

ÉCOLE DOCTORALE MATHÉMATIQUES, SCIENCES de l'INFORMATION et de l'INGÉNIEUR

Laboratoire ICube UMR 7357

THÈSE présentée par :

Marion SOURTY

soutenue le : **16 septembre 2016**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline / Spécialité : Signal, image, automatique, robotique / traitement du signal et des images

Analyse de la dynamique temporelle et spatiale des réseaux cérébraux spontanés obtenus en imagerie par résonance magnétique fonctionnelle

THÈSE dirigée par :

M. Jean-Paul ARMSPACH
M. Jack FOUCHER

Ingénieur de recherche, HDR, Université de Strasbourg
Maître de conférences, Praticien Hospitalier, HDR, Université de Strasbourg, CHU de Strasbourg

RAPPORTEURS :

Mme Su RUAN
M. Renaud JARDRI

Professeur, Université de Rouen
Professeur, Praticien Hospitalier, CHRU de Lille

MEMBRE INVITÉ :

M. Laurent THORAVAL

Maître de conférences, Université de Strasbourg

Remerciements

Beaucoup de personnes m'ont aidée à réaliser cette thèse, à la tenter d'abord, à l'achever enfin. J'espère que ces mots simples sauront témoigner de mon envie de dire merci à tous ceux qui ont permis ce travail.

La présence sur cette page des professeurs Su Ruan et Renaud Jardri marque la fin de cette thèse. Je les remercie d'avoir accepté le rôle de rapporteur et pour les idées et pistes de recherches permises par leur discussion.

Je remercie mes directeurs et encadrant de thèse, Jean-Paul, Jack et Laurent, pour avoir imaginé ce sujet qui ouvre la voie à de nombreux projets et m'avoir fait confiance pour l'entamer. C'est à vous qu'il revient l'opportunité de poursuivre nos idées.

Je tiens particulièrement à remercier Slavisa, qui, même sans l'avoir vu depuis le commencement de ma thèse a été décisif pour mon choix de me lancer dans ce travail. Ta curiosité, ta disponibilité et ton inventivité sans limite me donne une image du chercheur que je souhaite être.

Je remercie toute l'équipe IMIS au sein de laquelle j'ai passé mes années de thésarde ainsi que l'équipe MIV chez qui j'ai fait quelques apparitions. La bonne ambiance qui règne dans le laboratoire a fortement contribué au bon déroulement de cette thèse. Je tiens à remercier particulièrement Julien pour sa disponibilité au travail et pour m'avoir proposé des midis sportifs, Frédéric pour être toujours positif et motivé pour de nouvelles études, Franklin pour m'avoir fait découvrir les ficelles de la vie d'équipe, Aude pour sa gentillesse et son professionnalisme et Vincent pour les questions qu'il a soulevées. Je remercie également les autres doctorants Lucas, Laetitia, Meltem, Coralie, Elena : bon courage ! C'est à vous que revient de préparer les fêtes de labo maintenant !

Cette thèse aurait été bien différente si je ne l'avais entamée avec Alix, ma colocataire à l'IPB. Je la remercie du fond du cœur pour son soutien. Je te souhaite de trouver ta voie dans la recherche ou ailleurs... Je remercie d'ailleurs Madeleine² pour sa disponibilité entre 13h30 et 14h et sa capacité à me déconcentrer qui m'ont été tellement nécessaires.

Je remercie également Maroua et Camille que cette thèse à Strasbourg a permis de me faire rencontrer. Vous avez été un réel soutien pour les différents étapes de ces dernières années. Je sais qu'on se reverra très vite.

Je remercie mes amis disséminés dans toute la France qui m'ont proposé tant de distractions essentielles. Lyon, Mirecourt, Nîmes, Paris, Rennes, Nancy, Toulouse, Biarritz, je vais pouvoir recommencer mes tours de France.

Je remercie tout particulièrement Dan pour avoir tenté (en vain) de me faire entrer dans le monde de la neurobiologie. Merci de m'avoir portée et supportée durant ces dernières années, au niveau personnel comme professionnel.

Enfin, je remercie mes parents qui me soutiennent depuis toujours. Je ne doute pas qu'après la lecture de ce manuscrit, vous saurez expliquer l'objet de mon travail. Je remercie Jérémie, Clémence, Marjolaine et Côme par qui le mot fraternité a été inspiré. L'humour intarissable et la solidarité dont vous faites preuve nous donnent la capacité de réaliser tous nos projets.

À Clémence, Pour la force que tu as su trouver ces dernières années, et que tu m'as transmise sans t'en apercevoir. J'y ai puisé ma volonté tout au long de ce travail.

Soyez le changement que vous voulez voir dans ce monde
G.

Table des matières

Remerciements	iii
Table des figures	xi
Liste des tableaux	xiii
Introduction générale	xv
I Cadre de travail	1
Introduction	3
1 IRM fonctionnelle et réseaux de repos	5
1.1 Pré-requis anatomiques et fonctionnels de l'organisation cérébrale	5
1.1.1 Cerveau et structure connexes : point de vue anatomique	5
1.1.2 Physiologie du fonctionnement cérébral	6
1.2 Connectivité	7
1.2.1 Définition	7
1.2.2 Dynamique de la connectivité fonctionnelle	8
1.3 IRM fonctionnelle	9
1.3.1 Imagerie par Résonance Magnétique	9
1.3.2 Fondements physiologiques en IRM fonctionnelle	11
1.3.3 Principales étapes d'une étude en IRMf	13
1.4 Réseaux cérébraux spontanés - SAM	14
1.4.1 Définition	14
1.4.2 RSN : Réseaux communs mis en évidence en IRM fonctionnelle de repos	14
1.4.3 Réseaux idiosyncratiques	16
2 Analyse en Composantes Indépendantes en IRMf de repos	19
2.1 Méthode	19
2.1.1 Introduction	19
2.1.2 Formalisation du problème	20
2.2 Application en IRMf de repos	21
2.2.1 Généralités	21
2.2.2 Points difficiles	24
2.2.3 Évaluation de la méthode	24
3 Analyse markovienne cachée et processus en interaction : une synthèse	27
3.1 Modèle de Markov Caché standard	28
3.1.1 Principe de base	28
3.1.2 Formalisme mathématique	28

3.1.3	Problèmes fondamentaux et solutions	29
3.2	Modèles de Markov cachés multidimensionnels	32
3.2.1	Problématiques générales	32
3.2.2	Modèle de Markov Caché Produit (MMC-P)	32
3.2.3	Modèle de Markov Caché Couplés (MMC-C)	33
3.2.4	Autres MMC multidimensionnels	34
4	Bases de données utilisées	37
4.1	Composition des bases de données	37
4.2	Répartition des bases de données en fonction des études	39
4.2.1	Bases de données pour la sélection des SAM (Partie II)	39
4.2.2	Bases de données pour l'analyse de la DCF (Partie III)	40
	Conclusion partielle	41
II	Sélection automatique des cartes de co-activités spatiales	43
	Introduction	45
5	Méthodes de sélection des composantes liées à une activité cérébrale : Éléments d'état de l'art	47
5.1	Du classement des études	47
5.2	Méthodes de sélection de composantes	48
5.3	Point de départ : ce que disent les experts, article de Roquet et coll.	52
5.3.1	Critères de sélection	52
5.3.2	Critères de rejet	52
5.4	Descripteurs issus de la littérature	55
6	Méthode proposée	59
6.1	Vue d'ensemble de la méthode	59
6.2	Pré-traitements et outils	60
6.3	Descripteurs existants	61
6.4	Descripteurs nouveaux	64
6.5	Comparaison des descripteurs	65
6.5.1	Méthodes d'évaluation	65
6.5.2	Similitude empirique	66
6.6	Distributions des valeurs descriptives et règle de sélection	67
6.7	Règle de fusion	68
6.8	Alternatives possibles à chaque étape de la méthode	68
6.8.1	Identification des SAM	68
6.8.2	Règle de décision	68
6.8.3	Règle de décision et règle de fusion	69
7	Résultats et analyse des résultats	71
7.1	Analyse des résultats par base	71
7.2	Analyse des résultats par descripteur	72
7.2.1	Performance des descripteurs	72
7.2.2	Estimation des seuils	74
7.3	Analyses comparatives	74
7.3.1	Résultats	74
7.3.2	Discussion	76

7.4	Interface	77
7.5	Estimation du gain de temps	78
7.6	Applicabilité	78
Conclusion partielle		79
III Analyse statistique de la dynamique de la connectivité fonctionnelle		81
Introduction		83
8	Dynamique de la connectivité fonctionnelle	85
8.1	Principes et Problématiques	85
8.1.1	Variabilité de la DCF	85
8.1.2	Problématiques et motivations pour l'étude de la DCF	86
8.1.3	DCF et état de connectivité fonctionnelle	87
8.2	État de l'art des méthodes d'analyse de la DCF	88
8.2.1	Fenêtre glissante	88
8.2.2	Modèle reproductible avec approche par fenêtre glissante	89
8.2.3	Analyse en vecteurs indépendants (AVI) spatiale	90
8.2.4	ACI temporelle	91
8.2.5	Analyse temps-fréquence	92
8.2.6	Modèles de Markov	92
8.2.7	Bilan	93
8.3	Intérêt de l'analyse de la DCF pour les applications cliniques	95
8.4	Notre approche dans ce contexte	96
9	Analyse de données : observations et résultats préliminaires	99
9.1	Décours temporels vs. signaux de corrélations	99
9.2	Observations préliminaires	100
9.3	Résultats préliminaires	103
9.3.1	Mise en évidence de réseaux décorrélés	103
9.3.2	Médicaments vs. Placebo	105
10	Modélisation et analyse de la DCF par MMC-Produit et MMC-Couplés	107
10.1	Principes et vue d'ensemble	107
10.2	Modélisation	108
10.3	Initialisation	110
10.4	Apprentissage	111
10.5	Analyse	112
10.5.1	Comparaison de modèles	112
10.5.2	Matrice de transition, Décodage de Viterbi	113
11	Applications cliniques : Médicament vs placebo, Démence à corps de Lewy	115
11.1	Application 1 basée sujet Étude « Médicament vs Placebo »	115
11.1.1	Objectifs et contraintes	115
11.1.2	Mise en œuvre	115
11.1.3	Résultats	116
11.1.4	Discussion	124
11.2	Application 2 basée groupe : « Démence à corps de Lewy »	125
11.2.1	Objectifs et contraintes	125
11.2.2	Mise en œuvre	126

11.2.3	Analyses complémentaires	127
11.2.4	Résultats	128
11.2.5	Discussion	130
11.3	Bilan	132
Conclusion partielle		135
Conclusion générale et Perspectives		137
Liste des publications		145
A Cartes spatiales de SAM		147
A.1	Cartes spatiales des SAM sélectionnées	147
A.2	Cartes spatiales des SAM non sélectionnées : faux négatifs	148
B Distribution des valeurs des descripteurs par base		151
C Interface de sélection		153
D Cartes spatiales des RSN pour l'étude de la DCF sur la base B5		155
Bibliographie		157

Table des figures

1.1	Anatomie cérébrale	6
1.2	Images anatomiques du cerveau à partir de l'IRM	6
1.3	Les aires définies par Brodmann et quelques fonctions associées.	7
1.4	Illustration de la dynamique de la connectivité fonctionnelle temporelle	9
1.5	Mouvement de précession du spin	10
1.6	Illustrations des principaux RSN.	16
1.7	Exemple d'un réseau idiosyncratique chez un patient cérébrolésé.	17
2.1	Illustration de l'ACI spatiale	20
2.2	illustration de l'ACI en IRM fonctionnelle.	23
3.1	Graphe de dépendance d'un MMC standard	28
3.2	Graphe de dépendance d'un MMC-Produit	33
3.3	Graphe de dépendance d'un MMC-Couplés	34
5.1	Réseau par défaut	53
5.2	Un réseau idiosyncratique.	53
5.3	MON (Median Occipital Network).	53
5.4	Artefacts et bruits	54
6.1	Synoptique de la méthode	59
6.2	Illustration d'une composante	62
7.1	Distributions des valeurs descriptives <i>focusing degree</i> et <i>volume</i>	75
8.1	Illustration de la dynamique de la connectivité fonctionnelle temporelle	86
9.1	Corrélogramme temporel	102
9.2	Histogrammes des signaux de corrélation	104
9.3	Le réseau PPCN	105
10.1	Synoptique de la méthode	109
11.1	Matrices de transitions pour le sujet M7, avec les MCC-P et MCC-C et différences	120
11.2	Distribution temporelle des états : exemple 1	121
11.3	Distribution temporelle des états : exemple 2	122
11.4	Distribution temporelle des états : exemple 3	123
11.5	Illustration matrice de transition application groupe	129
11.6	Nouveau corpus d'apprentissage de L corrélogrammes $CC_{n,s,l}$	141
11.7	Graphe de dépendance du MMC-Couplés à seuils	142
A.1	B2 - sujet 9 - session 2 - composante 155	147
A.2	B5 - sujet 5 - composante 84	147

A.3	B1 - sujet 5 - composante 31	148
A.4	B1 - sujet 2 - composante 63	149
A.5	B1 - sujet 11 - composante 85	150
B.1	Distribution selon les bases et les descripteurs	152
C.1	Interface de sélection	154
D.1	Cartes spatiales des RSN	155

Liste des tableaux

1.1	Abréviations, fonctions et caractéristiques des RSN les plus communs.	15
4.1	Bases de données	38
4.2	Bases de données selon les études	39
4.3	Bases de données d'apprentissage et de test	39
5.1	Études sur sujet unique.	51
5.2	Descripteurs existants	56
6.1	Comparaisons des descripteurs	67
6.2	Test de Box (approximation asymptotique du χ^2).	69
6.3	Test de Box (approximation asymptotique du F de Fisher).	69
6.4	Test du Lambda de Wilks (approximation de Rao).	69
7.1	Résultats par base.	71
7.2	Spécificité par descripteur.	73
7.3	Seuil par descripteur.	74
7.4	Résultats par base avec l'algorithme des k-moyennes.	76
7.5	Résultats par base avec l'analyse factorielle discriminante.	76
7.6	Résultats par base avec l'algorithme des forêts aléatoires.	77
9.1	SAM communes aux sessions	100
9.2	SAM décorréliées	105
11.1	Dimension de S en fonction du nombre de SAM	116
11.2	SAM retenues pour la modélisation	116
11.3	Résultats des comparaisons entre sessions par sujet	117
11.4	Résultats des comparaisons entre groupe pour chaque RSN étudiés	128
11.5	Résultats positifs au t-test pour les 20 générations de signaux	130
11.6	Résultats spécifiques aux patients avec fluctuations cognitives	131

Introduction générale

Cadre de travail

L'imagerie par résonance magnétique fonctionnelle (IRMf) par contraste BOLD (blood oxygenation level-dependent) est un outil de choix pour cartographier d'une manière non invasive l'activité du cortex, donnant ainsi un accès à l'organisation fonctionnelle cérébrale. Cette organisation des aires cérébrales en réseaux complexes reste encore un vaste sujet d'étude, autant dans le domaine de la recherche fondamentale, pour mieux comprendre le développement et le fonctionnement du cerveau, que dans le domaine clinique, à des fins diagnostiques par exemple. Les réseaux cérébraux dits de repos, chez un sujet donné, peuvent être observés lors d'études IRMf, elles aussi dites de repos, c'est à dire lorsqu'aucune tâche motrice ou cognitive n'est imposée au sujet imagé, et quand le sujet ne dort pas. Ils peuvent être regroupés sous la dénomination de « SAM » pour Spontaneous co-Activity Maps et peuvent être divisés en deux groupes. D'une part, certaines SAM sont partagées par une population et sont appelées réseaux de repos, plus connus sous l'acronyme RSN pour resting-state networks. Ce sont les seuls réseaux qu'il est possible d'observer lors des études de groupe. D'autre part, sont présents les réseaux propres au sujet, les réseaux dits idiosyncratiques qui représentent environ un tiers des SAM chez un sujet.

La très grande majorité des travaux de recherche dans ce domaine s'intéresse aux RSN, donc aux études de groupes. Si ce type d'études reste pertinent dans le cadre de la recherche pour décrire ou distinguer des populations, en routine clinique les objectifs sont différents : l'analyse et l'interprétation doivent être faites au niveau du sujet afin de préserver ses spécificités, ce qui s'avère nécessaire dans un cadre diagnostique où l'individu prime. Dans ce but, la détection et l'identification des réseaux idiosyncratiques sont essentielles, certaines SAM pouvant être de potentiels biomarqueurs d'une pathologie ou d'un état (comme c'est le cas de certaines SAM idiosyncratiques liées spécifiquement aux hallucinations dans une population de patients schizophrènes). Au delà de la détection et de l'identification de SAM, l'étude des modes d'interaction de ces réseaux dans l'espace et dans le temps et plus généralement l'analyse de la dynamique de la connectivité fonctionnelle (DCF) est un champ de recherche très récent (2010). Cette analyse nécessite le développement de méthodes innovantes de traitement du signal et de l'image qui, pour l'heure, sont encore de nature exploratoire, les mécanismes qui sous-tendent ces interactions étant en effet mal connus et/ou à mettre au jour.

Cette thèse, menée au sein du laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie (ICube, UMR 7357, Université de Strasbourg-CNRS) dans les équipes « Imagerie Multimodale Intégrative en Santé » (IMIS) et « Modèles, Images et Vision » (MIV), à Strasbourg, présente des contributions dans les deux domaines de recherche que sont la détection et l'identification de SAM, et l'analyse de la dynamique de la connectivité fonctionnelle.

Problématiques

L'Analyse en Composantes Indépendantes (ACI ou Independent Component Analysis) est aujourd'hui l'outil privilégié pour distinguer les SAM. Entièrement guidée par les données, cette technique robuste permet de décomposer les volumes cérébraux, acquis lors de l'expérience IRMf, en cartes d'activation et décours temporels associés. Ces cartes spatiales, indépendantes les unes des autres, appelées composantes, sont majoritairement le reflet du bruit ou des artefacts. Cependant, elles peuvent aussi être représentatives de réseaux d'activations cérébrales spontanées (les SAM). Pour des études de connectivité fonctionnelle, on cherchera à sélectionner les composantes d'intérêt mettant en avant des SAM, pour analyse ultérieure. Notre première étude s'inscrit dans cette démarche.

La sélection et l'identification de SAM par l'ACI se heurtent à plusieurs difficultés. Tout d'abord, pour une étude sur sujet unique, le nombre de composantes à traiter est élevé comparativement à une étude de groupe (entre 120 et 250 composantes dans nos études contre 20 à 60 pour les études de groupe). Cette étude approfondie est en effet nécessaire pour révéler des réseaux singuliers. Si la plupart des RSN sont bien décrits dans la littérature, certains réseaux, idiosyncratiques ou partagés par une pathologie, sont peu connus alors même qu'ils pourraient être des biomarqueurs potentiels. En outre, les SAM ne représentent qu'une faible proportion des composantes (de 5 % à 10 % des composantes dans nos études). Les autres composantes sont principalement du bruit et peuvent être liées, entre autres, à des bruits physiologiques (respiration, battements cardiaques), des mouvements (de la tête, des yeux), des artefacts dus à l'imageur ou à la séquence choisie. Une analyse fine de ces nombreuses composantes est donc nécessaire pour détecter toutes les SAM d'un sujet, en d'autres termes, une sensibilité de détection de 100 % est recherchée. Ce processus fastidieux, nécessite un haut niveau d'expertise, même si des critères de sélection manuelle ont pu être établis pour accélérer le processus.

Les problématiques soulevées par l'analyse de la dynamique de la connectivité fonctionnelle (DCF) vont quant à elles bien au-delà de celles liées à la sélection de SAM. Il faut d'abord souligner que ce champ de recherche en est à ses balbutiements et que peu de méthodes ont été publiées. En effet, il a longtemps été considéré, de manière simplificatrice, que la connectivité fonctionnelle était stationnaire durant la durée de l'examen. Si les variations de la connectivité fonctionnelle entre sujets, ou entre sessions d'un même sujet ont rapidement été mises en avant, ce n'est que très récemment (2010) que la notion de DCF a été introduite. Aucune méthode ne semble faire encore l'unanimité, chaque type de données (sujets sains, patients, tâches, repos...) ayant ses propres particularités. L'analyse de la DCF permet par exemple de mettre en avant certaines configurations cérébrales, représentant différentes interactions entre les réseaux ou les régions qui les sous-tendent. Est ainsi apparue la notion d'« état » cérébral bien qu'aucun consensus sur sa définition n'existe à ce jour. De là, on peut envisager de représenter la dynamique de la connectivité fonctionnelle sous forme de séquence d'états et ainsi chercher la présence de structure ou de modèles reproductibles. Cependant, la très grande complexité combinatoire d'états et de séquences d'états, au regard de fenêtres d'observation courtes dans les études d'IRMf, soulève la question de la significativité des résultats. En outre, les variabilités intra-individu, inter-individus, entre sujets sains et pathologiques apportent aussi des difficultés supplémentaires dans la compréhension de la DCF. Pour finir, de nombreuses questions restent ouvertes concernant la façon d'aborder et d'interpréter la DCF : existe-t-il des séquences temporelles déterministes ? Peut-on dénombrer les configurations de liens entre les réseaux ? Quelles sont la représentation et la définition appropriées d'un réseau alors que celui-ci varie dans l'espace et dans le temps ?

Contributions

L'objectif de la première partie de la thèse a été de mettre au point un outil de pré-sélection automatique des SAM d'un sujet et ce de manière exhaustive. A partir d'une méthode manuelle, nous avons développé une chaîne de traitement permettant l'automatisation de la pré-sélection des SAM chez un sujet, avec pour contrainte forte la sélection pour chaque sujet de toutes les composantes ACI jugées d'intérêt par l'expert, c'est-à-dire avec une sensibilité de 100 % dans la détection des réseaux cérébraux spontanés.

La pré-sélection des SAM est basée sur des caractéristiques, précédemment identifiées, de leur carte spatiale et du décours temporel associé. Ainsi, cinq descripteurs ont été créés et mis en œuvre afin d'évaluer pour chaque composante : la cohérence spatiale (degré d'agrégation, rejet des cartes spatiales avec des coupes saturées, rejet des cartes spatiales présentant une alternance de connectivité entre les coupes), la localisation des agrégats formant le réseau et les fréquences présentes dans le décours temporel associé. La classification des composantes en bruit ou en réseaux cérébraux spontanés s'appuie sur la construction de statistiques de ces descripteurs et l'estimation de seuils de décision à partir de données d'apprentissage.

La méthode de pré-sélection automatique de réseaux cérébraux spontanés a été testée sur quatre bases de données issues d'imageurs, de populations, de résolutions et de contrastes différents. Les résultats obtenus ont été évalués suivant les critères standards de sensibilité, de spécificité et de prédictivité positive. Ainsi, la méthode développée permet une réduction de 70 % en moyenne du nombre de composantes à sélectionner tout en gardant une sensibilité de 100 % sur trois bases et 98.9 % sur une quatrième. De plus, les composantes retenues sont classées en fonction de leur qualité de réponse aux critères, facilitant ainsi la sélection finale des réseaux spontanés par l'expert. La sélection manuelle des composantes d'intérêt parmi les composantes retenues par notre méthode automatique est donc bien plus aisée suite à cette étape, permettant un gain de temps conséquent (entre 30 % et 87 % de temps gagné par rapport à une sélection entièrement manuelle selon le type de données) ainsi qu'une fiabilité augmentée comme en témoignent quelques révisions de jugement des deux experts.

Pour synthétiser l'affichage des résultats et faciliter la sélection restante, une interface a été développée sous Matlab (2012b, The MathWorks, Natick, États-Unis). Elle permet à l'expert de visualiser les composantes pré-sélectionnées ainsi que les scores aux différents descripteurs. Ces informations quantitatives sont une précision supplémentaire par rapport à une sélection manuelle classique et rendent la décision finale plus rapide et plus sûre.

Dans la seconde partie de cette thèse, consacrée à l'analyse de la dynamique de la connectivité fonctionnelle, nous proposons d'identifier les moments pendant lesquels les réseaux cérébraux spontanés se corrèlent ou s'anti-corrèlent. Deux applications ont été réalisées afin de se placer à l'échelle du sujet puis à l'échelle du groupe. La première application retenue vise à détecter et à caractériser le cas échéant le changement de la DCF avant et après prise de médicament, ainsi qu'avant et après prise d'un placebo. Pour ce faire, nous disposons d'une base de douze sujets avec quatre sessions de 20 minutes chacune, la dernière session étant avec prise de médicaments pour huit des douze sujets, ou prise de placebo pour les sujets restants. La deuxième application est une étude de groupe visant à comparer deux populations : des sujets sains et des sujets atteints de démence à corps de Lewy. L'analyse de la DCF peut permettre non seulement de différencier ces deux groupes mais aussi de cibler les réseaux responsables de cette différenciation. L'analyse de la DCF s'appuie sur les signaux d'inter-corrélation entre les décours temporels associés aux réseaux spontanés. D'autre part, l'approche proposée se démarque d'autres approches par le fait qu'elle est basée réseaux, et non basée régions, et qu'elle s'applique sur sujet unique aussi bien que sur un groupe.

Le cadre méthodologique retenu est celui des modèles de Markov cachés (MMC) multidimen-

sionnels et plus largement celui des réseaux bayésiens dynamiques. Ce type de modélisation probabiliste basée sur les graphes offre un cadre puissant d'analyse à la fois adapté à la problématique soulevée (aspect dual caché - observé) et potentiellement explicatif des mécanismes d'interactions entre réseaux cérébraux spontanés. Deux états ont été pris en compte (corrélé ou anti-corrélé) pour chaque signal observé (inter-corrélation). Chaque sujet ayant entre 6 à 10 réseaux à prendre en considération, cela mène à un très grand nombre d'états. Deux modélisations ont été proposées afin de prendre en compte différents aspects des liens de dépendances entre les réseaux. Par la suite, un état « neutre », ni-corrélé ni anti-corrélé, pourra être ajouté à la modélisation. Il est défini grâce à une variable « seuil » supplémentaire qui permet de prendre en compte le degré de corrélation entre signaux temporels dans l'identification des états corrélés et anti-corrélés entre réseaux. Cela rajoute évidemment de la complexité combinatoire mais qui a été réduite grâce à la restriction du champ d'observations aux RSN impliqués dans des fonctions de haut niveau et dont l'étude était plus particulièrement pertinente eu égard au médicament étudié.

Deux approches de modélisation de la DCF ont été successivement proposées : le MMC produit, les MMC couplés. Elles permettent d'appréhender les interactions de plusieurs réseaux simultanément en intégrant directement les modèles sous-jacents à chaque observation. Ces trois approches se distinguent principalement par leur graphe de dépendance entre états cachés de corrélation / d'anti-corrélation. La première approche (MMC produit) consiste à considérer que ce que l'on observe est généré par un état global qui est la concaténation des états de chaque signal observé. La seconde approche (MMCs couplés) permet de réduire le nombre de paramètres à estimer du modèle, en particulier les probabilités de transition, en réduisant le nombre de dépendances conditionnelles entre états cachés.

Ces approches de modélisation ont été mises en œuvre dans le cadre de la caractérisation de la DCF en présence de prise de médicament ou de placebo. Les résultats obtenus ont permis de montrer tout l'intérêt de ces modèles, non seulement de manière quantitative, par la possibilité de calcul de distances de (dis)similarité entre modèles de DCF, mais aussi de manière qualitative, par la possibilité de caractériser la DCF à travers la matrice de transition estimée des modèles, ou à travers les séquences d'états décodées, donnant accès à une transcription symbolique (corrélé, non corrélé) des décours temporels observés. A noter enfin que ces modèles, bien que développés sur sujets uniques dans le cadre de cette thèse, peuvent aisément être généralisés à l'étude d'un groupe, à l'étude d'un réseau particulier ou encore à l'analyse de la DCF d'une aire cérébrale ciblée au sein d'un réseau.

Organisation du manuscrit

Le mémoire de thèse est composée de trois parties.

La première, destinée à exposer les pré-requis généraux, décrit les principes aussi bien liés aux neurosciences qu'au traitement du signal. Ainsi, le premier chapitre survole les notions d'imagerie par résonance magnétique et de connectivité fonctionnelle pour mieux introduire le cœur d'étude de cette thèse : les réseaux cérébraux spontanés. L'analyse en composantes principales, outil de référence pour la mise en évidence de ces réseaux, est le sujet du deuxième chapitre. Nous y explicitons aussi bien la méthode générale de cette approche que sa mise en œuvre dans une étude de connectivité par IRM fonctionnelle. Enfin le dernier point de cette partie est l'introduction aux modèles markoviens cachés. Cette technique robuste sera utilisée de manière innovante dans le cadre de l'étude de la dynamique de la connectivité fonctionnelle. Ce chapitre est donc important afin d'introduire les algorithmes de bases de ces modèles ainsi que le formalisme mathématique associé.

La seconde partie de ce manuscrit détaille la première étude de cette thèse à savoir la sélection automatique des cartes de co-activités spatiales. Un premier chapitre fait un état de l'art des méthodes de sélection déjà publiées. Les chapitres suivants permettent de détailler les données uti-

lisées, la méthode proposée et les résultats obtenus. Un dernier chapitre discute de l'apport de ce travail ainsi que des améliorations et perspectives possibles.

La troisième et dernière partie de cette thèse porte sur l'étude de la DFC. Cette dernière étant une notion encore peu connue, le premier chapitre se charge de placer ce sujet d'étude au sein de la recherche en connectivité fonctionnelle en y montrant les problématiques associées et les applications cliniques envisagées. Avant de se pencher sur le cœur des développements de cette thèse, les données ont été longuement analysées et les observations et résultats préliminaires qui en découlent constituent le chapitre suivant. Les derniers chapitres se concentrent sur nos contributions pour évaluer la DFC. Le chapitre 10 expose les modélisations développées basées sur les MMC multidimensionnels (MMC-Produit et MMC-Coulés). Le dernier chapitre présente la mise en œuvre de ces modélisations sur les deux études dont nous avons parlé : étude pharmacologique au niveau sujet, étude de la démence à corps de Lewy au niveau du groupe.

Première partie

Cadre de travail

Introduction	3
1 IRM fonctionnelle et réseaux de repos	5
2 Analyse en Composantes Indépendantes en IRMf de repos	19
3 Analyse markovienne cachée et processus en interaction : une synthèse	27
4 Bases de données utilisées	37
Conclusion partielle	41

Introduction

Avant de présenter les travaux de cette thèse, il convient d'introduire les concepts fondamentaux qui seront utilisés dans les différentes parties qui vont suivre. Le premier chapitre expose ainsi le cadre général de nos études : l'imagerie par résonance magnétique fonctionnelle (IRMf). Seront exposés dans ce chapitre, aussi bien la physique de base régissant cette technique, que le principe de connectivité fonctionnelle obtenue par cette approche. Nous introduirons également les réseaux cérébraux de repos, cellule élémentaire clé étudiée dans cette thèse. Ces réseaux cérébraux de repos sont mis en évidence par une méthode d'analyse de données appelée ACI pour analyse en composantes indépendantes. Cette technique est présentée dans le chapitre 2. Le chapitre 3 introduit les modèles markoviens cachés, modèles probabilistes qui sont notre angle d'approche pour analyser la dynamique de la connectivité fonctionnelle. Enfin, le dernier chapitre de cette partie présente les bases de données utilisées pour les différentes études de ce travail de thèse.

IRM fonctionnelle et réseaux de repos

1

Ce chapitre est consacré aux rappels anatomiques et fonctionnels de la structure cérébrale et au concept de connectivité. L'imagerie par résonance magnétique fonctionnelle, aussi bien les principes physiques et biologiques qui la régissent que les principales étapes d'une étude via cette technique, seront présentés par la suite. Enfin, la dernière section de ce chapitre permettra d'exposer les réseaux cérébraux de repos, sujets d'étude de cette thèse.

1.1 Pré-requis anatomiques et fonctionnels de l'organisation cérébrale

Le cerveau, siège des facultés mentales, reste encore un vaste sujet d'études de part sa complexité et la diversité des méthodes pour appréhender son fonctionnement. Cette section se limitera à un rappel des éléments auxquels nos travaux font référence.

Anatomie et fonction sont très liées en ce qui concerne l'organisation du cerveau. Cependant, afin d'introduire les principales notions nécessaires à la compréhension de la suite des travaux, ces deux aspects sont présentés séparément.

1.1.1 Cerveau et structure connexes : point de vue anatomique

Avec le cervelet et le tronc cérébral, le cerveau forme le système nerveux central qui contrôle en permanence le fonctionnement de l'organisme. Celui-ci reçoit et analyse les informations provenant de l'intérieur (*e.g.* tension des muscles) ou de l'extérieur (*e.g.* stimulus sensoriel) et commande la réponse la mieux adaptée. Nous nous intéresserons plus particulièrement, dans la description qui va suivre, aux structures cellulaires et macrocellulaires responsables du traitement de ces informations.

Au niveau cellulaire, le cerveau est constitué entre autres de cellules nerveuses, les neurones. Ils sont entourés par un tissu de soutien composé de différents types cellulaires que l'on regroupe sous le nom de cellules gliales. Ce sont les neurones qui intègrent et transfèrent l'information sous forme d'influx nerveux électriques. Ils sont formés d'un corps cellulaire et de multiples prolongements : les corps cellulaires sont regroupées en couches et forment la partie externe du cortex appelée substance grise (ils peuvent également être regroupés en noyaux et former ainsi les noyaux dits sous-corticaux), le principal prolongement est appelé axone et sert à la conduction de l'influx nerveux. Ces axones constituent la substance blanche qui est la partie interne du cortex. Regroupés en faisceaux, ils établissent des connexions avec d'autres groupes de neurones.

L'organisation du cerveau se reflète aussi au niveau macrocellulaire. D'un point de vue anatomique, le cerveau se compose de deux hémisphères chacun partagé en quatre lobes externes : frontal, pariétal, temporal et occipital ; et un interne : l'insula (voir [FIGURE 1.1](#)). Le corps calleux,

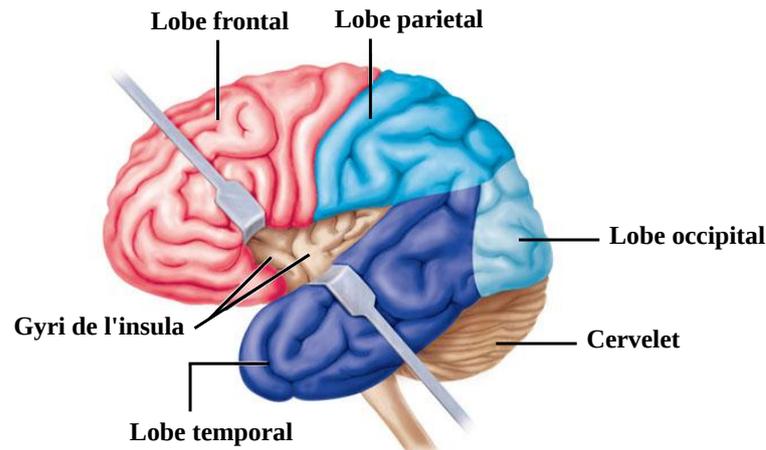
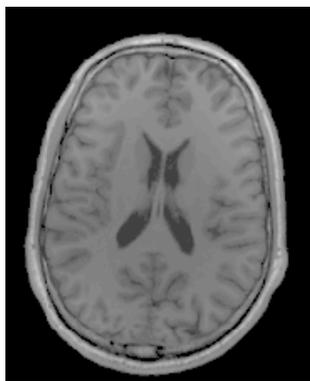


FIGURE 1.1 – Les lobes cérébraux externes, l’insula et les structures avoisinantes. En repoussant les lobes frontal et temporal le long du sillon latéral, il est possible d’apercevoir les gyri de l’insula (d’après [1] modifiée).



IRM cérébrale coupe axiale, pondération T1



segmentation de la matière grise à partir de l'image T1



segmentation de la matière blanche à partir de l'image T1

FIGURE 1.2 – Images anatomiques du cerveau à partir de l’IRM : mise en évidence des différentes structures du cerveau.

un faisceau d’axones, relie les hémisphères en assurant le transfert d’informations et permet la coordination entre les parties du cerveau. Afin de préserver ce système, le cerveau possède plusieurs niveaux de protection : le crâne, les méninges (membranes situées entre le crâne et la couche superficielle du cerveau, le cortex) et le liquide céphalo-rachidien. Ce fluide circule entre les couches des méninges et à travers une série de cavité appelées ventricules. Ces différentes couches sont visibles grâce à l’imagerie par résonance magnétique (IRM) (voir FIGURE 1.2).

1.1.2 Physiologie du fonctionnement cérébral

Deux principes fondamentaux permettent d’appréhender le fonctionnement cérébral : la ségrégation et l’intégration. Ils mettent en avant l’existence d’aires fonctionnelles, zones cérébrales sous-tendant des fonctions élémentaires différentes, qui peuvent interagir pour la réalisation d’une tâche.

La ségrégation fonctionnelle définit des aires fonctionnelles cérébrales comme étant spécialisées pour des aspects moteurs, sensitifs ou cognitifs : on suppose qu’une aire cérébrale réalise un traitement homogène de l’information en fonction de ses entrées et de ses sorties. Le fondement de ce principe vient du fait que des colonnes neuronales proches sont très connectées entre elles

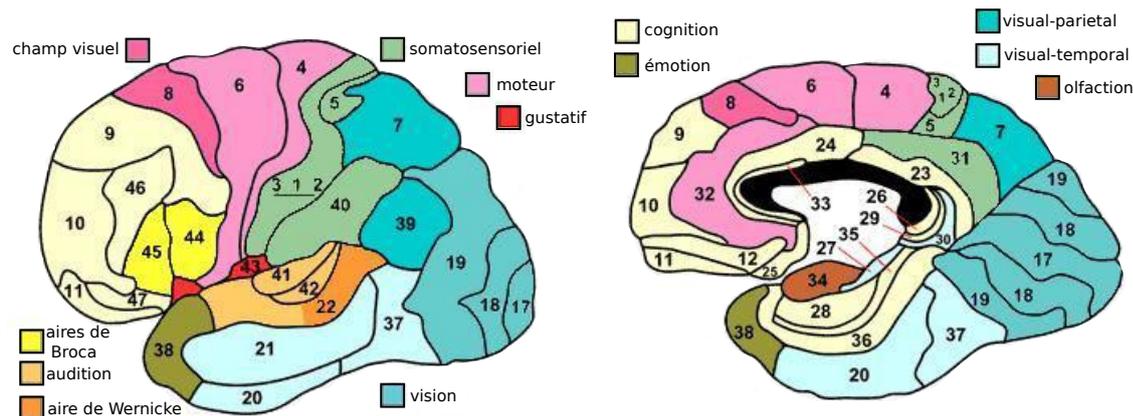


FIGURE 1.3 – Les aires définies par Brodmann et quelques fonctions associées.

et tendent donc à avoir des fonctions voisines. La finesse de la ségrégation fonctionnelle observée détermine la parcellisation du cortex, toujours sujette à de nombreuses études. À titre d'exemple, la parcellisation en aires de Brodmann [2] (voir FIGURE 1.3), couramment utilisée en neurosciences, est faite selon l'organisation des couches corticales mais ne correspond pas toujours à une réalité fonctionnelle. Pour faire le lien avec les structures anatomiques, nous pouvons citer quelques exemples d'aires fonctionnelles attribuées aux différents lobes. On retrouvera dans le lobe frontal, les aires relatives à la personnalité, la prise de décision, le contrôle des mouvements volontaires. Le lobe pariétal regroupe des aires sensibles qui reçoivent les stimuli provenant des jambes, bras, bouche etc. Le lobe temporal concentre aussi bien les aires des quatre autres sens que celles chargées de leur interprétation. Le lobe occipital quant à lui, est le siège de la vue. Enfin le système limbique réunit, entre autres, le système de régulation des émotions et les aires impliquées dans la formation de la mémoire.

Selon le principe de ségrégation fonctionnelle, chaque région élémentaire traite un type d'information très spécifique. Cette hypothèse a donné naissance au courant localisationniste qui ne prenait donc pas en compte les interactions possibles entre les aires. Pourtant, pour permettre une réponse adaptée de l'organisme, une information a besoin d'être contextualisée. Les informations provenant de plusieurs aires cérébrales doivent être intégrées en un ensemble cohérent pour que la réponse comportementale ou cognitive soit pertinente. C'est cette interaction entre les aires qui est caractérisée par l'intégration fonctionnelle. En plus des aires primaires (aires sensorielles ou motrices qui reçoivent directement les informations de structures sous-jacentes), secondaires et tertiaires (aires qui reçoivent les informations des aires primaires et qui ont la capacité d'intégrer et d'analyser ces informations), on définit aussi les aires associatives. C'est à ce niveau que se déroule l'intégration d'informations très élaborées mettant en jeu des inférences sensorielles, motrices et intellectuelles. Cette interaction cérébrale est la base de la notion de connectivité qui sera abordée dans la section suivante.

Le traitement d'un seul stimulus fait contribuer des milliards de connexions entre des cellules nerveuses hautement spécialisées, la ségrégation et l'intégration étant complémentaires pour définir le fonctionnement cérébral.

1.2 Connectivité

1.2.1 Définition

L'interaction entre les aires cérébrales permet l'étude du cerveau en tant que réseau justifiant l'émergence de la notion de connectivité. C'est un élément majeur en neurosciences, où la recherche

des liens entre structure et fonction reste en développement constant. Dans ce cadre, trois types de connectivité cérébrale sont à considérer : la connectivité structurelle, la connectivité effective et la connectivité fonctionnelle. Il n'existe donc pas à proprement parler une et une seule connectivité cérébrale, mais des connectivités.

La connectivité structurelle, ou anatomique, relève simplement d'une relation physique entre les neurones via les faisceaux de fibres qui constituent la matière blanche. Ces liens anatomiques peuvent être mis en évidence, par exemple, par l'imagerie du tenseur de diffusion (DTI), une technique basée sur l'IRM.

La connectivité effective permet d'étudier l'influence directe d'une aire cérébrale sur une autre. On s'intéresse alors aux interactions causales relatives à une tâche ou à des flux d'informations. En IRM, les mesures de connectivité effective dépendent en majorité d'un modèle déterministe et sont réduites à un ensemble de régions. Les principales techniques permettant cette analyse sont la modélisation en équations structurelles (SEM pour *structural equation modeling*, [3]) et les modèles causaux dynamiques (DCM pour *dynamic causal modeling*, [4]).

Enfin, la connectivité fonctionnelle mesure la corrélation temporelle entre des aires cérébrales spatialement distribuées [5]. Contrairement à la connectivité effective, on peut étudier ces dépendances statistiques sur l'ensemble du cerveau simultanément et de manière exploratoire. Sous l'hypothèse que le cerveau reste dans un état stable au cours de l'acquisition, la connectivité fonctionnelle peut être étudiée par de nombreuses méthodes [6], les plus utilisées étant les méthodes d'ensemencement (*seed base voxel*), les graphes ou encore l'analyse en composantes indépendantes. C'est cette méthode qui sera au cœur des travaux détaillés dans ce manuscrit.

1.2.2 Dynamique de la connectivité fonctionnelle

Les approches univariées et multivariées mentionnées ci-dessus supposent que le niveau d'interactions entre les régions est constant dans le temps et ce dans le cas où aucune tâche n'est requise durant l'examen. Par exemple, la méthode d'ensemencement représente la relation entre deux régions d'intérêt, exprimée par un coefficient de corrélation unique calculé à partir d'un signal temporel de l'ensemble de l'acquisition. De même, l'analyse spatiale en composantes indépendantes, bien qu'intégrant l'ensemble des interactions simultanément, est toujours réalisée sur toute la durée du signal enregistré. Les variations temporelles de ces interactions et corrélations ne seront pas perçues. Si cette hypothèse a été faite pour faciliter la mise en œuvre des techniques de mesure de connectivité fonctionnelle, les nouvelles approches tendent depuis peu à caractériser la dynamique de la connectivité fonctionnelle. Les fluctuations de cette connectivité fonctionnelle, ainsi que les méthodes employées pour les analyser, seront largement développées dans la troisième partie de ce manuscrit. Néanmoins, il convient à ce stade de commencer à appréhender, de manière graphique, en quoi consiste ces variations. C'est l'objectif de la FIGURE 1.4 qui représente les modifications possibles de la connectivité fonctionnelle, entre deux réseaux A et B et entre les aires cérébrales qui les composent (points gris). Ces variabilités sont observées en mesurant cette connectivité sur différentes fenêtres temporelles (fenêtres 1, 2 et 3, colonnes de droite) de durée inférieure à celle globale (colonne de gauche). La connectivité fonctionnelle est ici représentée par des liens rouges (un certain niveau de corrélation entre les décours temporels sur la fenêtre considérée), ou des liens bleus (un certain niveau d'anti-corrélation entre les décours temporels sur la fenêtre considérée). La force de cette connectivité fonctionnelle entre deux régions peut varier en amplitude (trait plus ou moins épais), en signe (trait rouge ou bleu) ou, du fait de ces deux cas, une région peut être associée/dissociée d'un réseau suite à un gain/perte d'amplitude de la relation.

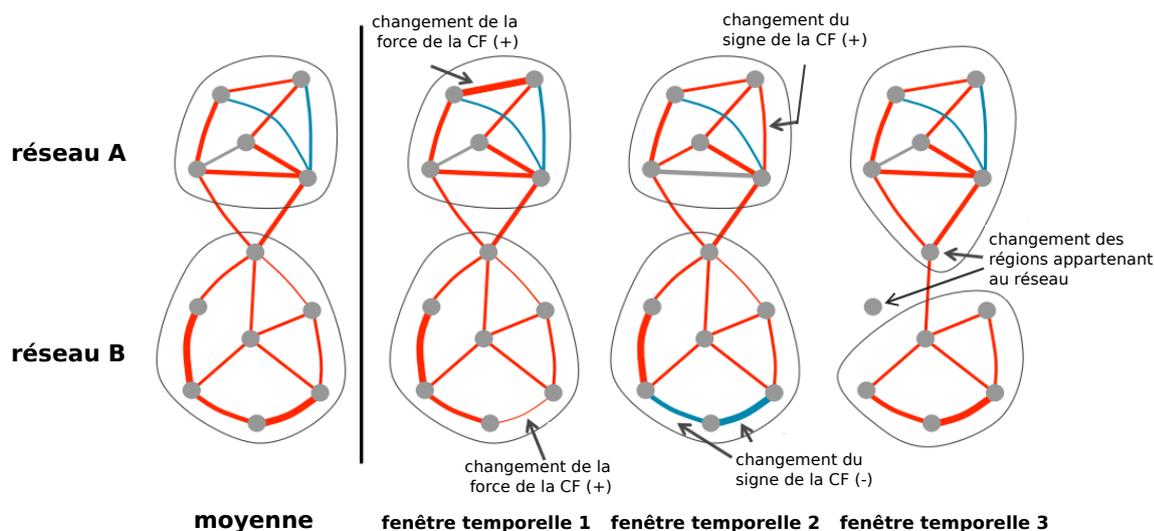


FIGURE 1.4 – Illustration de la dynamique de la connectivité fonctionnelle sur différentes fenêtres temporelles (colonnes 2, 3, 4) par rapport à celle obtenue sur l'ensemble de la session (colonne 1). La force de la connectivité fonctionnelle entre deux régions peut varier en amplitude (colonne 2), en signe (colonne 3) ou du fait de ces deux cas, une région peut être associée/dissociée d'un réseau suite à un gain/perte d'amplitude de la relation (colonne 4). En rouge : connexions positives ; en bleu : connexions négatives. (d'après [7] modifié).

1.3 IRM fonctionnelle

1.3.1 Imagerie par Résonance Magnétique

L'imagerie par résonance magnétique (IRM) est une modalité d'imagerie non-invasive largement utilisée en clinique comme en recherche grâce à ses nombreux domaines d'application. Elle permet de visualiser, entre autres, aussi bien les tissus mous du corps humain avec une haute résolution spatiale et un grand contraste tissulaire (imagerie morphologique), que l'activité et la connectivité cérébrale (imagerie fonctionnelle), les flux (angiographie par résonance magnétique), les faisceaux de fibres cérébrales (imagerie du tenseur de diffusion) ou encore le débit sanguin cérébral (imagerie de perfusion).

Cette méthode d'imagerie repose sur le principe de résonance magnétique nucléaire. La complexité des phénomènes physiques et mécaniques impliqués rend difficile une synthèse exhaustive, il sera donc simplement détaillé ci-dessous, les fondements de cette technique permettant la création d'une image.

1.3.1.1 Principe physique de la résonance magnétique nucléaire

La résonance magnétique nucléaire (RMN) va utiliser les propriétés magnétiques de certains atomes comme l'hydrogène, le fluor ou le carbone 13 par exemple. En pratique clinique, c'est l'imagerie du noyau d'hydrogène, constitué d'un proton, qui va être privilégiée en raison de la présence de ce dernier en grande quantité dans le corps humain (dans les molécules d'eau principalement). Celui-ci possède un moment magnétique de spin qui peut être représenté schématiquement par un vecteur en rotation sur lui-même. À l'état de repos, l'orientation de ces vecteurs est aléatoire ce qui induit une résultante magnétique de l'ensemble des protons nulle. Trois étapes centrales permettent l'émission du signal nécessaire pour former l'image : la polarisation, l'excitation et la relaxation.

En plaçant les protons dans un champ magnétique intense, appelé B_0 , les spins vont s'orienter suivant l'axe du champ (alignement selon \vec{B}_0 et $-\vec{B}_0$) : c'est la polarisation. La résultante magnétique \vec{M}_0 n'étant plus nulle et la proportion de spins dans le sens de \vec{B}_0 étant plus grande que celle des spins antiparallèles, \vec{M}_0 va être dans la même direction que \vec{B}_0 . À ce moment magnétique, va se superposer un moment cinétique qui permet aux spins de décrire un cône autour de l'axe de B_0 . Cette rotation, appelée mouvement de précession (voir FIGURE 1.5), est caractérisée par sa fréquence f_0 , la fréquence de Larmor. Elle est dépendante de l'intensité du champ magnétique B_0 selon la relation $f_0 = \frac{\gamma \cdot B_0}{2\pi}$. Le moment magnétique des protons peut être décomposé en une composante longitudinale et une composante transversale. Le mouvement de précession se traduit ainsi par une rotation de la composante transversale autour de l'axe longitudinal. À l'équilibre, la résultante de l'aimantation longitudinale \vec{M}_0 s'alignera avec \vec{B}_0 tandis que celle de l'aimantation transversale sera nulle puisque les spins ne précessent pas en phase.

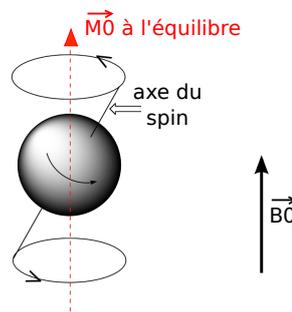


FIGURE 1.5 – Mouvement de précession du spin placé dans un champ magnétique B_0 .

La deuxième étape du processus de RMN est l'excitation. Elle repose sur le phénomène de résonance dû à l'application d'une onde radiofréquence de fréquence f_0 . Cet apport d'énergie aux spins va écarter \vec{M}_0 de l'axe \vec{B}_0 . Si l'on décompose l'aimantation résultante en une composante longitudinale (selon l'axe Z, dans l'axe de \vec{B}_0) et une composante transversale (dans le plan XY, orthogonal à \vec{B}_0), une diminution de la composante longitudinale et une augmentation de la composante transversale sont observées lors de cette bascule de \vec{M}_0 .

Enfin, lorsque l'émission de radio-fréquence est interrompue, le système va retourner à l'équilibre : c'est la relaxation. Le signal RMN enregistré est généré par l'émission d'énergie liée à cette étape de relaxation. Des caractéristiques telles que les temps de relaxation longitudinale (T1) et transversale (T2) déterminent respectivement le retour à l'équilibre longitudinale et le déphasage des spins et peuvent être mesurées. Les temps de relaxation sont différents pour chaque tissu et le contraste des images obtenues par IRM dépendra entre autres de ces temps de relaxation et des méthodes d'acquisition des images (durée de l'impulsion radio-fréquences, temps entre les impulsions, gradients etc.).

1.3.1.2 Formation d'une image par résonance magnétique

L'utilisation de la RMN pour la production d'une image, nécessite un codage spatial du signal. Celui-ci est réalisé par l'ajout de gradients de champ magnétique linéaires dans les trois axes de l'espace. Ainsi, l'intensité du champ magnétique le long d'un axe de l'espace va varier faiblement mais linéairement du fait de l'application de ces gradients. La vitesse de rotation des spins, et la phase, selon ce même axe dépendra de la position spatiale. L'application d'impulsions de radio-fréquences avec des gradients de champs magnétique linéaires variant temporellement, permettent de remplir l'espace fréquentiel (aussi appelé espace K) de l'objet dont on souhaite réaliser l'image. L'image de l'objet étudié sera reconstruite par une transformée de Fourier inverse.

Selon la séquence d'acquisition choisie (choix de la durée, de l'application des impulsions de

radio-fréquences, de l'application des gradients de champs), les tissus peuvent être observés avec des contrastes différents dépendant donc des temps de relaxation T1 et T2. Ces modifications des paramètres de la séquence permettent d'imager également d'autres caractéristiques des tissus comme la diffusion des molécules d'eau, la susceptibilité magnétique ou encore l'activité cérébrale.

1.3.2 Fondements physiologiques en IRM fonctionnelle

L'observation et la mesure du fonctionnement du cerveau ou tout du moins de son activité peuvent découler de l'activité neuronale (via l'électroencéphalographie, EEG ou la magnéto-encéphalographie, MEG), l'activité métabolique (via la tomographie par émission de positron, TEP) ou encore du système cérébrovasculaire. Ce dernier peut être observé par l'imagerie par résonance magnétique fonctionnelle (IRMf).

1.3.2.1 Imagerie fonctionnelle par contraste BOLD

Principe physique L'IRM fonctionnelle (IRMf) est une imagerie indirecte de l'activité cérébrale, basée sur la détection des variations locales de flux et d'oxygénation du sang secondaires à l'activité neuronale (contraste BOLD, Blood Oxygenation Level Dependent) [9].

Lors d'une activation neuronale, les besoins tissulaires en énergie augmentent, entraînant une hausse de la consommation de glucose (+20 à +40 %) et d'oxygène (+5 à +25 %), principales sources d'énergie du cerveau. À cette réponse métabolique, s'ajoute la réponse hémodynamique qui se traduit par une augmentation du débit sanguin cérébral (+20 à +70 %) et du volume sanguin cérébral (+5 à +30 %). En réalité, ce n'est pas directement l'évolution de l'oxygène que l'on va mesurer mais les variations de la concentration tissulaire en désoxyhémoglobine (hémoglobine sous forme désoxygénée). C'est cette dernière qui est responsable de la perturbation magnétique lors de l'activité neuronale, le sang partiellement désoxygéné ayant une susceptibilité magnétique plus élevée que le sang totalement oxygéné. Ainsi, la concentration en désoxyhémoglobine augmente avec la consommation en oxygène et diminue lors de l'accroissement du débit sanguin cérébral (apport de sang riche en oxyhémoglobine). Lors de l'activation neuronale, le système vasculaire surcompense la perte de sang oxygéné par une augmentation du débit sanguin cérébral (couplage neuro-vasculaire) qui est plus forte que la consommation d'oxygène. Par conséquent, la concentration relative de désoxyhémoglobine baisse, entraînant ainsi une légère augmentation du signal IRMf BOLD. Cette faible variation nécessite donc de répéter les acquisitions dans le temps afin de permettre des études statistiques aussi bien en repos que lorsque le sujet est soumis à une tâche.

À l'heure actuelle, un volume 3D est acquis toutes les 2 ou 3 secondes généralement, les processus métaboliques et hémodynamiques, plus lents que l'activité neuronale, étant de l'ordre de plusieurs secondes. Avec un séquence d'IRMf qui dure en moyenne une dizaine de minutes, plusieurs centaines de volumes sont acquis au cours d'une session. De récentes techniques [10] peuvent fournir une résolution temporelle de l'ordre de centaines de millisecondes et présagent l'accès à des analyses de variations plus rapides. D'un point de vue spatial, c'est donc la matière grise, riche en synapse, que l'on va imager au mieux par ce procédé. La résolution spatiale est de l'ordre de quelques millimètres. Cette technique présente néanmoins quelques limites. On peut citer entre autres la distance qui sépare neurones activés et variation vasculaire pouvant entraîner une erreur d'identification de la zone fonctionnelle activée. En outre, les mouvements physiques (tête, œil, ...) ou physiologiques (respiration, battements cardiaques, ...) ou encore le choix du type d'acquisition apportent des artefacts divers.

Sauf mention contraire, on parlera dans la suite de ce manuscrit, simplement d'IRMf pour suggérer l'IRMf par contraste BOLD.

Applications En pratique, l'IRM fonctionnelle peut être utilisée aussi bien pour des études de groupe que pour des études au niveau du sujet. En étude de groupe, il est ainsi possible de détecter

les aires cérébrales impliquées par une majorité de sujets lors d'une tâche bien définie [11] ou encore de mettre en évidence des comportements cérébraux pathologiques [12]. Le cas du sujet unique a des portées plus diagnostiques ou cliniques. L'IRMf a déjà un certain impact dans des applications spécifiques telle que la caractérisation d'une maladie [13], la planification neurochirurgicale [14], l'étude de la plasticité cérébrale lors de la récupération post-opératoire ou encore dans la meilleure prédiction des résultats d'un traitement thérapeutique [15].

On distingue deux types d'étude IRMf : l'imagerie par activation et l'imagerie de repos. L'imagerie par activation va permettre de déterminer les régions impliquées lors que le sujet est soumis à une tâche. Elle repose sur la différence entre deux conditions : les activations durant la tâche par rapport aux activations pendant les tâches dites « de contrôle » ou durant les phases de repos. On mesure ainsi l'activité cérébrale non pas de manière quantitative mais relative. En outre, des études de connectivité peuvent également être réalisées lors de telles tâches.

Les travaux de ce manuscrit reposent sur l'IRMf de repos, c'est-à-dire quand aucune tâche n'est demandée au sujet [16]. Ce dernier doit simplement arrêter son regard sur un point fixe pendant l'examen, ou fermer les yeux sans dormir. Dans ce cas, seule une mesure de connectivité fonctionnelle peut rendre compte du fonctionnement cérébral. Les études sous cette condition ont de nombreux avantages de par la relative simplicité de mise en place. Elles peuvent être appliquées à n'importe quel sujet et sont donc particulièrement adaptées à l'étude des populations cliniques et pédiatriques [17] qui ne sont pas toujours en capacité de réaliser certaines tâches. Par exemple, les patients souffrant du *locked-in syndrome*, une forme de paralysie généralisée, sont conscients mais incapables d'effectuer une réponse motrice. La connectivité fonctionnelle de repos mesurée par l'IRMf révèle des réseaux spontanés qui peuvent nous aider à comprendre l'organisation fonctionnelle du cerveau [18]. Il est d'autant plus important de bien connaître l'état de repos puisqu'il est souvent l'état de référence lors d'IRMf d'activation. Ainsi, la connectivité fonctionnelle permet de réaliser une topographie fonctionnelle du cerveau au repos [19] et nous aide dans la prédiction de la localisation de réponses fonctionnelles liées à une tâche [20].

1.3.2.2 L'imagerie fonctionnelle par marquage de spins

Lorsque l'on parle d'imagerie fonctionnelle, on sous-entend souvent « imagerie par contraste BOLD ». Or l'imagerie par marquage de spins (ASL pour *Arterial Spin Labeling*), ou imagerie de perfusion, est une modalité qui permet également de rendre compte de l'activité cérébrale via la mesure du débit sanguin cérébral. L'ASL est une technique différentielle au cours de laquelle deux types d'acquisitions sont réalisées : une acquisition avec marquage magnétique des protons artériels et une acquisition de contrôle. Le marquage magnétique des protons du sang artériel en amont du volume d'intérêt est effectué par une impulsion d'inversion spatialement sélective. Les protons marqués migrent ensuite via le réseau artériel vers les tissus cérébraux à perfuser, où ils passent du compartiment vasculaire vers le compartiment extra-vasculaire. À un temps TI (temps d'inversion) après l'impulsion de marquage, l'acquisition est réalisée. L'acquisition de contrôle est par définition réalisée sans marquage, les protons artériels du volume d'intérêt sont alors à l'équilibre, complètement relaxés. La soustraction de l'acquisition de marquage et de l'acquisition de contrôle permet d'éliminer la magnétisation statique et d'obtenir une cartographie pondérée en perfusion. De là, il est possible de mesurer le débit sanguin cérébral de manière absolue [21].

Une séquence développée au sein de notre équipe [22] nous permet actuellement d'acquérir simultanément un signal BOLD et un signal ASL. L'un ou l'autre de ces signaux pouvant être récupérés ultérieurement par filtrage si besoin.

1.3.3 Principales étapes d'une étude en IRMf

1.3.3.1 Acquisition

Les séquences qui peuvent mettre en évidence le phénomène BOLD sont les séquences en écho de gradient ($T2^*$), la technique d'acquisition la plus utilisée étant l'imagerie écho-planar (ou EPI pour *Echo-Planar Imaging*). Ces séquences ont des temps d'acquisition relativement longs (le temps de répétition variant entre 2 et 3 secondes dans la majorité des études) ce qui conditionne le nombre de volumes pouvant être acquis pendant une session.

Ainsi pour une étude faite au repos, une dizaine de minutes suffisent à fournir les données nécessaires pour assurer la robustesse des résultats et prendre en compte toutes les fréquences du signal BOLD. Par contre, lorsque l'on demande au sujet de réaliser une ou des tâches (cognitives, motrices ou sensorielles), la session peut durer plusieurs dizaines de minutes. Le choix des tâches, modélisé par le paradigme d'activation doit pouvoir mettre en évidence l'activité cérébrale recherchée en opposant différents états cognitifs (le repos peut en être un). La répartition temporelle des diverses tâches et moment de repos durant la session est préparée avec précision en amont. Il s'agit d'éviter l'anticipation ou habitude du sujet tout en s'assurant d'une durée d'examen confortable pour le sujet mais suffisante pour acquérir le nombre de volumes nécessaires à la robustesse de l'analyse statistique qui suivra.

L'acquisition du volume 3D ne se fait pas en une fois mais via l'acquisition successive des coupes qui le composent. Cette acquisition peut être continue ou entrelacée (les coupes paires puis impaires ou inversement) et impliquer des décalages temporels qui devront être pris en compte pour la suite des traitements.

1.3.3.2 Traitements des images

Qu'il s'agisse d'imagerie par activation ou de repos, une étude IRMf repose sur une étape importante de pré-traitement des images avant leur traitement statistique. L'ordre et les étapes de ces pré-traitements peuvent varier selon le type d'acquisition réalisée. Les pré-traitements effectués sur nos bases d'images sont :

- a) Correction du décalage temporel entre les coupes dite *slice-timing correction* [23] afin de prendre en compte l'acquisition multi-coupes (entrelacée) quand c'est le cas.
- b) Correction des distorsions spatiales dues à l'inhomogénéité locale du champ magnétique B_0 à partir de la carte de champ acquise durant l'examen. En effet, des distorsions spatiales dues à l'inhomogénéité locale du champ B_0 impliquent des artefacts sur les régions avec un fort gradient de susceptibilité magnétique (distorsion et perte de signal aux interfaces avec les os, l'air, les hématomes ou en post-opératoire).
- c) Correction des mouvements de tête du sujet : les mouvements de tête, même très légers, peuvent entraîner de fausses activations au niveau des cartes observées par la suite ; un recalage rigide entre les images acquises permet de s'affranchir de bon nombre d'artefacts de mouvements de la tête.

Des pré-traitements supplémentaires sont réalisés lors des études de groupes :

- d) Recalage des images fonctionnelles sur une image morphologique (image pondérée T1).
- e) Normalisation de ces images, c'est-à-dire recalage non rigide dans un espace commun. Cet espace peut être propre au groupe de sujets étudiés ou issu d'une base de données partagée par la communauté d'IRMf (comme l'espace MNI [24]). Cette normalisation, nécessaire pour les études de groupes, est effectuée sur des critères morphologiques mais ne garantit pas une précision parfaite dans la mise en correspondance des régions fonctionnelles entre les individus.
- f) Lissage spatial pour réduire les variabilités anatomiques et fonctionnelles entre sujets.

1.3.3.3 Analyse de la connectivité

Ces pré-traitements effectués, l'analyse de la connectivité peut être réalisée. Il s'agit dans cette étape de mettre en évidence des aires cérébrales corrélées temporellement, aussi appelées « réseaux fonctionnels cérébraux ». Pour extraire ces réseaux, différentes méthodes sont disponibles ; leur utilisation dépend du type d'étude réalisée (groupe ou sujet unique, sujet sain ou malade, repos ou activation). Les plus courantes sont : le modèle linéaire généralisé (GLM) quand on souhaite étudier l'activation que l'on a modélisée ; la connectivité par *seed-based* où l'on étudie la connectivité d'une région « graine » avec le reste du cerveau ; l'analyse en composantes indépendantes (ACI) qui permet une étude plus large du cerveau, y compris au repos. L'ACI étant l'analyse choisie pour les études de cette thèse, elle sera détaillée dans le chapitre suivant. Cette analyse est faite au niveau du sujet et/ou du groupe afin de mettre en avant la présence ou l'absence de certains réseaux, l'altération de la connectivité entre certaines aires cérébrales, la réponse cohérente à la tâche demandée ou encore des différences inter-sujets ou inter-sessions.

