained by looking at Fig. 3.26 where the dice coefficients between the masks of the different regions have been shown. By comparing Fig. 3.25 and Fig. 3.26 we can

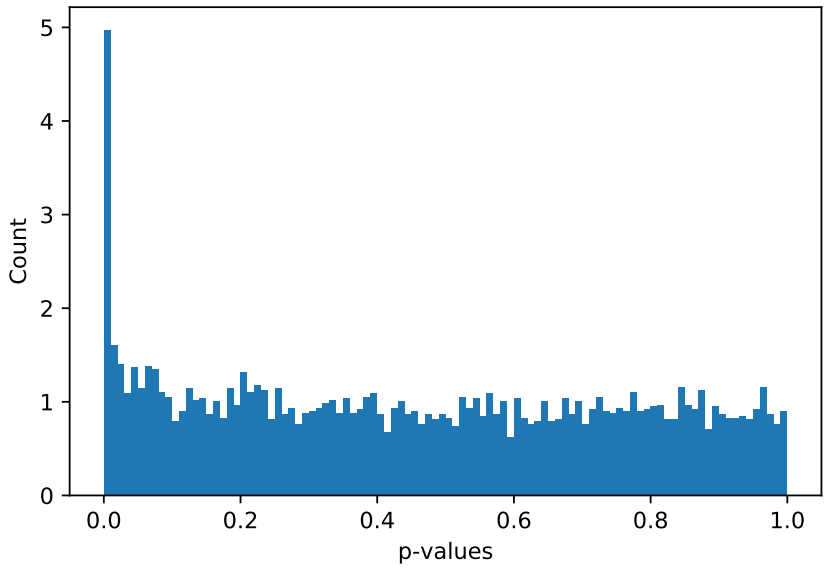


Figure 3.22: For the longitudinal test the p-values under $\mathcal{H}_0$ follow a uniform distribution and can be observed here. The small p-values towards the left most likely correspond to $\mathcal{H}_1$

evaluate that the regions where the signals were not added have fAllen under change detection because of some overlap between the regions. We observe that ACAd6a-ACAd6b overlap and the signal was modified just for ACAd6a; ILA1 and ILA2/3 overlap with PL1, ACAv1 and ACAv5 where signals were modified; ORBvl2/3 and ORBvl5 overlap with ORBl1 whose signals were changed. When we add signals in ORBl1, we can visualise that the signals would affect unmixing in the neighbouring regions such as Orbl 2/3, Orbl6a and Orbl6b. Other reasons for detections where signals were not added could also be due to not perfect unmixing as the problem is ill-posed, the non-stationarity due to dynamic connectivity could also be one of the causes. The unmixing was performed in the whole brain, so it possible that there is some overlapping of the regions in the prefrontal cortex to the bordering sub-regions of the neighbouring regions of the prefrontal cortex. If the q-value is further decreased, then some of the pairs with low correlation coefficients in table 3.1 fall under $\mathcal{H}_0$.

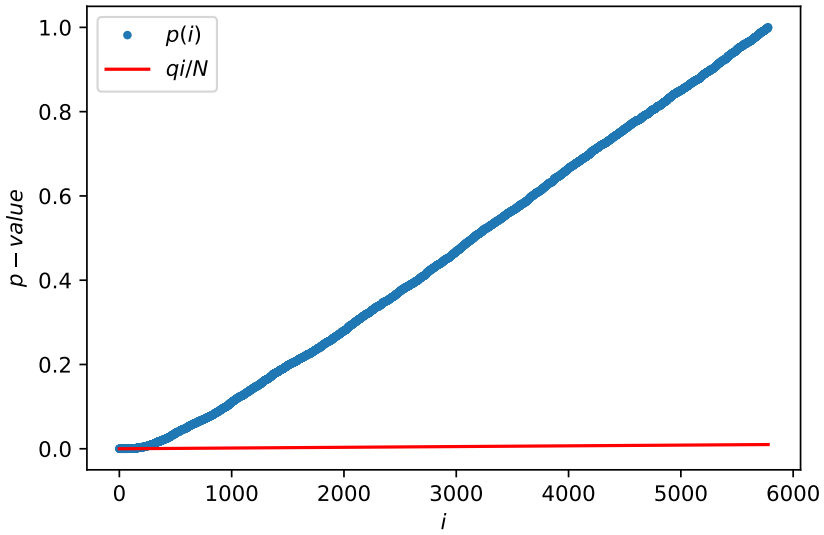


Figure 3.23: Longitudinal test. The plot for p-values arranged in an increasing order and the line $\frac{q_i}{N}$ cutting the p-values.

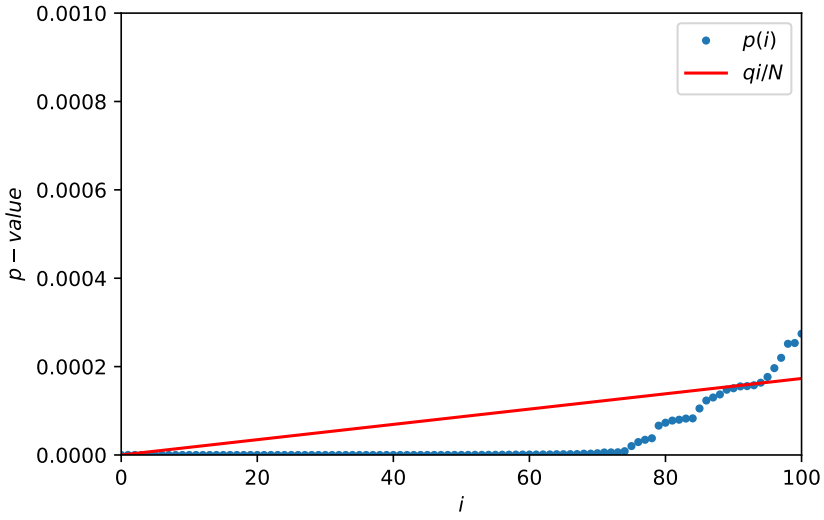


Figure 3.24: The zoomed-in portion showing the FDR line cutting the p-values, the values under the line qi/N are considered under $\mathcal{H}_1$ with the Benjamini-Hochberg procedure for controlling FDR.

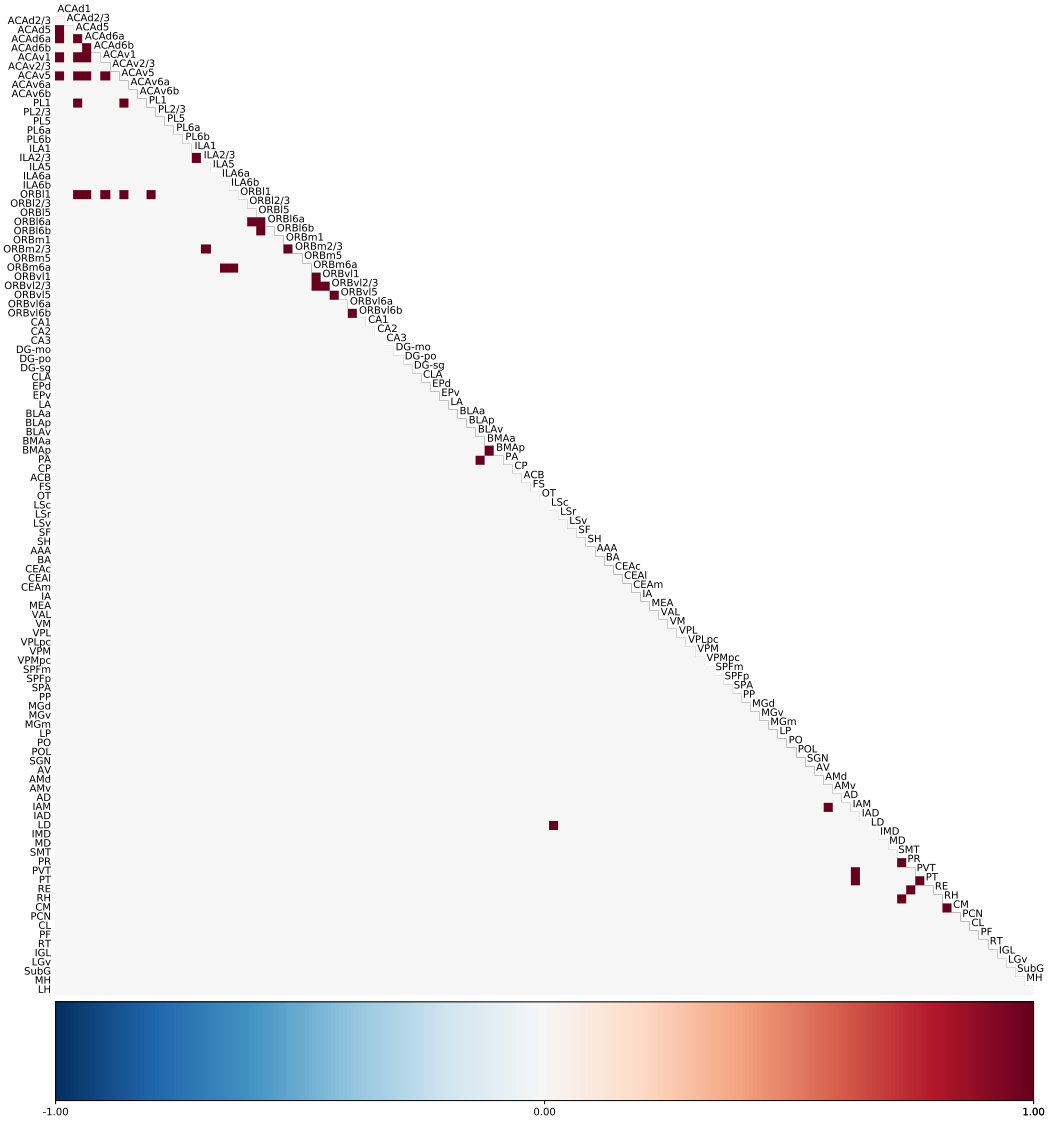


Figure 3.25: The correlation coefficients which significantly varied in the correlation matrix at timepoint A and timepoint B are marked here with a value equal to 1. The detections were obtained by thresholding with the FDR control of Benjamini-Hochberg at $q = 0.01$.

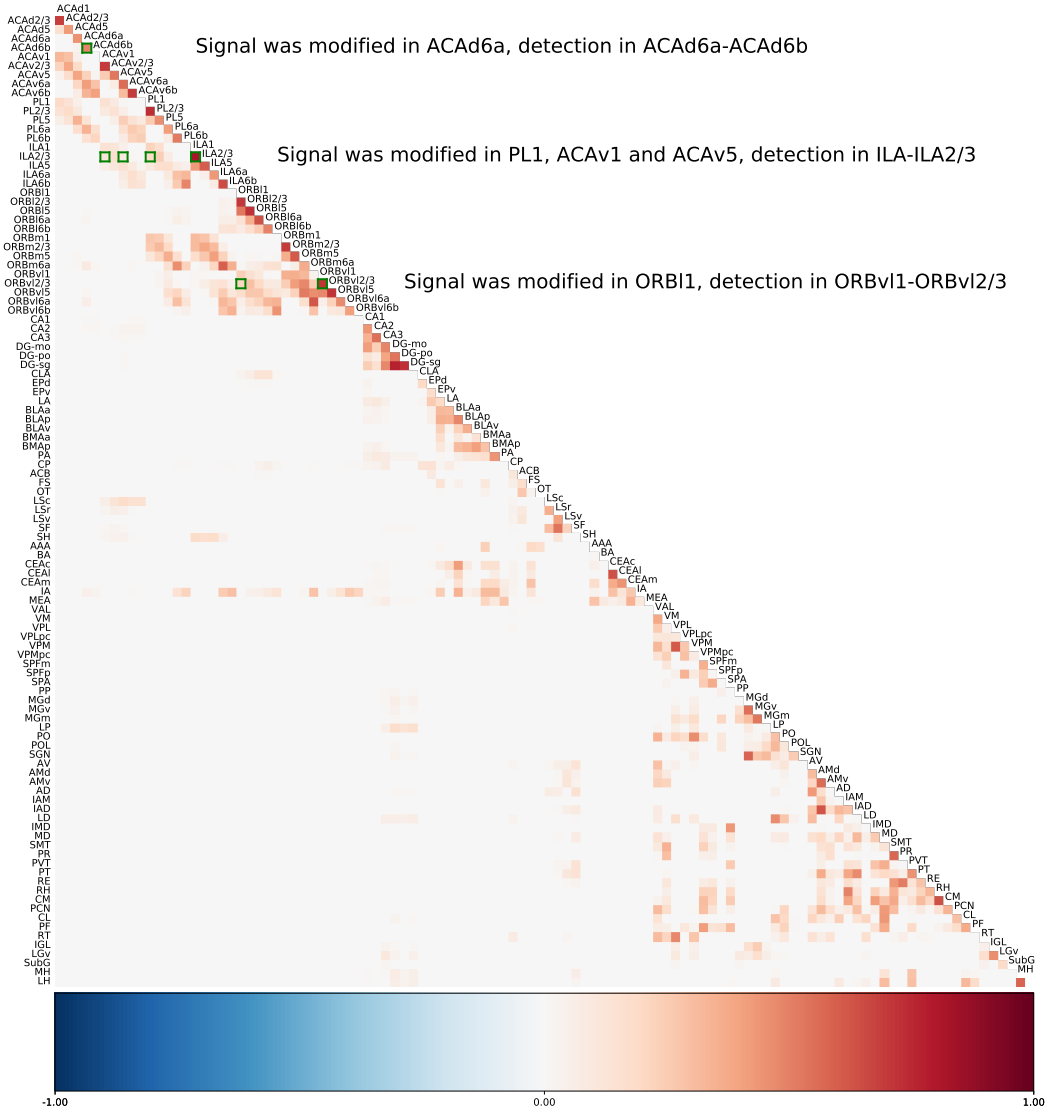


Figure 3.26: Dice coefficients between regions have been shown. Analysis of the dice coefficients enables better understanding of the false discoveries (shown in green).

3.6.2 Change detection in groups

Once the rs-fMRI data of single-subjects is unmixed and the correlation matrices calculated, change detection in groups can also be considered. Let us consider a group of synthetically generated correlation matrices of size 49×49, let us call these collections group 1 and group 2. One of the ideas in our approach on real data is to find the connectivity changes between the pairs of regions of a control group from a diseased group. If the connectivity of a pair of region is significantly different in the diseased mice from a pair in the control mice, we can say that it is a detection. To generate synthetic data, we generated 10 correlation matrices for the control for which the correlation coefficients follow a normal distribution $\mathcal{N}(0.3, 0.02)$ (group 1), and other 10 correlation matrices of the same size for the group 2 with the same mean and variances. Some correlation coefficients in the matrices for the second group were altered to follow $\mathcal{N}(0.35, 0.02)$. The changes were made only in the means of the distributions and not in the variances as the groups in real data passed the test for homoscedasticity. One such correlation matrix with the correlation coefficients in the last two rows following a different normal distribution is shown in Fig. 3.27. In a first approach, we will try to find these pairs which are different in the two groups. To that end, we use the Student's t-test and the permutations test. This was done to evaluate the performances of the two tests on the synthetic groups where the ground truth is known.

Student's t-test:

Once we have the correlation coefficients between different regions obtained from the timecourses estimated from the rs-fMRI data for a mouse of different groups, we can find connectivity changes by performing the Student's t-test (check A.1.2). The correlation coefficients were Fisher transformed (formula in (3.2) before performing the Student's test. In Fig. 3.28 we can observe the distribution of the test statistic. The bell curve on the right corresponds to the values that would most likely fall under $\mathcal{H}_0$, and the scattered values on the left would be the values which would fall under $\mathcal{H}_1$. The problem to detect changes is approached with the FDR test for multiple comparisons; in Fig. 3.29 we can see the expected p-values plot with the smaller p-values probably belonging to $\mathcal{H}_1$. In Fig. 3.30 the correlation coefficients that were modified in the group 2 mice can be seen. The Student's test was performed for the second time with $q = 0.05$ to look for quantitative analysis and to check the rate of different statistical errors explained in the appendix section A.1. The rate of errors calculated for 100 tests is given in table 3.4. It can be seen that when the q-value is decreased (0.05 rather than 0.1), the rate of false positives is lower. Depending on what we are interested in, the q-values can be adjusted. In this work, the interest is to find regions that actually undergo changes, so the q-values should be as low as possible. This would lead to fewer false positives or fewer discoveries which can mistakenly be considered as changes.

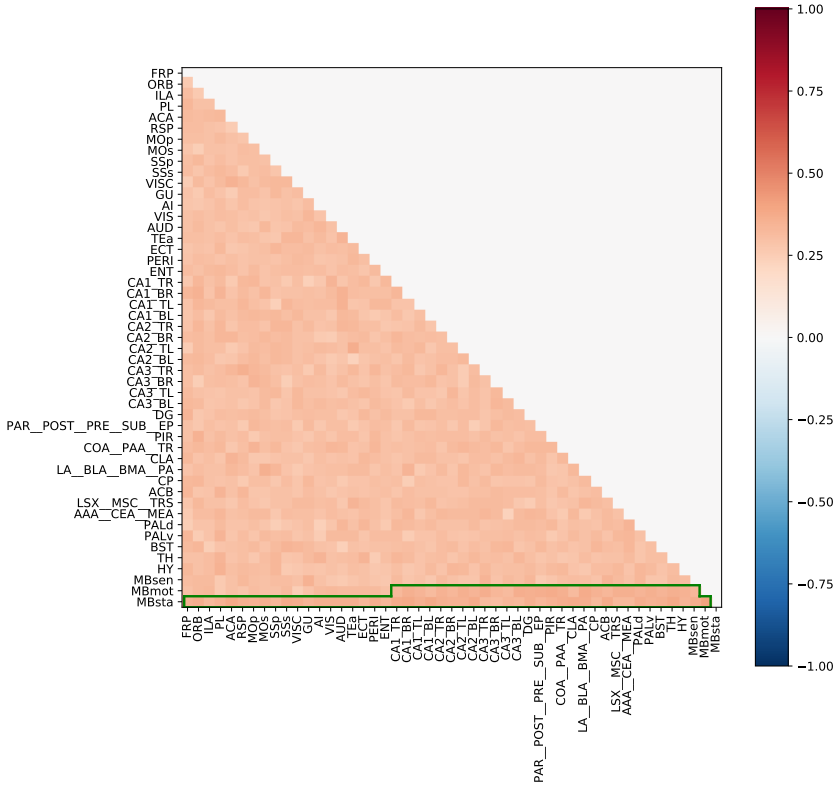


Figure 3.27: Correlation matrix for the second group. The correlation coefficients highlighted in green are the coefficients that were modified for the second group and should be detected as changes.

Proportions of:	False positives	True positives	False negatives	True negatives
q = 0.05	0.037	0.963	0.002	0.998
q = 0.1	0.083	0.917	0.001	0.999

Table 3.4: Student's test evaluation at two different FDR levels, $q = 0.05$ and $q = 0.1$

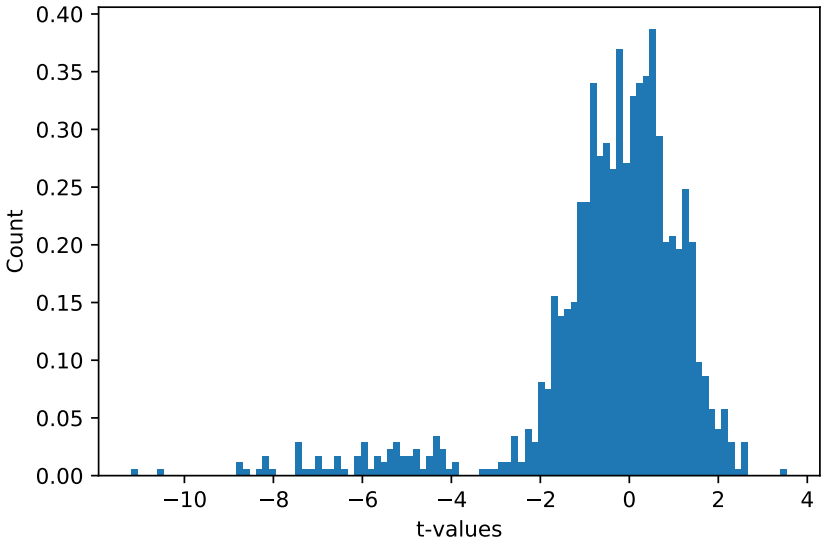


Figure 3.28: Histogram showing the distribution of the test statistic, the plot under $\mathcal{H}_0$ is supposed to follow a Student's t-distribution with the degrees of freedom(d.o.f.) equal to the 18 (Here $N_1 = 10$, $N_2 = 10$, d.o.f. $= N_1 + N_2 - 2$)

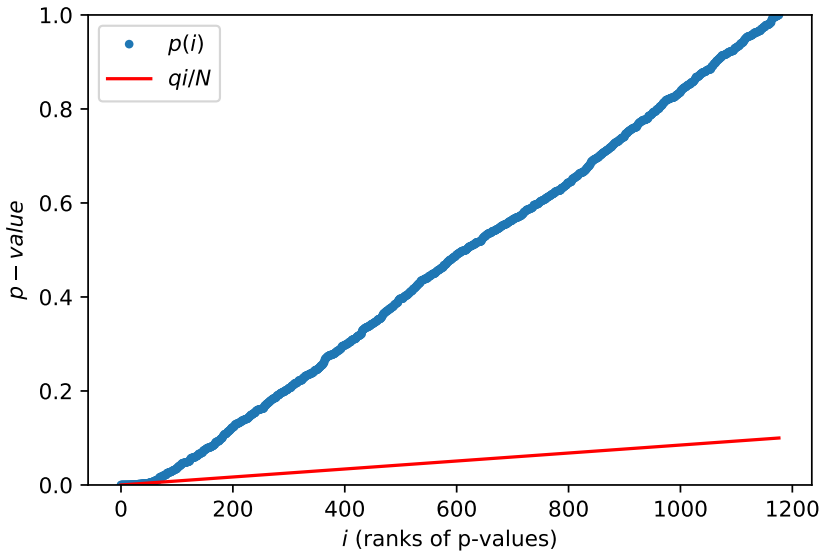


Figure 3.29: FDR line with q = 0.1 cutting the p values, the values under the line qi/N are considered under $\mathcal{H}_1$

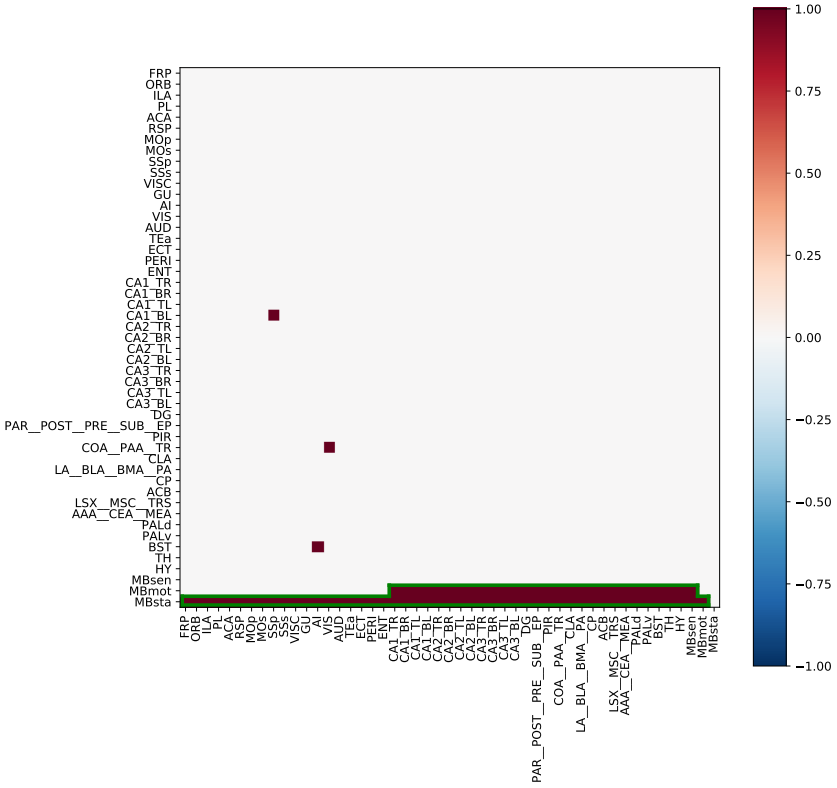


Figure 3.30: The pairs of correlation coefficients which were detected to have changed using the Student's t-test. The true positives have been highlighted in green.

Permutations testing

For real data, a distribution for the data under $\mathcal{H}_0$ may not always have a known form. In addition to that, for small sample sizes, the central limit theorem does not hold well. So, one strategy to perform statistics on real data in such a case is to use a non-parametric test in place of the Student's test explained earlier. One such non-parametric test which could be useful is the permutations test. It consists of estimating the null distribution empirically by randomly shuffling the individuals from the different groups. In the case of group studies, to detect changes in connectivity, we perform the random permutations tests to find the significance of the differences in the correlation coefficients for the two groups. The differences in the correlation coefficients are stored in the form of a vector. This is repeated to obtain a large number of samples, which is then used to generate a histogram representing the empirical $\mathcal{H}_0$. In our case, the groups were permuted 10000 times. In Fig. 3.31 shows the empirical distribution of the test statistic. The p-values were calculated by considering the proportions of the distribution higher than the test value. In Fig. 3.32 we can see that p-values under $\mathcal{H}_0$ follow a uniform distribution. The q-value was set to 0.1 for this case; in Fig. 3.33 we can see the FDR line (in red) and the p-values lying below it. In Fig. 3.34 the correlation coefficients changes detected in the two groups by using FDR have been shown. Rate of false discoveries, true discoveries, false missed findings and true missed findings for $q = 0.05$ and $q = 0.1$ are given in table 3.5.

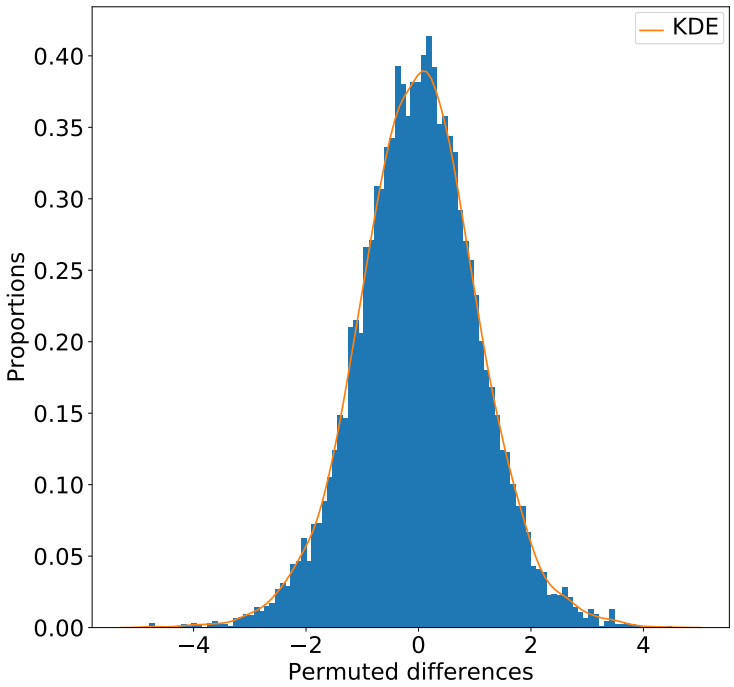


Figure 3.31: Histogram showing the distribution of permuted differences for the permutations test for the synthetic data and the kernel density estimation (KDE) in orange.

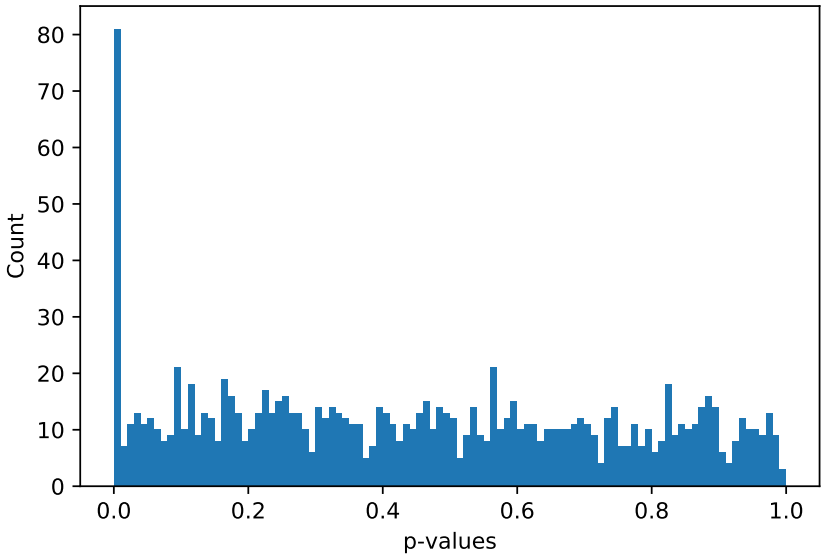


Figure 3.32: Distribution of p-values for the permutations tests, we can see that the p-values under $\mathcal{H}_0$ follow a uniform distribution.

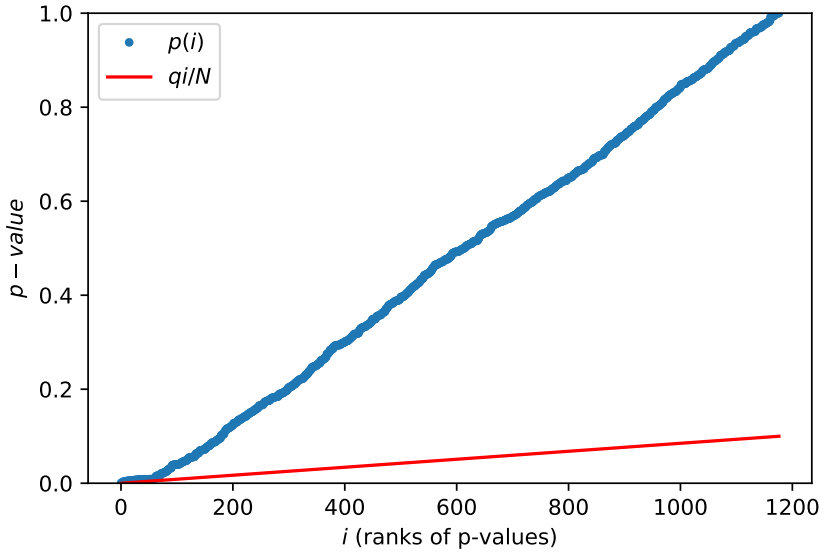


Figure 3.33: FDR thresholding of the p-values for the permutations test

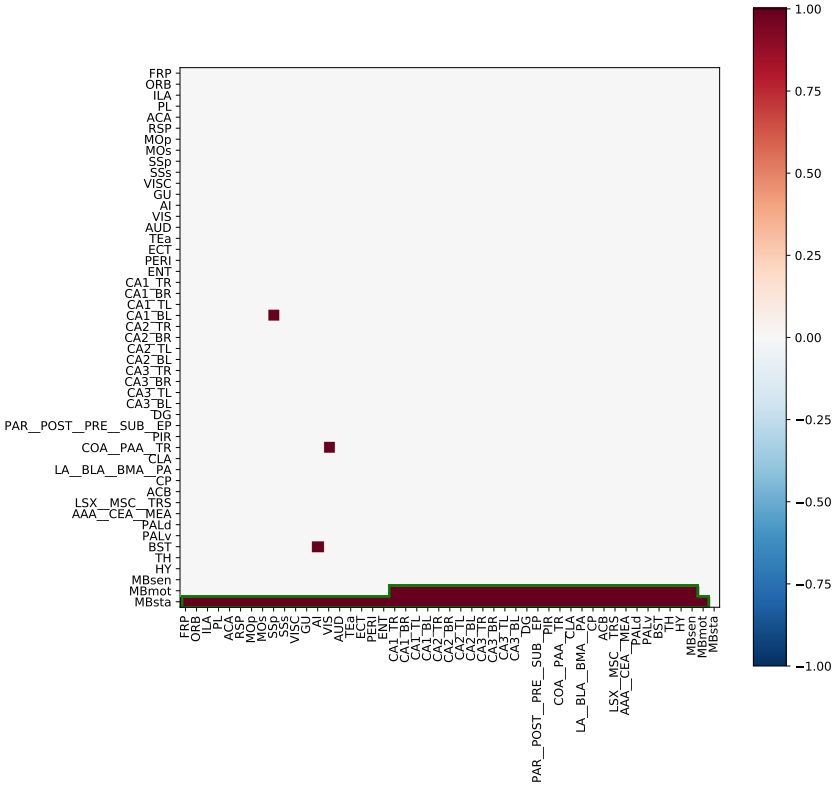


Figure 3.34: The pairs of correlation coefficients which were detected to have changed using the permutations test. The true positives have been highlighted in green.

Proportions of:	False positives	True positives	False negatives	True negatives
q = 0.05	0.0272	0.9728	0.0027	0.9973
q = 0.1	0.095	0.905	0.0	1.0

Table 3.5: Permutations test evaluation for different values of q.

The two tests explained here, i.e., the permutations test and the Student's test are used in the case of real data in the next section to check for detections for the group studies. It can be observed that the false detections for the cases shown for the permutations testing (Fig. 3.34) and the Student's t-test (Fig. 3.30) occur at the same place. However, from table 3.4 and table 3.4 it can be observed that it is not always the case as the values of the rates of false detections is not the same. Using the synthetic example, one test was not preferred over the other. Both tests were implemented for the study on real data.

3.7 Study on real data

Similar to the application on the quasi-real datasets, the dictionary learning method with spatial constraints proposed in the manuscript was used to unmix real fMRI data. It was done for two cases: in the first case, the small anatomical regions of the Allen brain atlas were grouped together to allow a regional-scale analysis of connectivity based on the correlation matrices obtained from the timecourses of the unmixed regions. The other case involved the unmixing at a finer scale, taking into account all the anatomical regions of the Allen brain atlas.

3.7.1 Group tests

Before applying the statistical tests directly on the real data for the first case, the correlation matrices are analysed. For this purpose, hierarchical clustering was employed on the means of the z-scores matrices to check clusters of regions showing similar brain activity for WT mice at 9 months. The metric for hierarchical clustering is the Euclidean distance between the different vectors formed by the z-scores pairs. The cerebral activity of the mice at 9 months is considered to be completely normal and would not show any signs of ageing. In Fig. 3.35, we can observe the results of the hierarchical clustering on the means of the z-scores of the correlation matrices at 9 months. The suffixes BR (Bottom Right) refers to the right dorsal, BL (Bottom Left) to the left dorsal, TR (Top Right) to the right superior, and TL (Top Left) to the left superior parts of the sub-regions of the hippocampus, i.e., CA1, CA2, CA3. The sub-regions of the hippocampus were divided so as to check the activity in the interior of these regions; the motivation behind this was the evidence provided in the neuroscience of the different activity between these partitions. The partitioned sub-regions of the hippocampus (a high functioning area of the brain), can be seen classified together; the left and right parts (TL-TR, BL-BR) can be observed clustered together for CA1, CA2 and CA3.

In Fig. 3.35, it can be seen that regions SSs (supplemental somatosensory area), SSp (primary somatosensory area), RSP (retrosplenial area) and, MOs (secondary motor area) belonging to the cortical network are belonging to the same high-level cluster. MBmot (midbrain, motor related) and MBsen (midbrain, sensory related) can be seen in the same cluster with AUD (audio) and VIS (visual) grouped not far away. For a comparison with the previous results on the same data in [97], there were some structures in the hierarchical clustering which were similar and show high correlation; these are ORB (Orbital area), MOp (primary motor area), MOs and SSp. MBSen and visual sub-regions form a network in [97] and can be observed to be clustered together in Fig. 3.35. The sub-regions ACB (Nucleus accumbens) and the LSX_MSC_TRS (lateral septal complex, medial septum, and triangular nucleus of the septum) belong to the basal ganglia; they are strongly connected and belong to the same cluster in the hierarchical clustering. These show a strong correlation to ORB which belongs to the prefrontal cortex. Even AI (agranular insula) and MoP

can be seen in the same cluster although there is no specific connections related to this in the literature. They have connections in general terms but functionally do not belong to the same cluster in the literature. In the broad sense, it can be analysed that the sub-cortical regions ACB and the LSX_MSC_TRS communicate with the cortical regions ORB, MOp, AI. The regions implied by the olfactory, somato-sensorial and motor regions exist in the block SSp, SSs, RSP, PIR (piriform), MOs and CP. This is coherent with other studies although the motor regions are less evident in the literature. RSP is the principal seed in the DMN (default mode network), which is a cluster of regions having highly correlated activity. It appears that RSP is in the same cluster as other regions from the DMN such as ACA (Anterior cingulate area), FRP (frontal pole), PL (prelimbic area) and TH (thalamus). In the regions of the hippocampus, DG (dentate gyrus) was found to have a similar activity to the sub-regions of CA2.

Analysis of clustered regions in such a way could be difficult to interpret the changes when it comes to comparing groups as the ordering of clustered regions for each group will be different. But for the group study we are interested to find the pairs of regions that vary in activity for different groups. For this purpose, we implement the approach and the statistical tests we presented in the previous section 3.6.2 on the correlation matrices of real data. This is shown in the next section.

3.7.2 Statistical tests for groups

Timecourses were estimated for the AD as well WT mice and the correlation matrices for each were calculated. Then, the permutations test and Student's t-test were performed using the method explained in section 3.6.2. The mean correlation matrices for AD as well as WT mice are shown in Fig. 3.36 and Fig. 3.37. It can be observed visually that the differences are not significant. In Fig. 3.39 and Fig. 3.38, we can see the variances of the correlation matrices with the values thresholded between -0.5 and $+0.5$, the variances did not differ significantly in the two groups. For the Student's t-test as well as permutations test, the q-values were set to 0.05 and 0.1. It was observed that the permutations tests and Student's t-test for groups did not result in significant detections when comparing the correlation pairs between the wild-type mice and the diseased mice as a group study. One main reason is the small difference between the means of the distributions for the different groups. The difference in the means of the z-values for the selected mice was of the order of 10^{-2}. Another likely reason the group tests didn't work is the number of subjects. The number of subjects in the case of real data was equal to 6. The correlation matrices with the correlation coefficients following $\mathcal{N}(0.3, 0.02)$ for the two groups were generated. Few of the correlation coefficients in one of the groups were changed according to the law $\mathcal{N}(\text{meanDifference}, 0.02)$, where meanDifference was changed. In the other study, group tests were performed by varying the number of subjects. Fig. 3.40 shows the power of the statistical tests for permutations as well as t-test for 6 mice with the differences in means varying from 0.01 to 0.19 with an interval

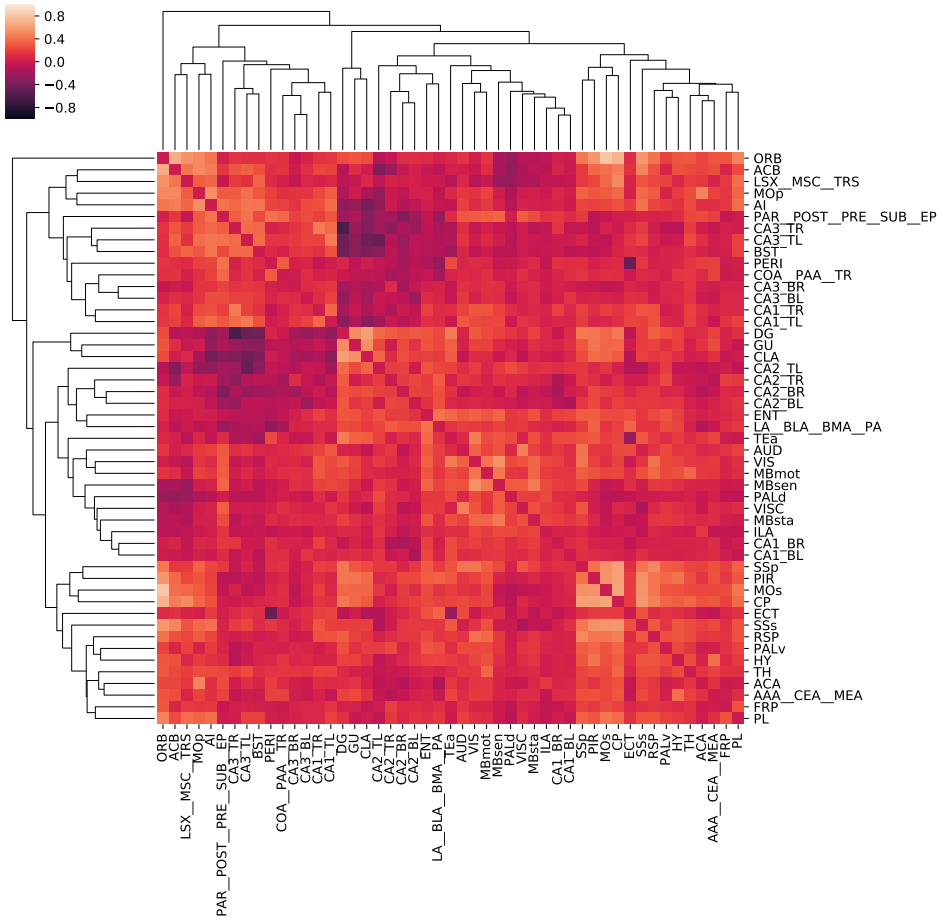


Figure 3.35: Hierarchical clustering for the means of z scores of the correlation matrices at 9 months of a WT mouse.

of 0.02. We can observe that the power of the test is not sufficient for our real data for group testing, and better tests and more data are required. The q-value for both t-tests and permutations tests was set as 0.1. It can be observed that when the mean differences are low, then the permutations tests offer better performance.

To test the influence of the number of subjects, we fix the mean difference to 0.01 as in the case for real data and consider the number of subjects as variables. The number of subjects was varied in the interval 6 to 206 with an interval of 50, and the q-value for FDR was set to 0.1. The plot for the power of the test for t-test and the permutations test is given in Fig. 3.41. The minimum number of subjects required for the power of the test to be more than 0.8 was found to be more than 106 subjects in each group for the permutations test and at least 106 to 125 subjects

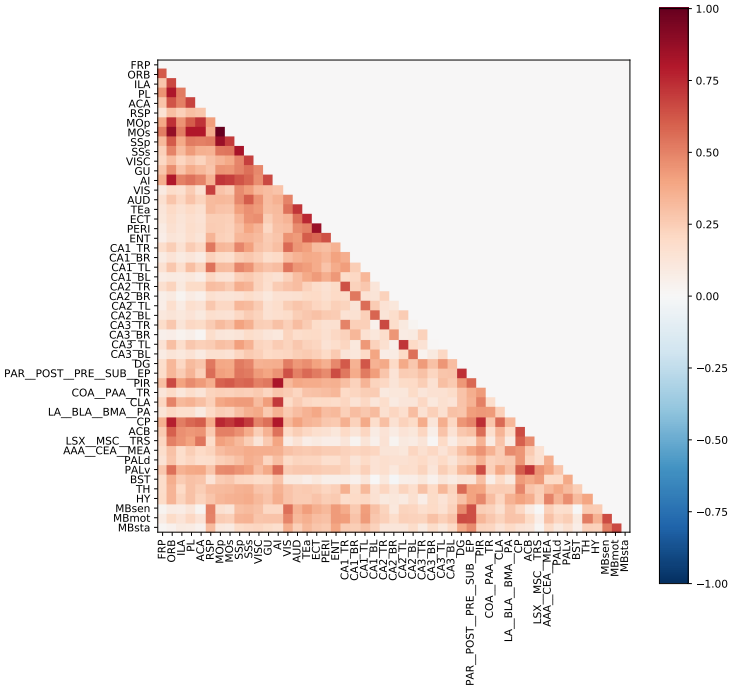


Figure 3.36: The mean correlation matrix for WT mice at 9 months.

in each group for the t-test.

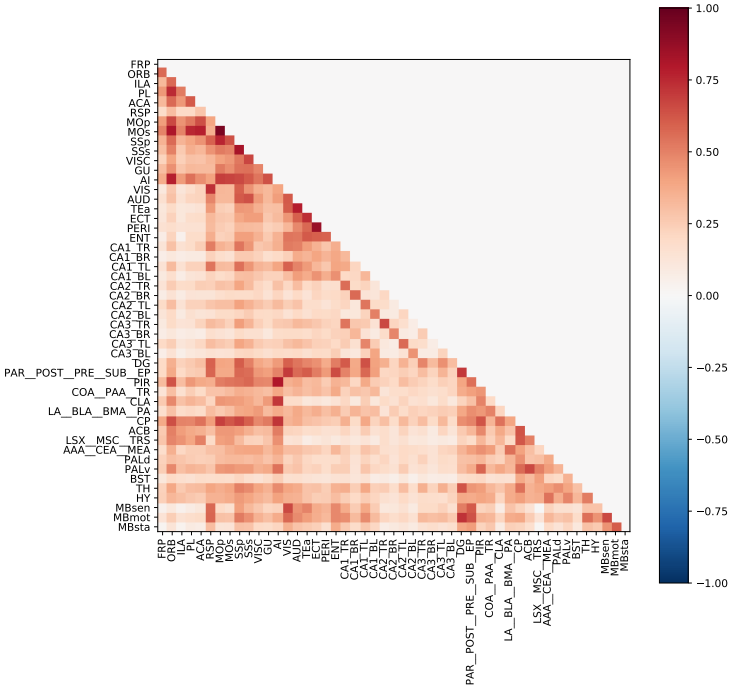


Figure 3.37: The mean correlation matrix for AD mice at 9 months.

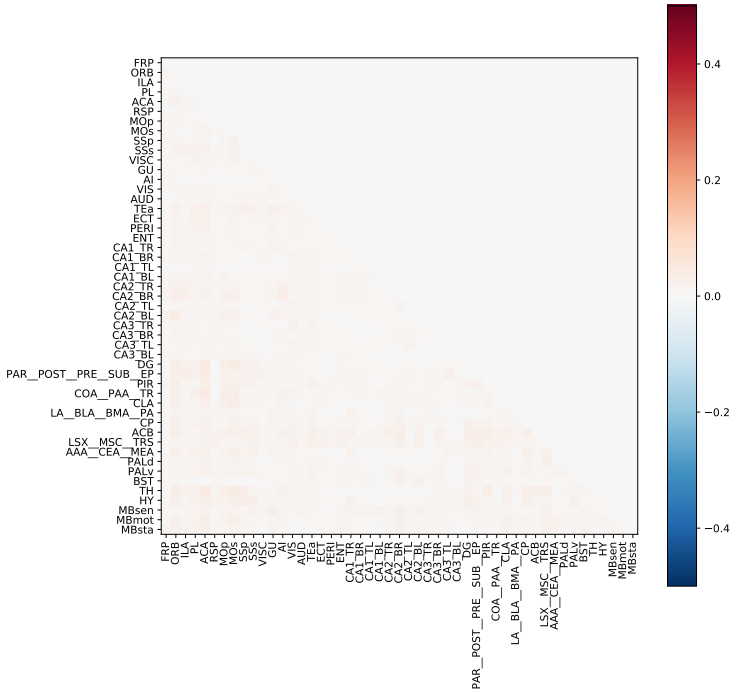


Figure 3.38: The variance of the correlation matrices for WT mice at 9 months.

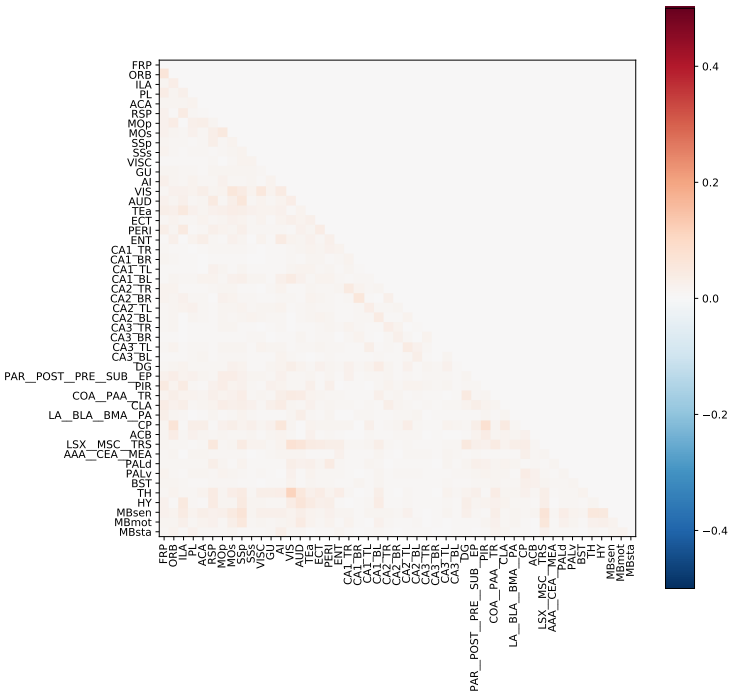


Figure 3.39: The variances of the correlation matrices for AD mice at 9 months.

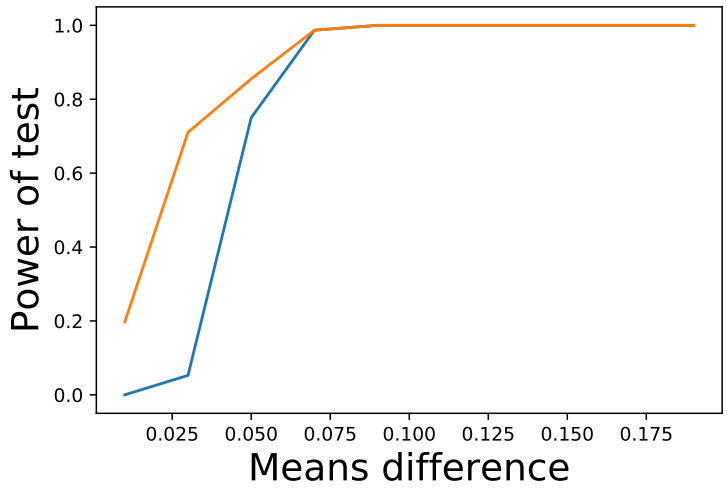


Figure 3.40: The power of the t-test and permutations test with multiple comparisons with difference in means of distributions varying from 0.01 to 0.19. The blue curve corresponds to the t-test and the orange curve to the permutations tests.