1.4 Réseaux cérébraux spontanés - SAM

1.4.1 Définition

On définit les cartes de co-activités spatiales (ou SAM pour *Spontaneous co-Activity Map*) comme des fluctuations conjointes, significatives et spontanées exhibant des réseaux spécifiques au repos [18, 25, 26, 27]. Cette appellation est large car elle regroupe aussi bien des réseaux communs à une population, les RSN pour *resting-state networks*, que ceux propres à un sujet ou à une session IRMf, les réseaux idiosyncratiques.

1.4.2 RSN : Réseaux communs mis en évidence en IRM fonctionnelle de repos

La plupart des études convergent vers des réseaux dont la configuration spatiale est commune d'un individu à l'autre, même à travers une variété d'états comportementaux (par exemple, les yeux fermés, ouverts, ou ouverts et fixateur [28, 29]). De même, entre les sujets sains et malades, on retrouve des réseaux communs bien qu'il existe des différences subtiles mais importantes entre les réseaux obtenus sur les sujets sains et ceux obtenus sur les patients [30, 31, 32]. Si on analyse au niveau du sujet, on remarque qu'une grande partie des SAM sont des RSN (environ 70 %) [33]. Cette reproductibilité permet, entre autres, les études de groupes et comparaisons entre sujets.

Il existe des classifications des RSN obtenus par ACI ainsi que des études sur leur reproductibilité [33, 34] mais nous présentons, ci-dessous, uniquement les plus connus. Pour certains RSN très étudiés, le nom est commun dans la communauté scientifique. C'est le cas du DMN par exemple qui fut l'un des premiers réseaux identifiés, d'où le nom de « réseaux par défaut » ou « Default Mode Network ». Pour la plupart des RSN, les appellations varient selon que les équipes désignent les réseaux par leur fonction ou par leur localisation. Les fonctions de certains réseaux n'étant pas toujours communément acceptées, leur appellation peut conduire à certaines confusions. C'est pourquoi nous présenterons ici les abréviations utilisées par notre équipe qui sont basées sur la localisation des régions cérébrales présentes dans le RSN [35]. Les réseaux suivant peuvent être caractérisés par :

- ▷ l'orientation de la cognition (Ori) :
 - tournée vers l'intérieur (I) : mémoire autobiographique, projection dans l'avenir, introspection...
 - ou vers l'extérieur (E) : mémoire de travail, perception, langage, attention
- ▷ le niveau (Niv) :
 - bas niveau (B) : pour les RSN incluant des aires primaires ou secondaires,
 - haut niveau (H) : pour les RSN incluant des aires tertiaires et associatives.

Le TABLEAU 1.1 liste les RSN identifiés par notre équipe. La FIGURE 1.6 représente ces réseaux dans l'espace normalisé MNI.

TABLEAU 1.1 – Abréviations, fonctions et caractéristiques des RSN les plus communs.

Abréviation	Nom anglais	Fonction(s)	Ori	Niv
ACFPN	Anterior Cingulo- Fronto-Polar Network	Cognitive branching	E/I	H
BG	Basal Ganglia			B
CN	Central Network	Traitement somatosensori-moteur	E	B
DMN	Default Mode Network	Initialement, état dans lequel le cerveau se mettrait lorsqu'il n'est pas engagé dans une tâche dirigée vers un but. Mémoire autobiographique...	I	H
ETN	External biTemporal Network	Multimodal significance extraction (language, facial and bodily expression...)	E	H
LFPN	Left Fronto-Parietal Network	Contrôle exécutif	E	H
MON	Median Occipital Network	Traitement visuel	E	B
OPN	Occipital Posterior Network	Traitement visuel	E	B
OPFN	Occipito-Parieto-Frontal Network	Traitement visuel	E	B
PN	Precuneal Network	Awareness (conscience primaire)	I	H
PCN	Posterior Cingulate Network	Awareness (conscience primaire)	I	H
RFPN	Right Fronto-Parietal Network	Contrôle exécutif	E	H

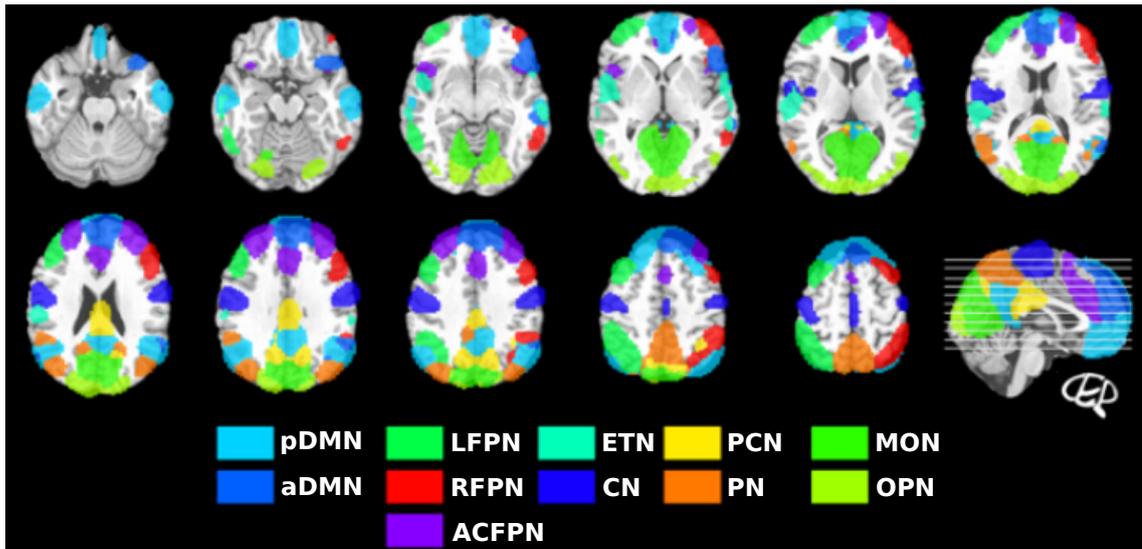


FIGURE 1.6 – Illustrations des principaux RSN.

Le DMN est ici scindé en deux réseaux : les DMN antérieur (aDMN) et postérieur (pDMN).
(d'après [35] modifié)

1.4.3 Réseaux idiosyncratiques

Les réseaux idiosyncratiques correspondent à une co-activité cérébrale propre à un sujet voire à une session. Ces SAM particulières peuvent résulter d'une activité mentale momentanée et, dans ce cas, il sera peu probable de les observer lors d'une autre acquisition. Elles peuvent à l'inverse être issues d'une topologie cérébrale spécifique, suite à une lésion ou une intervention chirurgicale par exemple. Elles sont alors potentiellement reproductibles d'une session à une autre et se manifestent par une configuration spatiale non connue ou partiellement partagée avec un RSN. Ce dernier cas est reconnaissable pour les RSN bilatéraux que l'on retrouve de manière clairement unilatérale. La FIGURE 1.7 illustre un de ces réseaux.

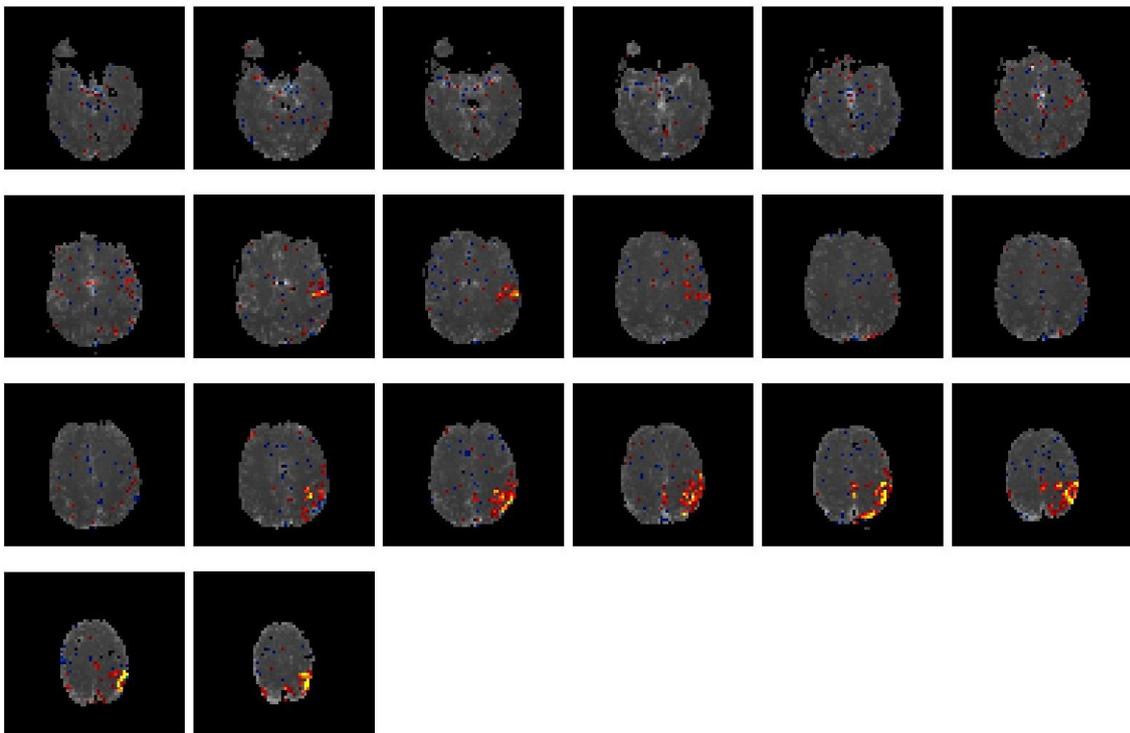


FIGURE 1.7 – Exemple d'un réseau idiosyncratique chez un patient cérébrolésé.

Analyse en Composantes Indépendantes en IRMf de repos

2

2.1 Méthode

2.1.1 Introduction

Des travaux communs au traitement du signal et à la neurobiologie ont permis à Jutten de mettre en évidence l'existence probable d'un système de séparation de sources dans le système nerveux central. Ses travaux de 1987 [36] étudient les réponses musculaires émises en fonction de différentes sortes d'excitation et soulèvent le problème de mélange de sources indépendantes. Formalisée dans le cadre linéaire par Comon en 1994 [37], l'analyse en composantes indépendantes, ou ACI (ICA en anglais), a été depuis fortement étudiée et développée ; elle reste une méthode phare en traitement du signal pour la séparation de sources à l'aveugle.

Le but de cette méthode est d'extraire, à partir d'observations qui peuvent être de différentes origines (signaux électriques, signaux sonores, images, etc.), des sources reflétant des informations « aussi indépendantes que possible ». On considère que les observations que l'on a à traiter ne sont en fait que des superpositions de processus sous-jacents ; l'idée est ici de rechercher une transformation linéaire qui minimise la dépendance statistique entre les sources finales. Ces sources, appelées composantes, sont supposées être statistiquement indépendantes ; c'est-à-dire que les valeurs d'une source ne nous donnent pas d'information par rapport aux valeurs des autres sources. Cette hypothèse d'indépendance statistique est très importante pour nous permettre une bonne séparation des composantes. On considère, en un sens, que les sources extraites sont générées par des processus physiques autonomes les uns par rapports aux autres. Pour mieux cerner les objectifs de cette technique, deux exemples sont proposés ci-dessous.

a - exemple ACI temporelle : Un exemple souvent utilisé pour illustrer l'analyse en composantes indépendantes est le cas d'enregistrements de voix au cours d'une soirée. Imaginons une salle avec N convives. P microphones sont disséminés dans la salle et enregistrent les conversations. Les enregistrements représentent des mélanges des différentes voix captées par les microphones aux alentours. L'ACI permet de retrouver, à partir de ces signaux sonores, les discours de chaque personne présente. Si aucune connaissance des voix au préalable n'est nécessaire, quelques hypothèses doivent être faites pour mener à bien cette analyse. Les discours sont supposés indépendants (temporellement) les uns des autres (les convives ne parlent pas à l'unisson (cas d'une chorale par exemple)) et le nombre d'enregistrements (P) doit être supérieur ou égal au nombre de voix (N).

b - exemple ACI spatiale : Le second exemple détaillé ici, permet de mieux appréhender le concept de l'ACI spatiale qui est celui utilisé dans le cadre de nos études en IRMf. Prenons le cas de

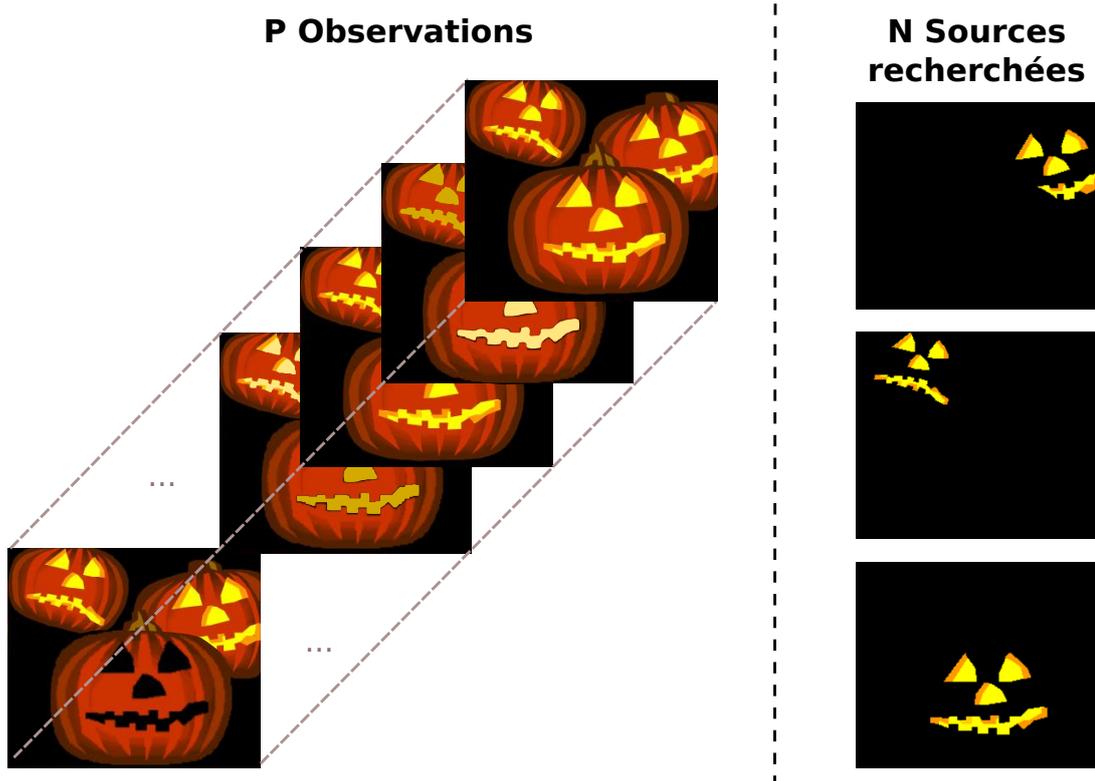


FIGURE 2.1 – Illustration de l’ACI spatiale. (Exemple d’après Calhoun modifié).

P photographies représentant N citrouilles ajourées, contenant chacune une bougie allumée, comme présentée sur la FIGURE 2.1. L’intensité de la lumière émise par la bougie varie dans le temps, donnant des photographies légèrement différentes. L’ACI va isoler, dans chaque composante, les aires éclairées correspondant à une citrouille. Les citrouilles n’étant pas disposées à la même place sur les photographies, elles répondent donc bien à l’hypothèse d’indépendance spatiale. L’analyse requiert cependant un nombre de photographies (P) supérieur ou égal au nombre de citrouilles (N).

Ces deux illustrations soulignent déjà deux points importants de la méthode à savoir : l’hypothèse d’indépendance des sources (indépendance temporelle dans l’exemple a et spatiale dans l’exemple b) et un nombre d’observations supérieur ou égal au nombre de sources (connues ou supposées).

2.1.2 Formalisation du problème

On se placera par la suite dans le contexte d’un mélange linéaire où le nombre de sources et le nombre d’observations sont les mêmes. Le modèle s’écrit donc :

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}, \quad (2.1)$$

avec

- \mathbf{X} un vecteur aléatoire de N observations tel que $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)^\top$ et $\mathbf{X}_i = (X_i^1, X_i^2, \dots, X_i^M)$, M étant la dimension des observations ;
- \mathbf{S} un vecteur aléatoire de N sources tel que $\mathbf{S} = (\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_N)^\top$ et $\mathbf{S}_i = (S_i^1, S_i^2, \dots, S_i^M)$, M étant la dimension des sources ;

- A une matrice $N \times N$, appelée matrice de mélange. Elle regroupe les pondérations des différentes sources qui permettent d’aboutir aux observations dans le cas d’un mélange linéaire.

Le but de l’ACI est de trouver la matrice de séparation W telle que $W = A^{-1}$ qui donnera \mathbf{Y} , la meilleure approximation de \mathbf{S} . Ainsi le modèle peut s’écrire sous la forme :

$$\hat{\mathbf{S}} = \mathbf{Y} = W \cdot \mathbf{X} \quad (2.2)$$

De manière vectorielle, on aura :

$$\begin{pmatrix} X_1^1 & X_1^2 & \dots & X_1^M \\ X_2^1 & X_2^2 & \dots & X_2^M \\ \vdots & \vdots & \dots & \vdots \\ X_N^1 & X_N^2 & \dots & X_N^M \end{pmatrix} = \begin{pmatrix} a_1^1 & a_1^2 & \dots & a_1^N \\ a_2^1 & a_2^2 & \dots & a_2^N \\ \vdots & \vdots & \dots & \vdots \\ a_N^1 & a_N^2 & \dots & a_N^N \end{pmatrix} \begin{pmatrix} S_1^1 & S_1^2 & \dots & S_1^M \\ S_2^1 & S_2^2 & \dots & S_2^M \\ \vdots & \vdots & \dots & \vdots \\ S_N^1 & S_N^2 & \dots & S_N^M \end{pmatrix} \quad (2.3)$$

Les coefficients (a_i^j) sont estimés suite à l’optimisation de W et représentent l’implication de chaque source dans les signaux d’origine.

Afin d’obtenir des composantes linéaires qui soient aussi « indépendantes que possibles », il faut minimiser la dépendance entre les sources \mathbf{S}_i . Différentes approches existent pour résoudre ce problème. Elles sont basées sur les propriétés des signaux sources et des signaux d’observations, à savoir :

- ▷ l’indépendance : bien que les signaux sources soient indépendants, leurs mélanges ne le sont pas puisque par définition, ce sont des mélanges issus des mêmes signaux ;
- ▷ la normalité : les signaux mélangés tendent à avoir des histogrammes plus gaussiens que n’importe quels signaux sources qui le composent ;
- ▷ la complexité : un signal mélangé a une complexité supérieure ou égale à celle la plus simple des signaux sources qui le constituent.

Les stratégies de résolution vont donc utiliser l’une de ces propriétés, les plus communes étant les algorithmes recherchant à minimiser l’information mutuelle et ceux visant la maximisation de la non-gaussianité des signaux sources. La méthode privilégiée pendant cette thèse est appelée « Infomax », pour « information-maximisation », et repose sur la maximisation de l’entropie [38]. L’optimisation de W se fait en recherchant la maximisation de l’entropie (dans le formalisme de Shannon) de $\mathbf{Y}(= \hat{\mathbf{S}})$ selon la dimension spatiale en ce qui nous concerne.

2.2 Application en IRMf de repos

2.2.1 Généralités

L’analyse en composantes indépendantes est un outil de plus en plus utilisé lors d’études de connectivité avec l’IRM fonctionnelle (voir chapitre 1). Les observations sont les volumes acquis au cours d’une session (se référer au paragraphe 1.3.3 pour le détail d’une étude en IRMf). On recherche les sources les plus indépendantes spatialement et dont les voxels qui les composent ont des décours temporels très proches au sein d’une même source. Les sources, ou composantes, trouvées par cette méthode mettent en avant des co-activités fonctionnelles. Elles peuvent correspondre à des réseaux cérébraux spontanés comme ceux décrits précédemment en section 1.4, à des réseaux liés à une tâche ou encore à des sources reflétant du bruit ou des artefacts.

En IRMf, il est possible de réaliser :

- ▷ une ACI spatiale, où les sources recherchées sont des cartes spatiales indépendantes et les observations sont les volumes acquis à chaque TR (temps de répétition). Dans l’équation (2.3), N est le nombre de volumes (ou points temporels) et M est le nombre de voxels d’un volume.

- ▷ ou une ACI temporelle, où les sources sont des signaux indépendants temporellement et les observations sont les décours temporels de chaque voxel des volumes. Dans l'équation (2.3), N est le nombre de signaux temporels observés (le nombre de voxels) et M le nombre de points temporels pour un voxel (le nombre de volumes).

L'ACI spatiale est préférée à l'ACI temporelle car le nombre de volumes acquis (quelques centaines) est très faible par rapport au nombre de voxels d'un volume (plus de 10^5). Ainsi, bien que pour les deux ACI le nombre d'observations est largement supérieur au nombre de sources attendu (quelques dizaines), dans le cas de l'analyse spatiale, la dimension de l'observation M est plus grande. En pratique, le nombre de volumes N étant élevé, on fait précéder l'ACI par une analyse en composantes principales (ACP) pour réduire l'ordre du modèle donc la dimension de la matrice W à estimer. Cet ordre est choisi en fonction du type d'étude : groupe ou sujet unique. Une étude de groupe ne recherchera que des sources communes donc par définition moins nombreuses que toutes celles présentes chez un sujet. Il s'agit de plus de faire un compromis entre la sur-division de sources (si l'ordre du modèle choisi est trop grand) et l'agglomération de différents réseaux au sein d'une même composante (dans le cas d'un ordre trop petit).

L'application de l'ACI en IRMf se fait, en pratique, en considérant le mélange entre les sources comme linéaire. Les hypothèses et propriétés des signaux requis, présentées dans les paragraphes 2.1.1 et 2.1.2, sont également réalisées, à savoir :

- ▷ Le nombre d'observations est supérieur au nombre de sources. Ce qui est majoritairement le cas puisque l'on s'attend à trouver entre 10 et 20 réseaux cérébraux pour plus de 100 volumes observés.
- ▷ Les observations ont des histogrammes de valeurs plus gaussiens que ceux des sources.

Ainsi, l'ACI peut être utilisée de manière totalement exploratoire en étudiant toutes les sources reflétant l'activité cérébrale, dans une situation de repos notamment, lors d'étude de connectivité avec un paradigme (réseaux isolés suite à une tâche spécifique par exemple) ou à des fins de débrouillage (en écartant les sources traduisant clairement un artefact). L'étude peut se faire aussi bien sur un groupe, pour mettre en avant les réseaux communs par exemple, ou sur un sujet unique qui est la seule approche pour en faire un instrument à visée diagnostique. La sélection des sources d'intérêt est le sujet de la première partie de cette thèse et sera développé ultérieurement.

2.2.1.1 ACI sujet unique

Après l'application de l'ACI sur les images fonctionnelles, on obtient les cartes spatiales des sources ainsi qu'un décours temporel associé à chacune d'entre elles. Ce décours temporel, correspondant aux coefficients de la matrice de mélange estimée, reflète l'implication d'une source spatiale dans les différentes images initiales, c'est-à-dire au cours du temps. La FIGURE 2.2 illustre les différents éléments de l'ACI spatiale appliquée en IRMf sujet unique. La matrice X correspond aux images fonctionnelles acquises ; la matrice A correspond aux décours temporels des composantes (les sources) ; la matrice S correspond aux cartes spatiales des composantes. Les valeurs obtenues, aussi bien pour la carte spatiale que pour les décours temporels, sont relatives. Aucune information suite à l'extraction n'est donc donnée sur l'intensité absolue des activations. Il s'agit ici uniquement de connectivité.

Un très faible pourcentage du nombre de composantes finales peut-être associé à un réseau cérébral. Les autres composantes, nommées « composantes de non-intérêt », peuvent refléter du bruit non structuré (de sources inconnues) ou des artefacts connus (liés à des bruits physiologiques ou des artefacts de la séquence par exemple).

2.2.1.2 ACI de groupe

Il est possible également de réaliser une ACI de groupe pour révéler les réseaux communs à une population. Les approches d'ACI de groupe peuvent être regroupées en différentes catégories

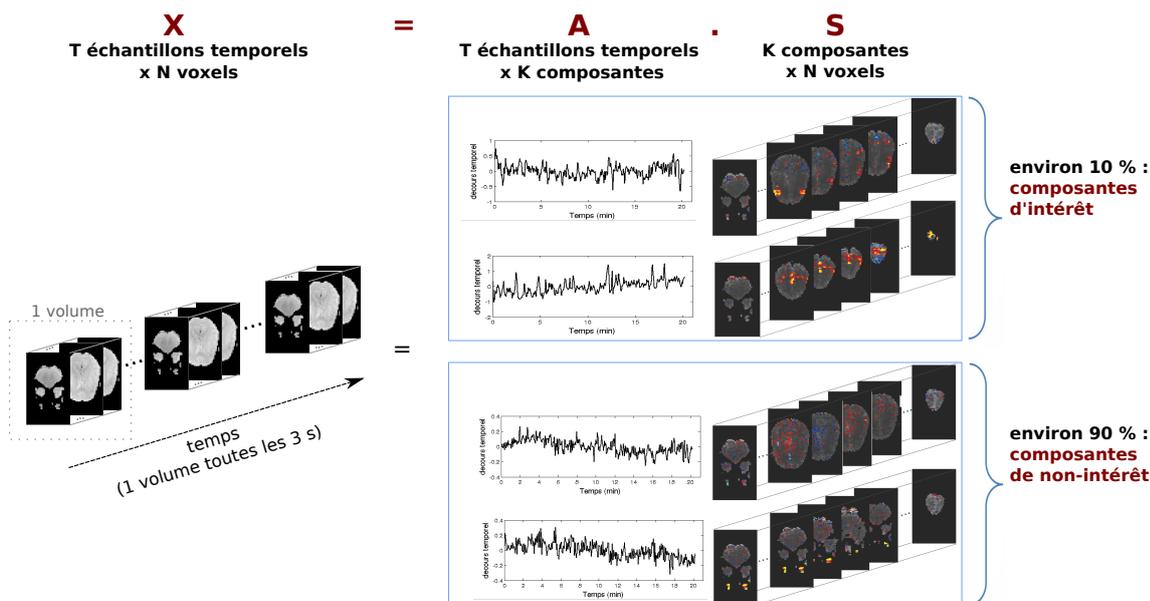


FIGURE 2.2 – illustration de l'ACI en IRM fonctionnelle.

selon l'organisation des données au préalable et les statistiques réalisées par la suite [39]. Certaines méthodes utilisent des ACI sujet unique puis combinent leurs résultats via une méthode de regroupement ou une recherche de corrélation spatiale [40, 41]. Les autres méthodes demandent un ré-agencement des données puisque l'ACI va être appliquée une seule fois sur le groupe entier. Pour ce faire, les images des différents sujets de l'étude vont être concaténées soit temporellement [42] soit spatialement [43].

La concaténation spatiale signifie qu'un volume représente non plus un cerveau d'un sujet mais les cerveaux de tous les sujets. Dans ce cas, une hypothèse temporelle est faite, c'est-à-dire que l'on va considérer que les réactions, en terme de connectivité, vont se dérouler au même moment. Cette hypothèse n'est évidemment pas applicable pour des études au repos, mais peut être intéressante dans le cas d'un paradigme très ciblé. On s'affranchit ici des variabilités anatomico-fonctionnelles des sujets.

Pour la concaténation temporelle, les volumes des nouveaux sujets vont être ajoutés à ceux du premier comme s'il s'agissait de volumes supplémentaires d'un même et unique sujet. Cette méthode implique une hypothèse de cohérence spatiale, c'est-à-dire que l'on considère que les régions fonctionnelles sont les mêmes chez tous les individus de l'étude. Ce qui évidemment nécessite une étape de normalisation (recalage des sujets les uns par rapports au autres). Cette normalisation est faite d'un point de vue anatomique qui ne prend pas en compte les variabilités des régions fonctionnelles inter-sujets (voir paragraphe 1.3.3.2 pour la préparation des données). C'est la méthode d'ACI de groupe la plus répandue. Elle semble donner de meilleurs résultats dans les études comparatives à ce sujet [44]. La toolbox GIFT [45], développée par l'équipe de Calhoun, est l'outil de référence pour réaliser une ACI de groupe et le post-traitement associé.

Il est possible par la suite d'estimer les images et décours temporels propres à chaque sujet de l'étude. Cette étape est appelée la reconstruction arrière ou *back-reconstruction* et peut se faire de manière directe (GICA [42, 46]) ou indirecte (régression spatio-temporelle ou *dual regression* [45, 47]).

2.2.1.3 Comparaison avec d'autres méthodes

Une méthode couramment utilisée durant les études d'IRMf est le modèle linéaire général (GLM). Pourtant on ne peut étudier par cette approche uniquement l'activation que l'on a mo-

délimitée contrairement à l'ACI qui est, elle, guidée par les données.

La principale alternative à l'ACI reste l'étude de la connectivité par la méthode d'ensemencement ou *seed-based correlation*. Dans ce cas, l'étude se fait par paires et demande le choix de la région graine. L'ACI, elle, considère la relation de tous les voxels simultanément.

2.2.2 Points difficiles

Afin de produire des résultats fiables, l'analyse en composantes indépendantes demandent de maîtriser les différentes étapes de pré- ou post-traitement. Plusieurs points requièrent donc une attention spéciale.

Avant même d'exécuter l'analyse, on peut se poser la question du filtrage des données afin d'augmenter le ratio signal sur bruit. L'ACI permet, idéalement, d'identifier également comme sources, les bruits et artefacts, particulièrement avec un nombre de composantes élevé (supérieur à 80). À nombre égal de composantes recherchées par l'ACI, un filtrage « excessif » pourrait mener à une diminution de l'indépendance entre les sources et nuirait donc à l'ACI. S'il est important de réaliser les étapes de recalage mentionnées précédemment (voir paragraphe 1.3.3.2) et d'écartier les images acquises lors d'un fort mouvement de tête, il n'est pas souhaitable de réaliser un filtrage supplémentaire des données. Dans les faits, le pré-traitement des images pour une étude de groupe contient une étape de lissage spatial (filtre passe-bas gaussien) (voir étape f) du paragraphe 1.3.3.2) qui se révèle suffisante [48].

Le nombre de composantes à estimer, choisi lors de l'étape d'ACP (précédant l'ACI, voir section 2.2.1) est différent selon que l'on se place en étude de groupe ou de sujet unique. Ainsi, en étude de groupe, l'ordre du modèle peut être faible par rapport au nombre d'images initiales (entre 15 et 60 en règle générale ; il peut être estimé par le logiciel GIFT [45]). Cela permet de restreindre le modèle à n'extraire que les réseaux partagés par tous les sujets. À un niveau individuel, quand on souhaite extraire également toutes les composantes spécifiques à un sujet, l'ordre du modèle doit être élevé sans pour autant mener à des sources représentant un réseau qui a été divisé [49, 50, 51].

Le dernier point qui vient après l'ACI est la sélection des composantes liées à une activité cérébrale. Cette étape est le sujet de la première étude de cette thèse et sera donc développée par la suite. Néanmoins, on peut déjà noter qu'il existe des méthodes de sélection manuelle ou semi-automatique, adaptées à différents cas de figures (sujet unique, groupe, tâche, débruitage).

2.2.3 Évaluation de la méthode

L'ACI est maintenant bien implantée dans les chaînes de traitement des protocoles d'IRMF. Afin d'interpréter correctement les résultats, de nombreuses études ont donc cherché à l'évaluer et à la développer [52, 53]

2.2.3.1 ACI temporelle

Diverses publications se sont intéressées, notamment, à l'implantation d'une ACI temporelle plutôt que spatiale pour mettre en avant des indépendances entre les décours temporels [11, 54]. Si certaines études comparatives arrivent à des résultats proches entre les deux types d'ACI quand il s'agit d'études avec paradigme de blocs, les résultats peuvent être moins prévisibles dans une étude où les sujets sont au repos. Dans le cas de l'ACI temporelle, la matrice W à estimer (voir section 2.1.2) est proportionnelle au nombre de voxels des volumes acquis contrairement à l'ACI spatiale dont l'ordre du modèle est proportionnel au nombre de volumes. La dimension spatiale étant largement supérieure à la dimension temporelle dans les jeux de données IRMF, cette méthode demande un très grand nombre de ressources informatiques calculatoires et reste marginale dans notre contexte d'étude.

2.2.3.2 Différents algorithmes développés

L'ACI ayant donné lieu à de nombreux développements, plusieurs algorithmes sont disponibles pour réaliser cette séparation de sources dans un contexte d'études IRMf. Il est donc intéressant d'examiner l'impact du choix de l'algorithme sur la séparation des sources qui nous concernent. Les deux algorithmes d'ACI les plus utilisés en IRMf sont dénommés « Infomax » [38] et « Fixed-Point algorithm » (ou « FastICA ») [55]. Il existe des liens significatifs entre indépendance et non-gaussianité des sources (propriétés recherchées pour les signaux sources sur lesquelles se basent respectivement les algorithmes « Infomax » et « FastICA »), il est donc peu étonnant que les études comparatives ne relèvent pas de différences importantes dans les résultats [56]. Esposito [56] a en effet montré que les deux algorithmes sont très performants, avec une supériorité de l'algorithme « Fixed-Point » concernant la précision spatiale et temporelle et une supériorité de l'algorithme « Infomax » en terme d'estimation globale et d'aptitude au débruitage. L'auteur conseille donc l'utilisation d'« Infomax » dans le cas d'études où les activations ne peuvent être modélisées *a priori*, ce qui le cas de l'IRMf de repos, notre champ d'application. Il est par ailleurs intéressant de noter dans cette étude, l'augmentation de la valeur de la vraisemblance de la décomposition en fonction du nombre croissant de composantes demandées par l'ACP qui précède l'ACI. Cette augmentation est d'autant plus importante avec l'algorithme « Infomax », ce qui conforte notre choix puisque travaillant sur sujet unique, l'ordre de nos modèles est élevé (voir section 2.2.2). Une étude comparative pour les études de groupe apporte des conclusions similaires [46].

En ce qui concerne les études avec tâches, l'algorithme « HYBICA » proposé par McKeown en 2000 [57], permet de mieux estimer les sources qui sont directement reliées au paradigme de tâche en fusionnant l'ACI, entièrement guidée par les données, et le modèle linéaire général, qui est une approche basée sur des hypothèses.

2.2.3.3 Reproductibilité des réseaux à travers l'ICA

Comme nous avons pu le voir dans le chapitre précédent, section 1.4, il existe de nombreux réseaux partagés par la population (les RSN). C'est d'ailleurs la base de l'ACI de groupe où l'on recherche les activations communes aux différents sujets (voir section 2.2). La question de la reproductibilité des sources trouvées par ACI se pose aussi lors des études intra-sujet ou inter-sujets pour ne pas influencer les interprétations des variabilités de l'activité ou de la connectivité cérébrale.

Afin de s'assurer de la fiabilité des composantes à étudier, des méthodes proposent de répéter l'ACI plusieurs fois consécutivement pour repérer les composantes les plus stables [58, 59]. Parmi elles, l'approche « ICASSO » publiée par Himberg en 2004 et implémentée dans la toolbox GIFT [45], est de plus en plus utilisée, particulièrement pour les études multi-sujets. L'ACI est lancée un nombre de fois défini par l'utilisateur (entre 5 et 15 fois généralement) avec des modifications dans les conditions initiales (modification aléatoire des valeurs initiales de W et/ou bootstrapping sur les données). Les composantes estimées sont agglomérées par une mesure de similarité et représentées sous forme d'agglomérats dans un espace 2D. Les agglomérats les plus compacts correspondent aux composantes les plus stables et donc fiables.

Analyse markovienne cachée et processus en interaction : une synthèse

3

Les modèles de Markov cachés, ou MMC, du nom du mathématicien Andreï Markov à l'origine de ces modèles dès 1906, permettent de modéliser l'évolution dynamique de processus stochastiques observés indirectement [60]. Introduits et étudiés à la fin des années 1960 par Baum [61, 62], ils ont connu un développement rapide grâce, d'une part, à Viterbi qui publie en 1967 un algorithme de décodage adaptable pour déchiffrer la séquence d'états cachés [63] et d'autre part, l'équipe de Baum à nouveau, qui propose une technique de ré-estimation des paramètres utilisés pour une meilleure modélisation [64]. Très largement utilisés avec succès en reconnaissance automatique de la parole, les MMC ont été rapidement adoptés par la communauté du traitement du signal. Les MMC trouvent aujourd'hui des applications dans de nombreux domaines allant de l'analyse du trafic dans les réseaux de communication à la cryptanalyse, en passant par l'étude de la génétique des populations ou bien encore dans le domaine des mathématiques financières.

De par leur succès précoce, les MMC ont fait, et font toujours, l'objet de nombreuses améliorations et extensions à la fois en matière de modélisation, pour mieux prendre en compte la réalité des phénomènes à analyser, que sur le plan algorithmique, dans la recherche, par exemple, de procédures d'optimisation des paramètres de ces modèles. Il n'est donc pas envisageable de traiter de tous ces aspects dans ce manuscrit. Nous nous limiterons à la présentation synthétique du MMC standard et à celle de quelques unes de ses variantes multidimensionnelles, nécessaires à la compréhension de cette thèse.

La présentation du MMC standard, unidimensionnel, est en effet incontournable car elle établit les bases des approches markoviennes cachées : processus caché, processus observable, états, lois d'observation, indépendance conditionnelle, graphe de dépendance, etc. Elle permet également d'aborder l'algorithmie de ces modèles suivant les trois problématiques classiques que sont l'évaluation, le décodage et l'apprentissage. Dans un second temps, nous nous intéresserons aux extensions multidimensionnelles du MMC standard, extensions qui nous intéressent particulièrement et pour lesquelles soit le processus observable, soit le processus caché, soit les deux, sont intrinsèquement multidimensionnels. Nous nous arrêterons en particulier sur le MMC-Produit, extension la plus directe du MMC standard, au cas multidimensionnel. Nous poursuivrons par les MMC-Couplés de moindre complexité pour terminer par un retour rapide sur d'autres modèles multidimensionnels non retenus pour cette thèse même si potentiellement applicables. Ces différentes présentations permettront aussi d'introduire les principales notations utilisées dans la partie III de ce manuscrit.

3.1 Modèle de Markov Caché standard

3.1.1 Principe de base

Une chaîne de Markov est une suite de variables aléatoires (X_t) , à valeur dans un espace fini d'états, permettant de modéliser l'évolution dynamique d'un système aléatoire. Deux propriétés régissent les chaînes de Markov (d'ordre 1 et homogène, cas le plus commun). La première propriété, dite propriété de Markov, stipule que l'état au temps t ne dépend que de l'état précédent au temps $t - 1$ (ordre 1) et aucunement des états antérieurs ($t - 2$, $t - 3$, etc.). Autrement dit, le futur du processus modélisé ne dépend que de son état présent, lequel contient toute l'histoire passée du processus. La seconde propriété, dite d'homogénéité, stipule que la probabilité de passage, ou de manière équivalente de transition d'un état i à $t - 1$ à un état j à t , ne dépend pas de t .

Une chaîne de Markov cachée λ est un double processus stochastique (X_t, Y_t) à temps discret. Il se compose de deux séquences :

- (X_t) : une chaîne de Markov, qui prend ses valeurs dans un ensemble fini d'états « cachés »,
- (Y_t) : un processus observable à valeurs dans \mathbb{R} , cas le plus usuel.

Le terme « caché » tient au fait que la chaîne de Markov est cachée, c'est à dire que les réalisations des variables aléatoires (X_t) ne sont pas directement observables. Ce qui l'est en revanche c'est l'observation Y_t émise par définition par l'état X_t à l'instant t . En particulier, la distribution statistique de Y_t dépend de l'état X_t occupé à l'instant t . Pour mieux appréhender l'ensemble des relations de dépendances et d'indépendances entre variables aléatoires X_t et Y_t au cours du temps, une chaîne de Markov cachée, appelée communément modèle de Markov caché (MMC), est souvent représentée par le graphe de dépendance de la FIGURE 3.1 :

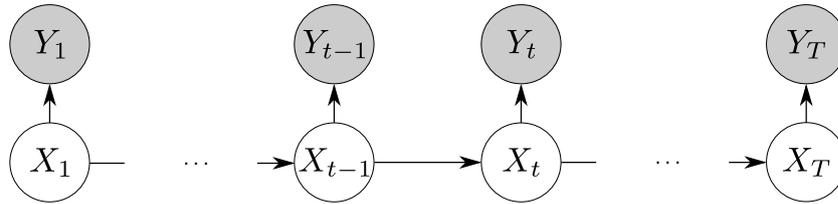


FIGURE 3.1 – Graphe de dépendance d'un MMC standard : en blanc les états cachés, en gris les observations.

Le graphe de dépendance d'un MMC est composé de nœuds représentant des variables aléatoires et d'arêtes représentant les relations de dépendance entre ces variables aléatoires. Les absences d'arêtes entre les observations Y_t (absence de flèches sur la FIGURE 3.1) dénotent une relation d'indépendance conditionnelle entre ces variables aléatoires. Les observations Y_t (en gris) sont ainsi indépendantes les unes des autres conditionnellement aux états cachés X_t (en blanc). Les arêtes du graphe de dépendance présenté ici sont orientées ; pour chaque $t > 1$, l'état caché X_t est conditionné à son parent X_{t-1} .

3.1.2 Formalisme mathématique

Un MMC λ est caractérisé par les éléments suivants :

- ▷ N , le nombre d'états du modèle. Bien que cachés, leur nombre est en pratique le plus souvent connu *a priori*. Pour plus de simplicité d'écriture, les états seront notés directement par leur indice i , $1 \leq i \leq N$. Avec cette écriture, on définit ainsi S l'espace d'état du modèle tel que $S = \{i\} = \{1, \dots, N\}$. L'état à l'instant t , sera noté X_t et prend ses valeurs dans S .
- ▷ $A = \{a_{ij}, 1 \leq i, j \leq N\}$, la matrice de transition entre états du modèle, où a_{ij} représente la probabilité de transition entre l'état i et l'état j :

$$a_{ij} = P[X_{t+1} = j | X_t = i] \quad \forall 1 \leq i, j \leq N \quad (3.1)$$

- ▷ $B = \{b_j(Y_t), 1 \leq j \leq N\}$, l'ensemble des lois d'observation du modèle, où $b_j(Y_t)$ est la loi d'observation associée à l'état j :

$$b_j(Y_t) = P[Y_t | X_t = j] \quad (3.2)$$

- ▷ $\pi = \{\pi_i, 1 \leq i \leq N\}$, l'ensemble des probabilités initiales, où π_i décrit la probabilité que le premier état occupé (à $t=1$) soit l'état i

$$\pi_i = P[X_1 = i] \quad \forall 1 \leq i \leq N \quad (3.3)$$

Un modèle $\lambda(\theta)$ est donc défini par son jeu de paramètres θ tel que

$$\theta = (A, B, \pi) \quad (3.4)$$

La matrice de transition A définie en (3.1) est stochastique. Les éléments $\{a_{ij}, 1 \leq i, j \leq N\}$ sont donc positifs et la somme des éléments de chaque ligne vaut 1 $\left(\sum_{j=1}^N a_{ij} = 1, 1 \leq i \leq N \right)$.

Les lois d'observations définies en (3.2) peuvent être discrètes, si le nombre d'observations distinctes est dénombrable, ou continues, ce qui sera notre cas. L'observation à l'instant t , sera noté Y_t et prend ses valeurs dans \mathbb{R} pour nos travaux.

Le cas le plus courant pour des lois d'observations continues est d'utiliser une gaussienne ou un mélange de gaussiennes dans le cas multi-modale. Pour une observation continue, uni-modale, $b_j(Y_t)$ aura pour expression :

$$b_j(Y_t) = \mathcal{N}(Y_t; \mu_j, \sigma_j^2) \quad (3.5)$$

avec μ_j et σ_j^2 la moyenne et la variance de la gaussienne associée à l'état j . Tout au long du manuscrit, les variables en gras représentent des matrices ou des vecteurs. Ainsi, \mathbf{X} peut dénoter un vecteur, X_t le t -ième élément du vecteur \mathbf{X} .

3.1.3 Problèmes fondamentaux et solutions

L'application pratique des MMC, quel que soit le domaine d'application, conduit à résoudre les trois problèmes fondamentaux suivants :

problème 1 : Évaluation : Étant donné une séquence de T observations $\mathbf{Y} = (Y_1, Y_2, \dots, Y_T)$ et un modèle λ , quelle est la probabilité que cette séquence ait été générée par ce modèle ? En d'autres termes, quelle est la vraisemblance $P[\mathbf{Y} | \lambda(\theta)]$?

problème 2 : Décodage : Étant donné une séquence de T observations $\mathbf{Y} = (Y_1, Y_2, \dots, Y_T)$ et un modèle λ , quelle est la séquence d'états correspondante $\mathbf{X} = (X_1, X_2, \dots, X_T)$ qui explique le mieux les observations ?

problème 3 : Apprentissage : Comment ajuste-t-on les paramètres $\theta = (A, B, \pi)$ du modèle λ pour maximiser la probabilité $P[\mathbf{Y} | \lambda(\theta)]$ pour une séquence d'observations \mathbf{Y} ?

3.1.3.1 Solution du problème 1

Afin d'éviter de calculer la vraisemblance $P[\mathbf{Y} | \lambda(\theta)]$ en énumérant tous les cas possibles d'une séquence d'états de T éléments, qui implique une complexité importante ($2T \cdot N^T$ opérations), il existe une procédure plus efficace. Elle s'appuie sur le calcul des probabilités avant-arrière (*forward-backward*) définies comme suit :

- probabilité avant $\alpha_t(i) = P[Y_1, Y_2, \dots, Y_t, X_t = i | \lambda(\theta)]$: probabilité d'observer la séquence partielle Y_1, Y_2, \dots, Y_t et d'être à l'état i au temps t , sachant le modèle $\lambda(\theta)$. Elle se calcule par induction de la manière suivante :

$$1) \text{ initialisation } \alpha_1(i) = \pi_i b_i(Y_1), \quad 1 \leq i \leq N,$$

$$2) \text{ récursivité } \alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(Y_{t+1}), \quad 1 \leq t \leq T-1, \quad 1 \leq j \leq N,$$

$$3) \text{ fin } P(\mathbf{Y}|\lambda(\theta)) = \sum_{i=1}^N \alpha_T(i).$$

– probabilité arrière $\beta_t(i) = P[Y_{t+1}, Y_{t+2}, \dots, Y_T | X_t = i, \lambda(\theta)]$: probabilité d'observer la séquence partielle $Y_{t+1}, Y_{t+2}, \dots, Y_T$ sachant que le modèle est à l'état i au temps t . Elle se calcule comme suit :

$$1) \text{ initialisation } \beta_T(i) = 1, \quad 1 \leq i \leq N,$$

$$2) \text{ récursivité } \beta_t(i) = \sum_{j=1}^N a_{ij} b_j(Y_{t+1}) \beta_{t+1}(j), \quad t = T-1, T-2, \dots, 1, \quad 1 \leq i \leq N.$$

Le calcul de $P[\mathbf{Y}|\lambda(\theta)]$ est donné par la formule :

$$P[\mathbf{Y}|\lambda(\theta)] = \sum_{i=1}^N \alpha_t(i) \beta_t(i) \quad 1 \leq t \leq T \quad (3.6)$$

En pratique, il est plus direct d'utiliser cette équation avec $t = T$. Cette procédure avant-arrière, basée sur la programmation dynamique, permet de réduire la complexité calculatoire de la vraisemblance $P[\mathbf{Y}|\lambda(\theta)]$ de $O(2T.N^T)$ à $O(N^2.T)$.

3.1.3.2 Solution du problème 2

Il s'agit ici d'inférer la séquence d'états la plus optimale ayant généré les observations. On peut ainsi considérer la séquence des états localement optimaux, ou bien la séquence d'états globalement optimale. La première solution, *a priori* acceptable, peut malheureusement conduire à une séquence d'états impossible, c'est à dire comportant des transitions interdites entre états, parce qu'à probabilités de transition nulles. La seconde solution ne présente pas ce défaut. C'est elle qui est généralement retenue et calculée suivant le critère MAP (Maximum A Posteriori), en utilisant l'algorithme de Viterbi [63]. Cet algorithme définit la variable $\delta_t(i)$ telle que :

$$\delta_t(i) = \max_{X_1, \dots, X_{t-1}} P[X_1, \dots, X_{t-1}, X_t = i, Y_1, \dots, Y_t | \lambda(\theta)] \quad (3.7)$$

qui est calculée de manière récursive par la procédure suivante :

$$1) \text{ initialisation } \delta_1(i) = \pi_i b_i(Y_1), \quad 1 \leq i \leq N,$$

$$\psi(i) = 0,$$

$$2) \text{ récursivité } \delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(Y_t), \quad 2 \leq t \leq T, \quad 1 \leq j \leq N,$$

$$\psi(i) = \operatorname{argmax}_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}], \quad 2 \leq t \leq T, \quad 1 \leq j \leq N,$$

$$3) \text{ fin } x_T^* = \operatorname{argmax}_{1 \leq i \leq N} [\delta_T(i)],$$

$$4) \text{ retour arrière } x_t^* = \psi_{t+1}(x_{t+1}^*), \quad t = T-1, T-2, \dots, 1.$$

La séquence optimale est obtenue par retour arrière en utilisant la variable de sauvegarde de l'indice ψ_t .

3.1.3.3 Solution du problème 3

L'apprentissage permet de maximiser $P[\mathbf{Y}|\lambda(\theta)]$ en ajustant les paramètres du modèle $\theta = (A, B, \pi)$. Là encore, il existe plusieurs méthodes permettant de s'approcher d'un modèle optimal. L'algorithme de Baum-Welch [64], qui repose sur l'algorithme EM (Expectation - Maximisation [65,66]), est celui communément utilisé pour appréhender ce problème. Par cette approche itérative, on cherche à maximiser la vraisemblance $L(\lambda(\theta)) = P[\mathbf{Y}|\lambda(\theta)]$. Pour plus de clarté, le modèle $\lambda(\theta)$ sera annoté par la suite de manière compact θ . Ainsi $L(\theta) = P[\mathbf{Y}|\theta]$. Il faut alors introduire

une fonction auxiliaire $Q(\theta, \hat{\theta})$ avec $\hat{\theta}$ le jeu de paramètres estimé :

$$Q(\theta, \hat{\theta}) = \sum_{\mathbf{X}} P[\mathbf{X}|\mathbf{Y}, \theta] \cdot \log(P[\mathbf{X}, \mathbf{Y}|\hat{\theta}]) \quad (3.8)$$

On montre par la relation suivante (3.9), en utilisant l'inégalité de Jensen, que tout accroissement de la fonction $Q(\theta, \cdot)$ obtenu en passant de θ à $\hat{\theta}$ conduit à accroître la vraisemblance $L(\theta)$:

$$Q(\theta, \hat{\theta}) - Q(\theta, \theta) \leq L(\hat{\theta}) - L(\theta) \quad \forall(\theta, \hat{\theta}) \quad (3.9)$$

$\hat{\theta}$ est donc construit de manière itérative afin d'accroître la vraisemblance.

En développant la fonction auxiliaire de manière à faire apparaître les paramètres à optimiser, l'équation (3.8) peut s'écrire :

$$\begin{aligned} Q(\theta, \hat{\theta}) &= \sum_{i=1}^N P[X_1 = i|\mathbf{Y}, \theta] \cdot \log(\hat{\pi}_i) \\ &+ \sum_{i=1}^N \sum_{j=1}^N \sum_{t=1}^{T-1} P[X_t = i, X_{t+1} = j|\mathbf{Y}, \theta] \cdot \log(\hat{a}_{ij}) \\ &+ \sum_{i=1}^N \sum_{t=1}^T P[X_t = i|\mathbf{Y}, \theta] \cdot \log(\hat{b}_i(Y_t)) \end{aligned} \quad (3.10)$$

Les trois termes de l'addition (3.10) peuvent être optimisés séparément pour conduire aux formules de ré-estimation suivantes :

$$\left\{ \begin{array}{l} \hat{\pi}_i = \gamma_1(i) \\ \hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \\ \hat{\mu}_i = \frac{\sum_{t=1}^T \gamma_t(i) Y_t}{\sum_{t=1}^T \gamma_t(i)} \end{array} \right. , \quad \hat{\sigma}_i^2 = \frac{\sum_{t=1}^T \gamma_t(i) (Y_t - \hat{\mu}_i)^2}{\sum_{t=1}^T \gamma_t(i)} \quad (3.11)$$

avec les probabilités ξ_t et γ_t définies grâce aux probabilités avant-arrière :

$$\begin{aligned} \xi_t(i, j) &= P[X_t = i, X_{t+1} = j|\mathbf{Y}, \theta] = \frac{1}{P(\mathbf{Y}|\theta)} \cdot \alpha_t(i) a_{ij} b_j(Y_{t+1}) \beta_{t+1}(j) \\ \gamma_t(i) &= P[X_t = i|\mathbf{Y}, \theta] = \sum_{j=1}^N \xi_t(i, j) \end{aligned} \quad (3.12)$$

Il est à noter que l'utilisation des probabilités avant-arrière pour les formules de ré-estimation, mène à des maxima locaux. L'initialisation du jeu de paramètres est donc importante aussi bien pour une bonne estimation de l'espace paramètre du modèle (en atteignant le maximum global), que pour une convergence rapide de l'algorithme d'apprentissage. Il est à noter également que l'implémentation des probabilités avant-arrière mène à des valeurs extrêmement faibles rendant les solutions instables ou du moins imprécises. Afin de palier à ces problèmes de précision numérique, les approches communément utilisées sont : le ré-échelonnement des valeurs de probabilités à l'aide d'un coefficient de normalisation [67] ou l'utilisation du logarithme des probabilités. C'est cette dernière solution qui a été utilisée lors de l'implémentation du travail de cette thèse.

Des alternatives existent pour l'apprentissage aussi bien sur la façon d'optimiser la vraisemblance (descente de gradients par exemple [68]) que sur le critère même à utiliser. En effet, le nombre d'observations n'est pas infini et il n'y a pas de solution exacte pour résoudre le problème 3. Cela dépend de la façon dont on considère ce qui est optimal. Dans un contexte de classification par exemple, le critère de vraisemblance est moins performant que celui de maximisation de l'information mutuelle (MMI) [69].

3.2 Modèles de Markov cachés multidimensionnels

3.2.1 Problématiques générales

Afin d'améliorer les estimations, des observations complémentaires peuvent être incluses dans le modèle. Il est donc possible de développer des MMC multidimensionnels qui modélisent les interactions entre plusieurs processus stochastiques non visibles. Ce type d'approche est fréquemment utilisé pour la reconnaissance de la parole où l'intégration des signaux visuels permettent une nette amélioration par rapport à la seule prise en compte des signaux audio [70]. L'ajout de différents canaux d'observation augmente, d'une part, les performances du modèle grâce à cette fusion d'informations et permet, d'autre part, l'analyse des interactions entre les processus aléatoires. Dans ce travail de thèse, nous nous intéressons aux extensions multidimensionnelles du MMC standard pour lesquelles soit le processus observable, soit le processus caché, soit les deux, sont intrinsèquement multidimensionnels. Ainsi avec K canaux d'observations, la façon la plus naïve d'aborder cette question serait de considérer que chaque processus évolue indépendamment, ce qui aboutirait à K MMC standards unidimensionnels indépendants. Cependant, un tel modèle ne permettrait pas de saisir les interactions temporelles entre les processus. Plusieurs solutions existent alors pour aborder cette problématique. Elles sont exposées dans la suite de cette section. Pour ce qui va être présenté, on considèrera qu'il y a plusieurs séquences d'états synchrones, et plusieurs séquences d'observables associées.

3.2.2 Modèle de Markov Caché Produit (MMC-P)

Le premier modèle présenté est le MMC-Produit (MMC-P). Un MMC-P λ est un MMC standard construit sur un jeu de MMC $\{\lambda^{(k)}\}$ qui prend en compte leurs interdépendances temporelles [70, 71, 72]. Chaque $\lambda^{(k)}$ est donc un processus doublement stochastique $(X^{(k)}, Y^{(k)})$ où $X^{(k)} = (X_1^{(k)}, \dots, X_T^{(k)})$ est une chaîne de Markov cachée observée à travers la séquence d'observations $Y^{(k)} = (Y_1^{(k)}, \dots, Y_T^{(k)})$ de longueur T . $X_t^{(k)} = i^{(k)}$ signifie que le modèle $\lambda^{(k)}$ est dans l'état $i^{(k)}$ au temps t . $S^{(k)} = \{i^{(k)}\}$ est l'espace d'état de $X^{(k)}$. Le MMC-P λ construit à partir des $\{\lambda^{(k)}\}$, $1 \leq k \leq K$, est ainsi lui aussi un processus doublement stochastique (\mathbf{X}, \mathbf{Y}) dont l'espace d'état S est par définition le produit cartésien des $S^{(k)}$, d'où l'appellation MMC-Produit :

$$S \triangleq S^{(1)} \times S^{(2)} \times \dots \times S^{(K)} \quad (3.13)$$

$\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_T)$ est une chaîne de Markov cachée où $\mathbf{X}_t = (X_t^{(1)}, \dots, X_t^{(K)})^\top$ désigne le vecteur d'état au temps t . $\mathbf{X}_t = \mathbf{i}$ signifie que λ est dans l'état $\mathbf{i} = (i^{(1)}, \dots, i^{(k)}, \dots, i^{(K)})$, $\mathbf{i} \in S$.

$\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_T)$ est la séquence d'observations de λ où $\mathbf{Y}_t = (Y_t^{(1)}, \dots, Y_t^{(K)})^\top$ est obtenue par la concaténation des observations $Y_t^{(k)}$.

Les paramètres d'un MMC-P λ sont :

$$\pi_{\mathbf{i}} = P[\mathbf{X}_1 = \mathbf{i}], \quad (3.14)$$

$$a_{\mathbf{ij}} = P[\mathbf{X}_t = \mathbf{j} | \mathbf{X}_{t-1} = \mathbf{i}], \quad (3.15)$$

$$b_{\mathbf{j}}(\mathbf{Y}_t) = P[\mathbf{Y}_t | \mathbf{X}_t = \mathbf{j}] \quad (3.16)$$

La matrice de transition $A = \{a_{\mathbf{ij}}\}$ reflète les dépendances temporelles entre les K chaînes de Markov cachées. Elle modélise de façon probabiliste l'évolution dans le temps de leurs états joints. En pratique, l'indépendance conditionnelle de $Y_t^{(k)}$ sachant $X_t^{(k)}$ est supposée telle que : $P[\mathbf{Y}_t | \mathbf{X}_t = \mathbf{j}] = \prod_k P[Y_t^{(k)} | X_t^{(k)} = j^{(k)}] = \prod_k b_{j^{(k)}}(Y_t^{(k)})$ dans l'équation (3.16). La représentation du graphe de dépendance d'un MMC-P est présentée FIGURE 3.2. Nous retrouvons sur ce graphe, si l'on considère les dépendances conditionnelles entre rectangles, le graphe de dépendance d'un MMC standard (FIGURE 3.1), mais où l'état \mathbf{X}_t est un élément du produit en croix des sous-espaces d'états $S^{(k)}$ (en gris) et l'observation \mathbf{Y}_t est un vecteur d'observations (en gris).

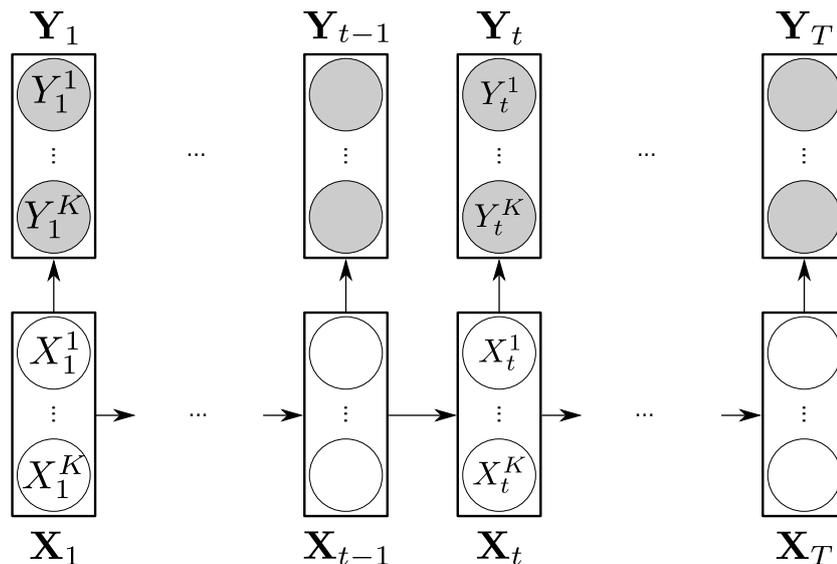


FIGURE 3.2 – Graphe de dépendance d'un MMC-Produit avec K canaux : en blanc les états cachés, en gris les observations. Les flèches modélisent les liens de dépendances.

De la même manière qu'un MMC standard, un MMC-P préserve les aspects algorithmiques d'un MMC concernant l'évaluation, le décodage et l'apprentissage [67,70]. La vraisemblance $P(\mathbf{Y}|\lambda)$ peut être calculée pour évaluer la manière dont un modèle correspond aux données observées (évaluation) et inversement, le jeu des paramètres du MMC-P, $\theta = \{\pi_{\mathbf{i}}, a_{\mathbf{ij}}, b_{j^{(k)}}(\cdot); \mathbf{i}, \mathbf{j} \in S, j^{(k)} \in S^{(k)}, 1 \leq k \leq K\}$, peut être appris de manière itérative en entraînant le modèle $\lambda(\theta)$ à produire \mathbf{Y} (*apprentissage*). La séquence d'états K -dimensionnelle optimale, $\hat{\mathbf{X}}$, et les K séquences d'états résultantes $\hat{X}^{(k)}$, peuvent être inférées ou décodées pour conduire à la transcription symbolique basée-état de \mathbf{Y} (*décodage*).

Par rapport à un MMC standard, le MMC-P possède toujours une matrice de transition A carrée, de dimension $N \times N$. Cependant, la taille de cette matrice croît exponentiellement avec le nombre K de canaux traités. En effet, N qui désigne le nombre de lignes et de colonnes de la matrice A du MMC-P, est aussi égal au cardinal de l'espace d'états résultant du produit en croix de K sous-espaces d'états. À raison de n états par sous-espace d'états, $N = n^K$. Le nombre de probabilités de transitions, $a_{\mathbf{ij}}$, à estimer est donc de $n^K \cdot n^K$.

3.2.3 Modèle de Markov Caché Couplés (MMC-C)

Le MMC Couplés, ou MMC-C [73] peut être considéré comme un cas particulier du MMC-P. Il peut être vu comme une collection de MMC, un par canal mais où chaque état caché au temps t sur un canal k dépend des états cachés au temps $t - 1$ sur tous les canaux. De cette manière les états peuvent interagir entre eux. On le retrouve pour la reconnaissance des gestes [74], de la parole (couplage audio-visuel) [70] ou encore des surfaces [75]. Le graphe de dépendance du MMC-C est illustré par la FIGURE 3.3. Les flèches modélisant les liens de dépendances entre les états et entre les états et les observations selon les canaux.

En reprenant les mêmes notations que précédemment, les paramètres du MMC-C sont définies ci-dessous. Ces équations mettent bien en évidence la distinction entre les canaux et l'indépendance des observations conditionnellement aux états.