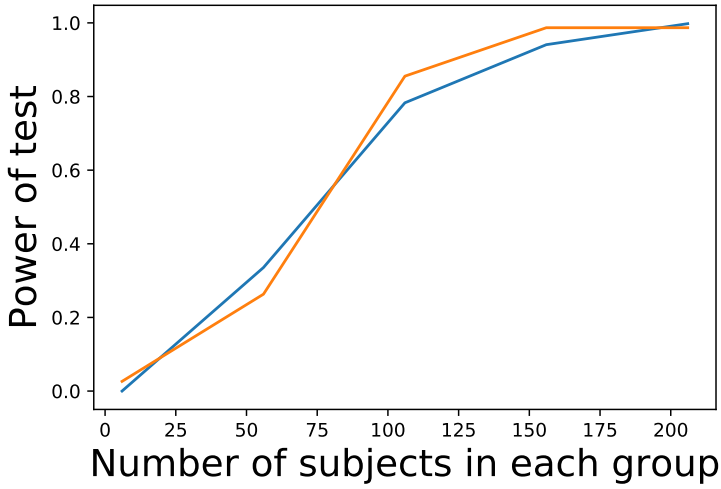


Figure 3.41: The power of the t-test and permutations test with multiple comparisons with difference in means of distributions varying from 0.01 to 0.19. The blue curve corresponds to the t-test and the orange curve to the permutations tests.

3.7.3 Longitudinal study

As explained earlier, for the longitudinal study, the Allen Brain atlas with 600 anatomical regions is used to define the regions, and the proposed unmixing algorithm is run on each mice individually. The estimated signals are used to construct correlation matrices for each mice. The correlation matrices are not shown here as the matrices are of size 600×600, and it is difficult to represent them here. After constructing the correlation matrices, the Gaussian test (as explained in 3.6.1), with the multiple comparisons is performed to look for the regions detected for each mice longitudinally. The regions detected for each mice in a longitudinal study for AD as well as WT mice are found, and then the regions that changed in longitudinal connectivity for at least 5 out of the 6 mice were kept. Then, the regions are compared between the two groups to check the regions which underwent connectivity changes in the longitudinal study AD mice but were not detected for a longitudinal study of WT mice. One of the significant results was the detection of the olfactory and entorhinal cortex, which can be validated in the literature for humans [98]. The olfactory bulb, olfactory areas, taenia tecta, dorsal peduncular area, entorhinal area all belong to the olfactory regions and can be seen together. It should be noted that the olfactory bulb is linked to Alzheimer's, but, is also subjected to artefacts in the literature. Also, the study was on Alzheimer's model mice and such results were not found in the previous study according to the team which created the same model [99]. The entorhinal area is a significant result here which is linked to the memory processes which are altered for the mice at 9 months in these mice (this has been confirmed in the behavioral tests in [97]). It is one of the first regions which are supposed to be affected by Alzheimer's. Other regions linked with the progression of AD were agranular insular areas, primary motor areas, visual areas, primary somatosensory areas, and orbital areas. Even agranular areas are connected to sensory areas as well as limbic systems related to memory and emotion. Other regions that were detected are given here in table 3.6.

Main olfactory bulb
Olfactory areas
Entorhinal area, medial part, dorsal zone, layer 1
Entorhinal area, medial part, dorsal zone, layer 2
Primary motor area, Layer 1
Primary motor area, Layer 2/3
Primary motor area, Layer 6a
Secondary motor area, layer 1
Primary somatosensory area, mouth, layer 4
Primary somatosensory area, mouth, layer 5
Primary somatosensory area, trunk, layer 2/3
agranular insular area, dorsal part, layer 5
Agranular insular area, ventral part, layer 6a

Primary visual area, layer 1
Posterolateral visual area, layer 5
Posterolateral visual area, layer 1
Anterolateral visual area, layer 1
Anterodorsal nucleus
CA2 Coronal Top Left
Crus 1
Dorsal peduncular area
Fields of Forel
Frontal pole, layer 1
Laterodorsal tegmental nucleus
Lobule III
Locus ceruleus
Medial mammillary nucleus, median part
Medulla
Nodulus (X)
Orbital area, lateral part, layer 1
Orbital area, medial part, layer 1
Pons
Substantia innominata
Taenia tecta, dorsal part

Table 3.6: Names of the regions detected to be linked with Alzheimer's.

Some of the results have been explained below:

- Motor: The Thy-Tau22 transgenic mice are found to be hyperactive, which can be due to the motor treatment being affected. Some layers of the primary motor regions have been detected as changes in table 3.6, which could be related to this behaviour. This result is still non-published but was observed in an experimental study at the IRIS (Imagerie, Robotique et Innovation en Santé) [83] platform of ICube.
- Somatosensory, visual and insula are the cortical regions that can be affected by Alzheimer's in particular somatosensory and insula.
- Locus coeruleus is known to be linked to Alzheimer's as it is involved in attention, emotions and sleep.
- Frontal pole and the orbital areas are the prefrontal region implied in decision-making and are altered in Alzheimer's.
- CA2, laterodorsal tegmental nucleus, medial mammillary nucleus, substantia innominata are not the regions known to be linked with Alzheimer's but still are linked to the memory (not strongly) and form the cholinergic system which

is linked to Alzheimer's. Fields of forel is also a region in the cholinergic system but is is a part of white matter and is maybe an artefact.

- Crus 1, Nodulus and Lobule III are the regions of the cerebellum which are usually not focussed upon for this kind of data as their connectivity is impacted by anaesthesia (given to mice during the acquisition) but are observed here.

- Pons and Medulla are the regions of the brainstem, and are relatively bigger in size than the other regions mentioned, and contain various kernels inside them. There may be certain regions inside these regions, which are affected by Alzheimer's, but this cannot be confirmed for now. Maybe a future study taking into account the sub-regions of these areas using a different atlas could be undertaken.

3.8 Discussion and conclusion

Application of the algorithm proposed in chapter 2 on fMRI data, i.e., the principal application in this thesis, has been presented in this chapter. Various spatial and temporal preprocessing techniques for fMRI have been detailed. The pipeline was adapted to the data and required the development of code for the various steps. The proposed algorithm was implemented on quasi-real data, where synthetic temporal signals were introduced in various sub-regions of the brain for validation. It was observed that the mean squared errors between the estimated temporal signals and the signals added were low. Then, the algorithm was implemented on real data at different scales of division of anatomical regions. For the real data in the first case, the fine anatomical regions were grouped together into major sub-regions of the brain, and in the second case, all the fine anatomical regions were considered. The usage of hierarchical clustering on estimated timecourse allowed to find functional networks coherent with literature.

Various statistical tests were performed on the correlation matrices constructed from the correlations between the estimated temporal signals. This was supported with a validation on the quasi-real dataset before being applied on the real dataset. To find the differences between the AD group and the WT group considering the regions grouped together, it was concluded that better tests are required. The possibility to increase the number of subjects and especially by the acquisition of mice brains with cryoscopes, will eventually improve the detections as the signals will contain less noise and, thus, stronger correlations. This was proven with some synthetic examples showing the lack of subjects and the minute difference between the means of the two groups. Both permutations test and Student's t-test with the multiple comparisons did not have enough power to perform the detections or the regions that were different in pair-connectivity (based on correlation coefficients). The approach based on the Gaussian test for longitudinal studies is interesting for single-subject

studies from an application point of view as it enabled to highlight the results validated by the neurobiologists. Certain regions were analysed to be linked to AD and have been equally provided in this chapter.

Chapter 4

Generalisation of the spatially constraint unmixing model

Contents

In this chapter, a generalisation of the unmixing model from 2 is presented. Equally, we present an extension to two new applications: scintigraphic imaging and hyperspectral astronomical data unmixing, with constraints that differ from those used in fMRI. The adaptability of the algorithm by the addition and dropping of certain constraints allowed us to deal with these applications from very different fields. The new applications involved quasi-real scintigraphic data with ground truth and the real astronomical hyperspectral data for which the approximate localisations of the sources are known. The latter type of data is extremely difficult to obtain as it requires the construction of ground truth maps. This is difficult and requires mobilisation of experts of the data, who can validate the spatial maps. Besides,

in both cases, comparisons to state-of-the-art domain-specific methods (RUDUR[19] and to expert manual evaluation [21]) have also been provided. It is observed that the spatial priors added to the unmixing model for these applications provide interesting results.

4.1 Generalised model

In the most general setting, the unmixing problem presented in chapter 2 is recast as:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) + g(\mathbf{A}) + h(\mathbf{U}), \tag{4.1}$$

where $g(\mathbf{A})$ summarises the constraints on the abundance matrix, except the localisation constraint that remains unchanged from the original model of chapter 2, and $h(\mathbf{U})$ are the temporal/spectral signature constraints.

A classical way to solve the joint estimation problem is to optimise the cost function (4.1) alternatively along $\mathbf{U}$ and $\mathbf{A}$ as presented in the algorithm:

```
Initialisation of binary matrix Ã
Initialisation of U^(0), l = 0
while STOPPING CRITERIA ≠ TRUE do
    Minimisation of problem (4.1) with respect to A
    Minimisation of problem (4.1) with respect to U
end
return U^(l+1), A^(l+1)
```

Algorithm 4: Alternated optimisation scheme of the dictionary learning algorithm to solve generic problem (4.1).

The minimisation with respect to $\mathbf{U}$ and $\mathbf{A}$ can be divided into two separate optimisation sub-problems with the specific constraints on $\mathbf{A}$ and $\mathbf{U}$.

4.1.1 Estimation of the temporal / spectral signatures matrix U

Considering that $\mathbf{A}$ is fixed, problem (4.1) becomes:

$$\min_{\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + h(\mathbf{U}). \tag{4.2}$$

In a generic case, $h(\mathbf{U})$ may be the sum of convex constraints but not necessarily differentiable. Let $h(\mathbf{U})$ be decomposed into the addition of a convex and differentiable term $h_d(\mathbf{U})$ and a convex but non differentiable term $h_{nd}(\mathbf{U})$. Problem (4.2) can be rewritten as:

$$\min_{\mathbf{U}} f_U(\mathbf{U}) + h_{nd}(\mathbf{U}), \tag{4.3}$$

where $f_U(\mathbf{U}) = \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + h_d(\mathbf{U})$ is convex. The resulting optimisation problem requires proximal gradient methods to estimate $\mathbf{U}$ such as the alternating-direction method of multipliers (ADMM)[71], projections onto convex sets (POCS) or proximal gradient descent algorithms, e.g. Fast Iterative Soft Thresholding Algorithm (FISTA)[72], depending on the form of the constraints in $h_d(\mathbf{U})$ and $h_{nd}(\mathbf{U})$. If $h_{nd}(\mathbf{U}) \neq 0$ then solution of problem (4.3) with proximal gradient descent is:

```
Initialisation of U^(0), k = 0
while STOPPING CRITERIA ≠ TRUE do
    for k ← 1 to proxsteps do
        U^(k+1) = prox_{h_nd}(U^(k) − λ∇f_U(U^(k)))
        k = k + 1;
    end
end
return U^(k+1)
```

Algorithm 5: Proximal gradient algorithm for estimation of $\mathbf{U}$, where $prox_{h_{nd}}$ is the proximal operator of h_{nd}, ∇f_U corresponds to the gradient of function f_U and λ is equal to the inverse of the Lipschitz constant of ∇f_U.

In the case of the FISTA algorithm, proximal operator and gradient are not evaluated at point $U^{(k)}$, but at an intermediate point $w^{(k)} = \mathbf{U}^{(k)} + \left(\frac{t^{(k-1)}-1}{t^{(k)}}\right)(\mathbf{U}^{(k)} - \mathbf{U}^{(k-1)})$ where expression of $t^{(k)}$ is given in [72] for increasing theoretically the convergence. It should be noted that the calculation of $prox_{h_{nd}}$ can be complicated if there are multiple constraints on h_{nd}.

4.1.2 Estimation of the abundance / mixing matrix A

Consider now that $\mathbf{U}$ is fixed, then problem (4.1) becomes:

$$\min_{\mathbf{A}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) + g(\mathbf{A}). \tag{4.4}$$

Note that if this function is separable according to the pixels/voxels $i \in \{1,\dots,\mathrm{P}\}$, it leads to:

$$\min_{\mathbf{a}_i} \frac{1}{2}\|\mathbf{y}_i - \mathbf{U}\mathbf{a}_i\|_F^2 + \mathcal{I}_{M(\tilde{\mathbf{a}}_i)}(\mathbf{a}_i) + g(\mathbf{a}_i), \tag{4.5}$$

where $\mathbf{a}_i$ is a column vector from the matrix $\mathbf{A}$. This allows to parallelise the minimisation step with respect to the matrix $\mathbf{A}$ (this step can be long if the matrix dimension is large). In that case, the minimisation according $\mathbf{A}$ can be parallelised w.r.t. the pixels to decrease the computation time. The set of all the vectors with a structure similar to $\mathbf{a}_i$ is given by $\tilde{\mathbf{a}}_i$, where $\tilde{\mathbf{a}}_i$ is a column of $\tilde{\mathbf{A}}$. In a generic case,

$g(\mathbf{A})$ may be the sum of convex constraints but not necessarily differentiable. Let the constraints $g(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ be decomposed into the addition of a convex and differentiable term $g_d(\mathbf{A})$ and a convex but non differentiable term $g_{nd}(\mathbf{A})$ that gathers non differentiable terms in $g(\mathbf{A})$ and $\mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A})$ which is also non differentiable. Problem (4.4) can be rewritten as:

$$\min_{\mathbf{A}} f_A(\mathbf{A}) + g_{nd}(\mathbf{A}), \tag{4.6}$$

where $f_A(\mathbf{A}) = \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + g_d(\mathbf{A})$.

The different differentiable and non-differentiable constraints for $\mathbf{A}$ and $\mathbf{U}$ in the model aid the estimation of the temporal/spectral signatures and the abundances. The addition of these constraints would require modification in the algorithms based on the different constraints. The generic model hereby developed is applied to different datasets in the next two sections.

4.2 Scintigraphic imaging dataset

The scintigraphic dataset is an example of realistic synthetic data in scintigraphy used in [9], for which the authors have proposed an unmixing method based on prior knowledge of the location of the regions of interest. In the specific application of scintigraphic imagery, robust unmixing of dynamic sequences using regions of interest (RUDUR) [19] is an unmixing method based on an objective function minimisation that promotes non-null abundances inside regions of interest (ROIs) while relaxing the model outside ROIs. The considered optimisation problem includes a weighted data fidelity term, the Tikhonov regularisation on the temporal signature, but no *sum-to-one* constraint. The integration of ROIs knowledge is formulated as a soft constraint based on the distance to the ROIs. This method has been compared in [19] to different ROI-based algorithms commonly used in scintigraphy such as FAMIS [20], FAROI [5], F_{PLS} [25], and another method based on a variational Bayesian approach [26]. Earlier existing ROI based unmixing methods mentioned in [9], but not compared to RUDUR, are [17], [70], [100]. Various results in [19] show that RUDUR performs better at estimating spatial maps and temporal signals than the other ROI-based algorithms. In scintigraphic imagery, our method will therefore only be compared to RUDUR. The unmixing methods in the article [19] are source separation methods. They estimate the time-activity curves (TACs) and emissions of a tracer (a radioactive element) in the different body organs.

4.2.1 Data description

In an effort to objectively evaluate the performances of our approach, we propose to test and compare our method on a physical model-based simulation of scintigraphic images (with ground truth) of scintigraphy images created for the evaluation of the

performances of the state-of-the-art RUDUR method [9]. Consequently, the ground truth for the timecourses and the abundances are at disposition. We have reused the dataset and the RUDUR code with the default parameters, as distributed by the authors of [9].

For this dataset, our method is confronted with a physical model-based simulation of scintigraphic images with ground truth. The dataset used in [9] is a dataset of scintigraphic data (with TACs and emissions of a tracer (a radioactive element) in the different body organs) that has been made available at [101]. This dataset is based on a Monte Carlo simulation of scintillation camera imaging [102]. The datacube comprises images of size 21×26, with N = 60. The dataset contains $R = 3$ regions, whose true abundance maps are shown in Fig. 4.1(a) and the associated time-activity curves in Fig. 4.1(f) shows the ground truths for the spatial maps. In the second row (b) the ROI initialisation is presented. This initialisation is not the same as in [9] as our method needs strictly greater initial ROIs than the regions which need to be unmixed.

4.2.2 Algorithm details

RUDUR algorithm has a soft constraint on the source locations, which allows the regions to unmix data even if the selected ROIs lie in the interior of the actual locations of the regions. This constraint is mentioned in (1.10). In our model, a hard constraint on the locations of the regions is used, through the regularisation term $\mathcal{I}_{M(\tilde{\mathbf{A}})}$ (see (4.1)), so initial ROIs should be strictly enclosing the regions for which we want to estimate the timecourses. To achieve this, the binary mask of ROIs used in [9] have been dilated with a 5 pixels square structuring element.

We should note that this application corresponds to an additive case of unmixing, so the sum-to-one constraint was dropped off in our algorithm. Further, as scintigraphy timecourses should be strictly positive (representing the emission of the tracer), (4.1) was changed to:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}) + \mathcal{I}_{\mathbb{R}^+}(\mathbf{U}). \tag{4.7}$$

In this problem we note:

$$\begin{aligned} f_{\mathbf{U}}(\mathbf{U}) &= \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2, \\ h_{nd}(\mathbf{U}) &= \mathcal{I}_{\mathbb{R}^+}(\mathbf{U}), \\ f_{\mathbf{A}}(\mathbf{A}) &= \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2, \\ g_{nd}(\mathbf{A}) &= \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \end{aligned}$$

Due to the addition of $\mathcal{I}_{\mathbb{R}^+}(\mathbf{U})$, the constraint of positivity on the TACs, the ridge regression given in step 4 of algorithm 3 (in chapter 2) to solve for $\mathbf{U}$ had to

be replaced by FISTA steps to estimate $\mathbf{U}$ in each alternate step. The initialisation of the algorithm was done with the help of ridge regression using the initial dilated ROIs as initialisation for matrix $\mathbf{A}$.

Initialisation of $\mathbf{A}^{(0)}, l = 0$
while *STOPPING CRITERIA* $\neq$ *TRUE* **do**
  Parallel minimisation w.r.t. the columns $\mathbf{u}_j$ of $\mathbf{U}$:
  for $\mathbf{u}_j$ *of* $\mathbf{U}$ **do**
    $\boldsymbol{\omega}_j^{(1)} = \mathbf{u}_j^{(l)}, \mathbf{c}_j^{(0)} = \mathbf{u}_j^{(l)}$
    for $k \leftarrow 1$ **to** $proxsteps_{\mathbf{u}_j}$ **do**
      $\mathbf{c}_j^{(k)} = prox_h(\boldsymbol{\omega}_j^{(k)} - \lambda_1 \nabla f_{\mathbf{U}}(\boldsymbol{\omega}_j^{(k)}))$
      $t^{(k+1)} = \frac{1+\sqrt{1+4\left(t^{(k)}\right)^2}}{2}$
      $\boldsymbol{\omega}^{(k+1)} = \mathbf{c}_j^{(k)} + \left(\frac{t^{(k)}-1}{t^{(k+1)}}\right)(\mathbf{c}_j^{(k)} - \mathbf{c}_j^{(k-1)})$
    end
    $\mathbf{u}_j^{(l+1)} = \mathbf{c}_j^{(proxsteps_{\mathbf{u}_j})}$
  end
  Parallel minimisation w.r.t. the columns $\mathbf{a}_i$ of $\mathbf{A}$:
  for $\mathbf{a}_i$ *of* $\mathbf{A}$ **do**
    $\boldsymbol{\omega}_i^{(1)} = \mathbf{a}_i^{(l)}, \mathbf{b}_i^{(0)} = \mathbf{a}_i^{(l)}$
    for $k \leftarrow 1$ **to** $proxsteps_{\mathbf{a}_i}$ **do**
      $\mathbf{b}_i^{(k)} = prox_g(\boldsymbol{\omega}_i^{(k)} - \lambda \nabla f_{\mathbf{A}}(\boldsymbol{\omega}_i^{(k)}))$
      $t^{(k+1)} = \frac{1+\sqrt{1+4\left(t^{(k)}\right)^2}}{2}$
      $\boldsymbol{\omega}_i^{(k+1)} = \mathbf{b}_i^{(k)} + \left(\frac{t^{(k)}-1}{t^{(k+1)}}\right)(\mathbf{b}_i^{(k)} - \mathbf{b}_i^{(k-1)})$
    end
    $\mathbf{a}_i^{(l+1)} = \mathbf{b}_i^{(proxsteps_{\mathbf{a}_i})}$
  end
  $l = l + 1$;
end
return $\mathbf{A}$, $\mathbf{U}$

Algorithm 6: Alternate optimisation algorithm to estimate $\mathbf{A}$ and $\mathbf{U}$ that combines three nested iterative algorithms. At each iteration l, the $\mathbf{A}$ and $\mathbf{U}$ matrices are updated. The estimation of $\mathbf{U}$ is parallelised, i.e., for a given TAC j, the index k refers to the iterations of the FISTA algorithm. In each iteration k, the calculation of the $prox_g$ is detailed in section 4.2.2. The estimation of $\mathbf{A}$ is pixel parallelised similarly.

The equation (4.7) can be divided into two sub problems described in the following paragraphs:

Minimisation of A:

As the sum-to-one condition is dropped, the proximal operator of the function g in algorithm 3 changes to:

$$prox_g(\mathbf{y}) = \underset{\mathbf{x}\in M(\tilde{\mathbf{a}})\cap R^+}{\operatorname{argmin}} \|\mathbf{x}-\mathbf{y}\|_2 = \mathcal{P}_{\mathbb{R}^+\cap M(\tilde{\mathbf{a}})}(\mathbf{y}), \tag{4.8}$$

where $\mathcal{P}$ is now the orthogonal projection operator on the set $\mathbb{R}^+ \cap M(\tilde{\mathbf{a}})$.

Minimisation of U:

The minimisation of $\mathbf{U}$ is also done with the proximal methods. For each timecourse $\mathbf{u}_j$ (j^{th} column of $\mathbf{U}$), related to the regions of interests, the steps 4 to 12 correspond to the FISTA steps. In algorithm 6, the gradient of $f(\mathbf{u}_i)$, i.e., $\nabla f(\mathbf{u}_i)$, is given by $\mathbf{U}^T(\mathbf{U}\mathbf{a}_i - \mathbf{y}_i)$. The step size λ_1 is set equal to the inverse of the Lipschitz constant of $\nabla f(\mathbf{a}_i)$, i.e., $1/L$, where $L = \|\mathbf{A}\mathbf{A}^T\|_F$.

$$prox_h(\mathbf{y}) = \underset{\mathbf{x}\in R^+}{\operatorname{argmin}} \|\mathbf{x}-\mathbf{y}\|_2 = \mathcal{P}_{\mathbb{R}^+}(\mathbf{y}), \tag{4.9}$$

4.2.3 Results and discussion

We ran our algorithm with 500 steps of alternate optimisation. At each iteration l, convergence is monitored by the optimisation gain κ_l defined as:

$$\kappa_l = \frac{\|\mathbf{Y}-\mathbf{U}^{(l)}\mathbf{A}^{(l)}\|_F - \|\mathbf{Y}-\mathbf{U}^{(l-1)}\mathbf{A}^{(l-1)}\|_F}{\|\mathbf{Y}-\mathbf{U}^{(l-1)}\mathbf{A}^{(l-1)}\|_F}, \tag{4.10}$$

which decreases to 10^{-15} at the 500th alternate step. The results of the proposed approach were compared to the DL-C-SUnSAL model defined in Chapter 2:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \|\mathbf{A}\|_1. \tag{4.11}$$

Estimated temporal signals and spatial maps were normalised by the criteria given in [9] for comparison with the provided ground truth. The normalised mean squared error (NMSE) and normalised mean absolute error (NMAE) for the estimated spatial maps and time-activity curves were calculated. In addition, we also provide spectral angle distances (SAD) (2.14) for the estimated time-activity curves. The formulae for NMAE and NMAE are given below:

$$NMSE(U_r^{GT}, U_r) = \frac{\sum_{n=1}^{N}(U_{r,n}^{GT} - U_{r,n})^2}{\sum_{n=1}^{N}(U_{r,n}^{GT})^2} \tag{4.12}$$

$$NMAE(U_r^{GT}, U_r) = \frac{\sum_{n=1}^{N} |U_{r,n}^{GT} - U_{r,n}|}{|\sum_{n=1}^{N} (U_{r,n}^{GT})|}. \tag{4.13}$$

Here U_r^{GT} is the ground truth temporal signal for r^{th} region, U_r is the estimated signal for the r^{th} region, N is the length of the temporal signal. NMAE and NMSE can be written in a similar manner to evaluate spatial maps.

The quantitative results on the synthetic scintigraphy data are given in table 4.1. We observe that the signals estimated by our method are close to those obtained by RUDUR, and DL-C-SUnSAL is slightly better for the first two ROIs but fails to estimate the TAC for ROI3. The NMSE and NMAE for spatial maps calculated using the two methods are given in table 4.2. The errors on the spatial maps were calculated by restricting the pixels of the estimated maps to the initial ROIs. Errors are generally lower in the case of DL-C-SUnSAL for the estimated spatial maps and globally well controlled for RUDUR. The methods were implemented on an Intel(R) Xeon(R) CPU E5-1620 v4 @ 3.50GHz. The calculation time for the proposed method is around 8 seconds, while for DL-SUnSAL it is around 1.5 seconds. RUDUR is the fastest of all taking around 0.7 seconds.

	ROI1	ROI2	ROI3
NMSE Proposed	0.049	0.046	0.011
NMSE RUDUR	0.031	0.047	**0.009**
NMSE DL-C-SUnSAL	**0.006**	**0.025**	0.089
NMAE Proposed	0.173	0.142	0.091
NMAE RUDUR	0.129	0.142	**0.085**
NMAE DL-C-SUnSAL	**0.070**	**0.129**	0.198
SAD Proposed	12.181	12.049	6.128
SAD RUDUR	9.561	12.362	**5.373**
SAD DL-C-SUnSAL	**4.180**	**7.150**	17.276

Table 4.1: Errors (NMAE and NMSE) and spectral angle distances (in degrees) between the estimated timecourses and the ground truth for the scintigraphy dataset. Best estimations for different cases have been highlighted in bold.

The spatial maps unmixed by our method are presented in Fig. 4.1 (c) and those estimated by RUDUR are displayed in Fig. 4.1 (d) and DL-C-SUnSAL in Fig. 4.1 (e). Visually, the maps are very similar although the errors are lower in the case of RUDUR. The last row (f) shows the TACs estimated by the proposed method (in blue), RUDUR (in magenta) and DL-C-SUnSAL. We observe in Fig. 4.1 that our solution is near the target solution (in dashed black) for all the sources and is close to the solution provided by RUDUR.

	ROI1	ROI2	ROI3
NMSE Proposed	0.189	0.084	0.139
NMSE RUDUR	0.156	0.069	**0.069**
NMSE DL-C-SUnSAL	**0.036**	**0.037**	0.341
NMAE Proposed	0.433	0.304	0.419
NMAE RUDUR	0.351	0.267	**0.343**
NMAE DL-C-SUnSAL	**0.158**	**0.196**	0.654

Table 4.2: Errors between the estimated spatial maps for the different regions and the ground truth. The errors given here were calculated by restricting the pixels of the estimated spatial maps to the initial ROIs. Best estimations for different cases have been highlighted in bold.

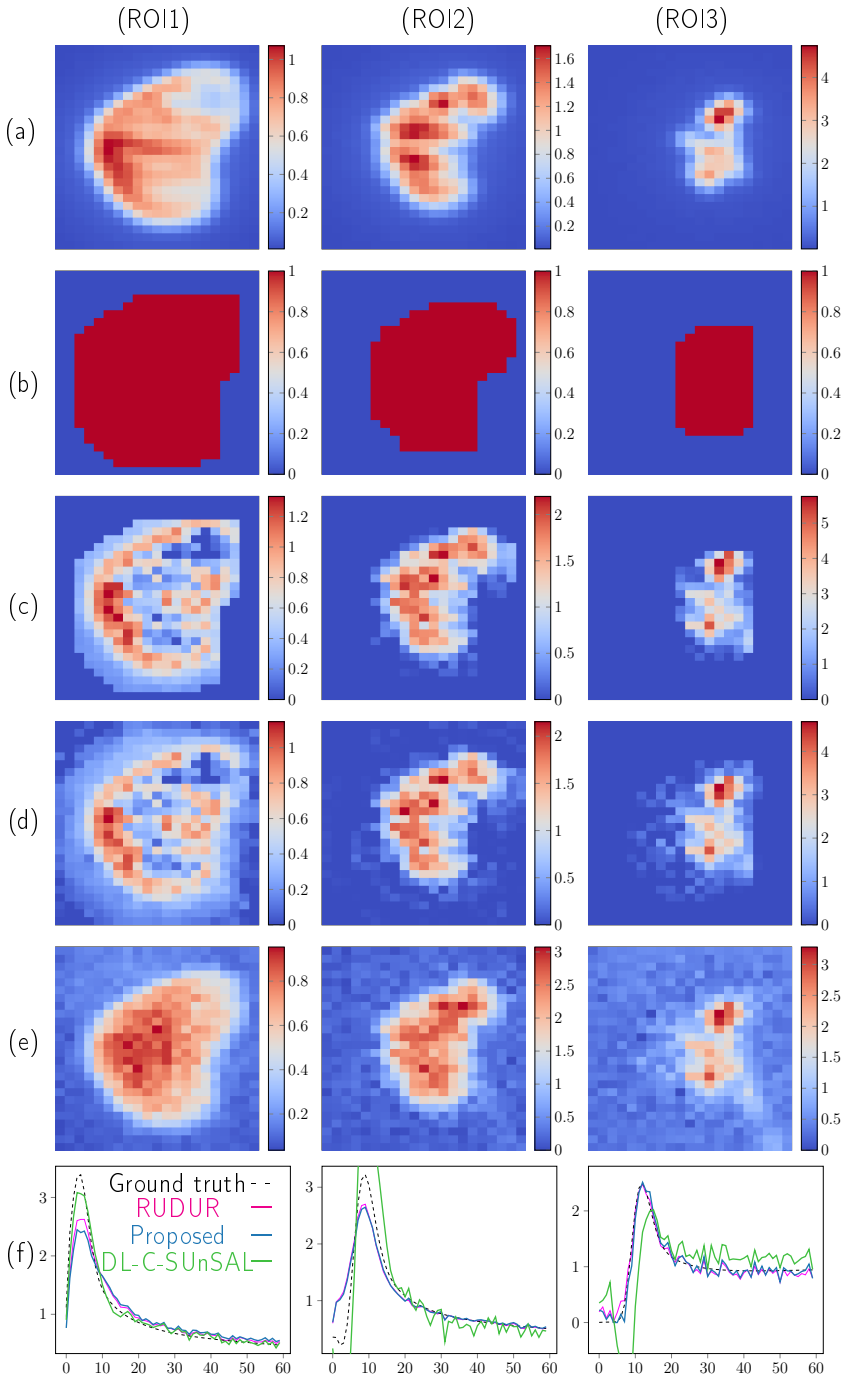


Figure 4.1: Scintigraphic data. (a) Ground truth for spatial maps, (b) initial ROIs, (c) spatial maps estimated by the proposed algorithm, (d) spatial maps estimated by RUDUR, (e) spatial maps estimated by DL-C-SUnSAL (f) TACs estimated by RUDUR and our method.

As far as the scintigraphic application is concerned, satisfactory results were obtained by the method proposed with respect to RUDUR. In the next section, we provide the application of our approach on another real dataset (the first application with real data being fMRI), i.e., hyperspectral astronomical imaging. The results were validated from the literature, and a pipeline to find new galaxies or celestial objects is suggested.

4.3 Hyperspectral astronomic data

The emergence of hyperspectral imaging has greatly benefited astronomy. Much of the details about the space not available a few decades back, due to celestial sources search in just a few bands, are available for certain portions of the sky and can be analysed now. We studied datacubes from the MUSE instrument, the Multi Unit Spectroscopic Explorer [103], installed at the Very Large Telescope, which produces hyperspectral observations of the deep sky. In these hyperspectral images, we can observe hundreds or even thousands of galaxies. Depending on their age, chemical composition, type, distance, and other physical factors, these galaxies have different spectra. These spectra may contain emission lines, continuous components, and nuisance components. One of the main objectives of MUSE data analysis is the detection of very distant galaxies, which therefore emit very low light flux. The spectrum of distant galaxies consists of a single emission line, the Lyman-α emission line, which is a marker of the strong presence of hydrogen in the galaxy. They are difficult to detect due to their distance and their very faint intensity compared to closer galaxies. In addition to that, significant noise affects the data. Moreover, two galaxies aligned in the direction of observation result in the blending of spectra inside pixels of MUSE images. The dataset recorded by the MUSE instrument explored in this work is Ultra Deep Field (UDF-10), which contains a 3D cube (2D images + spectral dimension) of data with spectra comprising of 3681 bands from 4750 to 9350 Angström ($1Ang = 0.1nm$). Much of these bands are in the visible wavelength range (3800 to 7500 Angström) and some in the near-infrared range. In order to confront our algorithm with an unmixing problem in astronomy, we consider the problem of unmixing sources in hyperspectral astronomic data with the ROI information from the Hubble space telescope. The UDF-10 field of view can be found in the Hubble Deep Field South (HDFS) data recorded with the Hubble telescope, using some of its bands. The HST observation is a spatially well-resolved image of spatial resolution 0.1×0.1 arcsec, for which there exists a segmentation map presented in [10]. Due to the difference in resolution of the MUSE data, which is only 0.7×0.7 arcsec, two distinct sources in the HST image, may overlap in the MUSE data creating a mix in the spectra. The advantage of using Hubble dataset is that the data is not containing atmospheric noise, and the images have a very high spatial resolution. On the other hand, the inconvenience is that all the galaxies are not visible because of the wide spectral bands spanning multiple wavelengths; this would lead to the

missing of galaxies with thin emission lines in the dataset. Some datasets in the astronomy domain comprise hyperspectral datacubes and external information on the spatial location of the sources. The data from Hubble was studied in a previous work [34] and a segmentation map is available.

One significant work which has focused on fusing the multimodal information from the Hubble and MUSE data to find the spectra of the galaxies is the thesis, [34]. The work is motivated by the lack of MUSE-data specific spectral unmixing method. The author of the thesis presents an algorithm that uses the field spread function (FSF) of the Hubble telescope images and carries it to the MUSE images to find the spatial localisation of the galaxies, and then using this spatial map in the MUSE data to find the spectra of galaxies. The field spread function for each of the HST bands is transferred to the MUSE data, and each is considered to be different.

Recently, two articles [104], [105] were published around the MUSE dataset called UDF-10, which corresponds to an area of the sky previously observed by the Hubble Space Telescope (HST). The proposed algorithm takes into account the a priori information present in the dataset; thus, the proposed unmixing algorithm is first run on a subcube of UDF-10 for validation in the literature and then on the whole UDF-10 cube. Equally, a pipeline is presented to find other sources in MUSE dataset with no correspondence in existing catalogues of the sky from lower spectral resolution but high spatial resolution HST images.

4.3.1 Validation on hyperspectral subcube

Data description

In [104], [105], the information provided in the Rafelski catalogue [10] is exploited to perform the deblending and prove that MUSE, despite its lower spatial resolution, allows, thanks to the spectral information, to unmix two spatially close or even superimposed sources. We have selected the same portion of the image that is presented in figure 21 of [104] where the objects identified by ID#4451, ID#4460 and ID#4465 in [10] are spatially superimposed in the MUSE observation. This gives a 25×25 pixels image with spectra composed of 3681 samples from 4750 to 9350 Angström (1 Ang = 0.1 nm). A total of 9 galaxies are present in this field of view, with three of them that are spatially close in the HST segmentation map represented at the middle of the first row in Fig. 4.2. The source ID#4465 is brighter than galaxies ID#4451 and ID#4460. Its contribution is visible on the white light image, obtained by averaging the datacube with respect to the wavelength axis. A visible source on the white light image indicates that its spectrum contains a continuous component plus, possibly, some emission lines. Contribution of source ID#4451 is embedded in the source ID#4465's. The objective of this section is to show that knowing the spatial location of such a blended source provides enough information to unmix spectra of the different superimposed sources with our algorithm.

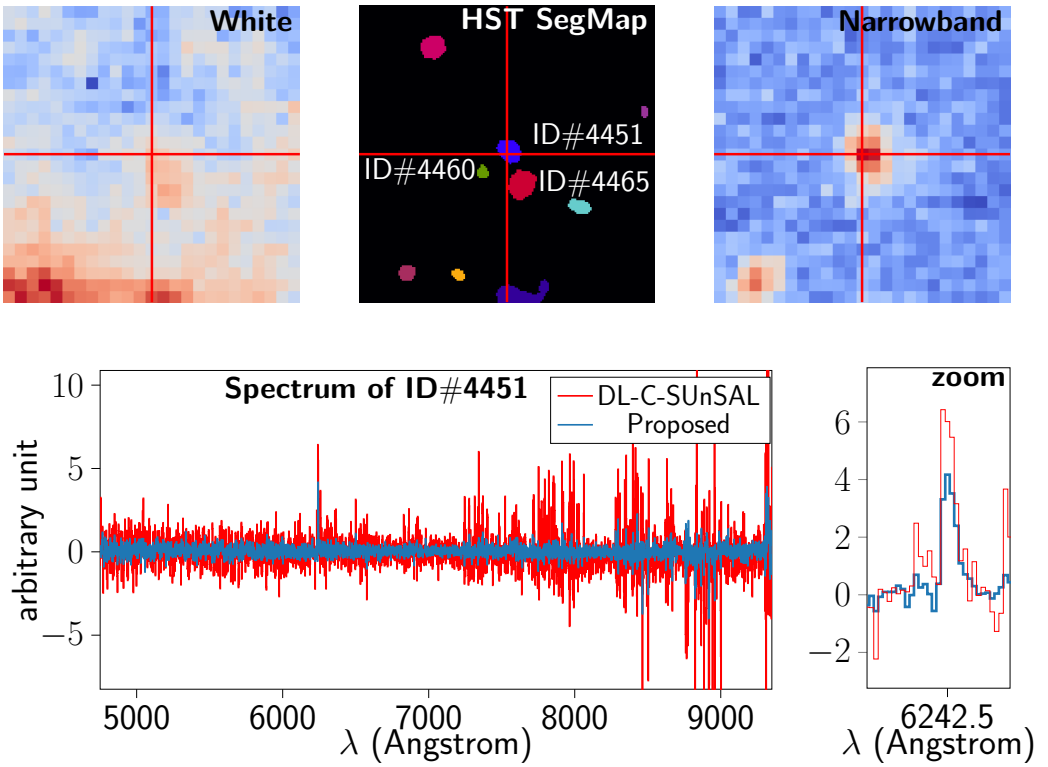


Figure 4.2: UDF Hyperspectral astronomic data. On top, from left to right, MUSE reconstructed white light image, HST Rafelski segmentation map and narrowband image centered on $\lambda = 6242.5$ Ang (position of the emission line in estimated spectrum of source ID#4451). The central Rafelski source denoted by red crosshair is ID#4451. Bottom, from left to right: estimated spectrum by the proposed method and its comparison to DL-C-SUnSAL for source ID#4451 over the whole wavelength range and zoom on the Lyα emission line estimated at $\lambda = 6242.5$.

Algorithm details

For defining the spatial constraint $\mathcal{I}_{M(\tilde{\mathbf{A}})}$ required by our model, the HST segmentation map provided by [10] is used. This map is perfectly registered on the MUSE data (4.3.1). In particular, there is no registration involved to superpose the images of HST to MUSE dataset. It is sufficient to know the equatorial coordinates in terms of right ascension (RA), and declination (DECA) angles to search for the corresponding field of view. The alignment for the data presented was done with the help of MUSE Python Data Analysis Framework (MPDAF) [106]. Then, by degrading the spatial resolution from 0.1 arcsec to 0.7 arcsec, we obtain binary masks for all the objects present in the field. The knowledge about the locations of ROIs (galaxies locations from Hubble) is more precise in this case than in the case of scintigraphic (manual region selection by an expert) or fMRI (registration of atlas) imaging. For the three central sources, the obtained binary masks are shown on the first line of Fig. 4.3.

Mixing of the galaxy spectra corresponds to an additive mixing, hence the *sum-to-one* constraint is dropped, and we note:

$$\begin{aligned} f_{\mathbf{U}}(\mathbf{U}) &= \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2, \\ h_{nd}(\mathbf{U}) &= 0, \\ f_{\mathbf{A}}(\mathbf{A}) &= \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2, \\ g_{nd}(\mathbf{A}) &= \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_{M(\tilde{\mathbf{A}})}(\mathbf{A}). \end{aligned}$$

The algorithm used for unmixing galaxy spectra is described in (7). One hundred alternated optimisation steps allow to reach a gain (4.10) equal to 10^{-10}. It should be noted that the background is considered here as a source, its mask is available in the segmentation map, and it is processed in the same way as for the galaxies to degrade its resolution to the resolution of the MUSE data.

Results and discussion

It is impossible to quantitatively compare the results obtained with ground truth for spectral signature and abundance matrix since no such information exists for the MUSE data. However, the same conclusions as in [104] can be drawn about the spectrum estimated by our algorithm for source #4451: at wavelength $\lambda = 6242.5$ Ang, there is an emission line corresponding to object #4451 of Rafelski's catalogue. This emission line has the characteristics of the Lyman-α line (Lyα), namely an asymmetric profile as illustrated in Fig. 4.2. These results are very similar to the ones presented in Figure 21 in [104] that is reproduced in Fig. 4.4 by courtesy of the authors. Similarity between results presented in Fig. 4.2 and Fig. 4.4 confirms the interest of our generic approach to solve this particular type of unmixing problem.

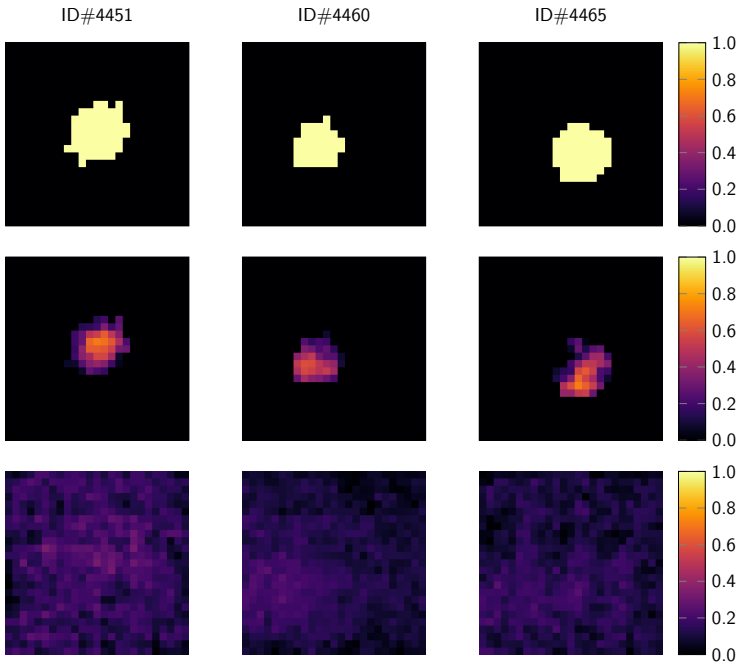


Figure 4.3: Hyperspectral astronomic data. Top row, from left to right, binary mask of sources ID#4451, ID#4460 and ID#4465. Middle row, from left to right, estimated abundance map of sources ID#4451, ID#4460 and ID#4465 by the proposed method. Bottom row, from left to right, estimated abundance maps of sources ID#4451, ID#4460 and ID#4465 by DL-C-SUnSAL method.

Initialisation of $\mathbf{A}^{(0)}, l = 0$
while *STOPPING CRITERIA* $\neq$ *TRUE* **do**
  Minimisation problem w.r.t $\mathbf{U}$:
  $\mathbf{U}^{(l+1)} = \mathbf{Y}\mathbf{A}^{(l)^T}(\mathbf{A}^{(l)}\mathbf{A}^{(l)^T} + \mu_\sigma \mathbf{I}_R)^{-1}$
  Parallel minimisation w.r.t. the columns $\mathbf{a}_i$ of $\mathbf{A}$:
  for $\mathbf{a}_i$ *of* $\mathbf{A}$ **do**
    $\boldsymbol{\omega}_i^{(1)} = \mathbf{a}_i^{(l)}, \mathbf{b}_i^{(0)} = \mathbf{a}_i^{(l)}$
    for $k \leftarrow 1$ **to** *proxsteps* **do**
      $\mathbf{b}_i^{(k)} = prox_g(\boldsymbol{\omega}_i^{(k)} - \lambda \nabla f(\boldsymbol{\omega}_i^{(k)}))$
      $t^{(k+1)} = \frac{1+\sqrt{1+4\left(t^{(k)}\right)^2}}{2}$
      $\boldsymbol{\omega}_i^{(k+1)} = \mathbf{b}_i^{(k)} + \left(\frac{t^{(k)}-1}{t^{(k+1)}}\right)(\mathbf{b}_i^{(k)} - \mathbf{b}_i^{(k-1)})$
    end
    $\mathbf{a}_i^{(l+1)} = \mathbf{b}_i^{(proxsteps)}$
  end
  $l = l + 1$;
end
return $\mathbf{A}$, $\mathbf{U}$

Algorithm 7: Alternate optimisation algorithm to estimate $\mathbf{A}$ and $\mathbf{U}$ that combines three nested iterative algorithms. At each iteration l, the $\mathbf{A}$ and $\mathbf{U}$ matrices are updated. The estimation of $\mathbf{A}$ is pixel-parallelised, i.e., for a given pixel i, the index k refers to the iterations of the FISTA algorithm.

The results were equally compared to DL-C-SUnSAL. The model for DL-SUnSAL can be written as:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2 + \mathcal{I}_{S^+}(\mathbf{A}) + \|\mathbf{A}\|_1. \tag{4.14}$$

The default parameters for [44] were used for the SUnSAL method in DL-C-SUnSAL. The DL-C-SUnSAL algorithm was tested on the dataset with the same initialisation and provided a much noisier spectrum. Even though similar conclusions can be drawn for the Lyα emission line in the estimated spectrum, its presence is not distinguishable from the other emission lines, especially at the end of the spectrum. The absence of a hard spatial constraint in DL-C-SUnSAL resulted in larger estimated spatial maps, and this can be observed in Fig. 4.3. The failure of DL-C-SUnSAL is due to the presence of a stronger noise at the end of the spectrum. The calculation time for the proposed method is around 5 seconds using the code in C++, while for using the code for DL-SUnSAL in MATLAB, it is around 50 seconds.

It should be added that by comparing the results with the scintigraphy dataset, we can see that the proposed method has a tendency to strongly remain localised

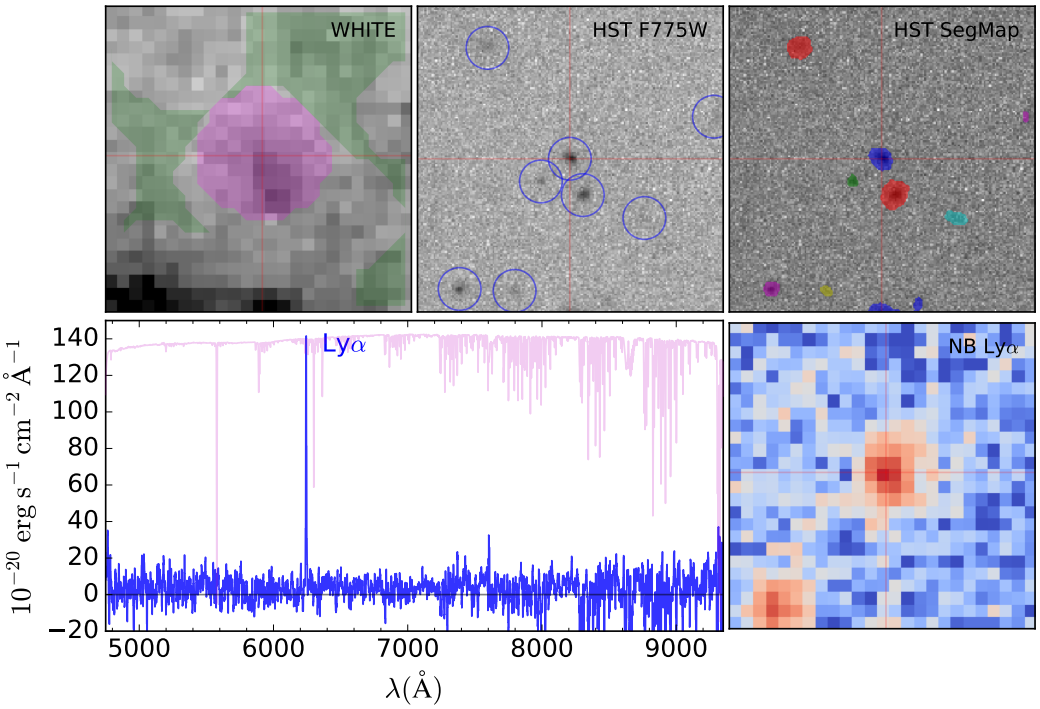


Figure 4.4: Hyperspectral astronomic data. Reproduction of figure 21 from paper [104] with the authors' permission and pending approval from A&A.

when the signal is stronger. This was observed for the scintigraphy case for the estimation of spatial maps for different ROIs. Here in the astronomy case, we can see that the non-overlapping pixels for ID #4465 have stronger intensities in the estimated spatial maps.

4.3.2 Astronomical whole cube

Decent unmixing results on the sub-cube, validation of the galaxy #4451 in the literature, and the possible scalability of the algorithm for large datasets led us to evaluate unmixing results on the whole MUSE cube at our disposition. The proposed method was implemented on the entire UDF-10 MUSE cube instead of the application on a small portion in the previous section. In a similar manner to the approach presented in the previous section on the small sub cube, the spatial maps were obtained from the Rafelski catalogue [10] having galaxies outlined from HST images. The implementation of algorithm 6 was run to estimate the spectra. It is difficult to analyse the results for the 1145 regions (1144 sources/galaxies and one additional source, i.e., the sky) for which the spectra are found, so in this section, we present the spectra of only some of the galaxies that overlap. The galaxies with significant overlaps were found by calculating the dice coefficients between the masks of the galaxies. A crop of the dice coefficients matrix has been shown in Fig. 4.5 where we can see that the galaxy #8222 overlaps with #8251 and #24692.