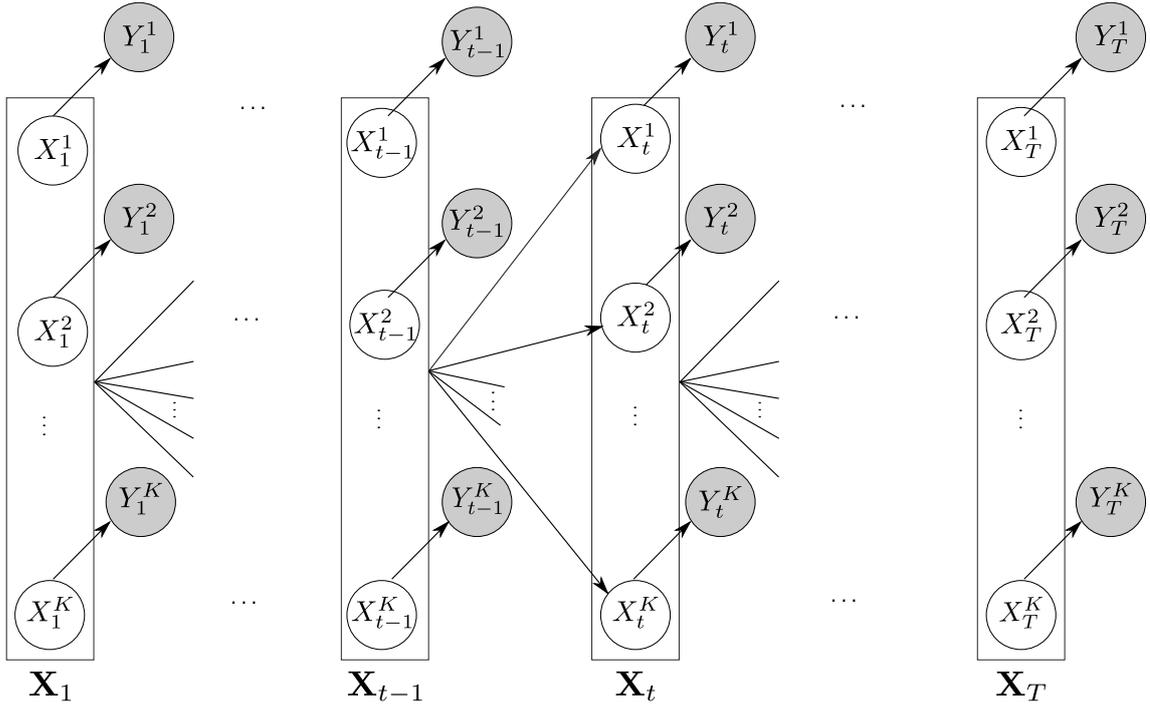


FIGURE 3.3 – Graphe de dépendance d'un MMC-Couplés avec $K = 2$ chaînes : en blanc les états cachés, en gris les observations. Les flèches modélisent les liens de dépendances.

$$\pi_{\mathbf{i}} = \prod_{k=1}^K \pi_{i^{(k)}}^{(k)} = \prod_{k=1}^K P[X_1^{(k)} = i^{(k)}], \quad (3.17)$$

$$a_{\mathbf{ij}} = \prod_{k=1}^K a_{ij^{(k)}}^{(k)} = \prod_{k=1}^K P[X_t^{(k)} = j^{(k)} | \mathbf{X}_{t-1} = \mathbf{i}], \quad (3.18)$$

$$b_{\mathbf{j}}(\mathbf{Y}_t) = \prod_{k=1}^K b_{j^{(k)}}^{(k)}(Y_t^{(k)}) = \prod_{k=1}^K P[Y_t^k | X_t^k = j^{(k)}] \quad (3.19)$$

De la même manière que le MMC-P, le MMC-C peut être appliqué en résolvant les mêmes problèmes que sont l'évaluation, le décodage et l'apprentissage. Pour ce modèle, ce sont les modélisation des paramètres $\pi_{\mathbf{i}}$ et $a_{\mathbf{ij}}$ qui diffèrent de ce qui est proposé par le MMC-P comme on peut le voir dans les équations (3.17) et (3.18). Par exemple, les paramètres les éléments $a_{\mathbf{ij}}$ de la matrice de transition seront calculés grâce à la formule (3.18) après estimation des $\{a_{ij^{(k)}}^{(k)}, \forall k\}$ au niveau de chaque canal k . Au final, la dimension de la matrice A d'un MMC-C construits sur N canaux aura la même dimension que celle d'un MMC-P basé sur le même nombre de canaux. Cependant, le MMC-C prend en compte des indépendances conditionnelles plus fortes que le MMC-P et doit estimer moins de probabilités de transitions que le MMC-P. En effet, il s'intéresse aux transitions d'un état I vers un état sur un canal $j^{(k)}$. Il y a donc $2K.n^K$ probabilités $a_{ij^{(k)}}^{(k)}$ à estimer.

3.2.4 Autres MMC multidimensionnels

Très utilisés en reconnaissance de la parole, notamment avec des canaux audio et visuels, les MMC multidimensionnels ont largement été développés pour s'adapter au mieux aux applications.

Sans expliciter ces autres possibilités, cette section les présente succinctement au regard des MMC-P et MMC-C privilégiés pour nos applications.

MMC multi-canaux Le MMC multi-canaux indépendants [76] modélise une indépendance entre les MMC disponibles sur chaque canal. La vraisemblance globale est le produit des vraisemblances de chaque MMC indépendant. Ainsi, contrairement au MMC-P ou au MMC-C, aucune dépendance statistique entre les canaux n'est prise en compte dans le calcul de la vraisemblance. De plus, le MMC multi-canaux ne requiert pas que les canaux aient la même longueur, ce qui peut être un avantage ou un inconvénient suivant que l'on souhaite synchroniser les canaux.

MMC factoriel Le MMC factoriel a été introduit par Ghahramani en 1997 [77]. Chaque séquence d'états cachés est indépendante des autres mais participe à toutes les séquences d'observations. La vraisemblance des observations est donc, comme pour le MMC-P, calculée de manière conjointe à partir de tous les canaux des états cachés. Par contre, à l'inverse des MMC-P et MMC-C, les probabilités de transitions sont construites sur l'hypothèse d'indépendance des séquences d'états cachés.

MMC hiérarchique Le MMC hiérarchique [78] est un peu plus éloigné des MMC multidimensionnels présentés précédemment puisque les différents MMC sont modélisés en « cascade ». Ainsi, les états cachés peuvent générer soit une observation unique soit une chaîne d'observations elle-même gouvernée par un sous-MMC. Ces sous-chaînes peuvent à leur tour générer d'autres observations ou sous-chaînes, elles-mêmes impliquant un sous-MMC et ainsi de suite. L'action récursive s'arrête quand on atteint un état qui génère une observation unique. Le MMC hiérarchique permet d'ajuster la modélisation à différents niveaux et longueurs de séquences.

Bases de données utilisées

4

4.1 Composition des bases de données

Les études réalisées dans cette thèse ont été développées et validées sur cinq bases d'images acquises au laboratoire dans des conditions différentes afin qu'elles soient les plus génériques possibles. Ces bases, B1, B2, B3, B4 et B5 forment ainsi une large base de données variées dont les caractéristiques sont détaillées ci-dessous et dans le TABLEAU 4.1.

Pour une étude de reproductibilité inter-sessions des données IRMf, douze volontaires sains ont effectué trois sessions espacées d'une semaine. Au cours de chaque session, une séquence EPI de 10 minutes (TR = 3 s) incluant une tâche visuelle ainsi qu'une séquence EPI-ASL de 20 minutes (TR = 3s) au repos ont été acquises et ont permis de constituer respectivement les bases B1 et B2. Lors de la dernière session, S4, huit sujets ont pris, une heure avant l'imagerie, un médicament, stimulant du système nerveux (30 mg de méthylphénidate) et quatre sujets ont pris, à la place, un placebo. Cette quatrième session n'étant pas utilisée pour toutes nos études, nous avons nommée B2_{3s} la sous-base de B2 contenant les données des 3 premières sessions (sans prise de médicaments ou placebo).

La base de données B3 est composée de seize volontaires sains recrutés pour une étude sur les troubles de la conscience suite à un coma. Pour chaque sujet, une session unique a permis l'acquisition d'une séquence EPI de 20 minutes (TR = 3 s) au repos.

Dix-huit patients constituent la base B4. Sélectionnés au cours d'une étude longitudinale sur deux types de démences neurodégénératives, ils répondent soit au diagnostic de maladie d'Alzheimer, soit au diagnostic de maladie à corps de Lewy ou présentent une pathologie mixte. Ces sujets, âgés, présentent des troubles cognitifs légers ou une démence légère. La séquence EPI-ASL utilisée dure 6 minutes (TR = 3 s) et se fait au repos.

Enfin la base B5 est formée de vingt sujets sains âgés. Ils constituent le groupe de sujets témoins pour une étude longitudinale sur les démences neurodégénératives mises en avant par les sujets de B4. La même séquence a donc été utilisée : EPI-ASL de 6 minutes (TR = 3 s) en condition de repos.

Ces bases d'images reflètent des conditions d'acquisition (intensité du champ de l'aimant, séquence, résolution), des protocoles (nombre de volume, tâche ou repos) ou encore des types de sujets (âge, sujet sain ou patient) différents. Au total, 66 patients ont été inclus dans ces bases de données, soit 138 acquisitions et 37 552 volumes. Les méthodes développées durant ce travail de thèse étant basées sur une combinaison de l'ACP et de l'ACI, le nombre de volumes (composantes) à sélectionner est réduit et il reste *in fine* 27 686 composantes à analyser et à pré-sélectionner. Le TABLEAU 4.1 ci-dessous décrit plus en détail les données qui composent notre base.

Les images anatomiques pondérées T1, de résolution $1 \times 1 \times 1 \text{ mm}^3$ pour chaque base, ont également été collectées au même moment.

TABLEAU 4.1 – Bases de données utilisées.

	B1	B2		B3	B4	B5
		B2 _{3s}	B2			
IRM	3 T Siemens	3 T Siemens		1,5 T Siemens	3 T Siemens	3 T Siemens
Sujets	sains	sains		sains	patients	sains
Condition	tâche	repos		repos	repos	repos
Nb de sujets	12	12		16	18	20
Ages (ans)	[22 - 43]	[22 - 43]		[19 - 51]	[54 - 85]	[46 - 76]
Nb ses / sujet	3	3	4	1	1	1
Nb total exa	36	36	48	16	18	20
Nb de vol / exa (durée (min))	200 (10)	404 (20,2)		400 (20)	120 (6)	120 (6)
Nb comp / exa	199	250		250	119	119
Rés (mm ³)	3×3×3	4×4×4		4×4×4	4×4×4	4×4×4
Séquence	EPI	EPI-ASL		EPI	EPI-ASL	EPI-ASL

Données classées selon : le type d'IRM (3 T Siemens ou 1,5 T Siemens) - les sujets (sains ou patients) - la condition (tâche ou repos) - le nombre de sujet - l'amplitude des âges - le nombre de sessions par sujet - le nombre total d'examens - le nombre de volumes acquis par examen (et la durée équivalente en minutes) - le nombres de composantes extraites avec l'ACI par examen - la résolution - le type de séquence d'acquisition des données IRMf.

4.2 Répartition des bases de données en fonction des études

Le TABLEAU 4.2 rend compte de l'utilisation des bases de données suivant les études de ce manuscrit.

TABLEAU 4.2 – Répartition des bases de données selon les études abordées dans ce manuscrit.

	Sélection des SAM (Partie II)	Analyse de la DCF (Partie III)	
		Étude 1 (sujet unique)	Étude 2 (groupe)
B1	×		
B2	B2 _{3s}	×	
	B2		×
B3	×		
B4	×		×
B5			×

Une croix symbolise l'utilisation de la base B_i dans l'une ou l'autre des parties du manuscrit.

4.2.1 Bases de données pour la sélection des SAM (Partie II)

La Partie II concernant la sélection des SAM s'appuie sur quatre bases à savoir B1, B2_{3s}, B3, B4. Au total, 46 patients ont été inclus dans cette étude, soit 106 acquisitions et 30 304 volumes.

Pour valider notre méthode, la base de données a été divisée en deux bases distinctes : une pour la phase d'apprentissage, l'autre pour la phase de test. Afin de préserver un maximum de données pour le test, le partage des composantes s'est fait selon un rapport d'environ 35 % (7 888) pour l'apprentissage et 65 % (14 418) pour le test. Les composantes d'une session ne peuvent appartenir qu'à l'une des deux bases, de test ou d'apprentissage. Le nombre de données étant suffisamment élevé pour que la méthode apprennent sur un nombre de données conséquent. Par ailleurs, la base B4 n'ayant été disponible que tardivement après le début de l'étude, il a été décidé de ne pas tenir compte des sujets de cette base lors de l'étape d'apprentissage. Cela a permis, *a posteriori*, de tester notre méthode sur une base n'étant pas impliquée dans le développement de la méthode. Ce choix a été d'autant plus intéressant que B4 présente plusieurs aspects particuliers concernant ces données (sujets avec pathologie, âgés, avec séquence EPI-ASL). Les deux bases sont présentées par le TABLEAU 4.3.

TABLEAU 4.3 – Bases de données d'apprentissage et de test.

	Base d'apprentissage	Base de test
Nombre de séquences issues de B1	10	20
Nombre de séquences issues de B2 _{3s}	10	20
Nombre de séquences issues de B3	10	6
Nombre de séquences issues de B4	-	18
Nombre total de séquences	30	64
Nombre total de composantes	7 888	14 418

4.2.2 Bases de données pour l'analyse de la DCF (Partie III)

Pour notre étude sur l'analyse de la dynamique de la connectivité fonctionnelle, nous avons utilisé les bases B2, B4 et B5. Pour une application à l'échelle du sujet, la base B2 est la plus adaptée puisqu'elle comprend quatre sessions dont la dernière se démarque par la prise d'un médicament ou d'un placebo.

Pour une application à l'échelle d'un groupe, nous avons utilisé les bases B4 et B5. Les patients de la base B4 sont atteints de la maladie à corps de Lewy ou DLB (pour *Dementia with Lewy Bodies*). Cette pathologie, qui touche 16 à 20 % des patients atteints de démence [79], est la deuxième forme la plus répandue de démence neurodégénérative après la maladie d'Alzheimer. Les principaux critères cliniques de la DLB sont des troubles cognitifs associés à des fluctuations cognitives, des syndromes parkinsoniens et des hallucinations visuelles. Tous les patients de la base B4 ont eu une évaluation formelle de leur diagnostic par trois cliniciens experts indépendants et les sujets sains ont été soumis à des évaluations cliniques, cognitives, psychiatriques et neurologiques de manière à exclure de l'étude les sujets qui pourraient avoir des troubles cognitifs non définis. Suite à cela, deux patients ont été exclus conduisant à une base B4 restreinte à 16 participants. Ces patients satisfaisaient aux critères de McKeith pour la démence à corps de Lewy probable (à savoir, au moins deux symptômes de base sur trois) [80]. Aucun de ces patients ne partageait à la fois les caractéristiques cliniques de la maladie à corps de Lewy et de la maladie d'Alzheimer. Les scores de MMSE (Mini Mental State Examination) [81], utilisés pour graduer les fonctions cognitives, étaient de 20,8 (écart-type : 3,2 ; compris entre 15 et 24) et 29,0 (écart-type : 1,0 ; compris entre 27 et 30) respectivement, pour les patients et les sujets sains. Les critères d'exclusion incluaient les contre-indications pour l'IRM, les antécédents de consommation abusive d'alcool et de drogues, les indices suggérant d'autres explications neurologiques ou psychiatriques pour leurs symptômes, les lésions cérébrales focales visibles par IRM et la présence d'une autre maladie médicale grave ou instable. La démence à corps de Lewy peut se manifester par différents symptômes tels que des hallucinations visuelles (9 patients concernés), des fluctuations cognitives (8 patients concernés) ou des symptômes parkinsoniens (10 patients concernés).

Au final, la base B5 est formée de vingt sujets sains âgés (11 femmes ; âge moyen 64,4 ans compris entre 46 ans et 76 ans) et la base B4 (pour l'étude de la partie III) de 16 patients atteints de démence à corps de Lewy (DLB) au stade de la démence légère (8 femmes ; âge moyen : 74,7 ans compris entre 54 ans et 89 ans). Des tests statistiques ont été réalisés pour vérifier l'appariement :

- en genre : les deux groupes sont appariés (test de Kolmogorov-Smirnov, $p < 0,05$);
- en âge : les patients sont plus âgés que les témoins (test de Kolmogorov-Smirnov, $p > 0,05$);
- en durée de scolarité : les témoins ont suivi une scolarité plus longue que les patients (test de Kolmogorov-Smirnov, $p > 0,05$).

Les sujets sains et les patients ne sont pas appariés pour les années d'études et l'âge car nous avons souhaité maximiser le nombre de sujets par groupe plutôt que de conserver un faible nombre de sujets par base.

Conclusion partielle

Cette partie a permis de présenter le contexte des travaux de cette thèse à savoir les études IRMf de repos au niveau du sujet unique comme du groupe ainsi que les bases de données utilisées. L'analyse en composantes indépendantes est l'outil privilégié pour extraire les SAM, réseaux de connectivité fonctionnelle, qu'elles soient idiosyncratiques ou partagées par plusieurs sujets (RSN). L'utilisation de cette méthode soulève néanmoins quelques problèmes quant à l'identification des composantes liées à un réseau d'aires cérébrales connectées fonctionnellement ainsi qu'au temps conséquent engendré par la sélection manuelle. Les interactions entre ces SAM sont encore peu connues et les modèles de Markov cachés multidimensionnels que l'on a présentés pourraient présenter une approche intéressante pour modéliser l'évolution temporelle de ces liens.

Deuxième partie

Sélection automatique des cartes de co-activités spatiales

Introduction	45
5 Méthodes de sélection des composantes liées à une activité cérébrale : Éléments d'état de l'art	47
6 Méthode proposée	59
7 Résultats et analyse des résultats	71
Conclusion partielle	79

Introduction

L'étude de la connectivité fonctionnelle en IRMf de repos est réalisée majoritairement par l'analyse en composantes indépendantes (ACI). Entièrement guidée par les données, cette technique robuste permet de décomposer les signaux, acquis lors de l'expérience IRMf, en cartes d'activation et décours temporels associés. Ces cartes spatiales, indépendantes les unes des autres, appelées composantes, reflètent majoritairement le bruit ou les artefacts. Cependant, elles peuvent être aussi représentatives de réseaux de co-activations cérébrales spontanées, ou SAM, entre 5% et 15% pour nos bases de données. La sélection des composantes ACI peut se faire alors dans deux buts distincts. Dans le cadre d'un débruitage, on cherchera à identifier les composantes artefactuelles afin de les retirer avant la reconstruction d'un signal BOLD « propre ». Dans l'étude de la connectivité fonctionnelle, on cherchera plutôt à identifier les composantes d'intérêt mettant en avant des réseaux cérébraux spontanés, pour analyse ultérieure.

C'est dans ce dernier contexte que s'inscrit notre première étude. La forte contrainte émanant de la demande des neuroscientifiques et cliniciens est de sélectionner, pour chaque sujet, toutes les composantes ACI jugées d'intérêt c'est-à-dire avec une sensibilité la plus proche des 100 % dans la détection des réseaux cérébraux spontanés. Ce type d'étude, réalisée sujet par sujet, a des visées médicales et pour cela aucun réseau ne doit être oublié afin de définir au mieux l'état du sujet. Une sélection entièrement manuelle a déjà été mise au point et validée dans l'article de Roquet et coll. [82].

Le but de ce premier travail est avant tout d'automatiser les critères de sélection manuelle existants, de manière à réduire le temps dédié à cette étape et ainsi envisager des études sur de nombreuses bases de données. La méthode se doit d'être applicable sur sujet unique, sans *a priori* et doit permettre la sélection des SAM liées au repos comme à la tâche. Les données utilisées pour cette partie ont été présentées dans le chapitre 4 et sont réparties en quatre bases B1, B2_{3s}, B3 et B4.

Un état de l'art des méthodes existantes est présenté dans un premier temps au chapitre 5. À partir de là, il a été possible de développer une méthode de sélection automatique en proposant de nouveaux critères et en adaptant d'autres issus de la littérature. La méthode est détaillée et discutée dans les chapitres suivants.

Méthodes de sélection des composantes liées à une activité cérébrale : Éléments d'état de l'art

5

5.1 Du classement des études

Dans le cadre de l'IRM fonctionnelle, de nombreuses méthodes permettant de sélectionner des composantes issues de l'analyse en composantes indépendantes ont déjà été proposées dans la littérature. Il est cependant difficile d'établir une classification unique de ces approches tant il existe de façons de les définir. Ainsi, on peut les différencier selon qu'elles :

- portent sur une étude de groupe [83, 84, 85] ou un sujet unique [57, 86, 87, 88, 89, 90, 91, 92] ;
- sont automatisées [12, 57, 83, 84, 85, 87, 88, 89, 90, 91, 93, 94, 95] ou manuelles [82, 86, 92] ;
- sont basées sur des caractéristiques spatiales [83, 88], temporelles [12, 57, 87, 90] ou les deux [85, 86, 89, 91] ;
- sont dirigées par les données [83, 84, 85, 89, 90, 91] ou requièrent des informations spatiales ou temporelles [12, 57, 86, 87, 88, 95] ;
- se focalisent sur une tâche [57, 83, 84, 85, 86, 87, 89, 91] ou s'intéressent aux composantes liées au repos également [12, 88, 90, 92, 93, 94, 95].

Lors de l'introduction à l'ACI au paragraphe 2.2.1, nous avons présenté le rôle de cette étape pour une étude en IRMf. Pour rappel, l'ACI peut être utilisée de manière totalement exploratoire en étudiant toutes les sources reflétant l'activité cérébrale, dans une situation de repos notamment, lors d'étude de connectivité avec un paradigme (réseaux isolés suite à une tâche spécifique par exemple) ou à des fins de débruitage (en écartant les sources traduisant clairement un artefact). L'étude peut se faire aussi bien sur un groupe, pour mettre en avant les réseaux communs par exemple, ou sur un sujet unique, à des fins diagnostiques. Les méthodes citées ci-dessus ont été développées dans un contexte bien délimité par le type d'étude et la place de l'ACI dans le traitement des données. Elles ne sont donc pas toutes adaptables à d'autres applications. Cependant, toutes ces méthodes ont en commun la volonté d'extraire des caractéristiques des composantes indépendantes, les CI, pour faciliter la sélection et c'est pourquoi nous les présentons dans ce chapitre. L'ordre chronologique de publication a été retenu dans la section suivante pour présenter ces méthodes.

Concernant le vocabulaire utilisé, nous parlerons de critère quand il s'agit de qualifier une CI. Le terme de descripteur renvoie à l'indicateur créé pour donner une mesure à ces critères, c'est-à-dire une quantification des caractéristiques des CI. Les méthodes automatiques font généralement appel à plusieurs descripteurs. Ces derniers sont souvent indépendants les uns des autres et c'est pourquoi certains sont redondants entre les méthodes. Ces descripteurs seront détaillés par la suite. Les méthodes publiées pour extraire uniquement la ou les composantes correspondant au réseau par défaut (DMN, décrit section 1.4) ne sont pas présentées par la suite car trop restrictives [12, 95].

5.2 Méthodes de sélection de composantes

En 1998, McKeown et coll. [86] figurent parmi les premiers à utiliser l'ACI pour une étude d'IRMf (avec tâches) dans le but d'extraire des réseaux en lien avec la tâche. Ils comparent, dans leur étude, l'ACI à l'analyse en composantes principales (ACP) et montrent à quel point cette méthode est prometteuse pour traiter les données d'IRMf et extraire les composantes reflétant un réseau. Leur méthode de sélection est simple puisqu'ils ne recherchent que les composantes liées à une tâche. Les CI liées à une tâche de manière permanente ou passagère sont retenues en observant la corrélation entre le décours temporel des composantes et celui du paradigme. Par la suite, ils ajoutent une étape de réduction de l'ordre du modèle en utilisant l'ACP, avant l'ACI, en plus des pré-traitements usuels en neurosciences [57] (voir paragraphe 1.3.3.2 page 13). Cette étape est constamment utilisée par toutes les autres méthodes qui suivront, l'objectif étant de réduire le nombre de composantes comme nous l'avons vu en section 2.2. De cette approche ont germées de nombreuses autres propositions.

Basée uniquement sur la densité spectrale de puissance des décours temporels, Thomas et coll. [87] ont développé, en 2002, une méthode automatique pour trier les composantes en trois classes : le signal, le bruit structuré et le bruit aléatoire. Leur méthode est néanmoins efficace uniquement sur les TR (temps de répétition) courts (310 ms dans leur étude) et ne considère comme bruits structurés que les bruits cardiaques et respiratoires.

La méthode proposée par Van de Ven et coll. [88] demande de définir des volumes d'intérêt lors de la première utilisation (avec des *a priori* sur les régions où l'on s'attend à du bruit structuré ou à une activité cérébrale de repos). Bien qu'applicable sur un sujet unique avec un nombre de composantes élevé, les *a priori* sont trop importants pour rechercher des réseaux peu connus et spécifiques d'un sujet.

En 2007, De Martino et coll. [89] ont proposé une classification automatique très complète. Après un apprentissage basé sur un SVM (*Support Vector Machine*, machine à vecteurs de support) six classes vont être décrites : 1) les composantes avec une activité liée à la tâche (en permanence ou de manière passagère), ou le réseau par défaut (DMN); 2) les composantes correspondant à un artefact de mouvement; 3) les composantes reflétant un artefact de susceptibilité typique en EPI; 4) les bruits physiologiques avec des pics très localisés (cas des gros vaisseaux); 5) les bruits avec hautes fréquences spatiales et enfin, 6) les composantes contenant des bruits avec des fréquences temporelles hautes. Après un pré-traitement classique, chaque composante va être représentée dans un espace multidimensionnel par des descripteurs qui constitueront l'empreinte de la composante. Onze descripteurs vont être ainsi calculés à partir de la distribution des valeurs des voxels, de l'agencement spatial, des propriétés temporelles et des propriétés spatiales. Cette méthode appelée *IC-fingerprint* est maintenant utilisée fréquemment lors d'étude avec tâche, y compris en sujet unique, dans un but de débruitage car elle permet de caractériser précisément les différents bruits. Cette méthode présente donc beaucoup d'avantages de par sa prise en compte de tous types d'artefacts, son absence de fortes hypothèses spatiales ou temporelles et sa catégorisation très poussée. Cependant, les recherches aussi bien des réseaux neuronaux que des bruits ou artefacts ne permettent pas une sensibilité maximum concernant la sélection des réseaux de repos.

La même année, Perlberg et coll. [90] s'intéressent à la sélection des composantes dans un but de débruitage pour des études au niveau du sujet. De manière automatique, seules les composantes liées à un bruit « connu », c'est-à-dire physiologique (respiration, battements cardiaques) vont être extraites. Un jeu de signaux représentatifs de bruits spécifiques est défini à partir des décours temporels issus de régions connues pour montrer de grosses fluctuations physiologiques (ventricules, vaisseaux sanguins, tronc cérébral). La sélection des composantes se fera en examinant la corrélation des décours temporels avec ces signaux de bruits caractéristiques. Cette méthode a l'avantage d'être facilement implémentable et nécessite peu de critères de sélection. Toutefois, la procédure de

sélection basée sur des *a priori* spatiaux, la sélection manuelle des masques d'intérêt (bien qu'elle puisse être automatisée) et la recherche de la spécificité uniquement sur les composantes de bruits structurés ne nous permettent pas d'utiliser directement cette technique pour notre application.

Dans la lignée des travaux précédents, Tohka et coll. [91] proposent une méthode avec pour objectif la reconstruction d'un signal BOLD sans bruit, pour une séquence avec tâche. Les composantes sont classées comme potentiellement liées à une tâche ou clairement liées à du bruit structuré. Le signal est reconstruit uniquement avec les composantes non étiquetées comme bruit. Des caractéristiques sont calculées pour chaque composante indépendante (CI) à partir des décours temporels et des cartes spatiales de chaque CI. Les seuils de détection sont appris sur un jeu de données pré-traitées. Cette méthode s'insère aussi dans le cadre d'études sur sujets uniques pourtant seulement 50 composantes sont extraites par l'ACI dans leur étude ce qui ne permet pas toujours d'isoler des composantes clairement liées à une activité cérébrale mais non décrites par la littérature car spécifiques au sujet.

La démarche de suppression d'artefacts plutôt que de sélection des SAM ou RSN est la plus courante dans les méthodes de sélection de composantes car certains artefacts sont bien décrits dans la littérature et peuvent être repérés facilement. Sui et coll. [83] utilisent cet aspect pour diviser leur méthode de sélection automatique en deux étapes consécutives : suppression des artefacts et sélection des composantes optimales (les RSN) parmi celles restantes après la première étape. Cette étude de 2009 est réalisée au niveau du groupe avec une séquence multitâche. Trois types d'artefacts sont retenus : les artefacts liés aux mouvements du LCR (liquide céphalo-rachidien), le bruit distribué de manière éparse et les artefacts liés au mouvement du sujet (tête, yeux). Ce choix se fait selon deux critères : la corrélation spatiale des CI avec la matière grise par rapport à leur corrélation avec le LCR et le degré d'agrégation qui prend en compte l'entropie de la composante et le nombre d'agrégats des voxels d'intérêt. Dans un second temps, les composantes restantes sont ordonnées selon des métriques de probabilités d'appartenance à un groupe. Celles en haut de la liste sont étiquetées comme composantes optimales. Cette méthode est particulièrement adaptée pour les études de groupe avec paradigmes de tâche puisque l'étude a testé cinq types de tâches différents.

Toujours dans le cadre d'étude avec une tâche, en étude de groupe, Zeng et coll. [84] ont classé les composantes en évaluant la ressemblance entre les CI obtenues à partir de toutes les données et celles obtenues à partir des volumes pairs (ou impairs). Ils s'appuient sur les travaux de McKeown [86] qui a montré que si on sous-échantillonnait les composantes liées à une tâche, en ne prenant que les paires ou que les impaires, on obtenait une composante qui ressemblait (décours temporel plus carte spatiale) aux originales. Les composantes liées à du bruit non-structuré seraient, elles, non reproductibles. L'idée est donc ici, de séparer les données des décours temporels des voxels pairs et impairs pour avoir deux jeux de données. Une ACI est ensuite réalisée sur chacun de ces jeux pris séparément ainsi qu'une ACI sur le jeu de données complet. Une corrélation spatiale et une corrélation temporelle sont faites entre chaque composante du jeu de données global et chaque composante du jeu de données « impair », le jeu « pair » servant de vérification. La moyenne entre les corrélations spatiales et temporelles est définie comme la corrélation moyenne. Les CI sont ainsi classées en fonction de la ressemblance à l'une des composantes estimées à partir des données sous-échantillonnées. Les composantes significatives se retrouvent en haut du classement, les composantes moins significatives en bas. Publiée sur une étude de groupe mais applicable pour le sujet unique (bien que coûteuse car de nombreuses ACI doivent être réalisées), cette méthode n'est cependant pas efficace pour détecter les composantes de bruit structuré.

Preuve s'il en est que beaucoup d'équipes se contentent encore d'une sélection manuelle, Kelly et coll. [92] proposent en 2010 une méthode de sélection manuelle avec des étapes bien définies. En effet, après avoir fait le point sur de nombreux articles présents dans la littérature, ils remarquent qu'il n'existe pas de critère commun pré-établi ni de méthode « gold standard » et proposent donc une méthode de sélection manuelle à l'aide de valeurs descriptives des composantes. Deux classes sont définies : les composantes liées à un signal neuronal d'intérêt et les composantes de bruit.

Les critères spatiaux deviennent prioritaires par rapport aux critères temporels. Ces derniers apparaissent uniquement quand il existe une indécision lors du classement. La décision est donc basée tout d'abord sur une inspection visuelle des cartes seuillées des composantes. Des améliorations restent possibles en apportant des informations sur la tâche du paradigme. Les algorithmes automatiques peuvent être utilisés en amont pour aider au classement visuel ; ces méthodes pouvant fournir des mesures de caractéristiques des CI qui sont autant d'informations supplémentaires pour la sélection manuelle.

Après le début des travaux décrits dans ce manuscrit, deux autres articles ont été publiés en 2013. Tout d'abord, une étude menée par Storti et coll. [93] ayant pour but de sélectionner les composantes liées à un réseau de repos sur sujet unique. Un second objectif de cette étude, déjà abordé au paragraphe 2.2.3.3, est d'identifier les différents réseaux sélectionnés afin de connaître leur reproductibilité d'une session à une autre. Leur algorithme de sélection se divise en quatre étapes successives basées chacune sur un descripteur : symétrie des données, agrégation spatiale, localisation spatiale, analyse spectrale. Pour les trois premières étapes, les descripteurs sont appliqués soit sur les cartes spatiales des composantes soit sur une matrice formée comme suit : chaque carte spatiale (volume 3D) d'une CI est mise sous forme de vecteurs ; lesquels sont ensuite concaténés pour former une matrice avec un nombre de lignes égal au nombre de composantes. Les trois premiers descripteurs, utilisés successivement, permettent de caractériser les voxels des cartes spatiales pour effectuer des mises à zéro de certains éléments de la matrice globale décrite ci-dessus. Dans la quatrième étape, la méthode ne prend en compte que les décours temporels des voxels restants (éléments de la matrice différents de zéro). Une analyse spectrale permet de rejeter les composantes en fonction de leur densité spectrale de puissance sur certaines bandes de fréquences. Au final, comme pour la plupart des méthodes décrites, on obtient deux classes : les composantes correspondant à un réseau de repos et les composantes artefactuelles ou de bruit. Cette méthode, bien qu'ayant des objectifs proches des nôtres (sujet unique, repos) produit un nombre de faux négatifs qui reste important (sensibilité entre 96,5 % et 98,8 % selon les sessions) au vu des différents tests présentés dans l'article et ne répond pas à notre objectif d'une sensibilité maximale.

Bhaganagarapu et coll. [94] ont également proposé une méthode de sélection des CI sur sujet unique mais à des fins d'extraction d'artefacts. Le nombre important de sujets étudiés (50) ainsi que les critères descriptifs pertinents de leur méthode, la rendent intéressante et prouvent son efficacité pour repérer les artefacts et bruits. Cette méthode utilise quatre descripteurs : mesure de lissage spatiale, mesure du bruit dans les ventricules et sur les bords du cerveau et mesure de la densité spectrale de puissance pour les fréquences au-dessous et au-dessus de 0,08 Hz. Les deux classes (artefacts ou non) sont décidées par un algorithme des k -moyennes ($k=2$) puis une règle de sélection est faite en donnant la priorité aux résultats de certains descripteurs. Cependant, les limites de cette méthode apparaissent si l'on souhaite l'étendre à une application clinique où une sensibilité maximale est requise. Elle semble en effet avoir des difficultés de décision lorsqu'une composante apparemment liée à une activité neuronale est fortement bruitée. Si l'impact d'un rejet de ce type de composantes reste mineur, il n'est pas souhaitable d'écarter des CI possiblement d'intérêt dans une étude sujet unique à portée clinique. De plus, leur méthode a été développée en extrayant, par l'ACP puis l'ACI, un nombre de composantes assez faible (54) pour des études sur sujet unique. Le type de composantes pour lesquelles leur algorithme est le moins performant, à savoir les composantes mêlant bruit et activité neuronale, sont plus nombreuses avec un ordre du modèle plus élevé.

Le principal désavantage des méthodes présentées ci-dessus est la sensibilité qui n'atteint pas 100 %. Par ailleurs, certains descripteurs proposés sont efficaces sur des données pré-traitées avec un lissage spatial, une normalisation et un faible nombre de composantes ce qui est contraire à notre souhait de rester à un niveau sujet et d'extraire les réseaux cérébraux idiosyncratiques. À titre d'exemple, le TABLEAU 5.1 expose les données utilisées, le nombre de composantes et les choix des

pré-traitements des autres méthodes destinées à des études pour sujet unique. Le choix de l'ordre du modèle (nombre de composantes étudiées), le lissage ou la normalisation modifient l'aspect des cartes spatiales et des décours temporels des composantes et ne permettent pas une application de ces méthodes sur tous types d'études. Les méthodes précédentes fournissent des descripteurs pertinents pour leurs objectifs (débruitage, groupe, tâche) mais aucune ne répond à toutes nos demandes et plus particulièrement à celle d'une sensibilité de sélection des SAM maximale.

TABLEAU 5.1 – Études sur sujet unique.

Auteur	sujets \times runs	vol / temps	CI	CI totales	N.	L.
McKeown [86]	4 \times 2	131-144 / \sim 6 min	131-144	1100		\times
Perlberg [90]	3 \times (7+3)	130-908 / 5 min	\sim 100 - \sim 900			
Thomas [87]	3 \times 1	1310 / 6,8 min	30	90		
Tohka	20 \times 6		\sim 50	6000		\times
	12 \times 5		\sim 50	3000		\times
Martino	7 \times 2	\sim 250 / \sim 8 min	60	840	\times	\times
Van de ven [88]	7 \times 2	400 / 10 min	200	2800	\times	\times
	9 \times 1	1200 / 60 min	\sim 54		\times	\times
Bhaganagarapu [94]	21 \times 1	166 / 10 min	\sim 54	2722	\times	\times
	20 \times 1	270 / 9 min	\sim 54		\times	\times
Storti [93]	25 \times 3	197 / 6,6 min	[16 ; 53]	1616		\times

Les données décrites sont, dans l'ordre des colonnes : le nombre de sujets et le nombre de sessions par sujet (runs), le nombre de volumes acquis (vol) et l'équivalent en minutes (temps), les nombres de composantes indépendantes retenues après ACP et ACI (CI), le nombre de composantes indépendantes totales à classer, s'il y a eu, ou non, une normalisation dans le pré-traitement (N) et s'il y a eu, ou non, un lissage dans le pré-traitement (L).

5.3 Point de départ : ce que disent les experts, article de Roquet et coll.

De précédentes recherches au sein de notre équipe avaient permis de mettre au point une méthode de sélection manuelle des CI avec pour objectifs, pour rappel :

- sur sujet unique,
- une sensibilité de 100 % dans la sélection des SAM,
- applicable à l'IRMF de repos ou de tâche,
- sans *a priori* spatial ou temporel.

L'article « Manual selection of spontaneous activity maps derived from independent component analysis : Criteria and inter-rater reliability study » de Roquet et coll. [82], issu de ces recherches, a donc été le point de départ de ce premier travail de thèse et sera donc décrit en détail. Par ailleurs, cette méthode présente aussi l'avantage d'avoir prouvé la fiabilité et l'indépendance de ses critères vis-à-vis des évaluateurs par un test de robustesse inter-évaluateur. Cette approche manuelle propose des critères d'acceptation et de rejet, rappelés ci-dessous, afin de classer les composantes comme d'intérêt ou non. Pour être évalués, ces critères supposent que la carte spatiale des composantes est seuillée avec un z-score de $\pm 1,6$. Nous reviendrons par la suite sur la notion de z-score. Le choix d'un z-score de 1,6 et d'un ordre du modèle (nombre de composantes en sortie) entre 150 et 250 composantes, est suggéré par Roquet et coll. dans une étude interne (et article en cours de soumission [13]). Cette valeur est un compromis entre fusion et division des SAM qui apparaissent avec respectivement moins ou plus de composantes demandées lors de l'ACP.

5.3.1 Critères de sélection

On attend d'un plausible signal BOLD qu'il réponde à chacun de ces critères pour le volume cérébral complet et sur l'ensemble du décours temporel de la composante :

- Fréquence : la majorité de la puissance du décours temporel doit se situer sous 0,08 Hz (fréquence de la réponse hémodynamique) et au-dessus de 0,004 Hz (oscillation maximum d'une période de 4 minutes). Même si ces limites ne peuvent être visualisées de manière nette, des dérives lentes du décours temporel, des valeurs aberrantes dans ce signal ou encore une fréquence très élevée permettent d'aider à visualiser les caractéristiques décrites par ce critère.
- Cohérence spatiale : les clusters 3D doivent être spatialement cohérents à l'intérieur et entre les coupes. Le cluster peut être suivi sur plusieurs coupes successives. Dans le cas d'une acquisition entrelacée, un signal apparaissant une coupe sur deux n'est pas considéré comme spatialement cohérent.
- Localisation : la distribution spatiale des clusters doit chevaucher principalement la matière grise.

Des exemples de cartes spatiales typiques devant être sélectionnées sont illustrées FIGURE 5.1, FIGURE 5.2 et FIGURE 5.3, accompagnées de leur décours temporel.

5.3.2 Critères de rejet

Afin d'éviter des composantes de bruit ou d'artefacts, aucun des aspects suivants ne doit être présent, ou de manière négligeable (où que ce soit dans la carte spatiale) ou le décours temporel de la composante.

- Un aspect d'alternance spatiale, c'est-à-dire une juxtaposition de voxels seuillés corrélés et anti-corrélés.
- Un aspect spatial bruité, c'est-à-dire quand les voxels sont légèrement significatifs et disséminés spatialement.
- Une rupture brutale ou une dérive lente sur le décours temporel.
- Un des aspects suivant :

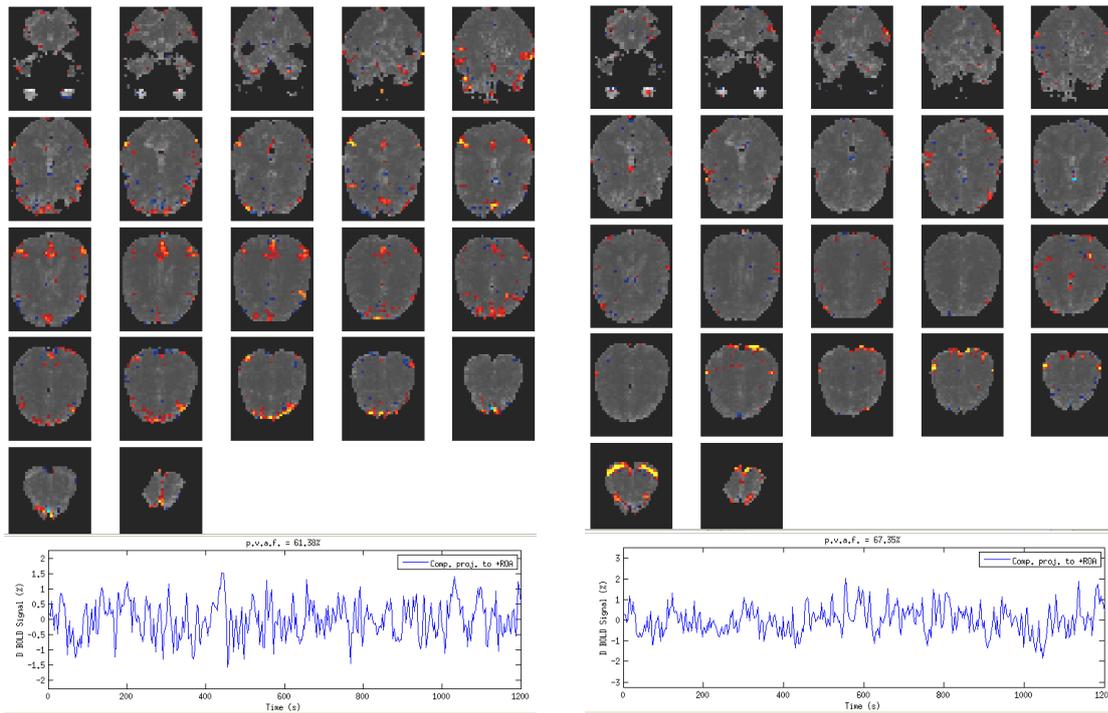


FIGURE 5.1 – DMN (Default Mode Network).

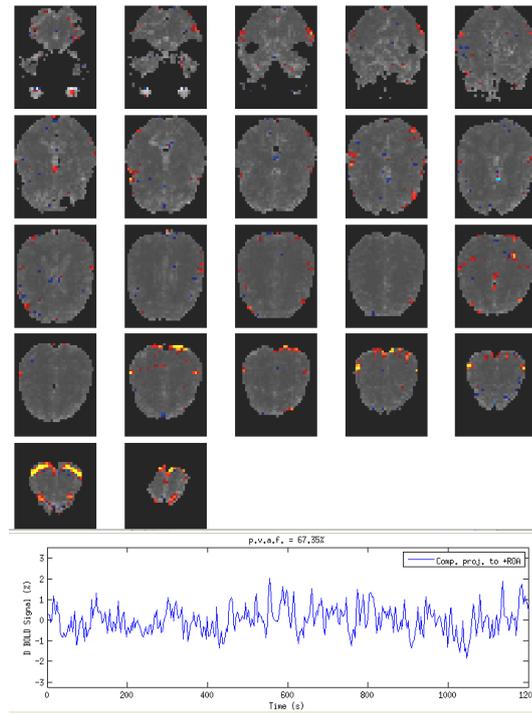


FIGURE 5.2 – Un réseau idiosyncratique.

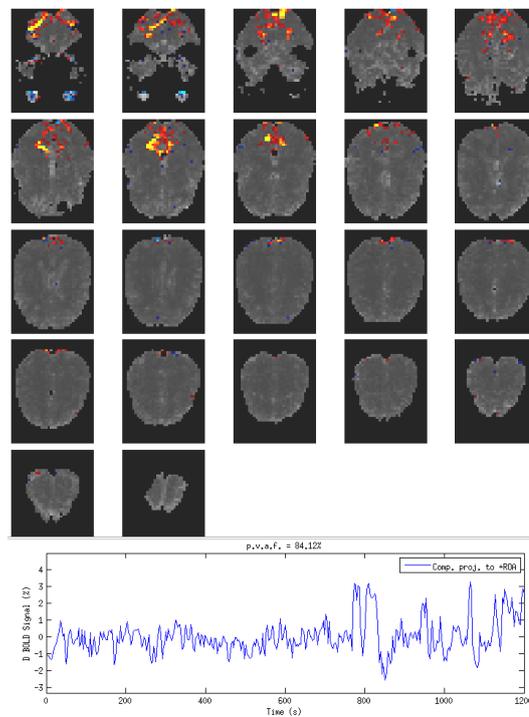


FIGURE 5.3 – MON (Median Occipital Network).

- Artefact des mouvements de la tête ;
- Artefact des mouvements oculaires ;
- Artefact dû à la pulsation dans le liquide céphalorachidien (LCR) ;
- Artefacts divers : acquisition entrelacée, instabilité sur une coupe.

Des exemples de coupes artefactées se trouvent FIGURE 5.4.

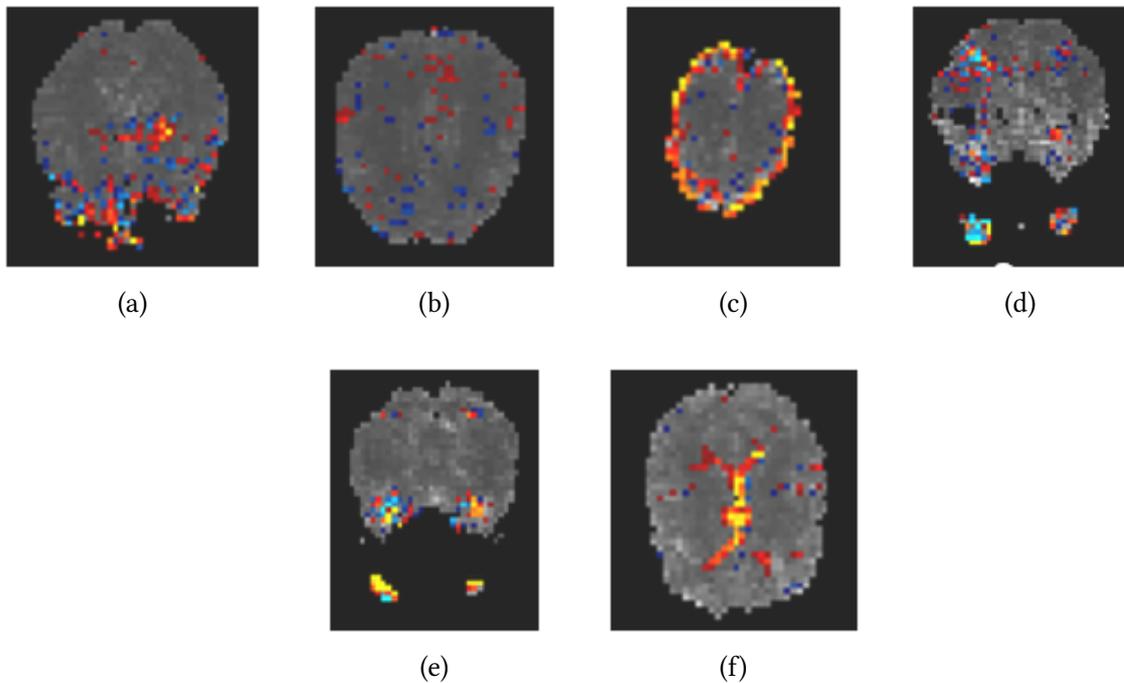


FIGURE 5.4 – (a) alternance spatiale, (b) bruit spatial, (c) artefact des mouvements de la tête, (d,e) artefacts des mouvements oculaires, (f) artefact dû à la pulsation dans le LCR

Le but de cette première partie de thèse est avant tout d'automatiser la méthode manuelle en prenant en compte les trois critères de sélection de la méthode détaillée précédemment, à savoir le critère fréquentiel, la cohérence spatiale et la localisation. Nous ne cherchons pas à quantifier tous les critères de rejet car certains sont redondants avec ceux de sélection : l'aspect bruité est modélisé par la cohérence spatiale, les ruptures ou dérives sur le décours temporel par le critère fréquentiel, les artefacts de mouvements se repèrent par la localisation. Les artefacts restants sont partiellement pris en compte par les critères de cohérence spatiale et de localisation. Bien que ces critères soient utiles pour améliorer la robustesse d'une sélection visuelle car chaque composante est mieux décrite, les descripteurs automatisant ces critères de rejet peuvent être similaires inclus dans ceux utilisés pour les critères de rejet.

Des descripteurs correspondant aux critères de sélection ou de rejet, ceux-ci se basant sur des informations admises en neurosciences, ont déjà été proposés dans la littérature. Ces derniers ont donc été étudiés et retenus dans le cas où ils correspondaient à un des critères énoncés dans cette section. Cette étude est proposée dans le paragraphe suivant. Cependant, les conditions d'acquisition des données et de pré-traitements ne permettent pas toujours une transposition des méthodes de la littérature à notre cas. Pour cela, de nouveaux descripteurs ont été établis pour répondre à la problématique de cette partie de thèse. Les descripteurs retenus, finalement, parmi ceux présentés ou développés au cours de la thèse seront détaillés par la suite (en section 6.4).

5.4 Descripteurs issus de la littérature

Comme nous venons de le signaler, la littérature présente plusieurs descripteurs en accord avec nos critères de sélection des CI. Nous allons donc les lister en les classant par étude. Avant cela, certains descripteurs peuvent être dès à présent écartés car les conditions de l'étude pour lesquels ils ont été développés sont trop éloignées des nôtres pour qu'ils soient transposables. C'est le cas de ceux qui demandent un apprentissage répété ou qui font référence à un paradigme [57, 84, 87].

Le TABLEAU 5.2 ci-dessous propose donc pour terminer ce chapitre une liste des descripteurs servant à la classification des composantes indépendantes lors des études en IRMf. Ils sont regroupés selon l'auteur qui les a proposés. Les descripteurs listés et étiquetés comme « retenus » dans le tableau sont retenus dans un premier temps. Certains ciblant le même critère de sélection, ils seront ensuite comparés pour ne retenir définitivement que les plus sélectifs.

TABLEAU 5.2: Descripteurs existants pour la sélection ou le rejet de composantes issues de l'ACI.

Descripteur	Auteur	Détails	Type de critère	Intérêt pour notre étude
corrélation	Perlberg [90]	sélection des composantes qui ont une forte corrélation avec des signaux de bruits caractéristiques	critère fréquentiel	non retenu : demande une pré-sélection manuelle des régions connues pour montrer des fluctuations physiologiques importantes.
kurtosis	De Martino [89]	mesure de l'aplanissement de la distribution de la valeur des voxels	critère de cohérence spatiale	retenu
skewness	De Martino [89]	mesure de l'asymétrie de la distribution de la valeur des voxels	critère de cohérence spatiale	retenu
entropie spatiale	De Martino [89]	mesure de l'information contenue dans la distribution spatiale	critère de cohérence spatiale	retenu : mesure prise en compte dans le descripteur de Sui [83].
degré de regroupement	De Martino [89]	mesure de la structure spatiale	critère de cohérence spatiale	retenu : mesure prise en compte dans le descripteur de Sui [83].
auto-corrélation	De Martino [89]	mesure de la structure temporelle		non retenu : ne fait pas partie des critères énoncés en 5.3.
entropie temporelle	De Martino [89]	mesure de l'information contenue dans le décours temporel		non retenu : ne fait pas partie des critères énoncés en 5.3.
densité spectrale de puissance	De Martino [89]	répartition de la puissance d'un signal suivant les fréquences	critère fréquentiel	retenu
index de Pearson	Storti [93]	mesure de l'asymétrie de la distribution de la valeur des voxels	critère de cohérence spatiale	retenu
clustering	Storti [93]	mesure de l'appartenance d'un voxel à un cluster	critère de cohérence spatiale	non retenu : l'article travaille sur les voxels des cartes spatiales et non sur la carte entière car ce critère s'insère dans un pipeline. Non utilisable en tant que tel dans notre étude.

segmentation	Storti [93]	rejet des voxels avec une probabilité d'appartenance à la matière blanche supérieure à 90 %	critère de localisation	non retenu : l'article travaille sur les voxels des cartes spatiales et non sur la carte entière car ce critère s'insère dans un pipeline. Non utilisable en tant que tel dans notre étude.
densité spectrale de puissance pour le bruit structuré	Thomas [87]	répartition de la puissance d'un signal dans les bandes de fréquences correspondant aux bruits structurés	critère fréquentiel	non-retenu : les fréquences recherchées ici sont celles correspondant aux bruits issus des signaux cardiaques et respiratoires.
rupture brutale dans le décours temporel	Tohka [91]	mesure du plus grand pic dans le décours temporel de chaque composante	critère de rejet	non-retenu : ne fait pas partie des critères énoncés en 5.3 (car redondant avec le critère fréquentiel).
localisation des voxels	Tohka [91]	mesure de la distribution des voxels dans les limites du cerveau et à l'intérieur du cerveau	critère de rejet (détecte les mouvements du cerveau)	non-retenu : ne fait pas partie des critères énoncés en 5.3 (car redondant avec le critère de localisation).
focusing degree	Sui [83]	mesure de la structure spatiale et de l'information contenue dans la distribution spatiale	critère de cohérence spatiale	retenu
corrélations spatiale et temporelle	Zeng [84]	mesure de la ressemblance entre les composantes obtenues à partir de toutes les données et celles obtenues à partir des volumes pairs		non-retenu : ne fait pas partie des critères énoncés en 5.3 et ne fonctionne que pour les détections de composantes liées à une tâche.
corrélation	McKeown [86]	mesure de la ressemblance entre un décours temporel et un paradigme	critère fréquentiel	non-retenu : tâche demandée ne fonctionne pas en resting-state.

Méthode proposée

6

6.1 Vue d'ensemble de la méthode

Nous avons développé une méthode automatique de pré-sélection de composantes indépendantes (CI) basée sur les critères définis dans l'article décrit précédemment. Après les pré-traitements décrits dans le paragraphe 6.2, cinq valeurs caractéristiques sont extraites pour chaque CI. Les seuils de décision pour chacun des descripteurs sont calculés sur la base d'apprentissage et reportés sur les CI de la base de test. Un ET logique est appliqué pour fusionner les décisions et le choix est fait de classer une CI comme SAM plausible ou non-SAM. Une dernière inspection est faite en suivant la méthode de sélection manuelle [82], voir section 5.3, sur les CI étiquetées comme plausibles SAM afin d'écarter les faux positifs restants. La FIGURE 6.1 offre une vue générale de la méthode.

Les données utilisées proviennent des bases B1, B2_{3s}, B3 et B4 décrites en Partie I, chapitre 4 page 37. L'apprentissage de la méthode a été effectué sur une partie de ces données ; la base d'apprentissage et celle de test utilisées par la suite ont été détaillées section 4.2.1 page 39.

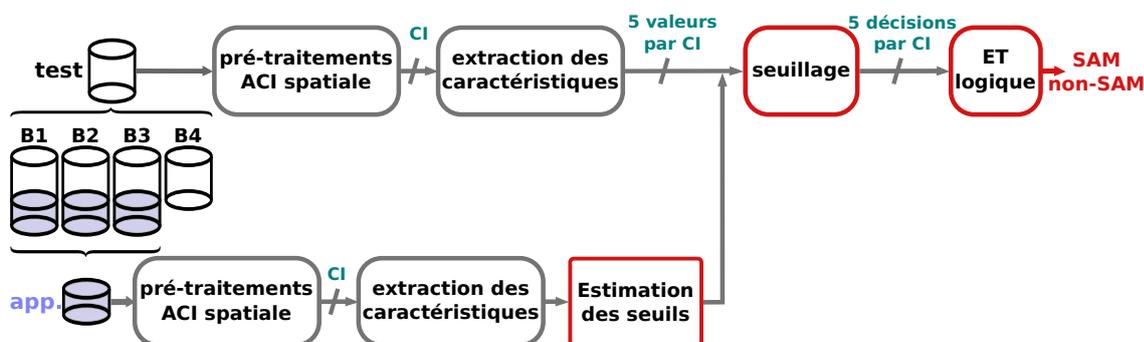


FIGURE 6.1 – Synoptique de la méthode : des données BOLD à la pré-sélection de SAM. 4 étapes sont nécessaires pour obtenir un jeu de composantes étiquetées comme SAM : pré-traitements des données et ACI ; extraction des caractéristiques à l'aide des descripteurs ; seuillage ; ET logique. Les valeurs des seuils sont obtenues sur une base dédiée à l'apprentissage (en bleu sur la figure).

De nombreux descripteurs ont été étudiés, certains développés afin de modéliser au mieux les critères de sélection d'autres issus de la littérature (voir TABLEAU 5.2). Le choix des cinq descripteurs utilisés au final va être explicité dans ce chapitre. Une fois les descripteurs établis, en écartant ceux redondants ou peu sélectifs, une règle de sélection ainsi qu'une règle de fusion des descripteurs sont appliquées pour classer chaque CI et feront l'objet des sections suivantes. Enfin, une dernière section permettra de proposer des comparaisons avec d'autres méthodes d'identification, d'analyse ou de sélection de CI.

6.2 Pré-traitements et outils

Avant la sélection et l'identification de SAM par l'ACI, plusieurs pré-traitements usuels lors des études d'IRMf (voir paragraphe 1.3.3.2) sont réalisés via la boîte à outils "Statistical Parametric Mapping" [96] (développée sous Matlab [97]) :

- Un filtrage passe-bas est réalisé sur les signaux issus d'une séquence ASL-BOLD (bases de données B2_{3s} et B4), afin de filtrer les fréquences spécifiques à l'ASL. La fréquence de coupure est déterminée selon l'article de Chuang [98], environ 0,11 Hz pour nos données.
- Une correction du décalage temporel inter-coups (*slice-timing*) permet de corriger les effets dus au délai inhérent à l'acquisition des coupes, effet d'autant plus important lors d'acquisitions entrelacées [23].
- Un recalage rigide des volumes permet de compenser partiellement les artefacts dus aux mouvements du sujet.
- Enfin, pour réduire le nombre de composantes, une ACP est réalisée comme cela a été décrit en section 2.2.1 page 21. Le choix du nombre de composantes, aussi appelé ordre du modèle, a été fait selon les recommandations de Roquet et coll. [49,99]. Cette étude interne sur le sujet unique avait mis en évidence qu'un très grand nombre de composantes devait être conservées suite à l'ACP afin de maximiser la sensibilité et la reproductibilité des co-activations cérébrales. La sensibilité a été calculée en fonction de l'ordre du modèle à partir du nombre de SAM, de nouvelles SAM et du pourcentage de volume cérébral impliqué dans ces SAM. La reproductibilité a été mesurée sur la base du nombre de divisions d'une SAM en plusieurs SAM, ainsi que sur la variation du volume des SAM. Le nombre de composantes retenues pour chaque base a été présenté dans le TABLEAU 4.1 du chapitre 4.

L'ACI est réalisé par FMRLAB [100] (développé également sous Matlab); l'algorithme utilisé étant InfoMax (voir section 2.1.2 page 20). En sortie de ces traitements, on obtient le nombre de CI souhaité; cela se traduit par une carte spatiale et un décours temporel par CI comme illustré FIGURE 6.2. La carte spatiale reflète les variations spatiales du signal BOLD exprimées en z-score. Ce dernier représente l'écart-type par rapport à la moyenne, sur la distribution des voxels. En appliquant un certain seuil de z-score, on retrouve dans les couleurs rouges, les voxels dont les décours temporels sont corrélés entre eux et dans les couleurs bleues, les voxels dont les décours temporels sont anti-corrélés aux décours temporels des voxels rouges. On les dénommera par la suite comme, respectivement, les voxels seuillés corrélés et les voxels seuillés anti-corrélés. Le décours temporel fournit une information sur l'évolution de l'activité de ce réseau au cours du temps (au cours de la session d'IRMf plus précisément). L'analyse des cartes spatiales et décours temporels se fait principalement sur les voxels seuillés corrélés, l'ACI effectué par FMRLAB (et les autres logiciels d'ACI dédiés à l'IRMf) privilégiant les valeurs positives (réorientation de chaque composante pour que la plus grande variance soit expliquée par la partie positive). Les voxels seuillés anti-corrélés donnent des informations supplémentaires mais non indispensables à la sélection des composantes. Le seuil appliqué sur les cartes spatiales afin de localiser les activations cérébrales significatives est fixé à $\pm 1,6$. Il correspond, en théorie, à $\pm 1,6$ fois l'écart-type soit une marge d'erreur de 5 % sur les données positives.

Il est important de faire un point sur le décours temporel fourni par FMRLAB puisque ce ne sont pas exactement les coefficients de mélange comme nous avons pu le voir en section 2.1.2 page 20. FMRLAB propose deux types de signaux pour les décours temporels :

- La première façon d'obtenir le décours temporel d'une composante *comp* en fonction du temps *t* est donnée par l'équation (6.1). On calcule donc une moyenne sur les signaux BOLD

d'origine des voxels seuillés corrélés.

$$DT1_{comp}(t) = \frac{1}{Nvsc} \sum_{i=vsc_{comp}} \mathbf{X}(t, i) \quad (6.1)$$

avec vsc_{comp} les indices des voxels seuillés corrélés pour la composante, $Nvsc$ le nombre de voxels seuillés corrélés pour la composante et \mathbf{X} la matrice des signaux BOLD d'observation de taille $(T \times M)$, T le nombre de volumes et M le nombre de voxels (voir Partie I section 2.1.2 pour la formalisation de l'ACI).

- La seconde possibilité, présentée par équation (6.2), est d'utiliser la matrice de mélange estimée. (voir $\hat{A} = W^{-1}$ dans la Partie I section 2.1.2). Dans ce cas, le décours temporel est une moyenne sur les signaux BOLD reconstruits à partir des coefficients calculés par l'ACI et de la source de la composante étudiée. Le moyennage se fait encore uniquement sur les voxels significatifs, les voxels seuillés corrélés.

$$\begin{aligned} S_{comp}(k, i) &= S(k, i) \text{ si } k \text{ est la composante étudiée, } 0 \text{ sinon, } \forall i \in [1; M], \\ & \quad M \text{ le nombre de voxels d'un volume} \\ \hat{\mathbf{X}} &= W^{-1} \cdot S_{comp} \\ DT2_{comp}(t) &= \frac{1}{Nvsc} \sum_{i=vsc_{comp}} \hat{\mathbf{X}}(t, i) \end{aligned} \quad (6.2)$$

avec \mathbf{S} la matrice des signaux sources $(K \times M)$, K le nombre de sources (les composantes) et M le nombre de voxels, W la matrice de séparation $(T \times K)$.

Bien que ces deux signaux soit très proches, nous avons retenu $DT2$ pour la suite de nos travaux afin que l'information contenue dans les coefficients de mélange soit prise en compte. L'amplitude des décours temporels correspond à l'écart à la moyenne du signal BOLD. Il est exprimé en pourcentage.

6.3 Descripteurs existants

Le chapitre précédent nous a permis de retenir certains descripteurs de la littérature qui nous paraissaient pertinents au regard de notre cahier des charges (voir TABLEAU 5.2).

Les descripteurs de la littérature retenus à ce stade de l'étude sont donc :

- **Focusing degree** : Le calcul du degré d'agrégat, formule (6.3), est proposé par De Martino [89]. Il permet une mesure de la structure de la carte spatiale de la composante évaluée en dénombrant les voxels seuillés corrélés faisant partie d'un agrégat. Sui [83] propose, par la suite, d'observer le rapport entre cette mesure et l'entropie de Shannon (formule (6.4)) sur la valeur des voxels pour créer le descripteur *focusing degree*, formule (6.5). Ce descripteur capture, quant à lui, l'information contenue dans la distribution spatiale. On s'attend à un score plus fort pour les composantes que l'on souhaite sélectionner, les réseaux présents dans ces composantes se manifestant sous forme d'agrégats. Pour ce descripteur, nous appelons agrégat un ensemble 3D (sur plusieurs coupes consécutives) de voxels voisins 6-connexes présentant un volume supérieur à 400 mm³ (soit entre 6 et 15 voxels pour nos bases de

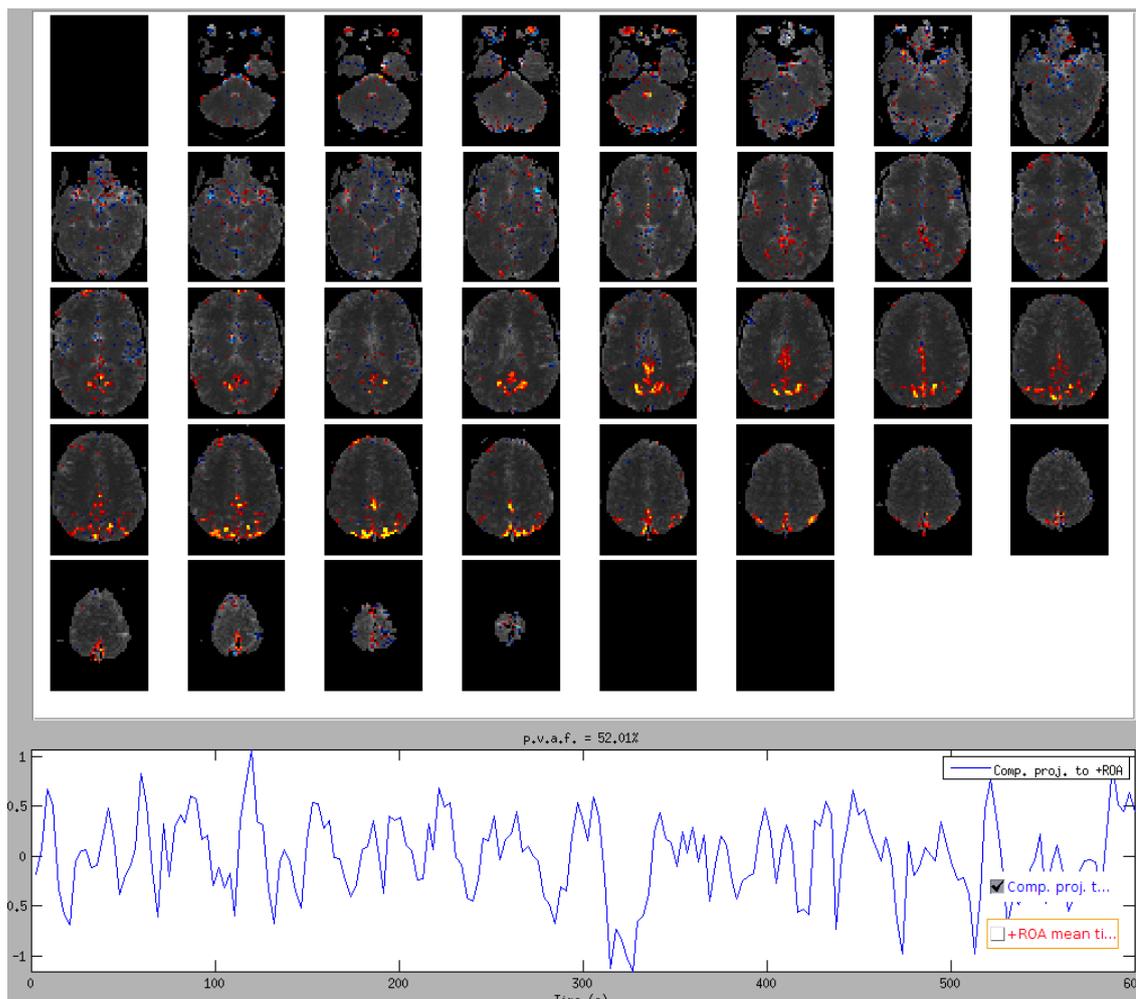


FIGURE 6.2 – La carte spatiale d’une composante et le décours temporel associé (DT^2).

données).

$$DegreeCl = \frac{Na}{Nvsc} \quad (6.3)$$

$$EntroDis = \left| \ln \left(\sum_{b=1}^{Nb} hs(b) \cdot \log_2(hs(b)) \right) \right| \quad (6.4)$$

$$focusing\ degree = \frac{DegreeCl}{EntroDis} \quad (6.5)$$

avec Na , le nombre de voxels appartenant à un agrégat

$Nvsc$, le nombre de voxels seuillés corrélés

hs , l’histogramme de la valeur des voxels de la composante étudiée

- **Kurtosis** : Cette mesure nous renseigne sur l’aplatissement de la distribution des valeurs en z-score des voxels de la carte spatiale. On s’attend à un kurtosis plus élevé pour les composantes d’intérêt : alors que les voxels appartenant au réseau (voxels seuillés corrélés) sont faibles en nombre et élevés en valeur de z, les valeurs des voxels du reste de l’encéphale sont plus proches de la moyenne (nulle). Il y a donc moins de déviation que pour les composantes de bruit.

$$kurtosis = \frac{\sum_{k=1}^{Nvs} z_k^4}{Nvs} - 3 \quad (6.6)$$

avec Nvs , le nombre de voxels seuillés
 z , les valeurs en z-score de la composante étudiée

- **Skewness** : Nous évaluons, avec cette mesure, l'asymétrie de la distribution de la valeur des voxels de la carte spatiale de la composante. On s'attend à des valeurs proche de zéro pour les composantes de bruit (distribution faiblement asymétrique). Pour les composantes d'intérêt présentant un réseau de voxels seuillés corrélés, ce score est positif puisque la distribution est plus longue à droite de la moyenne, c'est à dire pour les valeurs élevées. Il est à noter que certains artefacts structurés (comme la pulsation du LCR) peuvent présenter également un skewness élevé.

$$skewness = \frac{\sum_{k=1}^{Nvs} z_k^3}{Nvs} \quad (6.7)$$

avec Nvs , le nombre de voxels seuillés
 z , les valeurs en z-score de la composante étudiée

- **Indice de skewness de Pearson** : Cet indice est un autre moyen d'évaluer l'asymétrie d'une distribution. Comme pour le descripteur *skewness*, on s'attend à un score plus élevé pour les composantes avec une structure spatiale que pour les composantes liées à du bruit. Le calcul se fait également sur les voxels seuillés de la carte spatiale des composantes.

$$Pearson = \frac{3 \cdot (\bar{x} - \tilde{x})}{\sigma} \quad (6.8)$$

avec x , les voxels seuillés
 \bar{x} , la moyenne de la distribution
 \tilde{x} , la médiane
 σ , l'écart-type

- **Densité spectrale de puissance** : Calculée à partir du décours temporel de la composante, cette mesure nous renseigne sur la répartition de la puissance d'un signal suivant la fréquence. La bande de fréquences où se situe l'information des signaux d'intérêt est [0,004 Hz ; 0,08 Hz] (voir section 5.3).

$$p(f) = \frac{1}{T} |DT2(f)|^2 \quad (6.9)$$

$$dsp = \frac{\sum_{f \in F_{cible}} p(f)}{\sum_{f \in F} p(f)} \quad (6.10)$$

avec $DT2$, la transformée de Fourier du décours temporel,
 f , la fréquence évaluée
 F_{cible} , la bande de fréquence [0,004 Hz ; 0,08 Hz]

6.4 Descripteurs nouveaux

Ces descripteurs sont réalisés afin de modéliser, au plus proche, les critères établis par l'article de sélection manuelle [82].

- **Fréquence cumulée** : Ce descripteur s'intéresse exclusivement au décours temporel de la composante. Le calcul permet de dégager le pourcentage du spectre en fréquence du décours temporel qui se situe entre 0,004 Hz et 0,08 Hz, conformément aux critères de sélection décrits en 5.3.

$$fréquence\ cumulée = \frac{\sum_{f=0,004}^{0,08} DT2(f)}{\sum_f DT2(f)} \quad (6.11)$$

avec $DT2$, la transformée de Fourier du décours temporel et f , la fréquence

- **Localisation** : Ce descripteur permet de quantifier le chevauchement entre des voxels seuillés corrélés et la matière grise. On réalise une segmentation du cerveau à partir de l'image anatomique T1. La toolbox « NewSegment », fournie par SPM, produit 5 cartes : la matière grise (c1), la matière blanche (c2), le liquide céphalo-rachidien (c3), les méninges (c4) et le crâne (c5). Les corps cellulaires des neurones étant localisés dans la matière grise, on s'attend à obtenir pour les composantes d'intérêt un score plus élevé pour ce descripteur :

$$localisation = \frac{Nvsc_{c1}}{Nvsc} \quad (6.12)$$

avec $Nvsc_{c1}$, le nombre de voxels seuillés corrélés dans la matière grise
 $Nvsc$, le nombre de voxels seuillés corrélés

- **Alternance** : Ce descripteur met en évidence les composantes dont les voxels seuillés corrélés ne sont présents qu'une coupe sur deux. Le slice-timing ne permet pas de compenser entièrement les effets de l'ordre d'acquisition des coupes, ce descripteur vient donc pallier ce problème. Quand l'alternance n'est que partielle, les composantes avec cet artefact ne sont pas éliminées par le descripteur de cohérence spatiale *focusing degree*. Cela ne concerne, dans les faits, qu'un nombre très restreint de composantes (moins d'une dizaine sur l'ensemble des composantes du sujet) et uniquement les données obtenues par une acquisition entrelacé (bases B1 et B3 pour notre étude) mais permet néanmoins d'améliorer la spécificité. Plus le score à ce descripteur (6.13) est grand, plus l'alternance entre les coupes paires et impaires est marquée :

$$alternance = \frac{\left| \sum_i Nvsc_{2i} - \sum_i Nvsc_{2i-1} \right|}{N_{vox}} \quad (6.13)$$

avec $1 \leq i \leq \frac{C}{2}$ et C le nombre de coupes dans un volume
 $Nvsc_i$, le nombre de voxels seuillés corrélés dans la coupe i
 N_{vox} , le nombre de voxels qui forment l'image du cerveau

- **Volume** : Ce dernier descripteur a pour but d'éliminer les composantes de non-intérêt qui présentent une instabilité sur une coupe. On remarque pour ces composantes, un agrégat de voxels seuillés corrélés ou un agrégat de voxels seuillés anti-corrélés sur une coupe seulement. De même que pour le descripteur précédent, ce type d'artefact n'est présent que sur

un nombre limité de composantes (moins d'une dizaine sur l'ensemble des composantes du sujet). Ici, on regarde indifféremment les voxels seuillés corrélés et anti-corrélés, puisque ce qui nous importe est un changement brutal de ce nombre de voxels d'une coupe à celle qui la suit.

$$volume = \max_{i \in Coupes} \left(\frac{Nvs_i}{Nvox_i} - \frac{Nvs_{i-1}}{Nvox_{i-1}} \right) \quad (6.14)$$

avec Nvs_i , le nombre de voxels seuillés dans la coupe i

$Nvox_i$, le nombre de voxels qui forment l'image du cerveau dans la coupe c

$Coupes$, l'ensemble des coupes de la composante

6.5 Comparaison des descripteurs

Parmi les neuf descripteurs, certains paraissent répondre aux mêmes critères de sélection manuelle (fréquence, cohérence spatiale ou localisation). Il convient donc dans un premier temps d'évaluer la similitude des descripteurs et de ne retenir que les plus spécifiques parmi ceux redondants. Deux descripteurs sur les neuf (*alternance* et *volume*) ont un statut particulier puisqu'ils concernent deux critères de rejet et ciblent des composantes de non-intérêt particulières. Très peu de composantes sont concernées par ces critères et ces descripteurs sont efficaces pour les repérer. Ils sont donc directement gardés pour la suite du processus de sélection et ne sont pas comparés aux autres dans ce paragraphe. La matrice de corrélation entre les descripteurs ou des tests statistiques de multi-colinéarité peuvent nous renseigner sur la similitude ou la dépendance entre deux descripteurs. Cependant, ils ne sont pas adaptés pour évaluer les descripteurs dans le cas très particulier d'une sensibilité (Se) de 100%. Cette contrainte particulière nous a menés vers une approche empirique dans un premier temps afin de comparer les descripteurs.

6.5.1 Méthodes d'évaluation

Les études de classification automatique ou semi-automatique, publiées récemment et décrites dans le chapitre 5, ont été validées en comparant leurs résultats avec ceux obtenus via une inspection visuelle par des experts [83, 90, 91], les sélections des experts étant considérées comme un *gold standard*. Cette validation est d'autant plus adaptée à notre cas puisque la méthode d'évaluation utilisée par nos experts a été publiée et sa robustesse inter-évaluateur a été démontrée [82].

On souhaite donc sélectionner les composantes correspondant à des réseaux cérébraux. Elles seront nommées, ci-après, composantes d'intérêt (à ne pas confondre avec l'abréviation CI pour composante indépendante). Les composantes de bruit et d'artefacts, qu'on souhaite rejeter, seront qualifiées par la suite de composantes de non-intérêt. Les résultats sont évalués via les outils statistiques standards. On calcule les valeurs suivantes :

- VP (vrai positif) représente le nombre de composantes d'intérêt sélectionnées ;
- FP (faux positif) représente le nombre de composantes de non-intérêt sélectionnées ;
- FN (faux négatif) représente le nombre de composantes d'intérêt rejetées ;
- VN (vrai négatif) représente le nombre de composantes de non-intérêt rejetées.

On peut ensuite calculer :

- La sensibilité : taux de composantes d'intérêt correctement sélectionnées.

$$Se = \frac{VP}{VP + FN}$$

- La spécificité : taux de composantes de non-intérêt correctement rejetées.

$$Sp = \frac{VN}{VN + FP}$$

- La valeur prédictive positive : parmi les composantes sélectionnées, taux de composantes d'intérêt.

$$Pp = \frac{VP}{VP + FP}$$

Ici, on souhaite obtenir une sensibilité maximale ($Se = 100\%$). En effet, l'étude se faisant sur sujet unique, tous les réseaux doivent être sélectionnés afin de réaliser le meilleur diagnostic. La valeur prédictive positive permet une quantification de la qualité de la sélection.

6.5.2 Similitude empirique

Nous choisissons donc de partir du résultat d'une sélection qui garantit $Se=100\%$ afin de comparer les descripteurs deux à deux. Une telle sélection peut se faire de manière efficace en seuillant chaque descripteur (*focusing degree*, *kurtosis*, *skewness*, *dsp*, *Pearson*, *fréquence cumulée*, *localisation*) de manière à obtenir toutes les composantes d'intérêt. Le seuil est donc défini comme le minimum des scores des composantes d'intérêts pour chaque descripteur.

$$seuil_{descripteur} = \min_{SAM}(descripteur) \quad (6.15)$$

avec *descripteur* parmi : *focusing degree*, *kurtosis*, *skewness*, *dsp*, *Pearson*, *fréquence cumulée* et *localisation*.

On s'attend en effet à des valeurs plus élevées pour les composantes d'intérêt que pour celles de non-intérêt pour ces descripteurs. Seuls les descripteurs *alternance* et *volume* fourniront un score logiquement plus faible pour les composantes d'intérêt mais comme nous l'avons déjà dit ils ne sont pas pris en compte dans cette comparaison car ils ciblent des composantes de non-intérêt particulières. Ces seuils sont les seuls qui nous garantissent une sensibilité de 100%. Ils sont calculés sur les trois bases d'apprentissage disponibles soit 7 888 composantes (voir TABLEAU 4.3).

Soit S_i l'ensemble des composantes sélectionnées par le descripteur i , le taux de similitude entre deux descripteurs i et j se calcule :

$$taux_{similitude}(i, j) = \frac{Card(S_i \cap S_j)}{Card(S_i)} \quad \forall i, j \text{ pour } i \neq j \quad (6.16)$$

Si le $taux_{similitude}$ est élevé, cela signifie que la sélection du descripteur i est en grande partie redondante avec celle de j . De plus si le $taux_{similitude}(i, j)$ est plus élevé que $taux_{similitude}(j, i)$ alors le descripteur i est plus spécifique que le j puisqu'il sélectionne peu d'autres composantes (de non-intérêt) en dehors de cette intersection.

Le TABLEAU 6.1 met en avant les paires qui ont un score supérieur à 95% (en gras supérieur à 98%). On en déduit le descripteur inutile car redondant et moins spécifique. Comme on pouvait s'y attendre le descripteur *dsp* est fortement similaire au descripteur *fréquence cumulée* mais moins spécifique (le nombre de composantes de non-intérêt sélectionnées est plus important). Il est donc écarté. Les descripteurs *skewness* et *Pearson* mesurant tous deux l'asymétrie de la distribution, il paraît évident qu'ils fournissent des résultats très proches. Cependant ces deux descripteurs ayant de fortes similitudes de résultats avec d'autres descripteurs (*fréquence cumulée*, *focusing degree*) tout en étant moins spécifiques, ils sont aussi écartés. Il en est de même pour le descripteur *kurtosis* redondant avec le descripteur *localisation*. Les descripteurs restants, les plus significatifs, sont donc la *fréquence cumulée*, la *localisation* et le *focusing degree*.

TABLEAU 6.1 – Similitude entre les descripteurs et descripteurs redondants à retirer.

paire i - j	$taux_s(i, j) > 95\%$	$taux_s(j, i) > 95\%$	descripteur à retirer
$f. cum.$ - $f. deg.$	0	0	
$f. cum.$ - localisation	0	0	
$f. cum.$ - dsp	1	0	dsp
$f. cum.$ - kurtosis	0	0	
$f. cum.$ - Pearson	0	0	
$f. cum.$ - skewness	1	0	skewness
$f. deg$ - localisation	0	0	
$f. deg$ - dsp	0	0	
$f. deg$ - kurtosis	0	0	
$f. deg$ - Pearson	1	0	Pearson
$f. deg$ - skewness	1	0	skewness
localisation - dsp	0	0	
localisation - kurtosis	1	0	kurtosis
localisation - Pearson	0	0	
localisation - skewness	0	0	
dsp - kurtosis	0	1	dsp
dsp - Pearson	0	0	
dsp - skewness	1	0	skewness
kurtosis - Pearson	0	0	
kurtosis - skewness	1	0	skewness
Pearson - skewness	1	0	skewness

$taux_{similitude}(i, j)$, défini en équation (6.16), abrégé $taux_s(i, j)$ dans le tableau. Un $taux_s(i, j)$ plus élevé qu'un $taux_s(j, i)$ signifie que le descripteur i est plus spécifique que le descripteur j .

6.6 Distributions des valeurs descriptives et règle de sélection

Chaque composante possède à ce stade un vecteur de valeurs descriptives composé des scores aux descripteurs *fréquence cumulée*, *localisation*, *focusing degree*, *alternance* et *volume*. Pour chacun des descripteurs, nous allons pouvoir estimer un seuil de sélection qui permettra de pré-classer la composante comme composante d'intérêt ou non avant de fusionner ces différentes classification pour décider d'un choix final.

Les seuils ont été estimés de la même manière que précédemment, sur la base de données d'apprentissage, soit 7 888 composantes provenant de 34 sessions différentes. Les scores obtenus par les CI de la base d'apprentissage nous permettent de tracer la distribution des composantes d'intérêt (les SAM) et les composantes de non-intérêt (les non-SAM) pour chaque descripteur. Afin de respecter la forte contrainte d'une sensibilité de 100 %, le seuil a été choisi comme les extrema des distributions des SAM avec une marge d'erreur de 2,5 %. Pour le *focusing degree*, la *fréquence cumulée* et la *localisation*, on s'attend à un score plus fort pour les SAM. C'est donc le minimum des distributions des SAM qui a permis de calculer le seuil. Concernant les deux autres critères, l'*alternance* et le *volume*, un score élevé dénote une plus grande altération de la composante d'un point de vue spatiale. On a donc retenu, pour les seuils de ces descripteurs, le maximum des

distributions associées aux SAM.

La FIGURE B, en annexe, présente les distributions des valeurs des SAM des cinq descripteurs pour la base d'apprentissage et pour chaque sous-base de données qui la compose. On observe déjà certains critères qui paraissent plus discriminants entre les deux groupes (composantes d'intérêt, composantes de non-intérêt). Les descripteurs *fréquence cumulée* et *focusing degree* semblent se détacher et mieux décrire les deux populations.

6.7 Règle de fusion

C'est en concertation avec les experts que s'est décidée la méthode utilisée pour fusionner les décisions prises sur chaque descripteur après l'étape du seuillage. En effet, aucun critère ne devait être plus considéré qu'un autre, de manière à retranscrire au mieux la méthode manuelle. Aucune pondération n'a donc été envisagée. Pour ce faire, on applique un ET logique entre les descripteurs. Une composante est considérée comme d'intérêt si elle sélectionnée comme d'intérêt par chacun des descripteurs.

6.8 Alternatives possibles à chaque étape de la méthode

Aucune autre méthode répondant totalement à nos objectifs définis chapitre 5.3 n'a été publiée. Il ne nous est donc pas possible de comparer notre résultats directement. Cependant, trois étapes de notre approche (définition de SAM avec l'ACI sujet unique, règle de décision par seuillage, règle de fusion par ET logique) peuvent être aisément remplacées pour comparaison. Ainsi, trois autres approches ont remplacées tour à tour les nôtres pour un ou plusieurs étapes et nous ont permis d'évaluer nos résultats.

6.8.1 Identification des SAM

6.8.1.1 ACI de groupe

Comme première comparaison, l'ACI de groupe a été utilisée pour extraire les SAM [42] communes en lieu et place de l'ACI sujet unique de notre méthode. L'ACI de groupe, décrite au paragraphe 2.2.1.2, permet d'extraire les RSN de la population étudiée. Même si par définition, elle ne permet pas d'extraire les SAM idiosyncratiques, il est néanmoins intéressant d'évaluer cette méthode afin de valider (ou non) l'approche sujet unique défendue pour les applications cliniques, c'est à dire dans notre cas de figure.

Pour sa mise en œuvre, l'ACI de groupe nécessite que les images soit normalisées (recalage sur un modèle commun) et lissées spatialement (avec un noyau Gaussien de 8 mm). L'ACI de groupe a été réalisée avec le logiciel GIFT [45] (sous Matlab). Une reconstruction arrière permet d'estimer les composantes au niveau de chaque sujet [46]. Seules les bases B1 et B3 (BOLD) ont été soumises à cette comparaison puisque le but n'était pas tellement une comparaison base par base mais plutôt la présentation par l'exemple de ce que peut donner une étude de groupe.

6.8.2 Règle de décision

6.8.2.1 Algorithme des k -moyennes

Plus adaptée à notre cadre de travail, nous pouvons comparer une autre façon de décider des composantes d'intérêt ou de non-intérêt suite à l'extraction des caractéristiques par les cinq descripteurs. Notre méthode propose un seuillage brut par le maximum ou le minimum pour chaque descripteur. Nous allons la comparer avec l'algorithme bien connu des k -moyennes [101]. Cet algorithme est utilisé par deux autres méthodes de sélection des SAM en sujet unique [90, 94].

L'algorithme est appliqué séparément sur les valeurs en sortie de chacun des cinq descripteurs. Pour chaque descripteur, avec $k = 2$, la méthode permet de diviser les valeurs en deux classes, une pour les CI d'intérêt et une autre pour les CI de non-intérêt. La décision finale avec le ET logique a été conservée. De cette manière, une CI est considérée comme SAM si elle a été classée comme d'intérêt pour les cinq descripteurs.

6.8.3 Règle de décision et règle de fusion

Des algorithmes de classification peuvent permettre de remplacer la règle de décision (le seuillage dans notre cas) et la règle de fusion (le ET logique dans notre cas). Deux d'entre eux, l'analyse factorielle discriminante et l'algorithme des forêts aléatoires, ont été implémentés. Ils sont appliqués à la suite de l'extraction des cinq caractéristiques par composante.

6.8.3.1 Analyse factorielle discriminante

L'analyse factorielle discriminante (AFD) en pondérant chaque descripteur pour améliorer sa classification, ne correspond pas à la demande initiale de considérer de manière égale les différents critères énoncés (cf section 5.3). Cependant, cette pondération nous renseigne sur l'importance de chacun des descripteurs dans la décision finale. L'analyse factorielle discriminante (AFD) est à la fois explicative, car elle permet d'identifier les caractéristiques de chaque groupe suivant les variables, et prédictive, puisqu'on souhaite en déduire une règle de décision afin de définir le groupe d'appartenance pour une composante. Ce dernier point n'est cependant pas évident, car encore une fois la contrainte de sensibilité maximale nous empêche de nous contenter d'un compromis, entre spécificité et sensibilité, basé sur un coût minimum (nombre d'erreurs, fausses alarmes ou faux négatifs minimums).

Nous avons dans notre cas deux groupes (composantes d'intérêt et composantes de non-intérêt) et cinq variables explicatives quantitatives (*fréquence cumulée*, *focusing degree*, *volume*, *alternance* et *localisation*). Plusieurs hypothèses sont vérifiées avant d'appliquer l'AFD :

- Le modèle de l'AFD choisi dépend d'une hypothèse fondamentale. Si l'on suppose que les matrices de covariances sont identiques, on retiendra l'analyse factorielle discriminante linéaire, sinon on choisira le modèle quadratique. Le test de Box permet de vérifier cette hypothèse (voir TABLEAU 6.2 et TABLEAU 6.3 :

TABLEAU 6.2 – Test de Box (approximation asymptotique du χ^2).

p -valeur	< 0,0001
alpha	0,05

TABLEAU 6.3 – Test de Box (approximation asymptotique du F de Fisher).

p -valeur	< 0,0001
alpha	0,05

▷ Pour nos données, la p -value calculée est inférieure au niveau de signification alpha (=0,05). Il faut rejeter l'hypothèse H_0 d'égalité des covariances intra-classe et se diriger vers un modèle quadratique.

- Pour déterminer si les groupes présentent des moyennes significativement différentes, condition nécessaire à l'AFD, on procède au test du lambda de Wilks (voir TABLEAU 6.4).

TABLEAU 6.4 – Test du Lambda de Wilks (approximation de Rao).

p -valeur	< 0,0001
alpha	0,05

▷ Pour nos données, l'hypothèse alternative est retenue : les vecteurs moyens des deux groupes sont significativement différents. L'analyse discriminante a donc un sens.

Calculée sur la base d'apprentissage, l'AFD fournit en sortie les coefficients à appliquer à chaque terme des fonctions quadratiques de classification. Les valeurs de ces coefficients sont adaptées aux valeurs des descripteurs et ne sont pas comparables en l'état puisque les plages de valeurs diffèrent les unes des autres. L'AFD fournit néanmoins des valeurs de projection factorielle qui donnent une indication sur l'importance de chaque descripteur pour la classification.

6.8.3.2 Algorithme des forêts aléatoires

La seconde méthode de classification implémentée comme alternative à l'estimation des seuils et à la règle de fusion est l'algorithme des forêts aléatoires (plus connus sous le nom anglais de *random forest*) [102]. Cette procédure réalise l'apprentissage sur plusieurs arbres de décision, chacun étant entraîné avec un échantillon de données extraites aléatoirement de la base d'apprentissage. Pour chacun de ces échantillons, on estime un arbre dont le choix des variables pour découper chaque nœud de l'arbre est également aléatoire. La collection d'arbres obtenus est enfin agrégée pour donner les classifications finales.

Résultats et analyse des résultats

7

Ce chapitre présente et discute les résultats obtenus sur notre base de test. Ils sont analysés au regard des différentes bases utilisées, des différents descripteurs développés et des résultats obtenus en changeant certaines étapes par des méthodes alternatives (exposées dans la section précédente). Des perspectives d'amélioration sont présentées en fin de partie.

7.1 Analyse des résultats par base

Les résultats de la sélection sont présentés dans le TABLEAU 7.1. Au total, 14 418 composantes ont été étiquetées. Parmi elles, 1 080 ont été annotées comme SAM par deux experts (voir paragraphe 6.5.1), contre 4 366 pour notre méthode. Cela représente 92,5 % de composantes rejetées avec la sélection manuelle contre environ 70 % avec la méthode proposée dans cette partie. Notre méthode automatique a permis d'atteindre une sensibilité très proche des 100 % désirés (sensibilité moyenne sur l'ensemble de la base de test de 99,7 %) tout en réduisant de manière substantielle le nombre de composantes restant à sélectionner. Avec une valeur prédictive positive globale de 23,3 % par rapport à 7,5 % sans pré-sélection, la pré-sélection automatique a permis une augmentation du rapport SAM/non-SAM, parmi les composantes qui restent à classer. Des exemples de cartes spatiales de SAM bien sélectionnées sont montrées en annexe A.1. Cinq SAM n'ont cependant pas été sélectionnées par la méthode. Une ré-examen de ces CI par les experts a révélé que deux d'entre elles étaient finalement de non-intérêt, en accord avec la classification de notre méthode. Les trois autres restent cependant des détections manquées. Leur profil sera discuté dans la section suivante.

TABLEAU 7.1 – Résultats par base.

Bases	CI	Experts	Notre méthode						
		<i>VP</i>	<i>VP</i>	<i>FP</i>	<i>VN</i>	<i>FN</i>	<i>Se</i>	<i>Sp</i>	<i>P+</i>
B1	4 776	271	268	214	4 291	3	98,9 %	95,2 %	55,6 %
B2 _{3s}	6 000	621	621	2 262	3 117	0	100 %	54,3 %	21,2 %
B3	1 500	60	60	71	1 369	0	100 %	93,8 %	40,3 %
B4	2 142	96	96	774	1 272	0	100 %	59,3 %	10,4 %
ΣB	14 418	1 048	1 045	3 321	10 049	3	99,7 %	73,3 %	23,3 %

Parmi les quatre bases étudiées, les meilleurs résultats sont observés pour les données issues des bases B1 et B3. Pour B1 à partir de 4 776 CI extraites par l'ACI, 519 ont été pré-sélectionnées parmi lesquelles 249 étaient effectivement des SAM. Cela peut se traduire par une moyenne par sujet de 22 CI retenues et 177 CI rejetées sur les 199 au départ. Cette base est néanmoins la seule dont des SAM

n'ont pas été retenues (3 faux négatifs). Appliquée sur les bases B2_{3s} et B4, la méthode est moins performante en terme de spécificité mais permet quand même une importante réduction du nombre de composantes restant à classer pour l'expert, tout en s'assurant d'une sensibilité maximale. La sensibilité est maintenue à 100 % et la spécificité varie pour B2_{3s} et B4 entre 54,3 % et 59,3 %.

L'origine du signal semble être le principal facteur de ces variations. Nous obtenons de bien meilleurs résultats sur les données issues de séquences « BOLD pur » que sur les données issues de séquences ASL-BOLD. En fonction du type du signal, 48 % à 90 % des composantes sont rejetés. Parmi les CI de non-intérêt sélectionnées dans les bases B2_{3s} et B4, nous remarquons que certaines remplissent les critères de sélection proposés par Roquet et coll. [82]. Ces derniers ont été définis sur la base de données « BOLD pur » or quelques artefacts sont spécifiques des signaux provenant d'une séquence ASL-BOLD (un seul gros agglomérat plutôt qu'un réseau de plus petits agglomérats).

Le cas de B4 est particulier pour plusieurs raisons qui peuvent expliquer ces résultats plus faibles. Tout d'abord, les données issues de cette base n'ont pas été incluses pour l'estimation des seuils de décisions lors de l'apprentissage mais uniquement utilisées pour la phase de test. Cependant, bien que la spécificité soit moindre que pour les autres bases de données, la sensibilité est bien de 100 % et permet de valider notre méthode sur de nouvelles données n'ayant pas nécessités d'étape d'apprentissage préalable. En plus des artefacts décrits précédemment et qui sont spécifiques à la séquence ASL-BOLD, le type même des sujets a probablement joué un rôle dans ces résultats. B4 était en effet composée de patients âgés (entre 52 et 84 ans avec un âge moyen de 67 ans) atteints de troubles cognitifs légers ou de démence légère. La réponse hémodynamique des sujets âgés est connue pour être d'amplitude plus faible que pour les autres tranches d'âges [103]. Les artefacts de mouvements peuvent être aussi plus importants [104]. Enfin, le nombre de composantes extraites, pratiquement égal au nombre de volumes d'IRMf, était faible (120 volumes, 119 CI) par rapport aux autres bases de données. Par conséquent, la réduction de l'ordre du modèle effectuée par la PCA, qui permet également un premier filtrage, a eu peu d'effet et a laissé la majorité du bruit présent initialement dans les données. L'association de ces conditions a mené vers des CI où les caractères d'intérêt (aussi bien dans les cartes spatiales que dans les décours temporels) étaient de faible intensité et fréquemment accompagnés de bruit.

7.2 Analyse des résultats par descripteur

7.2.1 Performance des descripteurs

Les résultats de la sélection ont aussi été examinés selon chaque descripteur (les spécificités sont affichées TABLEAU 7.2). Chacune des cinq caractéristiques permet une sensibilité égale à 100 % pour B2_{3s}, B3 et B4 et de 99,8 % pour B1. On remarque néanmoins que le descripteur *focusing degree* se détache en terme de spécificité. La *fréquence cumulée* permet, elle aussi, une bonne classification mais principalement pour les bases B1 et B3 (signaux « BOLD pur »). Les valeurs prédictives positives (non présentées) restent comparables pour tous les descripteurs, principalement à cause du faible nombre de SAM parmi toutes les composantes des bases de données initiales.

Les CI de bruits inconnus, c'est-à-dire présentant un aspect bruité caractérisé par des voxels seuillés disséminés sur l'ensemble du cerveau, sont bien plus nombreuses que les CI d'artefacts connus qui peuvent, elles, présenter des agglomérats de voxels semblables à ceux des SAM. Le *focusing degree* qui mesure le niveau d'agglomération des voxels au-dessus du seuil dans la carte spatiale de la composante est capable de les repérer et fournit donc logiquement une meilleure spécificité. Cependant, c'est par ce même descripteur que les trois faux négatifs (SAM non sélectionnées) ont été produits. Ces SAM (leur carte spatiale est montrée en annexe A.2) ont pour particularité de présenter un réseau comprenant peu de voxels. Des voxels activés corrélés isolés sont eux assez nombreux sur les composantes et ont entraîné une mauvaise classification. *A posteriori*, nous avons pu remarquer que la majorité des voxels impliqués dans ces trois SAM mal classées

TABLEAU 7.2 – Spécificité par descripteur.

descripteurs	B1 (%)	B2 _{3s} (%)	B3 (%)	B4 (%)	ΣB (%)
<i>focusing degree</i>	83,1	36,2	79,5	49,7	53,7
<i>alternance</i>	15,2	1,0	8,2	0,3	6,5
<i>volume</i>	20,7	4,6	10,1	2,8	9,4
<i>localisation</i>	0,9	0,7	5,4	1,7	1,5
<i>fréquence cumulée</i>	54,8	2,2	60,6	5,7	18,6

Les colonnes présentent les résultats pour chacune des bases de l'étude et les lignes, les résultats pour chacun des descripteurs. La dernière colonne, intitulée ΣB, correspond aux résultats de la base de test complète, c'est-à-dire comprenant toutes les données destinées au test issues des sous-bases B1, B2_{3s}, B3 et B4.

faisaient partie d'autres SAM sélectionnées. C'est d'ailleurs ce qui explique que ces SAM soient si bruitées, et relativise le problème posé par la non sélection de ces SAM au niveau du sujet, en vue d'une application clinique. Il est intéressant de noter que le deuxième descripteur le plus performant, la *fréquence cumulée*, l'est plus particulièrement pour les bases B1 et B3 (signaux « BOLD pur »). Pour rappel, les bases B2_{3s} et B4 avaient subi un pré-traitement supplémentaire (filtrage passe-bas) afin d'isoler le signal BOLD du signal ASL.

De manière surprenante, la *localisation* ne fournit pas de bons résultats en terme de classification. Pourtant, comme nous l'avons mentionné plus haut, la majorité des CI de non-intérêt présente un aspect bruité de voxels dispersés sur l'ensemble du cerveau (matière blanche, LCR, crâne compris) qui pourrait laisser croire que ce descripteur mènerait à de meilleurs résultats. En réalité, les images n'ayant subi que peu de pré-traitements spatiaux (pas de normalisation, ni de lissage) et l'analyse se faisant sur sujet unique avec beaucoup de composantes, même les SAM extraites peuvent parfois présenter de nombreux voxels isolés assimilables à du bruit. Cela conduit d'une part à des détections manquées par le *focusing degree* mais aussi un seuil relativement bas pour la *localisation*. Cela est visible également sur les bases d'apprentissage en annexe B. Ce descripteur permet néanmoins d'écarter les CI spécifiques aux artefacts de mouvements des yeux, de la tête ou de la pulsation du LCR dans les ventricules (voir les illustrations au chapitre 5 FIGURE 5.4). Le *volume* et l'*alternance*, bien qu'ils soient peu discriminants, restent essentiels pour augmenter la spécificité. Ces descripteurs sont en réalité moins axés sur la sélection des SAM que sur le rejet de certaines composantes très particulières (instabilité du signal sur une coupe ou aspect d'alternance sur signal entre les coupes paires et impaires). Celles-ci sont peu nombreuses par rapport à l'ensemble des composantes d'un sujet mais sont aisément identifiables par ces descripteurs. Ces derniers étant facilement implémentables, leur intégration à la méthode n'en est que bénéfique.

Nous retrouvons, comme dans la section précédente, de meilleurs résultats sur les bases « BOLD pur » pour lesquelles les critères manuels étaient destinés initialement. Au vu des résultats, nous pourrions éventuellement nous passer de quelques descripteurs pour certaines bases (l'*alternance* pour B4, la *localisation* pour B1 et B2_{3s}). Cependant tous les descripteurs amènent à une amélioration pour l'une ou l'autre des 4 bases testées, il est donc important de garder ces descripteurs dans l'optique d'une application la plus générique possible voire en routine clinique.

La règle de fusion utilisée, le ET logique, pour combiner les décisions binaires en sortie des cinq descripteurs peut être considérée comme particulièrement simpliste. Cependant elle a été retenue pour correspondre au mieux à celle demandée par la sélection manuelle conçue par des experts du domaine [82]. Les critères de sélection manuelle que nous avons cherchés à automatiser étaient fondés sur un raisonnement biologique qui excluait des règles de combinaison plus sophistiquées.

Les critères testaient essentiellement des caractéristiques indépendantes. Si dans la pratique, lors de la sélection manuelle, les experts favorisent la carte spatiale au décours temporel, aucune pondération ne devait être attribuée lors de l'automatisation. Ce choix est renforcé par nos résultats. Par exemple, le *focusing degree* pourrait être celui sur lequel une plus forte crédibilité serait apportée à ses décisions. Il est cependant celui qui amène les trois seuls faux négatifs. En choisissant d'accorder le même poids à chaque descripteur, les critères tels que le décours temporels ou la localisation peuvent être intégrés à la décision bien plus facilement que par estimation visuelle. Cela ne peut rendre que plus robuste la méthode par rapport à une sélection manuelle.

7.2.2 Estimation des seuils

Les seuils obtenus par l'apprentissage et appliqués sur chaque valeur descriptive de la base de test sont donnés dans le TABLEAU 7.3.

TABLEAU 7.3 – Seuil par descripteur.

descripteur	valeur du seuil
<i>focusing degree</i>	0,7822
<i>alternance</i>	0,0159
<i>volume</i>	0,1603
<i>localisation</i>	0,3195
<i>fréquence cumulée</i>	0,4914

Les distributions des valeurs descriptives de la base d'apprentissage et de test sont illustrées FIGURE 7.1 pour le descripteur *focusing degree* et *volume*. Pour rappel, la base d'apprentissage comprenait 7 888 CI et celle de test, 14 418. Les distributions pour chaque descripteur et chaque classe (SAM, non-SAM) des bases d'apprentissage et de test ont été comparées en utilisant un test de Kolmogorov-Smirnov. L'hypothèse nulle (les deux distributions testées sont semblables) n'a jamais été rejetée (seuil de signification de 5 %). Les distributions « apprentissage » et « test » sont donc comparables. Sur nos jeux de données « tests », les seuils choisis ont donc du sens.

Parmi les différents descripteurs, nous remarquons deux types de comportement de distributions. Le premier où les deux classes se distinguent, c'est le cas pour le *focusing degree* et la *fréquence cumulée*. Ce comportement est illustré par le *focusing degree* sur la FIGURE du haut 7.1; toutes les distributions pour l'apprentissage étant disponibles en annexe B. Le deuxième cas concerne l'*alternance*, le *volume* (voir FIGURE 7.1, en bas) et la *localisation*. Pour ces descripteurs, les distributions des deux classes se chevauchent largement laissant toutefois un écart suffisant au rejet des composantes artefactuelles. Cela rejoint les résultats décrits précédemment concernant les écarts de spécificité selon les descripteurs.

7.3 Analyses comparatives

7.3.1 Résultats

7.3.1.1 ACI de groupe

Le point de comparaison avec l'ACI de groupe pour l'identification des SAM nous permettait plus de renforcer l'utilité d'une approche sujet unique que de véritablement proposer une comparaison. Appliquée sur les bases « BOLD pur », les experts ont identifié respectivement 29 et 13 RSN pour les bases B1 et B3 après une ACI de groupe avec 50 et 20 composantes. Bien que cela diminue grandement le temps de sélection (une seule sélection par base), la sensibilité est faible.

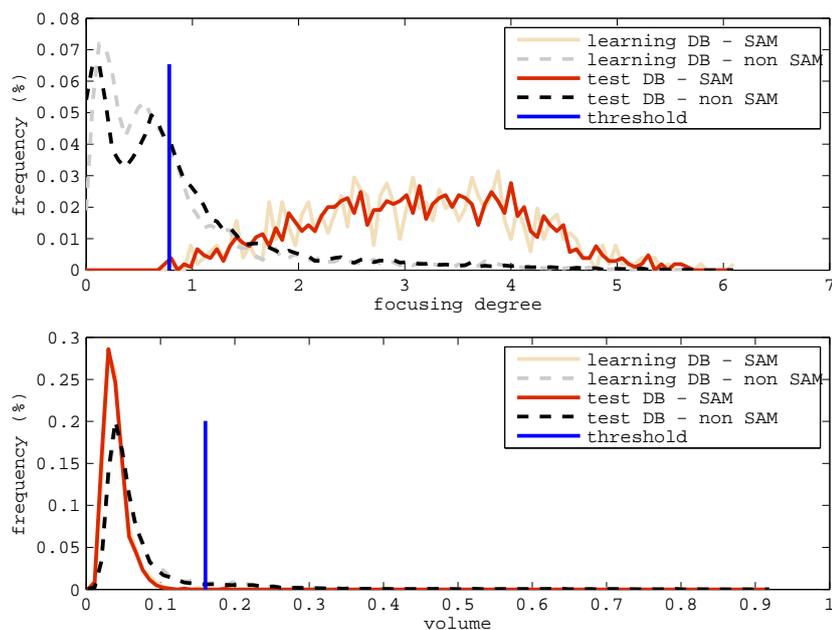


FIGURE 7.1 – Distributions des valeurs descriptives *focusing degree* (en haut) et *volume* (en bas). La base d'apprentissage est représentée par les couleurs claires (gris et beige) et celle de test par les couleurs foncées (noir et rouge). Les lignes continues sont les distributions des SAM, celles en pointillés représentent les distributions des composantes de non-intérêt (artefacts et bruits). Le recouvrement des distributions des bases pour la figure du bas ne permet pas la distinction entre les deux. La ligne bleue représente le seuil appris sur la base d'apprentissage (beige)

La reconstruction arrière permet de revenir dans l'espace du sujet, rendant possible le calcul de la correspondance entre les composantes identifiées et celles de notre base de données obtenues par l'ACI en sujet unique. Cette similitude a été mesurée par l'analyse de la qualité de l'ajustement (plus connue sous l'appellation anglaise *goodness of fit*) utilisée par Greicius dans un cas similaire [12]. Parmi toutes les SAM de notre base sujet unique, seules 77,1 % pour B1 et 75,3 % pour B3 ont été retrouvées par l'ACI de groupe.

7.3.1.2 Algorithme des k-moyennes

De manière prévisible, puisque la contrainte d'une sensibilité de 100 % n'était pas indiquée dans cette procédure, les résultats obtenus grâce à l'algorithme des k-moyennes suivi du ET logique montrent une forte spécificité et valeur prédictive positive. Cela se fait au détriment de la sensibilité, le nombre de SAM manquantes augmentant de manière inappropriée pour notre objectif. Avec plus de 190 SAM non sélectionnées comme telles, 18 % des SAM manquent sur le nombre de SAM totales. Les résultats sont présentés dans le TABLEAU 7.4.

7.3.1.3 Analyse discriminante factorielle

Les projections sur l'axe factoriel (entre -1 et 1) sont les suivantes :

Nous retrouvons ainsi dans l'ordre du plus au moins discriminant : le *focusing degree*, la *fréquence cumulée*, la *localisation*, le *volume* et l'*alternance*. Les signes sont négatifs pour les deux derniers car les valeurs descriptives des composantes d'intérêt sont plus petites que celles de non-intérêt contrairement aux autres descripteurs. Ces observations concordent avec celles faites sur les distributions des bases de données dans la section précédente (section 6.6). Les différentes performances de ces descripteurs sont dans la lignée de ce qui avait été observé paragraphe 7.2.1. Les

TABLEAU 7.4 – Résultats par base avec l’algorithme des k-moyennes.

Bases	CI	méthode avec k-moyennes							
		Experts		<i>VP</i>	<i>FP</i>	<i>VN</i>	<i>FN</i>	<i>Se</i>	<i>Sp</i>
B1	4 776	271	224	23	4 482	47	82,7 %	99,4 %	90,7 %
B2 _{3s}	6 000	621	452	80	5 299	169	72,8 %	98,5 %	85,0 %
B3	1 500	60	58	10	1 430	2	96,7 %	99,3 %	85,3 %
B4	2 142	96	73	72	1 974	23	76,0 %	96,5 %	50,3 %
ΣB	14 418	1 048	807	13 185	185	241	82,0 %	98,5 %	77,8 %

<i>focus. deg.</i>	<i>fréq. cum.</i>	<i>localisation</i>	<i>volume</i>	<i>alternance</i>
0,850	0,715	0,505	-0,184	-0,174

coefficients de l’équation quadratique de classification permettent d’obtenir les résultats présentés dans le TABLEAU 7.5. Encore une fois, la sensibilité n’étant pas maximale (35 faux négatifs), la spécificité augmente pour toutes les bases. Forcer l’algorithme à une sensibilité de 100 % (ou au moins à 99,99 %) mène à une spécificité quasi nulle. Ces résultats ne sont donc pas présentés.

TABLEAU 7.5 – Résultats par base avec l’analyse factorielle discriminante.

Bases	CI	AFD							
		Experts		<i>VP</i>	<i>FP</i>	<i>VN</i>	<i>FN</i>	<i>Se (%)</i>	<i>Sp (%)</i>
B1	4 776	271	244	153	4 352	27	90,0 %	96,6 %	61,5 %
B2 _{3s}	6 000	621	620	2 029	3 350	1	99,8 %	62,3 %	23,4 %
B3	1 500	60	54	30	1 410	6	90,0 %	97,9 %	64,3 %
B4	2 142	96	95	1 074	972	1	99,0 %	47,5 %	8,1 %
ΣB	14 418	1 048	1 013	3 286	10 084	35	96,7 %	75,4 %	23,6 %

7.3.1.4 Algorithme des forêts aléatoires

Avec l’algorithme des forêts aléatoires, il a été possible de jouer sur les seuils de décision (*a posteriori*) pour obtenir les mêmes sensibilités et spécificités que celles de notre méthode. La comparaison est donc directe. Les résultats présentés dans le TABLEAU 7.6 montrent que notre méthode est légèrement meilleure pour B2_{3s} et B4 (respectivement 21,5 % et 11,0 % de valeur prédictive positive avec notre méthode contre 17,9 % et 10,4 %), nettement plus efficace sur B1 (55,6 % contre 15,3 %) mais moins performante pour B3 (45,8 % contre 55,0 %).

7.3.2 Discussion

Comme déjà mentionné, la différence majeure entre les méthodes publiées et la nôtre réside dans le fait qu’elles ne recherchent pas une sensibilité de 100 %, même dans le cadre d’études sur sujet unique. Nos comparaisons se sont donc portées sur la modification de certaines étapes de notre chaîne de traitement plutôt que sur l’implémentation complète d’une méthode publiée. Si l’on cherche une sensibilité de 100 % et ce quelle que soit l’alternative choisie (ACI de groupe,

TABLEAU 7.6 – Résultats par base avec l’algorithme des forêts aléatoires.

Bases	CI	forêts aléatoires							
		Experts	<i>VP</i>	<i>VP</i>	<i>FP</i>	<i>VN</i>	<i>FN</i>	<i>Se (%)</i>	<i>Sp (%)</i>
B1	4 776	271	268	1 486	3 019	3	98,9 %	67,0 %	15,3 %
B2 _{3s}	6 000	621	621	2 846	2 533	0	100 %	47,1 %	17,9 %
B3	1 500	60	60	49	1 391	0	100 %	96,6 %	55,0 %
B4	2 142	96	96	828	1 218	0	100 %	59,5 %	10,4 %
ΣB	14 418	1 048	1 045	7 746	5 624	3	99,7 %	41,6 %	11,8 %

k-moyennes et ET logique, AFD, forêts aléatoires), les résultats sont meilleurs avec la méthode que nous proposons. L’identification des SAM avec la méthode de référence pour sélectionner les composantes indépendantes d’intérêt, à savoir l’ACI de groupe, n’a pas réussi à identifier environ 25 % des SAM identifiées lors d’une analyse sujet unique. Cela est d’autant plus gênant que les SAM identifiées avec l’ACI de groupe sont des RSN et non des SAM idiosyncratiques, certaines de ces dernières ayant le potentiel pour être spécifiques à un symptôme (tels que les réseaux d’épilepsie, des hallucinations ou des pensées obsessionnelles). En effet, non seulement on ne s’attend pas à ce que les sujets sains les présentent mais de plus la différence entre les patients et les sujets sains peut résider dans la présence ou non de SAM idiosyncratiques, par définition non recherchées dans une étude de groupe. Cette comparaison illustre l’utilité d’étudier la connectivité fonctionnelle à un niveau individuel que cela soit à des fins cliniques comme à des fins exploratoires. Des études rétrospectives menées sur un groupe de sujets (patients) à un niveau sujet unique peuvent permettre de révéler de nouvelles informations qu’une étude de groupe n’aurait pu mettre en avant (SAM idiosyncratiques présentes uniquement chez certains sujets d’un groupe de patients par exemple). Les études utilisant l’algorithme des k-moyennes pour la classification composantes d’intérêt vs. composantes de non-intérêt [90, 94] utilisent des composantes lissées et/ou normalisées et avec un ordre du modèle plus petit (entre 54 et 100). Les nombres de composantes idiosyncratiques et de composantes de bruit diminuent avec ces pré-traitements. Quoi qu’il en soit, ces méthodes ne recherchent pas de sensibilité maximale. Les deux autres méthodes utilisées dans la comparaison, l’AFD et les forêts aléatoires, restent moins performantes que notre méthode. Il est intéressant de remarquer que la base B4 obtient les résultats les plus faibles pour chacun de ces algorithmes. Les mêmes remarques que pour notre méthode peuvent être faites ici pour expliquer la difficulté de cette base.

7.4 Interface

Pour améliorer et accélérer encore plus le processus de sélection finale, une interface fonctionnant sous Matlab a été développée. Après avoir appliqué la méthode de pré-sélection, l’interface présente le nombre de CI restant à sélectionner par session et par sujet. La carte spatiale ainsi que le décours temporel sont affichés comme cela est déjà le cas sous FMRLAB (interface avec laquelle l’ACI est faite). Un aperçu de l’interface est visible en annexe C.1.

L’ajout principal vient dans le classement des composantes restantes et l’affichage des valeurs descriptives. Pour chaque composante, nous calculons un indice global descriptif qui correspond à la somme non pondérée des valeurs descriptives normalisées. Les composantes sont triées dans l’ordre croissant de cet indice. Plus l’indice est fort, plus il est probable que ce soit une SAM. Les composantes en bas du classement sont souvent des composantes de non-intérêt. Pour chaque composante, il est possible de voir également rapidement les valeurs descriptives. Utiles particu-

lièrement pour la *fréquence cumulée* et la *localisation*, l'utilisateur pourra s'appuyer sur ces informations pour s'aider dans sa décision (ces deux descripteurs décrivent des pourcentages et sont donc facilement interprétables).

7.5 Estimation du gain de temps

Bien que notre approche de pré-sélection automatique n'ait pas écarté le rôle de l'expert humain, il a fourni un important facteur de gain de temps en réduisant le nombre de CI restant à inspecter à 30 % du nombre de données d'origine. Pour un expert qualifié, nous avons estimé à 20 minutes le temps nécessaire pour sélectionner les composants d'intérêt par la méthode de sélection manuelle, pour un sujet et une session de 250 CI. Avec une sensibilité très proche du maximum, notre méthode offre un gain de temps de deux à dix fois par rapport à une sélection manuelle. Cependant, on ne s'attend pas à ce que le gain de temps corresponde parfaitement à la proportion de CI rejetées. D'une part, les CI restantes peuvent être plus difficiles à classer et donc nécessiter plus de temps d'inspection. De plus, afin d'optimiser la sélection finale par l'expert, les CI sont classées selon un indice global qui, en cas de difficulté, peut aider et accélérer la décision des experts en fin de processus. Enfin et surtout nous pouvons nous attendre à moins d'erreurs humaines grâce à une sélection manuelle courte (voir section 7.1 page 71, les 2 erreurs de classification par les experts lors de l'établissement du *gold standard* par la sélection manuelle). Un nombre réduit de composantes à classer limite le nombre d'erreurs manuelles qui peuvent être dues à la fatigue, particulièrement lorsque le nombre de sujets est important.

7.6 Applicabilité

Étant donné que les seuils ont été obtenus manuellement sur une base d'apprentissage, notre méthode est performante et donc applicable uniquement sur des données ayant été soumises aux mêmes pré-traitements que la base d'apprentissage (voir section 6.2). Pour les images EPI-BOLD ou ASL-BOLD, avec une taille de voxel isotrope entre 3 mm et 4 mm et avec les mêmes pré-traitements, les seuils que nous avons déterminés peuvent être réappliqués tels quels (voir TABLEAU 7.3). Si les résultats se révèlent insatisfaisants, une nouvelle estimation de ces seuils peut être faite sur une base de données d'apprentissage plus proche de l'environnement étudié.

Si l'on souhaite rechercher les RSN communs à un groupe avec une sensibilité maximale, une utilisation de cette méthode avec une ACI de groupe pourrait être envisagée. Les pré-traitements étant différents dans ce cas, un nouvel apprentissage doit être effectué. Il permettrait d'être plus sélectif que l'ACI en sujet unique et améliorerait d'autant plus la pré-sélection. Si l'application vise exclusivement des données issues de la séquence ASL-BOLD, pour lesquelles un filtrage supplémentaire est effectué au début des pré-traitements, il est également possible d'adapter le seuil du descripteur *fréquence cumulée*. Le filtrage modifiant le profil fréquentiel des décours temporels des composantes, un seuil plus ciblé pour ces données peut être appliqué. À court terme, il est également possible d'améliorer la sélection en s'éloignant un peu plus des critères manuels et en ajoutant de nouveaux descripteurs. Par exemple, le rejet de certaines composantes particulières peut se faire via des critères spatiaux additionnels comme des masques adaptés aux types d'artefacts engendrés par la séquence ASL-BOLD.

Actuellement, notre méthode ayant été développée et testée sur une base de données variées, elle est applicable en l'état en suivant les recommandations mentionnées ci-dessus et adaptable aux spécificités de certaines séquences pour des résultats encore meilleurs.

Conclusion partielle

Nous avons présenté dans cette partie une méthode de pré-sélection de SAM obtenues à partir de l'IRMf. La méthode proposée, basée sur l'ACI spatiale, automatise des critères manuels définis dans l'article de Roquet et coll. sous la contrainte d'une sensibilité de 100 % comme exigé pour des études sur sujet unique. L'approche, entièrement guidée par les données, extrait cinq caractéristiques par composantes indépendantes pour les classer comme SAM ou non-SAM en fonction de seuils estimés sur une base de données d'apprentissage. La procédure de pré-sélection conduit à une forte réduction du nombre de CI et fournit la sensibilité requise tout en économisant un temps considérable pour les scientifiques et/ou les médecins. En outre, la méthode, évaluée sur quatre bases d'images avec des conditions d'acquisition, des séquences, des résolutions et des types de sujets différents, a permis de mettre en évidence la reproductibilité de notre méthode. Ainsi la méthode proposée pourrait être appliquée dans le cas d'études multi-centriques sans nécessiter un nouvel ensemble de données d'apprentissage. Avec un classement des composantes par ordre décroissant de pertinence, l'interface développée pour la validation finale des CI permet également un gain de temps et une prise en main plus aisée de la méthode. Pour clore cette seconde partie, la méthode automatique de pré-sélection SAM que nous venons de présenter devrait faciliter l'intégration de l'ACI sur sujet unique dans les études en neurosciences. Elle devrait aussi favoriser le développement d'études sur les réseaux idiosyncratiques.

Troisième partie

Analyse statistique de la dynamique de la connectivité fonctionnelle

Introduction	83
8 Dynamique de la connectivité fonctionnelle	85
9 Analyse de données : observations et résultats préliminaires	99
10 Modélisation et analyse de la DCF par MMC-Produit et MMC-Couplés	107
11 Applications cliniques : Médicament vs placebo, Démence à corps de Lewy	115
Conclusion partielle	135

Introduction

La notion de DCF (pour dynamique de la connectivité fonctionnelle) a été introduite récemment afin d'expliquer les variations de connectivité fonctionnelle observée sur un temps bien inférieur à la durée d'une session d'IRMf. Les SAM activées au cours d'une session voient leurs liens de connectivité varier dans le temps laissant suggérer une configuration fonctionnelle cérébrale changeante sur des fenêtres temporelles très courtes (de l'ordre de la seconde). Le travail que nous présentons dans cette partie s'inscrit dans la recherche de nouvelles approches pour décrire et analyser la DCF. L'extraction de caractéristiques spatio-temporelles de la CF doit pouvoir mener à une meilleure compréhension des propriétés dynamiques des réseaux de repos qui permettent le fonctionnement normal du cerveau. En outre, ces études peuvent également apporter des réponses sur la façon dont une pathologie ou des traitements médicamenteux peuvent altérer la DCF dans un but de devenir idéalement un indicateur diagnostique voire pronostique.

Plusieurs approches révélant la DCF en IRMf ont été développées afin de commencer à répondre aux interrogations concernant les origines neurobiologiques des variations de la connectivité fonctionnelle, le type de modélisation qui serait le plus pertinente selon l'étude ou encore concernant l'influence des bruits connus ou non sur les observations. Ces études seront présentées dans le premier chapitre de cette partie. À partir des bases de données à notre disposition (B2, B4 et B5, voir chapitre 4 page 37), des observations simples décrites dans le chapitre 9 nous ont confirmé l'existence de modifications de la connectivité fonctionnelle. Afin de proposer des outils plus ciblés vers la dynamique des interactions entre les SAM, nous avons souhaité élaborer de nouvelles méthodes d'analyse de la DCF. Nos méthodes, exposées dans le chapitre 10 nous ont permis d'établir ces changements dans un cadre statistique grâce à l'utilisation de modèles probabilistes : les modèles de Markov cachés multidimensionnels.

Dynamique de la connectivité fonctionnelle

8

8.1 Principes et Problématiques

8.1.1 Variabilité de la DCF

Bien que longtemps considérée comme statique pour simplifier les approches, la connectivité fonctionnelle (CF) renferme en réalité de nombreuses variabilités qui peuvent se manifester à différentes échelles. Il a été montré que la CF changeait sous certaines conditions ou contextes, par exemple : durant le sommeil [105], la sédation [106], selon les activités cognitives [107], l'apprentissage ou encore les troubles neurologiques [12]. Des variations de la CF corrélées à des variabilités au niveau des mesures individuelles (QI, personnalité, âge, sexe) ont également été observées [108, 109]. Il en est de même au niveau intra-sujet où des variations peuvent apparaître lorsque l'on compare différents examens espacés dans le temps. La question de changements rapides de la CF dans le temps se pose donc naturellement, d'autant que les activités électriques observées par l'EEG et le MEG ont mis en avant des interactions entre les régions pouvant rapidement se reconfigurer. En IRMf de repos, les changements à la fois de force et de direction des CF semblent être sur des échelles de temps plus petites (secondes/minutes) que celles étudiées actuellement (session d'IRMf de 5 à 30 minutes). En considérant maintenant non pas des conditions ou contextes différents (tâche, troubles neurologiques, âge etc..) mais une échelle d'analyse réduite, serions-nous en mesure de mettre en avant des changements de la connectivité fonctionnelle en IRMf de repos ? Plusieurs études récentes se sont penchées sur cette question et ont permis de démontrer l'existence de multiples variations.

Une réflexion intuitive nous amène à penser que les réseaux de CF, observés jusqu'à présent durant les études, seraient en réalité des moyennes de ces « mêmes » réseaux sur la durée de l'enregistrement. Kiviniemi et coll. [110] ont illustré cette hypothèse en suivant l'évolution du DMN (facilement identifiable chez de nombreux sujets même avec une faible résolution temporelle) en extrayant ce réseau sur différentes fenêtres temporelles tout au long de la séquence. Ils ont ainsi pu comparer les DMN obtenus par cette approche dynamique à celui obtenu à partir des données de la séquence prise dans son ensemble. Aucun DMN extrait avec l'approche dynamique n'avait exactement le même profil spatial que le DMN « statique ». Cette constatation soulève d'emblée un ensemble de questions concernant la définition même d'un réseau puisque l'on remarque qu'il dépend, en plus d'autres facteurs connus pour le cas statique, de l'échelle de temps sur laquelle il est défini.

Dans leur étude, Gonzalez-Castillo et coll. [111] ont eu l'avantage d'avoir pu travailler sur une longue séquence d'IRMf (54 minutes). En faisant varier la fenêtre temporelle d'observation par différentes subdivisions de leur signal (30+24 minutes ou 15+15+15+9 minutes), ils ont suivi l'évolution du profil spatial de certains réseaux. Ils ont ainsi pu remarquer différents types de changements

entre les régions impliquées dans un réseau selon que l'on regarde une fenêtre plus ou moins large. Une illustration de leurs observations a déjà été introduite en partie I de ce document (voir FIGURE 1.4 page 9, section 1.2.2). Pour autant, cette figure est sciemment introduite une nouvelle fois ci-dessous (FIGURE 8.1), tant elle nous paraît capitale pour percevoir pleinement cette notion de variation de la connectivité fonctionnelle.

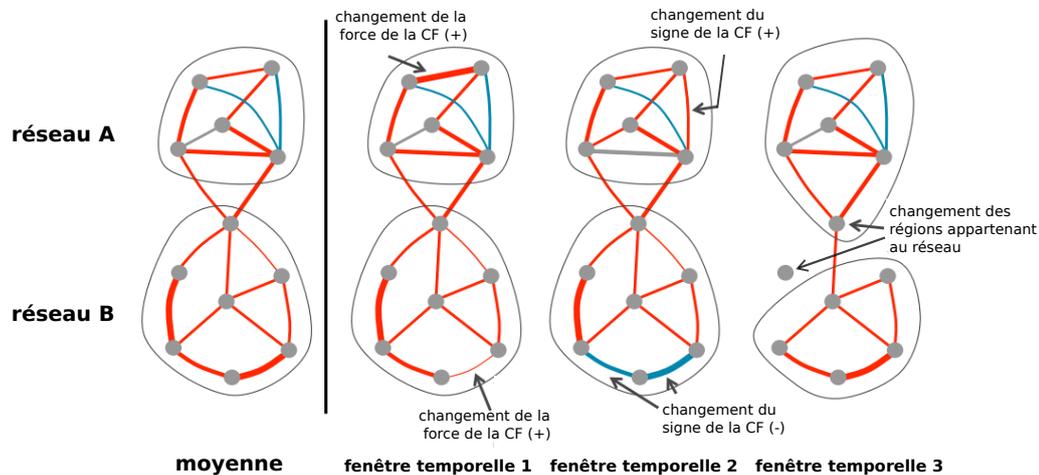


FIGURE 8.1 – Illustration de la dynamique de la connectivité fonctionnelle sur différentes fenêtres temporelles (colonnes 2, 3, 4) par rapport à celle obtenue sur l'ensemble de la session (colonne 1). La force de la connectivité fonctionnelle entre deux régions peut varier en amplitude (colonne 2), en signe (colonne 3) ou du fait de ces deux cas, une région peut être associée/dissociée d'un réseau suite à un gain/perte d'amplitude de la relation (colonne 4). En rouge : connexions positives ; en bleu : connexions négatives. (d'après [7] modifié).

Dans la suite de ce manuscrit, nous utiliserons donc le terme « dynamique » de la connectivité fonctionnelle (ou DCF) pour se référer spécifiquement aux variations de la CF pour un sujet au repos (éveillé). La CF « statique » se rapporte, elle, à la CF estimée sur l'ensemble des acquisitions d'une session (sous l'hypothèse de stationnarité temporelle des réseaux). En apportant des informations différentes, telles que le profil moyenné temporellement des réseaux, la CF statique est complémentaire à la DCF et son étude reste pertinente.

8.1.2 Problématiques et motivations pour l'étude de la DCF

Les deux études mentionnées précédemment ont véritablement ouvert la voie à la recherche sur la DCF, tout en soulevant dans le même temps de nombreuses questions liées par exemple aux bases neuronales de la DCF ou à la façon d'analyser cette DCF et d'en interpréter les résultats. Les questions les plus récurrentes dans la littérature ont été regroupées ci-dessous :

- Neurobiologie :
 - ▷ Quels sont les origines et mécanismes neuronaux de la DCF ?
 - ▷ Les variabilités de la CF sont-elles liées à une pathologie ? Pourrait-on utiliser les caractéristiques temporelles de la CF comme biomarqueur d'une maladie ?
 - ▷ Est-ce que la DCF est modulée par un exécutif central comme le cortex pré-frontal ou est-ce un processus auto-organisé ?
 - ▷ À quel degré la DCF représente-t-elle des processus conscients par rapport à des processus automatiques ou inconscients ?