The proposed method was run on the entire astronomical cube after the preprocessing step of standardising the data with the estimated variance of noise provided with the cube UDF-10 in Flexible Image Transport System (FITS) format. The denoising results in a lower number of peaks due to the noise at the end of the spectrum, which are mainly due to the sky and can cause misinterpretations of peaks. These peaks could be confused as emission lines. The initialisations of the masks for these galaxies are shown in the first row of Fig. 4.6, and the estimated spatial maps are shown in the second row of Fig. 4.6. We see that the algorithm refined the spatial maps in the interior of the provided ROIs (from HST). The spectra estimated for the maps have been shown in the figure Fig. 4.7. We see that there are some emission lines that could be studied in the future with distinct emission lines appearing in the spectra estimated for galaxy #9706, whereas there is a distinct adsorption line for source #9708 near 6000 Angstrom. The spectrum of #9706 is most likely due to the presence of a large number of pure pixels for #9706. This characteristic continuous spectrum of #9706 is not observable in #9708 although the galaxies superpose.

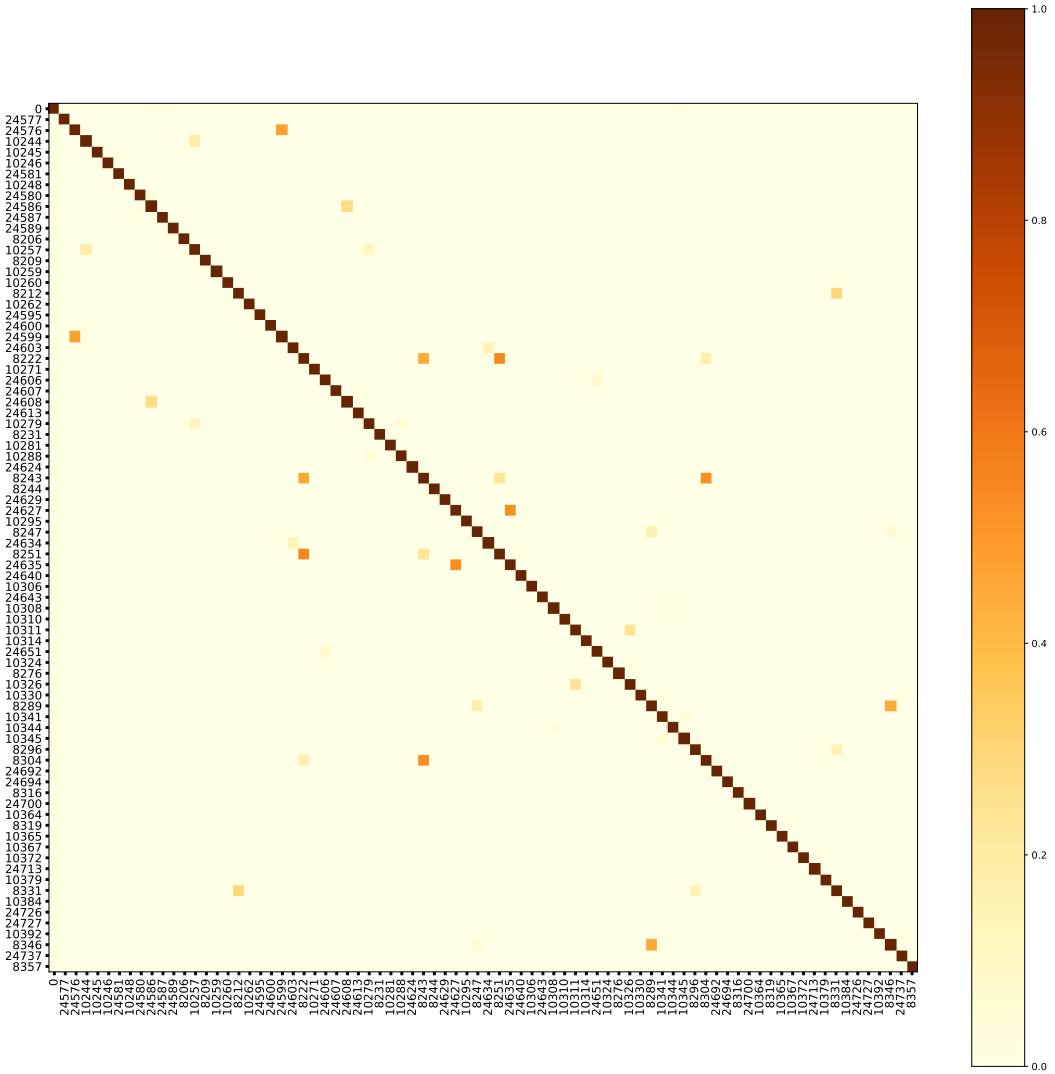


Figure 4.5: In this dice score matrix, the labels correspond to the IDs of the galaxies in the MUSE dataset. The darker is the element of the matrix, more is the overlap.

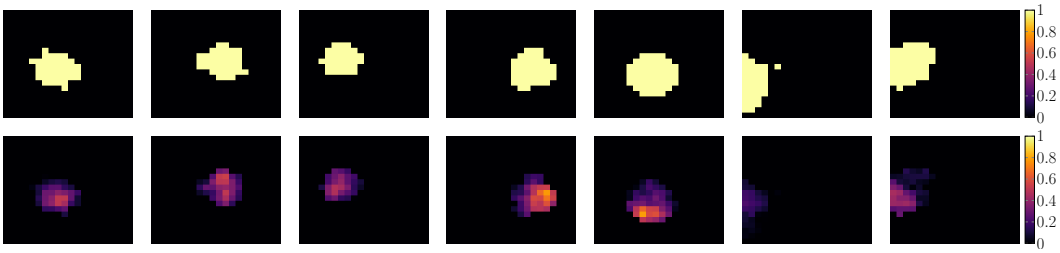


Figure 4.6: Entire UDF-10 data. On top, from left to right, binary mask of sources ID#8222, ID#8243, ID#8251, ID#8304, ID#9679, ID#9706 and ID#9708. Bottom line shows the estimated spatial maps for the same sources.

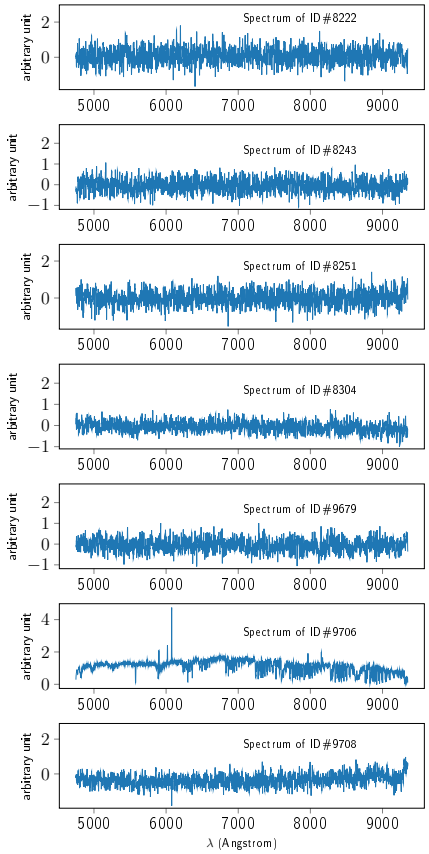


Figure 4.7: Entire UDF-10 data. The spectra estimated for #8222, ID#8243, ID#8251, ID#8304, ID#9679, ID#9706 and ID#9708 have been shown here.

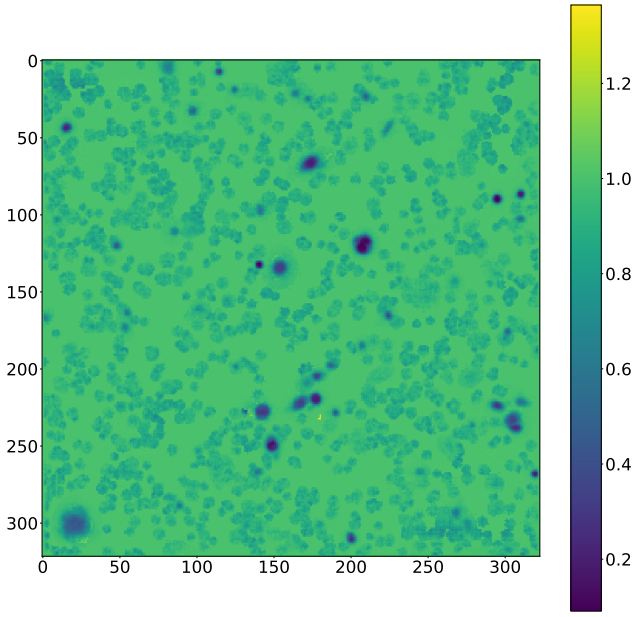


Figure 4.8: Entire UDF-10 data. Variance of the spectral residues after the unmixing.

4.3.3 Analysis of the estimation residues

The hyperspectral data may contain certain sources with faint emission spectra which were not observed in the data from the Hubble telescope and thus their contributions. Their spectra were not estimated as only the sources in the Rafelski catalogue were taken into account for unmixing. Such sources need to be found by looking at the narrow band spectra, which is time consuming and could be impossible to detect. It is thus suggested to analyse the residual variances after unmixing the cube once. This could provide information about the presence of such sources. In Fig. 4.8, the residual variance can be seen; it can be observed that the variance is lower at locations where the spatial maps were initialised. An image of the overlapping of the binary masks of the galaxies observed by Rafelski is presented in Fig. 4.9. By comparing Fig. 4.8 to Fig. 4.9 we can see that generally, residues are lower for ROIs for which the pure pixel assumption is valid. To look for the faint sources not observed by the Hubble telescope in the residue, a Gaussian PSF with a size of $3\times3\times3$ was convolved with the whole residual image ($\mathbf{Y}-\mathbf{UA}$). This operation helps in the localisation of galaxies by spreading the emission lines and the spatial map to a certain number of voxels. This also reduces the noise due to the small peaks in the spatial as well as the spectral dimension. The resultant image helped to distinguish the probable galaxies in the following step. In the next step, the maxima in the spectral dimension are displayed in the form of a 2D image (see figure Fig. 4.10); this is done to spatially locate the galaxies with faint emission lines. Instead of maxima in the whole spectral

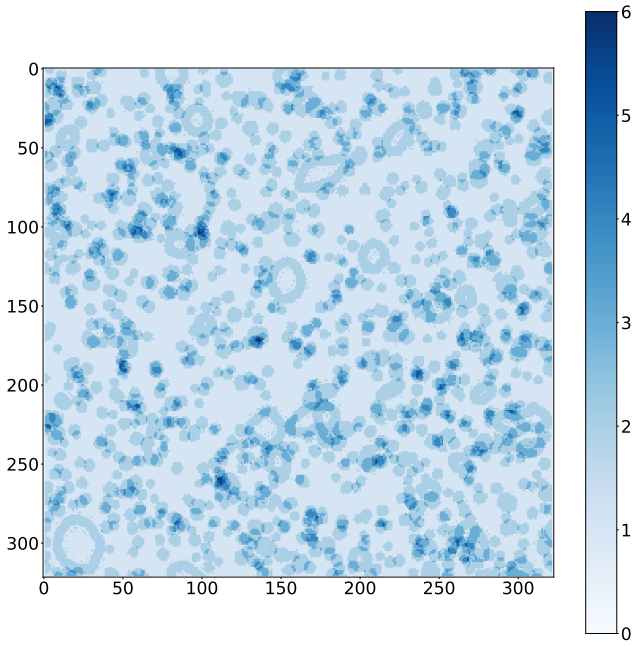


Figure 4.9: Overlapping sources of binary masks of Rafelski have been shown here. The indices on the colorbar refer to the amount of overlapping regions in the initial A matrix.

dimensions, if some *a priori* knowledge is available (presence of emission lines in a particular band), the maxima could be obtained for that specific spectral interval. The possible galaxies not existing in the initial catalogue could be located in this manner.

Once the galaxies are located, they need to be assigned labels and added to the initial spatial maps matrix. The spatial maps for the probable galaxies can be retrieved from the maxima found by a thresholding operation in Fig. 4.10. Using some geometric knowledge about the galaxies, new labels can be assigned to such structures. Then, by performing the alternate optimisation using the dictionary learning method on the data again, the refined spatial maps and spectra for the galaxies can be obtained. This procedure of hierarchical unmixing to add the located galaxies, adding them to the initial maps, and running the algorithm will improve the unmixing. In Fig. 4.11 the thresholded regions are shown. Automatic labelling of these thresholded regions can be done by object detection algorithms. The method skimage.measure.label is used for automatic labelling of the connected regions with skimage.measure.regionprops method to measure properties of the labelled regions, which is used to reject objects below a certain size and those that are not enough circular or elliptic. In Fig. 4.12 different colours represent different labels automatically assigned.

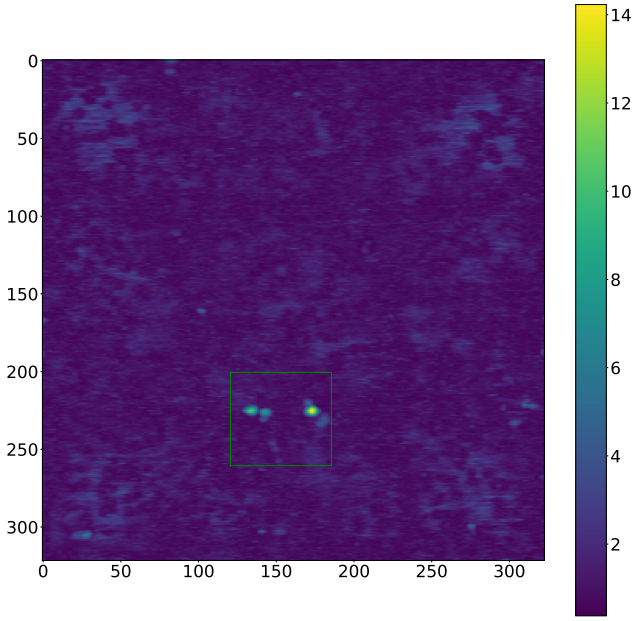


Figure 4.10: Maxima in the spectral dimension of the residue convolved with a 3 voxel cube in UDF-10 after unmixing. The colorbar represents the amplitudes of the emission lines. The galaxies in the green box were further analyed.

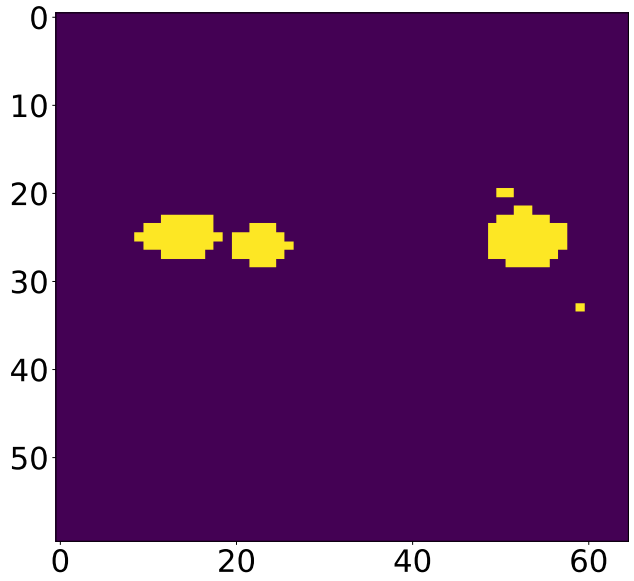


Figure 4.11: Thresholding the residual variance in a cropped section of Fig. 4.10 to find probable galaxies.

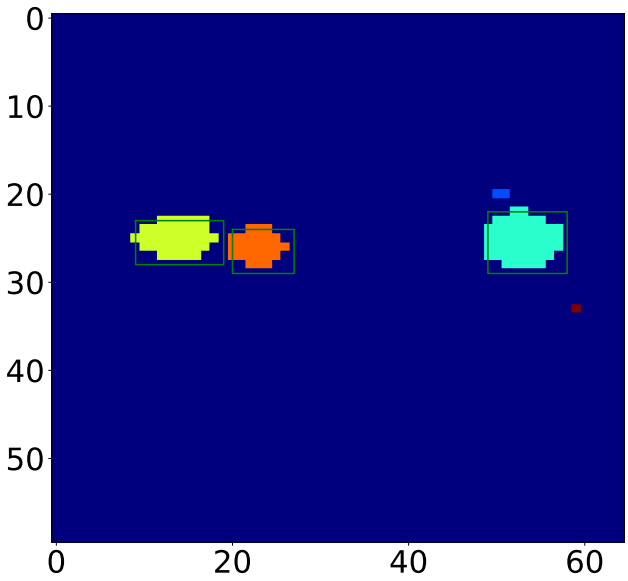


Figure 4.12: Labelling the galaxies according to some geometric criteria based on the shape or area of the galaxy (in terms of pixels square). Different colours represent different labels

The spectra for two sources on the left in Fig. 4.12 were further evaluated. A subcube considering the probable galaxies was cropped and the two galaxies were added to the other sources present in this subcube. Alternate optimisation method was run on this subcube and the results were evaluated.

It can be seen that there are some unique spectra attributed to these probable sources. By investigating the spectra, it can be observed that the emission lines present in the left source (in orange) actually corresponded to the spectra of source #24874 (in blue). This can be verified in Fig. 4.13, where we see that the spectra corresponding to the two sources have an overlap of certain emission lines. The estimated spatial maps are shown in Fig. 4.14. It was observed that the new source added may, in fact, be the source #24874 with a problem of precise correspondence in the initial catalogue because of the absence of the knowledge of the MUSE spectral bands in the HST images.

For the other source it was seen that it certainly corresponds to the core of the galaxy #24348 and it is the continuous component of the spectrum which lead to its presence in the residue.

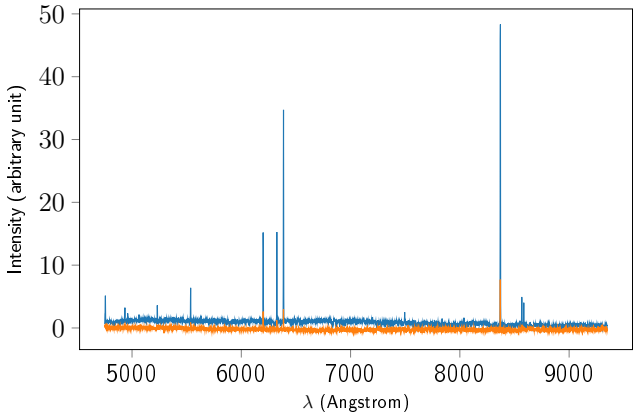


Figure 4.13: Spectra of source #24348 in blue and the spectra of the source added from the residue in orange.

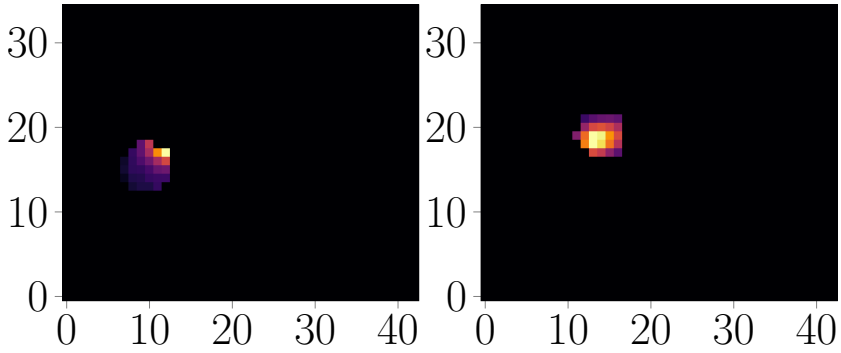


Figure 4.14: Estimated spatial maps for source #24874 and the first source that was added as a ROI from the residue.

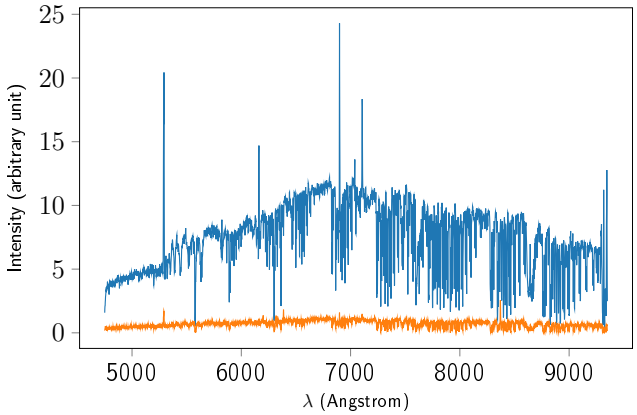


Figure 4.15: Spectra of source #24874 in orange and the spectra of the source which was investigate in blue.

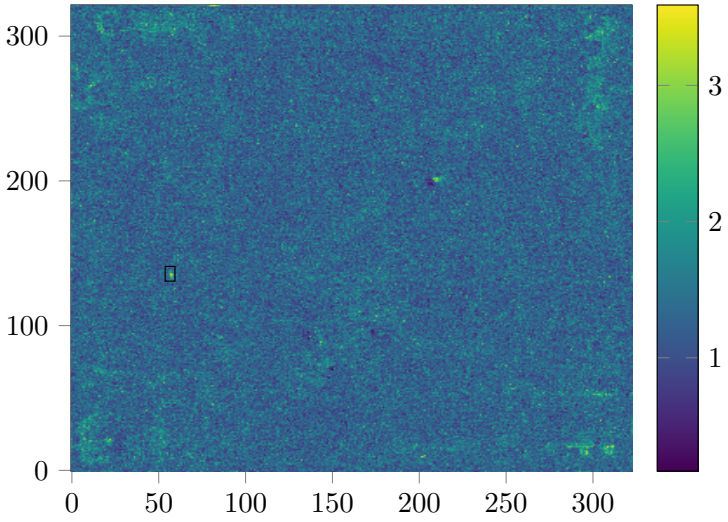


Figure 4.16: Maxima found in the convolved residue by inspecting the wavelengths by brackets of 40 samples. A probable galaxy can be seen in the black box in the figure.

4.3.4 An interesting finding

Using the pipeline provided to examine the residues in the previous section, it was tested whether it was possible to look for new sources that do not exist in the Rafelski catalogue [10]. The source that was inspected in the residue is the source 329 in the article [107] which is the same source as 6317 in the article [108]. The initialisation for the galaxy #329 and the estimated spatial map is given in Fig. 4.17. This source has an emission line at $\lambda = 7782$ that was found using the maxima in the estimated spectra. The estimated spectra and a zoom of the Lyman α line are provided in Fig. 4.18. The spectra of the same galaxy given in [107] can be observed in Fig. 4.19. It can be observed that the peak for the Lyman α line is situated at the same wavelength in both cases.

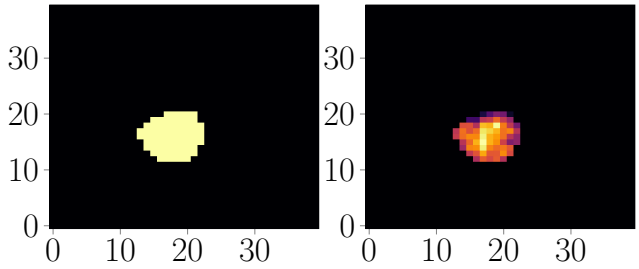


Figure 4.17: Mask extracted by thresholding the maxima of the convolved residue (on the left) and the estimated spatial map for one of the sources that does not exist in the Rafelski map and was found in [107] (on the right).

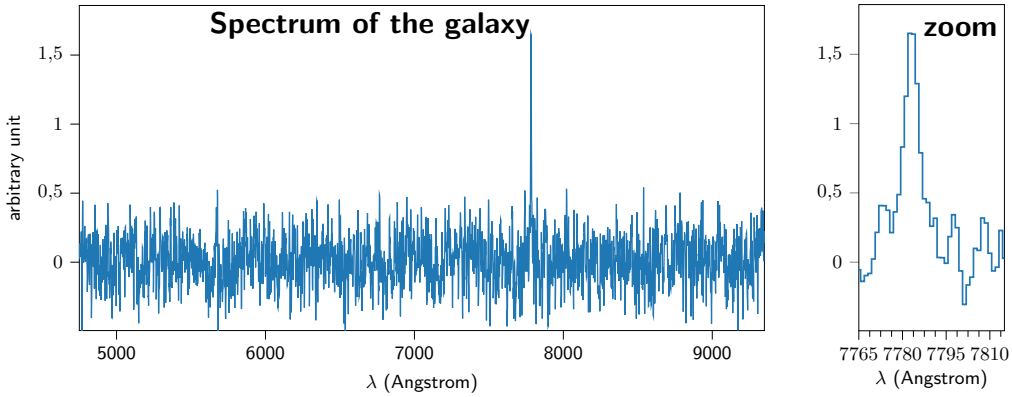


Figure 4.18: ROIs and the estimated spectrum for the source after smoothing with a boxcar of 3 pixels and zoom on frequencies showing the presence of $Ly\alpha$ spectral line.