- ▶ Acquisition :
 - ▷ Quelles sont les contributions à cette dynamique observée, des mouvements, des bruits physiologiques et artefacts de la machine ?
 - ▷ Quelle est la correspondance des variations de la CF avec le niveau d'éveil du sujet pendant la durée d'un examen au repos ?
- ▶ Modélisation :
 - ▷ Quelle est la définition la plus pertinente d'un réseau et de ses nœuds pour quantifier la DCF ?
 - ▷ Est-ce que la DCF suit des séquences temporelles déterministes durant la tâche ou le repos ?
 - ▷ Existe-t'il un jeu discret de configurations neuronales (un « répertoire fonctionnel ») qui est continuellement revisité ? Sont-elles reproductibles selon les sujets ou liées à un génotype sous-jacent ? De nouvelles configurations émergent-elles durant le développement ou à l'âge adulte ?

Ces interrogations montrent, s'il en était besoin, l'engouement suscité et les problématiques soulevées par l'analyse de la DCF dans la quête d'une meilleure compréhension du fonctionnement cérébral. Dans la suite de cette thèse, nous allons nous intéresser plus spécialement aux approches de modélisation de la DCF. Des revues complètes existent sur les bases neuronales de la DCF ; pour plus d'informations sur ces points, nous renvoyons le lecteur aux articles de Keilholz [112] et Tagliazucchi [113].

8.1.3 DCF et état de connectivité fonctionnelle

Un concept aussi récurrent que les questions abordées précédemment dans la communauté DCF est celui d'« état de connectivité fonctionnelle ». Même si tout le monde semble s'accorder sur la pertinence de ce concept, il n'existe pas encore aujourd'hui de consensus ou de définition arrêtée sur ce terme. Un état est souvent représenté par une carte spatiale d'aires cérébrales connectées fonctionnellement à un instant t ou sur une période temporelle très courte (quelques secondes) et qui est reproductible. La notion de reproductibilité peut varier selon l'étude. Ainsi on peut définir un état *a posteriori* selon les données présentes durant la session et chercher une reproductibilité inter-session et intra-sujet si l'étude est faite au niveau du groupe. On peut également définir les états *a priori*, en envisageant un large panel de possibilités et s'intéresser ainsi à des états reproductibles d'une session à l'autre, pour un même sujet par exemple. Quelque soit l'approche choisie, c'est la dynamique globale du cerveau qui peut être modélisée. Le terme de *chronnectome* a été proposé par Calhoun [114] pour définir la dynamique, non pas à l'intérieur d'une région, mais des couplages de connectivité entre deux (ou plus) régions ou réseaux neuronaux. La force de la connectivité entre plusieurs réseaux est donc exprimée en fonction du temps. Cela peut inclure les couplages temporels (la connectivité fonctionnelle telle qu'elle est décrite actuellement) mais aussi les couplages spatiaux voire spatio-temporels. Ainsi les propriétés spatiales de certains réseaux peuvent aussi être fonction du temps. L'étude du chronnectome a donc pour but d'identifier des modèles de couplage entre les régions du cerveau qui varient dans le temps mais qui sont reproductibles. C'est aussi dans ce contexte que s'inscrivent nos études.

Le concept d'état de connectivité fonctionnel est intimement lié à la conceptualisation de la DCF. Deux hypothèses sont à l'étude : d'une part on peut voir la DCF comme un espace d'états où plusieurs modèles discrets réapparaissent, ainsi à un instant t le cerveau se trouve dans un état, une configuration spatiale, parmi un nombre fini et celui-ci pouvant être répété dans le temps ; d'autre part la DCF peut être vue comme variant le long d'un espace continu d'états. À l'heure actuelle, c'est principalement la première solution qui est à la base des approches de modélisation de la DCF, car plus aisée à mettre en œuvre et à interpréter dans une première approche d'analyse. Parmi les méthodes déjà publiées, qui seront présentées ci-dessous, la plupart estime et caractérise les états de manière à n'autoriser au sujet qu'à être dans un seul état à un instant t parmi plusieurs

prédéfinis. Cependant, d'autres approches ont restreint le nombre d'états de connectivité tout en les autorisant à se chevaucher dans le temps. À ce jour, le recul ne semble pas être suffisant pour juger de la meilleure approche. Les études menées en DCF commençant cette année à ré-utiliser et appliquer les méthodes développées (que l'on présentera dans les paragraphes suivants) sur des bases de données plus larges, on peut s'attendre à des investigations sur la robustesse des méthodes et leur interprétabilité ou encore sur leur résistance aux bruits. Des travaux seront nécessaires pour voir quelle métrique est la plus informative pour une population et une hypothèse données.

8.2 État de l'art des méthodes d'analyse de la DCF

Malgré le caractère très récent de ce nouveau champ de recherche, il existe déjà nombre de méthodes pour quantifier et qualifier la DCF et nous allons nous y consacrer dans ce chapitre. Leurs objectifs sont variables selon qu'elles cherchent à différencier deux populations, à quantifier la DCF, à l'illustrer (c'est-à-dire en donner une représentation graphique) ou encore à interpréter les observations générées par les méthodes. Elles sont recensées ici selon le type d'outils qui est à la base de la mise en évidence de la dynamique. En outre, comme nous l'avons déjà évoqué, les méthodes diffèrent aussi selon qu'elles captent l'évolution d'un même réseau et ses régions qui le composent ou l'évolution des interactions entre les réseaux. Si nous avons choisi, dans notre cas, de nous intéresser aux réseaux cérébraux, les méthodes ciblant les régions seront aussi présentées ; d'autant que le pont entre les deux est aisément franchissable rendant l'application de certaines méthodes adaptable au deux cas de figures. Les approches bayésiennes seront davantage détaillées eu égard aux approches markoviennes cachées que nous proposons dans ce manuscrit pour analyser la DCF.

8.2.1 Fenêtre glissante

L'étude de la DCF est récente et il n'existe pas encore de stratégie d'analyse qui se soit imposée. Certaines sont néanmoins partagées par plusieurs publications : c'est le cas de l'analyse par fenêtre glissante qui semble être à la base de nombreuses études [110, 111, 115, 116, 117, 118, 119, 120, 121, 122]. Cette méthode est la façon la plus populaire et intuitive d'aborder la DCF. L'enregistrement est divisé en fenêtres d'une durée fixe (les limites de la fenêtre peuvent être atténuées) et on calcule la CF des données qui sont dans cette fenêtre. Classiquement la CF dans une fenêtre est calculée par corrélation entre les décours temporels de deux régions/réseaux, donnant ainsi une matrice de corrélation entre tout(e)s les régions/réseaux étudié(e)s. On déplace ensuite la fenêtre d'un intervalle temporel allant d'un pas à la taille de la fenêtre : cet intervalle définit le chevauchement entre les fenêtres.

La taille de la fenêtre doit être un compromis entre un nombre suffisant de points pour assurer un calcul robuste et une longueur réduite pour une analyse de la dynamique. Il ne faut pas perdre de vue les fréquences d'intérêt dans le signal BOLD qui sont, pour des études au repos, comprises entre 0,004 Hz et 0,08 Hz (voir page 52 Partie II, paragraphe 5.3), soit des périodes entre 12,5 et 250 secondes. Les intervalles entre chaque volume acquis étant classiquement de l'ordre de quelques secondes en IRMf (TR de 2 à 3 secondes dans la majorité des études), les tailles des fenêtres sont aisément comparables d'une étude à l'autre.

Hutchison et coll. [115] ont, par exemple, utilisé cette technique lors d'une étude comparative entre les macaques anesthésiés et les humains éveillés. Cette étude a permis de mettre en avant des périodes de forte corrélation entre toutes les régions d'un réseau alternant avec des périodes de faible corrélation ainsi que l'inclusion passagère de nouvelles aires dans le réseau. Ce fut le cas pour les deux macaques anesthésiés et les humains éveillés. Cette observation permet d'inférer que les fluctuations observées avec cette méthode ne peuvent pas être générées uniquement par des processus conscients tels que les déplacements de l'attention, le traitement sensoriel, le souvenir

ou encore la planification. Il était important de vérifier que la variabilité des interactions entre les zones cérébrales représente bien un phénomène intrinsèque et spontané, indépendant des processus conscients. En effet, une des limitations dans la caractérisation de la dynamique de l'IRMF de repos (déjà mise en avant par Chang [120]) est l'incapacité d'assurer une absence de processus conscients lors de l'acquisition des images. Sans contrainte, les sujets humains éveillés se livrent à un large éventail d'activités mentales qui peuvent modifier l'organisation fonctionnelle du cerveau (vagabondage mental, surveillance de l'environnement, changements liés à la vigilance, l'attention et l'excitation, formation de la mémoire). Mais ce qui a le plus attiré notre attention dans cette étude est l'utilisation de cinq tailles de fenêtres différentes (15 s - 30 s - 60 s - 120 s - 240 s). Si les observations sont cohérentes et globalement similaires pour les quatre fenêtres les plus petites, la décorrélation ponctuelle de certaines aires des réseaux étudiés n'était pas possible avec les fenêtres les plus grandes (120, 240). Shen et coll. [117] ont poursuivi l'étude précédente (mêmes données et méthode d'analyse) et ont montré le lien entre connexion anatomique et variabilité de la CF : les régions avec la CF la plus stable dans le temps semblaient être celles avec des connexions anatomiques bidirectionnelles.

8.2.2 Modèle reproductible avec approche par fenêtre glissante

La corrélation par fenêtre glissante n'est qu'une première étape dans l'étude de la DCF. Pour une modélisation plus approfondie, cette étape est généralement suivie d'une approche analytique pour déterminer d'une part les états de connectivité et d'autre part leur évolution. Cette section présente ces différentes approches.

Graphes : Plutôt utilisée dans les études sur la dynamique en EEG (électroencéphalographie) ou MEG (magnétoencéphalographie), la représentation graphique est particulièrement intéressante dans le cadre de la DCF pour les métriques qu'elle propose. Le coefficient d'agrégation, l'efficacité globale et la force des connectivités permettent à Yu et col [123] de souligner les performances moindres du cerveau de patients atteints de schizophrénie. Une autre étude utilise, elle, la mesure de modularité des représentations graphiques des états de connectivité pour observer les différences entre sujets sains et patients atteints d'Alzheimer [124]. Ces états de connectivité fonctionnelle représentés par des graphes sont obtenus à partir des corrélations calculées par fenêtres glissantes.

Algorithme des k-moyennes : Allen et coll. [119] ont réalisé leur étude sur un examen de durée certes assez courte (5 minutes) mais dont le nombre de patients inclus (405 volontaires sains) a permis une analyse robuste et autorisé une interprétation « fondée » de la dynamique observée. À partir des décours temporels, issus des RSN obtenus par une ACI de groupe, ils réalisent une corrélation à fenêtre glissante pour extraire des matrices de covariances en chaque point temporel. Cette première étape leur a permis de partitionner ces matrices en deux groupes en fonction des réseaux qui montraient plus de variabilité de CF que d'autres. Ils ont ainsi déterminé des zones de stabilité (CF la moins variable) et d'instabilité (CF la plus variable) sur l'ensemble du cerveau. Après l'évaluation de cette dynamique, cette étude a identifié sept états de connectivité au repos par la recherche d'agrégats via l'algorithme des k-moyennes. Les auteurs ont ainsi montré que ces états différaient fortement des modèles de CF observés dans les analyses statiques. Ces états peuvent être représentatifs de tendances moyennes desquelles et vers lesquelles les fluctuations de la CF se font dans le temps. L'étude a été poussée un peu plus loin en modélisant la CF par une chaîne de Markov afin de caractériser le comportement des transitions entre états. Les résultats sont présentés au niveau du sujet et leurs mesures, telles que le temps passé dans chaque état, l'occurrence de chaque état en fonction du temps ou les transitions les plus probables, font émerger de possibles interprétations. Ainsi, un état possiblement représentatif de la somnolence est plus fréquemment sollicité en fin d'examen et il est plus probable d'y rester une fois atteint. Cependant, la durée de la séquence (5 minutes) paraît particulièrement courte pour établir avec certitude une

telle conclusion. L'application de cette méthode sur une durée plus longue permettrait de rendre plus sûres les interprétations, les états ayant plus de possibilité de se répéter (ou non). Cette étude étant issue de l'équipe de Calhoun, « The Mind Research Network », qui regroupent de nombreux chercheurs dédiées à la CF, plusieurs publications ont déjà suivi et utilisé la méthode détaillée pour des applications cliniques (voir plus bas, section 8.3, [118, 125]).

Analyse en composantes principales (ACP) : À l'opposé de l'approche par agrégation employée précédemment par Allen pour identifier des états de connectivité fonctionnelle, Leonardi et coll. [122] ont utilisé l'ACP pour déterminer les profils spatiaux qui peuvent au mieux modéliser la CF à travers le temps et les sujets. C'est une nouvelle manière d'appréhender la DCF puisque, plutôt que de la considérer comme une suite d'états discrets, elle est modélisée ici comme différents profils de connectivité qui se chevauchent dans le temps. Ils considèrent ainsi que la DCF peut être représentée, à n'importe quel instant t , comme une somme pondérée de plusieurs profils de connectivité. La DCF est construite avec la méthode de la corrélation par fenêtre glissante suivie d'une ACP globale sur tous les sujets de l'étude permettant de dégager dix composantes principales. Le décours temporel associé à chacun de ces dix profils spatiaux peut être reconstruit au niveau de chaque sujet. Pour chaque sujet, on a ainsi accès à la contribution de chaque profil de connectivité à travers le temps. Leur méthode est employée par la suite pour évaluer l'altération de la DCF chez des patients atteints de sclérose en plaques. Ils remarquent que deux des dix composantes étudiées contribuent à la variation de la DCF de manière significativement différente entre les sujets sains et patients et peuvent ainsi déduire quelles connexions semblent affectées chez les patients.

ACI spatiale : Plutôt que de s'intéresser aux variations des interactions entre des réseaux dont la définition spatiale est fixe, la DCF peut aussi se concevoir comme l'étude des variations à l'intérieur d'un réseau et donc de la configuration variable d'un réseau au cours d'une session d'IRMf de repos. Kiviniemi et coll. [110] se sont ainsi intéressés au DMN (voir Partie I section 1.6), un des réseaux les plus partagés par la population et facilement identifiable. Si l'ACI spatiale est fréquemment utilisée pour faire émerger des réseaux étudiées en DCF, elle est habituellement appliquée sur l'ensemble de la session donnant ainsi accès à la configuration spatiale « moyenne » de ces réseaux (c'est le cas de cinq études présentées dans ce chapitre [118, 119, 123, 124, 126]). Dans l'étude de Kiviniemi et coll., l'ACI est appliquée par fenêtre glissante sur les volumes présents dans chaque fenêtre. Ainsi, il n'y a plus d'hypothèse sur les voxels faisant partie du réseau, ici le DMN. Bien qu'ils étudient une fenêtre de plus de 100 secondes pour avoir un nombre de volumes suffisamment important pour le calcul de l'ACI, ils ont mis en évidence la nature variable du profil spatial du DMN. Celui-ci diffère sensiblement selon les fenêtres observées sans jamais complètement correspondre au gabarit que l'on obtient par l'ACI appliquée sur toute la session. Ces conclusions vont dans le sens d'une hypothèse déjà émise qui stipule que la CF « statique » nous offre une image moyenne des profils de CF qui ont émergé pendant la session. Si cette approche a l'avantage d'être guidée par les données, elle nécessite un nombre de données important dans chaque fenêtre pour assurer la robustesse de l'analyse. Il serait donc intéressant d'appliquer cette technique sur des données obtenues avec un TR très court et ainsi descendre à des fenêtres temporelles plus fines. En outre, Kiviniemi et coll. utilisent une très faible dimension pour leur ACI (15 composantes) ce qui peut être acceptable au vu du réseau qu'ils recherchent mais paraît trop limitée pour travailler sur le cerveau entier *a fortiori* au niveau du sujet.

8.2.3 Analyse en vecteurs indépendants (AVI) spatiale

Une alternative à l'ACI est l'AVI qui est dédiée aux études de groupe car elle fournit des composantes (appelés vecteurs) indépendantes au sein d'un même jeu de données mais qui peuvent avoir des dépendances avec les composantes d'autres différents jeux de données (d'autres sujets). Ma et col [121] ont eu l'idée d'utiliser cette approche pour étudier la DCF en l'appliquant à des signaux

fenêtrés. Ainsi, à partir d'une session de 5 minutes, ils ont divisé chaque signal acquis en sept fenêtres. L'AVI est appliquée sur un ensemble de signaux regroupant les sept fenêtrages de chaque sujet de l'étude (sujet sains et patients). Cela permet de retrouver les vecteurs indépendants, communs à chaque sujet et fenêtre. La reconstruction de ces vecteurs au niveau sujet permet plusieurs analyses statistiques : d'une part ils mettent en évidence certaines paires de composantes qui ont un comportement différent d'un groupe à l'autre, d'autre part, une modélisation Markovienne est réalisée pour rendre compte au mieux des transitions entre les états (trois états de connectivité sont définis à partir des composantes observées). Cette approche est intéressante pour l'extraction des réseaux (ou vecteurs indépendants) qu'elle propose. En effet, ils sont calculés non plus sur l'ensemble de la session mais à partir de signaux fenêtrés ce qui permet d'éviter l'effet moyenné de l'ACI de groupe et de mettre en avant de plus grandes variations entre les sujets et fenêtres. Néanmoins, le faible nombre d'états de connectivité étudiés et de fenêtres extraites rend l'approche insuffisamment détaillée pour une étude approfondie de la DCF.

8.2.4 ACI temporelle

Dans la première partie de ce manuscrit (section 2.2.1), nous avons également introduit l'ACI temporelle. Celle-ci est peu utilisée en IRMf en étude de groupe puisqu'elle fait l'hypothèse de profils temporels similaires à travers les sujets. De plus, elle est basée sur moins de données que l'ACI spatiale car elle est limitée par le nombre de volumes acquis (donc par la durée de l'examen et le TR utilisé).

Pourtant, une étude a utilisé cette analyse pour identifier les réseaux sur la base de leur indépendance temporelle. Smith et coll. [126] sont partis du principe que les réseaux peuvent se chevaucher spatialement soit parce que certaines régions comprennent plusieurs unités fonctionnelles qui ne peuvent être séparées par la résolution spatiale limitée que l'on a à disposition, soit parce que certaines régions appartiennent réellement à plusieurs réseaux fonctionnels distincts. Ainsi, sur une durée suffisante de la session IRMf, des réseaux distincts fonctionnellement devraient être aussi distincts temporellement même s'ils se chevauchent spatialement. Si l'ACI temporelle est plutôt utilisée pour l'IRMf de tâche à cause de l'hypothèse d'indépendance temporelle entre les réseaux, elle semble ici plus adaptée pour autoriser les chevauchements spatiaux. Consciente des limites dues à l'utilisation de l'ACI temporelle (moins robuste car moins de données) cette équipe a utilisé une séquence d'imagerie accélérée ($TR = 0,8s$) pour acquérir un plus grand nombre de volumes. Une ACI spatiale avec un ordre élevé (200 composantes) permet d'écarter les artefacts et l'ACI temporelle est réalisée ensuite sur les décours temporels des composantes restantes. Vingt et un états en résultent, chacun associé à un vecteur qui décrit la force d'implication de chaque parcelle (les réseaux fonctionnels issus de l'ACI spatiale) dans chaque état fonctionnel temporel. Par cette approche, Smith et coll. décomposent les décours temporels de l'activité cérébrale et non de la connectivité comme c'est le cas pour les autres travaux cités jusqu'alors. Bien que cette analyse permette de souligner des états indépendants temporellement et donc évite le questionnement sur la taille de la fenêtre d'observation, elle ne fournit pas d'information sur le moment de chaque état. La stationnarité temporelle de chaque état implicitement supposée par l'utilisation de l'ACI temporelle rend difficile l'interprétation et n'est pas directement explicite pour des études sur la dynamique temporelle [127]. Leur approche a néanmoins permis, grâce une étude complémentaire, d'estimer que 25% de la non-stationnarité apparente dans la corrélation des données brutes est en fait attribuable au fait que plusieurs réseaux fonctionnels se chevauchent dans le temps avec des amplitude variables (avec le temps). Cela donne du poids à l'approche par ACP, présentée par Leonardi, qui autorise de tels chevauchements.

8.2.5 Analyse temps-fréquence

Comme cela est récurrent dans l'analyse de la DCF, un compromis est recherché concernant l'échelle temporelle d'observation de la CF : d'une part assez longue pour récupérer les basses fréquences du signal BOLD et estimer la CF avec un rapport signal sur bruit suffisant, d'autre part assez courte pour être sensible aux changements transitoires de la connectivité au sein ou entre les réseaux. Si une fenêtre glissante impose une taille fixe, les analyses temps-fréquence avec des outils comme la transformée par ondelette semblent particulièrement adaptées pour une telle problématique.

Chang et Glover [120] sont les premiers à avoir proposé une approche pour visualiser la DCF basée sur l'analyse par cohérence en ondelettes. Cette analyse consiste en une mesure de l'inter-variation entre deux signaux ou entre deux représentations de ces signaux suivant différentes échelles (fréquences) au cours du temps. Les valeurs de cohérence permettent de trouver des régions dans l'espace temps-fréquence où les deux signaux covarient mais n'ont pas nécessairement une haute puissance commune. Leur étude est particulièrement complète car ils ont cherché à vérifier statistiquement leurs résultats (cohérence et amplitude de la cohérence), et à analyser l'influence du mouvement du sujet sur les variabilités observées et l'impact de l'imageur utilisé (sujets acquis sur différents IRM) afin d'éviter toute interprétation hâtive. Chang et Glover ont confirmé, via cette étude, que les variations dans la connectivité fonctionnelle de repos pouvaient se produire à des échelles très courtes et que les anti-corrélations apparaissent être plus des phénomènes transitoires que stables. Cette étude n'est pas axée vers la recherche d'états de connectivité, il s'agissait de mettre en évidence la variabilité de la connectivité fonctionnelle. Des mesures de la variabilité globale de la cohérence à des bandes de fréquences sélectionnées ou à partir de séries suffisamment longues, et de la vitesse à laquelle la cohérence et la phase sont modulées, permettent en outre de quantifier les fluctuations de DCF observées par cette approche. L'analyse ayant été faite dans cette étude sur les interactions (par paire) entre 5 ROI, appliquer cette approche au cerveau complet représenterait un défi important concernant l'analyse des données ainsi que la visualisation.

8.2.6 Modèles de Markov

Les réseaux bayésiens dynamiques (RBD) ont déjà été utilisés pour la connectivité effective (effet d'un réseau sur un autre) pour mettre en évidence les relations non synchrones par exemple. En IRMf, néanmoins, ils sont encore peu mis à profit alors même que les notions d'états, de transitions ou encore d'inter-relations utilisés en DCF font écho au champ lexical des RBD. Parmi eux, les modèles de Markov ont commencé à être appliqués pour la DCF principalement en fin de chaîne de traitements, en supplément d'autres analyses statistiques. Les méthodes développées par Allen [119] et Ma [121], déjà détaillées dans les paragraphes précédents, ont su tirer avantage des modèles de Markov pour extraire des informations sur les transitions entre les états ou les durées d'occupation des états. Il existe néanmoins quelques publications, principalement des actes de conférences [128, 129], où les modèles de Markov sont utilisés comme outil principal de l'analyse de la DCF.

Eavani et coll. [128] ont voulu passer outre l'étape de corrélation afin de s'affranchir des questions inhérentes au choix de la fenêtre. Ils ont développé une méthode basée sur les modèles de Markov cachés (MMC) associés à une méthode de *Sparse Dictionary Learning* (apprentissage de dictionnaire pour les représentations parcimonieuses). Chaque état est caractérisé par une matrice de covariance unique construite sur un ensemble de régions d'intérêt prenant en compte tout le cerveau. Ils font l'hypothèse qu'un nombre relativement petit de processus sous-jacents (et donc de régions associées) entraînent des variations dans les signaux d'IRMf. Cela se caractérise par des changements subtils dans les matrices de covariances à partir desquelles des états cachés peuvent être identifiés. Afin de modéliser cette hypothèse, ils ont cherché à approximer au mieux les matrices représentant les états cachés (matrice de covariance) par une combinaison linéaire non négative.

tive des matrices de bases (méthodes de *Dictionary Learning*). Cela se fait dans un processus itératif qui permet de trouver, à partir des données, aussi bien les matrices de bases (pour l'apprentissage de dictionnaire) que les états cachés. Le signe des éléments des matrices de covariance dans les états permet de décider si les ROI sont corrélées ou anti-corrélées. Leur méthode permet de décoder la DCF en séquences temporelles d'états cachés. Leur algorithme a été optimisé pour un nombre d'états cachés égal au nombre de matrices de base. Leurs résultats sont présentés pour trois états cachés (matrices de covariances distinctes). Leur proposition permet de modéliser directement la DCF sans nécessiter une étape de corrélation fenêtrée avec un faible nombre d'états cachés, ceux-ci étant appris sur les données.

Une seconde étude, par Ou et col [129], a fait le choix des MMC pour décrire au mieux les phénomènes fluctuants d'engagement ou de désengagement dans certains états fonctionnels chez les enfants atteints de TDAH (trouble de déficit de l'attention avec hyperactivité). Des matrices de corrélation sont calculées sur une fenêtre temporelle de 38 secondes à l'aide de corrélation à fenêtre glissante comme cela a déjà été décrit dans les paragraphes précédents. La modélisation qui suit est faite en deux temps. Une première étape permet la génération de séquences d'observations symboliques à partir de ces matrices de corrélation de toutes les fenêtres et de tous les sujets. La deuxième étape consiste en la modélisation par MMC des processus des sujets. Un modèle est créé pour les sujets contrôles et un autre pour les enfants atteints de TDAH. Le nombre de symboles est déterminé expérimentalement à 15. Le nombre d'états est également défini empiriquement pour les sujets contrôles et pour les patients respectivement à 24 et 25. En plus d'apporter de nouvelles connaissances sur cette pathologie cognitive, cette étude a réalisé une classification entre les sujets (sains et patients) en comparant la vraisemblance d'une observation selon le modèle choisi (MMC des sujets sains ou MMC des sujets TDAH). Si cette étude est encore sujette à plusieurs paramétrages empiriques, et apprise sur un nombre trop restreint de sujets pour pouvoir être étendue telle quelle comme méthode de classification, elle laisse suggérer que la modélisation par MMC de la DCF est un bon outil pour apporter de nouvelles connaissances sur les variations temporelles du fonctionnement cérébral.

8.2.7 Bilan

Alors que la fenêtre glissante reste l'outil privilégié pour évaluer la DCF, son utilisation soulève quelques questions dont, en premier lieu, le choix de la taille de la fenêtre. Assez large pour une estimation robuste et capturer les basses fréquences mais suffisamment petite pour déceler les effets transitoires, la taille de la fenêtre est souvent décidée de manière empirique, en lien toutefois avec les fréquences du signal BOLD qui sont connues. Beaucoup des études mentionnées ont rapporté avoir utilisé, dans des travaux initiaux, différentes tailles de fenêtre (entre 15 secondes et 4 minutes) mais leurs résultats sont généralement donnés pour une échelle entre 30 et 60 secondes. Dans les faits, peu d'impacts sur la dynamique sont observés en fonction de ce choix au-delà du résultat attendu que les plus grandes fenêtres mettent moins de variabilité en évidence. Étant donné ces similarités entre les états de connectivité extraits à plusieurs échelles temporelles, il semble peu probable que la DCF soit uniquement déterminée par des processus de bruits aléatoires. Cet intervalle ([30 s ; 60 s]) est donc souvent présenté pour réaliser le compromis déjà soulevé ; il est de plus appuyé par une étude de Shirer et coll. en 2012 [130] qui cherche à identifier des états cognitifs pré-appris par l'algorithme de reconnaissance utilisé. Cette étude arrive à identifier des états cognitifs à partir de matrice de covariance entre régions cérébrales dès 30 secondes de données. En plus des alternatives proposées par les méthodes de Chang et Glover (temps-fréquence), Leonardi (ACP) ou encore Smith (ACI temporelle et acquisition accélérée), Cribben et coll. [131] proposent d'aborder la question non par la recherche d'états, mais par la recherche de points de changements de CF. De cette manière, la délimitation des fenêtres via ces points de changements, donc leur taille, peut varier le long du signal étudié.

Les problématiques liées à la fenêtre glissante sont fortement associées à la métrique utilisée

en aval pour analyser la DCF sur la fenêtre choisie. On peut se poser la question du choix de la corrélation par fenêtre glissante comme outil pour mettre en avant la DCF. Parmi les premières études qui se sont intéressées à la DCF, certaines considéraient la simple présence de fluctuations dans ce signal comme une évidence de la présence de DCF. Si cette corrélation permet de quantifier dans une certaine mesure cette dynamique, la connectivité n'est pas directement observable par la corrélation car non exempte de bruit ou variations connexes. Certaines études qui ont suivi ont donc prouvé l'existence de dynamique dans leurs données en comparant leurs signaux d'origine à ces mêmes signaux dont la dynamique de la corrélation entre les réseaux a été modifiée mais la connectivité statique conservée. Plusieurs algorithmes sont proposés lors des études de DCF pour réaliser ce test [120,122,132] et ils ont à chaque fois permis de montrer que les fluctuations observées n'étaient pas dues uniquement au bruit ou à la variabilité inter-sujet (la plupart des études se font au niveau du groupe).

De manière plus générale, concernant toutes les métriques de la CF, la part du bruit est une question récurrente lors de l'analyse de la DCF puisque la mesure est faite sur un faible nombre de points temporels. Les mouvements du sujet, notamment physiologiques, doivent être identifiés pour éviter de les prendre en compte lors de la mesure de la DCF. Puisque, par nature, ils ont un caractère transitoire, leurs effets sur la CF sont atténués quand on utilise une grande fenêtre temporelle d'analyse (comme dans l'analyse « statique » par exemple). L'impact peut être considérable pour des études de la dynamique, un mouvement de faible amplitude ou une respiration courte introduiront des fluctuations dans le signal BOLD qui peuvent se manifester de la même manière que les changements de connectivité. Il est possible par exemple d'enregistrer simultanément les signaux cardiaques et respiratoires. Mais ces données sont rarement disponibles et ne peuvent donc pas être utilisées à des fins de débruitage. Les études intègrent donc de nombreuses étapes de filtrage, particulièrement lors d'études de groupe. La plupart des pré-traitements utilisés en CF statique peuvent l'être aussi en DCF (filtrage spatial, régression des paramètres de ré-alignement). Certains, par contre, nécessitent une attention particulière. Par exemple, les valeurs anormalement élevées ou faibles sur le décours temporel sont, dans le cas des études de CF statique, retirées ou remplacées par interpolation. Cette étape doit être faite de manière précise en DCF puisqu'elle aura une incidence sur l'analyse dynamique en raison de l'altération de la structure temporelle des données. Concernant le filtrage temporel, certaines études choisissent d'appliquer un filtrage passe-haut supplémentaire avant la fenêtre glissante afin d'observer uniquement les changements de la CF dont les fréquences ont des périodes plus petites que la taille de la fenêtre. Allen, dans son étude de la DCF [119], obtient des états de connectivité presque similaires avec un pré-traitement agressif (remplacement des valeurs aberrantes par une estimation, filtrage passe-bas, filtrage de la composante tendancielle linéaire, quadratique et cubique, régression multiples des paramètres de ré-alignement) et un pré-traitement plus standard (filtrage passe-bas, filtrage de la composante tendancielle linéaire). Ces étapes de débruitage ne semblent donc pas modifier fondamentalement les propriétés dynamiques des signaux.

La limite générale à toutes ces études ne concerne non pas la méthode mais les applications souvent réalisées sur des enregistrements relativement courts (entre 5 et 10 minutes). Le temps total de l'acquisition paraît bien court pour arrêter des profils de connectivité et étudier la variabilité au niveau sujet lors d'étude de groupe. La recherche de séquences reproductibles d'états de CF doit être nuancée au regard du temps observé.

Si nous avons souhaité rester dans le domaine de la connectivité fonctionnelle au repos, la DCF a aussi déjà commencé à être étudiée sur des séquences d'IRMf incluant des tâches cognitives ou motrices. Nous pouvons citer par exemple Basset et coll. [133] qui se sont intéressés à des tâches motrices simples ou encore l'équipe de Kinnison [134] qui a cherché à caractériser les états engendrés par des menaces ou des encouragements. Ces deux études ont privilégié les graphes pour définir au mieux leurs réseaux et utiliser ainsi les nombreux outils proposés par ce type d'approche.

Il est clair que les états de DCF contiennent des informations très intéressantes et potentielle-

ment utiles dans la caractérisation des populations. Pourtant, ce champ de recherche encore jeune, requiert de nouvelles techniques pour valider les résultats antérieurs mais aussi pour proposer de nouveaux outils de mesure, de modélisation et de représentation pour une meilleure interprétation des observations.

8.3 Intérêt de l'analyse de la DCF pour les applications cliniques

Au-delà des méthodes d'analyse de la DCF, il nous paraît important d'apporter dans ce manuscrit un éclairage (purement) clinique sur la pertinence de l'analyse de la DCF. Pour commencer, il est à noter que des altérations de la connectivité fonctionnelles ont été soulignées dans la plupart des maladies psychiatriques. Cause ou conséquence, le lien entre les changements de la DCF et les dysfonctionnements constatés chez ces populations cliniques sont un enjeu majeur des recherches actuelles dans ce domaine. L'extraction de caractéristiques spatio-temporelles de la CF doit pouvoir mener à une meilleure compréhension de la maladie, à des traitements médicamenteux plus ciblés et, idéalement, être un indicateur diagnostique voire pronostique. En outre, ces études apportent aussi des réponses sur la façon dont les propriétés dynamiques des réseaux de repos permettent le fonctionnement normal du cerveau. Déjà, de multiples applications ont pu montrer l'utilité d'aborder la CF via sa dynamique et la question se pose de l'utilisation de la DCF comme potentiel biomarqueur de maladies mentales. Ces études ont déjà été présentées en filigrane des descriptions méthodologiques dans les paragraphes précédents mais nous les reprenons ici succinctement afin de faire le point sur les maladies déjà étudiées sous l'angle de la DCF.

La schizophrénie et les troubles bipolaires semblent être les pathologies où l'utilisation de la DCF est la plus avancée, le point de vue statique ayant été largement étudié [135]. Différentes approches ont été testées : Sakoglu par une analyse temps-fréquence [136], Damaraju [118] et Rashid [125] en utilisant la méthode d'agrégation développée par leur équipe [119] et enfin Ma [121] via différents outils statistiques. Bien qu'ils aient tous réussi à différencier les sujets sains des patients, leurs résultats quant aux types de changements observés varient. Ainsi, suivant la manière de définir les états, certains ont pu mettre en avant des différences spatiales et temporelles des profils de certains RSN [136] alors que d'autres obtiennent des états de connectivité similaires [118, 125]. Ces derniers ont cependant observés des changements d'états plus fréquents chez les sujets sains. Ils suggèrent ainsi que les sujets sains seraient probablement plus rapides dans l'engagement de ressources nécessaires face à l'évolution des exigences dues à une tâche. Bien que ces résultats aient été corroborés par d'autres études de cette même équipe, les résultats de Ma et coll. [121] montrent l'inverse, à savoir plus de fluctuations des profils spatiaux chez les patients schizophrènes. Des études restent donc encore à faire, autant pour valider les méthodes de quantification de la DCF que pour distinguer les variations entre les réseaux de celles entre les aires.

La maladie d'Alzheimer a été également largement étudiée d'un point de vue statique, où des altérations de la CF ont été rapportées pour ces patients par rapport aux sujets sains [30]. L'étude de Jones [124], déjà mentionnée dans le paragraphe ci-dessus concernant les graphes, a mesuré des différences dans les temps d'occupation des états entre sujets sains et patients. Plus précisément, les états de connectivité comportant la région du DMN antérieur étaient bien plus souvent atteints que les états comportant les régions du DMN postérieur pour les patients, cela n'était pas le cas chez les sujets sains. Ces résultats concordent avec des observations faites en CF statique mais l'apport de la dynamique nous permet de visualiser plus précisément les différentes configurations spatiales occupées par les patients plutôt qu'un état moyen.

Des altérations de la CF ont déjà été soulignées chez des patients atteints de dépression majeure. Récemment c'est via la connectivité effective que des changements sont apparus chez ces sujets par rapport aux sujets sains [137]. La dynamique (causale) des aires cérébrales du réseau par défaut (DMN) paraît fortement altérée et il serait intéressant d'aborder ces données par la DCF pour explorer les interactions entre ces aires à différentes échelles temporelles.

Ces observations n'auraient pu être faites en utilisant une approche statique et soulignent l'importance d'évaluer la DCF voire même de combiner les deux approches pour une vision plus complète et robuste. D'une part, l'approche statique fournit une vue d'ensemble des réseaux apparus ou non durant une session ce qui permet déjà une distinction entre sujets sains et patients. D'autre part, l'approche dynamique permet de rentrer plus en détails pour observer des variations plus fines telles que les transitions entre les états de connectivité ou encore les modifications de l'implication de certains réseaux dans la dynamique globale. Ces observations qui peuvent être « diluées » si l'on ne dépasse pas l'analyse statique sont pourtant porteuses d'informations permettant de distinguer sujets sains et patients. Ces résultats pourraient permettre la caractérisation de la DCF d'un symptôme ou d'une pathologie. S'ils sont avérés, cela peut être déterminant pour la mise en place d'une méthodologie basée sur la DCF, appliquée à l'échelle du sujet, capable de différencier les symptômes de l'information propre au patient. Améliorer le diagnostic à un moment précoce de la maladie reste un objectif majeur des études en cours. Il convient de noter qu'une limitation importante pour les résultats présentés ici réside dans le fait que les patients sont sous traitement médicamenteux ; un travail supplémentaire est nécessaire pour découpler les effets des médicaments à partir des véritables effets de la maladie. À mesure que ces méthodes vont devenir parties intégrantes des études de CF, la DCF va pouvoir élargir son champ d'application. Déjà, des états bien particuliers tels que le sommeil profond [138, 139], la rêverie [140] et les états psychédéliques [141] ont commencé à être observés sous l'éclairage de la DCF.

8.4 Notre approche dans ce contexte

À travers ces lectures bibliographiques, les notions d'organisation temporelle et spatiale des réseaux cérébraux spontanés (SAM) et plus largement l'organisation des processus cérébraux d'un point de vue fonctionnel, apparaissent comme les inspirations et hypothèses partagées des études de DCF. Le nombre de régions cérébrales ou de SAM à observer et donc le nombre d'interactions possibles sont très nombreuses dans les études de DCF. En outre, la liberté d'évolution de l'activité mentale d'un sujet durant l'état non contraint de repos rend la recherche de structure dans la DCF délicate. Cette remarque doit être particulièrement prise en considération étant donné que les études IRMf fournissent un nombre limité d'observations, la robustesse devenant plus difficile à garantir.

Pour résoudre ce problème, les procédés mentionnés ci-dessus reposent sur une étape de réduction des données. Les états de connectivité sont définis par agrégation en fonction de ce qui a été observé au cours d'une première étape d'analyse DCF. Cette stratégie conduit à un petit nombre d'états et permet de se concentrer exclusivement sur les données observées. Les états définis sur un groupe de sujets, représentent des configurations spatiales de la connectivité entre les RSN qui apparaissent de manière reproductible dans le temps et/ou entre les sujets. La dynamique entre ces états est ensuite modélisée par les différentes méthodes que nous avons décrites dans ce chapitre. Cette procédure suppose de prendre suffisamment en compte, dans la réduction et dans la modélisation de la dynamique, les interactions spatiales et temporelles entre RSN. Cependant, cette approche, bien plus proche des données, ne permet pas d'observer les états qui sont plus éphémères ou moins représentatifs de la DCF d'un groupe de sujets. De plus, pour une étude au niveau du sujet, l'étude des SAM idiosyncratiques (particulières au sujet) dans la DCF pourrait être importante pour permettre de mieux appréhender leur rôle dans le fonctionnement normal ou non du cerveau au repos. Une autre approche consiste donc à définir les états de connectivité, directement à partir des états possibles pour chaque SAM, sans la connaissance des interactions qui apparaîtront dans la séquence observée. De cette façon, la complexité de ces interactions entre les SAM est transférée dans un modèle multidimensionnel comme nous souhaitons le proposer.

En parallèle de cette définition des états de connectivité, il est nécessaire de définir le cadre de notre travail puisque des choix sont à faire au préalable. Tout d'abord, concernant la profondeur

de l'analyse, nous avons choisi d'examiner la dynamique à l'échelle des réseaux et non des régions. Même s'il est vrai que les connections entre les aires formant un réseau peuvent varier dans le temps, la connectivité fonctionnelle statique permet déjà de fournir de l'information quant à leurs connections pour former le réseau. En revanche, les interactions entre les réseaux cérébraux, qui plus est, sur l'ensemble du cerveau, sont, elles, inconnues ou peu suffisamment décrites par la CF statique. Par ailleurs, compte tenu des bases de données à disposition (voir chapitre 4), plusieurs buts peuvent être envisagés. Notre méthode doit pouvoir permettre d'identifier et de caractériser, à l'échelle de la seconde, les mécanismes conversationnels entre réseaux cérébraux spontanés chez un sujet particulier. En admettant que l'activité fonctionnelle cérébrale n'est pas un processus totalement libre, sans structure temporelle, il est donc envisageable de rechercher des configurations d'inter-dépendances entre les SAM et leurs enchaînements dans le temps. Dans une visée de différenciation, la méthode doit pouvoir proposer des outils comparables afin de confronter la DCF évaluée sur deux sessions distinctes. Une extension de la méthode pour différencier deux populations (sujets sains vs. patients) sur la base de leur DCF peut être également considérée.

Notre contribution à ce champ de recherche est donc motivée par le besoin d'apporter de nouveaux outils qui permettraient de mieux caractériser le comportement de chaque SAM ou RSN par rapport aux autres. Les méthodes disponibles dans la littérature permettent jusqu'à présent soit de caractériser la DCF de manière globale sur le cerveau soit de cibler en amont un réseau pour étudier ses changements dans le temps quitte à ne pas prendre en compte le reste de réseaux. Les modèles de Markov cachés multidimensionnels nous paraissent particulièrement pertinents 1) pour saisir les dépendances temporelles entre les réseaux cérébraux, 2) sans nécessiter une réduction des données en amont, 3) tout en offrant des outils de mesure et différenciation entre modèles.

En ce sens, nous avons développé deux déclinaisons d'une même approche basées sur les modèles de Markov cachés produit (voir section 3.2.2) puis sur les modèles de Markov cachés couplés (voir section 3.2.3). Ces déclinaisons seront présentées dans le chapitre 10. Avant cela, une primo-analyse des données brutes est présentée au chapitre suivant, comme nous avons pu la conduire en pratique, afin de mieux circonscrire et appréhender les données, et avant même d'exposer les approches markoviennes cachées, développées dans le chapitre 10.

Analyse de données : observations et résultats préliminaires

9

Avant de développer toute nouvelle méthode d'analyse de la DCF, nous avons souhaité « prendre en main » les données à disposition. L'objectif est ici de tenter d'en dégager d'éventuels comportements singuliers, des tendances, des informations permettant potentiellement de mieux orienter nos futurs développements. Pour cette première étape d'analyse des données, nous disposons de trois bases, l'une pour une application orientée sujet et composée de sujets sains (B2), les deux autres pour une application orientée groupe et composées de patients et de sujets sains (respectivement B4 et B5). Les informations concernant ces bases de données sont disponibles au chapitre 4, section 4.2.2 page 40. Les analyses se sont concentrées uniquement sur la première base (B2), ceci afin d'éviter l'effet moyenneur des études de groupe. Il s'agissait en effet de rester à l'échelle du sujet et d'observer l'interaction des différentes SAM, la dynamique spatiale de chaque SAM n'étant pas l'objectif premier. L'étude est faite sur les réseaux et non sur les régions qui les composent ; les réseaux reflétant déjà des aires cérébrales connectées entre elles. Évidemment, ces réseaux ont été définis d'un point de vue statique, à partir de la fenêtre complète de la séquence IRMf. Ce choix est nécessaire, dans un premier temps, pour réaliser, avec plus de recul, une étude s'intéressant à la variabilité de la connectivité entre les réseaux qui se sont manifestés pendant la séquence.

Pour rappel la base B2, utilisée pour cette étude préliminaire, regroupe douze volontaires sains ayant réalisé chacun quatre sessions d'IRMf (Ses1, Ses2, Ses3, Ses4) espacées d'une semaine. Au cours de chaque session, une séquence EPI-ASL de 20 minutes (TR = 3 s) au repos a été acquise, soit 404 volumes. Lors de la dernière session, Ses4, huit sujets ont pris, une heure avant l'imagerie, un médicament, stimulant du système nerveux (30 mg de méthylphénidate) et quatre sujets ont pris, à la place, un placebo. Dans la suite, les huit sujets sous médicament en session 4 seront dénommés par l'acronyme M_i (i de 1 à 8). Les quatre sujets sous placebo en session 4 seront notés quant à eux P_j (j de 1 à 4).

9.1 Décours temporels vs. signaux de corrélations

L'analyse en composantes indépendantes fournit, pour chaque SAM, le profil spatial du réseau (via la carte spatiale) et l'implication de ce réseau au cours de la session (via le décours temporel), (voir FIGURE 6.2). Ce dernier apporte donc déjà une information temporelle en lien avec la session étudiée. Néanmoins, nous n'avons ni accès aux évolutions de la connectivité interne à une SAM ni, en ce qui nous concerne, à la dynamique des interactions entre les SAM. Pour cela, deux possibilités sont envisageables : d'une part, modéliser la DCF directement à partir des décours temporels des SAM, d'autre part, générer des signaux de corrélation avec des fenêtres glissantes, entre des paires de SAM.

L'utilisation des décours temporels des SAM comme entrées de l'analyse de la DCF a l'avant-

tage de fournir des résultats facilement interprétables, du fait que l'on travaille immédiatement sur chaque SAM sans passer par des étapes d'appariement, comme c'est le cas avec la corrélation. Les modèles résultants fournissent une DCF globale de tous les réseaux pris ensemble. Par ailleurs, cette stratégie permet également de s'affranchir de l'application d'un fenêtrage qui engendre des questions délicates (limites soulevées en section 8.2.7, appuyées par l'étude d'Hindriks [132]). Cependant elle ne permet pas de se focaliser sur le comportement et les interactions d'un réseau en particulier par rapport aux autres.

La deuxième méthode, c'est à dire l'utilisation de signaux de corrélation, est facile à implémenter et donne directement une indication sur l'interaction entre deux SAM au cours du temps. Des informations sur l'évolution de la CF sont donc déjà accessibles à ce stade. C'est pourquoi elle est largement répandue, comme nous avons pu le voir en section 8.2.2. Dans les méthodes qui utilisent cet outil, les signaux de corrélation ne sont souvent qu'une première étape pour avoir accès à la DCF. Elle permet ensuite de définir des états de connectivité, une seconde étape étant réalisée pour trouver des modèles reproductibles. L'utilisation des signaux de corrélations entre décours temporels de SAM présente un autre avantage : ils permettent de concentrer l'analyse sur une SAM en particulier et de rechercher comment évolue ses relations avec les autres SAM. Cela peut-être intéressant dans le cadre de la caractérisation d'une pathologie par exemple, où des *a priori* peuvent guider l'étude ou lorsque le médecin souhaite se concentrer sur certaines SAM spécifiquement. Les signaux de corrélation obtenus par fenêtre glissante restent d'ailleurs l'outil privilégié pour mettre en avant la DCF. Par la suite, nous avons décidé de travailler sur ces signaux à la fois facilement mise en œuvre et interprétable.

9.2 Observations préliminaires

Les pré-traitements effectués sur les données ont déjà été détaillés dans le paragraphe 6.2. Pour rappel, il s'agissait d'un filtrage passe-bas (filtrage de la composante ASL), une correction du décalage temporel inter-coupes, un recalage rigide et une ACP-ACI. Quelques pré-traitements supplémentaires ont été effectués sur les décours temporels, avant corrélations, pour d'une part limiter les effets qui pourraient influencer la DCF et d'autre part pour pouvoir comparer les signaux issus de différentes sessions : suppression de la tendance, vérification des valeurs aberrantes sur le signal, normalisation. Ces types de pré-traitements sont standards dans la plupart des études de la DCF. Une méthode d'analyse de DCF ([119], présentée en 8.2.2) utilisant beaucoup de pré-traitements sur les décours temporels (suppression de la tendance, régression des 6 paramètres de ré-alignement et de leur dérivées, enlèvement des valeurs aberrantes avec remplacement par estimation, passe-bas avec $F_c = 0,15\text{Hz}$) a également été réalisée avec très peu de pré-traitements (suppression de la tendance et passe-bas) et a obtenu des résultats similaires. Il est donc intéressant de noter que les pré-traitements « agressifs » appliqués n'ont pas fondamentalement altéré la structure dynamique des signaux.

Afin de faire une étude comparative entre les sessions, nous avons décidé de ne retenir comme SAM parmi celles sélectionnées après l'ACI spatiale (voir Partie II), uniquement celles qui étaient présentes sur chacune des quatre sessions. Pour les sujets de la base B2 décrite précédemment, ce nombre varie entre 6 et 21 selon les sujets (comme indiqué dans le TABLEAU 9.1).

TABLEAU 9.1 – N : nombre de SAM communes aux 4 sessions par sujet, pour la base B2

Sujet	M1	M2	M3	M4	M5	M6	M7	M8	P1	P2	P3	P4
N	10	9	9	8	9	21	6	8	19	15	9	8

En accord avec ce qui est fait classiquement dans la littérature, nous avons défini la fenêtre glis-

sante, utilisée pour délimiter les décours temporels en vue de la corrélation, comme proposé par Allen et col [119]. La résolution temporelle (définie par TR, le temps de répétition) est de 3 secondes pour les données que nous avons à disposition. La fenêtre a été obtenue par la convolution entre un signal carré (15 TR, soit 45 secondes) et une gaussienne (écart-type de 2 TR, soit 6 secondes). La fenêtre finale présente ainsi des bords atténués et couvre 51 secondes. Elle est déplacée d'un TR, soit 3 secondes, à chaque pas. La corrélation par fenêtre glissante entre deux signaux temporels x et y est définie par

$$\rho_{x,y}(t) = Cor(x_t^{t+w-1}, y_t^{t+w-1}) \quad (9.1)$$

où w est la taille de la fenêtre, x_t^{t+w-1} et y_t^{t+w-1} représente la portion des signaux x et y entre les temps t et $t + w - 1$ et $Cor()$ le calcul du coefficient de corrélation de Pearson. Au total, pour chaque session initialement de 20 minutes (404 échantillons temporels), nous obtenons 19 minutes et 21 secondes de signaux de corrélation soit 387 échantillons temporels ($T = 387$). Si N est le nombre de SAM retenues pour un sujet donné, alors $N \times (N - 1)/2$ est le nombre de signaux de corrélation obtenus pour ce sujet. Afin de faire ressortir d'éventuels comportements spécifiques à une SAM, les paires sont regroupées de manière à présenter les signaux de corrélation de chaque SAM avec toutes les autres. Ainsi, pour chaque SAM n , $1 \leq n \leq N$, et chaque session s , $1 \leq s \leq 4$, sont regroupés les signaux de corrélations avec les autres $K = N - 1$ SAM qui permettent de former les jeux de signaux de corrélations, ou corrélogrammes temporels, définis par :

$$CC_{n,s} = \{\rho_{x,y}(t); 1 \leq n \leq N, 1 \leq m \leq N, m \neq n, 1 \leq t \leq T\} \quad (9.2)$$

Un exemple de tel corrélogramme est représenté FIGURE 9.1, en utilisant une échelle de couleur : une version non seuillée à gauche avec échelle de couleur, la même version seuillée à 0,3 à droite (seuil arbitraire, pour la visualisation) en version tricolore, marron ($\rho_{x,y}(t) > 0,3$), vert ($-0,3 < \rho_{x,y}(t) < 0,3$) et bleu ($\rho_{x,y}(t) < -0,3$). Ce mode de représentation a pour double avantage de pouvoir visualiser d'un coup d'œil les fluctuations de la corrélation pour le jeu de SAM considéré et de faire ressortir au travers de couleurs foncées les interdépendances (corrélation, anti-corrélation) entre SAM ainsi que la distribution temporelle de ces interdépendances. La version seuillée à droite permet de mieux appréhender visuellement ces inter-dépendances et leurs enchaînements. Sur cette double figure, neuf SAM sont étudiées ($N = 9$) soient 36 signaux de corrélation différents (les signaux des paires de corrélation étant regroupés par SAM, chacun apparait deux fois sur la figure, pour chacune des SAM de la paire considérée). L'ordre de visualisation des paires étant arbitraire, nous avons retenu l'ordre dans lequel nous apparaissaient les composantes lors de la sélection.

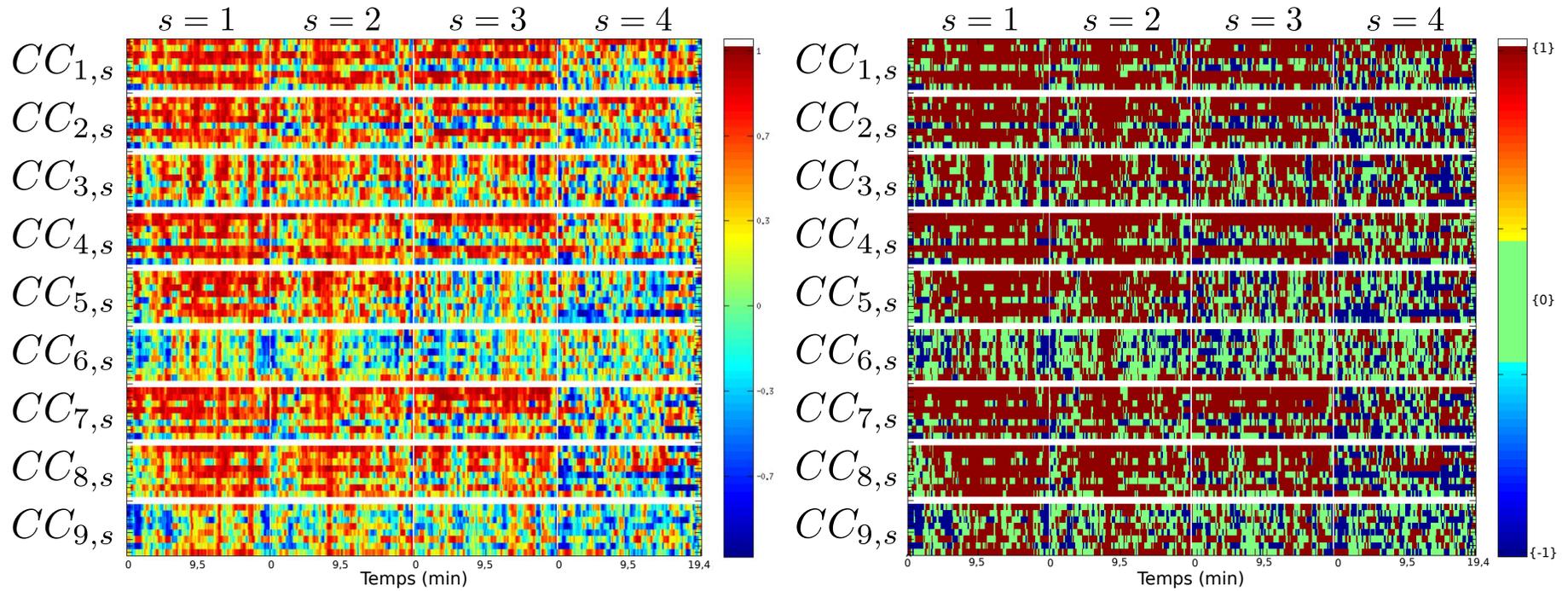


FIGURE 9.1 – Corrélogramme temporel sujet M3, 4 sessions s , $N = 9$ SAM, $K = 8$ paires de corrélation par SAM. La figure de gauche représente le signal de corrélation non seuillés, celle de droite, seuillés à $\pm 0,3$ (voir texte). En ligne, les signaux de corrélation regroupés par SAM. En colonne, les quatre sessions. Les 9 SAM étudiées pour ce sujet sont de 1 à 9 : un réseau idiosyncratique, l'OPN, un réseau idiosyncratique, le CN, le DMN, le réseaux des noyaux gris, l'OPFN, l'ETN et le MON (voir Partie I, section 1.4 pour leur description).

Nous distinguons, sur cette figure relative au sujet M3, des groupements de lignes (correspondant aux signaux de corrélation de certaines SAM) comme les signaux appartenant au $CC_{6,s}$ et dans une moindre mesure au $CC_{9,s}$, $\forall s$, et une colonne (celle de droite, $s = 4$, correspondant à la quatrième session) qui semblent avoir une plus grande fluctuation de la corrélation dans le temps que la plupart des autres lignes ou colonnes. Une première analyse de la FIGURE 9.1 met donc en lumière des différences de comportement entre les SAM et les sessions qui demandent à être évaluées de manière plus approfondie. Le seuillage effectué sur les corrélogrammes à $\pm 0,3$ permet d'accroître ces différences entre SAM et entre sessions. Ainsi, nous remarquons aisément que la session 4 du sujet M3 présente une DCF sensiblement différente des sessions 1 à 3, et ceci quel que soit le $CC_{n,s}$ excepté pour le $CC_{6,\cdot}$ et $CC_{9,\cdot}$. Pour ceux-ci, la DCF semble en première approximation (visuelle) être similaire sur les 4 sessions. À noter enfin que l'ensemble de ces observations est tiré des signaux de corrélation observés, c'est à dire de la partie observable des interactions entre SAM, sans présager des interactions effectives entre ces SAM, qui représentent la partie cachée de la DCF, comme nous le verrons au chapitre 10.

Pour approfondir ces premières observations, nous avons choisi d'analyser les histogrammes des valeurs de corrélations de ces paires de réseaux. Deux études sont réalisées en parallèle : la première entre les SAM, la deuxième entre les sessions. Pour la première étude, un histogramme est réalisé pour chaque réseau sur les trois premières sessions en prenant en compte les valeurs comprises dans ces $N - 1$ paires. La quatrième session étant particulière du fait de la prise d'un médicament ou d'un placebo, elle est écartée en ce qui concerne la différenciation des SAM. Sujet par sujet, nous calculons ensuite la distance entre chaque histogramme par corrélation d'enveloppe. Pour deux histogrammes h_1 et h_2 ayant le même nombre de classes (N), la mesure de similarité d entre h_1 et h_2 est définie par

$$d(h_1, h_2) = \frac{\sum_{i=1}^N \bar{h}_1(i) \bar{h}_2(i)}{\sqrt{\sum_{i=1}^N \bar{h}_1(i)^2 \sum_{i=1}^N \bar{h}_2(i)^2}}, \text{ où } \bar{h}(i) = h(i) - \frac{1}{N} \sum_{i=1}^N h(i).$$

Un réseau est considéré comme décorréolé s'il obtient des scores distants de plus de 2 écart-types par rapport à la moyenne. Pour la seconde étude, un histogramme est réalisé pour chaque session séparément avec l'ensemble des valeurs de corrélation des N SAM pour le sujet considéré. De la même manière que pour la première étude, la distance est calculée par corrélation d'enveloppe entre la session 4 (prise de médicament/placebo) et chacune des 3 autres sessions. La session 4 est considérée comme présentant une dynamique différente des autres si les 3 distances (Ses1 vs. Ses4, Ses2 vs. Ses4, Ses3 vs. Ses4) obtiennent des scores distants de plus de 2 écart-types par rapport à la moyenne.

9.3 Résultats préliminaires

La FIGURE 9.2 illustre une vue des histogrammes par paires et par session pour un sujet. De la même manière que pour la FIGURE 9.1 présentant les corrélogrammes $CC_{n,s}$ pour le sujet M3, nous remarquons que l'histogramme de $CC_{6,\cdot}$ et celui de $CC_{9,\cdot}$ montrent des valeurs de corrélation plus dispersées entre -0,5 et 0,5 que les histogrammes des autres SAM. Ces derniers étant largement centrés vers des valeurs supérieures à 0,5. La session 4 se détache aussi puisque les valeurs de corrélation semblent bien plus variables que pour les 3 premières sessions (excepté les SAM 6 et 9 qui semblent ne pas modifier leur comportement suivant les sessions).

9.3.1 Mise en évidence de réseaux décorréolés

Les premières observations des histogrammes font émerger une forte corrélation entre les SAM (moyenne des maxima des histogrammes = 0,65). Les histogrammes prenant en compte 387 valeurs

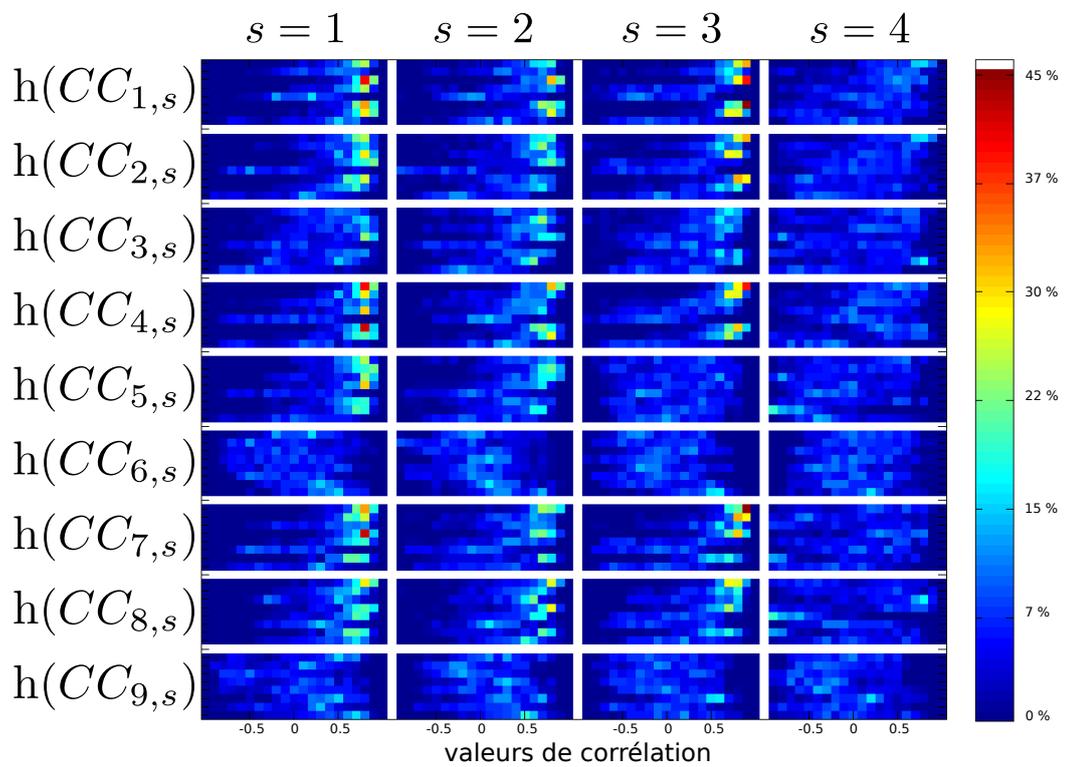


FIGURE 9.2 – Histogrammes des signaux de corrélation, sujet M3, 4 sessions s , $N = 9$ SAM, $K = 8$ paires de corrélation par SAM. En ligne, les histogrammes des signaux de corrélation regroupés par SAM. En colonne, les quatre sessions.

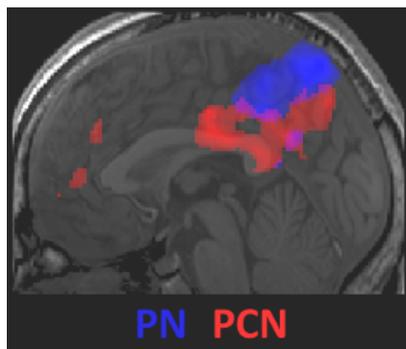


FIGURE 9.3 – Coupe sagittale du réseau PPCN, composé du cortex cingulaire postérieur (le réseau PCN en rouge) et du précunéus (le réseau PN bleu).

de corrélation (fenêtrée), ces fortes valeurs laissent présager une corrélation entre les SAM assez constante tout au long de la session et ce durant les 3 sessions de 20 minutes. Cela concorde avec des publications récentes qui évoquaient une tendance des réseaux à survenir en bouffée [126].

Cependant, le TABLEAU 9.2 des résultats des dissimilarités entre histogrammes permet de mettre en évidence des SAM qui sont significativement décorréliées des autres. Il s'agit plus fréquemment du réseau cingulaire postérieur. Dans la littérature, il est fréquent de retrouver les réseaux PN et PCN fusionnés, formant un seul réseau PPCN pour réseau cingulaire postérieur et précunéal (voir FIGURE 9.3). Avec cette information, il est d'autant plus intéressant de remarquer que le PN apparaît également parmi ces SAM décorréliées (de manière générale, le PN n'est présent, décorrélié ou non, que pour les sujets M1 et P1). Le réseau « noyaux gris », principalement constitué du striatum, ressort également de ces observations. Pour les sujets chez qui ces réseaux sont présents mais ne ressortent pas comme significativement décorréliés, il est noté qu'ils ont cependant une tendance à avoir des valeurs de corrélations plus faibles. La présence de ces réseaux décorréliés plaide pour une hypothèse neuronale dans la genèse des SAM. De plus, contrairement à ce qui a été observé lors de tâches d'activation, le DMN n'est retrouvé décorrélié que chez un seul sujet sur les neuf où il était présent. Il est aussi possible que la très longue durée d'enregistrement (3 fois 20 min, vs. 5 à 7 min dans la littérature) soit une des raisons de la particularité de ce résultat. Ceci illustre l'importance qu'il y a d'étudier la dynamique des interactions entre les SAM dans différents contextes.

TABLEAU 9.2 – SAM étant significativement décorréliées des autres réseaux au cours des sessions 1, 2 et 3.

Sujet	M1	M2	M3	M4	M5	M6	M7	M8	P1	P2	P3	P4
SAM	.	PCN	BG	PCN	PCN	DMN	.	BG	PN,BG	PCN	PCN	.

PCN : réseau cingulaire postérieur ; PN : réseau précunéal ; BG : réseau des ganglions de la base ; DMN : réseau par défaut, voir section 1.1 page 15 pour la carte spatiale et la fonction des réseaux. Les sujets n'ayant pas de réseau se démarquant sont désignés par un point dans la colonne correspondante.

9.3.2 Médicaments vs. Placebo

Si l'on observe une décorrélation pour certains sujets ayant pris un médicament lors de la quatrième session, les tests de dissimilarité entre histogrammes n'ont pas permis de souligner une différence significative entre cette session et les autres sessions pour les sujets M_i . En considérant les signaux de corrélation comme étant la représentation directe de la DCF, il n'a pas été non plus possible de tirer des conclusions quant à une différence dans la dynamique des processus cérébraux suite à la prise ou non d'un médicament pour un sujet donné alors même que les corrélogrammes

temporels laissent entrevoir le contraire (exemple visible en FIGURE 9.1). Une étude plus approfondie est donc nécessaire pour voir si l'on peut modéliser cette dynamique de manière à en savoir plus et peut-être ainsi dévoiler des différences non quantifiables jusqu'à présent. C'est tout l'objectif du prochain chapitre.

Modélisation et analyse de la DCF par MMC-Produit et MMC-Couplés

10

10.1 Principes et vue d'ensemble

Pour faire suite à nos premières observations, nous avons souhaité développer une méthode permettant de mettre en avant et de caractériser l'évolution des interactions entre les réseaux cérébraux s'étant manifestés dans une condition de repos. Ainsi, l'étude de la DCF peut être abordée par l'analyse de la dynamique des changements d'états de connectivité fonctionnelle cérébrale. Le concept d'état de connectivité fonctionnelle pour modéliser la DCF est largement partagé par les différentes méthodologies développées (voir section 8.1.3). Pour rappel, nous définissons, dans ce manuscrit, un état comme une configuration d'aires cérébrales connectées fonctionnellement à un instant t ou sur une période temporelle très courte (quelques secondes) et qui est potentiellement reproductible (au cours de la session, à travers les sessions, voire entre les sujets). Nous considérons donc que les SAM interagissent au cours de la session formant ainsi des configurations spatiales. De cette manière, la DCF peut être conçue comme une séquence d'états de connectivité préalablement définis. Les fluctuations des interactions entre les SAM peuvent déjà être perçues via les variations du niveau de corrélation entre les décours temporels des SAM au cours d'une session. L'observation directe de ces signaux de corrélations, appelés corrélogrammes et définis dans le chapitre précédent, n'est pas suffisante pour extraire des informations sur les propriétés dynamiques des processus cérébraux.

Nous proposons donc d'analyser ces corrélogrammes en les considérant comme la partie observable du processus conversationnel entre les réseaux, qui lui est considéré comme caché. Nous cherchons ainsi à inférer les corrélations « réelles » (inconnues) à partir des données observables bruitées. Chaque interaction entre deux SAM activées durant la session est observée via un signal de corrélation. Nous avons donc une multitude de processus observables pour lesquels nous voulons retrouver les séquences d'états cachés sous-jacentes. Ainsi, nous distinguons deux niveaux de modélisation des états :

- à l'échelle d'un signal de corrélation, donc de l'interaction entre deux SAM, un état représente la « corrélation » ou l'« anti-corrélation » entre ces deux réseaux ;
- à l'échelle globale du cerveau, un état représente une configuration spatiale de toutes les interactions d'une SAM avec les autres. Ce supra-état est construit à partir des états de l'échelle inférieure.

Comme nous l'avons annoncé dans le chapitre 8, l'approche que nous avons privilégiée est celle des modèles de Markov cachés. Chaque corrélogramme $CC_{n,s}$ associé à une SAM n et une session s (défini par l'équation (9.2)), est composé de K signaux de corrélation, les séquences d'observations de longueur T . Un modèle $\lambda_{n,s}$ est supposé en être à l'origine.

À partir de ces éléments de base, nous pouvons représenter une vue d'ensemble, FIGURE 10.1,

des étapes de modélisation et d'analyse de la DCF telles que nous les avons conçues. Pour cette figure, le corrélogramme représentant les signaux d'observation est composé de trois signaux de corrélations $\mathbf{Y}^{(1)}$, $\mathbf{Y}^{(3)}$, $\mathbf{Y}^{(3)}$ (corrélations d'une SAM avec trois autres).

Notre méthode est composée de trois étapes principales illustrées par les parties a), b) et c) de la FIGURE 10.1. Ces étapes seront décrites dans les sections suivantes dans cet ordre : la modélisation (FIGURE 10.1.a) afin de placer nos données dans le cadre markovien ; l'initialisation et l'apprentissage (FIGURE 10.1.b) pour représenter au mieux nos données et en extraire leurs caractéristiques, et enfin l'analyse des résultats (FIGURE 10.1.c) qui va nous aider à évaluer et comparer les modèles notamment les uns par rapport aux autres.

10.2 Modélisation

Notre approche étant exploratoire, nous avons souhaité la développer avec le moins d'*a priori* possible. C'est pourquoi, nous n'avons pas fait d'hypothèse sur l'importance d'un canal par rapport aux autres (pas de hiérarchisation entre les SAM) et donc aucune pondération n'a été mise en place. D'autre part, nous souhaitons capturer toutes les dépendances possibles entre les réseaux, bien que cela augmente le nombre de paramètres à estimer, afin d'englober le maximum d'observations dans la modélisation. La dimension de nos modèles est, quoi qu'il en soit, limitée par les hypothèses faites lors du choix du nombre de réseaux étudiés.

Pour reprendre les notations établies dans les chapitres précédents, nous avons :

N Le nombre de SAM étudiées.

K Le nombre de canaux d'observation.

Dans le cas de l'étude des paires de signaux de corrélation, $K = N - 1$.

\mathbf{Y} = $\{Y_t^{(k)}, 1 \leq k \leq K; 1 \leq t \leq T\}$ les séquences d'observation sur les K canaux, d'une durée T .

\mathbf{X} = $\{X_t^{(k)}, 1 \leq k \leq K; 1 \leq t \leq T\}$ les séquences d'états sur les K canaux, d'une durée T .

\mathbf{i} = $[i^{(1)}, i^{(2)}, \dots, i^{(K)}]^T$ un vecteur d'états

S L'espace d'états et $S^{(k)} = \{i^{(k)}\}$ l'espace d'états sur le canal k .

A = $\{a_{\mathbf{i}\mathbf{j}}, \mathbf{i}, \mathbf{j} \in S\}$, les probabilités de transition.

B = $\{b_{\mathbf{j}}(\mathbf{Y}_t), \mathbf{j} \in S\}$, les probabilités d'observation (conditionnelles aux états).

Π = $\{\pi_{\mathbf{i}}, \mathbf{i} \in S\}$, les probabilités initiales.

Ses Le nombre de sessions.

À chaque corrélogramme temporel $CC_{n,s}$, défini par l'équation (9.2), est associé un modèle $\lambda_{n,s}(\theta)$, $1 \leq n \leq N$, $1 \leq s \leq Ses$. Ce corrélogramme temporel est considéré comme le processus observable de $\lambda_{n,s}(\theta)$, c'est-à-dire $\mathbf{Y} \equiv CC_{n,s}$ (voir FIGURE 10.1.a) et $Y_t^{(k)} = \rho_{n,m}(t)$, $1 \leq t \leq T$; $1 \leq k \leq K$. Ainsi, pour une SAM considérée et une session donnée, un modèle va être décrit par ses paramètres $\theta_{n,s} = (A, B, \Pi)$.

Nous avons choisi, dans un premier temps de définir l'espace d'états $S^{(k)}$ de $\lambda_{n,s}^{(k)}$ par deux états, l'état d'anti-corrélation "-1" et l'état de corrélation "+1", d'où $S^{(k)} = \{"-1"; "+1"\}$ et $\text{Card}(S^{(k)}) = 2$. La dimension de la matrice de transition A , directement liée au cardinal de $S^{(k)}$ ainsi qu'au nombre K de canaux modélisés, possède ainsi une dimension de $2^K \times 2^K$.

Les MMC multidimensionnels utilisés ont été décrits en Partie I, section 3.2.2 (page 32) pour les MMC-Produit (MMC-P) et section 3.2.3 (page 33) pour les MMC-Couplés (MMC-C). Un rappel de leurs paramètres est fait ci-dessous.

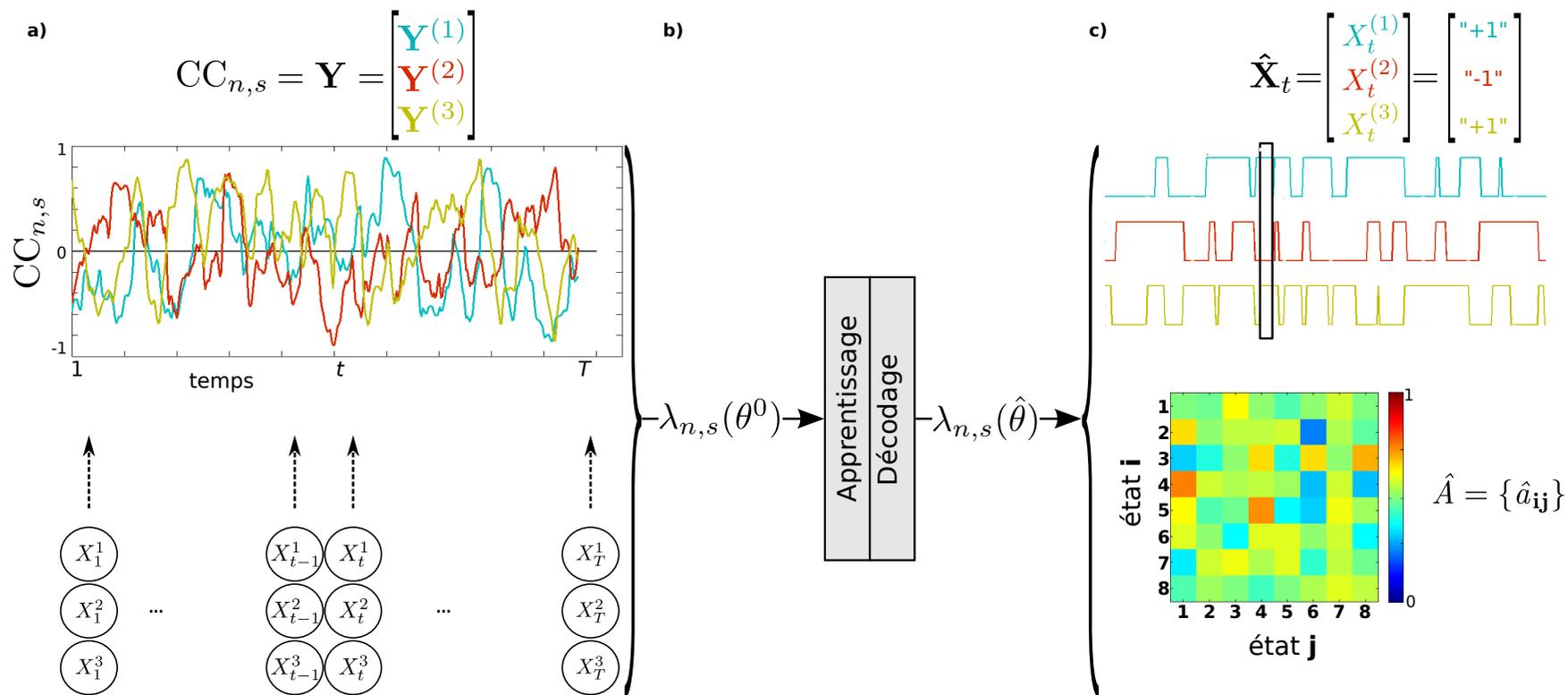


FIGURE 10.1 – Vue d’ensemble de la méthode développée pour analyser la DCF. a) Modélisation avec un MMC multidimensionnel du corrélogramme $CC_{n,s}$. b) Initialisation des paramètres, apprentissage et décodage. c) Analyse des résultats : séquence d’états $\hat{\mathbf{X}}$ décodée en haut, matrice de transition \hat{A} estimée par apprentissage en bas. Exemple pour un sujet, pour une SAM n parmi 4 ($K = 3$) et une session s .

Concernant le MMC-P, les probabilités d'observations et des états sont modélisées de manière jointe sur tous les canaux. les paramètres du MMC-P s'écrivent :

$$\pi_{\mathbf{i}} = P[\mathbf{X}_1 = \mathbf{i}], \quad (10.1)$$

$$a_{\mathbf{ij}} = P[\mathbf{X}_t = \mathbf{j} | \mathbf{X}_{t-1} = \mathbf{i}], \quad (10.2)$$

$$b_{\mathbf{j}}(\mathbf{Y}_t) = P[\mathbf{Y}_t | \mathbf{X}_t = \mathbf{j}], \quad \mathbf{i}, \mathbf{j} \in S; 1 \leq t \leq T \quad (10.3)$$

Dans le second modèle présenté, le MMC-C, chaque état $X_t^{(k)}$, sur chaque canal k , à un temps t , va générer sa propre observation mais les états peuvent interagir entre eux. Nous n'avons pas choisi de limiter les dépendances entre les états aux canaux les plus proches comme cela est fait parfois dans un but de simplification. Ainsi, pour le MMC-C, un état sur un canal va dépendre des états précédents sur chaque canal. Les paramètres du MMC-C s'écrivent :

$$\pi_{\mathbf{i}} = \prod_{k=1}^K \pi_{i^{(k)}}^{(k)} = \prod_{k=1}^K P[X_1^{(k)} = i^{(k)}], \quad (10.4)$$

$$a_{\mathbf{ij}} = \prod_{k=1}^K a_{i^{(k)}j^{(k)}}^{(k)} = \prod_{k=1}^K P[X_t^{(k)} = j^{(k)} | \mathbf{X}_{t-1} = \mathbf{i}] \quad (10.5)$$

Le nombre de données étant limité, on va supposer l'indépendance entre les canaux pour estimer les $b_{\mathbf{j}}(\mathbf{Y}_t)$. D'où

$$\begin{aligned} b_{\mathbf{j}}(\mathbf{Y}_t) &= P[\mathbf{Y}_t | \mathbf{X}_t = \mathbf{j}] \\ &= \sum_{m=1}^{M_{\mathbf{j}}} w_{\mathbf{j},m} \prod_{k=1}^K P[Y_t^{(k)} | X_t^{(k)} = j^{(k)}] \\ &= \sum_{m=1}^{M_{\mathbf{j}}} w_{\mathbf{j},m} \prod_{k=1}^K b_{j^{(k)}}^{(k)}(Y_t^{(k)}) \end{aligned} \quad (10.6)$$

Les $b_{j^{(k)}}^{(k)}(\mathbf{Y}_t)$ modélisent ainsi la distribution de $Y_t^{(k)} = \rho_{n,m}(t)$ sachant l'état $X_t^{(k)} = j^{(k)}$, $j^{(k)} \in \{-1, +1\}$. Ces probabilités sont classiquement modélisées par un mélange de gaussiennes. Nos observations $\mathbf{Y}^{(k)}$ étant uni-dimensionnelles et chaque canal ayant le même poids, l'équation (10.6) peut être écrite :

$$b_{\mathbf{j}}(\mathbf{Y}_t) = \prod_{k=1}^K \left[\mathcal{N}(Y_t^{(k)}, \mu_{j^{(k)}}, \sigma_{j^{(k)}}) \right] \quad (10.7)$$

Par exemple, pour $K = 3$, et $\mathbf{i} = [-1, +1, +1]^T$:

$$b_{\mathbf{i}}(\mathbf{Y}_t) = \mathcal{N}(Y_t^{(1)}, \mu_{-1^{(1)}}, \sigma_{-1^{(1)}}) \cdot \mathcal{N}(Y_t^{(2)}, \mu_{+1^{(2)}}, \sigma_{+1^{(2)}}) \cdot \mathcal{N}(Y_t^{(3)}, \mu_{+1^{(3)}}, \sigma_{+1^{(3)}})$$

Avec K canaux et $\text{Card}(S^{(k)}) = 2$, il y a donc $2K$ moyennes et $2K$ variances.

10.3 Initialisation

Les paramètres d'un modèle MMC-O ou d'un modèle MMC-C doivent être initialisés avant l'apprentissage. Cette étape est particulièrement importante puisque la procédure itérative d'apprentissage utilisée (algorithme de Baum-Welch, voir page 30 paragraphe 3.1.3.3) converge vers des valeurs de paramètres qui assurent un maximum local de $P[\mathbf{Y} | \lambda(\theta)]$. On définit ainsi un vecteur paramètre initial $\theta = \theta^0$ pour un MMC-P ou pour un MMC-C.

Matrice de probabilités de transition A Pour les MMC-P comme pour les MMC-C, les probabilités de transitions $\{a_{ij}, \mathbf{i}, \mathbf{j} \in S\}$ ont été fixées initialement à $\frac{1}{2^{K+1}}$ excepté pour les éléments de la diagonale, $\{a_{ii}\}$, qui ont été initialisés à $\frac{2}{2^{K+1}}$. Ce choix a été fait afin de n'injecter aucun *a priori* sur les transitions générales avant l'apprentissage mises à part les transitions de bouclage des états sur eux-mêmes. Pour ces transitions particulières, la probabilité est légèrement augmentée. En effet, les premiers résultats publiés en DCF et s'intéressant aux transitions entre les états [119] révèlent un faible nombre de changements d'état. Bien que notre étude diffère dans sa modélisation (nombre d'états et longueur de la session), l'augmentation des transitions de bouclage dans un état (transition d'un état vers lui-même) permet de prendre en compte cet aspect de la DCF.

Concernant les MMC-C, les probabilités $\{a_{ij}, \mathbf{i}, \mathbf{j} \in S\}$ sont données par l'équation (10.5). Cependant pour l'initialisation, elles sont définies directement comme expliqué dans le paragraphe précédent. Les probabilités $a_{ij}^{(k)}$ qui permettent d'estimer ces probabilités a_{ij} sont calculées par la suite, lors de l'apprentissage.

Matrice des probabilités d'émission des observations B L'équation (10.7), qui définit les probabilités $b_j(\mathbf{Y}_t)$ telles que nous les avons modélisées, requiert une initialisation des moyennes et variances pour chaque état ("−1" ou "+1") et chaque canal K . Les moyennes $\mu_{j^{(k)}}$ et variances $\sigma_{j^{(k)}}$ sont définies à partir des données en considérant les états comme directement observables. Ainsi, pour chaque canal k , la moyenne et la variance des valeurs positives des signaux d'observation $\mathbf{Y}^{(k)}$ permettent d'initialiser $\mu_{+1^{(k)}}$ et $\sigma_{+1^{(k)}}$. Il en est de même avec les valeurs négatives pour définir $\mu_{-1^{(k)}}$ et $\sigma_{-1^{(k)}}$. Cette initialisation, mise en place pour cette première démarche de modélisation de la DCF avec les MMC multidimensionnels, nous permet d'être au plus proche des données. Les probabilités d'émission des observations sont initialisées de la même manière pour les MMC-P et MMC-C.

Matrice de probabilités initiales Π Les matrices de probabilités initiales ont été fixées de manière uniforme selon les états, soit $\pi_{\mathbf{i}} = \frac{1}{2^K}$, pour les MMC-P et MMC-C. Pour les MMC-C, nous retrouvons bien les probabilités $P[X_1^{(k)} = i^{(k)}] = \frac{1}{\sqrt{1/2^K}} = 1/2$ qui exprime bien une probabilité uniforme d'être dans l'un des deux états ("−1" ou "+1") sur un canal k à $t = 1$. Cette initialisation ne donne ainsi aucun *a priori* sur un état initial.

10.4 Apprentissage

Les $\lambda_{n,s}(\theta)$ sont ensuite entraînés de manière itérative jusqu'à convergence ($\theta = \hat{\theta}$). Les probabilités d'émission des observations $b_j(\mathbf{Y}_t)$ telles que nous les avons initialisées sont déjà apprises au plus proche des données et ne seront pas ré-estimées pour ces modélisations.

Matrice de probabilités de transition A Le MMC-P pouvant se ramener à un MMC uni-dimensionnel, les formules de ré-estimation sont reprises des formules eq. (3.11) en y intégrant simplement les K canaux :

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(\mathbf{i}, \mathbf{j})}{\sum_{t=1}^{T-1} \gamma_t(\mathbf{i})} = \frac{\sum_{t=1}^{T-1} \alpha_t(\mathbf{i}) a_{ij} b_j(\mathbf{Y}_{t+1}) \beta_{t+1}(\mathbf{j})}{\sum_{t=1}^{T-1} \alpha_t(\mathbf{i}) \beta_t(\mathbf{j})} \quad (10.8)$$

Concernant le MMC-Couplés, les formules sont données par Nefian et coll. [70] avec :

$$\hat{a}_{ij}^{(k)} = \frac{\sum_{\mathbf{j}} \sum_{t=1}^{T-1} \alpha_t(\mathbf{i}) a_{ij} b_{\mathbf{j}}(\mathbf{Y}_{t+1}) \beta_{t+1}(\mathbf{j})}{\sum_{t=1}^{T-1} \alpha_t(\mathbf{i}) \beta_t(\mathbf{j})}, \quad \forall \mathbf{j} \in S \text{ tel que } j^{(k)} = j \quad (10.9)$$

Les éléments \hat{a}_{ij} de la matrice de transition \hat{A} sont ensuite calculés grâce à la formule détaillée en équation (10.5).

Matrice de probabilités initiales Π Quel que soit le modèle (MMC-P ou MCC-C), les probabilités initiales Π sont estimées selon l'équation (10.10) suivante :

$$\hat{p}_{i\mathbf{i}} = \gamma_1(\mathbf{i}) = \frac{\alpha_1(\mathbf{i}) \beta_1(\mathbf{i})}{\sum_{i=1}^N \alpha_1(\mathbf{i}) \beta_1(\mathbf{i})} \quad (10.10)$$

10.5 Analyse

Les approches Markoviennes cachées en général et les MMC-Produit ou les MMC-Couplés en particulier offrent à l'utilisateur dans un même cadre probabiliste, de nombreux outils d'analyse des données observées. Il est ainsi possible de comparer de manière globale des modèles entre eux, sur la base de distances. Cet aspect est particulièrement intéressant dans le contexte de l'analyse de la DCF puisqu'il permet de comparer globalement des DCF entre elles à partir de leur modèle statistique ; ces DCF pouvant provenir d'un même sujet imagé à des moments différents, ou de sujets différents appartenant ou non à un même groupe (de patients atteints d'un même trouble ou porteurs de la même pathologie). Il est aussi possible, au travers de la matrice de transition d'un modèle de DCF, de capturer la dynamique d'ensemble de la DCF observée sous une forme matricielle et probabiliste. Cette même matrice permet aussi de cibler certains comportements particuliers du processus caché modélisé. Enfin, le décodage de Viterbi des séquences d'observation permet d'analyser la DCF en préservant la dimension temporelle, tout en permettant là aussi de cibler des comportements spécifiques de la DCF dans l'espace des processus cachés et/ou observables et dans le temps. Nous développons maintenant plus en détail ces différents aspects de l'analyse multidimensionnelle et markovienne cachée de la DCF.

10.5.1 Comparaison de modèles

Il est possible de comparer des MMC grâce à des mesures de distances entre deux modèles λ_1 et λ_2 . Le calcul de la distance tel qu'il est proposé par Juang et Rabiner [67, 142] est défini ci-dessous :

$$D(\lambda_1, \lambda_2) = \frac{1}{T} [\log P(\mathbf{Y}_{(2)}|\lambda_1) - \log P(\mathbf{Y}_{(2)}|\lambda_2)] \quad (10.11)$$

où $\mathbf{Y}_{(2)}$ est une séquence de T observations générées par le modèle λ_2 . L'équation (10.11) mesure la vraisemblance avec laquelle les observations générées par λ_2 peuvent correspondre au modèle λ_1 par rapport à la vraisemblance avec laquelle ces mêmes observations sont bien générées par le modèle λ_2 . Cela permet de comparer deux modèles par leur capacité à générer séparément la même séquence d'observations. Cette mesure n'étant pas symétrique, elle le devient en calculant :

$$D_{sym}(\lambda_1, \lambda_2) = \frac{D(\lambda_1, \lambda_2) + D(\lambda_2, \lambda_1)}{2} \quad (10.12)$$

L'utilisation de cette distance nécessite de faire générer des observations par les modèles à comparer. Aucune information n'est apportée à l'utilisateur quant au nombre T d'observations à générer, nombre que nous appellerons T_{gen} dorénavant. En pratique, une distance sera calculée pour deux valeurs de T_{gen} .

10.5.2 Matrice de transition, Décodage de Viterbi

Après l'apprentissage, le modèle $\lambda_{n,s}(\hat{\theta})$ est utilisé pour décoder \mathbf{Y} et inférer la séquence d'états cachés $\hat{\mathbf{X}}$ qui explique le mieux les observations $CC_{n,s}$. Pour illustrer ces étapes, la FIGURE 10.1.c montre la matrice de transitions \hat{A} obtenue après apprentissage et les séquences décodées à partir des observations $CC_{n,s}$ de cette même figure (FIGURE 10.1.a).

La matrice de transition \hat{A} d'un modèle fournit une représentation compacte et globale du comportement statistique de la DCF observée. Il nous permet d'identifier rapidement les transitions qui sont privilégiées, ignorées ou modifiées pour un sujet donné ou entre les sujets. La matrice de transition nous permet également de nous concentrer sur des informations plus précises. En particulier, le terme diagonal $\{a_{ii}\}$ correspond à la probabilité de boucler dans l'état i de t à $t + 1$. Sa grandeur reflète la stabilité temporelle de l'état de connectivité fonctionnelle i qui, rappelons le, représente une configuration particulière de réseaux corrélés et anti-corrélés couvrant une large part du cerveau.

Quant aux séquences d'états décodées par l'algorithme de Viterbi, $\hat{\mathbf{X}}$ pour le MMC-P, et sur chaque canal k , $\hat{\mathbf{X}}^{(k)}$ pour le MMC-C, elles donnent accès à la transcription symbolique des données d'entrée observées (la corrélation inter SAM), exprimée comme une séquence des états visités. En préservant la dimension temporelle, la séquence $\hat{\mathbf{X}}$ pourra nous permettre par la suite d'analyser des fenêtres temporelles spécifiques au cours de l'examen d'IRMf. De telles séquences peuvent également être le point de départ pour des analyses simples ou avancées qui enrichissent la description de la DCF. Ces analyses peuvent être effectuées sur une seule séquence $\hat{\mathbf{X}}$, devenant spécifiques de $\hat{\mathbf{X}}$, ou réalisées sur un ensemble $\{\hat{\mathbf{X}}\}$ de séquences d'états décodées, vis-à-vis d'un sujet, d'une SAM ou d'un groupe de sujets. Des statistiques simples qui mettent en évidence les états les plus communs, ceux jamais atteints, le temps passé dans chaque état ou encore le nombre de transitions, sont également des mesures intéressantes pour caractériser au mieux la DCF. De manière plus avancée, il peut être envisagé de rechercher ou de reconnaître des « *patterns* » de connectivité, et d'en mesurer la reproductibilité dans le temps et entre sujets.

Il convient de souligner que les statistiques extraites de ces outils sont intimement liées à la durée des séquences observations disponibles en premier lieu. La reproductibilité des états ou leur absence sont liées au nombre d'états et de transitions qu'il nous est permis de voir sur la durée de la session. Ainsi, et cela est vrai quel que soit les techniques de mesure de la DCF présentes dans la littérature, la durée des sessions est un paramètre important à prendre en compte. Elle peut vite s'avérer être une limite à la généralisation des résultats en IRMf lors de sessions de durée parfois inférieure à 5 minutes.

Applications cliniques : Médicament vs placebo, Démence à corps de Lewy **11**

11.1 Application 1 basée sujet : « Médicament vs Placebo »

11.1.1 Objectifs et contraintes

Cette première application d'analyse de la DCF par MMC-P et MMC-C a été réalisée sur la base de données B2. Nous avons présenté dans le chapitre précédent des observations préliminaires sur cette base. Pour rappel, l'extraction des SAM a été faite au niveau du sujet. Le nombre et le type de SAM étudiées ne sont donc pas les mêmes selon les sujets. Une comparaison directe des interactions concernant une SAM en particulier, entre les sujets, n'est pas possible.

L'étude faite au paragraphe 9.3.2 ne nous a pas permis de différencier de manière statistique la quatrième session, au cours de laquelle un médicament ou un placebo avait été pris, par rapport aux trois sessions précédentes sans médicament ni placebo. Il s'agit donc dans cette application centrée sur le sujet, de regarder s'il est possible de distinguer cette quatrième session des trois premières à l'aide de mesures effectuées sur la DCF modélisée par MMC-P ou MMC-C. De la même manière que pour le chapitre précédent, nous allons analyser les interactions entre les SAM communes aux quatre sessions d'un même sujet.

Nous disposons d'un nombre restreint d'observations ($T = 387$ par signal de corrélation) qui doit être mis en parallèle avec le nombre d'états rendus possibles par notre approche de modélisation multidimensionnelle de la DCF. Le nombre de SAM étudiées est ainsi limité par l'ordre de grandeur de la modélisation pour une interprétation solide et par le temps de calcul nécessaire. À titre d'exemple, le TABLEAU 11.1 donne une estimation de la dimension des paramètres du modèle en évaluant le nombre d'états-produit possibles selon le nombre N de SAM étudiées et le cardinal de $S^{(k)}$. Au vu de ce tableau et du nombre d'échantillons temporels par session, des choix pour circonscrire l'étude à moins de 10 SAM doivent être faits. Par ailleurs, le nombre de SAM prises en compte dans la modélisation markovienne influe aussi sur le temps de calcul qui peut aller de quelques minutes ($N = 6$) à quelques jours ($N = 10$).

11.1.2 Mise en œuvre

Pour les trois sujets ayant plus de 15 SAM communes aux quatre sessions (sujets M6, P1, P2, voir TABLEAU 9.1), le nombre de réseaux a été restreint. Pour ces sujets, seuls les RSN (réseaux décrits par la littérature, que l'on peut retrouver chez d'autres sujets) ont été retenus. Ces réseaux ont été choisis car ils sont mieux connus donc l'interprétation qui peut découler des résultats pourra être facilitée. Pour P1 et P2, le nombre final de SAM qui découle de cette sélection est de 9. Pour M6, le nombre de RSN reste trop élevé ($N = 12$) même après cette étape de sélection. Nous avons donc mis de côté les RSN qui étaient présents plusieurs fois pour ne garder que le plus corrélé

TABLEAU 11.1 – Nombre d'états-produit en fonction du nombre N de SAM considérées, et du cardinal du sous-espace d'états $S^{(k)}$.

$\text{Card}(S^{(k)}) \backslash N$	3	5	7	8	9	10	11	16
2	4	16	64	128	256	512	1 024	32 768
3	9	81	729	2 187	6 561	19 683	59 049	$>1.10^7$

Par exemple, pour $\text{Card}S^{(k)}=2$ avec $N=3$ SAM prises en compte dans l'étude, il existe $K=N-1$ paires de corrélation se rapportant à chaque SAM (dimension du corrélogramme : $K \times T$) soit 2^3 états-produits possibles.

spatialement aux RSN des modèles de la littérature.

Le nombre final de SAM retenues pour les études par sujet est donné dans le TABLEAU 11.2. Pour rappel, les signaux de corrélation obtenus sont de longueur $T = 387$ et nous travaillons sur quatre sessions d'où $ses = 4$. Pour $1 \leq n \leq N$ et $1 \leq s \leq 4$, le nombre de corrélogrammes temporels $CC_{n,s}$ est de $N \times 4$ par sujet, soit autant de modèles MMC multidimensionnels à développer.

TABLEAU 11.2 – Nombre de SAM retenues pour la modélisation

Sujet	M1	M2	M3	M4	M5	M6	M7	M8	P1	P2	P3	P4
N	10	9	9	8	9	9	6	8	9	9	9	8
K	9	8	8	7	8	8	5	7	8	8	8	7

N : nombre de SAM communes aux 4 sessions par sujet, pour la base B2. $K = N - 1$: nombre de canaux par modèles.

Afin d'évaluer des modifications de la DCF entre sessions, les modèles $\lambda_{n,s}$ ont été comparés deux à deux par la mesure de distance définie par les équations (10.11) et (10.12). Pour chaque sujet, deux jeux de données sont créés :

- le premier comprend les distances entre les sessions 1, 2 et 3, pour toutes les SAM : $D_{1-3} = \{D_s(\lambda_{n,s}, \lambda_{n,s'}), (s, s') \in \{(1, 2), (1, 3), (2, 3)\}, 1 \leq n \leq N\}$;
- le second se compose des distances entre la session 4 et les sessions 1, 2 et 3, pour toutes les SAM : $D_4 = \{D_s(\lambda_{n,s}, \lambda_{n,s'}), (s, s') \in \{(1, 4), (2, 4), (3, 4)\}, 1 \leq n \leq N\}$.

Pour le calcul de ces distances (voir paragraphe 10.5.1), la longueur des séquences générées T_{gen} a été choisie pour se rapprocher des séquences initiales, d'où $T_{gen} = 400$. L'approche a aussi été testée pour $T_{gen} = 1000$ et $T_{gen} = 1500$ afin de prendre en compte plusieurs cas de figure. Afin de confirmer ou d'infirmer l'existence d'une différence significative de la DCF dans la session 4 par rapport aux trois autres sessions, un test de Student à deux échantillons (*two samples t-test*) est réalisé, pour chaque sujet, entre les jeux de distances D_{1-3} et D_4 . Des analyses complémentaires décrivant la DCF sont réalisées grâce aux outils présentés en 10.5.2.

11.1.3 Résultats

La convergence avec la modélisation MMC est rapide, après trois itérations nous obtenons les paramètres estimés « optimaux » \hat{A} et \hat{X} .

11.1.3.1 Comparaison basée sujet des DCF

Le TABLEAU 11.3 présente les p -valeurs des tests de Student. Avec la modélisation par MMC-P ainsi qu'avec la modélisation par MMC-C, cinq sujets (M2, M3, M4, M6, M8) du groupe ayant pris un médicament lors de la dernière session affichent une DCF significativement différente ($p < 0,05$)

pour cette session par rapport aux trois précédentes. De telles variations n'ont pas été observées chez les sujets ayant pris un placebo. Ces résultats se retrouvent pour différentes générations de signal ($T_{gen} = 400, 1000$ ou 1500) pour le MMC-P comme pour le MMC-C.

La mesure de la distance nous a permis simplement de distinguer différentes sessions sans donner pour autant les caractéristiques des signaux qui ont permis ces résultats. Pour essayer de caractériser ce changement de comportement, nous avons regardé l'évolution du nombre de transitions au cours des quatre sessions ($\hat{\mathbf{X}}_t \neq \hat{\mathbf{X}}_{t-1}, \forall t$). Les sujets M2, M3, M4 et M8 présentent une augmentation du nombre de changements d'états en session 4 par rapport aux autres sessions. Les autres sujets, M1, M5, M6, M7, P1-P4, n'ayant d'ailleurs pas montré de signe de modifications de leur DCF, ne présentent pas de variation notable en termes de nombre de transitions entre les sessions. En moyenne, les sujets présentent 81 transitions par session (pour rappel $T = 387$ pour une session). Cela représente environ 4 changements d'états par minute. Évidemment, les transitions ne sont pas réparties de manière homogène dans le temps, comme nous allons le voir dans le paragraphe suivant, mais ce chiffre donne une estimation de la variabilité de la connectivité cérébrale fonctionnelle telle que nous l'étudions.

TABLEAU 11.3 – Résultats des comparaisons entre sessions par sujet, pour les MMC-P en haut et les MMC-C en bas, et différents T_{gen} .

Sujet	T_{gen}	MMC-P				MMC-C			
		400	1000	1500	h	400	1000	1500	h
M1		0,18	0,51	0,71	0	0,29	0,64	0,62	0
M2		<0,000	<0,000	<0,000	1	<0,000	<0,000	<0,000	1
M3		0,040	0,003	<0,000	1	0,078	<0,000	<0,000	1
M4		<0,000	0,005	0,030	1	<0,000	0,030	0,005	1
M5		0,99	0,16	0,70	0	0,18	0,09	0,86	0
M6		0,003	<0,000	<0,000	1	<0,000	0,008	0,006	1
M7		0,32	0,45	0,37	0	0,049	0,08	0,37	0
M8		<0,000	<0,000	<0,000	1	<0,000	<0,000	<0,000	1
P1		0,59	0,72	0,75	0	0,91	0,38	0,50	0
P2		0,22	0,54	0,21	0	0,21	0,35	0,56	0
P3		0,25	0,32	0,15	0	0,35	0,15	0,25	0
P4		0,63	0,46	0,49	0	0,73	0,97	0,73	0

p -valeur pour les 12 sujets (8 avec prise de médicament : M1-M8, 4 avec prise de placebo : P1-P4), pour les deux approches MMC-P et MMC-C, et $T_{gen}=400, 1000$ ou 1500 observations. Les p -valeurs significatives $<0,05$ (en gras) permettent de décider que la DCF de la session 4 est significativement différente de celles des autres sessions ($h=1$).

11.1.3.2 Caractérisation de la DCF d'un sujet : matrices de transition, décodages de Viterbi

Nous présentons ici pour un sujet donné, les résultats typiquement apportés par la matrice de transition d'un modèle de DCF (MMC-P ou MMC-C) et par le décodage de Viterbi des séquences d'observations. Le sujet considéré est le sujet M7, lequel présentait, parmi tous les sujets, le nombre de SAM communes aux quatre sessions le plus faible ($N = 6$ SAM, $K = 5$ signaux de corrélation par corrélogramme $CC_{n,s}$, $1 \leq n \leq N$, $1 \leq s \leq 4$), et par conséquent le nombre d'états de connectivité obtenus par modélisation également le plus faible ($2^K = 32$ états).

La FIGURE 11.1 révèle les probabilités de transitions après apprentissage en présentant les matrices A pour les quatre sessions (en colonne) et les deux modélisations (en ligne) pour la SAM 1 du sujet M7. La dernière ligne met en lumière les différences entre ces matrices en fonction de la modélisation choisie (MMC-P ou MMC-C). Pour plus de lisibilité, les états i ont été noté pour cette figure de 1 à 32 ($=2^5$); l'état 1 correspond à l'état ["-1" "-1" "-1" "-1" "-1"] et l'état 32 (["+1" "+1" "+1" "+1" "+1"], la progression de l'état 1 à l'état 32 s'effectuant à l'image du binaire naturel). On observe des probabilités fortes sur la diagonale révélant des probabilités de bouclage importantes. Les différences observées entre les probabilités des deux modélisations sont majoritairement faibles (en dessous de 0,01) néanmoins certaines transitions voient leur probabilité varier jusqu'à 0,95. Malgré la faible différence des probabilités de transition entre MMC-P et MMC-C, il est intéressant de remarquer que pour une grande majorité ($> 95\%$) des matrices observables (quelque soit les sujets, les sessions et les réseaux), le MMC-P fournit systématiquement un nombre plus important de transitions dont les probabilités sont supérieures à 0,5. À l'inverse si l'on dénombre les transitions probables à plus de 0,1, c'est le MMC-C qui se démarque pour la très grande majorité ($> 85\%$) des matrices estimées.

Ces légères différences se retrouvent sur les séquences d'états décodées. En effet, les décodages sont similaires pour 90 % à 99 % des états décodés sur une session et un canal. Les différences se présentent généralement au moment des transitions qui vont se faire plus tôt ou plus tard selon le décodage fourni par l'une ou l'autre des modélisations. Aucune distinction allant dans un sens ou dans l'autre d'une modélisation n'est cependant observable. Les FIGURES 11.2 et 11.3 représentent la répartition temporelle de quelques états décodés superposés sur leurs observations, c'est-à-dire sur les corrélogrammes $\mathbf{Y} = CC_{n,s}$. Les deux figures concernent les interactions de la SAM numéro 3 (l'OPN, un réseau visuel) avec les autres SAM pour les sessions 1 et 3. La couleur bleue est utilisée pour représenter les vecteurs d'états i où 2 canaux sont à "+1" et deux canaux sont à "-1" parmi les 5 canaux. Ce sont les états où il y a sensiblement le même nombre de SAM qui sont corrélées et anti-corrélées à la SAM de référence c'est-à-dire celle étudiée (la SAM 3). À l'inverse, les couleurs rouge et rose pâle correspondent à l'état de connectivité où toutes les SAM sont corrélées avec celle de référence (en rouge), ou aux états où seule une SAM sur 5 n'est pas corrélée avec celle de référence (en rose pâle). Nous remarquons notamment sur la FIGURE 11.2 une répartition particulière puisque les états de connectivité fortement différenciés se manifestent principalement en première partie de session et les états de grande corrélation inter-SAM en deuxième partie de session. La deuxième figure présentée (FIGURE 11.3) (même sujet, même SAM de référence mais session 3) met en évidence un autre type de phénomène courant lors des analyses de repos : des états stables et un faible nombre de transitions. Les données extraites des séquences d'états décodés sont aussi visibles sur cette figure à savoir : 1) le sujet est principalement dans un état où toutes les SAM sont corrélées à celle de référence, 2) cet état de super-connectivité dure parfois jusqu'à 6 minutes consécutives. Cela se traduit aussi dans la matrice de transition estimée \hat{A} du modèle correspondant $\lambda_{3,3}$ où l'on trouve la probabilité de transition de rester dans cet état particulier $i\mathbf{C} = ["+1" "+1" "+1" "+1" "+1"]^T$:

- pour le MMC-P : $a_{i\mathbf{C} \rightarrow i\mathbf{C}} = 0,978$

- pour le MMC-C : $a_{i\mathbf{C} \rightarrow i\mathbf{C}} = \prod_{k=1}^5 a_{i\mathbf{C} \rightarrow "+1^{(k)}"}^{(k)} = 0,979$

Pour ce sujet, le nombre moyen de transitions sur toutes les SAM et toutes les sessions étudiées est

de 45 ± 13 ce qui, au regard du nombre de points considérés ($T = 387$ soit environ 20 minutes) et du nombre d'états définis ($2^5 = 32$), reflète une évolution plutôt lente des interactions entre les SAM. Cette transcription de la connectivité fonctionnelle en états peut permettre par exemple de retrouver des suites d'états reproduites durant une session voire à travers les sessions. La FIGURE 11.4 présente par exemple des suites d'états qui se sont manifestées sur plusieurs sessions et parfois plusieurs fois dans un même session.

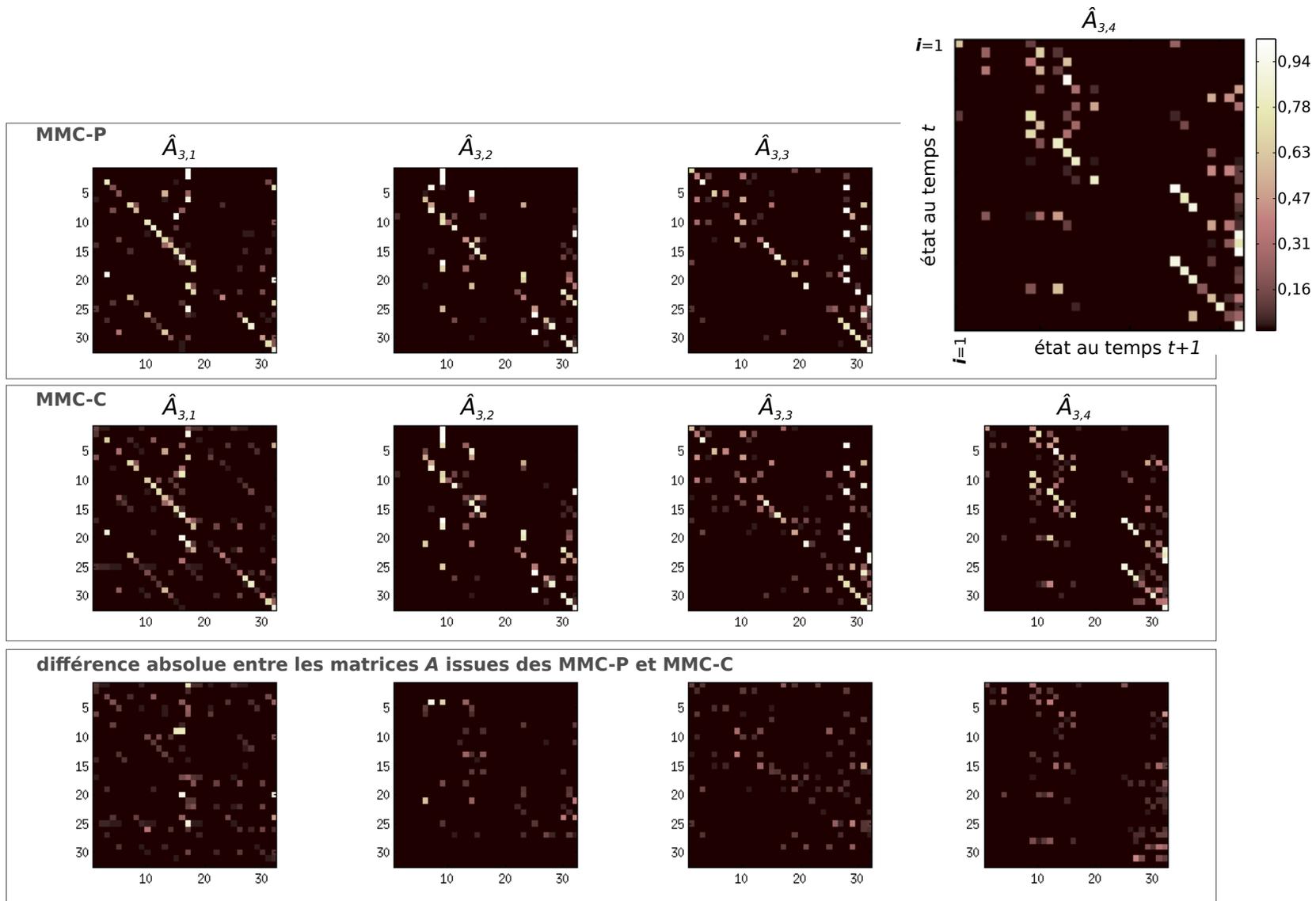


FIGURE 11.1 – Matrices de transitions pour le sujet M7, SAM 1 (réseau OPFN) pour les 4 sessions s (en colonne). Lignes de haut en bas : matrices issues de la modélisation et estimation par MMC-P; matrices issues de la modélisation et estimation par MMC-C; différences absolues, session par session, entre les matrices apprises par MMC-P et celles apprises par MMC-C. L'échelle de couleur et le sens de lecture des matrices sont donnés pour la matrice en haut à droite. Les probabilités les plus fortes sont de couleurs claires. Les lignes de chaque matrice dénotent les états de départ au temps t , les colonnes les états d'arrivée au temps $t + 1$ (voir texte).

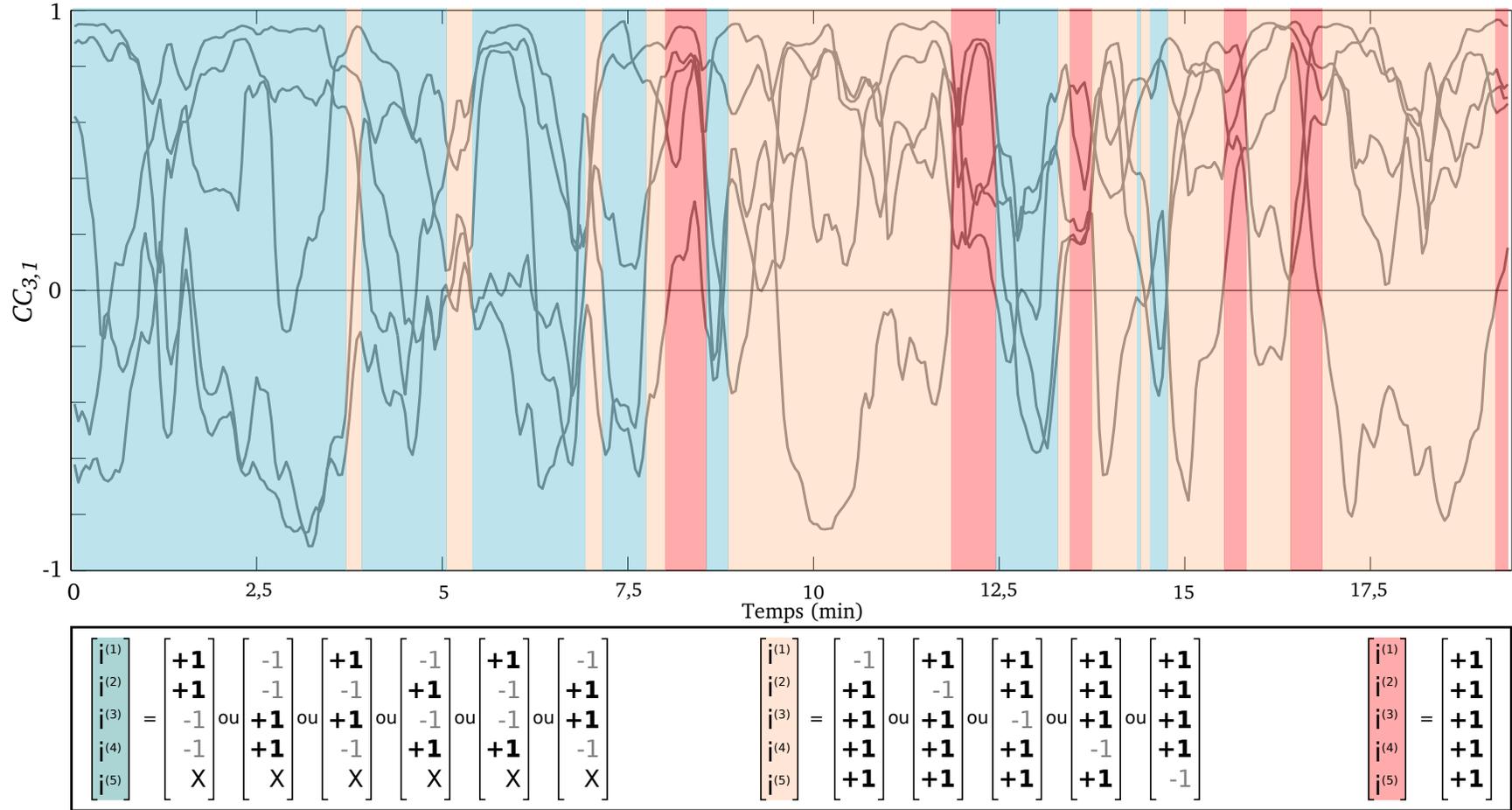


FIGURE 11.2 – $CC_{3,1}$ du sujet M7 (SAM 3 réseau OPN, session 1) et répartition de plusieurs états décodés par le MMC-P $\lambda_{3,1}$. Les états rouges correspondent à un état où toutes les SAM sont corrélées à la SAM 3 (rouge foncé) ou 4 SAMs sur 5 sont corrélées à la SAM 3 (orange). Le bleu représente ici les états les plus différenciés (2 SAMs corrélées, 2 SAMs anti-corrélées, 1 corrélée ou anti-corrélée par rapport à la SAM 3)

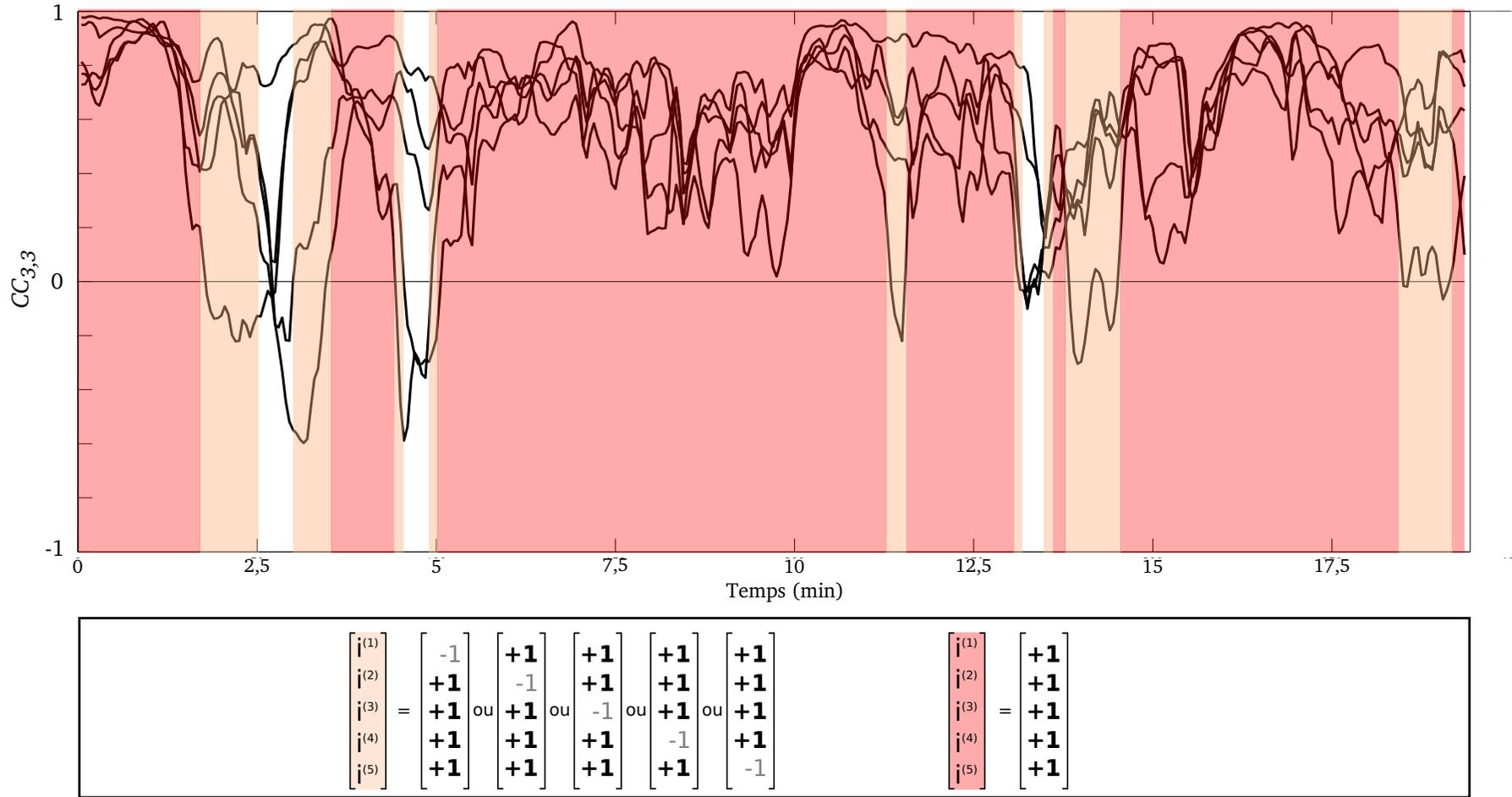


FIGURE 11.3 – $CC_{3,3}$ du sujet M7 (SAM 3, session 3) et répartition de 4 états décodés par le MMC-P $\lambda_{3,3}$. Les états rouges correspondent à un état où toutes les SAM sont corrélées à la SAM 3 (rouge foncé) ou 4 SAMs sur 5 sont corrélées à la SAM 3 (orange).

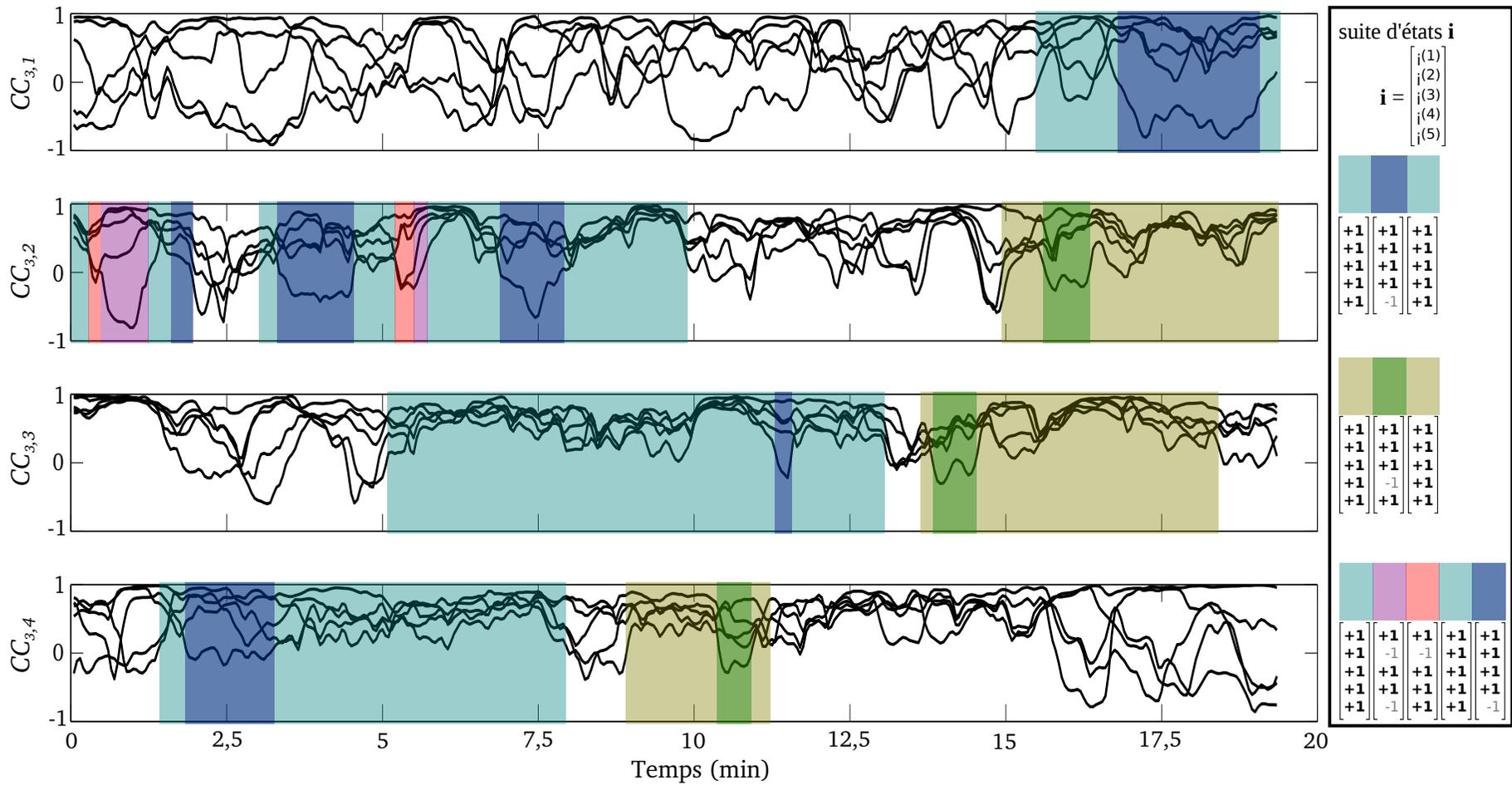


FIGURE 11.4 – $CC_{3,s}$ du sujet M7 (SAM 3) pour les 4 sessions (en ligne) et mise en évidence de 3 suites d'états décodés par le MMC-P $\lambda_{3,s}$ et reproduites au cours d'une session ou dans les différentes sessions. Chaque couleur correspond à un état décodé.

11.1.4 Discussion

L'analyse inter-sessions de la DCF avec entre des sessions avec prise de médicament ou non a permis de mettre en évidence une modification dans l'évolution des interactions entre les SAM pour cinq sujets sur huit ayant pris un médicament. En revanche, de tels changements dans les propriétés dynamiques de la connectivité fonctionnelle n'ont jamais pu être observés chez les sujets ayant pris un placebo lors de la dernière session. Il est important de noter qu'une modification de la DCF ne signifie pas forcément une augmentation ou une diminution des transitions entre états. Dans la notion de DCF, le mot « dynamique » reflète l'évolution et la modification de la connectivité fonctionnelle au cours du temps. Il n'est pas question ici de qualifier un signal de plus ou moins « dynamique ». Une modification de la DCF est perçue par nos modèles comme une modification plus large et à plusieurs niveaux : changements du type d'interactions (lien entre les réseaux), changements des états visités, changements du type ou nombre de transitions. Ainsi, l'étude de l'évolution du nombre de transitions au cours des quatre sessions a révélé une augmentation notable chez les sujets M2, M3, M4, M8. Le changement de DCF observé chez le sujet M6 en session 4 grâce à la mesure de la distance, ne semble pas être expliqué au niveau du nombre de transitions. Le méthylphénidate étant un stimulant du système nerveux central, il était attendu qu'il provoque des états plus différenciés, c'est-à-dire plus de décorrélations ou d'anti-corrélations. De plus, nos observations préliminaires (paragraphe 9.3.2) nous avaient déjà laissé présager une modification dans ce sens pour certains sujets ayant pris un médicament. Pour rappel, les réseaux surviennent généralement en bouffée et sont donc fortement corrélés dans le temps [126]. Cela avait été remarqué dans le chapitre précédent pour la plupart des sujets (voir les résultats préliminaires paragraphe 9.3.1). Les variations observées en session 4 pour les sujets M2, M3, M4, M6 et M8 sont donc conformes aux attentes. Ce qui n'avait pu être montré directement à partir des corrélations, est rendu possible grâce à l'approche basée sur les MMC multidimensionnels. Pour M1, le sujet présentait déjà des corrélations entre SAM fluctuantes (histogramme des valeurs de corrélations centré sur 0) dès les trois premières sessions. Il n'a peut-être pas été réceptif au médicament, ayant déjà des fortes fluctuations de la connectivité fonctionnelle. M5 et M7 n'ont pas répondu au médicament non plus peut-être pour cause de dosage insuffisant ou pour des raisons qui leur sont spécifiques ; un médicament pouvant présenter des effets variables selon la sensibilité du sujet.

Les données étudiées pour cette application sont issues d'une étude pharmacologique plus large pour laquelle les sujets ont rempli des grilles d'auto-évaluation pour jauger, avant et après chaque session, leur état vis-à-vis de différents critères (éveillé, calme, fort, vaseux, adroit, mou, content, inquiet, lent, tendu, attentif, incapable, heureux, hostile, intéressé, renfermé). Ces scores d'auto-évaluation ne sont pas présentés car ils s'éloignent trop du cadre de notre travail mais nous avons pu néanmoins y avoir accès. Par exemple, pour chacune des quatre sessions, la différence entre les scores avant et après l'acquisition des données IRM peut être calculée. L'évolution de cette différence au travers des quatre sessions donne un indice sur la façon dont le sujet a ressenti l'effet du médicament ou du placebo au cours de la dernière session comparativement à une session où il n'avait rien pris. En se focalisant sur le critère « éveillé », nous remarquons une nette augmentation en session 4 de la différence entre le score avant et le score après par rapport à cette même différence pour les trois autres sessions. Cela a été remarqué chez tous les sujets (M_i comme P_i) excepté pour M4 et M6. L'effet placebo a donc été important ce qui n'a donc pas permis de distinguer un effet du médicament en se basant uniquement sur les retours du sujet. Pourtant, l'effet sur la dynamique de la connectivité fonctionnelle n'est visible que sur certains sujets médicamentés.

Les résultats obtenus suggèrent le potentiel de notre méthode à mettre en évidence les sujets réceptifs ou non à ce médicament. Cela doit être encore confirmé mais cette application laisse fortement suggérer l'intérêt d'une telle approche pour une application clinique ou une étude pharmacologique afin de connaître *a priori* quels sujets pourraient présenter un bénéfice à l'introduction du traitement.