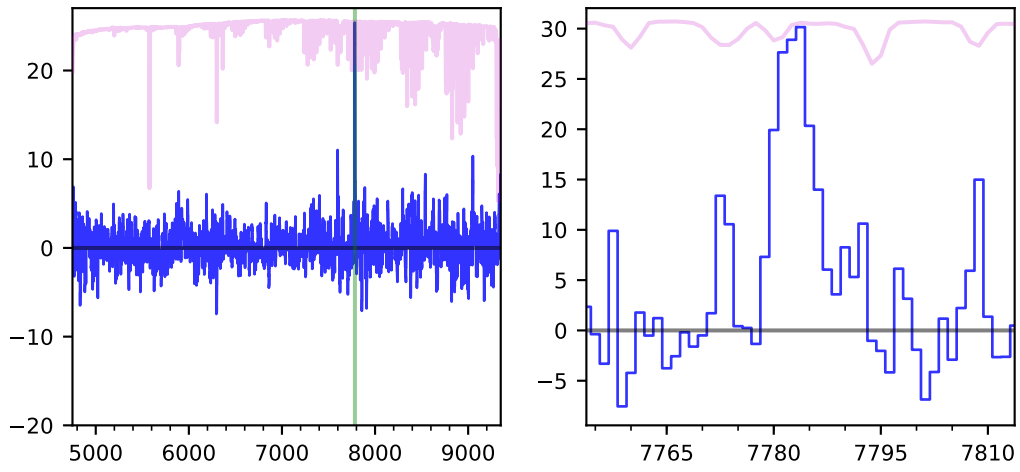


Figure 4.19: Reproduction of crop of figure 7 from paper [107]. The figure shows the spectrum of source 329 in the figure.

4.3.5 Discussion

The unmixing model has been generalised so as to make it applicable to other applications in this chapter. A region-of-interest based dataset on which the algorithm is tested is the scintigraphic dataset. Another application to which the proposed region-of-interest based approach has proven effective is an astronomical application. A dictionary learning method based on the general ℓ_1 sparsity state-of-the-art method mostly used for hyperspectral unmixing, i.e., DL-C-SUnSAL, has thus been compared to our method by respecting the constraints for the two applications.

For the quasi-real scintigraphic dataset, quantitative analysis revealed that the performance of the proposed algorithm was similar to the state-of-the-art algorithms. Better defined ROIs by experts may lead to better results using the proposed spatial constraint for real case scenarios. The results have been compared to the results from RUDUR algorithm [9] and DL-C-SUnSAL. The reason why the comparisons were not made to other ROI based algorithms is that RUDUR was the closest algorithm to the kind of problems we deal with in the article, and it performed the best out of the other state-of-the-art algorithms that have been compared to in [9].

To compare the estimation of the spectrum for the astronomical data, we did not find any method which takes into account the priors from a high-resolution segmentation map to find the spectra of different galaxies so the algorithm was compared to a semi-automatic method in the literature. The proposed approach effectively unmixed the galaxies and DL-C-SUnSAL failed in this case. It was observed that the spectrum estimated by DL-C-SUnSAL contains a lot of noise due to the contributions of non-zero abundances outside the defined boundaries of the galaxy. Even though the spectral lines are partly recovered by the DL-SUnSAL method, they are lost in the spectrum due to a lot of noise.

To locate galaxies with faint emission, the details are provided in this chapter. One of the galaxies, labelled as source 329 in the article [107] and 6317 in the article [108], was also found using this pipeline detailed in this chapter. This confirmed that the algorithm unmixes the data efficiently. It was also observed that in the residue, it is not only the noise that exists, but it is also possible to find other galaxies present in the MUSE or other future hyperspectral datasets not present in the spatial maps a priori.

Conclusion, discussion and perspectives

General conclusion

The framework of this work evolved over the course of time. Initially, the focus was on the development of an optimisation method with fMRI aspects in mind. Considering the context of fMRI, the main application was to find the difference between the AD mice and the WT mice. Under the hypothesis considered, each functional network is comprised of fine anatomical regions. An atlas with a segmentation map was considered to be used for the localisation of these small regions on the fMRI images. To perform unmixing, special constraints were added for the optimisation problem to aid unmixing. For the fMRI data, statistical studies were required to find the differences between the diseased and the control mice. In the later stage of this work, the generalisation of the unmixing model allowed us to apply it to different datasets such as scintigraphy and astronomy. The proposed spatially constraint dictionary learning algorithm was compared to the state-of-the-art for the different applications.

To check for differences in the AD and WT mice the Gaussian test on the longitudinal data provided interesting results. Regions which had significantly different changes in connectivity for AD mice but had not undergone changes for the WT mice have been unveiled. It was observed that some of these regions like the entorhinal and the olfactory regions are linked to AD in humans as well.

Contributions

The contributions in different areas are given as follows:

- The fMRI images inherently do not have any localisation information of the regions available thus, it can be difficult to quantitatively analyse the changes between the control mice and the diseased mice. To solve this purpose, a high resolution atlas with a highly detailed segmentation map (with around 600 defined anatomical regions) was registered to the fMRI. It should be noted

that the fMRI data have a really low resolution, and the projection of the atlas on the data causes the regions to overlap. This led us to consider a more sophisticated unmixing technique than the least squares method.

- Classically, an ℓ_1 constraint is used to induce sparsity in the unmixing model. For the proposed unmixing method, the classical ℓ_1 constraint for the spatial/abundance maps was replaced with a spatial constraint in the form of an indicator function. This allows unmixing respecting the boundaries of the initial spatial maps, whereas the classical ℓ_1 norm does not spatially constrain the sources. The performance was validated on a synthetic example with different cases of overlaps that can be found in real data. Once errors were evaluated for the synthetic data, the algorithm was applied to real human fMRI data from a popular fMRI analysis software, SPM. The results validated the proposed model for fMRI data. Finally, the method was applied to real mice data and interpreted by an expert neurobiologist, with outcomes compatible with what was expected for the pathologies considered in this animal model.

- The hypothesis behind using a detailed segmentation atlas was that the functional networks comprise of small anatomical regions. Classically, the fMRI data is registered to the atlas and the unmixing is performed. In this work, the atlas was registered to the fMRI data. A pipeline, different from the state-of-the-art, was developed to register the highly detailed segmentation map to the fMRI data without much altering the original temporal signals. The hypothesis and the pipeline introduced led to the development of codes that are adaptable for different datasets and were internally used in the team.

  In order to compare the brain connectivity, analysis of correlation matrices is the preferred approach in this study. For validation, quasi-real signals with known correlations were introduced in the pre-frontal regions, and the unmixing was performed. The estimated signals and the correlation matrices confirmed the good performance of the unmixing approach. Different kinds of statistical validations were performed to look for longitudinal changes on the correlation matrices of quasi-real and synthetic data. The validation of the statistical tests allowed us to perform the statistical tests on real data.

- In the last chapter, the generalisation of the model has been detailed to adapt the spatially constraint unmixing model to other applications. Results similar to state-of-the art have been provided for the scintigraphic and astronomical application. The work on these two applications required an understanding of the various modalities involved. Promising results on quasi-real data in scintigraphy provides evidence of the applicability of the algorithm on real scintigraphic data. The results could be discussed with specialists from the domain in the future. Since the proposed problem is adaptable, scintigraphy-based priors on the temporal signals could be added if required. The hyperspectral

data for the astronomical application is produced by the MUSE instrument, and the external knowledge on sources locations comes from a catalogue developed from Hubble telescope observations in [10]. The algorithm was able to provide efficient unmixing results for the whole MUSE UDF-10 dataset, thus proving its scalability. Various tools to analyse the data have been provided with a pipeline to estimate the spectra of galaxies with faint emission lines or retrieve galaxies that do not exist in the initial catalogue but are confirmed by other approaches. The generalisation of the model led to a publication in IEEE Transactions on Image Processing.

Perspectives

The perspectives are divided into model-based perspectives and application-based ones.

Model based perspectives

- Introducing spatial smoothing: The model proposed unmixing model in the manuscript performs unmixing considering the pixels independently and besides, abundance estimations are parallelised according to the pixels in the implementation of the algorithm. In order to introduce more spatial coherence within each abundance map, a constraint for spatial smoothing within the regions could be introduced in a future variant of the algorithm. However, depending upon the nature of the added constraint, it would increase the complexity of the algorithm. One way such a constraint could be added is in the form of total variation (T.V. regularisation). The spatially constraint optimisation problem taking into account the neighbouring pixels for cases involving 2D images with a temporal or spectral information can be written (in the 3D case, there would be a supplementary term $\|\mathbf{A}\mathbf{D}_z\|$). Recalling the general model:

$$\min_{\mathbf{A},\mathbf{U}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+\cap M(\tilde{\mathbf{A}})}(\mathbf{A}) + g(\mathbf{A}) + h(\mathbf{U}), \tag{4.15}$$

for which the convex optimisation problem in the framework of Total Variation (TV) problem can be formulated as:

$$\begin{aligned}
f_{\mathbf{U}}(\mathbf{U}) &= \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2,\\
h_{nd}(\mathbf{U}) &= \mathcal{I}_{\mathbb{R}^+}(\mathbf{U}),\\
f_{\mathbf{A}}(\mathbf{A}) &= \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2,\\
g_{nd}(\mathbf{A}) &= \mathcal{I}_{\mathbb{R}^+}(\mathbf{A}) + \mathcal{I}_{S^+\cap M(\tilde{\mathbf{A}})}(\mathbf{A}) + \|\mathbf{A}\mathbf{D}_x\|_1 + \|\mathbf{A}\mathbf{D}_y\|_1,
\end{aligned}$$

where the last two terms represent the directional total variation terms in the x, i.e. the horizontal direction and y, i.e. the vertical direction.

The problem can be divided into two different sub-problems to solve for $\mathbf{U}$ and $\mathbf{A}$. The sub-problem for $\mathbf{U}$ can be solved by following the proximal method explained in chapter 4. It can be observed that the estimation of $\mathbf{A}$ is more complex than the original problem as the estimation for each pixel cannot be done independently. One of the techniques that permit to resolve such a problem, by dividing the problem into multiple sub-problems defined for each of the constraints on $\mathbf{A}$, is the ADMM (alternating direction method of multipliers) approach [71]. The corresponding ADMM problem to calculate $\mathbf{A}$ where the original optimisation is divided into a set of sub-problems can be written as:

$$\min_{\mathbf{A}} \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2+\frac{\mu_\sigma}{2}\|\mathbf{U}\|_F^2+\mathcal{I}_{S^+\cap M(\tilde{\mathbf{A}})}(\mathbf{A})+\|\mathbf{A}\mathbf{D}_x\|_1+\|\mathbf{A}\mathbf{D}_y\|_1, \quad (4.16)$$

subject to:

$$\mathbf{B}=\mathbf{A}, \quad (4.17)$$

$$V_x=\mathbf{A}\mathbf{D}_x, \quad (4.18)$$

$$V_y=\mathbf{A}\mathbf{D}_y. \quad (4.19)$$

Some details to solve the problem 4.16 are given in C.

- Concerning the calculatory aspects of the proposed approach in the second chapter, certain improvements could be made to improve the speed of the estimation. A parallelised GPU implementation can vastly improve the speed and could be implemented. Projection onto the convex sets was currently done using the Michelot algorithm [11]. One of the improvements that could be done is the implementation of [12] for faster projection onto the convex sets, which could eventually reduce the calculation time.

Application based perspectives

- fMRI: In order to avoid modifying the fMRI data, the data was kept in its original resolution and an atlas was projected onto the fMRI data. This step is different from the studies in literature where the fMRI data is registered to the atlas, and then an unmixing algorithm is run over the data. Our approach involved a preprocessing pipeline which could form the basis of a future article with the addition of further technical analysis.

  A few regions in the hippocampus were divided into multiple sub-regions based on evidence of different functional activities at the interior of such regions.

This deeper scale seed analysis could equally be done for other regions for which such pieces of evidence can be made. One way this could be done is by performing ICA in the interior of a particular region and then checking the estimated spatial maps. In highly detailed segmentation atlas, at its finest resolution, certain fine anatomical regions are further divided into multiple thin layers. These layers overlap each other at the lower fMRI resolution scale and in this case, it is less appropriate to conserve such fine segmentation. The idea to fusion the anatomical regions to stay on a superior anatomical scale would improve the difficulty of the problem by reducing the dimensions of the problem of estimation (for example, by improving the conditioning of the matrix $\mathbf{A}$).

During the course of this work, the targeted fMRI applications were for resting-state mice data. Although a human task-based application is presented, future applications could involve unmixing on resting-state human fMRI data. Other datasets involving a cryoscope have low noise, and our proposed approach could be interesting for such data. Lower noise would lead to better estimated time signals and thus better correlation matrices, which may improve the clustering of functional regions into functional networks. Either more powerful statistical tests could be studied and eventually performed to find the difference between the groups, or group studies on data with more subjects could be performed.

- Astronomy: Various results corroborate that the proposed method can be applied for unmixing hyperspectral data where a segmentation map of the galaxies is available. Theoretically, the proposed method should provide better-estimated spectra than the spectra estimated by manual selection by experts where the signals are contaminated with the signals from other sources. The analysis of residues after unmixing of the known sources thanks to the [10] catalogue enabled to find a new galaxy, also detected by other methods developed in the framework of the analysis of the data for project MUSE. The entire MUSE cube must now be explored to search for other sources in the residues and running the algorithm again by adding the new sources in the initial maps in order to refine the results. The pipeline to find galaxies with faint emission lines from the residues could be used to validate other galaxies after discussion with astronomers.

Appendix A

Hypothesis testing theory

Hypothesis testing is useful to test whether a statistical hypothesis is valid for a given observation x (in scalar or vector form). This is done by defining a test, which is a function T of the observation, which would enable us to accept or reject the considered hypothesis by comparing it with a threshold of decision. The first part section A.1 is about statistical tests to check whether the results of the test are significant or not. Both parametric tests and non-parametric test are considered here. The second part section A.2 is about multiple tests where the tests are no longer binary, and the acceptation and rejection of an observation are defined taking into account all observations.

A.1 Binary tests

A hypothesis is an ideal representation of what is expected in the form of a statement for a given data. In statistical terms, the hypothesis would either fall under a null distribution ($\mathcal{H}_0$) or the alternative distribution ($\mathcal{H}_1$). If a binary decision needs to be taken to check whether the observations fall under $\mathcal{H}_0$ or $\mathcal{H}_1$, a binary statistical test is sufficient. In this manuscript, we consider the case of correlation coefficients x from a correlation matrix. Suppose that we have two groups of mice, healthy and diseased. Let us assume that the correlation coefficients for different regions lie under $\mathcal{H}_0$ for the normal case, and the correlation coefficients the diseased case fall under $\mathcal{H}_1$.

$$\begin{cases} \mathcal{H}_0 : x \sim p(x|\mathcal{H}_0) \text{ (correlation coefficients for a normal case)} \\ \mathcal{H}_1 : x \sim p(x|\mathcal{H}_1) \text{ (correlation coefficients for an alternative case)}, \end{cases}$$

where $p(x|\mathcal{H}_i)$ is the distribution for x_i under the hypothesis $\mathcal{H}_i$. To decide whether the observations fall under the null or the alternative hypothesis, a test statistic $T(.)$ and a threshold for such a test is required. The decision rule can be

Null hypothesis / Decision	$\mathcal{H}_0$ is true	$\mathcal{H}_1$ is true
H0 is accepted	Correct decision (1 - p_{FA})	Type II error (p_M)
H0 is rejected	Type I error (p_{FA})	Correct decision (1-p_M)

Table A.1: Probabilities associated with different possible decisions

explained as:

$$T(x) \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\gtrless}} \epsilon, \tag{A.1}$$

where ϵ is the threshold that enables to control the decision. Fig. A.1 shows $T(x)$ under the different hypothesis and the representation of the probability of false alarm P_{FA} and of missing detections P_M corresponding to a threshold ϵ. The test statistic chosen could be parametric or non-parametric. These have been explained in A.1.1

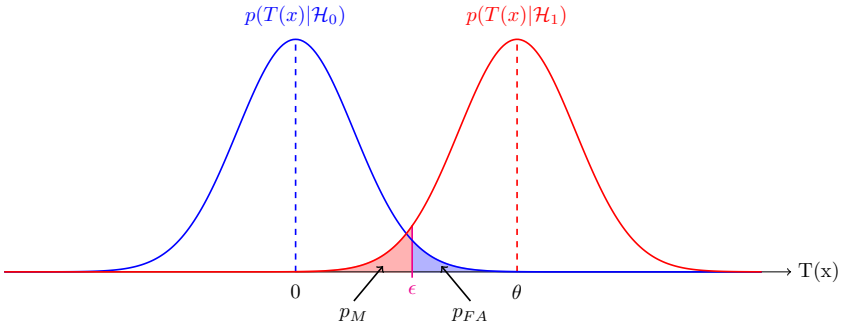


Figure A.1: Hypothesis testing, for example T(x) = $D = Z_A - Z_B$ differences between z-values of a certain parameter of two groups under the hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ (Figure inspired from [109] figure A.1)

The probability of false alarm (p_{FA}) characterises the probability $P(.)$ of rejecting $\mathcal{H}_0$ given that it is true. For our problem, it can be expressed as:

$$p_{FA} = P(T(x) > \epsilon|\mathcal{H}_0) = \int_{\epsilon}^{+\infty} p(T(x)|H_0)dx. \tag{A.2}$$

If the distribution of the test is known under $\mathcal{H}_1$ then it is also possible to calculate p_M :

$$p_M = P(T(x) < \epsilon|\mathcal{H}_1) = \int_{-\infty}^{\epsilon} p(T(x)|H_1)dx. \tag{A.3}$$

P-values

P-values are an inseparable tool in statistical analysis. The p-value associated to a test value $T(x_i)$ on the observation x_i refers to the probability of T(x) to achieve at least that extreme value for the observation x_i knowing that the hypothesis $\mathcal{H}_0$ is true. For a single-tailed test, i.e., where $\mathcal{H}_1$ is assumed to be further on the positive side, the p-value for observation x_i can be written as :

$$p_{x_i} = P(T(x) > T(x_i)|\mathcal{H}_0) = \int_{T(x_i)}^{+\infty} p((T(x)|\mathcal{H}_0))dx. \tag{A.4}$$

P-values can also be expressed as:

$$p_{x_i} = 1 - P(T(x) < T(x_i)|\mathcal{H}_0) \tag{A.5}$$

Given that $\Phi_{\mathcal{H}_0}$ is the CDF of the test statistic under $\mathcal{H}_0$ then the last equation can be written as:

$$p_{x_i} = 1 - \Phi_{\mathcal{H}_0}(T(x_i)). \tag{A.6}$$

Being a probability, p-value is a variable belonging to the interval [0,1]. The p-values obtained are considered as a transformation of the test-statistic. The characteristics of this random variable can be summarised as:
- p_{x_i} is distributed according to the uniform law $\mathcal{U}([0,1])$ under $\mathcal{H}_0$
- p_{x_i} is stochastically smaller than $\mathcal{U}([0,1])$ if $x_i \sim \mathcal{H}_1$, i.e. $Pr(p_{x_i} < t) > t$ for all $t \in [0,1]$, and independently of the initial distribution of $T(x_i)$ (which means that if $x_i \sim \mathcal{H}_1$, the distribution of p_{x_i} has a positive skew).

Example

For the binary hypothesis test if $T(x_i)$ under $\mathcal{H}_0$ follows a zeo-mean normal distribution, Fig. A.2 illustrates the calculation of p-value associated with the observations x_i. The p-value is denoted by:

$$p_{x_i} = 1 - \Phi\left(\frac{x_i}{\sigma}\right), \tag{A.7}$$

where Φ is the CDF of a standard normal distribution.

A.1.1 Parametric tests

Such tests are used when the distributions of the test statistic are known. There are different kinds of parametric tests in the literature, such as the Student's t-test, chi-squared test [110], [111], Pearson correlation, Gaussian test. Here, Student's test for groups was used and is explained here.

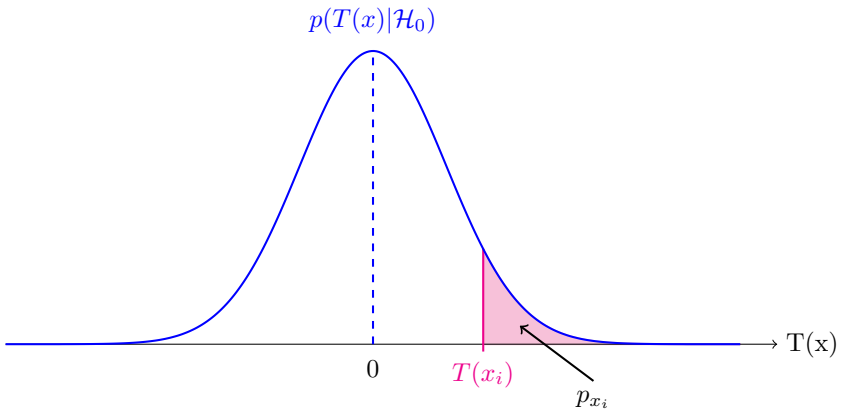


Figure A.2: Calculation of p-values (Figure inspired from [109] figure A.2)

A.1.2 Student's t-test

This test looks for differences in the means of two populations with unknown standard deviations. The family of Student's distribution depends on the degrees of freedom, which are equal to the sum of the total number of samples in each group minus two. Variants of the Student's test exist: e.g., for samples greater than 30, the difference of means follows a normal distribution. When there are not many samples (as in this study), the Student's test adapted for small-sample size could be performed. The test statistic t for a small sample size with two populations of size N_1 and N_2 having equal variances is given by:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{S^2\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}}, \tag{A.8}$$

with

$$S^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}. \tag{A.9}$$

In (A.8) $\bar{X}_1$ and $\bar{X}_2$ are the arithmetic means of the two groups.

The variances can be verified by Fisher's test for homoscedasticity [112]. A general rule of thumb is to perform a test for a small sample size when N_1 and N_2 are less than or equal to 30. The statistical distribution, in this case, is the Student's distribution with a degree of freedom equal to $N_1 + N_2 - 2$.

A.1.3 Non-parametric tests

Non-parametric tests are useful when the null distribution cannot be parametrically expressed. Bootstrapping, permutations testing, Wilcoxon signed-rank test [113] are some non-parametric tests. Permutations testing for groups has been explained as follows.

Permutations testing for groups

Suppose we have two groups with N_1 elements in the first group and N_2 elements in the second group. To perform the permutations test, first of all, a mix is generated with elements from the two groups pooled together, i.e., two groups are generated by randomly selecting N_1 elements without replacement from the mix and then selecting N_2 elements and putting them in group 2. The null distribution is estimated by calculating the metric chosen for the two generated groups many times. Permutations testing is used in order to test the significance of a test statistic. This is different from the using p-values (refer to section A.1), which are interesting but do not provide any insight into the significance. The advantage of permutations testing is that the test statistic could be any metric. Another advantage is to generate the null distribution empirically if no knowledge about the statistics is available. Compute the test statistic by finding the value of the metric used for the comparison of two populations to be tested. The procedure considering the metric as differences can be explained as follows: Repeat step 1 and step 2 'B' number of times.

1. Group the populations together and, without replacement, randomly choose individuals to form synthetic populations of the original group sizes.
2. Compute the histogram of the 'B' values obtained.
3. Estimate the p-value of the original test value obtained on the population.

Ideally, while performing the first step, each new group generated should be unique. This would guarantee the performance of the test.

A.2 Introduction to multiple comparisons/multiple testing problem

Given N observations $x = [x_1, ..., x_N]$, each associated with a system of binary hypotheses:

$$\mathcal{H}_0^i : x_i \sim p(x_i|\mathcal{H}_0^i) \tag{A.10}$$

$$\mathcal{H}_1^i : x_i \sim p(x_i|\mathcal{H}_1^i) \tag{A.11}$$

We consider the case where the hypothesis $\mathcal{H}_0^i$ (respectively $\mathcal{H}_1^i$) are identical for the N observations, so $\mathcal{H}_0^i = \mathcal{H}_0$ (respectively $\mathcal{H}_1^i = \mathcal{H}_1$). In multiple comparisons, N

Decision \ Null hypothesis	H0 is true	H1 is true	Total
H0 is accepted	$N_0 - a$	$N_1 - b$	$N_0 + N_1 - R$
H0 is rejected	a	b	R
Total	N_0	N_1	N

Table A.2: Decision table associated with N tests

tests corresponding to N observations are required; table A.2 represents the possible outcomes. The number N_0 (respectively N_1) represents the number of cases that fall under hypothesis $\mathcal{H}_0$ (respectively $\mathcal{H}_1$). The proportion of false alarms over the whole set of N decisions is represented by the fraction a/N and the number of $R = a+b$ represents the total number of discoveries or cases for which $\mathcal{H}_0$ is rejected.

The problem in the case of N tests is formulated in a different fashion than the case of a unique test. It is possible to control each test individually by checking the probability of false alarms or missed detection; however, it may be much more interesting to have a global error control criterion for all N tests. To understand the problem, let us consider the case where the probability of false alarms is controlled individually for the N tests at a level α. The average number of false alarms over the whole set of tests will then be $N\alpha$. We can observe that if the number of tests is very large, then the number of false alarms will be large but will be independent of the number of correctly identified detection, b, falling under the hypothesis $\mathcal{H}_1$. Two main criteria for global error control in the case of multiple tests have been introduced here: control of the FWER for family-wise error rate and FDR control for false discovery rate. This section has been inspired by [109] appendix A.2.