L'analyse de la DCF pour cette application est concluante mais ce n'était pas le seul objec-

tif. En effet, en complément de l'étude inter-sessions, notre méthode fournit également des outils permettant de caractériser certaines propriétés dynamiques de la connectivité fonctionnelle chez les sujets sains. Ainsi, l'étude des matrices de transitions et des décodages de Viterbi ont mis en évidence une forte probabilité de rester dans un état, associée à nombre relativement faible d'états différents observés par session par rapport au nombre de données. Le cerveau semble rester dans la même configuration cérébrale plusieurs dizaines de secondes voire minutes durant cette condition de repos. En outre, il a été observé chez les sujets que les états les plus atteints étaient ceux où les SAMs étaient les plus corrélés entre elles. L'exemple donné avec le sujet M7 permet de visualiser la répartition de certains états au cours du temps. Ainsi, ces états de forte connectivité (l'état rouge et les états oranges) peuvent être interprétés comme des états où le sujet est moins éveillé puisque que les réseaux ne sont pas différenciés et forment un supra-réseau (regroupement de plusieurs réseaux) peu spécifique pour une tâche. Apparaissant sur toute la durée de la session ou en deuxième partie de session pour ce sujet, ces états de connectivité peuvent résulter d'une diminution de l'éveil du sujet due peut-être à la durée importante de la session (20 minutes) ou à l'habitué d'une expérience d'IRMF en condition de repos. De telles interprétations avaient été proposées par Allen et al. [119] mais sur une session unique de durée bien inférieure à celle à notre disposition (5 minutes contre 20 minutes dans notre cas). En outre, l'observation via le décodage de Viterbi de suites d'états qui se répètent dans et entre les sessions laisse suggérer des états privilégiés par le sujet et une connectivité fonctionnelle de repos plus structurée qu'aléatoire.

Du point de vue de la modélisation, les approches par MMC-P ou MMC-C permettent toutes deux de distinguer la quatrième session pour les sujets M2, M3, M4, M6 et M8. Les résultats en terme de décodage ou probabilité de transitions sont d'ailleurs très proches, ce qui permet au calcul de la distance de mener à la même conclusion. Cependant les légères variations observées pour les probabilités de transitions suggèrent que la modélisation est plus riche avec le MMC-P qu'avec le MMC-C, lequel résulte de simplifications (liens de dépendance conditionnelle) du MMC-P.

La mesure de distance entre modèles a été réalisée pour des observations générées avec différentes durées. Aucune différence n'a été observée à ce sujet, excepté pour le sujet M7 qui a montré un résultat positif au test avec le MMC-C et $T=400$. Cependant, la p -valeur obtenue était à la limite du seuil choisi et aucune des autres générations de signaux ou modélisation n'a pu confirmer cette variation pour ce sujet. Il est donc pertinent d'effectuer plusieurs générations de différentes durées afin de s'assurer que les résultats du test de Student ne soient pas le fait d'un seul jeu de données générées. Une étude plus approfondie sur ce point sera d'ailleurs effectuée pour l'application qui suit.

11.2 Application 2 basée groupe : « Démence à corps de Lewy »

11.2.1 Objectifs et contraintes

Cette seconde application d'analyse de la DCF par MMC-P et MMC-C a été réalisée sur les bases de données B4 et B5 (description disponible dans au paragraphe 4.2.2). Pour cette étude, nous disposons de deux groupes de sujets : 20 sujets sains pour le premier groupe (base B4) et 16 sujets atteints de démence à corps de Lewy (DLB pour *Dementia with Lewy Bodies*) pour le second (base B5). Pour rappel, la démence à corps de Lewy est la deuxième forme la plus répandue de démence neurodégénérative après la maladie d'Alzheimer (voir paragraphe 4.2.2). Les principaux critères cliniques de la DLB sont des troubles cognitifs associés à des fluctuations cognitives, des syndromes parkinsoniens et des hallucinations visuelles. Il est intéressant de remarquer que la notion de fluctuations cognitives englobent de nombreuses situations puisqu'elle a été évaluée pour cette étude par les réponses du patient à un même jeu de tests neuropsychologiques ciblant différentes fonctions cognitives, répété à intervalles de temps variable (à l'échelle de la semaine ou du mois). Les fluctuations de la vigilance pouvant être sur des échelles beaucoup plus petites (minutes, secondes),

les fluctuations cognitives évaluées entre deux sessions de tests neuropsychologiques peuvent aussi inclure ces troubles de l'attention. Pour mieux comprendre les anomalies cérébrales dans la DLB, de nombreuses études ont été réalisées par IRM, mais très peu se sont intéressées à la connectivité fonctionnelle de repos et, le cas échéant, seulement en « statique » [143, 144]. L'étude des variations des interactions entre les réseaux est donc entièrement nouvelle pour cette pathologie et semble particulièrement pertinente au vu des symptômes : les fluctuations cognitives et les hallucinations présentes chez certains patients laissent suggérer que les anomalies cérébrales fonctionnelles pourraient être à caractère passager et *de facto* seraient de « bonnes candidates » pour des approches basées sur la DCF.

Dans ce contexte, nous souhaiterions voir s'il est possible comme nous l'avons déjà fait dans l'application 1, de capturer dans les modèles MMC-P ou MMC-C de DCF, l'information permettant de différencier les patients atteints de DLB des sujets sains. Dans un second temps, en s'appuyant sur les outils d'analyse offerts par ces approches de modélisation, nous tenterons d'aller plus loin dans la caractérisation de la DLB.

11.2.2 Mise en œuvre

Cette application se faisant au niveau du groupe, certains pré-traitements doivent être réalisés sur l'ensemble des sujets, en plus de ceux déjà effectués au niveau individuel (détaillés paragraphe 1.3.3.2). Il s'agit d'un recalage des images fonctionnelles sur les images morphologiques, d'une normalisation des images dans l'espace standard MNI et d'un lissage spatiale (les raisons de ces traitements ont été détaillées paragraphe 1.3.3.2).

L'ACI est réalisée au niveau du groupe [45] pour extraire les SAM communes à tous les sujets (patients et sujets sains inclus). Les composantes ont été produites à l'aide de l'algorithme ICASSO qui s'assure de la fiabilité des réseaux trouvés (voir paragraphe 2.2.1.2 page 22 et paragraphe 2.2.3.3 page 25). Les SAM étant partagées par les sujets de l'étude, il s'agit donc de RSN (réseaux connus et détaillés dans la littérature, partagés par plusieurs sujets, voir paragraphe 1.4.2 page 14) et nous utiliserons cette terminologie pour parler des réseaux étudiés dans cette étude. Parmi les 30 composantes en sortie de l'ACI, 10 ont pu être identifiées comme des RSN. La reconstruction arrière (voir paragraphe 2.2.1.2 page 22) nous permet d'obtenir la carte spatiale et son décours temporel associé pour chaque RSN et chaque sujet.

Compte tenu du nombre de volumes acquis par sujet (120 volumes pour 6 minutes d'acquisition) et des considérations en relation avec le TABLEAU 11.1, nous nous sommes concentrés sur sept RSN spécifiques présents durant la session. Ils ont été choisis par un neurologue et un neurobiologiste en fonction de leur pertinence par rapport à la maladie :

- le RFPN et le LFPN pour prendre en compte les potentiels troubles dysexécutifs et attentionnels [145, 146, 147, 148] ;
- l'OPFN, le MON et l'OPN pour prendre en compte les déficiences visuelles [147, 148] ;
- le BG, impliqué parfois dans les troubles moteurs chez les patients atteints de parkinsonisme [80].

Le DMN a lui aussi été inclus dans l'étude étant donné qu'il reste un réseau majeur dans les études de connectivité fonctionnelle de repos. Les trois réseaux restants et non sélectionnés pour l'étude de la DCF étaient un réseau frontal et deux réseaux centraux. Pour rappel, le TABLEAU 1.1 présenté en Partie I page 15 décrit succinctement les RSN. Les cartes spatiales des RSN étudiés sont présentées en annexe, FIGURE D.1.

La même fenêtre de corrélation (51 s) est choisie pour cette étude, donnant lieu à des signaux comportant 103 échantillons temporels. Ainsi, pour cette étude, $N = 7$, $K = 6$, $T = 103$ et $Ses = 1$. L'indice s relatif à la session n'a plus lieu d'être dans cette étude et sera donc omis pour le reste des notations ($CC_n \equiv CC_{n,1}$, $\lambda_n \equiv CC_{n,1}$).

Pour chacun des 36 sujets de cette base (20 sujets sains, 16 patients), leurs modèles λ_n , $1 \leq n \leq 7$, sont initialisés comme décrit dans la méthode section 10.2. De la même manière que pour

l'application précédente, ces modèles λ_n ont été comparés deux à deux entre les sujets, avec la mesure de distance déjà présentée (voir paragraphe 10.5.1), afin d'évaluer les modifications de la connectivité fonctionnelle entre les patients et les sujets sains. Pour chacun des 7 réseaux, deux jeux de distances ont été créés : D_{c-c} rassemblant les distances entre toutes les paires possibles entre sujets sains (c pour contrôle), soit $20 \times 19/2 = 190$ paires, et D_{c-p} pour les distances entre les paires comprenant un sujet sain et un patient, soit $16 \times 20 = 320$ paires. Pour le calcul de ces distances, la longueur des séquences générées T_{gen} (voir paragraphe 10.5.1) a été choisie pour se rapprocher des séquences initiales, d'où $T_{gen} = 150$. L'approche a aussi été testée pour $T_{gen} = 300$.

Un test de Student à deux échantillons (*two samples t-test*) est réalisé entre ces deux jeux de distances D_{c-c} et D_{c-p} pour chacun des 7 réseaux. Un test positif ($p < 0,05$) permet de révéler un changement significatif des propriétés dynamiques des signaux de connectivité fonctionnelle entre les sujets sains et les patients atteints de DLB et, concernant les interactions d'un réseau par rapport aux 6 autres étudiés.

11.2.3 Analyses complémentaires

Afin d'aller plus loin dans l'interprétation de nos résultats, des analyses complémentaires ont été réalisées.

Multiple générations de séquences : La mesure de la distance entre deux modèles telle que proposée par Rabiner et Juang (équation (10.11)) demande la génération de signaux basée sur les modèles appris. Ces nouveaux signaux sont créés à partir des probabilités estimées (probabilités de transitions et d'émission des observations). La répétition de cette étape de génération mène ainsi à de multiples signaux non identiques rendant la mesure de la distance entre deux modèles sensible à ces signaux générés. En générant pour chaque modèle λ_n , M séquences d'observation, nous pouvons calculer à nouveau les distances sur ces séquences synthétisées de manière à obtenir M jeux de distances D_{c-c} et D_{c-p} . Cela nous permet de vérifier la reproductibilité des p -valeurs, et donc des décisions quant aux RSN. Pour des raisons de temps de calcul, nous avons généré $M = 20$ signaux différents et donc 20 fois les distances qui suivent pour $T_{gen} = 150$ ainsi que pour $T_{gen} = 300$

Durée des séquences générées : Une étude a également été réalisée pour juger l'impact de la durée de la séquence générée sur la mesure de la distance entre deux modèles. Dans ce but, nous avons fait varier le nombre T_{gen} d'observations générées par séquence, avec $T_{gen} = 150, 300, 600, 900, 1200, 1500$ et 2000 . Le cas considéré est celui des distances entre des paires patient-sujet sain pour un type de modélisation (le MMC-P) et pour un réseau (le MON). Ces choix arbitraires représentent néanmoins 320 mesures de distances. Ils ont pour objectif d'apporter un premier éclairage sur l'influence du choix de la durée des signaux générés.

Choix des patients : Dans cette seconde application, certains patients sont sujets à des fluctuations cognitives. Nous pensons qu'il est intéressant de focaliser l'analyse sur le comportement de ces patients en particulier. En effet, peu d'informations sont connues sur l'échelle temporelle de ces fluctuations et donc si cela peut se traduire par une modification de la DCF pour ces patients en particulier par rapport aux patients ne présentant pas de fluctuations cognitives. Des comparaisons supplémentaires sont donc effectuées entre les sujets sains et les patients avec fluctuations ainsi qu'entre les sujets sains et les patients ne présentant pas de fluctuations cognitives.

11.2.4 Résultats

11.2.4.1 Comparaison des DCF entre groupe « contrôle » et groupe « patients DLB »

Le TABLEAU 11.4 présente les résultats des t -tests pour chacun des 7 RSN étudiés. Le réseau occipito-pariétal (OPN), le réseau occipital médian (MON) ainsi que le réseau fronto-pariétal gauche (LFPN) affichent des différences entre les DCF des sujets sains et des patients particulièrement significatives, quel que soit le modèle utilisé pour décrire les réseaux ou les différentes générations du signal utilisées (de 150 ou 300 points).

TABLEAU 11.4 – Résultats des comparaisons entre groupe pour chaque RSN étudiés, pour les différents MMC modélisés et différents T_{gen} .

RSN	T_{gen}	MMC-P			MMC-C		
		150	300	h	150	300	h
OPN		<0,000	<0,000	1	0,005	0,03	1
MON		0,001	0,001	1	0,02	0,001	1
RFPN		0,74	0,07	0	0,93	0,55	0
LFPN		<0,000	<0,000	1	0,03	<0,000	1
OPFN		0,44	0,08	0	0,03	0,26	0
DMN		0,38	0,76	0	0,95	0,30	0
BG		0,45	0,71	0	0,20	0,10	0

Les p -valeurs significatives $<0,05$ (en gras) permettent de décider que la DCF du réseau considéré est significativement différente entre le groupe des sujets sains et celui des patients ($h=1$).

11.2.4.2 Caractérisation de la DCF : matrices de transition, décodages de Viterbi

Les résultats précédents ont mis en avant le rôle des réseaux visuels et attentionnels dans la différenciation de la DCF du groupe patients DLB par rapport au groupe contrôle. Dans cette section, on se propose d'apporter des informations complémentaires sur ces réseaux à l'aide des outils d'analyse de la DCF que sont les matrices de transitions et les décodages de Viterbi.

Matrices de transition Un examen attentif des matrices de transition des réseaux visuels et attentionnels montre que les états ayant la plus forte probabilité d'être atteints sont ceux avec l'OPN, le OPFN, et/ou le MON dans l'état de corrélation "+1". Une comparaison des matrices de transitions entre les deux groupes de sujets ont mis en avant que les vecteurs d'états avec le MON et le RFPN dans un état corrélé ont été plus probablement atteints par les sujets sains que par les patients ($p = 0,044 < 0,05$). La FIGURE 11.5 illustre les matrices de transitions pour un sujet (patient 2), pour le réseau MON avec les deux modélisations présentées.

Séquences d'états décodées L'étude des séquences d'états décodées \hat{X}_{OPFN} , \hat{X}_{MON} et \hat{X}_{RFPN} ne révèle aucune différence concernant les changements d'états (nombre de transitions, répartition des transitions au cours de la séquence, durée dans les états), ni entre les groupes ni à l'intérieur des groupes. Vingt-deux transitions et 18 états distincts ont été observés en moyenne pour chaque session de 6 minutes. Seulement 3 à 4 états en moyenne par sujet sont à la fois longs (durent plus

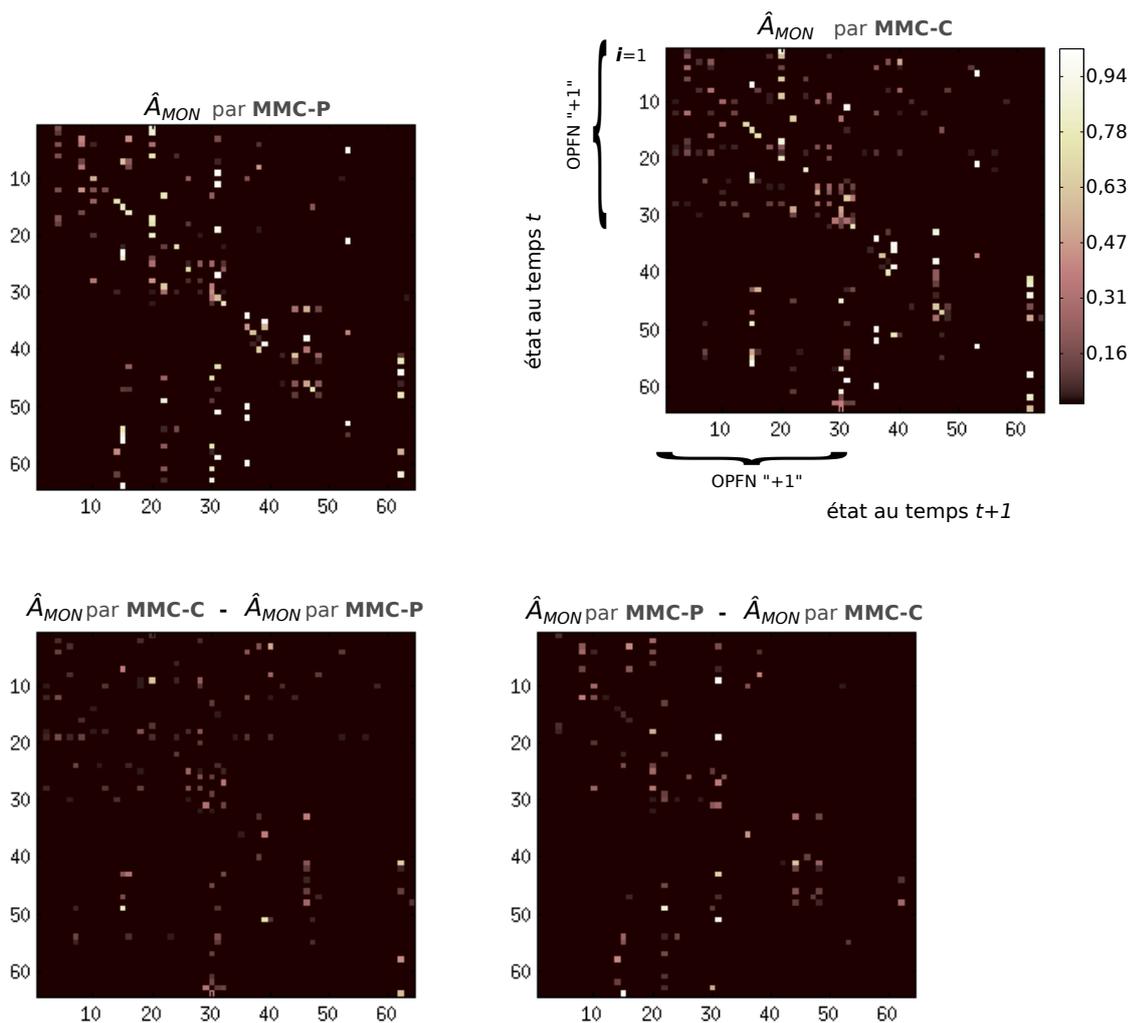


FIGURE 11.5 – Matrice de transition pour le patient 2, le réseau MON avec MMC-P et MMC-C. Les matrices du bas représentent les différences entre les matrices issues des deux types de modélisation. Les états de départ et ceux d’arrivée sont représentés de manière synthétique allant de 1 à 64 ($=2^6$). Les états notés OPFN “+1” correspondent aux 32 états dans lequel le OPFN est corrélé avec le MON.

d’un point temporel de 3 secondes) et répliqués au cours de la session à des moments différents. Ces états reproduits variaient entre les sujets, même si nous avons remarqué que l’état le plus visité (long et répliqué pendant la session) pour \hat{X}_{OPN} et \hat{X}_{MON} a été le même pour les patients et les contrôles, à savoir l’état [“+1” “-1” “+1” “+1” “+1” “-1”] c’est-à-dire l’état {OPN/MON -rfpn-LFPN-OPFN-DMN-bg }, (avec des états « corrélés » en lettres majuscules et états « anti-corrélés » en minuscules par rapport au réseau de référence, l’OPN ou le MON). La moyenne de la durée d’occupation d’un état est de 14 secondes pour les patients comme pour les contrôles.

Analyse combinée des résultats Précédemment, une relation entre le réseau visuel MON et le réseau attentionnel RFPN a d’abord été soulignée au travers des matrices de transitions. L’analyse qui suit montre comment la combinaison des résultats issus des MMC-P ou des MMC-C permet de fournir des informations supplémentaires sur certaines interactions spécifiques entre RSN. En se restreignant aux interactions entre le MON et les réseaux attentionnels (RFPN, LFPN) nous pouvons montrer que les patients se trouvent bien plus souvent dans un état où le LFPN est corrélé avec le MON et le RFPN décorrélé avec le MON par rapport aux sujets sains, quel que soit l’état des

autres réseaux ($p = 0,021 < 0,05$, test non-paramétrique de Kruskal-Wallis). Le taux d'occupation moyen dans une séquence par ces états était de 40 % pour les patients contre 22 % pour les sujets sains. En outre, les résultats de la matrice de transition ont révélé que les états impliquant une corrélation entre le MON et le RFPN, quels que soient les autres réseaux, sont beaucoup moins susceptibles d'être atteints chez les patients que chez les témoins.

11.2.4.3 Résultats des analyses complémentaires

Les résultats issus des multiples générations de signaux sont présentés FIGURE 11.5. Les deux groupes sont systématiquement différenciés avec le MMC-P comme avec le MMC-C, pour différentes tailles de signaux générés. En moyenne, ce sont bien les 3 réseaux soulignés (OPN, MON et LFPN) qui ressortent de l'étude. Dans environ 40 % des cas, on retrouve un quatrième réseau (le RFPN) qui se démarque entre les deux groupes de sujets.

TABLEAU 11.5 – Résultats positifs au t-test pour les 20 générations de signaux modélisés par le MMC-P et le MMC-C, pour $T_{gen} = 150$ et $T_{gen} = 300$.

	RSN	OPN	MON	RFPN	LFPN	OPFN	DMN	BG
MMC-P	150	12/20	15/20	9/20	16/20	6/20	5/20	9/20
	300	15/20	16/20	3/20	12/20	3/20	5/20	4/20
Total		68 %	80 %	30 %	70 %	23 %	20 %	33 %
MMC-C	150	14/20	17/20	2/20	17/20	6/20	7/20	2/20
	300	15/20	18/20	5/20	16/20	8/20	7/20	6/20
Total		73 %	88 %	18 %	83 %	35 %	35 %	20 %

Une étude sur la durée de la séquence générée ($T_{gen} \in [150, 300, 600, 900, 1200, 1500, 2000]$) a montré une faible variation des scores de distances selon le T_{gen} choisi. Les mêmes RSN (l'OPN, le MON et le LFPN) se démarquent systématiquement.

Un dernier point inspecté lors de cette étude est l'impact des symptômes des patients par rapport à l'effet observé pour l'ensemble des patients. En scindant les patients entre ceux manifestant des fluctuations cognitives et ceux qui n'en présentent pas, il nous a été possible de souligner l'effet de ces fluctuations sur la DCF. Le TABLEAU 11.6 présente les résultats de cette étude où l'on remarque que ce sont les patients avec des fluctuations qui ont le plus de différences avec les sujets sains du point de vue de la DCF.

11.2.5 Discussion

Comme nous l'avons précisé lors du choix des réseaux, le MON (aires primaires) et l'OPN (aires secondaires et associatives) concernent le traitement visuel [149] et le LFPN concerne le contrôle exécutif de l'attention [150, 151]. Les patients atteints de DLB souffrant principalement de déficiences attentionnelles [145, 146, 147, 148] et visuelles [147, 148], les changements de DCF de ces trois réseaux sont donc cohérents avec le profil cognitif des patients atteints de DLB. De précédentes études avaient également souligné une diminution de l'activité neuronale dans le cortex occipital et les zones frontales et pariétales via des mesures du métabolisme [152] et de perfusion cérébrale en IRM [153, 154]. Bien que le cortex occipital soit souvent désigné comme étant perturbé dans la DLB, la connectivité fonctionnelle statique [143, 144] n'a révélé aucune différence, contrairement à notre approche en dynamique. L'avantage de cette approche par réseaux permet non seulement

TABLEAU 11.6 – Résultats spécifiques aux patients présentant des fluctuations cognitives par rapport aux sujets sains (lignes F) et résultats spécifiques aux patients ne présentant pas de fluctuations cognitives par rapport aux sujets sains (lignes non-F), pour $T_{gen} = 150$ et $T_{gen} = 300$ et pour les modélisations MMC-P et MMC-C. En gras les p -valeurs en-dessous du seuil 0,05.

		RSN	OPN	MON	RFPN	LFPN	OPFN	DMN	BG
MMC-P (p)	150	F	< 0,000	< 0,000	0,65	< 0,000	0,52	0,33	0,09
		non F	0,66	0,32	0,26	0,02	0,66	0,63	0,50
	300	F	< 0,000	< 0,000	0,80	< 0,000	0,01	0,66	0,26
		non F	0,48	0,38	0,005	0,01	0,89	0,28	0,52
MMC-C (p)	150	F	< 0,000	0,09	0,08	0,03	0,004	0,40	0,15
		non F	0,50	0,02	0,02	0,23	0,54	0,41	0,53
	300	F	< 0,000	0,001	0,19	< 0,000	0,27	0,13	0,15
		non F	0,58	0,02	0,007	0,04	0,42	0,91	0,21

de différencier les groupes mais également de cibler les réseaux qui participent le plus à cette différence.

Ces résultats se retrouvent si l'on réitère l'étape de génération de signal nécessaire à la comparaison de modèles. Pour les 80 répétitions du test (20 fois par modèle et par T_{gen}), le groupe des patients se différencie systématiquement. Les réseaux impliqués les plus fréquemment dans cette différenciation sont ceux que nous avons décrits (MON, OPN et LFPN), ils se détachent largement. Pourtant, il est intéressant de remarquer l'apparition de quelques uns des autres réseaux lors de certains tests. Nous remarquons ainsi que le RFPN et l'OPFN semblent aussi se détacher légèrement suggérant qu'ils participent également à la différenciation observée entre les DCF des sujets sains et des patients. Cela n'est pas surprenant puisque le RFPN est un réseau attentionnel liés au contrôle exécutif (au même titre que le LFPN) et que l'OPFN est un réseau visuel (comme le MON et l'OPN), deux types de fonctions dont nous avons montré les altérations en DCF. Le DMN, par ailleurs, montre aussi une tendance à se différencier chez les patients. Ce réseau est souvent décrit dans la littérature comme étant anti-corrélé aux réseaux attentionnels. Notre étude s'appuyant sur la connectivité fonctionnelle, et non l'activité, il semble ainsi logique que ce réseau soit aussi impacté par une modification de la connectivité. Cependant l'altération du DMN dans la DCF des patients atteints de DLB est vraisemblablement plus une conséquence due à l'altération des réseaux attentionnels qu'une des causes des modifications de la DCF. Nos autres analyses complémentaires vont également dans le sens d'une différence entre les sujets sains et les patients due à la DCF des réseaux visuels et attentionnels. Ainsi le calcul de la distance pour différentes durées de génération de séquence produit des résultats dans la suite de ceux déjà décrits. Notre choix d'un T_{gen} proche de la durée initiale n'est donc pas remis en cause. La multiplicité de la synthèse de signaux pour le calcul de la distance est donc intéressante si l'on souhaite connaître les poids de chaque réseaux dans la différence trouvée entre les deux groupes. Le fait de cibler des différences au niveau des réseaux (et non juste entre deux groupes ou deux sessions) amènent logiquement plus de variations car moins de données sont prises en compte dans les jeux de distances. Néanmoins peu de variations sont observées concernant les réseaux les plus altérés ce qui nous permet d'appuyer nos conclusions concernant le rôle des réseaux visuels et attentionnels dans les études de DCF chez les patients atteints de DLB.

L'étude de statistiques simples ne nous a pas permis de remarquer à quel niveau (nombre de

transitions, durée dans les états) se situait la différence entre les sujets sains et les patients. L'hétérogénéité des symptômes des patients souffrant de DLB joue probablement un rôle dans la difficulté à trouver une caractéristique commune. De plus, les sujets sains, comme les patients, étaient relativement âgés (moyenne à 70 ans tous sujets confondus, voir TABLEAU 4.1 page 38) ce qui peut soulever la question de l'influence de l'âge sur la variabilité de la connectivité fonctionnelle. L'analyse plus ciblée de certaines transitions, ou états permettent néanmoins quelques remarques. Analysés conjointement, ces résultats obtenus via la matrice de transition ou le décodage de Viterbi, suggèrent un manque d'interaction spécifiquement entre le réseau droit attentionnel et les réseaux visuels. Dans la démence à corps de Lewy, la DCF laisse à penser que les réseaux visuels et en particulier le MON pourraient moins interagir avec le RFPN qu'avec le LFPN. Un tel résultat pourrait refléter un phénomène compensatoire par rapport au déficit observé des interactions entre le réseau attentionnel droit et les réseaux visuels. Il est intéressant de rappeler que la méthode, bien qu'appliquée à l'échelle du groupe, peut être personnalisée à des fins de caractérisation ou surveillance individuelle.

Sans aller jusqu'à ce niveau individuel, nous avons tout de même cherché à scinder le groupe de patients pour dissocier ceux ayant des fluctuations cognitives des autres patients. Notre analyse sur ces patients montre que ce sont principalement les patients avec des fluctuations cognitives qui présentent le plus de différences en terme de nombre de réseaux altérés en DCF. Ils présentent également des p -valeurs bien plus faibles pour les réseaux qui sont sous le seuil choisi. Les autres patients présentent toutefois également des différences de DCF par rapport aux sujets sains et montre bien que seules les fluctuations cognitives ne sont pas à l'origine d'une modification de la DCF chez les patients atteints de DLB. Cette étude est particulièrement intéressante d'un point de vue médical du fait que l'échelle de temps des fluctuations cognitives décrites par les médecins ou les patients n'est pas bien établie. Ainsi, le patient n'est pas toujours en mesure de se souvenir de fluctuations de vigilance se manifestant sur des temps très courts et ce phénomène n'est pas toujours décrit lors de la rencontre avec le médecin ou le neuropsychologue. Le fait de mettre en avant des différences plus prononcées pour ces sujets laisse suggérer que les fluctuations peuvent apparaître sur des temps courts (inférieurs à la durée de la session d'IRMf de 5 minutes dans notre cas). Des analyses plus spécifiques sur les nombre de transitions ou la durée dans les états pour ces sujets-ci par rapport aux sujets sans fluctuations et aux sujets sains pourraient nous aider à mieux caractériser ce symptôme dans cette pathologie et aider à mieux comprendre les troubles qui en résultent.

11.3 Bilan

Que ce soit pour l'application 1 à l'échelle du sujet pour comparer des sessions, ou pour l'application 2 à l'échelle du groupe pour comparer deux populations, la méthode développée a montré des résultats prometteurs quant à sa capacité à mettre en évidence des changements de DCF. La prise d'un médicament stimulant le système nerveux central ou une pathologie affectant l'attention modifient la dynamique globale de la connectivité fonctionnelle cérébrale. Nos résultats permettent également une première caractérisation de la DCF en condition de repos. Pour notre étude de la base B2 à l'échelle du sujet comme pour celle des bases B4 et B5 à l'échelle du groupe, les changements d'états semblent peu nombreux et la connectivité fonctionnelle au repos fluctue relativement lentement (un état pouvant durer jusqu'à des dizaines de secondes voire minutes). Les SAM restent majoritairement toutes corrélées entre elles au cours des sessions de repos. Cela reste cohérent avec les fréquences des signaux BOLD au repos ($[0.004; 0.08 \text{ Hz}]$) ainsi qu'avec les premières études de DCF remarquant l'apparition conjointe des réseaux [126] et s'intéressant au nombre de transitions [119]. Bien que la durée de la session peut sembler courte pour conclure sur le caractère de la DCF entre nos groupes dans notre deuxième étude, les multiples analyses complémentaires réalisées permettent de conforter les choix méthodologiques effectués dans la modélisation et l'analyse

de la DCF.

Nous avons souhaité, dans une première tentative de modélisation markovienne cachée multidimensionnelle de la DCF, développer une approche simple. D'une part, l'espace d'états par canal $S^{(k)}$ autorise deux états : « corrélé » et « anti-corrélé » ce qui restreint le cardinal de l'espace d'états S . D'autre part, les lois d'observation n'étaient pas ré-estimées car considérées comme suffisamment proches des données pour une première modélisation. Pourtant, malgré ces simplifications, la modélisation de la DCF par les MMC multidimensionnels est suffisamment efficace pour différencier les observations de DCF et apporter des précisions sur le plan clinique. Tous les résultats présentés montrent les capacités des MMC-P et MMC-C à modéliser et analyser des processus multiples qui interagissent dans le temps, comme la DCF entre les réseaux observés lors d'une IRMf de repos. Ils démontrent également la façon dont il est possible de capturer par apprentissage statistique, les relations spatiales et temporelles entre SAM (ou RSN). Enfin, ils illustrent la gamme d'outils d'analyse qui sont immédiatement disponibles dans le cadre d'une analyse par MMC multidimensionnel c'est-à-dire pour comparer des modèles de DCF et les caractériser via leur matrice de transition et leurs séquences d'états décodées. Du point de vue analytique, les outils fournis par la modélisation markovienne sont nombreux. La matrice de transition donne des informations sur les transitions les plus probables, les états ayant la plus forte probabilité d'être atteints ou sur la stabilité des états de connectivité fonctionnelle, exprimés en termes de probabilité. La séquence d'états décodée apporte des informations à différentes échelles d'analyse : à l'échelle globale avec $\hat{\mathbf{X}}$, à une échelle plus petite avec $\hat{X}^{(k)}$ si l'on cherche des interactions plus spécifiques. D'une part, $\hat{\mathbf{X}}$ nous donne accès au nombre de changements d'état, aux états les plus reproduits, à la durée moyenne dans un état et surtout à la répartition temporelle des états au cours de la session. Ce dernier aspect est particulièrement intéressant si l'on cherche à interpréter certains états de connectivité fonctionnelle présents uniquement en début ou fin de session pour des études du repos ou avec des tâches d'activation. D'autre part, le $\hat{X}^{(k)}$ se concentre sur une interaction particulière. Par exemple, le rapport entre le nombre d'états « corrélés » et « anti-corrélés » offre des informations sur la relation entre les deux SAM spécifiques.

Des résultats similaires sont obtenus à partir des modélisations par MMC-P et MMC-C bien que l'on distingue des p -valeurs plus petites avec le MMC-P. Cependant, comme nous ne disposons pas de vérité terrain, il est difficile de donner une priorité sur l'un des deux résultats. Le MMC-C semble toutefois plus robuste au calcul des distances comme nous l'avons montré lors de notre seconde application via l'étude de la reproductibilité des résultats avec de multiples générations de signaux. Par ailleurs, il est intéressant de noter la complexité de la matrice de transition pour le MMC-C est bien moindre que pour le MMC-P au fur et à mesure que le nombre de canaux K augmente. En effet, avec $n = 2$ le nombre d'états possibles par canal et N le nombre de vecteurs d'états possibles, on obtient : pour le MMC-P, $N.N = n^K.n^K = n^{2K} = 2^{2K}$ et pour le MMC-C, $K.N.n = K.n^K.n = Kn^{K+1} = K.2^{K+1}$.

Il est important de revenir aussi sur la problématique soulevée par la taille de l'espace paramètres comparativement aux nombres de données disponibles. Nos modélisations proposent un grand espace d'états, construit *a priori*, et représentant tous les états possibles de connectivité entre les SAM étudiées. La plupart d'entre eux ne sont pas, ou ne peuvent pas, être observés dans les fenêtres temporelles limitées qui nous sont données à étudier à partir d'une session d'IRMf. La création de ces états en amont de l'étude de la DCF conduit à un nombre d'états bien plus important que s'ils avaient été définis directement sur les observations comme c'est le cas dans certaines méthodes décrites en chapitre 8.2 [119, 121, 128]. Il est vrai qu'un grand espace d'états peut conduire à des états très proches les uns des autres en terme de configuration générale des SAM. Pour autant, c'est aussi cela qui permet l'émergence de configurations spatiales plus subtiles, de courte durée ou de faible fréquence. Cette approche par état, rendu possible par la modélisation par MMC multidimensionnels, nous semble plus intéressante car des comparaisons plus directes peuvent être établies entre sessions, sujets ou groupes à partir du moment où les mêmes SAM sont considérées. Il est par

ailleurs toujours possible de se concentrer sur certains états *a posteriori* si l'on souhaite cibler certains RSN. Cette analyse peut donc être rendue spécifique à un RSN tout en s'assurant qu'un grand nombre de RSN ait été pris en compte lors de l'étape de modélisation. Le grand nombre d'états définis s'avère donc intéressant pour approfondir la connaissance de la CF et sa dynamique. Toutefois, il impose une forte demande en terme de données nécessaires pour une bonne estimation des paramètres du modèle. L'espace des paramètres croît de façon exponentielle avec le nombre de processus cachés ce qui nous conduit à limiter l'approche de l'analyse multidimensionnelle à un sous-ensemble de SAM. Pour rester avec un nombre de SAM important et étant donné les durées moyennes d'acquisition de séquences d'IRMf (entre 5 et 10 minutes), il reste pertinent pour une première approche de n'étudier que 2 états possibles par interaction (« corrélié » et « anti-corrélés »). L'introduction d'un 3^{ème} état « neutre », de décorrélation entre paires de SAM, tout en conservant le même nombre de SAM imposerait un espace d'états trop important au vu du nombre de données comme d'un point de vue calculatoire. Avec 7 SAM par exemple, le nombre d'états serait de $3^{(7-1)} = 729$ soit 531 441 transitions possibles.

Conclusion partielle

Au travers des chapitres de cette partie, nous avons pu découvrir et définir plus précisément la notion de dynamique de la connectivité fonctionnelle encore peu décrite dans la littérature et introduire le concept d'« états de connectivité fonctionnelle ». De simples observations permettent parfois de mettre en évidence les fluctuations des relations inter-SAM, comme la corrélation à fenêtre glissante a pu nous le prouver en ciblant certains réseaux décorrélés de manière quasi permanente. Cependant, des approches plus poussées ont été nécessaires pour différencier plus finement des sessions ou des populations.

Ainsi, le développement de méthode de modélisation par MMC-multidimensionnels nous a permis d'une part de comparer des sessions ou individus sur la base de leurs DCF et d'autre part d'apporter des outils permettant de mieux caractériser les processus d'interactions entre SAM. Le rôle clé de l'étude de la DCF est manifeste pour une meilleure compréhension du cerveau sain ou malade.

Nos outils trouvent entièrement leur place parmi les méthodes déjà proposées pour aider à appréhender la DCF. L'approche multidimensionnelle a l'avantage en effet de permettre d'étudier la DCF à différents niveaux : d'un point de vue global en prenant en compte toutes les interactions ou au niveau d'un réseau spécifique pour se concentrer sur ses relations avec un panel d'autres réseaux. Le cadre probabiliste fourni par les approches MMC offre un cadre puissant d'analyse pour modéliser les transitions entre états de connectivité et décoder les relations non observables en premier abord.

Conclusion générale et Perspectives

Conclusion générale

Nous avons présenté dans cette thèse de nouveaux outils pour l'étude de la connectivité fonctionnelle de repos : une méthode de pré-sélection automatique des SAM, au niveau du sujet, et de nouvelles modélisations de la dynamique des interactions entre ces SAM. Le nombre d'études de la dynamique de la connectivité fonctionnelle (DCF) augmente de manière importante ces dernières années, montrant l'intérêt suscité par ce nouveau champ de recherches. Au fur et mesure que des études sont présentées, de nouvelles problématiques apparaissent concernant l'origine neuronale des variations de la connectivité et leur interprétation. Pour aider à répondre à ces questions, la modélisation des interactions entre les réseaux, idiosyncratiques ou communs à un groupe de sujets, est un besoin important. Notre travail a donc été orienté par la nécessité de proposer une réponse à cette problématique de modélisation et caractérisation de la DCF. Cette thèse a été réalisée en ayant toujours le souci de l'applicabilité des méthodes développées par le plus grand nombre d'expérimentateurs du domaine des neurosciences (neuroscientifiques, traiteurs du signal/images, médecins,...). L'objectif clinique est important lors de la conception de nouvelles méthodes, si ce n'est pour une utilisation immédiate, *a minima* cela peut permettre une meilleure compréhension de certaines pathologies ou comportement. Dans cette perspective, les travaux de ce manuscrit ont été tournés vers une application à l'échelle du sujet comme du groupe.

La première partie de cette thèse a porté sur l'automatisation de l'étape de sélection de composantes d'intérêt extraites de l'analyse en composantes indépendantes. C'est une étape chronophage et fastidieuse lorsqu'elle est entièrement faite manuellement. La méthode proposée permet de pré-sélectionner toutes les SAM à l'échelle du sujet. Notre approche, guidée par les données, repose sur une caractérisation des composantes via cinq descripteurs spatiaux et fréquentiels permettant de mettre en avant aussi bien les réseaux connus que ceux spécifiques à un sujet. La procédure de pré-sélection conduit à une forte réduction du nombre de composantes à sélectionner et fournit une sensibilité maximale, requise pour des études sur sujet unique, tout en économisant un temps considérable pour les scientifiques et/ou les médecins.

D'un point de vue applicatif, la méthode a été évaluée sur quatre bases de données issues de protocoles différents (condition d'acquisition, type de séquence, résolution, sujets sains et malades) et montre ainsi son applicabilité dans de multiples études. De plus, une interface a également été développée et permet une sélection finale plus aisée car riche d'informations supplémentaires caractérisant chaque composante restante à classifier. La méthode automatique de pré-sélection SAM que nous proposons devrait faciliter le développement d'études à l'échelle du sujet et particulièrement sur les réseaux idiosyncratiques, peu connus et pourtant particulièrement importants dans le cadre d'une médecine personnalisée.

La deuxième partie de cette thèse présente de nouvelles approches pour modéliser la DCF. Elle permet ainsi de mettre en avant l'importance de l'analyse de la variabilité de la connectivité fonctionnelle sur des échelles de temps plus courtes que celles couramment utilisées. Dans le cadre

probabiliste des réseaux bayésiens dynamiques, le travail de cette partie propose la modélisation, l'apprentissage et l'analyse statistique des signaux de corrélation entre SAM. Deux approches de modélisation et d'analyse markovienne cachée de la DCF ont été proposées : le MMC-Produit et les MMC-Couplés.

La définition *a priori* des états, pour les deux modélisations, fournit un espace d'états riche permettant de représenter, de manière fine, les configurations spatiales des relations entre une SAM et les autres, quelle que soit leur durée ou fréquence durant la session. Cette approche par état, rendu possible par les MMC multidimensionnels, permet la comparaison de modèles entre sessions, entre sujets et entre groupes de sujets. De plus les modélisations multidimensionnelles proposées ont l'avantage de permettre d'étudier la DCF à différents niveaux : d'un point de vue global en prenant en compte toutes les interactions ou au niveau d'un réseau spécifique pour se concentrer sur ses relations avec un panel d'autres réseaux. Le cadre probabiliste fourni par ces approches MMC offre un cadre puissant d'analyse pour modéliser les transitions entre états de connectivité et décoder les relations non observables en premier abord. Avec une première application au niveau du sujet pour une étude pharmacologique et une seconde application au niveau du groupe pour une étude de la démence à corps de Lewy, ces approches ont permis de montrer tout l'intérêt de l'étude de la DCF et le type d'informations que l'on peut en extraire pour aider à la compréhension du cerveau en condition de repos.

La différence entre les deux modèles utilisés se situe plutôt au niveau de la modélisation et de la mise en œuvre de la méthode qu'en termes de résultats au niveau de nos applications. Le MMC-P dispose en effet d'un espace paramètre plus grand, notamment pour la matrice de transitions qui modélise les interactions entre SAM à travers le temps. Les transitions sont décrites entre chaque état multidimensionnel. Le MMC-C intègre, lui, des hypothèses d'indépendances conditionnelles plus fortes puisqu'il modélise les transitions entre un état multidimensionnel et un état sur un seul canal. Si le MMC-C peut paraître plus limité par ces hypothèses, la complexité est moindre pour ce modèle ce qui le rend intéressant dans une optique d'extension de cette modélisation.

Perspectives

Du point de vue de la modélisation, des investigations supplémentaires pourraient ainsi renforcer nos travaux et permettre une plus large application.

- Différents choix d'initialisation des paramètres des modèles, bien que justifiés, ont été effectués. Ils seraient intéressants de les étudier plus en détails (initialisation des probabilités de transitions et d'émission des observations) afin de mesurer leur impact sur les résultats.
- Les résultats obtenus sont nombreux, matrices de transition et décodages des séquences d'états sont disponibles par canal, par réseaux, par sujet, et laissent donc une porte ouverte pour des analyses ultérieures plus approfondies. La caractérisation statistique des états les plus communs, ou encore le repérage automatique de *patterns* de connectivité dans les séquences décodées comme dans les matrices de transitions pourraient apporter encore de nombreuses informations, potentiellement capables d'aider à différencier des sessions, des sujets ou des groupes et à comprendre des processus neuronaux en interaction. Pour cela, des algorithmes de fouille de données ou encore de reconnaissance de formes (structures, schémas dans les décodages ou dans les matrices de transitions), dans une perspective exploratoire, seraient particulièrement profitables.
- D'autres approfondissements des modèles MMC développés peuvent être également envisagés comme extension des travaux de cette thèse. Pour nos études, nous avons considéré que les canaux d'observation, c'est-à-dire les corrélations d'une SAM avec toutes les autres, avait le même poids les unes par rapport aux autres. Pourtant des études récentes ont mis en avant l'aspect hiérarchique du fonctionnement cérébral suggérant ainsi que certaine(s) région(s) ou certain(s) réseau(x) pourraient être supérieurs sur le plan fonctionnel c'est-à-dire régulateurs

ou modulateurs de l'activité de repos [155]. L'adaptation d'un MMC hiérarchique intègre cette structure particulière entre les processus. Il serait intéressant d'observer la DCF sous cet angle pour peut-être répondre à la question de l'existence d'un exécutif central gérant la connectivité fonctionnelle de repos ou de règles cognitives à la base des relations temporelles entre les réseaux.

- De manière plus générale, il serait intéressant d'envisager de synthétiser des signaux de DCF dont on pourrait maîtriser les caractéristiques temporelles afin de tester nos outils. Actuellement il existe des techniques pour évaluer si oui ou non une méthode est apte à détecter la présence de DCF dans un signal (par rapport à du bruit aléatoire) [116, 120]. Mais connaître et contrôler la vérité terrain serait un autre enjeu pour assurer la robustesse des outils de DCF.

Du point de vue des applications, nos méthodes peuvent être déployées facilement à d'autres cas de figures.

- Ces travaux se sont basés sur l'analyse des corrélations de réseaux définis en amont par l'ACI sur l'ensemble de la session (ou des sessions dans le cas de groupe). Pourtant cela peut paraître paradoxal d'utiliser la session complète alors même qu'une étude [110] a montré les variations du profil spatial des réseaux au cours d'une session. Notre méthode pourrait être appliquée pour analyser ces interactions à l'échelle du réseau, c'est-à-dire les liens entre les aires formant le réseau « moyen » obtenu par étude de connectivité statique. Ainsi, en utilisant les corrélations entre les décours temporels des régions formant un réseau *a priori* en entrée de nos modèles, nous pourrions ainsi déterminer le rôle plus précis de certaines aires cérébrales dans la dynamique globale du réseau considéré. En outre, certains réseaux ne peuvent être extraits en étude de groupe incluant des patients du fait de l'altération de ces réseaux par la pathologie étudiée. Cette approche pourrait aider à étudier les origines de cette altération du réseau du point de vue de la DCF.
- En marge de notre modélisation par MMC multidimensionnel, l'étude de Kiviniemi et coll. [110] citée précédemment, montrant les variations du profil spatial d'un réseau au cours du temps, présentent quelques limites comme nous l'avons souligné en paragraphe 8.2.1 page 90. Le faible nombre de composantes extraites par fenêtres temporelles examinées est restreint d'une part par le nombre de données nécessaires pour réaliser l'ACI sur une fenêtre réduite et d'autre part par le temps de sélection des composantes qui augmente avec le nombre de fenêtres à étudier. Cela limite le choix des réseaux à examiner. Notre méthode d'aide à la sélection des SAM alliée aux nouvelles séquences d'IMRF disponibles permettant d'acquérir de nombreuses données en un temps très court (TR pouvant descendre à 100 ms) pourraient résoudre ce problème et rendre possible l'étude de la DCF spatiale au niveau du sujet. Par ailleurs, dans le cadre de la publication de notre méthode de pré-sélection, il est envisagé de rendre disponible au téléchargement, l'interface d'aide à la sélection des SAM.
- Nous nous sommes concentrés jusqu'à présent sur la condition de repos. Notre modélisation serait également pertinente pour une étude comprenant un paradigme de tâches précis qui contraint le sujet à certains états cognitifs répartis dans le temps.
- Enfin, si nous nous sommes restreints pour ces travaux à l'étude des corrélations entre SAM, un atout de ces méthodes est qu'elles peuvent être appliquées directement sur les décours temporels. Cela permettrait de s'affranchir des choix relatifs à l'approche par fenêtre glissante et à la corrélation. Si cette approche ne permet pas de se concentrer sur certains réseaux comme c'est le cas avec les corrélogrammes qui représentent un réseau par rapport à tous les autres, elle a peut suffire pour une caractérisation globale de la DCF ou pour une comparaison de sessions, sujets ou groupes.

Vers une modélisation de l'état de décorrélation entre SAM

Pour terminer ce manuscrit, nous avons souhaité discuter d'une dernière perspective de travail, celle de l'introduction d'un troisième état de connectivité, l'état « neutre » de décorrélation entre deux SAM, dans la modélisation et l'analyse par MMC-P ou MMC-C de la DCF. Cette perspective est volontairement traitée en dehors de la section précédente car nous y apportons une réponse formelle, une nouvelle approche de modélisation : le MMC-P ou le MMC-C à seuil (MMC-PàS, MMC-CàS). Nous n'avons pas encore mis en application cette dernière approche, néanmoins, elle constitue à nos yeux une ultime contribution à ces travaux de thèse.

Principe et vue d'ensemble

Dans les modélisations précédentes, les interactions entre deux SAM étaient associées à un état parmi deux possibles : « corrélé » ou « anti-corrélé ». Cette approche nous a permis de poser les bases pour modéliser la dynamique de ces interactions à l'aide des MMC multidimensionnels (MMC-P ou MMC-C). Les premiers résultats sont intéressants et semblent aller dans le sens de ce qui était attendu. Pour autant, l'ajout d'un troisième état de connectivité qui rendrait possible une décorrélation entre deux SAM, c'est-à-dire où deux SAM ne seraient ni corrélées ni anti-corrélées, apparaît bien plus réaliste.

En pratique, l'approche proposée pour prendre en compte ce troisième état "neutre", de décorrélation entre SAM, s'appuie toujours sur les observables que sont les corrélogrammes $CC_{n,s}$. Pour introduire ce nouvel état, nous avons fait appel à la notion de seuil, variable, de valeur z . Le principe est le suivant. Au delà de la valeur z , le signal de corrélation entre deux SAM sera considéré comme dans l'état « corrélé ». En deçà de la valeur $-z$, ce même signal sera considéré comme dans l'état « anti-corrélé ». Entre ces deux valeurs, $-z$ et z , les deux SAM en question seront considérées comme dans l'état « neutre », c'est-à-dire décorrélées. Les valeurs de seuils $-z$ et z permettent ainsi de définir entre les états « corrélé » et « anti-corrélé », un état intermédiaire, l'état "neutre". Mais attention, dans ce schéma, les valeurs $-z$ et z ne sont pas à considérer comme des seuils de décision durs appliqués directement aux signaux observés : ils sont à considérer au sens statistique, car mis en œuvre indirectement dans la modélisation probabiliste des lois d'observation des états. En effet, le seuil z est probabilisé, c'est-à-dire qu'il est supposé être la réalisation de la variable aléatoire Z , laquelle est considérée comme observable, discrète, et indépendante du temps sur la session considérée. Z intervient dans cette approche pour faire apparaître l'état « neutre » et conditionner les lois d'observations des états cachés de la DCF.

La FIGURE 11.6 permet de donner une vue d'ensemble de la nouvelle problématique à traiter. Antérieurement (côté gauche de la FIGURE 11.6), un modèle de DFC, MMC-P ou MMC-C, avait en charge la modélisation et l'analyse d'un corrélogramme $CC_{n,s}$ représentant la partie observable de la DFC d'une SAM avec K autres SAM. Dans la nouvelle approche (côté droit de la FIGURE 11.6), un nouveau modèle multidimensionnel, que nous avons appelé MMC-P à seuil (MMC-PàS) ou MMC-C à seuil (MMC-CàS), se charge de modéliser globalement L corrélogrammes $CC_{n,s,l}$. Chaque corrélogramme $CC_{n,s,l}$ résulte de l'adjonction d'un seuil z_l au $CC_{n,s}$. Dans notre approche, nous avons considéré L seuils z_l distincts, chaque seuil n'étant observé qu'une seule fois, avec la probabilité d'observation $c_l = P[Z = z_l | \lambda]$.

Avant de préciser le formalisme du MMC-CàS (nous considérons dans la suite que ce cas de figure, le MMC-PàS pouvant être déduit du MMC-P suivant le même principe), il convient de souligner quelques aspects de cette nouvelle approche de modélisation. Comme nous l'avons dit précédemment (voir TABLEAU 11.1), l'introduction du troisième état « neutre » dans les sous espaces d'états $S^{(k)}$ va augmenter de manière significative la complexité du modèle MMC-CàS. Par suite, la nouvelle approche de modélisation de la DFC est dès lors contrainte à limiter l'analyse à un nombre de SAM inférieur à celui des chapitres précédents. En revanche, l'introduction de la variable aléatoire Z et de ses L réalisations z_l , permet, de manière quelque peu artificielle, de multiplier par

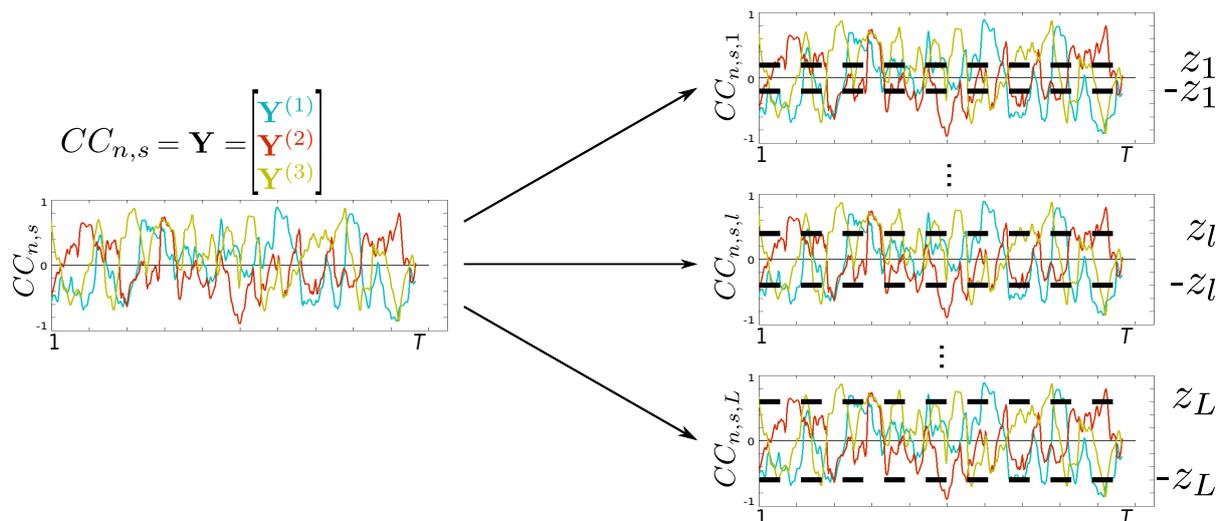


FIGURE 11.6 – Création d'un nouveau corpus d'apprentissage de L correlogrammes $CC_{n,s,l}$ pour les modèles à seuils à partir du correlogramme $CC_{n,s}$.

L le volume des données observées et utilisées pour l'apprentissage du modèle MMC-CàS. Cet avantage *a priori* est néanmoins à pondérer par le fait que, si l'on ne considère que le cas de la matrice de transition, la dimension de cette matrice passe de $2^K \times 2^K$ à $3^K \times 3^K$ (cas du MMC-PàS). Enfin, soulignons la possibilité donnée à l'utilisateur, au travers de la distribution de probabilité de Z , de prendre en compte de manière souple et probabiliste différents seuils d'interprétation des données observées.

MMC-CàS : aspects formels

Un MMC-CàS λ est caractérisé par son jeu de paramètres $\theta = (\Pi, A, B, C)$ avec

$\Pi = \{\pi_{\mathbf{i}}, 1 \leq \mathbf{i} \leq N\}$, l'ensemble des probabilités initiales, où $\pi_{\mathbf{i}}$ décrit la probabilité que le premier état occupé (à $t=1$) soit l'état \mathbf{i} :

$$\pi_{\mathbf{i}} = \prod_{k=1}^K \pi_{i^{(k)}}^{(k)} = \prod_{k=1}^K P[X_1^{(k)} = i^{(k)}] \quad (\text{a})$$

$A = \{a_{\mathbf{i}\mathbf{j}}, 1 \leq \mathbf{i}, \mathbf{j} \leq N\}$, la matrice de transition entre états du modèle, où $a_{\mathbf{i}\mathbf{j}}$ représente la probabilité de transition entre l'état \mathbf{i} et l'état \mathbf{j} :

$$a_{\mathbf{i}\mathbf{j}} = \prod_{k=1}^K a_{i^{(k)}j^{(k)}}^{(k)} = \prod_k P[X_t^{(k)} = j^{(k)} | \mathbf{X}_{t-1} = \mathbf{i}, \lambda] \quad (\text{b})$$

$B = \{b_{\mathbf{j},l}(\mathbf{Y}_t), 1 \leq \mathbf{j} \leq N, 1 \leq l \leq L\}$, l'ensemble des lois d'observation du modèle, où $b_{\mathbf{j}}(\mathbf{Y}_t)$ est la loi d'observation associée à l'état \mathbf{j} observé conditionnellement au seuil $Z = z_l$:

$$b_{\mathbf{j},l}(\mathbf{Y}_t) = P[\mathbf{Y}_t | \mathbf{X}_t = \mathbf{j}, z_l, \lambda] = \prod_{k=1}^K P[Y_t^{(k)} | X_t^{(k)} = j^{(k)}, z_l, \lambda] = \prod_{k=1}^K b_{j^{(k)},l}(Y_t^{(k)}) \quad (\text{c})$$

$C = \{c_l, 1 \leq l \leq L\}$, les probabilités associées à chacun des L seuils :

$$c_l = P[Z = z_l | \lambda] \quad (\text{d})$$

L'équation (c) met en avant dans le MMC-CàS le conditionnement supplémentaire des lois d'observation du MMC-C par la variable aléatoire Z . Ceci conduit à un graphe de dépendance dérivé de celui du MMC-C où l'on voit apparaître les liens orientés entre Z , modélisant le seuil d'interprétation des observations, et les $Y_t^{(k)}$ représentant ces mêmes observations.

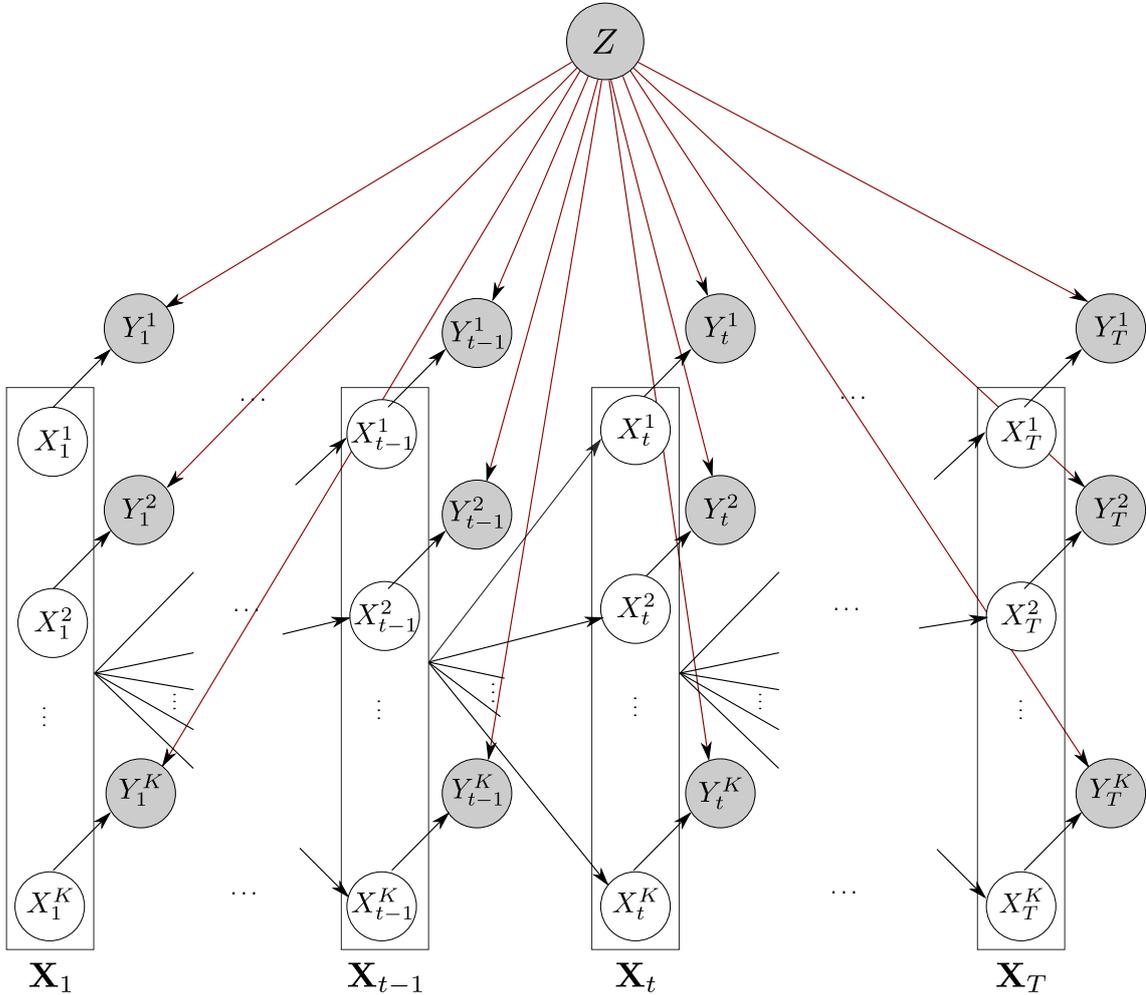


FIGURE 11.7 – Graphe de dépendance du MMC-CàS. En gris les observations, en blanc les états cachés. Les flèches représentent les liens de dépendances. Les couleurs des flèches ont été mises pour plus de clarté et n'ont pas de signification particulière.

Il est à noter que d'un point de vue complexité, rien ne change côté modélisation des processus cachés en interaction. La matrice de transition d'un MMC-CàS reste identique formellement à celle d'un MMC-C. En revanche, elle sera apprise sur l'ensemble des L corrélogrammes $CC_{n,s,l}$, reflétant ainsi de manière statistique l'évolution des interactions entre SAM pour les différents seuils considérés, et non pas pour un seul seuil, le seuil 0 en l'occurrence, comme dans les approches MMC-P/C.

Côté processus observables, la complexité de modélisation est plus importante dans le MMC-CàS puisqu'en lieu et place de chaque loi $b_{j^{(k)}}(Y_t^{(k)})$ apparaissent L lois d'observations $b_{j^{(k)},l}(Y_t^{(k)})$. Cette complexité plus élevée, conduisant à accroître le nombre de paramètres d'observation à estimer, peut néanmoins être réduite substantiellement par des techniques de liaison de paramètres (*parameter tying*) bien connues en reconnaissance automatique de la parole, notamment.

Liste des publications et communications

Publiées/Acceptées

Articles dans des revues internationales à comité de lecture

M. Sourty, L. Thoraval, D. Roquet, J. Armspach, J. Foucher, F. Blanc : "Identifying dynamic functional connectivity changes in dementia with Lewy bodies based on product hidden Markov models", *Frontiers in Comput. Neurosci.*, vol. 10, pp. 60, juin 2016.

A. Ernst, M. Sourty, D. Roquet, V. Noblet, D. Gounot, F. Blanc, J. De Seze, L. Manning : "Functional and structural cerebral changes in key brain regions after a facilitation programme for episodic future thought in relapsing-remitting multiple sclerosis patients ", *Brain and Cognition*, vol. 105, pp 34-45, Juin 2016.

D. Roquet, M. Sourty, A. Botzung, J-P Armspach, F. Blanc : "Brain perfusion in dementia with Lewy bodies and Alzheimer's disease : an arterial spin labeling MRI study on prodromal and mild dementia stages", *Alzheimer's Research and Therapy* (Accepté)

Conférences internationales avec comité de lecture et actes

M. Sourty, L. Thoraval, J. Armspach, J. Foucher : "Product Hidden Markov Models for Subject-based Dynamic Functional Connectivity Analysis of the Resting State Brain", *International Symposium on Biomedical Imaging*, Prague, République Tchèque, Avril 2016.

M. Sourty, L. Thoraval, D. Roquet, J. Armspach, J. Foucher : "Towards an Automated Selection of Spontaneous co-Activity Maps in Functional Magnetic Resonance Imaging", *SPIE Medical Imaging*, Orlando, États-Unis, Février 2015, p. 94170K (présentation orale).

Conférences nationales avec comité de lecture

M. Sourty, L. Thoraval, D. Roquet, J. Armspach, J. Foucher : "Modèles de Markov Cachés Produits pour l'Analyse de la Dynamique de la Connectivité Fonctionnelle du Cerveau au Repos", *4èmes Journées Scientifiques de la Fédération de Médecine Translationnelle de Strasbourg*, Strasbourg, France, Mai 2016 (prix de la meilleure présentation orale).