A.2.1 Control of FWER

Different procedures exist in the literature to control FWER. It is the criteria used to control the probability of having at least one false discovery out of the N tests at a threshold α:

$$FWER = P(a \geq 1) \leq \alpha. \tag{A.12}$$

Different methods to control FWER exist, two of the classical methods have been detailed here: Bonferroni correction and Holm-Bonferroni method.

A.2.2 Bonferroni correction

[114] The Bonferroni method consists of rejecting all the cases where the p-value $p_{x_i} < \frac{\alpha}{N}$ makes it possible to maintain the $FWER \leq \alpha$. This probability can be

interpreted with the help of the Boole's inequality:

$$FWER = P\left(\bigcup_{i\in I_0}(p_{x_i} \leq \frac{\alpha}{N})\right) \leq \sum_{i\in I_0} P\left(p_{x_i} \leq \frac{\alpha}{N}\right) \tag{A.13}$$

where the set I_0 represents the set of cases where $\mathcal{H}_0$ is true. $\mathcal{H}_0$ is true when the associated p-value follows a uniform distribution in the interval [0,1] and therefore $P(p_{x_i} \leq \frac{\alpha}{N}) = \frac{\alpha}{N}$. Finally:

$$FWER \leq \sum_{i\in I_0} \frac{\alpha}{N} = N_0 \frac{\alpha}{N} \tag{A.14}$$

This procedure is conservative as the threshold depends on N, the number of tests taken into account.

In [114], the control of FWER was improved. The procedure permitted to control FWER by keeping it less than or equal to α ($FWER \leq \alpha$).

A.2.3 Controlling FDR

If a large number of observations needs to be tested, where the discoveries are very few in nature, FWER could lead to missing detection. The control of FDR was introduced by Benjamini and Hochberg [96] and avoids missing important detection. The false discoveries proportions (FDP) corresponding to N tests for which the decision table is written in table A.2, is given by :

$$FDP = \frac{a}{R}. \tag{A.15}$$

By convention, $FDP = 0$ when $R = 0$ or in other words when none of the null hypothesis is rejected. The false discovery rate, FDR, is given by:

$$FDR = E[FDP] = E\left[\frac{a}{R}\right]. \tag{A.16}$$

Controlling the FDR under a certain level q consists of keeping, on average, the FDP lower than the threshold q. This guarantees that the proportion of hypothesis rejected by hazard is on average less than q. The procedure is given in [96] permits the control of FDR in the case of N independent tests at a level $\pi_0 q$ where $\pi_0 = \frac{N_0}{N}$ is the proportion of tests absolutely under the null hypothesis and $0 \leq q \leq 1$ is the control parameter. Even if the proportion π_0 is not known, the control is always guaranteed at a level q. The procedure of Benjamini-Hoschberg is described as:

$$FDR = E[FDP] = E\left[\frac{a}{R}\right]. \tag{A.17}$$

Since FDR is a criterion less conservative than FWER, the detection procedure performed with the control of FDR will have more power. It should be noted that

all procedures that control FWER control equally the FDR but in a sub-optimal manner.

The performance can be measured on synthetic data by measuring False Discovery Proportion(FDP). Lesser is the power of the test, more is the FDP. For a simulated dataset, it can be seen that the rate of FDP approaches FDR. The algorithm of Benjamini Hochberg can be written as:

1. Arrange the p values in the ascending order.
2. Fix an FDR level q.
3. $p_{thresh} = argmax_i\{p_i < q\frac{i}{N}\}$ where i is the rank of the sorted p-values p_i.
4. All p-values lying less than p_{thresh} are considered under $\mathcal{H}_1$ or as discoveries or detections.

Example

Given N = 9000 samples generated with 8900 samples lying under $\mathcal{H}_0$ according to a Gaussian distribution $\mathcal{N}(0,1)$ and 100 samples generated with $\mathcal{N}(3,1)$ under $\mathcal{H}_1$. For each observation, the following binary hypothesis model is associated:

$$\mathcal{H}_0^i : x_i \sim \mathcal{N}(0,\sigma^2) \textit{(normal case)} \tag{A.18}$$

$$\mathcal{H}_1^i : x_i \sim \mathcal{N}(\theta,\sigma^2), \theta > 0 \textit{(alternative case)} \tag{A.19}$$

The detections or discovereies in Fig. A.3 are the samples under the values 0.05 for FWER and under qi/N for FDR.

The detections or discovereies in Fig. A.3 and the zoomed in figure Fig. A.4 are the samples under the values 0.05 for FWER and under qi/N for FDR.

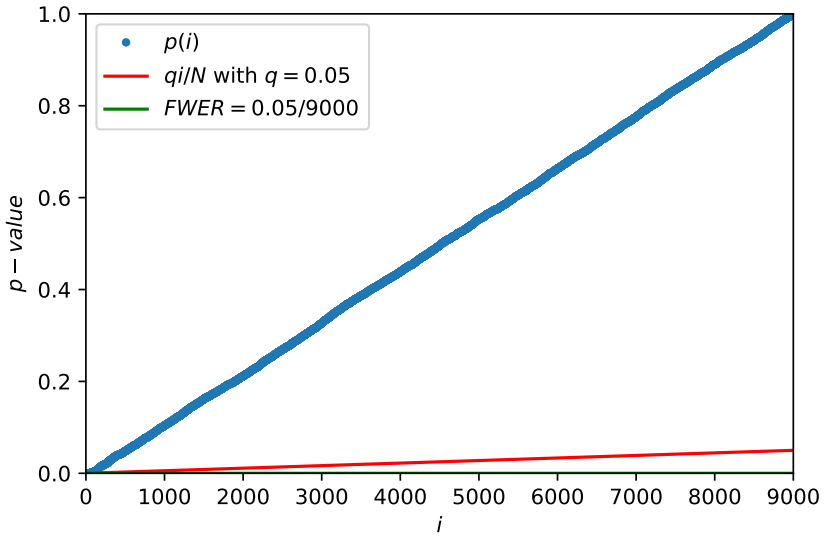


Figure A.3: Graphical representation of the thresholding of p-values with the Bonnferroni procedure with 9000 Gaussian samples. 33 detections were made for q = 0.05 whereas 8 for the case of FWER = 0.05.

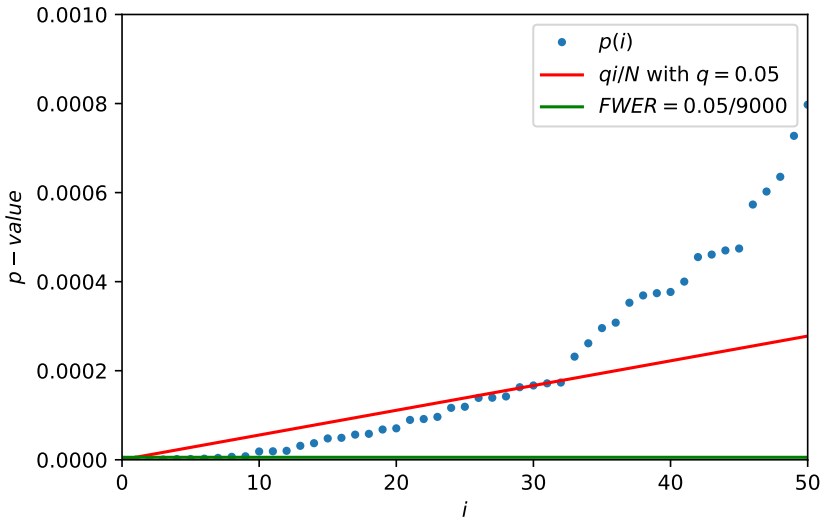


Figure A.4: Graphical representation of the thresholding of p-values with the Bonnferroni procedure showing the first 50 samples.

Appendix B

Allen atlas brain annotation details

B.1 Names of Brain regions

B.1.1 Prefrontal cortex regions

ACAd1	Anterior cingulate area, dorsal part, layer 1
ACAd2/3	Anterior cingulate area, dorsal part, layer 2/3
ACAd5	Anterior cingulate area, dorsal part, layer 5
ACAd6a	Anterior cingulate area, dorsal part, layer 6a
ACAd6b	Anterior cingulate area, dorsal part, layer 6b
ACAv1	Anterior cingulate area, ventral part, layer 1
ACAv2/3	Anterior cingulate area, ventral part, layer 2/3
ACAv5	Anterior cingulate area, ventral part, layer 5
ACAv6a	Anterior cingulate area, ventral part, 6a
ACAv6b	Anterior cingulate area, ventral part, 6b
PL1	Prelimbic area, layer 1
PL2/3	Prelimbic area, layer 2/3
PL5	Prelimbic area, layer 5
PL6a	Prelimbic area, layer 6a
PL6b	Prelimbic area, layer 6b
ILA1	Infralimbic area, layer 1
ILA2/3	Infralimbic area, layer 2/3
ILA5	Infralimbic area, layer 5
ILA6a	Infralimbic area, layer 6a
ILA6b	Infralimbic area, layer 6b
ORBl1	Orbital area, lateral part, layer 1
ORBl2/3	Orbital area, lateral part, layer 2/3
ORBl5	Orbital area, lateral part, layer 5

ORBl6a	Orbital area, lateral part, layer 6a
ORBl6b	Orbital area, lateral part, layer 6b
ORBm1	Orbital area, medial part, layer 1
ORBm2/3	Orbital area, medial part, layer 2/3
ORBm5	Orbital area, medial part, layer 5
ORBm6a	Orbital area, medial part, layer 6a
ORBvl1	Orbital area, ventrolateral part, layer 1
ORBvl2/3	Orbital area, ventrolateral part, layer 2/3
ORBvl5	Orbital area, ventrolateral part, layer 5
ORBvl6a	Orbital area, ventrolateral part, layer 6a
ORBvl6b	Orbital area, ventrolateral part, layer 6b
CA1	Field CA1
CA2	Field CA2
CA3	Field CA3
DG-mo	Dentate gyrus, molecular layer
DG-po	Dentate gyrus, polymorph layer
DG-sg	Dentate gyrus, granule cell layer
CLA	Claustrum
EPd	Endopiriform nucleus, dorsal part
EPv	Endopiriform nucleus, ventral part
LA	Lateral amygdalar nucleus
BLAa	Basolateral amygdalar nucleus, anterior part
BLAp	Basolateral amygdalar nucleus, posterior part
BLAv	Basolateral amygdalar nucleus, ventral part
BMAa	Basomedial amygdalar nucleus, anterior part
BMAp	Basomedial amygdalar nucleus, posterior part
PA	Posterior amygdalar nucleus
CP	Caudoputamen
ACB	Nucleus accumbens
FS	Fundus of striatum
OT	Olfactory tubercle
LSc	Lateral septal nucleus, caudal (caudodorsal) part
LSr	Lateral septal nucleus, rostral (rostroventral) part
LSv	Lateral septal nucleus, ventral part
SF	Septofimbrial nucleus
SH	Septohippocampal nucleus
AAA	Anterior amygdalar area
BA	Bed nucleus of the accessory olfactory tract
CEAc	Central amygdalar nucleus, capsular part
CEAl	Central amygdalar nucleus, lateral part
CEAm	Central amygdalar nucleus, medial part
IA	Intercalated amygdalar nucleus

MEA	Medial amygdalar nucleus
VAL	Ventral anterior-lateral complex of the thalamus
VM	Ventral medial nucleus of the thalamus
VPL	Ventral posterolateral nucleus of the thalamus
VPLpc	Ventral posterolateral nucleus of the thalamus, parvicellular part
VPM	Ventral posteromedial nucleus of the thalamus
VPMpc	Ventral posteromedial nucleus of the thalamus, parvicellular part
SPFm	Subparafascicular nucleus, magnocellular part
SPFp	Subparafascicular nucleus, parvicellular part
SPA	Subparafascicular area
PP	Peripeduncular nucleus
MGd	Medial geniculate complex, dorsal part
MGv	Medial geniculate complex, ventral part
MGm	Medial geniculate complex, medial part
LP	Lateral posterior nucleus of the thalamus
PO	Posterior complex of the thalamus
POL	Posterior limiting nucleus of the thalamus
SGN	Suprageniculate nucleus
AV	Anteroventral nucleus of thalamus
AMd	Anteromedial nucleus, dorsal part
AMv	Anteromedial nucleus, ventral part
AD	Anterodorsal nucleus
IAM	Interanteromedial nucleus of the thalamus
IAD	Interanterodorsal nucleus of the thalamus
LD	Lateral dorsal nucleus of thalamus
IMD	Intermediodorsal nucleus of the thalamus
MD	Mediodorsal nucleus of thalamus
SMT	Submedial nucleus of the thalamus
PR	Perireunensis nucleus
PVT	Paraventricular nucleus of the thalamus
PT	Parataenial nucleus
RE	Nucleus of reuniens
RH	Rhomboid nucleus
CM	Central medial nucleus of the thalamus
PCN	Paracentral nucleus
CL	Central lateral nucleus of the thalamus
PF	Parafascicular nucleus
RT	Reticular nucleus of the thalamus
IGL	Intergeniculate leaflet of the lateral geniculate complex
LGv	Ventral part of the lateral geniculate complex
SubG	Subgeniculate nucleus
MH	Medial habenula

LH	Lateral habenula

Table B.1: Table for acronyms and their full forms for sub-regions of the prefrontal cortex

Appendix C

Details of ADMM to introduce spatial smoothing

Rewriting the minimisation problem for $\mathbf{A}$ given in 4.16:

$$\min_{\mathbf{A}} \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+ \cap M(\tilde{\mathbf{A}})}(\mathbf{A}) + \|\mathbf{A}\mathbf{D}_x\|_1 + \|\mathbf{A}\mathbf{D}_y\|_1, \tag{C.1}$$

subject to:

$$\mathbf{B} = \mathbf{A}, \tag{C.2}$$
$$V_x = \mathbf{A}\mathbf{D}_x, \tag{C.3}$$
$$V_y = \mathbf{A}\mathbf{D}_y. \tag{C.4}$$

Posing the problem as a variant of ADMM, let us consider:

$$\mathbf{A}\Sigma + \mathbf{Z}\Lambda = 0, \tag{C.5}$$

where $\Sigma = [\mathbf{I}_p|\mathbf{A}\mathbf{V}_x|\mathbf{A}\mathbf{V}_y]$, $\mathbf{Z} = [\mathbf{B}; \mathbf{V}_x; \mathbf{V}_y]$ and $\Lambda = \mathbf{I}_{3P}$.

The augmented Lagrangian for ADMM, if Θ, Φ_x and Φ_y denote the Lagrange multipliers and ρ is the penalty parameter, can be written as:

$$\begin{aligned} L_\rho(\mathbf{A}, \mathbf{B}, \mathbf{V}_x, \mathbf{V}_y, \Theta, \Phi_x, \Phi_y) = \frac{1}{2}\|\mathbf{Y} - \mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+ \cap M(\tilde{\mathbf{A}})}(\mathbf{B}) + \|\mathbf{V}_x\|_1 + \|\mathbf{V}_y\|_1 + \\ \langle \Theta; (\mathbf{A} - \mathbf{B}) \rangle + \langle \Phi_x^T; (\mathbf{A}\mathbf{D}_x - \mathbf{V}_x) \rangle + \langle \Phi_y^T; (\mathbf{A}\mathbf{D}_y - \mathbf{V}_y) \rangle + \\ \frac{\rho}{2}(\|\mathbf{A} - \mathbf{B}\|_F^2 + \|\mathbf{A}\mathbf{D}_x - \mathbf{V}_x\|_F^2 + \|\mathbf{A}\mathbf{D}_y - \mathbf{V}_y\|_F^2), \end{aligned} \tag{C.6}$$

if

$$L_\rho(\mathbf{A},\mathbf{B},\mathbf{V}_x,\mathbf{V}_y,\Theta,\Phi_x,\Phi_y) = \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+\cap M(\tilde{\mathbf{A}})}(\mathbf{B}) + \|\mathbf{V}_x\|_1 + \|\mathbf{V}_y\|_1 + \langle[\Theta;\Phi_x;\Phi_y];\mathbf{A}\Sigma+\mathbf{Z}\Lambda\rangle + \frac{\rho}{2}(\|\mathbf{A}\Sigma+\mathbf{Z}\Lambda\|_F^2). \tag{C.7}$$

$$L_\rho(\mathbf{A},\mathbf{B},\mathbf{V}_x,\mathbf{V}_y,\Theta,\Phi_x,\Phi_y) = \frac{1}{2}\|\mathbf{Y}-\mathbf{U}\mathbf{A}\|_F^2 + \mathcal{I}_{S^+\cap M(\tilde{\mathbf{A}})}(\mathbf{B}) + \|\mathbf{V}_x\|_1 + \|\mathbf{V}_y\|_1 + \mathrm{Tr}(\Theta^T(\mathbf{A}-\mathbf{B})) + \mathrm{Tr}(\Phi_x^T(\mathbf{A}\mathbf{D}_x-\mathbf{V}_x)) + \mathrm{Tr}(\Phi_y^T(\mathbf{A}\mathbf{D}_y-\mathbf{V}_y)) + \frac{\rho}{2}(\|\mathbf{A}-\mathbf{B}\|_F^2 + \|\mathbf{A}\mathbf{D}_x-\mathbf{V}_x\|_F^2 + \|\mathbf{A}\mathbf{D}_y-\mathbf{V}_y\|_F^2). \tag{C.8}$$

C.0.1 A sub-problem

The $\mathbf{A}$ sub-problem can be written as:

$$\min_{\mathbf{A}} \frac{1}{2}\|Y-\mathbf{U}\mathbf{A}\|_F^2 + \mathrm{Tr}(\Theta^T(\mathbf{A})) + \mathrm{Tr}(\Phi_x^T(\mathbf{A}\mathbf{D}_x)) + \mathrm{Tr}(\Phi_y^T(\mathbf{A}\mathbf{D}_y)) + \frac{\rho}{2}(\|\mathbf{A}-\mathbf{B}\|_F^2 + \|\mathbf{A}\mathbf{D}_x-\mathbf{V}_x\|_F^2 + \|\mathbf{A}\mathbf{D}_y-\mathbf{V}_y\|_F^2). \tag{C.9}$$

In order to find the value of $\mathbf{A}$ that minimises C.9, the first derivative should be equal to 0 and can be written as:

$$\mathbf{U}^T\mathbf{U}\mathbf{A} - \mathbf{U}^T\mathbf{Y} + \Theta^T + \Phi_x^T\mathbf{D}_x + \Phi_y^T\mathbf{D}_y + \rho(\mathbf{A}-\mathbf{B}+\mathbf{A}\mathbf{D}_x\mathbf{D}_x^T - \mathbf{V}_x\mathbf{D}_x^T + \mathbf{A}\mathbf{D}_y\mathbf{D}_y^T - \mathbf{V}_y\mathbf{D}_y^T) = 0 \tag{C.10}$$

Rearranging the terms:

$$\mathbf{U}^T\mathbf{U}\mathbf{A} + \rho\mathbf{A} + \rho\mathbf{A}\mathbf{D}_x\mathbf{D}_x^T + \rho\mathbf{A}\mathbf{D}_y\mathbf{D}_y^T = \mathbf{U}^T\mathbf{Y} - \Theta^T - \Phi_x^T\mathbf{D}_x - \Phi_y^T\mathbf{D}_y + \rho\mathbf{B} + \rho\mathbf{V}_x\mathbf{D}_x^T + \rho\mathbf{V}_y\mathbf{D}_y^T \tag{C.11}$$

Regrouping similar terms to represent it in the form of Sylvester equation ($\mathbf{M}_1\mathbf{X} + \mathbf{X}\mathbf{M}_2 = \mathbf{C}$):

$$(\mathbf{U}^T\mathbf{U}+\rho\mathbf{I})\mathbf{A} + \mathbf{A}(\rho\mathbf{D}_x\,\mathbf{D}_x^T + \rho\mathbf{D}_y\mathbf{D}_y^T) = \mathbf{U}^T\mathbf{Y} - \Theta^T - \Phi_x^T\mathbf{D}_x - \Phi_y^T\mathbf{D}_y + \rho\mathbf{B} + \rho\mathbf{V}_x\mathbf{D}_x^T + \rho\mathbf{V}_y\mathbf{D}_y^T \tag{C.12}$$

Using the solution of the Sylvester equation:

$$\mathbf{vecA} = (\mathbf{I}\otimes(\mathbf{U}^T\mathbf{U}+\rho\mathbf{I}) + (\rho\mathbf{D}_x\,\mathbf{D}_x^T + \rho\mathbf{D}_y\mathbf{D}_y^T)\otimes\mathbf{I})^{-1}\mathbf{vec}(\mathbf{U}^T\mathbf{Y} - \Theta^T - \Phi_x^T\mathbf{D}_x - \Phi_y^T\mathbf{D}_y + \rho\mathbf{B} + \rho\mathbf{V}_x\mathbf{D}_x^T + \rho\mathbf{V}_y\mathbf{D}_y^T), \tag{C.13}$$

where $\mathbf{vecA}$ is a stack of columns of $\mathbf{A}$ and $\otimes$ is the Kroenecker product.

C.0.2 B sub-problem

The **B** subproblem can be written as:

$$\min_{\mathbf{B}} \mathcal{I}_{S^+ \cap M(\tilde{\mathbf{A}})}(\mathbf{B}) + \langle (\Theta^T; (\mathbf{B} - \mathbf{A}) \rangle + \frac{\rho}{2} \|\mathbf{B} - \mathbf{A}\|_F^2 \tag{C.14}$$

As $\min_{\mathbf{B}} \mathcal{I}_{S^+ \cap M(\tilde{\mathbf{A}})}(\mathbf{B})$ would be a simple projection, expanding the second and third term to make it suitable for the minimisation problem:

$$\min_{\mathbf{B}} \frac{\rho}{2} \operatorname{Tr}(\frac{2}{\rho} \Theta^T (\mathbf{B} - \mathbf{A}) + (\mathbf{B} - \mathbf{A})^T (\mathbf{B} - \mathbf{A})) \tag{C.15}$$

Simplifying the equation:

$$\min_{\mathbf{B}} \frac{\rho}{2} \|\mathbf{B} - \mathbf{A} + \rho^{-1} \Theta\|_F^2 - \frac{\rho}{2} \|\rho^{-1} \Theta\|_F^2 \tag{C.16}$$

As minimisation is with respect to **B**, only the first term is to be considered:

$$\min_{\mathbf{B}} \frac{\rho}{2} \|\mathbf{B} - \mathbf{A} + \rho^{-1} \Theta\|_F^2 \tag{C.17}$$

Introducing C.17 back into the original **B** subproblem.

$$\min_{\mathbf{B}} \mathcal{I}_{S^+ \cap M(\tilde{\mathbf{A}})}(\mathbf{B}) + \frac{\rho}{2} \|\mathbf{B} - \mathbf{A} + \rho^{-1} \Theta\|_F^2 \tag{C.18}$$

The minimisation is in fact an orthogonal projection given by:

$$B^{k+1} = \mathbf{P}_{\tilde{A} \cap S^+}(B^k + \rho^{-1} \Theta^k) \tag{C.19}$$

C.0.3 $\mathbf{V}_x$ sub-problem

$$\min_{\mathbf{V}_x} \|\mathbf{V}_x\|_1 + \langle \Phi_x; (\mathbf{A}\mathbf{D}_x - \mathbf{V}_x) \rangle + \frac{\rho}{2} \|\mathbf{A}\mathbf{D}_x - \mathbf{V}_x\|_F^2. \tag{C.20}$$

$$\min_{\mathbf{V}_x} \|\mathbf{V}_x\|_1 + \frac{\rho}{2} \|\mathbf{A}\mathbf{D}_x - \mathbf{V}_x + \rho^{-1} \Phi_x\|_F^2 - \frac{\rho}{2} \|\rho^{-} 1 \Phi_x\|_F^2 \tag{C.21}$$

As minimisation is with respect to $\mathbf{V}_x$, taking into account just the first two terms:

$$\min_{\mathbf{V}_x} \|\mathbf{V}_x\|_1 + \frac{\rho}{2} \|\mathbf{A}\mathbf{D}_x - \mathbf{V}_x + \rho^{-1} \Phi_x\|_F^2 \tag{C.22}$$

$$\min_{\mathbf{V}_x} \|\mathbf{V}_x\|_1 + \frac{\rho}{2} \|\mathbf{V}_x - (\mathbf{A}\mathbf{D}_x + \rho^{-1} \Phi_x)\|_F^2 \tag{C.23}$$

From the shrinkage formula for a similar term given in [115]:

$$\mathbf{V}_x^{k+1} = \max\{\|\mathbf{V}_x^k\| - \rho^{-1} \Phi_x, 0\}.sign(\mathbf{V}_x^k), \tag{C.24}$$

where *sign* is the signum function.

C.0.4 $\mathbf{V}_y$ sub-problem

$$\min_{\mathbf{V}_y} \|\mathbf{V}_y\| + \Phi_y^T(\mathbf{A}\mathbf{D}_y - \mathbf{V}_y) + \frac{\rho}{2}\|\mathbf{A}\mathbf{D}_y - \mathbf{V}_y\|_F^2 \tag{C.25}$$

Similarly to the case of $\mathbf{V}_x$:

$$\mathbf{V}_y^{k+1} = \max\{\|\mathbf{V}_y^k\| - \rho^{-1}\Phi_y, 0\}.sign(\mathbf{V}_y^k), \tag{C.26}$$

C.0.5 Θ sub-problem

$$\Theta^{k+1} = \Theta^k + \frac{\rho}{2}(\|\mathbf{A}^k - \mathbf{B}^k\|_F^2) \tag{C.27}$$

C.0.6 Φ_x sub-problem

$$\Phi_x^{k+1} = \Phi_x^k + \frac{\rho}{2}(\|\mathbf{A}^k\mathbf{D}_x - \mathbf{V}_x^k\|_F^2) \tag{C.28}$$

C.0.7 Φ_y sub-problem

$$\Phi_x^{k+1} = \Phi_y^k + \frac{\rho}{2}(\|\mathbf{A}^k\mathbf{D}_y - \mathbf{V}_y^k\|_F^2). \tag{C.29}$$

C.0.8 Structure of $\mathbf{D}_x$ and $\mathbf{D}_y$

Suppose a 2D image $\mathbb{R} \in l \times m$ contains p total pixels with $m = 3$. The example for matrices $\mathbf{D}_x$ and $\mathbf{D}_y$ for an image (with 3 pixels in each row) can be written as:

$\mathbf{D}_x$:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}}_{\mathbf{P}}$$

$\mathbf{D}_y$:

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ -1 & 0 & 0 & 1 & \dots & 0 \\ 0 & -1 & 0 & 0 & \dots & \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix}}_{\mathbf{P}}$$

Bibliography

[1] A. Bhanot, C. Meillier, F. Heitz, and L. Harsan, "Spatially constrained online dictionary learning for source separation," *IEEE Transactions on Image Processing*, vol. 30, pp. 3217–3228, 2021 (cit. on pp. xviii, xxxv).

[2] ——, "Online dictionary learning for single-subject fMRI data unmixing," in *2019 27th European Signal Processing Conference (EUSIPCO)*, IEEE, 2019, pp. 1–5 (cit. on pp. xviii, xxxv, 17, 23).

[3] A. Bhanot, C. Meillier, F. Heitz, and L. Harsan, "Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf," in *Gretsi*, 2019. [Online]. Available: `http://icube-publis.unistra.fr/5-BMHH19` (cit. on pp. xviii, xxxv).

[4] A. Bhanot, C. Meillier, F. Heitz, and L. Harsan, "Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen Mouse Brain en IRMf chez la souris," in *SFRMBM, Strasbourg, France*, 2019 (cit. on pp. xviii, xxxv).

[5] Bhanot, Argheesh and Meillier, Céline and Heitz, Fabrice and Harsan, Laura, " Online dictionary learning for single-subject fMRI data unmixing: présentation orale," in *DSAI, Strasbourg, France*, 2019 (cit. on pp. xviii, xxxv).

[6] A. Bhanot, C. Meillier, F. Heitz, and L. Harsan, "Apprentissage par dictionnaire pour le démélange de signaux temporels en IRMf; poster," in *FMTS, Strasbourg, France*, 2019 (cit. on pp. xviii, xxxv).

[7] ——, "Estimation de l'activité au repos des régions anatomiques extraites de l'atlas Allen mouse brain en IRMf chez la souris.," in *Journée poster des doctorants ICube, Strasbourg, France*, 2019 (cit. on pp. xix, xxxv).

[8] S. W. O. et al., "A mesoscale connectome of the mouse brain," *Nature*, vol. 508, pp. 207–214, 2014 (cit. on pp. xxiv, 45).

[9] M. Filippi, M. Desvignes, and E. Moisan, "Robust unmixing of dynamic sequences using regions of interest," *IEEE transactions on medical imaging*, vol. 37, no. 1, pp. 306–315, 2017 (cit. on pp. xxvii, 8, 18, 96, 97, 99, 122).

[10] M. Rafelski, H. I. Teplitz, J. P. Gardner, D. Coe, N. A. Bond, A. M. Koekemoer, N. Grogin, P. Kurczynski, E. J. McGrath, M. Bourque, *et al.*, "UVUDF: ultraviolet through near-infrared catalog and photometric redshifts of galaxies in the Hubble Ultra Deep Field," *The Astronomical Journal*, vol. 150, no. 1, p. 31, 2015 (cit. on pp. xxix, xxxii, xxxiv, 103, 104, 106, 111, 119, 125, 127).

[11] C. Michelot, "A finite algorithm for finding the projection of a point onto the canonical simplex of r^n," *Journal of Optimization Theory and Applications*, vol. 50, no. 1, pp. 195–200, 1986 (cit. on pp. xxxiii, 22, 126).

[12] L. Condat, "Fast projection onto the simplex and the l1 ball," *Mathematical Programming*, vol. 158, no. 1-2, pp. 575–585, 2016 (cit. on pp. xxxiii, 22, 126).

[13] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987 (cit. on p. 1).

[14] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal processing*, vol. 24, no. 1, pp. 1–10, 1991 (cit. on pp. 1, 2).

[15] P. Comon, "Independent component analysis, a new concept?" *Signal processing*, vol. 36, no. 3, pp. 287–314, 1994 (cit. on pp. 1, 2).

[16] J. M. Bioucas-Dias, "A variable splitting augmented Lagrangian approach to linear spectral unmixing," in *2009 First workshop on hyperspectral image and signal processing: Evolution in remote sensing*, IEEE, 2009, pp. 1–4 (cit. on pp. 1, 10).

[17] V. Šmídl and O. Tichý, "Automatic regions of interest in factor analysis for dynamic medical imaging," in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, IEEE, 2012, pp. 158–161 (cit. on pp. 1, 96).

[18] K. Pearson, "Liii. on lines and planes of closest fit to systems of points in space," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 2, no. 11, pp. 559–572, 1901 (cit. on p. 1).

[19] H. Hotelling, "Analysis of a complex of statistical variables into principal components.," *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933 (cit. on p. 1).

[20] J.-P. Roll, "Contribution à la proprioception musculaire, à la perception et au contrôle du mouvement chez l'homme," PhD thesis, 1981 (cit. on p. 1).

[21] M. S. Lewicki and T. J. Sejnowski, "Learning overcomplete representations," *Neural computation*, vol. 12, no. 2, pp. 337–365, 2000 (cit. on p. 1).

[22] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural computation*, vol. 13, no. 4, pp. 863–882, 2001 (cit. on pp. 1, 2).

[23] C. Jutten and J Hérault, "Une solution neuromimétique au problème de séparation de sources," *Traitement du signal*, vol. 5, no. 6, pp. 389–403, 1988 (cit. on p. 2).

[24] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010, p. 642 (cit. on p. 2).

[25] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE transactions on Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999 (cit. on p. 2).

[26] M. J. McKeown, T.-P. Jung, S. Makeig, G. Brown, S. S. Kindermann, T.-W. Lee, and T. J. Sejnowski, "Spatially independent activity patterns in functional MRI data during the Stroop color-naming task," *Proceedings of the National Academy of Sciences*, vol. 95, no. 3, pp. 803–810, 1998, ISSN: 0027-8424. DOI: 10.1073/pnas.95.3.803. eprint: https://www.pnas.org/content/95/3/803.full.pdf. [Online]. Available: https://www.pnas.org/content/95/3/803 (cit. on p. 2).

[27] J. Xu, M. Potenza, and V. Calhoun, "Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses," *Frontiers in Neuroscience*, vol. 7, p. 154, 2013, ISSN: 1662-453X. DOI: 10.3389/fnins.2013.00154. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2013.00154 (cit. on p. 2).

[28] M. Sourty, L. Thoraval, D. Roquet, J.-P. Armspach, and J. Foucher, "Towards an automated selection of spontaneous co-activity maps in functional magnetic resonance imaging," in *Medical Imaging 2015: Biomedical Applications in Molecular, Structural, and Functional Imaging*, B. Gimi and R. C. Molthen, Eds., International Society for Optics and Photonics, vol. 9417, SPIE, 2015, pp. 144 –151. DOI: 10.1117/12.2075643. [Online]. Available: https://doi.org/10.1117/12.2075643 (cit. on p. 2).

[29] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011 (cit. on pp. 2, 36, 37, 39, 46).

[30] J.-L. Starck, D. Donoho, and M. Elad, "Redundant multiscale transforms and their application for morphological component separation," CM-P00052061, Tech. Rep., 2004 (cit. on p. 2).

[31] J. Bobin, J.-L. Starck, J. Fadili, and Y. Moudden, "Sparsity and morphological diversity in blind source separation," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2662–2674, 2007 (cit. on p. 2).

[32] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011 (cit. on p. 2).

[33] J. Bieniarz, R. Müller, X. Zhu, and P. Reinartz, "Sparse approximation, coherence and use of derivatives in hyperspectral unmixing.," 2012 (cit. on p. 2).

[34] R. Bacher, "Méthodes pour l'analyse des champs profonds extragalactiques muse : Démélange et fusion de données hyperspectrales ;détection de sources étendues par inférence à grande échelle," PhD thesis, Communauté Université Grenoble Alpes, 2016, pp. 1282–1285 (cit. on pp. 2, 104).

[35] C.-I. Chang, *Hyperspectral data processing: algorithm design and analysis*. John Wiley & Sons, 2013 (cit. on p. 2).

[36] Z. Yang, G. Zhou, S. Xie, S. Ding, J. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1112–1125, Apr. 2011, ISSN: 1941-0042. DOI: 10.1109/TIP.2010.2081678 (cit. on p. 2).

[37] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012 (cit. on pp. 2, 8, 12, 17).

[38] H. Eavani, R. Filipovych, C. Davatzikos, T. D. Satterthwaite, R. E. Gur, and R. C. Gur, "Sparse dictionary learning of resting state fMRI networks," in *Pattern Recognition in NeuroImaging (PRNI), 2012 International Workshop on*, IEEE, 2012, pp. 73–76 (cit. on pp. 2, 3).

[39] A. Mensch, G. Varoquaux, and B. Thirion, "Compressed online dictionary learning for fast resting-state fMRI decomposition," in *International Symposium on Biomedical Imaging*, ser. 13th International Symposium on Biomedical Imaging, IEEE, Prague, Czech Republic: IEEE, Apr. 2016, pp. 1282–1285. DOI: 10.1109/ISBI.2016.7493501. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01271033 (cit. on pp. 2, 3, 8, 17).

[40] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent component analysis and applications*. Academic press, 2010, p. 643 (cit. on p. 2).

[41] V. Abolghasemi, S. Ferdowsi, and S. Sanei, "Blind separation of image sources via adaptive dictionary learning," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 2921–2930, Jun. 2012, ISSN: 1941-0042. DOI: 10.1109/TIP.2012.2187530 (cit. on p. 2).

[42] A. Abraham, E. Dohmatob, B. Thirion, D. Samaras, and G. Varoquaux, "Extracting brain regions from rest fMRI with total-variation constrained dictionary learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2013, pp. 607–615 (cit. on p. 3).

[43] G. Varoquaux, A. Gramfort, F. Pedregosa, V. Michel, and B. Thirion, "Multi-subject dictionary learning to segment an atlas of brain spontaneous activity," in *Information Processing in Medical Imaging*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 562–573 (cit. on p. 3).

[44] Bioucas-Dias, José M and Figueiredo, Mário AT, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, IEEE, 2010, pp. 1–4 (cit. on pp. 3, 8, 25, 108).

[45] M. Udell, C. Horn, R. Zadeh, and S. Boyd, "Generalized low rank models," *arXiv preprint arXiv:1410.0342*, 2014 (cit. on p. 5).

[46] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE TRANS. NEURAL NETW*, vol. 10, no. 3, pp. 626–634, 1999 (cit. on p. 5).

[47] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995 (cit. on p. 5).

[48] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," in *IEE proceedings F (radar and signal processing)*, IET, vol. 140, 1993, pp. 362–370 (cit. on p. 5).

[49] F. R. Bach and M. I. Jordan, "Kernel independent component analysis," *Journal of machine learning research*, vol. 3, no. Jul, pp. 1–48, 2002 (cit. on p. 5).

[50] V. D. Calhoun, T Adali, G. Pearlson, and J. J. Pekar, "Spatial and temporal independent component analysis of functional mri data containing a pair of task-related waveforms," *Human brain mapping*, vol. 13, no. 1, pp. 43–53, 2001 (cit. on p. 6).

[51] M. Kowalski, "Sparse regression using mixed norms," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 303–324, 2009 (cit. on p. 7).

[52] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on geoscience and remote sensing*, vol. 26, no. 1, pp. 65–74, 1988 (cit. on p. 9).

[53] J. W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," 1993 (cit. on p. 9).

[54] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of aviris data," 1995 (cit. on p. 9).

[55] M. E. Winter, "N-findr: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Imaging Spectrometry V*, International Society for Optics and Photonics, vol. 3753, 1999, pp. 266–275 (cit. on p. 10).

[56] J. M. Nascimento and J. M. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005 (cit. on p. 10).

[57] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data," in *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, vol. 3, 2008, pp. III–250 (cit. on p. 10).

[58] 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*, vol. 57, no. 11, pp. 4355–4368, 2009 (cit. on p. 11).

[59] P. J. Green, "Reversible jump markov chain monte carlo computation and bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995 (cit. on p. 11).

[60] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 721–741, 1984 (cit. on p. 11).

[61] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953 (cit. on p. 11).

[62] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," 1970 (cit. on p. 11).

[63] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, "Hybrid monte carlo," *Physics letters B*, vol. 195, no. 2, pp. 216–222, 1987 (cit. on p. 11).

[64] N. Dobigeon, Y. Altmann, N. Brun, and S. Moussaoui, "Linear and nonlinear unmixing in hyperspectral imaging," in *Data handling in science and technology*, vol. 30, Elsevier, 2016, pp. 185–224 (cit. on p. 12).

[65] B. Yang, X. Fu, N. D. Sidiropoulos, and K. Huang, "Learning nonlinear mixtures: Identifiability and algorithm," *IEEE Transactions on Signal Processing*, pp. 1–1, 2020 (cit. on p. 12).

[66] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966 (cit. on p. 12).

[67] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *Journal of Mathematics and Physics*, vol. 6, no. 1-4, pp. 164–189, 1927 (cit. on p. 12).

[68] V. V. Vesselinov, B. S. Alexandrov, and D. O'Malley, "Nonnegative tensor factorization for contaminant source identification," *Journal of contaminant hydrology*, vol. 220, pp. 66–97, 2019 (cit. on p. 13).

[69] L. Loncan, L. B. De Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, *et al.*, "Hyperspectral pansharpening: A review," *IEEE Geoscience and remote sensing magazine*, vol. 3, no. 3, pp. 27–46, 2015 (cit. on p. 18).

[70] H. Benali, I. Buvat, F. Frouin, J. Bazin, and R. Di Paola, "Foundations of factor analysis of medical image sequences: A unified approach and some practical implications," *Image and Vision Computing*, vol. 12, no. 6, pp. 375–385, 1994 (cit. on pp. 18, 96).

[71] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011 (cit. on pp. 21, 95, 126).

[72] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009 (cit. on pp. 21, 95).

[73] S. Boyd, J. Dattorro, *et al.*, "Alternating projections," *EE392o, Stanford University*, 2003 (cit. on p. 22).

[74] K. C. Kiwiel, "Breakpoint searching algorithms for the continuous quadratic knapsack problem," *Mathematical Programming*, vol. 112, no. 2, pp. 473–491, 2008 (cit. on p. 22).

[75] M. Held, P. Wolfe, and H. P. Crowder, "Validation of subgradient optimization," *Mathematical programming*, vol. 6, no. 1, pp. 62–88, 1974 (cit. on p. 22).

[76] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter estimation and inverse problems*. Elsevier, 2018 (cit. on p. 22).

[77] D. L. Collins, "3d model-based segmentation of individual brain structures from magnetic resonance imaging data," 1994 (cit. on p. 36).

[78] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson, *et al.*, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178, 2016. [Online]. Available: https://identifiers.org/neurovault.image:30759 (cit. on p. 36).

[79] N. K. Logothetis and B. A. Wandell, "Interpreting the BOLD signal," *Annu. Rev. Physiol.*, vol. 66, pp. 735–769, 2004 (cit. on p. 42).

[80] M. H. Lee, C. D. Smyser, and J. S. Shimony, "Resting-state fMRI: A review of methods and clinical applications," *American Journal of Neuroradiology*, vol. 34, no. 10, pp. 1866–1872, 2013 (cit. on pp. 42, 43).

[81] V. Zerbi, J. Grandjean, M. Rudin, and N. Wenderoth, "Mapping the mouse brain with rs-fmri: An optimized pipeline for functional network identification," *NeuroImage*, vol. 123, pp. 11 –21, 2015, ISSN: 1053-8119. DOI: https://doi.org/10.1016/j.neuroimage.2015.07.090. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811915007399 (cit. on p. 43).

[82] J. P. Owen, Y.-O. Li, F. G. Yang, C. Shetty, P. Bukshpun, S. Vora, M. Wakahiro, L. B. Hinkley, S. S. Nagarajan, E. H. Sherr, *et al.*, "Resting-state networks and the functional connectome of the human brain in agenesis of the corpus callosum," *Brain connectivity*, vol. 3, no. 6, pp. 547–562, 2013 (cit. on p. 43).

[83] Accueil, *Plateforme IRIS: Imagerie, robotique et innovation en santé,* https://plateforme.icube.unistra.fr/iris/index.php?title=Accueil&oldid=649, [Online; accessed 14-May-2020], 2020 (cit. on pp. 43, 55, 89).

[84] I. Havsteen, A. Ohlhues, K. H. Madsen, J. D. Nybing, H. Christensen, and A. Christensen, "Are movement artifacts in magnetic resonance imaging a real problem?—a narrative review," *Frontiers in neurology*, vol. 8, p. 232, 2017 (cit. on p. 43).

[85] S. M. Smith, "Fast robust automated brain extraction," *Human brain mapping*, vol. 17, no. 3, pp. 143–155, 2002 (cit. on p. 46).

[86] M. Jenkinson, M. Pechaud, S. Smith, *et al.*, "Bet2: Mr-based estimation of brain, skull and scalp surfaces," in *Eleventh annual meeting of the organization for human brain mapping*, Toronto., vol. 17, 2005, p. 167 (cit. on p. 46).

[87] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *Neuroimage*, vol. 17, no. 2, pp. 825–841, 2002 (cit. on p. 46).

[88] B. B. Avants, N. Tustison, and G. Song, "Advanced normalization tools (ANTS)," *Insight Journal*, vol. 2, pp. 1–29, 2009 (cit. on p. 47).

[89] W. Luo and T. Nichols, "Diagnosis and exploration of massively univariate neuroimaging models," *NeuroImage*, vol. 19, pp. 1014–1032, 2003 (cit. on p. 47).

[90] N. Chou, J. Wu, J. B. Bingren, A. Qiu, and K.-H. Chuang, "Robust automatic rodent brain extraction using 3-d pulse-coupled neural networks (pcnn)," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2554–2564, 2011 (cit. on p. 48).

[91] I. Oguz, H. Zhang, A. Rumple, and M. Sonka, "Rats: Rapid automatic tissue segmentation in rodent brain mri," *Journal of neuroscience methods*, vol. 221, pp. 175–182, 2014 (cit. on p. 48).

[92] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka, "Logismos—layered optimal graph image segmentation of multiple objects and surfaces: Cartilage segmentation in the knee joint," *IEEE transactions on medical imaging*, vol. 29, no. 12, pp. 2023–2037, 2010 (cit. on p. 48).

[93] R. Sladky, K. J. Friston, J. Tröstl, R. Cunnington, E. Moser, and C. Windischberger, "Slice-timing effects and their correction in functional mri," *Neuroimage*, vol. 58, no. 2, pp. 588–594, 2011 (cit. on p. 54).

[94] C. Lau, L. Ng, C. Thompson, S. Pathak, L. Kuan, A. Jones, and M. Hawrylycz, "Exploration and visualization of gene expression with neuroanatomy in the adult mouse brain," *BMC bioinformatics*, vol. 9, no. 1, p. 153, 2008 (cit. on p. 56).

[95] E. Damaraju, E. A. Allen, A. Belger, J. M. Ford, S McEwen, D. Mathalon, B. Mueller, G. Pearlson, S. Potkin, A Preda, *et al.*, "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia," *NeuroImage: Clinical*, vol. 5, pp. 298–308, 2014 (cit. on pp. 60, 63).

[96] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: A practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995 (cit. on pp. 64, 135).

[97] L. Degiorgis, M. Karatas, M. Sourty, C. Po, T. Bienert, H. Lei Lee, D. v. Elverfeldt, C. Mathis, A.-L. Boutillier, F. Blanc, J.-P. Armspach, and L.-A. Harsan, "Patterns of resting-state functional connectivity in the prodromal phase of alzheimer's disease: Insights from a tauopathy mouse model (thy-tau22)," in *ISMRM 25th Annual Meeting and Exhibition*, 2017 (cit. on pp. 80, 88).

[98] Y.-m. Zou, L.-p. L. Da Lu, H.-h. Zhang, and Y.-y. Zhou, "Olfactory dysfunction in alzheimer's disease," *Neuropsychiatric disease and treatment*, vol. 12, p. 869, 2016 (cit. on p. 88).

[99] G. Martel, A. Simon, S. Nocera, S. Kalainathan, L. Pidoux, D. Blum, S. Leclère-Turbant, J. Diaz, D. Geny, E. Moyse, *et al.*, "Aging, but not tau pathology, impacts olfactory performances and somatostatin systems in thy-tau22 mice," *Neurobiology of aging*, vol. 36, no. 2, pp. 1013–1028, 2015 (cit. on p. 88).

[100] K. Nijran and D. Barber, "The importance of constraints in factor analysis of dynamic studies," in *Information processing in medical imaging*, Springer, 1988, pp. 521–529 (cit. on p. 96).

[101] M Šámal, A Piepsz, G Brolin, J. Heikkinen, and J Valoušek, *Database of dynamic renal scintigraphy*, http://www.dynamicrenalstudy.org/, [Online; accessed 3-Mars-2020], 2012 (cit. on p. 97).

[102] G. Brolin, K. Sjögreen Gleisner, and M. Ljungberg, "Dynamic (99m)Tc-MAG3 renography: images for quality control obtained by combining pharmacokinetic modelling, an anthropomorphic computer phantom and Monte Carlo simulated scintillation camera imaging.," English, *Physics in Medicine and Biology*, vol. 58, no. 10, pp. 3145–3161, 2013, ISSN: 1361-6560. DOI: 10.1088/0031-9155/58/10/3145 (cit. on p. 97).

[103] R Bacon, M Accardo, L Adjali, H Anwand, S Bauer, I Biswas, J Blaizot, D Boudon, S Brau-Nogue, J Brinchmann, *et al.*, "The MUSE second-generation VLT instrument," in *Ground-based and Airborne Instrumentation for Astronomy III*, International Society for Optics and Photonics, vol. 7735, 2010, p. 773 508 (cit. on p. 103).

[104] R. Bacon, S. Conseil, D. Mary, J. Brinchmann, M. Shepherd, M. Akhlaghi, P. M. Weilbacher, L. Piqueras, L. Wisotzki, D. Lagattuta, *et al.*, "The MUSE Hubble Ultra Deep Field Survey-I. Survey description, data reduction, and source detection," *Astronomy & Astrophysics*, vol. 608, A1, 2017 (cit. on pp. 104, 106, 109).

[105] H. Inami, R. Bacon, J. Brinchmann, J Richard, T. Contini, S. Conseil, S Hamer, M. Akhlaghi, N. Bouché, B Clément, *et al.*, "The MUSE Hubble Ultra Deep Field Survey-II. Spectroscopic redshifts and comparisons to color selections of high-redshift galaxies," *Astronomy & Astrophysics*, vol. 608, A2, 2017 (cit. on p. 104).

[106] R. Bacon, L. Piqueras, S. Conseil, J. Richard, and M. Shepherd, *MPDAF: MUSE Python Data Analysis Framework*, Nov. 2016. ascl: 1611.003 (cit. on p. 106).

[107] D. Mary, R. Bacon, S. Conseil, L. Piqueras, and A. Schutz, "Origin: Blind detection of faint emission line galaxies in muse datacubes," *Astronomy & Astrophysics*, vol. 635, A194, 2020 (cit. on pp. 119–122).

[108] M. V. Maseda, R. Bacon, M. Franx, J. Brinchmann, J. Schaye, L. A. Boogaard, N. Bouché, R. J. Bouwens, S. Cantalupo, T. Contini, *et al.*, "Muse spectroscopic identifications of ultra-faint emission line galaxies with m uv - 15," *The Astrophysical Journal Letters*, vol. 865, no. 1, p. L1, 2018 (cit. on pp. 119, 122).

[109] C. Meillier, "Détection de sources quasi-ponctuelles dans des champs de données massifs," PhD thesis, Grenoble Alpes, 2015 (cit. on pp. 130, 132, 134).

[110] K. Pearson, "X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 50, no. 302, pp. 157–175, 1900 (cit. on p. 131).

[111] W. G. Cochran, "The chi2 test of goodness of fit," *The Annals of mathematical statistics*, pp. 315–345, 1952 (cit. on p. 131).

[112] G. W. Snedecor and W. G. Cochran, "Statistical methods, 8thedn," *Ames: Iowa State Univ. Press Iowa*, vol. 54, pp. 71–82, 1989 (cit. on p. 132).

[113] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945, ISSN: 00994987. [Online]. Available: http://www.jstor.org/stable/3001968 (cit. on p. 133).

[114] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian journal of statistics*, pp. 65–70, 1979 (cit. on pp. 134, 135).

[115] C. Li, "An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing," PhD thesis, 2010 (cit. on p. 145).