M. Sourty, L. Thoraval, D. Roquet, J. Armspach, J. Foucher : "Des réseaux non-corrélés révélés par la connectivité fonctionnelle dynamique en IRM fonctionnelle de repos", *SFRMBM*, Grenoble, France, Mars 2015.

En cours de soumission

Articles dans des revues internationales à comité de lecture

M. Sourty, L. Thoraval, D. Roquet, J. Armspach, J. Foucher : "An Automated Pre-Selection Method of Spontaneous co-Activity Maps in Functional Magnetic Resonance Imaging", *Journal of Neurosciences Methods*.

A. Ernst, M. Sourty, D. Roquet, V. Noblet, D. Gounot, F. Blanc, J. De Seze, L. Manning : "Benefits from an autobiographical memory facilitation programme in relapsing-remitting multiple sclerosis patients : a clinical and neuroimaging study", *Neurorehabilitation and Neural Repair*.

Cartes spatiales de SAM

A

A.1 Cartes spatiales des SAM sélectionnées

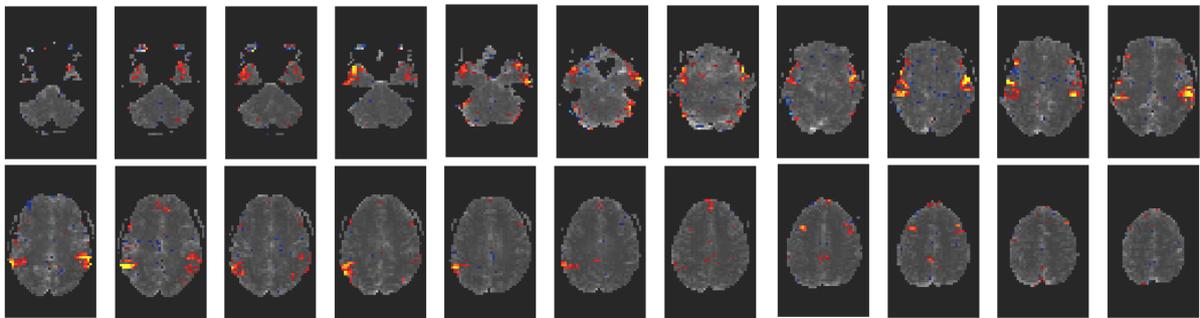


FIGURE A.1 – B2 - sujet 9 - session 2 - composante 155

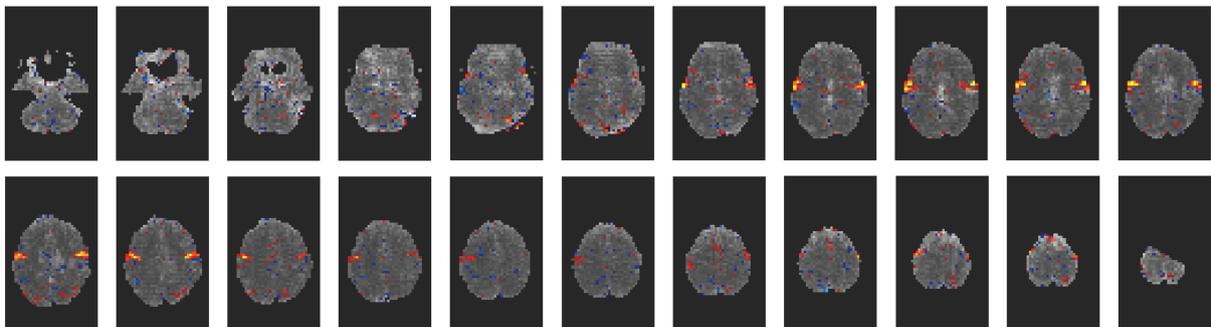


FIGURE A.2 – B5 - sujet 5 - composante 84

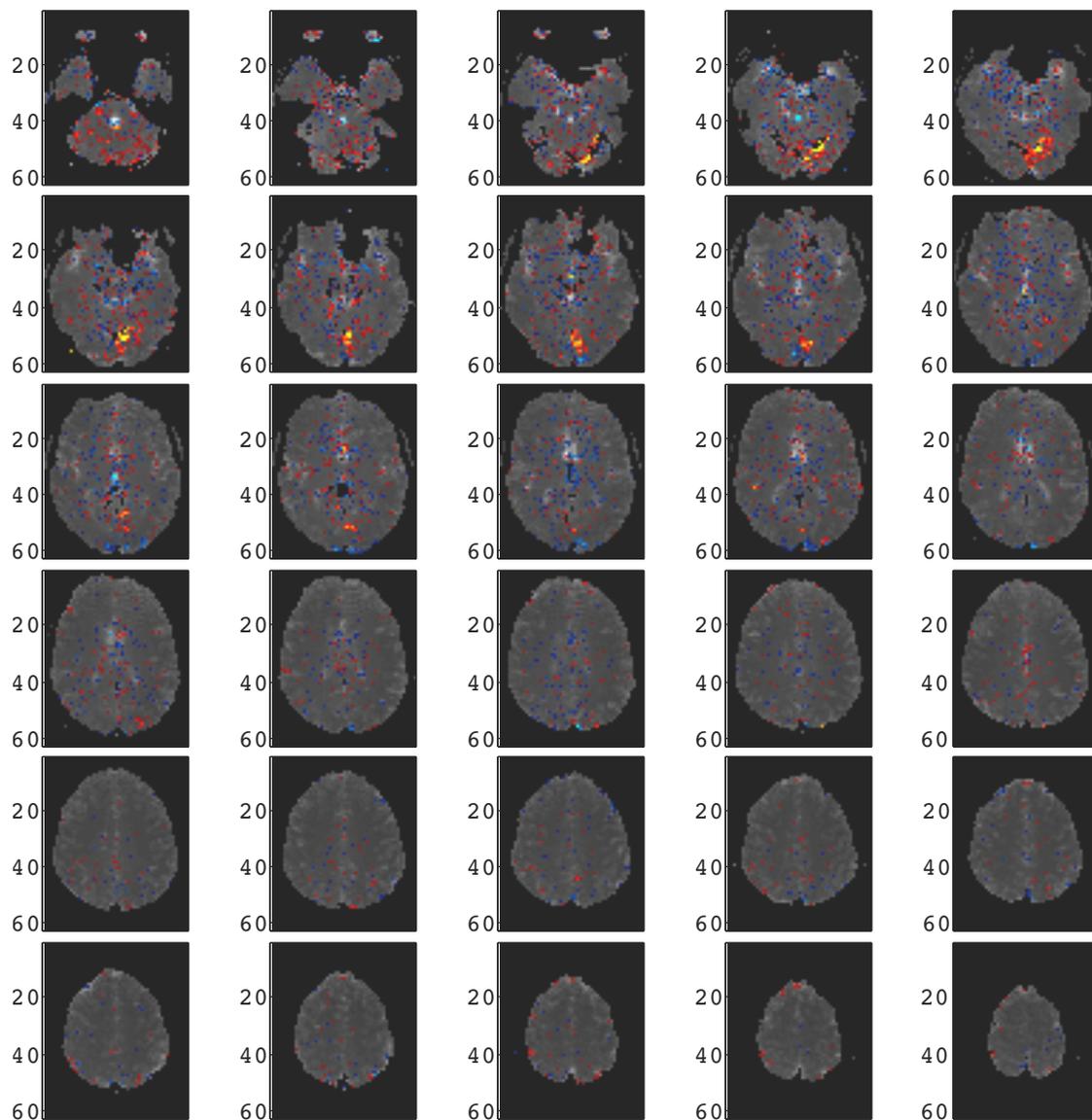
A.2 Cartes spatiales des SAM non sélectionnées : faux négatifs

FIGURE A.3 – B1 - sujet 5 - composante 31

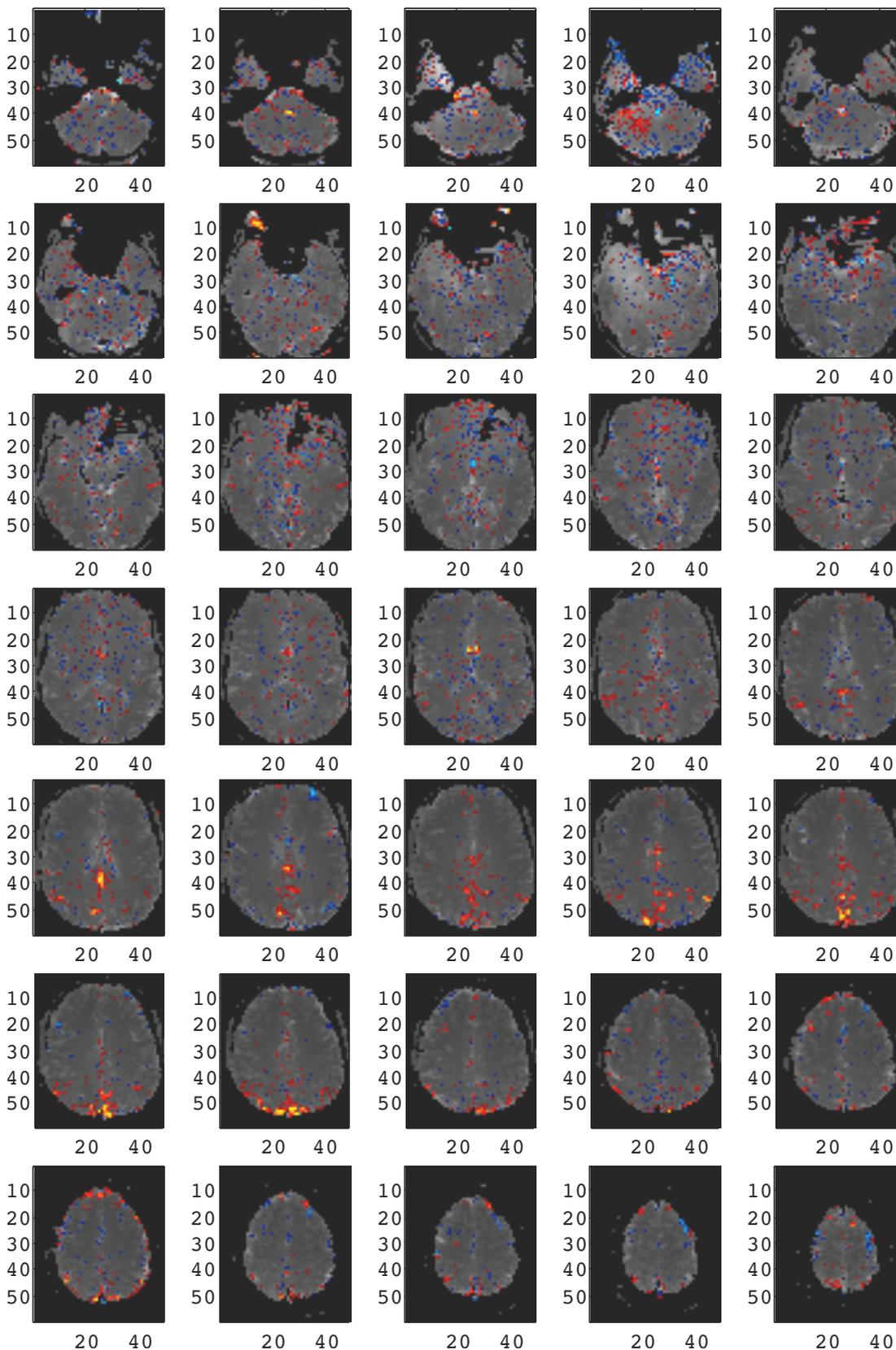


FIGURE A.4 – B1 - sujet 2 - composante 63

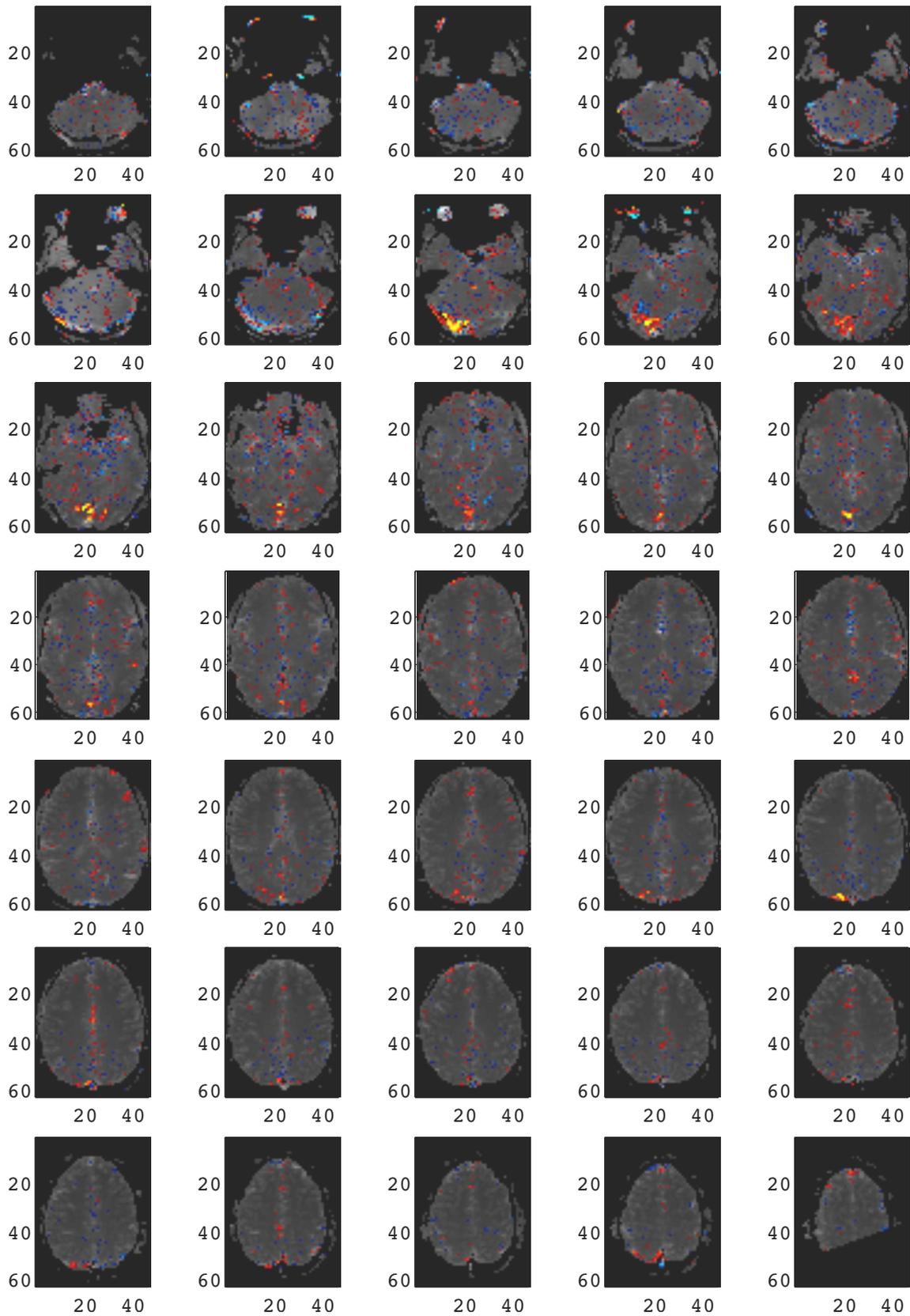


FIGURE A.5 – B1 - sujet 11 - composante 85

Distribution des valeurs des descripteurs par base

B

La figure qui suit illustre la distribution des valeurs aux descripteurs en fonction des différentes bases d'apprentissage.

On retrouve en ligne, dans l'ordre suivant, les descripteurs *fréquence cumulée*, *focusing degree*, *alternance*, *volume*, *localisation* et en colonne les bases de données B1, B2, B3 et l'union des 3 bases de données précédentes.

Les distributions des SAMs sont en rouge, celles des artefacts ou bruits sont en noir. Le seuil bleu est celui calculé sur l'union des trois bases d'apprentissage.

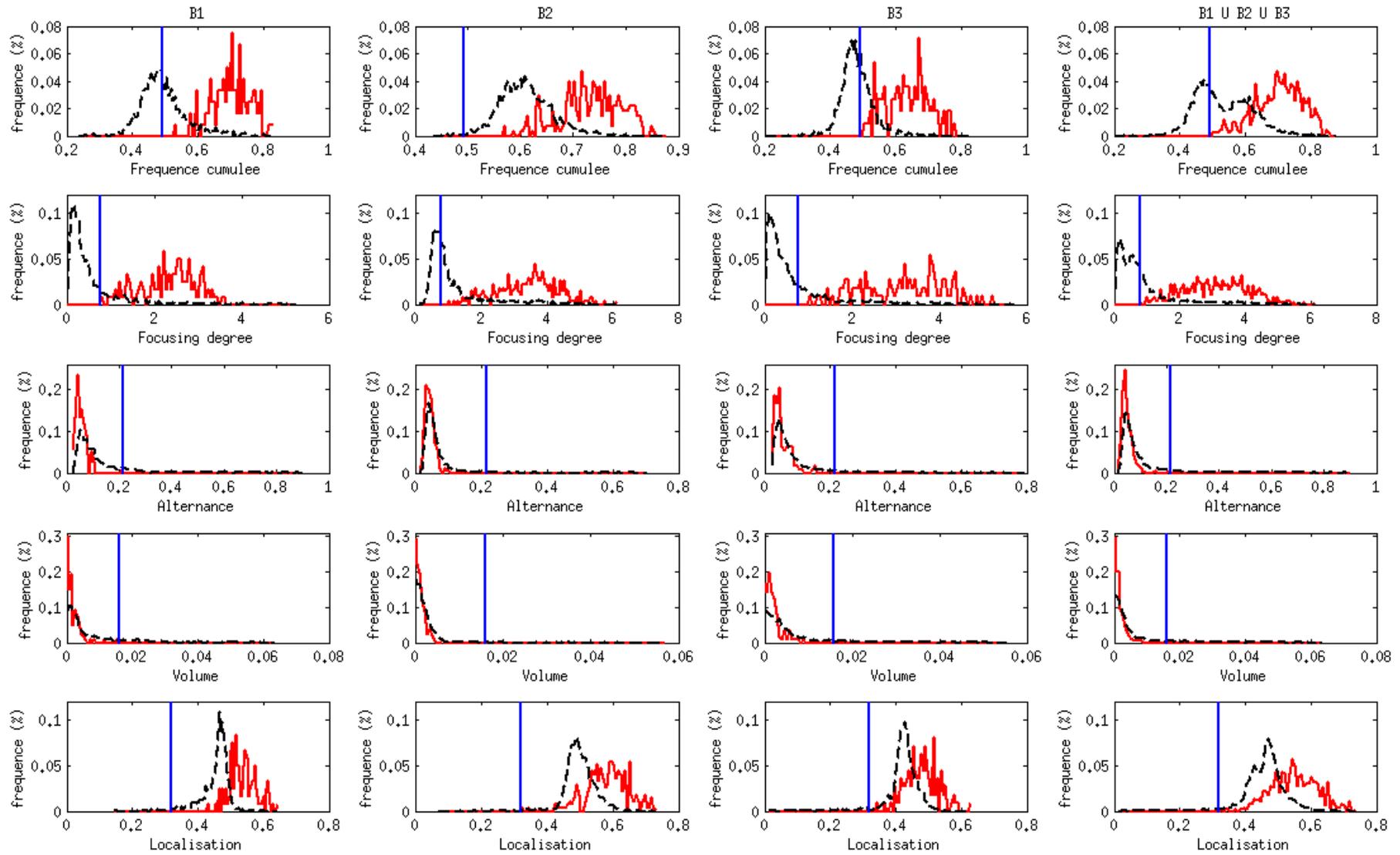


FIGURE B.1 – Distribution selon les bases et les descripteurs

Interface de sélection

C

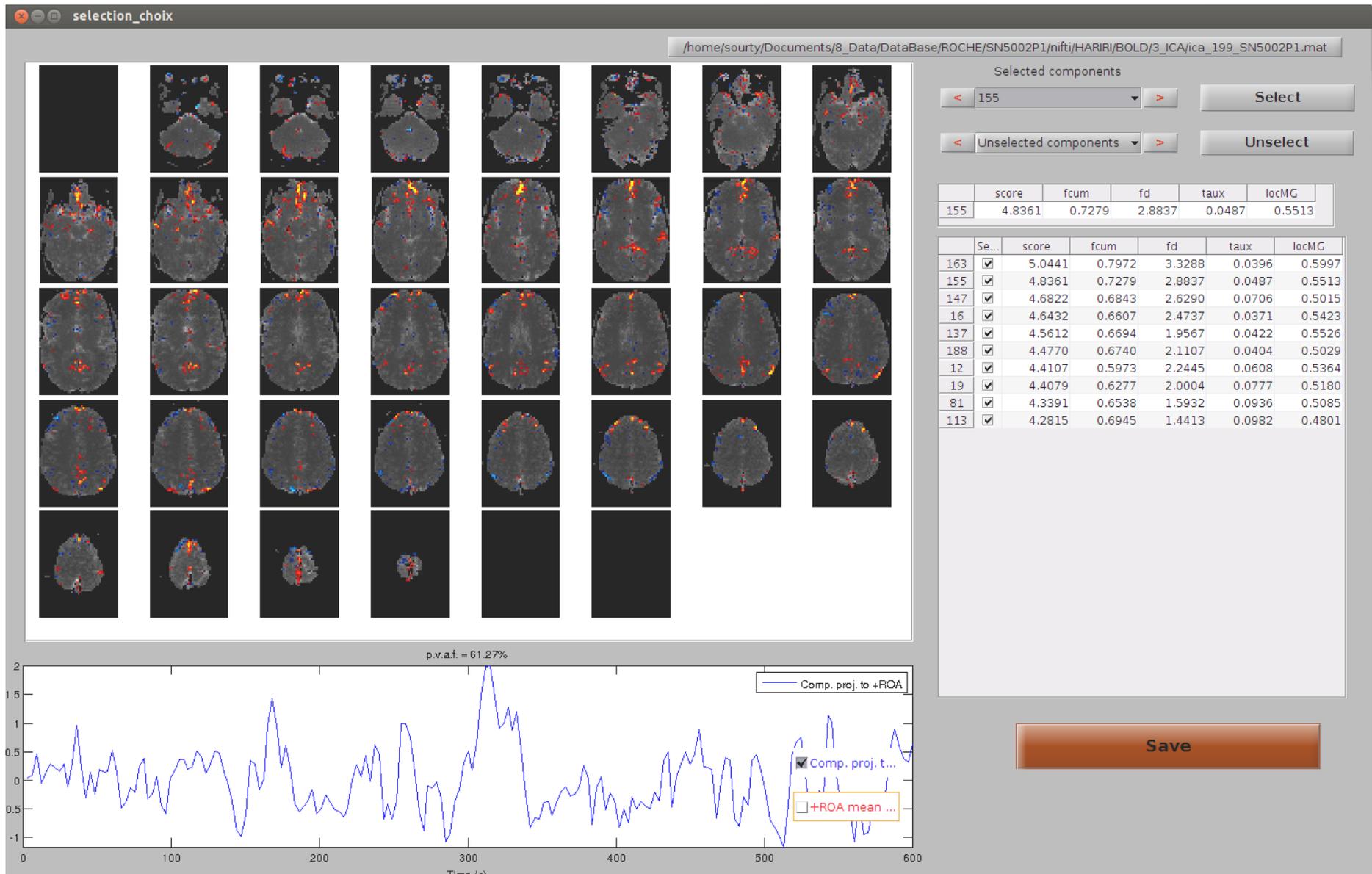


FIGURE C.1 – Interface de sélection

Cartes spatiales des RSN pour l'étude de la DCF sur la base B5

D

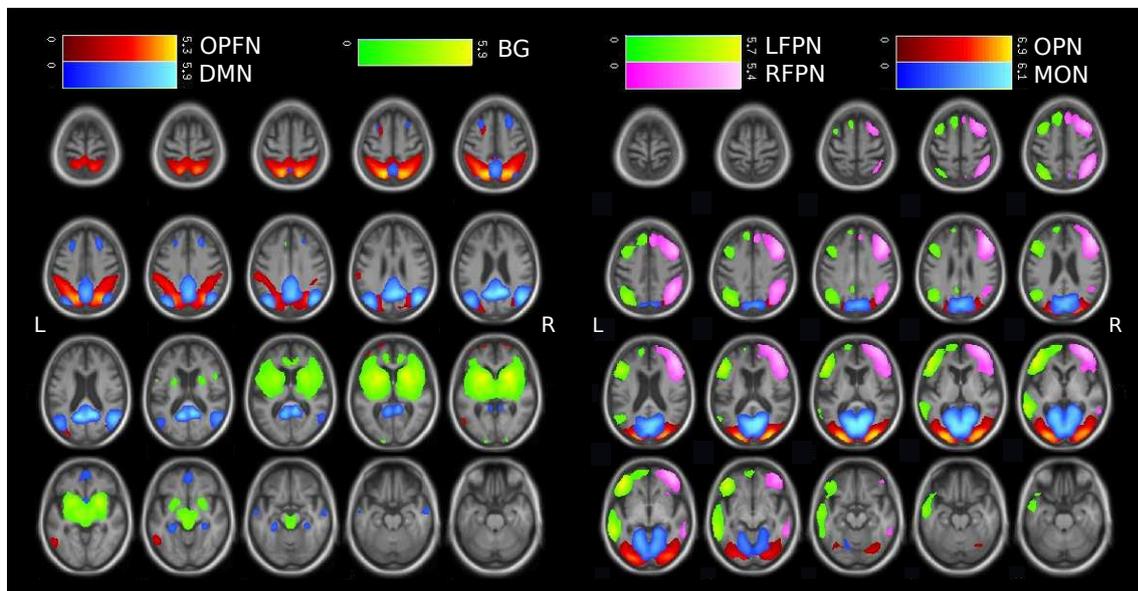


FIGURE D.1 – Cartes spatiales des 7 RSN obtenus par l'ACI de groupe sélectionnés pour l'étude de la base B5. Cartes avec échelle en z-score superposées sur l'image anatomique moyenne sur tous les sujets. DMN : Default Mode Network ; LFPN : Left Fronto-Parietal Network ; RFPN : Right Fronto-Parietal Network ; OPFN : Occipito-Parieto-Frontal Network ; OPN : Occipital Posterior Network ; MON : Medial Occipital Network ; BG : Basal ganglia

Bibliographie

- [1] <http://antranik.org/the-cerebral-hemispheres/>. [connection le 01/06/2016].
- [2] K. Brodmann, *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig : Johann Amb Barth, johann amb barth ed., 1909.
- [3] A. R. McIntosh and F. Gonzalez-Lima, "Structural equation modeling and its application to network analysis in functional brain imaging," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 2–22, 1994.
- [4] K. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, pp. 1273–1302, aug 2003.
- [5] K. J. Friston, C. D. Frith, P. F. Liddle, and R. S. Frackowiak, "Functional connectivity : the principal-component analysis of large (PET) data sets.," *J Cereb Blood Flow Metab*, vol. 13, pp. 5–14, Jan. 1993.
- [6] D. M. Cole, S. M. Smith, and C. F. Beckmann, "Advances and pitfalls in the analysis and interpretation of resting-state fMRI data.," *Frontiers in systems neuroscience*, vol. 4, p. 8, Jan. 2010.
- [7] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, "Dynamic functional connectivity : Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, 2013.
- [8] "e-MRI." <https://www.imaios.com/fr/e-Cours/e-MRI/RMN/spin-nucleaire>. [connection le 01/08/2015].
- [9] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 24, pp. 9868–9872, 1990.
- [10] S. Moeller, E. Yacoub, C. A. Olman, E. Auerbach, J. Strupp, N. Harel, and K. Ugurbil, "Multi-band multislice ge-epi at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fmri," *Magnetic Resonance in Medicine*, vol. 63, no. 5, pp. 1144–1153, 2010.
- [11] V. D. Calhoun, T. Adahi, G. D. 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, pp. 43–53, may 2001.
- [12] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, "Default-mode network activity distinguishes Alzheimer's disease from healthy aging : Evidence from functional MRI," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, pp. 4637–4642, mar 2004.
- [13] D. Roquet, J. Foucher, and S. Kremer, "Resting-state networks distinguish locked-in from vegetative state patients," *NeuroImage : Clinical*, vol. à paraître, 2016.

- [14] M. P. Heilbrun, J. N. Lee, and L. Alvord, "Practical application of fMRI for surgical planning.," *Stereotactic and functional neurosurgery*, vol. 76, pp. 168–74, jan 2001.
- [15] P. M. Matthews, G. D. Honey, and E. T. Bullmore, "Applications of fMRI in translational medicine and clinical practice," *Nature Reviews Neuroscience*, vol. 7, pp. 732–744, sep 2006.
- [16] M. D. Fox and M. E. Raichle, "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," *Nature Reviews Neuroscience*, vol. 8, pp. 700–711, jan 2007.
- [17] M. Fox and M. Greicius, "Clinical applications of resting state functional connectivity," *Frontiers in systems neuroscience*, vol. 4, p. 19, 2010.
- [18] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, pp. 9673–8, July 2005.
- [19] A. L. Cohen, D. A. Fair, N. U. F. Dosenbach, F. M. Miezin, D. Dierker, D. C. Van Essen, B. L. Schlaggar, and S. E. Petersen, "Defining functional areas in individual human brains using resting functional connectivity MRI.," *NeuroImage*, vol. 41, pp. 45–57, may 2008.
- [20] J. L. Vincent, A. Z. Snyder, M. D. Fox, B. J. Shannon, J. R. Andrews, M. E. Raichle, and R. L. Buckner, "Coherent spontaneous activity identifies a hippocampal-parietal memory network.," *Journal of neurophysiology*, vol. 96, pp. 3517–31, dec 2006.
- [21] R. B. Buxton, "Quantifying CBF with arterial spin labeling.," *Journal of magnetic resonance imaging : JMRI*, vol. 22, pp. 723–6, dec 2005.
- [22] J. R. Foucher, D. Roquet, C. Marrer, B.-T. Pham, and D. Gounot, "Correcting for the echo-time effect after measuring the cerebral blood flow by arterial spin labeling.," *Journal of magnetic resonance imaging : JMRI*, vol. 34, pp. 785–90, oct 2011.
- [23] 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, pp. 588–94, Sept. 2011.
- [24] A. Evans, D. Collins, S. Mills, E. Brown, R. Kelly, and T. Peters, "3D statistical neuroanatomical models from 305 MRI volumes," in *1993 IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference*, pp. 1813–1817, IEEE, 1993.
- [25] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI.," *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, vol. 34, pp. 537–41, Oct. 1995.
- [26] M. J. Lowe, M. Dzemidzic, J. T. Lurito, V. P. Mathews, and M. D. Phillips, "Correlations in low-frequency BOLD fluctuations reflect cortico-cortical connections.," *NeuroImage*, vol. 12, pp. 582–7, Nov. 2000.
- [27] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain : A network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, pp. 253–258, July 2003.
- [28] M. Bianciardi, M. Fukunaga, P. van Gelderen, S. G. Horovitz, J. A. de Zwart, and J. H. Duyn, "Modulation of spontaneous fMRI activity in human visual cortex by behavioral state.," *NeuroImage*, vol. 45, pp. 160–8, Mar. 2009.
- [29] M. McAvoy, L. Larson-Prior, M. Ludwikow, D. Zhang, A. Z. Snyder, D. L. Gusnard, M. E. Raichle, and G. D'Avossa, "Dissociated mean and functional connectivity BOLD signals in visual cortex during eyes closed and fixation.," *Journal of neurophysiology*, vol. 108, pp. 2363–72, Nov. 2012.

- [30] M. Greicius, "Resting-state functional connectivity in neuropsychiatric disorders.," *Current opinion in neurology*, vol. 21, pp. 424–30, Aug. 2008.
- [31] L. Heine, A. Soddu, F. Gómez, A. Vanhaudenhuyse, L. Tshibanda, M. Thonnard, V. Charland-Verville, M. Kirsch, S. Laureys, and A. Demertzi, "Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states," *Frontiers in Psychology*, vol. 3, no. AUG, 2012.
- [32] V. Menon, "Large-scale brain networks and psychopathology : a unifying triple network model.," *Trends in cognitive sciences*, vol. 15, pp. 483–506, Oct. 2011.
- [33] J. S. Damoiseaux, S. Rombouts, F. Barkhof, P. Scheltens, C. J. Stam, S. M. Smith, and C. F. Beckmann, "Consistent resting-state networks across healthy subjects," *Proceedings of the national academy of sciences*, vol. 103, no. 37, pp. 13848–13853, 2006.
- [34] K. Kalcher, W. Huf, R. N. Boubela, P. Filzmoser, L. Pezawas, B. Biswal, S. Kasper, E. Moser, and C. Windischberger, "Fully exploratory network independent component analysis of the 1000 functional connectomes database.," *Frontiers in human neuroscience*, vol. 6, p. 301, jan 2012.
- [35] J. Foucher, "Cercle d'excellence sur les psychoses." <http://www.cercle-d-excellence-psy.org>.
- [36] C. Jutten, *Calcul neuromimétique et traitement du signal : analyse en composantes indépendantes*. PhD thesis, Grenoble INPG, 1987.
- [37] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, pp. 287–314, Apr. 1994.
- [38] A. J. Bell and T. J. Sejnowski, "An Information-Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, nov 1995.
- [39] V. D. Calhoun, J. Liu, and T. Adalı, "A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data," *Neuroimage*, vol. 45, pp. S163–S172, Mar. 2009.
- [40] V. D. Calhoun, T. Adalı, V. B. McGinty, J. J. Pekar, T. D. Watson, and G. D. Pearlson, "fMRI activation in a visual-perception task : network of areas detected using the general linear model and independent components analysis.," *NeuroImage*, vol. 14, pp. 1080–8, Nov. 2001.
- [41] F. Esposito, T. Scarabino, A. Hyvarinen, J. Himberg, E. Formisano, S. Comani, G. Tedeschi, R. Goebel, E. Seifritz, and F. Di Salle, "Independent component analysis of fMRI group studies by self-organizing clustering.," *NeuroImage*, pp. 193–205, Mar. 2005.
- [42] V. D. Calhoun, T. Adalı, G. D. Pearlson, and J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis.," *Human brain mapping*, vol. 14, pp. 140–51, Nov. 2001.
- [43] M. Svensén, F. Kruggel, and H. Benali, "ICA of fMRI Group Study Data," *NeuroImage*, vol. 16, pp. 551–563, July 2002.
- [44] V. J. Schmithorst and S. K. Holland, "Comparison of three methods for generating group statistical inferences from independent component analysis of functional magnetic resonance imaging data.," *Journal of magnetic resonance imaging : JMRI*, vol. 19, pp. 365–8, Mar. 2004.
- [45] C. VD., "Group ICA of fMRI toolbox (GIFT)." <http://mialab.mrn.org/software/>, 2004.
- [46] E. B. Erhardt, S. Rachakonda, E. J. Bedrick, E. A. Allen, T. Adalı, and V. D. Calhoun, "Comparison of multi-subject ICA methods for analysis of fMRI data.," *Human brain mapping*, vol. 32, pp. 2075–95, Dec. 2011.
- [47] C. F. Beckmann, C. E. Mackay, N. Filippini, and S. M. Smith, "Group comparison of resting-state fMRI data using multi-subject ICA and dual regression," in *Proc. Human Brain Mapping*, (San Francisco), 2009.

- [48] J. M. Pignat, O. Koval, D. Van De Ville, S. Voloshynovskiy, C. Michel, and T. Pun, "The impact of denoising on independent component analysis of functional magnetic resonance imaging data.," *Journal of neuroscience methods*, vol. 213, pp. 105–22, feb 2013.
- [49] D. R. Roquet and J. R. Foucher, "Empirical estimation of the dimensions to retain in independent component analysis," 2012.
- [50] A. Abou-Elseoud, T. Starck, J. Remes, J. Nikkinen, O. Tervonen, and V. Kiviniemi, "The effect of model order selection in group PICA.," *Human brain mapping*, vol. 31, pp. 1207–16, Aug. 2010.
- [51] E. A. Allen, E. B. Erhardt, Y. Wei, T. Eichele, and V. D. Calhoun, "Capturing inter-subject variability with group independent component analysis of fMRI data : a simulation study.," *NeuroImage*, vol. 59, pp. 4141–4159, Feb. 2012.
- [52] M. J. McKeown, L. K. Hansen, and T. J. Sejnowsk, "Independent component analysis of functional MRI : what is signal and what is noise?," *Current Opinion in Neurobiology*, vol. 13, pp. 620–629, Oct. 2003.
- [53] V. Kiviniemi, J.-H. Kantola, J. Jauhiainen, A. Hyvärinen, and O. Tervonen, "Independent component analysis of nondeterministic fMRI signal sources," *NeuroImage*, vol. 19, pp. 253–260, June 2003.
- [54] K. Petersen and L. Hansen, "On the independent components of functional neuroimages," in *Second International Symposium on Independent Component Analysis and Blind Signal Separation*, no. September 2000, pp. 615–620, 2000.
- [55] A. Hyvarinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Transactions on Neural Networks*, vol. 10, pp. 626–634, may 1999.
- [56] F. Esposito, E. Formisano, E. Seifritz, R. Goebel, R. Morrone, G. Tedeschi, and F. Di Salle, "Spatial independent component analysis of functional MRI time-series : To what extent do results depend on the algorithm used?," *Human Brain Mapping*, vol. 16, pp. 146–157, jul 2002.
- [57] M. J. McKeown, "Detection of consistently task-related activations in fMRI data with hybrid independent component analysis.," *NeuroImage*, vol. 11, pp. 24–35, Jan. 2000.
- [58] J. Himberg, A. Hyvärinen, and F. Esposito, "Validating the independent components of neuroimaging time series via clustering and visualization.," *NeuroImage*, vol. 22, jul 2004.
- [59] Z. Yang, S. LaConte, X. Weng, and X. Hu, "Ranking and averaging independent component analysis by reproducibility (RAICAR)," *Human Brain Mapping*, vol. 29, no. 6, pp. 711–725, 2008.
- [60] A. Markov, "Extension of the law of large numbers to dependent quantities [in Russian]," *Izvestiia Fiz.-Matem. Kazan Univ.*, vol. 15, no. (2nd Ser.), pp. 135–156, 1906.
- [61] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state markov chains," *Annals of Mathematical Statistics*, vol. 37, pp. 1554–1563, 1966.
- [62] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bulletin of the American Mathematical Society*, vol. 73, pp. 360–363, May 1967.
- [63] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, pp. 260–269, Apr. 1967.
- [64] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," in *Inequalities III : Proceedings of the Third Symposium on Inequalities* (O. Shisha, ed.), (University of California, Los Angeles), pp. 1–8, Academic Press, 1972.

- [65] A. Dempster, N. Laird, and D. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [66] J. A. Bilmes, "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," tech. rep., 1998.
- [67] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," in *Proceedings of the IEEE*, pp. 257–286, 1989.
- [68] S. E. Levinson, L. R. Rabiner, and M. M. Sondhi, "An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition," *Bell System Technical Journal*, vol. 62, pp. 1035–1074, apr 1983.
- [69] L. Bahl, P. Brown, P. de Souza, and R. Mercer, "Maximum mutual information estimation of hidden Markov model parameters for speech recognition," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, no. 0, pp. 49–52, 1986.
- [70] A. V. Nefian, L. Liang, X. Pi, X. Liu, and K. P. Murphy, "Dynamic Bayesian networks for audio-visual speech recognition," *Journal on Applied Signal Processing*, vol. 2002, no. 11, pp. 1274–1288, 2002.
- [71] L. White, "Cartesian hidden Markov models with applications," *Signal Processing, IEEE Transactions on*, vol. 40, no. 6, pp. 1601–1604, 1992.
- [72] Mukundhn N. and T. Sreevinas, "Product HMM - A novel class of HMMS for sub-sequence modelling," 2002.
- [73] M. BRAND, "Coupled Hidden Markov models for modeling interacting processes," tech. rep., MIT Media Lab Perceptual Computong, 1997.
- [74] M. Brand, N. Oliver, and a. Pentland, "Coupled hidden Markov models for complex action recognition," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 994–999, 1997.
- [75] F. Pernkopf, "3D surface analysis using coupled HMMs," *Machine Vision and Applications*, vol. 16, no. 5, pp. 298–305, 2005.
- [76] J. Luettin, G. Potamianos, and C. Neti, "Asynchronous stream modeling for large vocabulary audio-visual speech recognition," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.01CH37221)*, vol. 1, pp. 169–172, IEEE, 2001.
- [77] Z. Ghahramani and M. I. Jordan, "Factorial hidden Markov models," *Machine learning*, vol. 29, pp. 245–273, 1997.
- [78] S. Fine, Y. Singer, and N. Tishby, "The Hierarchical Hidden Markov Model : Analysis and Applications," *Maschine Learning*, vol. 32, no. 1, pp. 41–62, 1998.
- [79] D. Aarsland, A. Rongve, S. Piepenstock Nore, R. Skogseth, S. Skulstad, U. Ehrh, D. Hopreksstad, and C. Ballard, "Frequency and case identification of dementia with Lewy bodies using the revised consensus criteria," *Dementia and Geriatric Cognitive Disorders*, vol. 26, no. 5, pp. 445–452, 2008.
- [80] I. G. McKeith, D. W. Dickson, J. Lowe, M. Emre, J. T. O'Brien, H. Feldman, J. Cummings, J. E. Duda, C. Lippa, E. K. Perry, D. Aarsland, H. Arai, C. G. Ballard, B. Boeve, D. J. Burn, D. Costa, T. Del Ser, B. Dubois, D. Galasko, S. Gauthier, C. G. Goetz, E. Gomez-Tortosa, G. Halliday, L. A. Hansen, J. Hardy, T. Iwatsubo, R. N. Kalaria, D. Kaufer, R. A. Kenny, A. Korczyn, K. Kosaka, V. Lee, A. Lees, I. Litvan, E. Londos, O. L. Lopez, S. Minoshima, Y. Mizuno, J. A. Molina, E. B. Mukaetova-Ladinska, F. Pasquier, R. H. Perry, J. B. Schulz, J. Q. Trojanowski, and M. Yamada, "Diagnosis and management of dementia with Lewy bodies : Third report of the DLB consortium," *Neurology*, vol. 65, pp. 1863–1872, dec 2005.

- [81] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "“Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician.,” *Journal of psychiatric research*, vol. 12, no. 3, pp. 189–198, 1975.
- [82] D. R. Roquet, B.-T. Pham, and J. R. Foucher, “Manual selection of spontaneous activity maps derived from independent component analysis : criteria and inter-rater reliability study.,” *Journal of neuroscience methods*, vol. 223, pp. 30–4, Feb. 2014.
- [83] J. Sui, T. Adalı, G. D. Pearlson, and V. D. Calhoun, “An ICA-based method for the identification of optimal fMRI features and components using combined group-discriminative techniques,” *NeuroImage*, vol. 46, pp. 73–86, May 2009.
- [84] W. Zeng, A. Qiu, B. Chodkowski, and J. J. Pekar, “Spatial and temporal reproducibility-based ranking of the independent components of BOLD fMRI data,” *NeuroImage*, vol. 46, pp. 1041–1054, July 2009.
- [85] S. Malinen and R. Hari, “Data-based functional template for sorting independent components of fMRI activity,” *Neuroscience Research*, vol. 71, pp. 369–376, Dec. 2011.
- [86] M. J. Mckeown, S. Makeig, G. G. Brown, T.-P. Jung, S. S. Kindermann, A. J. Bell, and T. J. Sejnowski, “Analysis of fMRI data by blind separation into independent spatial components,” *Human Brain Mapping*, vol. 6, no. 3, pp. 160–188, 1998.
- [87] C. G. Thomas, R. A. Harshman, and R. S. Menon, “Noise Reduction in BOLD-Based fMRI Using Component Analysis,” *NeuroImage*, vol. 17, pp. 1521–1537, Nov. 2002.
- [88] V. G. van de Ven, E. Formisano, D. Prvulovic, C. H. Roeder, and D. E. Linden, “Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest,” *Human Brain Mapping*, vol. 22, no. 3, pp. 165–178, 2004.
- [89] F. D. Martino, F. Gentile, F. Esposito, and M. Balsi, “Classification of fMRI independent components using IC- fingerprints and support vector machine classifiers,” *Neuroimage*, vol. 34, pp. 177–194, Jan. 2007.
- [90] V. Perlberg, P. Bellec, J.-L. Anton, M. Péligrini-Issac, J. Doyon, and H. Benali, “CORSICA : correction of structured noise in fMRI by automatic identification of ICA components,” *Magnetic Resonance Imaging*, vol. 25, pp. 35–46, Jan. 2007.
- [91] J. Tohka, K. Foerde, A. R. Aron, S. M. Tom, A. W. Toga, and R. A. Poldrack, “Automatic independent component labeling for artifact removal in fMRI,” *NeuroImage*, vol. 39, pp. 1227–1245, Feb. 2008.
- [92] R. E. Kelly Jr., G. S. Alexopoulos, Z. Wang, F. M. Gunning, C. F. Murphy, S. S. Morimoto, D. Kanellopoulos, Z. Jia, K. O. Lim, and M. J. Hoptman, “Visual inspection of independent components : Defining a procedure for artifact removal from fMRI data,” *Journal of Neuroscience Methods*, vol. 189, pp. 233–245, June 2010.
- [93] S. F. Storti, E. Formaggio, R. Nordio, P. Manganotti, A. Fiaschi, A. Bertoldo, and G. M. Toffolo, “Automatic selection of resting-state networks with functional magnetic resonance imaging.,” *Frontiers in neuroscience*, vol. 7, p. 72, Jan. 2013.
- [94] K. Bhaganagarapu, G. D. Jackson, and D. F. Abbott, “An automated method for identifying artifact in independent component analysis of resting-state FMRI.,” *Frontiers in human neuroscience*, vol. 7, p. 343, Jan. 2013.
- [95] F. Esposito, A. Aragri, I. Pesaresi, S. Cirillo, G. Tedeschi, E. Marciano, R. Goebel, and F. Di Salle, “Independent component model of the default-mode brain function : combining individual-level and population-level analyses in resting-state fMRI,” *Magnetic Resonance Imaging*, vol. 26, pp. 905–913, Sept. 2008.
- [96] S. P. Mapping, “Spm8, wellcome trust centre for neuroimaging, ucl, uk.” <http://www.fil.ion.ucl.ac.uk/spm/>, 2009.

- [97] MATLAB, *version 8.0.0.783 (R2012b)*. Natick, Massachusetts : The MathWorks Inc., 2012.
- [98] K.-H. Chuang, P. van Gelderen, H. Merkle, J. Bodurka, V. N. Ikonomidou, A. P. Koretsky, J. H. Duyn, and S. L. Talagala, "Mapping resting-state functional connectivity using perfusion MRI," *NeuroImage*, vol. 40, pp. 1595–605, May 2008.
- [99] D. R. Roquet, M. Sourty, and J. R. Foucher, *Etude et application de la connectivité fonctionnelle chez le sujet sain et dans la pathologie - article : Empirical estimation of the dimensions to retain in independent component analysis*. PhD thesis, Université de Strasbourg, 2014.
- [100] I. for Neural Computation, "FMRLAB, University of California San Diego." <http://sccn.ucsd.edu/fmrlab/>.
- [101] J. MacQueen, "Some Methods for Classification and Analysis of Multivariate Observations," in *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, Berkeley : University of California Press, 1967.
- [102] L. Breiman, "Random Forests," *Machine Learning*, vol. 45, pp. 5–32, oct 2001.
- [103] C. J. Gauthier, C. Madjar, L. Desjardins-Crépeau, P. Bellec, L. Bherer, and R. D. Hoge, "Age dependence of hemodynamic response characteristics in human functional magnetic resonance imaging," *Neurobiology of aging*, vol. 34, pp. 1469–85, may 2013.
- [104] L. Griffanti, O. Dipasquale, M. M. Laganà, R. Nemni, M. Clerici, S. M. Smith, G. Baselli, and F. Baglio, "Effective artifact removal in resting state fMRI data improves detection of DMN functional connectivity alteration in Alzheimer's disease," *Frontiers in Human Neuroscience*, vol. 9, pp. 1–11, aug 2015.
- [105] S. G. Horowitz, A. R. Braun, W. S. Carr, D. Picchioni, T. J. Balkin, M. Fukunaga, and J. H. Duyn, "Decoupling of the brain's default mode network during deep sleep," *Proceedings of the National Academy of Sciences*, vol. 106, no. 27, pp. 11376–11381, 2009.
- [106] M. D. Greicius, V. Kiviniemi, O. Tervonen, V. Vainionpää, S. Alahuhta, A. L. Reiss, and V. Menon, "Persistent default-mode network connectivity during light sedation," *Human Brain Mapping*, vol. 29, pp. 839–847, jul 2008.
- [107] F. Esposito, A. Bertolino, T. Scarabino, V. Latorre, G. Blasi, T. Popolizio, G. Tedeschi, S. Cirillo, R. Goebel, and F. Di Salle, "Independent component model of the default-mode brain function : Assessing the impact of active thinking," *Brain research bulletin*, vol. 70, pp. 263–9, oct 2006.
- [108] M. Song, Y. Zhou, J. Li, Y. Liu, L. Tian, C. Yu, and T. Jiang, "Brain spontaneous functional connectivity and intelligence," *NeuroImage*, vol. 41, pp. 1168–1176, jul 2008.
- [109] R. Sala-Llonch, D. Bartrés-Faz, and C. Junqué, "Reorganization of brain networks in aging : a review of functional connectivity studies," *Frontiers in Psychology*, vol. 6, p. 663, may 2015.
- [110] V. Kiviniemi, T. Vire, J. Remes, A. A. Elseoud, T. Starck, O. Tervonen, and J. Nikkinen, "A sliding time-window ICA reveals spatial variability of the default mode network in time," *Brain connectivity*, vol. 1, pp. 339–47, jan 2011.
- [111] J. Gonzalez-Castillo, D. A. Handwerker, M. E. Robinson, C. W. Hoy, L. C. Buchanan, Z. S. Saad, and P. A. Bandettini, "The spatial structure of resting state connectivity stability on the scale of minutes," *Frontiers in neuroscience*, vol. 8, p. 138, jan 2014.
- [112] S. D. Keilholz, "The Neural Basis of Time-Varying Resting-State Functional Connectivity," *Brain Connectivity*, vol. 4, pp. 769–779, dec 2014.
- [113] E. Tagliazucchi, F. von Wegner, A. Morzelewski, V. Brodbeck, and H. Laufs, "Dynamic BOLD functional connectivity in humans and its electrophysiological correlates," *Frontiers in Human Neuroscience*, vol. 6, p. 339, jan 2012.

- [114] V. D. Calhoun, R. Miller, G. Pearlson, and T. Adalı, "The Chronnectome : Time-Varying Connectivity Networks as the Next Frontier in fMRI Data Discovery," *Neuron*, vol. 84, pp. 262–274, oct 2014.
- [115] R. M. Hutchison, T. Womelsdorf, J. S. Gati, S. Everling, and R. S. Menon, "Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques.," *Human Brain Mapping*, vol. 00, no. November 2011, 2012.
- [116] A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, and M. Breakspear, "Time-resolved resting-state brain networks.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 28, pp. 10341–6, 2014.
- [117] K. Shen, R. M. Hutchison, G. Bezgin, S. Everling, and a. R. McIntosh, "Network Structure Shapes Spontaneous Functional Connectivity Dynamics," *Journal of Neuroscience*, vol. 35, no. 14, pp. 5579–5588, 2015.
- [118] E. Damaraju, E. A. Allen, A. Belger, J. M. Ford, S. McEwen, D. H. Mathalon, B. A. Mueller, G. D. Pearlson, S. G. Potkin, A. Preda, J. A. Turner, J. G. Vaidya, T. G. van Erp, and V. D. Calhoun, "Dynamic functional connectivity analysis reveals transient states of dysconnectivity in schizophrenia.," *NeuroImage. Clinical*, vol. 5, pp. 298–308, jan 2014.
- [119] E. a. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, "Tracking Whole-Brain Connectivity Dynamics in the Resting State.," *Cerebral cortex (New York, N.Y. : 1991)*, pp. 1–14, Nov. 2012.
- [120] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fMRI.," *NeuroImage*, vol. 50, pp. 81–98, mar 2010.
- [121] S. Ma, V. D. Calhoun, R. Phlypo, and T. Adalı, "Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis.," *NeuroImage*, vol. 90, pp. 196–206, apr 2014.
- [122] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville, "Principal components of functional connectivity : A new approach to study dynamic brain connectivity during rest.," *NeuroImage*, vol. 83, pp. 937–50, dec 2013.
- [123] Q. Yu, E. B. Erhardt, J. Sui, Y. Du, H. He, D. Hjelm, M. S. Cetin, S. Rachakonda, R. L. Miller, G. Pearlson, and V. D. Calhoun, "Assessing dynamic brain graphs of time-varying connectivity in fMRI data : application to healthy controls and patients with schizophrenia.," *NeuroImage*, vol. 107, pp. 345–55, feb 2015.
- [124] D. T. Jones, P. Vemuri, M. C. Murphy, J. L. Gunter, M. L. Senjem, M. M. Machulda, S. A. Przybelski, B. E. Gregg, K. Kantarci, D. S. Knopman, B. F. Boeve, R. C. Petersen, and C. R. Jack, "Non-stationarity in the "resting brain's" modular architecture.," *PLoS one*, vol. 7, p. e39731, jan 2012.
- [125] B. Rashid, E. Damaraju, G. D. Pearlson, and V. D. Calhoun, "Dynamic connectivity states estimated from resting fMRI Identify differences among Schizophrenia, bipolar disorder, and healthy control subjects.," *Frontiers in human neuroscience*, vol. 8, p. 897, jan 2014.
- [126] S. M. Smith, K. L. Miller, S. Moeller, J. Xu, E. J. Auerbach, M. W. Woolrich, C. F. Beckmann, M. Jenkinson, J. Andersson, M. F. Glasser, D. C. Van Essen, D. A. Feinberg, E. S. Yacoub, and K. Ugurbil, "Temporally-independent functional modes of spontaneous brain activity.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, pp. 3131–6, feb 2012.
- [127] V. D. Calhoun, T. Eichele, T. Adalı, and E. A. Allen, "Decomposing the brain : components and modes, networks and nodes.," *Trends in cognitive sciences*, vol. 16, pp. 255–6, may 2012.
- [128] H. Eavani, T. D. Satterthwaite, R. E. Gur, R. C. Gur, and C. Davatzikos, "Unsupervised learning of functional network dynamics in resting state fMRI.," in *Information processing in medical imaging*, vol. 23, pp. 426–37, jan 2013.

- [129] J. Ou, L. Xie, P. Wang, X. Li, D. Zhu, R. Jiang, Y. Wang, Y. Chen, J. Zhang, and T. Liu, "Modeling brain functional dynamics via hidden Markov models," in *International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 569–572, IEEE, nov 2013.
- [130] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns.," *Cerebral cortex (New York, N.Y. : 1991)*, vol. 22, pp. 158–65, jan 2012.
- [131] I. Cribben, R. Haraldsdottir, L. Y. Atlas, T. D. Wager, and M. a. Lindquist, "Dynamic connectivity regression : determining state-related changes in brain connectivity.," *NeuroImage*, vol. 61, pp. 907–20, jul 2012.
- [132] R. Hindriks, M. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. Logothetis, and G. Deco, "Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI?," *NeuroImage*, vol. 127, pp. 242–256, nov 2015.
- [133] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Academy of Sciences*, vol. 108, pp. 7641–7646, may 2011.
- [134] J. Kinnison, S. Padmala, J.-M. Choi, and L. Pessoa, "Network analysis reveals increased integration during emotional and motivational processing.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 32, pp. 8361–72, jun 2012.
- [135] V. D. Calhoun, T. Eichele, and G. Pearlson, "Functional brain networks in schizophrenia : a review," *Frontiers in Human Neuroscience*, vol. 3, p. 17, 2009.
- [136] Ü. Sakoğlu, G. D. Pearlson, K. a. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun, "A method for evaluating dynamic functional network connectivity and task-modulation : application to schizophrenia," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 23, pp. 351–366, dec 2010.
- [137] J. P. Hamilton, G. Chen, M. E. Thomason, M. E. Schwartz, and I. H. Gotlib, "Investigating neural primacy in Major Depressive Disorder : multivariate Granger causality analysis of resting-state fMRI time-series data.," *Molecular psychiatry*, vol. 16, pp. 763–72, jul 2011.
- [138] E. Tagliazucchi and H. Laufs, "Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep.," *Neuron*, vol. 82, pp. 695–708, may 2014.
- [139] E. Tagliazucchi, F. von Wegner, A. Morzelewski, V. Brodbeck, K. Jahnke, and H. Laufs, "Breakdown of long-range temporal dependence in default mode and attention networks during deep sleep," *Proceedings of the National Academy of Sciences*, vol. 110, pp. 15419–15424, sep 2013.
- [140] A. Kucyi and K. D. Davis, "Dynamic functional connectivity of the default mode network tracks daydreaming.," *NeuroImage*, vol. 100, pp. 471–80, oct 2014.
- [141] E. Tagliazucchi, R. Carhart-Harris, R. Leech, D. Nutt, and D. R. Chialvo, "Enhanced repertoire of brain dynamical states during the psychedelic experience.," *Human brain mapping*, vol. 35, pp. 5442–56, nov 2014.
- [142] B.-H. Juang and L. R. Rabiner, "A Probabilistic Distance Measure for Hidden Markov Models," *AT&T Technical Journal*, vol. 64, pp. 391–408, feb 1985.
- [143] E. R. Lowther, J. T. O'Brien, M. J. Firbank, and A. M. Blamire, "Lewy body compared with Alzheimer dementia is associated with decreased functional connectivity in resting state networks," *Psychiatry Research : Neuroimaging*, vol. 223, pp. 192–201, sep 2014.
- [144] L. R. Peraza, M. Kaiser, M. Firbank, S. Graziadio, L. Bonanni, M. Onofri, S. J. Colloby, A. Blamire, J. O'Brien, and J.-P. Taylor, "fMRI resting state networks and their association with cognitive fluctuations in dementia with Lewy bodies," *NeuroImage : Clinical*, vol. 4, pp. 558–565, 2014.

- [145] E. K. Johns, N. A. Phillips, S. Belleville, D. Goupil, L. Babins, N. Kelner, B. Ska, B. Gilbert, G. Inglis, M. Panisset, C. de Boysson, and H. Chertkow, "Executive functions in frontotemporal dementia and Lewy body dementia.," *Neuropsychology*, vol. 23, pp. 765–77, nov 2009.
- [146] J. H. Yoon, M. Kim, S. Y. Moon, S. W. Yong, and J. M. Hong, "Olfactory function and neuropsychological profile to differentiate dementia with Lewy bodies from Alzheimer's disease in patients with mild cognitive impairment : A 5-year follow-up study," *Journal of the Neurological Sciences*, vol. 355, no. 1-2, pp. 174–179, 2015.
- [147] T. J. Ferman, G. E. Smith, B. F. Boeve, N. R. Graff-Radford, J. A. Lucas, D. S. Knopman, R. C. Petersen, R. J. Ivnik, Z. Wszolek, R. Uitti, and D. W. Dickson, "Neuropsychological Differentiation of Dementia with Lewy Bodies from Normal Aging and Alzheimer's Disease," *The Clinical Neuropsychologist*, vol. 20, no. 4, pp. 623–636, 2006.
- [148] T. J. Ferman, G. E. Smith, K. Kantarci, B. F. Boeve, V. S. Pankratz, D. W. Dickson, N. R. Graff-Radford, Z. Wszolek, J. V. Gerpen, R. Uitti, O. Pedraza, M. E. Murray, J. Aakre, J. Parisi, D. S. Knopman, and R. C. Petersen, "Nonamnesic mild cognitive impairment progresses to dementia with Lewy bodies," *Neurology*, vol. 81, no. 23, pp. 2032–2038, 2013.
- [149] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, "Investigations into resting-state connectivity using independent component analysis," *Philosophical Transactions of the Royal Society B : Biological Sciences*, vol. 360, pp. 1001–1013, may 2005.
- [150] S. Markett, M. Reuter, C. Montag, G. Voigt, B. Lachmann, S. Rudorf, C. E. Elger, and B. Weber, "Assessing the function of the fronto-parietal attention network : Insights from resting-state fMRI and the attentional network test," *Human Brain Mapping*, vol. 35, pp. 1700–1709, apr 2014.
- [151] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius, "Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control," *Journal of Neuroscience*, vol. 27, pp. 2349–2356, feb 2007.
- [152] K. Kantarci, V. J. Lowe, B. F. Boeve, S. D. Weigand, M. L. Senjem, S. A. Przybelski, D. W. Dickson, J. E. Parisi, D. S. Knopman, G. E. Smith, T. J. Ferman, R. C. Petersen, and C. R. Jack, "Multimodality imaging characteristics of dementia with Lewy bodies," *Neurobiology of Aging*, vol. 33, pp. 2091–2105, sep 2012.
- [153] S. J. Colloby, J. D. Fenwick, D. E. Williams, S. M. Paling, K. Lobotesis, C. Ballard, I. McKeith, and J. T. O'Brien, "A comparison of 99mTc-HMPAO SPET changes in dementia with Lewy bodies and Alzheimer's disease using statistical parametric mapping," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 29, pp. 615–622, may 2002.
- [154] K. Lobotesis, J. Fenwick, A. Phipps, A. Ryman, A. Swann, C. Ballard, I. McKeith, and J. O'Brien, "Occipital hypoperfusion on SPECT in dementia with Lewy bodies but not AD," *Neurology*, vol. 56, pp. 643–649, mar 2001.
- [155] D. Meunier, R. Lambiotte, and E. T. Bullmore, "Modular and Hierarchically Modular Organization of Brain Networks," *Frontiers in Neuroscience*, vol. 4, p. 200, 2010.

Analyse de la dynamique temporelle et spatiale des réseaux cérébraux spontanés obtenus en imagerie par résonance magnétique fonctionnelle

Résumé

L'imagerie par résonance magnétique fonctionnelle (IRMf) est un outil de choix pour cartographier d'une manière non invasive l'activité du cortex, donnant ainsi un accès à l'organisation fonctionnelle cérébrale. Cette organisation des aires cérébrales en réseaux complexes reste encore un vaste sujet d'étude, autant dans le domaine de la recherche fondamentale, pour mieux comprendre le développement et le fonctionnement du cerveau, que dans le domaine clinique, à des fins diagnostiques par exemple. Les réseaux cérébraux dits de repos, chez un sujet donné, peuvent être observés lors d'études IRMf lorsqu'aucune tâche motrice ou cognitive n'est imposée au sujet imagé. La première partie de cette thèse a permis le développement d'une méthode automatique d'identification de ces réseaux. Réalisée à l'échelle du sujet, cette méthode permet de sélectionner tous les réseaux spécifiques au sujet ce qui s'avère nécessaire dans un cadre diagnostique où l'individu prime. Au delà de la détection et de l'identification de ces réseaux, l'étude de leurs modes d'interaction dans l'espace et dans le temps et plus généralement l'analyse de la dynamique de la connectivité fonctionnelle (DCF) fait l'objet d'un intérêt grandissant. Cette analyse nécessite le développement de méthodes innovantes de traitement du signal et de l'image qui, pour l'heure, sont encore de nature exploratoire. La deuxième partie de cette thèse présente donc de nouvelles approches pour caractériser la DCF en utilisant le cadre probabiliste de modèles de Markov cachés multidimensionnels. Les mécanismes conversationnels entre réseaux cérébraux peuvent ainsi être identifiés et caractérisés à l'échelle de la seconde. Deux applications, au niveau du sujet puis du groupe, ont permis de mettre en avant les modifications des propriétés dynamiques des interactions entre réseaux sous certaines conditions ou pathologies.

**IRMf – réseaux de repos – dynamique de la connectivité fonctionnelle – MMC
multidimensionnels**

Abstract

The functional magnetic resonance imaging (fMRI) is a perfect tool for mapping in a non-invasive manner the activity of the cortex, giving access to the functional organization of the brain. This organization of brain areas into complex networks remains a large topic of study, both from a fundamental research perspective, to better understand the development and function of the brain, and from a clinical perspective, for diagnostic purposes for instance. The resting-state networks in a given subject can be observed in fMRI studies where no motor or cognitive tasks are imposed to the subject. The first part of this thesis focused on the development of an automatic identification method of these networks. Performed at the subject level, this method selects all the resting-state networks proper to the subject. Beyond the detection and identification of these networks, the study of interactions between these networks in space and time, and more generally the analysis of the dynamic functional connectivity (DFC), is the subject of growing interest. This analysis requires the development of innovative methods of signal or image processing that, for now, are still exploratory. The second part of this thesis thus presents new approaches to characterize the DFC using the probabilistic framework of multidimensional hidden Markov models. Conversational mechanisms between brain networks can be identified and characterized at the resolution of the second. Two applications, first on a single subject then on a group, helped to highlight the changes of dynamic properties of interaction between networks under certain conditions or diseases.

**fMRI – resting-state networks – dynamic functional connectivity – multidimensional
HMM**